리버밴스 중점 사업 추진 목록

2024 년 5 월 27 일

1. 발표자료
    a. 2024 년 5 월: WorldLand My AI Network Consensus 2024
    b. 2024 년 4 월: 계명회 강의 (왜 나만의 AI 가 필요한가)
    c. 2024 년 1 월: 리버밴스 IR DECK (STO 를 위한 WorldLand Rollup)
    d. 2024 년 1 월: WorldLand My AI Network 기술 상세 자료 정리
    e. 2023 년 12 월: Foundation AI 모델개발세미나 @마음 AI 사옥
    f. 2023 년 12 월: AI-DEX 개발 최종보고 발표 자료
    g. 2023 년 8 월 28 일: 인공지능의 역사 와 LLM 현황
    h. 2023 년 5 월 15 일: WorldLand AI Network 풀 버전
    i. 2022 년 8 월: TIPS 발표자료 및 기술 상세
2. 논문 특허 과제 목록 Excel.



Consensus 2024
May 29th – 31st, Austin, Texas, USA
Heung-No Kerry Lee
@fighterkerry

LiberVance

01. Problem

De-globalization fragments world markets.

People suffer!

- Inflated asset prices
- Inflated food and energy prices

People need goods & services!



Dall-E generated by HNLee's prompt

01. Problem

What future would centralized planning create?

Government spending







Asset Price Inflation



01. Problem

Collaboration of empowered individuals creates miracles.



Dall-E generated by HNLee's prompt



02. Solution

Bitcoin is a solution.

Bitcoin

Empowering individuals

- Store of value
- Self custody
- Inflation hedging
- Safe haven to geopolitical risks



02. Solution

Bitcoin is a solution **but getting old!**

Bitcoin

Empowering individuals

- Store of value
- Self custody
- Inflation hedging
- Safe heaven to geopolitical risks

BUT

- Not a currency!
- Difficult to deliver digital products!

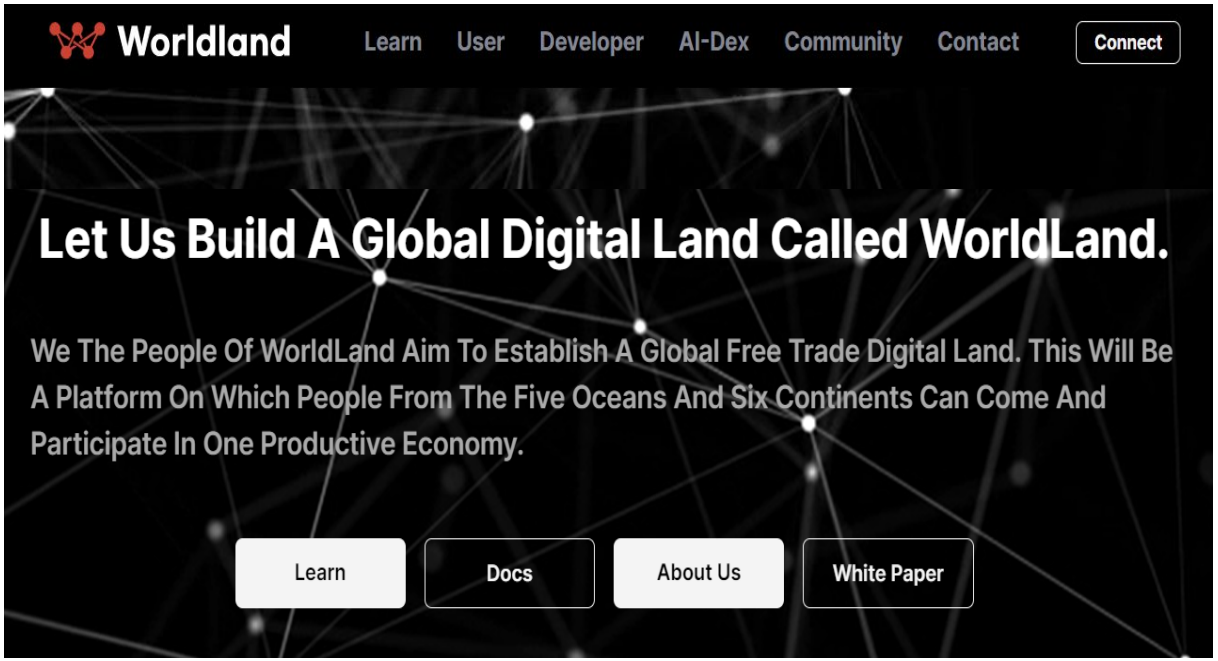


03. Solution

WorldLand is a new solution.

https://worldland.foundation

- Aug. 2023
- 2.5M Blocks
- 20K Nodes



03. Solution

WorldLand advances sound money!

- Offers sound money
  - 2 year halving period
  - 50M WLCs

"I am more convinced than ever that if we ever again have sound money, it will not come from the government. This is issued by a private enterprise."

Denationalization of Money 1976



Friedrich Hayek (1899 – 1992)
Novel Prize in Economic Sciences (1974)

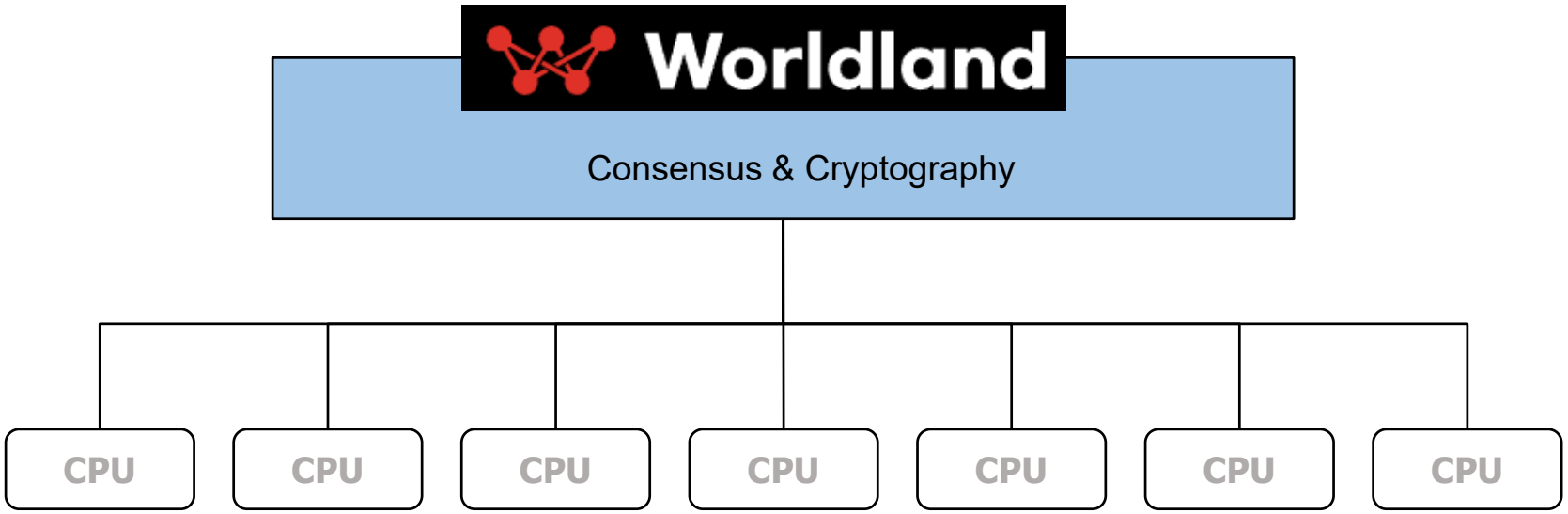
03. Solution

WorldLand is EVM compatible!

- Ethereum Virtual Machine
  - Turing complete
  - Easy to program

03. Solution

Worldland is ready for post-quantum computers!



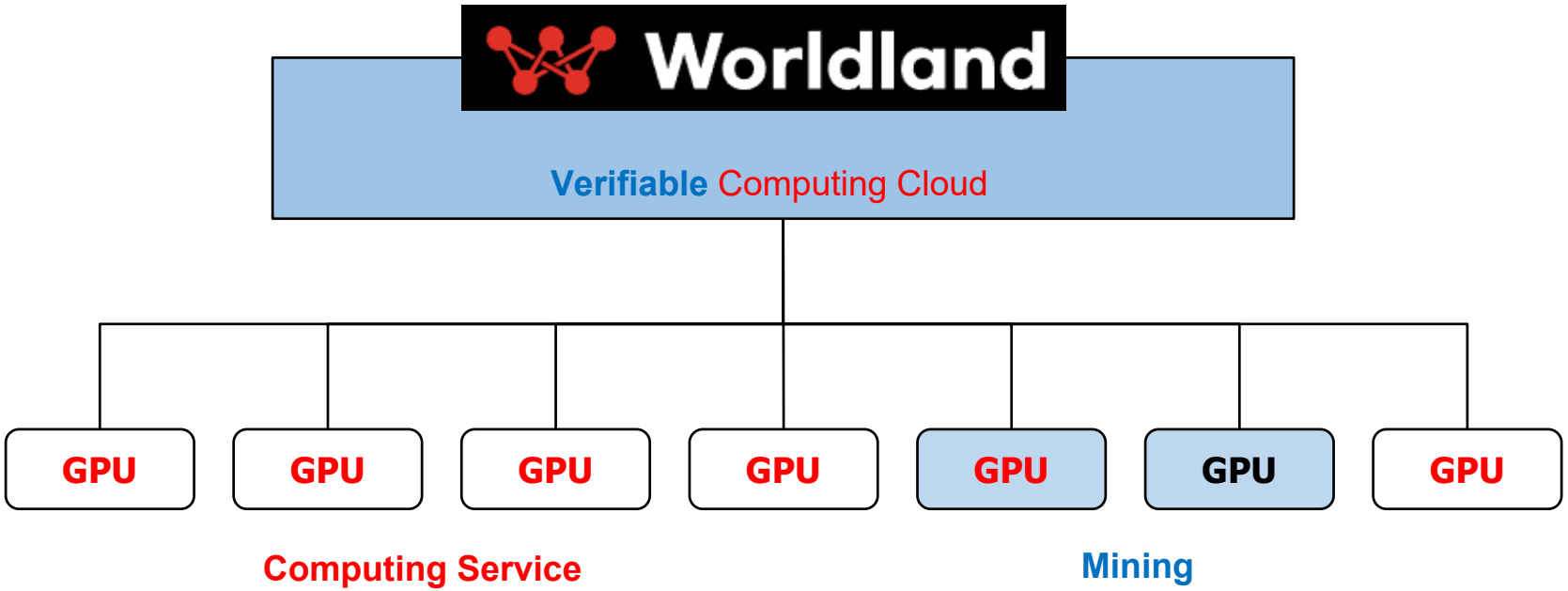
Post-Quantum Security

Error Correction Code Consensus

Lattice Codes Cryptography

03. Solution

Worldland extends best decentralized consensus!



Error Correction Code PoW

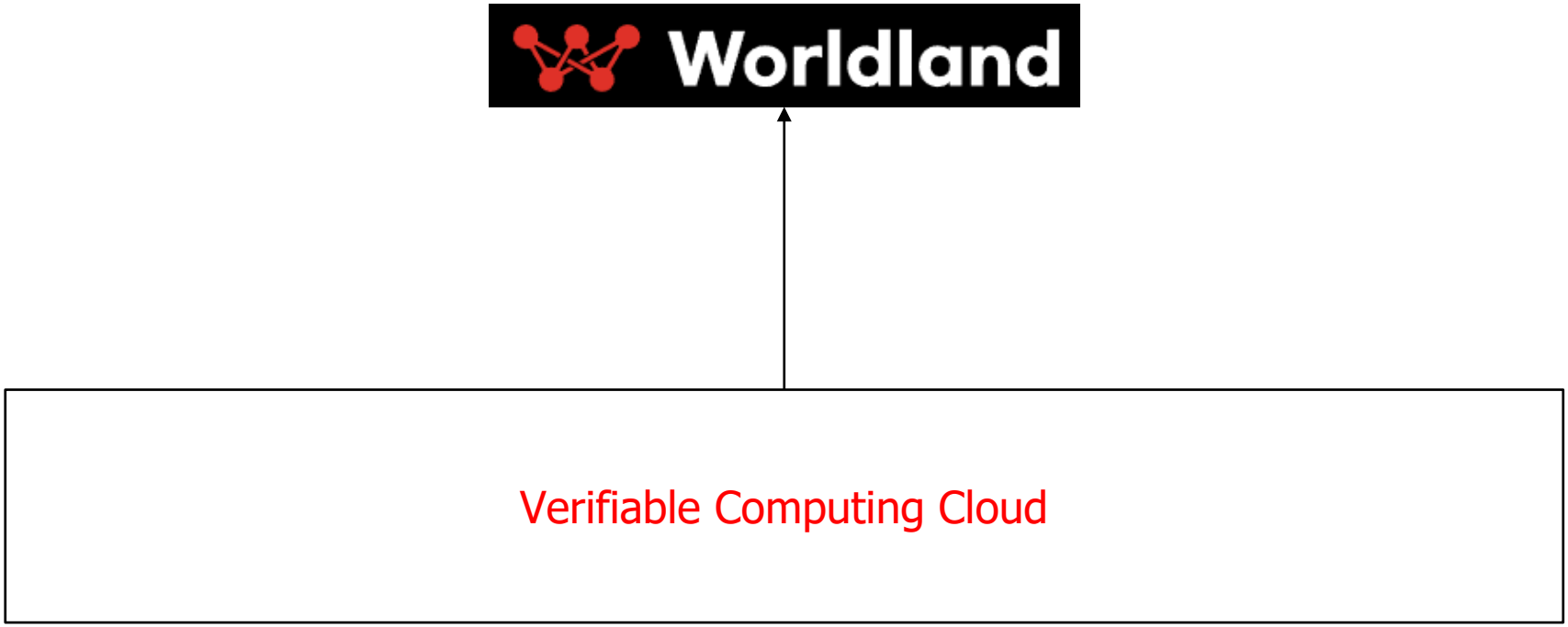
- **ASIC resistance**

Verifiable Coin Toss

- **Random self selection to do either mining or computing service**

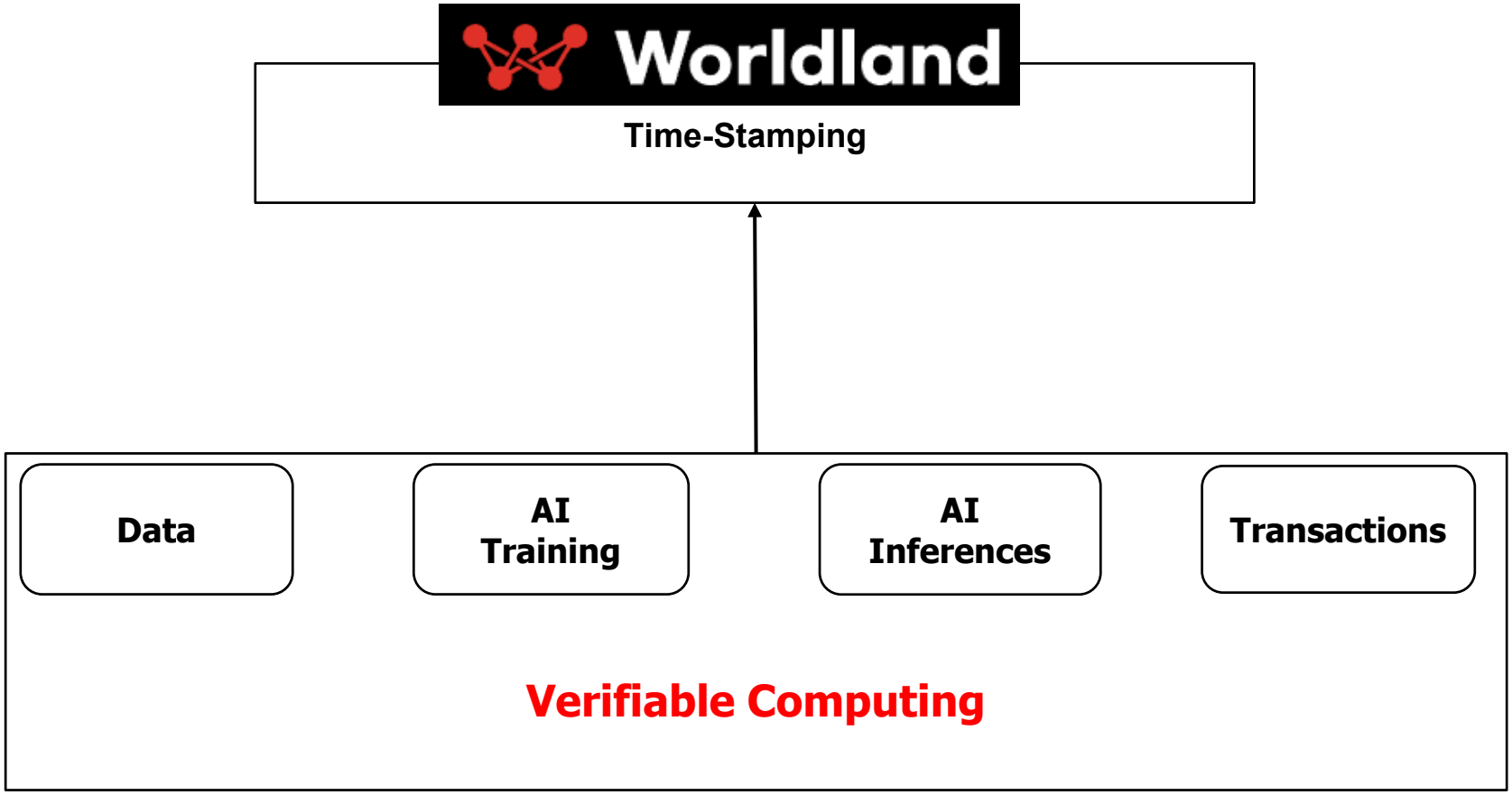
03. Solution

Worldland offers verifiable computing cloud!



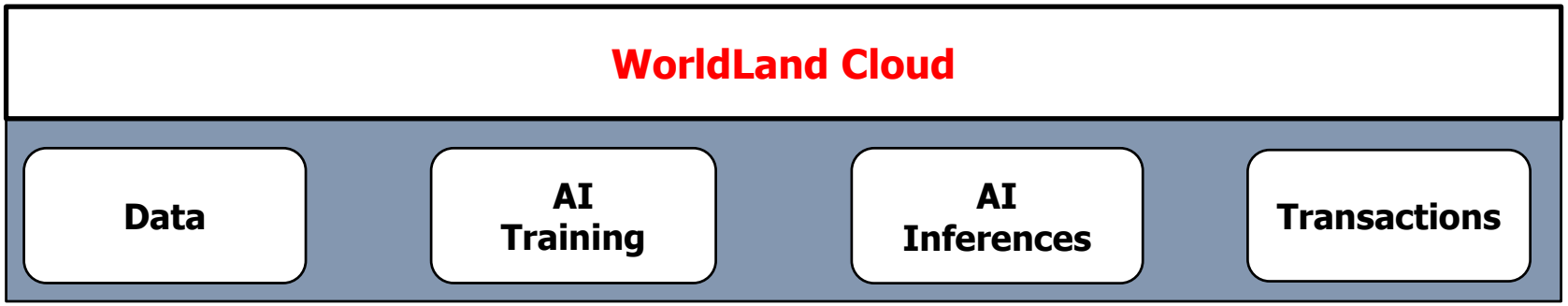
03. Solution

Worldland enables verifiable AI computing!



03. Solution

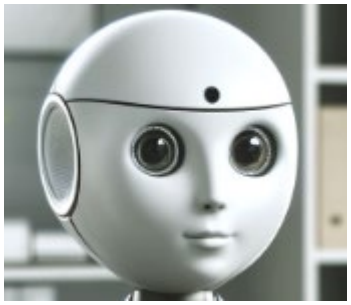
Worldland is a Web3 cloud!



03. Solution: My AI Network

Human vs. AI





03. Solution

In society, one teaches, others learn.

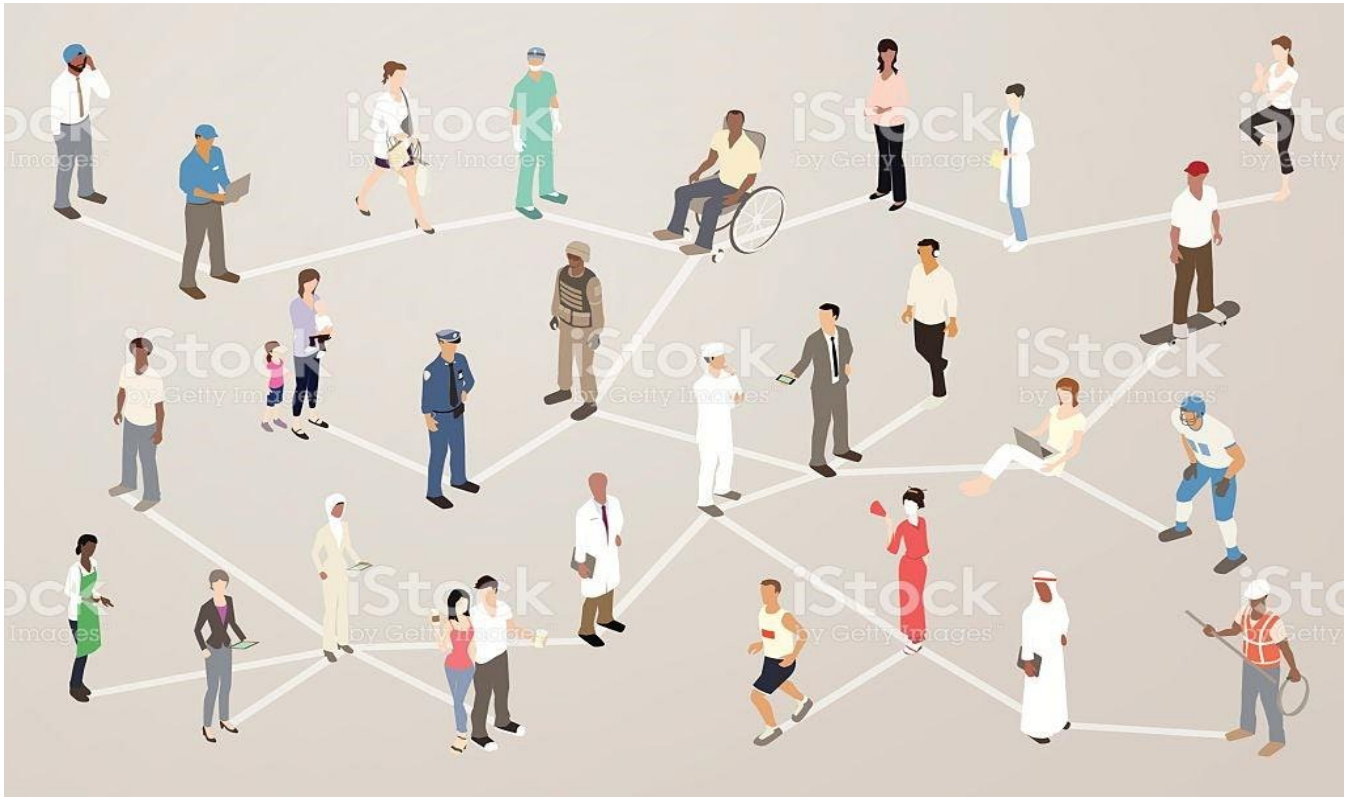






03. Solution

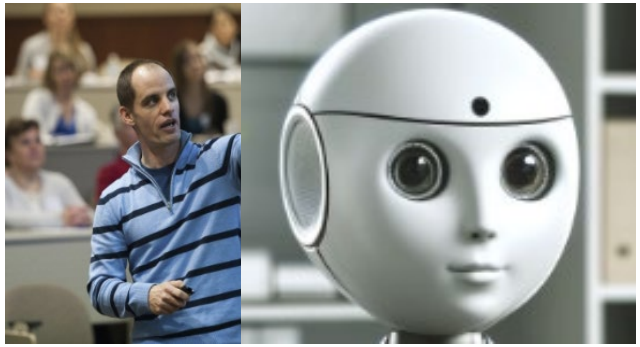
Interactions make network intelligence.

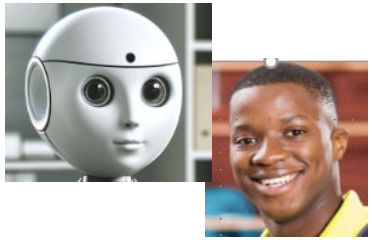


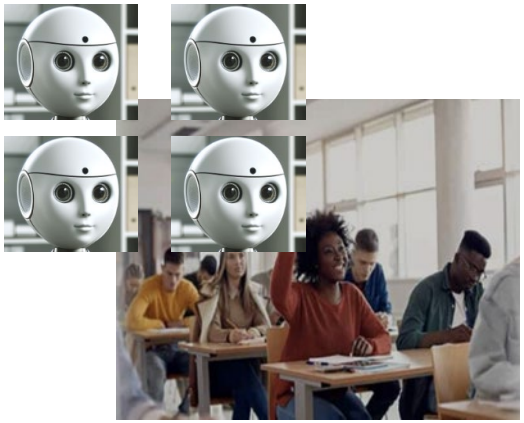
03. Solution

A person can grow his AI and have it interact with others.

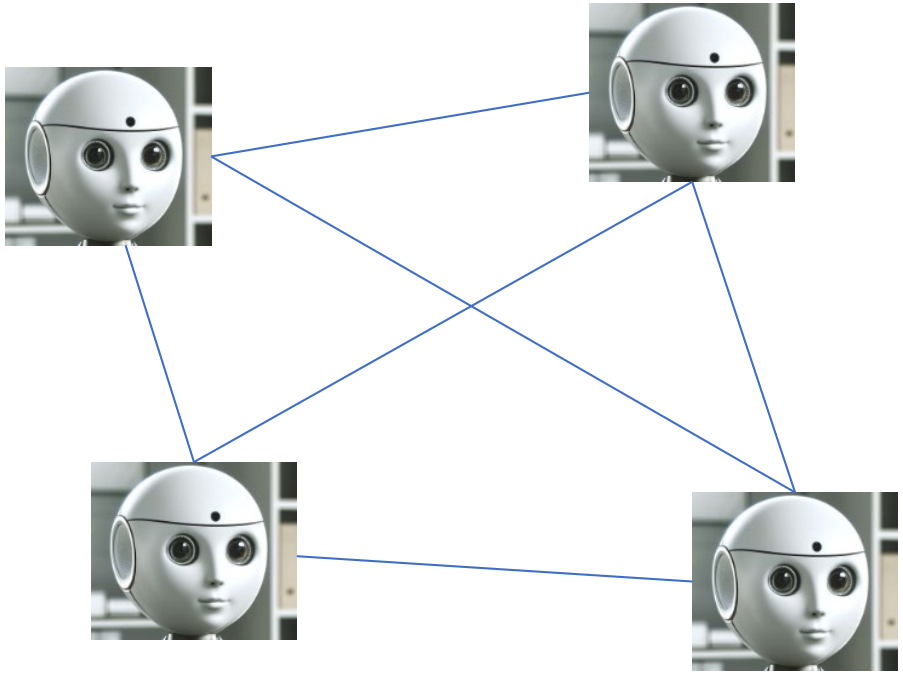
My AI







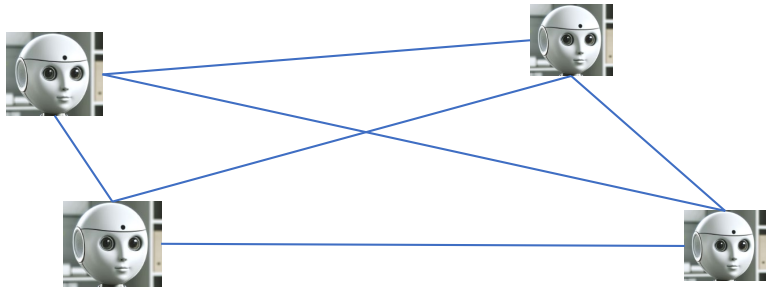
My AI Network

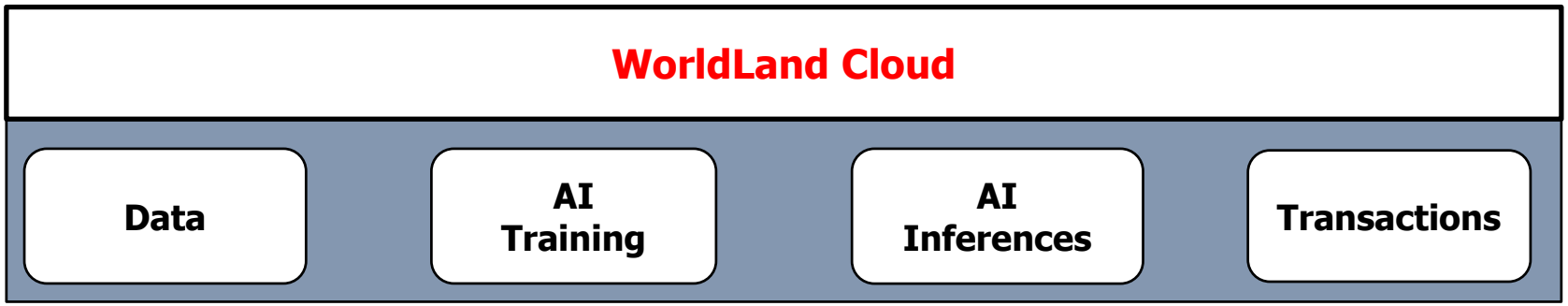


03. Solution

Worldland supports My AI Network!

Give digital products and take tokens!





03. Solution

You can have your AI make money for you.

WorldLand My AI Network







04. Market

Market is growing fast.

- Generative AI:
  CAGR 38%, USD266 by 32’
- Blockchain:
  CAGR 30%, USD 766B by 30’
- AI Cloud:
  CAGR 20%, USD 274B by 29’

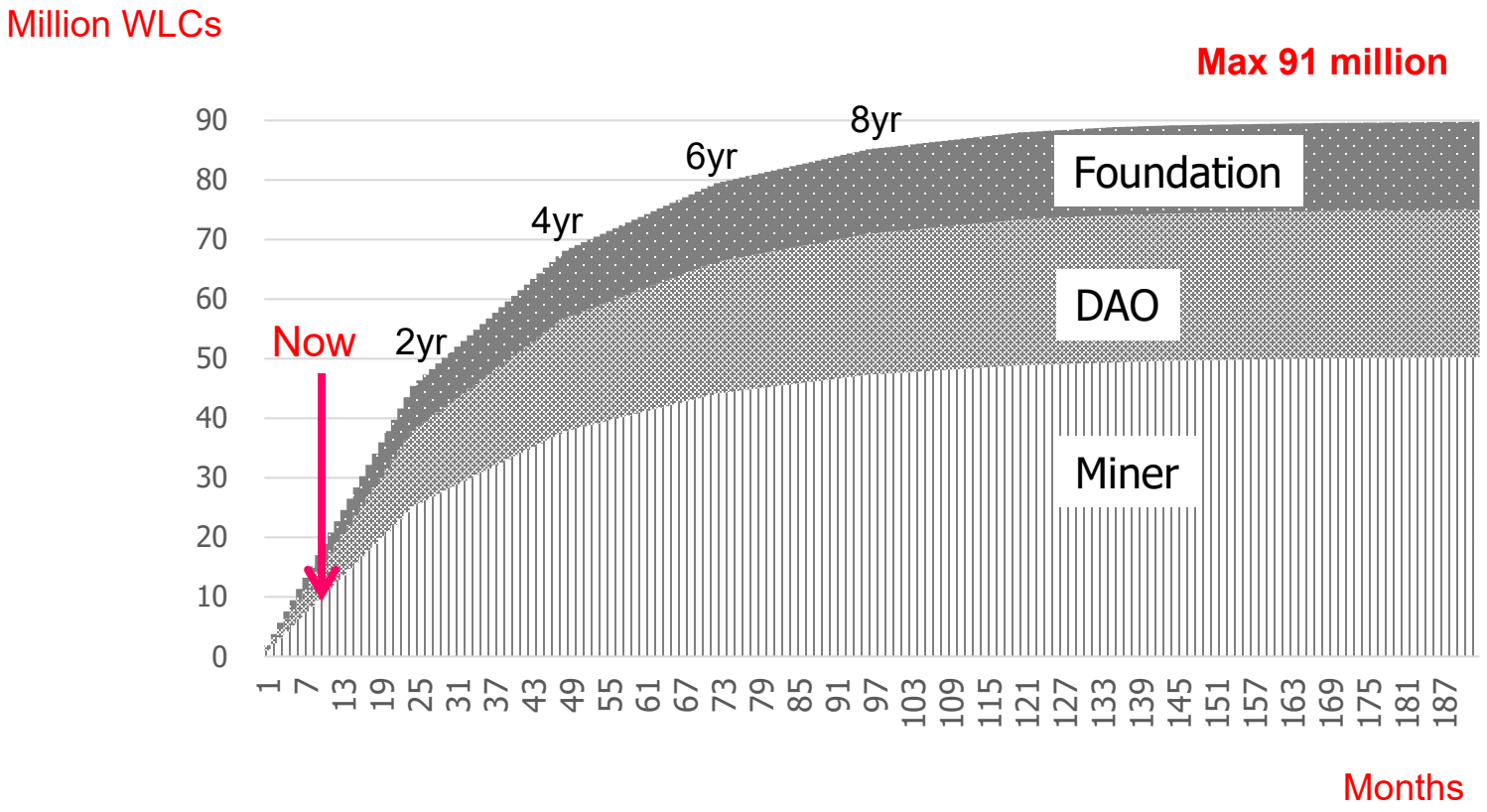
04. Market

Our solution is unique, PQ secure and decentralized!

Blockchain-Infra-for-AI Players

- DeepBrain Chain(THETA)
  - Market cap: USD 2.1B
  - POS
  - AI Computing
- Fetch.AI (FET)
  - Market cap: USD 1.8B
  - POS
  - AI Agent Interaction
- SingularityNET (AGIX)
  - Market cap: USD 1.1B
  - AI Market
  - No mainnet

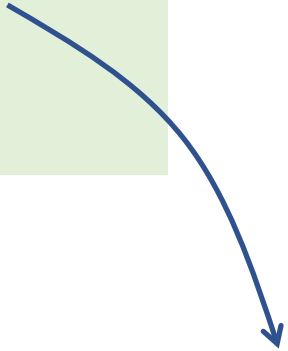
Worldland Coin Minting Schedule



05. Business

WorldLand coinomics!

- ❑ Total is 91M WLCs.
  - Total Minted Genesis TMG is 41M WLCs.
  - Total Mining is 50M WLCs.



Why TMGs?

We aim to raise developers and projects!

Educators
Developers
Entrepreneurs

05. Business

We plan to fund new projects on WorldLand!

TMG 41M WLCs over years.

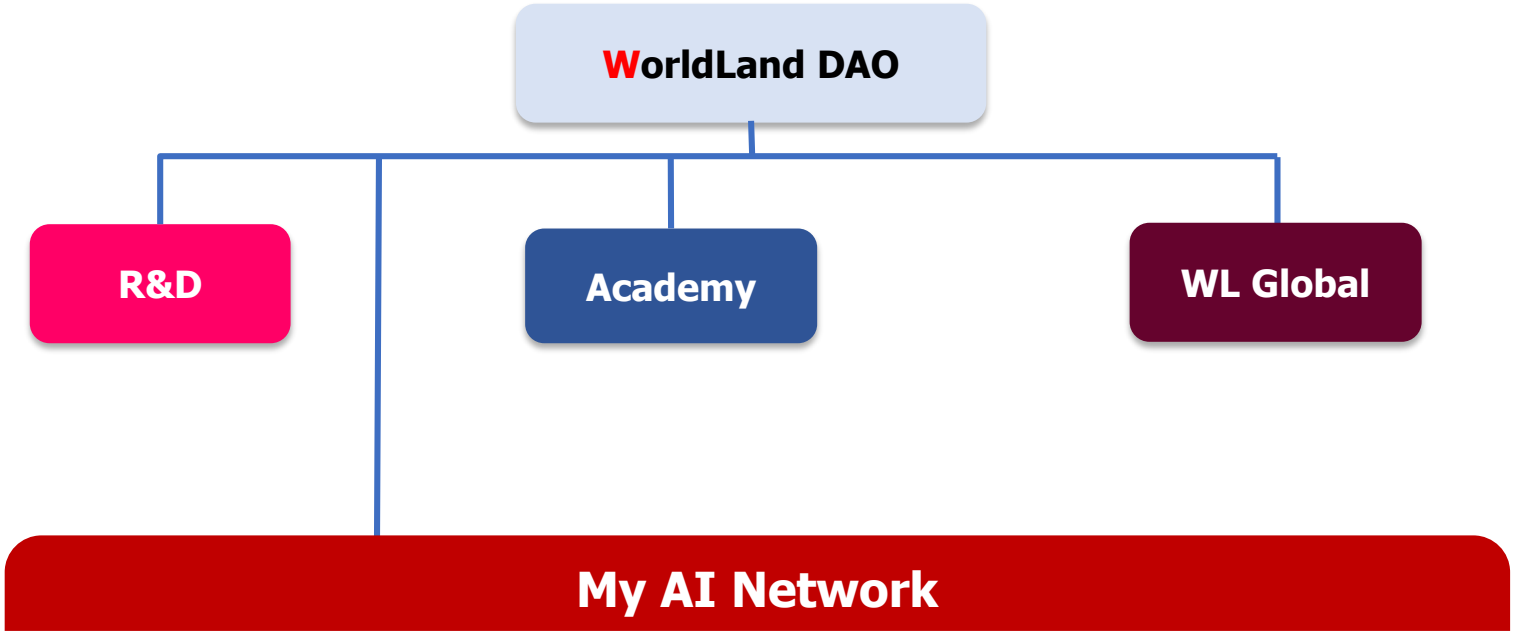
- 80% of TMG is for ECO.

- 20% of TMG is for LTE.

LTE : Long-Term Endowment
ECO Projects: My AI Network, Wallets, RnDs, Investments

05. Business

Send us your proposals, DAO evaluates & funds them.

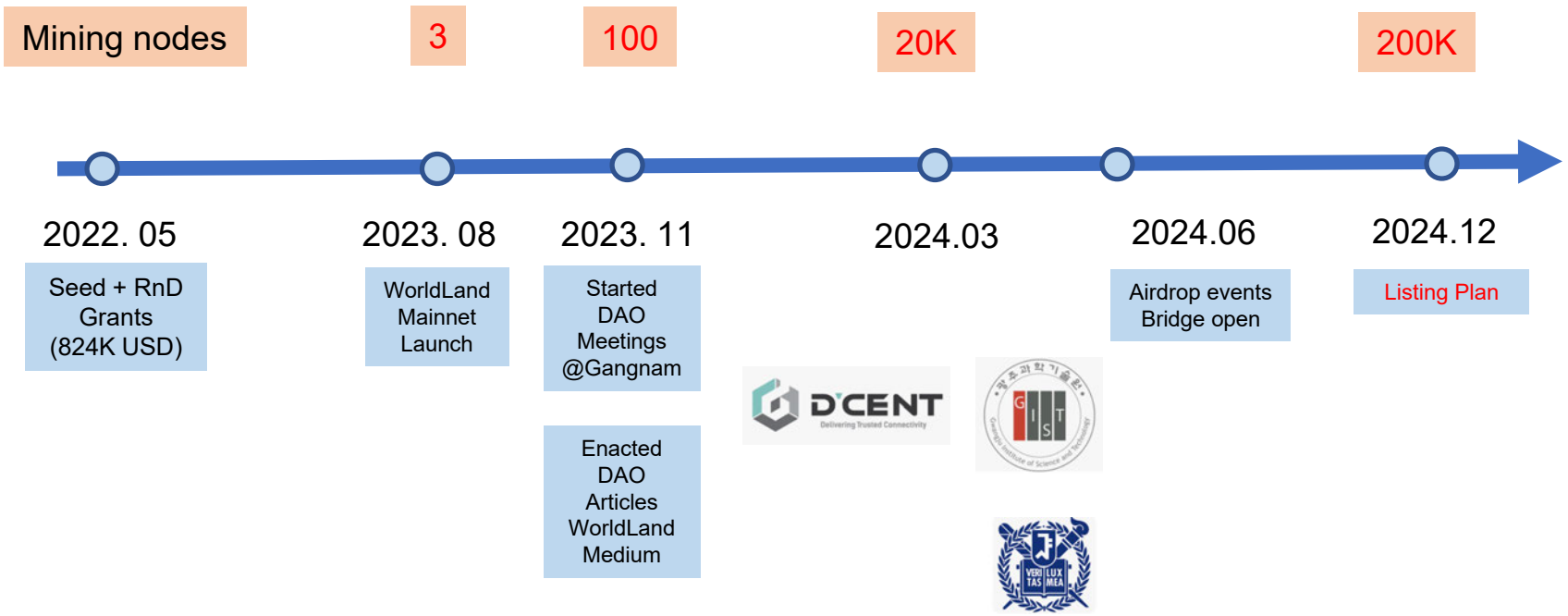


05. WorldLand Status

WorldLand is growing explosively!

No ICO
No coin sale

What will occur when ECO fund begins distribution?



LiberVance

06. Founder & CEO

LiberVance

- **Origin**: University Lab Spin Off
- **Headquarter**: Academic-Industrial Cooperation Research Building, GIST, Korea
- **Seoul Office**: Gangnam-gu, Seoul, Korea
- **Employees**: Total 13

Major Milestones

- **January 2020**: Established as a GIST Faculty Startup
- **April 2022**: Investment from Sunbo Angel Partners
- **July 2022**: Secured TIPS RnD Project



Heung-No Lee

IEEE

- Editorial Board Member of IEEE Transactions on Cybernetics (Top 4%)
- Tenured Professor at GIST
- Director of Blockchain Intelligence Convergence

- Former Professor at the University of Pittsburgh, PA, USA
- Former Researcher at Hughes Research Lab, Malibu, CA, USA
- Graduated from UCLA, USA (BS 93', MS 94', PhD 99')

- Authored over 100+ international journal papers

- Awards
- Prime Minister's Commendation (April 2022)
- IEIE Haedong Academic Award (2019)
- NRF Scientist of the Month Award (January 2014)

Total fund raised so far is $824K.

Angel Investment: $145K
Total R&D Support: $679K

Market Caps

- No ICO ever!

- In circulation: 10M WLCs
- OTC market price today: $0.2/WLC

- Current Market Cap = 10M *$0.2 = $2M
- Fully Diluted Market Cap = $0.2 *91M = $18M

Fund raising goal this round

- Goal: $1M or 15 BTCs
- Need:
  - Roadmap to support My AI Network
- How
  - LV Stocks
  - WLC Block Sale

We aim to raise 15 bitcoins for roadmap and scale up!



WorldLand Block Sale Event

Each is redeemable for 500K WLCs.

- Class A: price > 1 btc, 1yr lock 1yr vesting
- Class B: price > 2 btc, 1yr vesting

07. IR Competitiveness

Please find below our 9 point summary:

1. Launched new mainnet on August 8, 2023, which has never stopped producing blocks.
2. Developed WorldLand technology via Korean national RnD funding.
3. Possesses patents and intellectual properties for the project.
4. Draws high interest from Seoul and around the world for the new POW consensus algorithm.
5. Seen explosive recent increase in mining nodes starting March 2024.
6. Pursuing listing on internationally recognized exchanges by Dec. 2024.
7. Aim to scale WorldLand through execution of ecosystem expansion plan.
8. Aim to incubate new projects such as My AI Network and AI-DEX.
9. Please join us to make better future!

Thank you

Worldland



Contact me for more information!

- @fighterkerry
- heungno@gist.ac.kr
- Research Lab: https://heungno.net
- LiberVance: https://libervance.com
- WorldLand Project: https://worldland.foundation



인공지능, GPT, Law AI

계명회 초청 강의

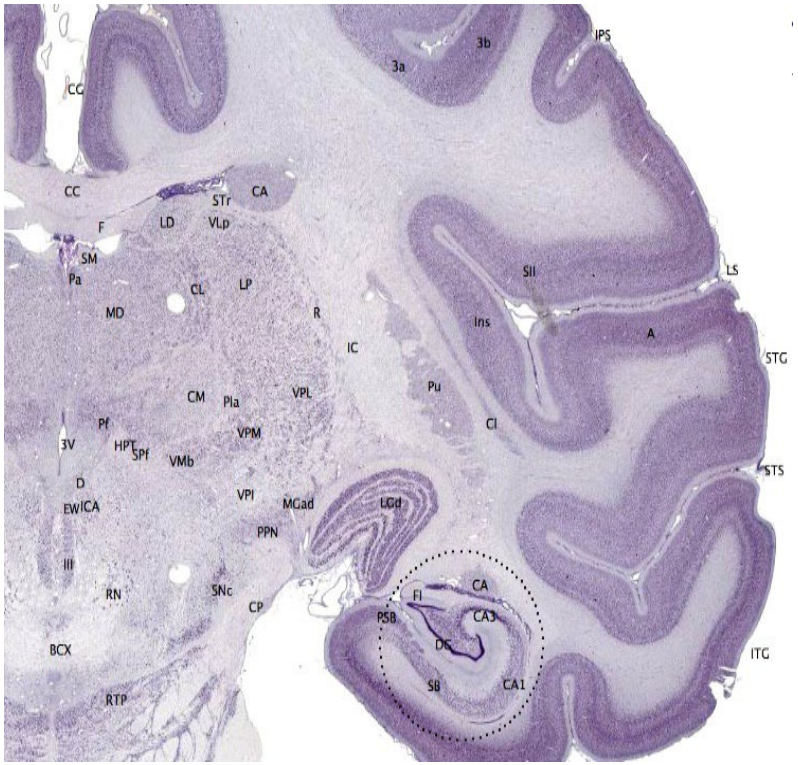
2024.04.06 (17:30) | GIST/리버밴스㈜

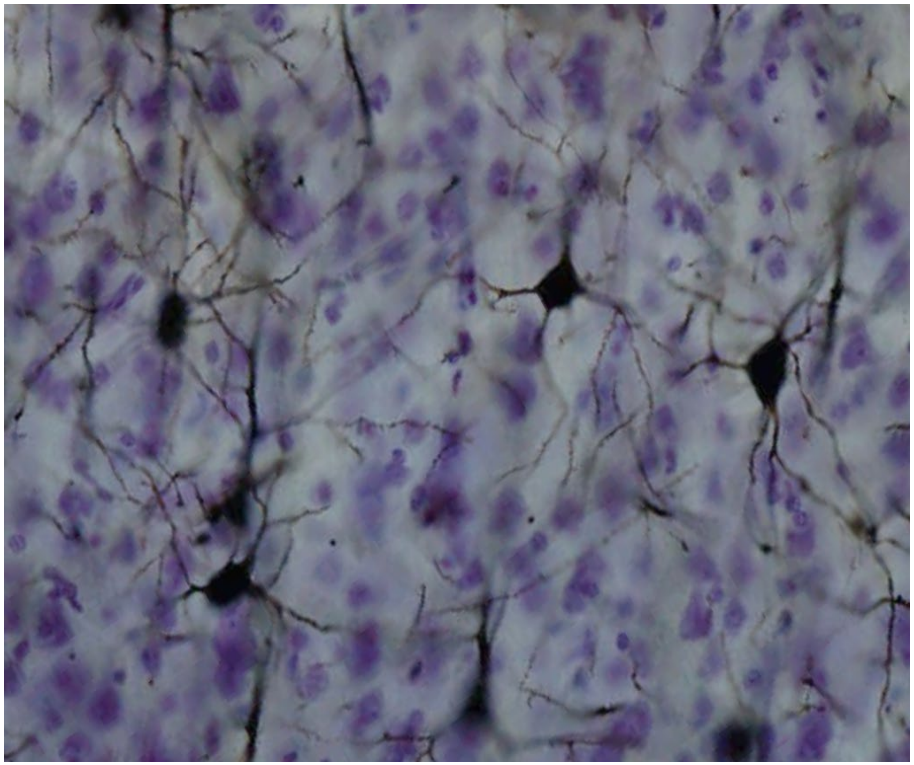
서울 시청 컨퍼런스하우스 달개비

https://heungno.net

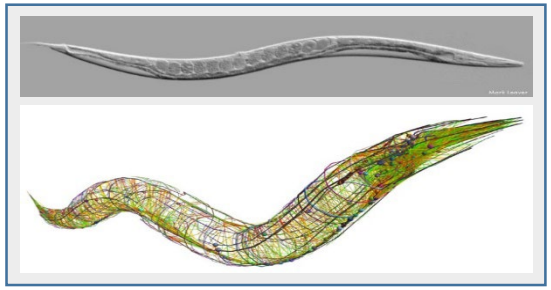
지능

인간의 두뇌 – network of cells

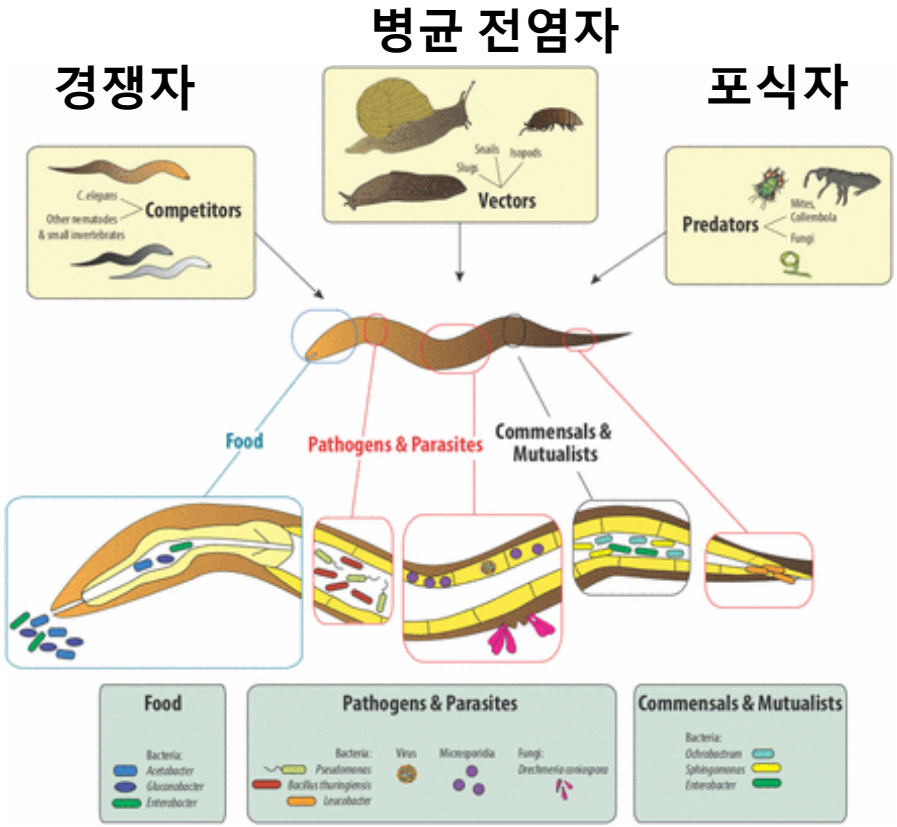




꼬마선충(C elegans), 302개의 뉴론으로 늪지에서 산다.



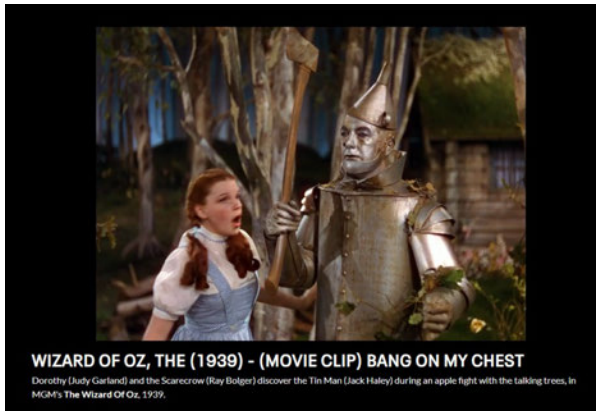


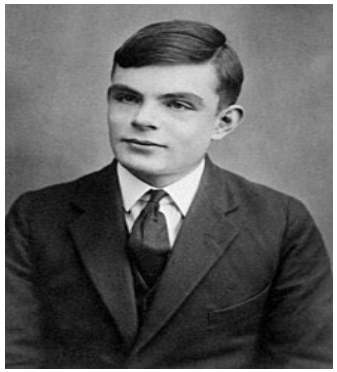


인공지능

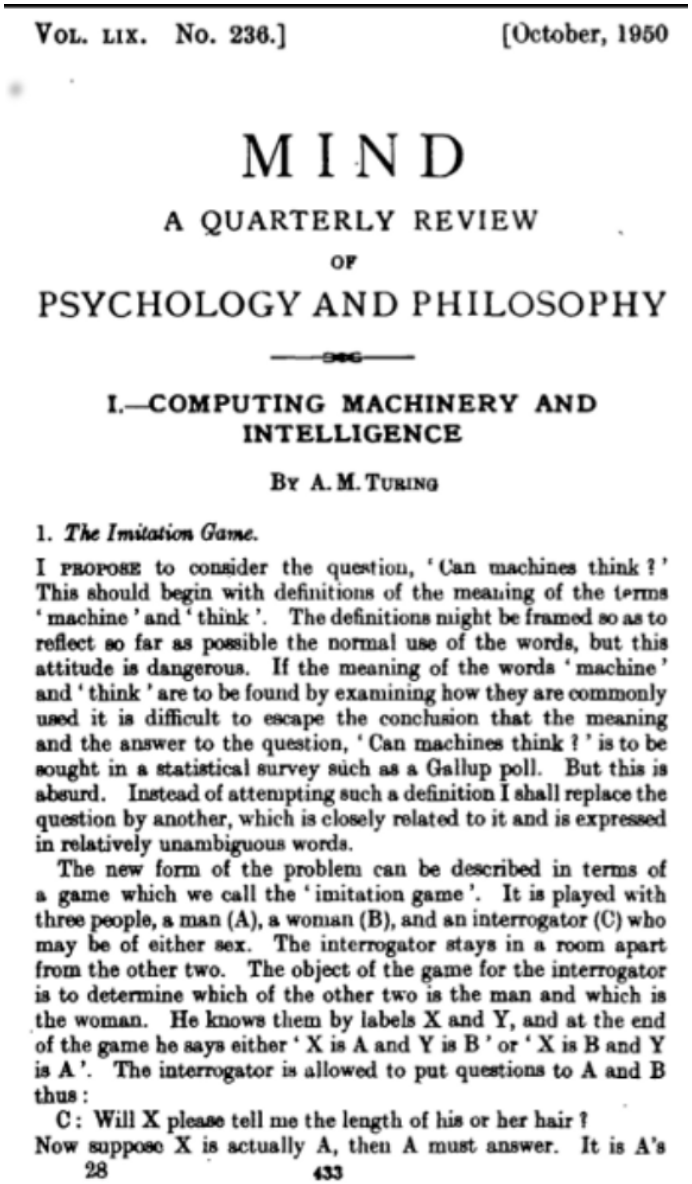
03. 인공지능

Start: Can machines think?





1950: Alan Turing proposes the Turing Test as a measure of machine intelligence in his paper "Computing Machinery and Intelligence."



VOL. LIX. No. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the 'imitation game'. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either 'X is A and Y is B' or 'X is B and Y is A'. The interrogator is allowed to put questions to A and B thus:

C: Will X please tell me the length of his or her hair?

Now suppose X is actually A, then A must answer. It is A's

28 433

2000s

- **2006**: Geoffrey Hinton ~ “deep learning” to fix the “vanishing gradient” problem.

A fast learning algorithm for deep belief nets *

Geoffrey E. Hinton and **Simon Osindero**
Department of Computer Science University of Toronto
10 Kings College Road
Toronto, Canada M5S 3G4
{hinton, osindero}@cs.toronto.edu

Yee-Whye Teh
Department of Computer Science
National University of Singapore
3 Science Drive 3, Singapore, 117543
tehyw@comp.nus.edu.sg

Abstract

We show how to use “complementary priors” to eliminate the explaining away effects that make inference difficult in densely-connected belief nets that have many hidden layers. Using complementary priors, we derive a fast, greedy algorithm that can learn deep, directed belief networks one layer at a time, provided the top two layers form an undirected associative memory. The fast, greedy algorithm is used to initialize a slower learning procedure that fine-tunes the weights using a contrastive version of the wake-sleep algorithm. After fine-tuning, a network with three hidden layers forms a very good generative model of the joint distribution of handwritten digit images and their labels. This generative model gives better digit classification than the best discriminative learning algorithms. The low-dimensional manifolds on which the digits lie are modelled by long ravines in the free-energy landscape of the top-level associative memory and it is easy to explore these ravines by using the directed connections to display what the associative memory has in mind.

1 Introduction

Learning is difficult in densely-connected, directed belief nets that have many hidden layers because it is difficult to infer the conditional distribution of the hidden activities when given a data vector. Variational methods use simple approximations to the true conditional distribution, but the approximations may be poor, especially at the deepest hidden layer where the prior assumes independence. Also, variational learning still requires all of the parameters to be learned together and makes the learning time scale poorly as the number of parameters increases.

We describe a model in which the top two hidden layers form an *undirected* associative memory (see figure 1) and the remaining hidden layers form a directed acyclic graph that converts the representations in the associative memory into observable variables such as the pixels of an image. This hybrid model has some attractive features:

1. There is a fast, greedy learning algorithm that can find a fairly good set of parameters quickly, even in deep networks with millions of parameters and many hidden layers.
2. The learning algorithm is unsupervised but can be applied to labeled data by learning a model that generates both the label and the data.
3. There is a fine-tuning algorithm that learns an excellent generative model which outperforms discriminative methods on the MNIST database of hand-written digits.
4. The generative model makes it easy to interpret the distributed representations in the deep hidden layers.
5. The inference required for forming a percept is both fast and accurate.
6. The learning algorithm is local: adjustments to a synapse strength depend only on the states of the presynaptic and post-synaptic neuron.
7. The communication is simple: neurons only need to communicate their stochastic binary states.

Section 2 introduces the idea of a “complementary” prior which exactly cancels the “explaining away” phenomenon that makes inference difficult in directed models. An example of a directed belief network with complementary priors is presented. Section 3 shows the equivalence between restricted Boltzmann machines and infinite directed networks with tied weights.

Section 4 introduces a fast, greedy learning algorithm for constructing multi-layer directed networks one layer at a time. Using a variational bound it shows that as each new layer is added, the overall generative model improves. The greedy algorithm bears some resemblance to boosting in its repeated use of the same “weak” learner, but instead of reweighting each data-vector to ensure that the next step learns something new, it re-represents it. The “weak” learner that

*To appear in Neural Computation 2006

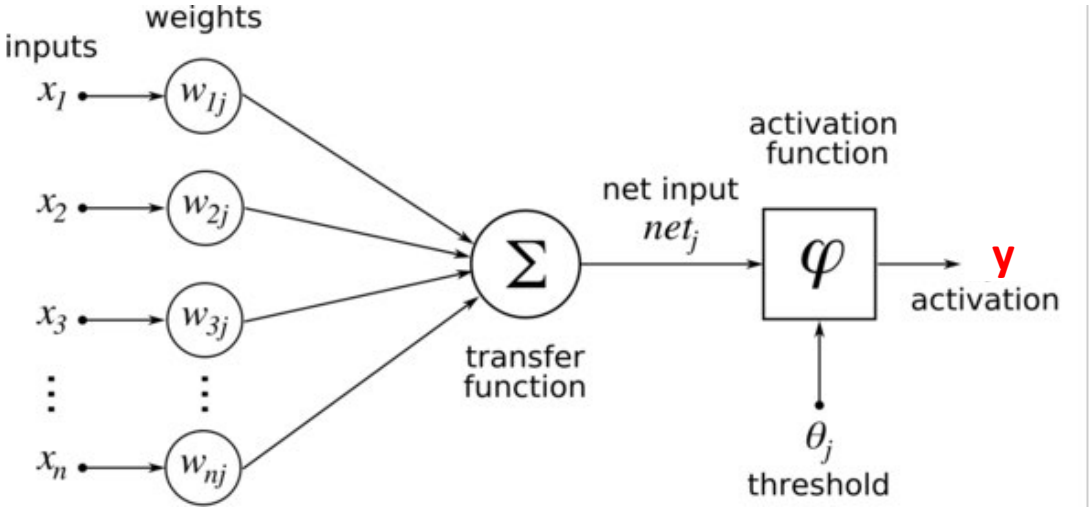
03. 인공지능

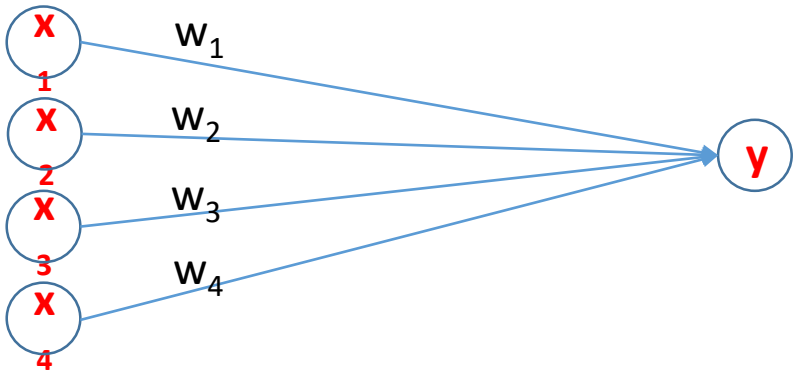
2010s:

- **2011**: IBM's Watson wins the quiz show Jeopardy!, defeating human champions.
- **2012**: AlexNet, a deep neural network, achieves a breakthrough in image recognition accuracy in the ImageNet competition.
- **2014**: Chatbot Eugene Goostman reportedly passes the Turing Test, though this claim is widely debated.
- **2015**: Google DeepMind’s AlphaGo defeats a human professional Go player.
- **2018**: OpenAI’s GPT-2 shows that large language models can generate coherent and contextually appropriate text.

뉴론 Neuron

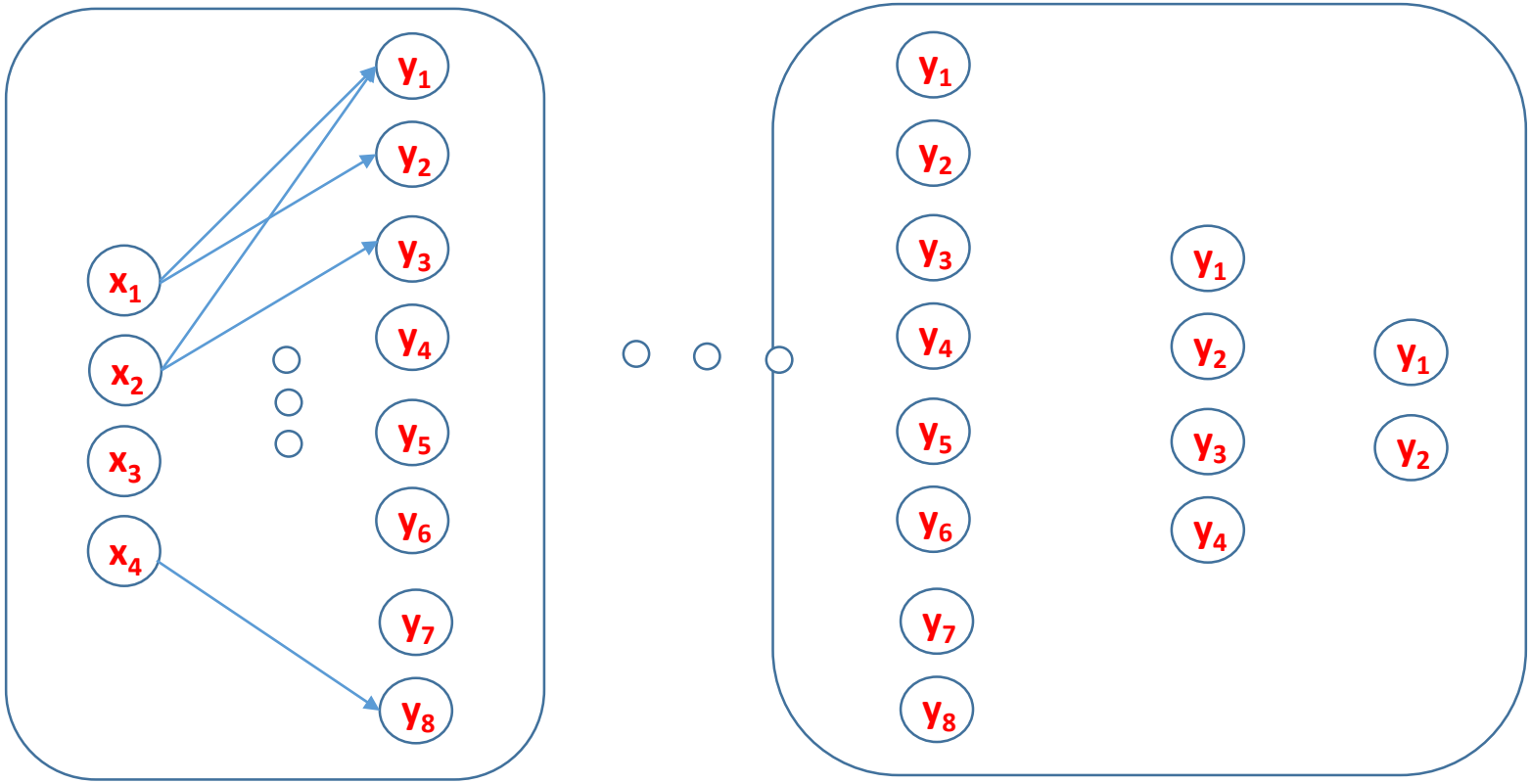
- 입력 x
  - 가중치 w
  - 편향
  - 합계
  - 활성함수
- 출력 y





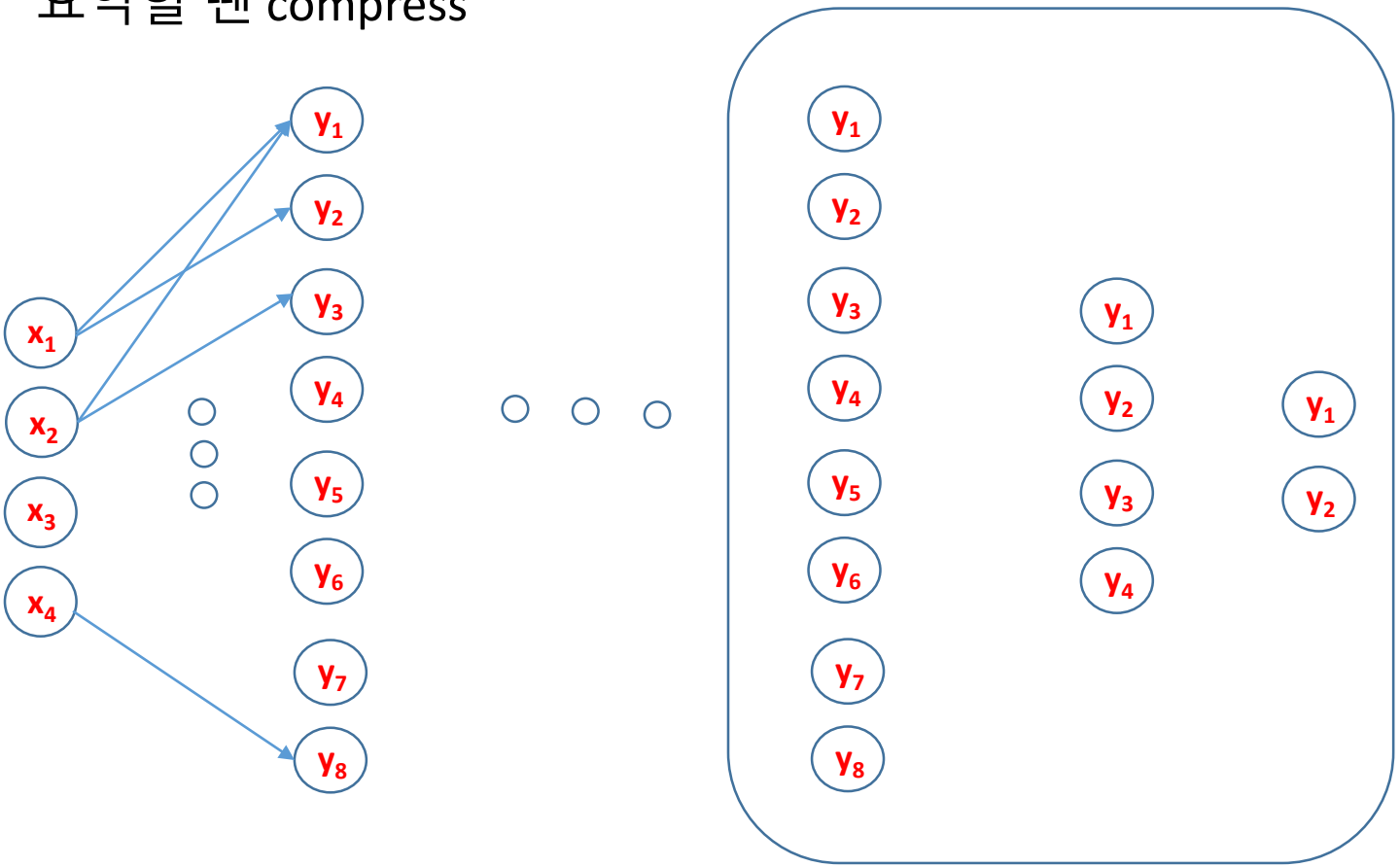
분석과 요약

- 분석할 땐 expand



분석과 요약

- 요약할 땐 compress



Large Language Models

심층신경망 자연어 처리 NLP 발전 주요 이력

	paper	비고
2013.01	Word2vec: “Efficient estimation of word representations in vector space,” Google **단어를 고차원 벡터로 표현. 대규모데이터로 학습. 유사한 연관 단어 벡터는 서로 가깝게 배치.**	BoW: 연관단어 → 타겟단어 Skip-gram: 타겟단어 → 연관단어
2014.02	“Glove: Global vectors for word representation,” by Pennington et.al., Stanford University	전역적인 count-base 통계 모델 및 예측 모델 활용
2016.07	“Enriching word vectors with subword information,” Bojanowski et. al. 단어를 하위어 단어로 분해하고 임베딩, 형태론적 변형과 유사성 포착	단어를 Subword로 쪼개어 형태론적 변형 포착
2017.06	Transformer. “*Attention is all you need*,” Vaswani et.al. **주의집중 메카니즘 활용, 다중 계층 인코더로 문맥 파악, 다중 계층 디코더로 문맥화된 단어 생성**	**주의 집중 메카니즘**
2018.02	Elmo. “Deep contextualized word representations,” Peters et.al. 양방향 LSTM사용, 좌우 문맥을 고려, 각 단어의 문맥화된 표현 생성	양방향 LSTM문맥
2018.10	BERT: bidirectional encoder representation from Transformers, Devlin et.al. Transformer를 활용 단어 문맥 양방향 고려, 단어 임베딩	양방향 Transformer문맥

Transformer

- Attention is all you need! (NIPS2017)
  - GER-ENG 번역기
  - Input 문장을 받아 Output 문장을 생성
  - IN 문장에서 단어들간의 관계 및 주목할 단어 파악, OUT에서 한 단어씩 생성
- 수십 층의 심층 신경망
  - 학습 매개 변수

$$O(layers \times d_{\text{model}}{}^2) \sim \text{ 10s million}$$

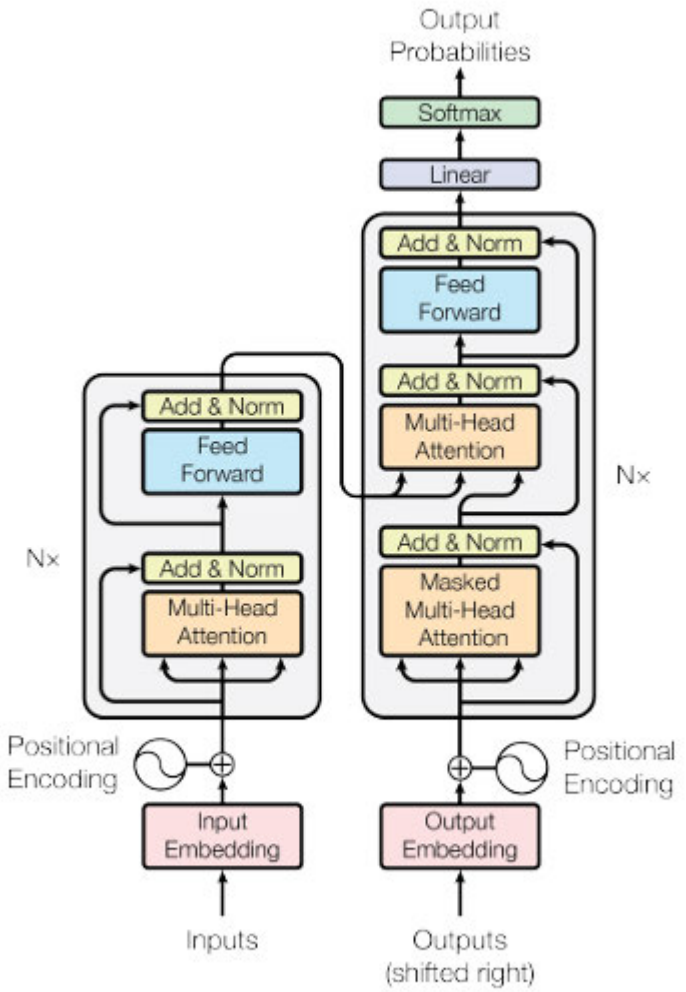


Figure 1: The Transformer - model architecture.

Transformer는 언어 변환 모델로 시작하였다.



04. Large Language Models

OpenAI는 Transformer를 맥락 인지 문장 생성기로 발전시킴.

- GPT-3, July 2020
- GPT-3.5, March 2022
  - 1750억 개 매개변수
  - 800 GB storage
- Emergent Abilities:
  - Few-shot learning
  - Comm. language
- Computer codes, poetry, prose, "Terrifying ability," NYT 2020.

Table 2.2: Datasets used to train GPT-3.

Dataset	Quantity (tokens)	Weight in training mix	Epochs elapsed when training for 300B tokens
Common Crawl (filtered)	410 billion	60%	0.44
WebText2	19 billion	22%	2.9
Books1	12 billion	8%	1.9
Books2	55 billion	8%	0.43
Wikipedia	3 billion	3%	3.4

1. **500 billion 토큰 사전, 비지도 학습**
2. **휴먼 피드백 세부 튜닝 학습**

04. Large Language Models

LLMs 은 few-shot 학습자다.

- **언어 모델의 규모 확장**, 소수 데이터 학습(few-shot learning) 성능 크게 향상.
- 번역, 질의 응답, 빈칸 채우기 작업, 3자리 산술 연산 등 즉석 추론이나 도메인 적응 필요 작업도 처리 가능.
- GPT-3가 뉴스 기사 작성, **인간이 쓴 기사와 구별하기 어려움**.

arXiv:2005.14165v4 [cs.CL] 22 Jul 2020

Language Models are Few-Shot Learners

Tom B. Brown* Benjamin Mann* Nick Ryder* Melanie Subbiah*
Jared Kaplan† Prafulla Dhariwal Arvind Neelakantan Pranav Shyam Girish Sastry
Amanda Askell Sandhini Agarwal Ariel Herbert-Voss Gretchen Krueger Tom Henighan
Rewon Child Aditya Ramesh Daniel M. Ziegler Jeffrey Wu Clemens Winter
Christopher Hesse Mark Chen Eric Sigler Mateusz Litwin Scott Gray
Benjamin Chess Jack Clark Christopher Berner
Sam McCandlish Alec Radford Ilya Sutskever Dario Amodei
OpenAI

Abstract

Recent work has demonstrated substantial gains on many NLP tasks and benchmarks by pre-training on a large corpus of text followed by fine-tuning on a specific task. While typically task-agnostic in architecture, this method still requires task-specific fine-tuning datasets of thousands or tens of thousands of examples. By contrast, humans can generally perform a new language task from only a few examples or from simple instructions – something which current NLP systems still largely struggle to do. Here we show that scaling up language models greatly improves task-agnostic, few-shot performance, sometimes even reaching competitiveness with prior state-of-the-art fine-tuning approaches. Specifically, we train GPT-3, an autoregressive language model with 175 billion parameters, 10x more than any previous non-sparse language model, and test its performance in the few-shot setting. For all tasks, GPT-3 is applied without any gradient updates or fine-tuning, with tasks and few-shot demonstrations specified purely via text interaction with the model. GPT-3 achieves strong performance on many NLP datasets, including translation, question-answering, and cloze tasks, as well as several tasks that require on-the-fly reasoning or domain adaptation, such as unscrambling words, using a novel word in a sentence, or performing 3-digit arithmetic. At the same time, we also identify some datasets where GPT-3's few-shot learning still struggles, as well as some datasets where GPT-3 faces methodological issues related to training on large web corpora. Finally, we find that GPT-3 can generate samples of news articles which human evaluators have difficulty distinguishing from articles written by humans. We discuss broader societal impacts of this finding and of GPT-3 in general.

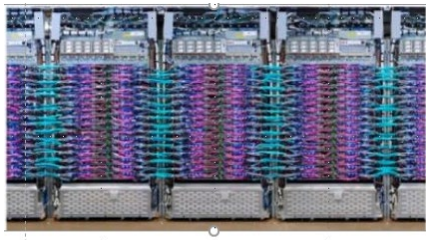
*Equal contribution
†Johns Hopkins University, OpenAI

Author contributions listed at end of paper.

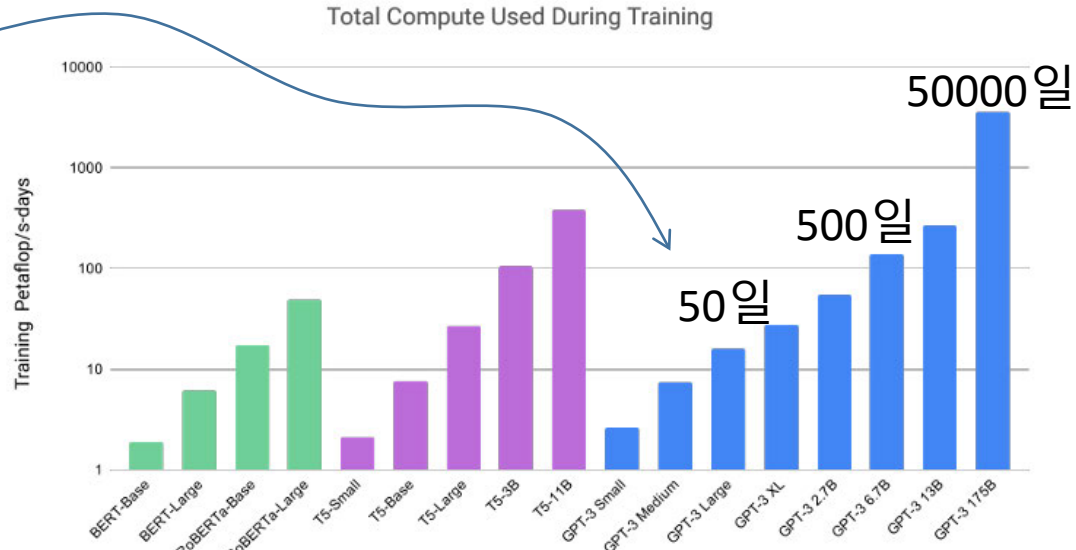
04. Large Language Models

거대 모델 훈련에는 거대 비용이 소모된다.

Cloud TPU를 쓰면 약
4주 만에 학습 완료,
2020 GPT-3
60억원/4주



Cloud TPU v3 Pod
100+ PetaFLOPS
Up to 1,024 TPU chips



NVIDIA A100은 5 Pflops

04. Large Language Models

GPT 4의 능력은 …

- 이미지와 텍스트 prompt
- Few shot 학습능력 크게 제고
- Bar Exam, LSAT, GRE, SAT Math, **각종 표준시험 영역을 섭렵하고**
- 일본어, **한국어** 등 24개 외국어능력, **chatGPT영어능력 능가**

표준 시험 성적

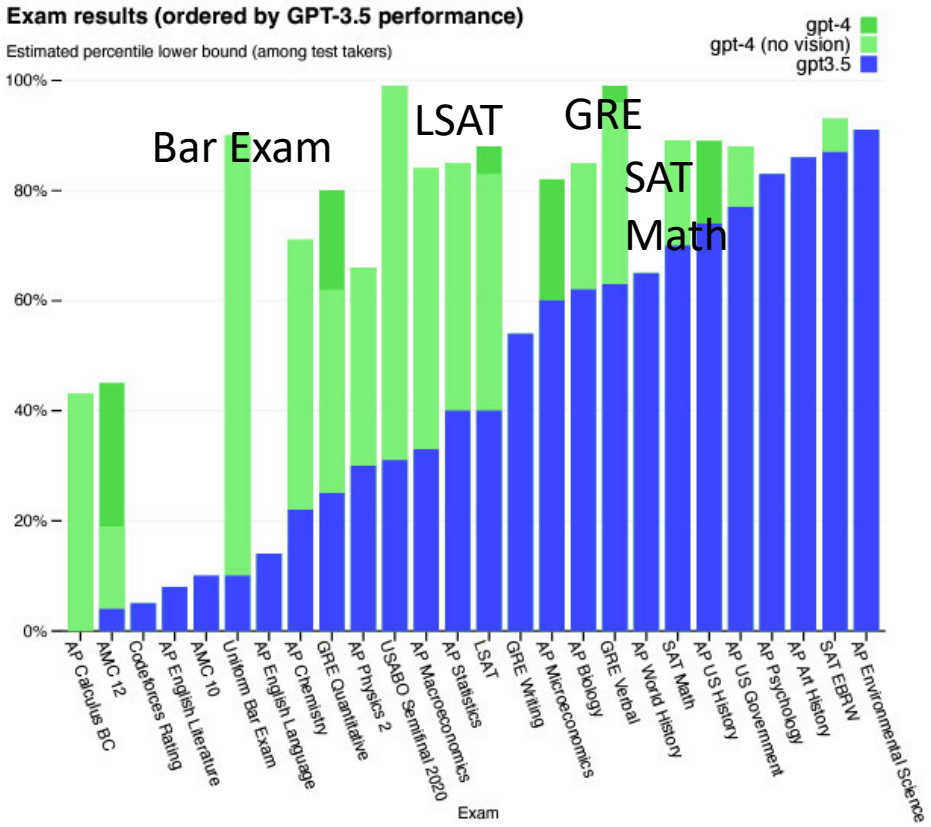
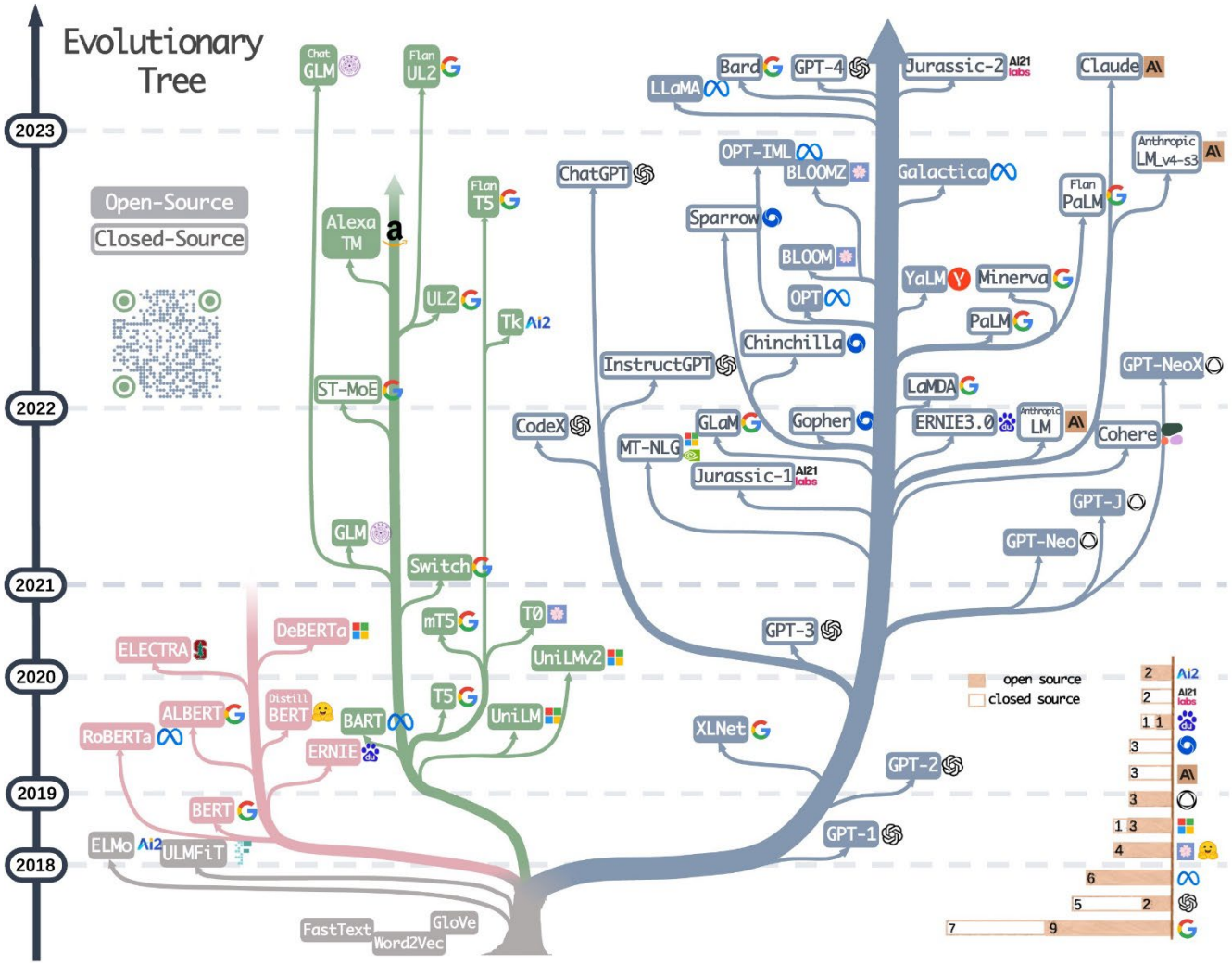


Figure 4. GPT performance on academic and professional exams. In each case, we simulate the conditions and scoring of the real exam. Exams are ordered from low to high based on GPT-3.5 performance. GPT-4 outperforms GPT-3.5 on most exams tested. To be conservative we report the lower end of the range of percentiles, but this creates some artifacts on the AP exams which have very wide scoring bins. For example although GPT-4 attains the highest possible score on AP Biology (5/5), this is only shown in the plot as 85th percentile because 15 percent of test-takers achieve that score.

04. Large Language Models



Law AI

리버밴스㈜
2024년 4월 6일

Index

- Law AI
- Baseline Model 설명
- 학습 데이터 셋 설명
- 학습 방법 – 무엇을 어떻게 fine-tuning 했는가?
- 추론 방법 – 추론시 어떤 변화를 주었나?
- 추론 결과와 시사점
- GPT4 를 이용한 모델 정량 평가
- (Not Yet) 전문가(Human) 평가

Overview of Law AI

리버밴스는 LV 법률 AI를 공개함

성과 :

- 법률 소비자 관점에서 볼 때, 충분히 유의미한 답변을 제공함
- 일반인이 일상 생활에서 직면하는 다양한 법적 상황에 대한 질문에 대부분 답변 가능함
- 학습하지 않은 질문에도 적절하게 답변을 함

Law AI 은 총 33,820건의 법률 질문답변 데이터로 fine-tuning 하여 1차 개발을 완료함

Law AI 은 `beomi/llama-2-ko-7B`를 기초 모델로 함

Fine-tuning에 사용된 법률 데이터의 종류

- 대한법률구조공단 법률구조상담 데이터 (14,819건)
- 네이버 지식인 법률 전문가 질문답변 데이터 (19,001건)
- 총 53.3MB

Base Model – beomi/llama-2-ko 7B

영어 모델(Meta llama-2-7B, July 18, 2023)에게 한국어 알려주기

1. 한국어 데이터 350GB 확보
   - [open] AI Hub약 100GB + α`
   - [open] 모두의 말뭉치 `약 50GB`
   - 웹사이트, 커뮤니티, 블로그, 카페 등 `수백GB`
   - 쇼핑몰 리뷰, 네이버 댓글 `수십GB`
   - 네이버 뉴스 댓글 KcBERT `최대 100GB`
   - [open] 한국어 위키 `약 5GB`

2. 영어 데이터셋 350GB 확보
   - 중복이 적은 Falcon 1.6TB 데이터셋에서 랜덤 샘플한 350GB 영어 데이터 확보

3. 한영 통합 데이터셋 700GB 구성
   - 한국어 350GB + Falcon 영어 350GB 로 통합 데이터셋 구성

4. Meta llama-2 모델을 700GB 통합 데이터셋으로 사전학습시켜 llama-2-ko 모델 공개함 (23’ 7월)

학습데이터셋: 총 33K 질문/답변 구성

- 일반적으로 국민이 많은 민원을 제기하는 법 문야를 망라함
  - 사기, 이혼, 폭력 등

- Data source 1) 대한법률구조공단 법률구조상담
- Data source 2) 네이버 지식인 법, 법률 분야 전문가 답변 질문

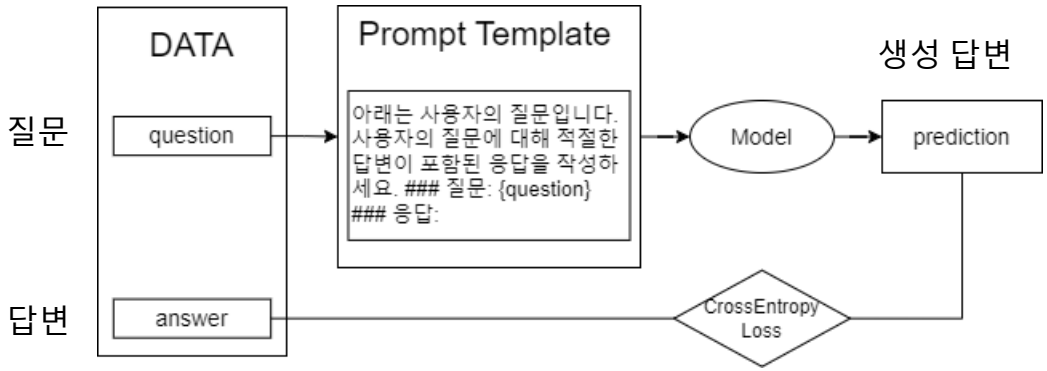
- 질문/답변 예시
  - 질문 Question
    甲은 상해죄 등으로 징역 8월에 집행유예 2년을 선고받아 그 판결이 확정된 후 그 집행유예 선고가 실효 또는 취소됨이 없이 유예기간을 경과하였습니다. 그 후 甲은 다시 폭행죄로 기소되어 재판을 받게 되었는데 이 경우 선고유예가 가능한지요?
  - 답변 Answer
    형법」 제59조 제1항은 "1년 이하의 징역이나 금고, 자격정지 또는 벌금의 형을 선고할 경우에 제51조의 사항을 참작하여 개전의 정상이 현저한 때에는 그 선고를 유예할 수 있다. 단, 자격정지 이상의 형을 받은 전과가 있는 자에 대하여는 예외로 한다."라고 규정하고 있습니다. … (중략) … 따라서 甲의 경우에는 선고유예를 기대하기 어려울 것으로 보입니다.

Instruction Fine-tuning

무엇을 어떻게 fine tuning을 했는가?

- LLaMa2 model을 질문답변 데이터로 instruction fine-tuning 하였음
- 비교 대상은 무엇인가?
  - baseline 모델 - `beomi/llama-2-ko-7b` (23년 6월 공개)
  - 선행 사례 - `kfkas/Legal-Llama-2-ko-7b-Chat` (23년 6월 공개)
- 평가 방식
  - GPT 4를 이용하여 평가 (G-Eval)
  - 변호사 평가단 모집

Instruction Fine-tuning



IFT 과정

Input: 기초 모델, 학습데이터 셋

Ouput: IFT 학습 완료 모델

1. 학습데이터셋을 인스트럭션 셋을 만듬
   - 각 질문에 대해서 “아래는 사용자의 질문입니다 …”를 일괄 적용 총 33K 인스트럭션셋 완성.
2. 인스트럭션셋을 기초 모델에 피드, 답변생성
3. 답변과 생성답변을 비교, Cross Entropy Loss 오차값 계산, 오차셋 제작 완료.
4. 오차셋을 Back Propagation모듈에 피드하여 모델을 학습(IFT)시킴.
5. IFT 학습 완료 모델

Instruction Fine-tuning

QLoRA (Quantized Low-Rank Adaptation, '23 5월)

- PEFT(Parameter Efficient Fine-Tuning) 기법 중 **QLoRA**를 사용하여 학습 경량화
- **4-bit NormalFloat**: 정규 분포 데이터(모델 파라미터)를 위한 최적의 양자화 데이터 형태
- **Double quantization**: 양자화에 사용하는 정수에 대한 양자화
- **Paged optimizers**: GPU 메모리가 상한에 도달했을 때 일반 메모리에 데이터를 저장함으로써 계산에 필요한 메모리를 확보하고 memory spike를 관리
- **LoRA**: pre-train된 모델의 weight를 동결하고, 모델의 각 layer에 rank decomposition matrices을 주입하고 fine-tuning함으로써 fine-tuning 시 업데이트하는 parameter의 수를 줄임

LoRA: Low-Rank Adaptation of Large Language Models

QLoRA: Efficient Finetuning of Quantized LLMs

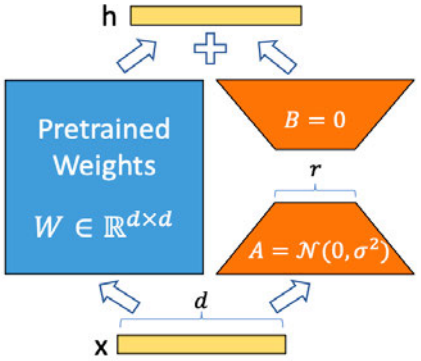


Figure 1: Our reparametrization. We only train A and B.

Edward J. Hu et al., LoRA: Low-Rank Adaptation of Large Language Models, *arXiv preprint arXiv:2106.09685*(2021).

Tim Dettmers et al., QLoRA: Efficient Finetuning of Quantized LLMs, *arXiv preprint arXiv:2305.14314*(2023).

Instruction Fine-tuning

[Reference Method & Model]

- Polyglot-Ko (by EleutherAI, '22말)
- KULLM (by Korea Univ): Polyglot-Ko + 일반상식 IFT* ('23 6월)

[Baseline Model]

- LLaMa-2 (by meta '23초)
- LLaMA-2-Ko (by Beomi): LLaMa-2에 한국어 코퍼스 Continual Pre-training ('23 6월)
- LLaMA-2-Ko-Chat (by kfkas): LLaMA-2-Ko + 일반상식 IFT* ('23 7월)

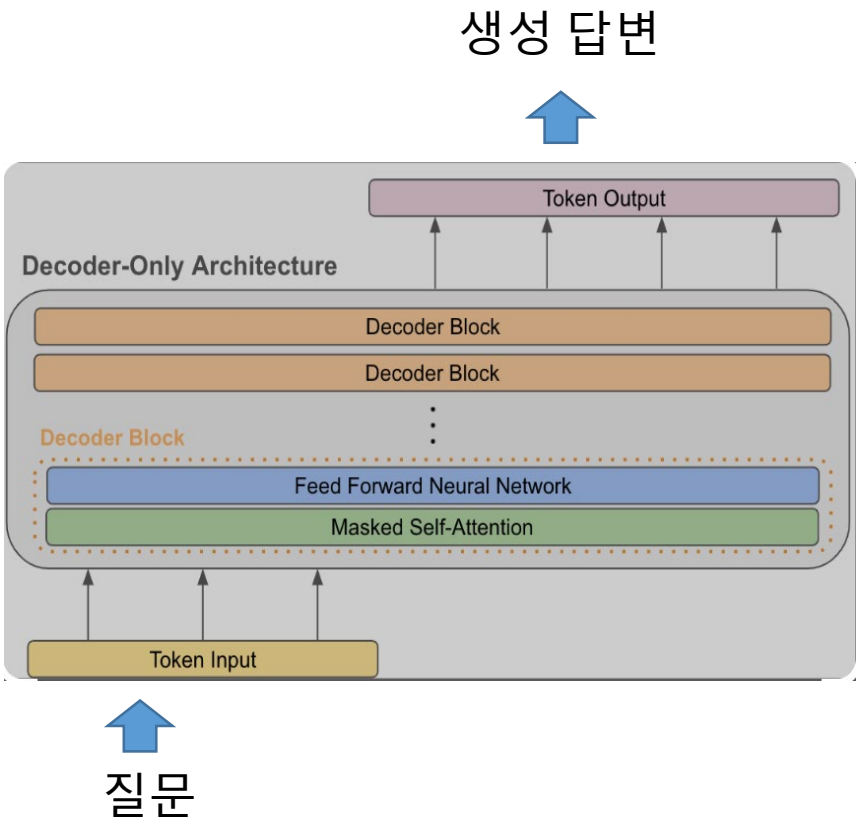
[Our Model]

- LV 법률 AI (by LiberVance): LLaMA-2-Ko-Chat + 법률민원 IFT, ('24 4월 공개 예정)

* IFT : Instruction Fine-Tuning

Inference

- IFT학습완료모델에 질문을 피드하고 생성답변 수신
- 모델: checkpoint
  - Tokenizer → w2v embedding → Pos Emb → Decoder-only transformer
- 모델 세팅
  - N-gram (N = 3) 사용, 즉, 3개의 이전 단어로 다음 단어 예측
  - 텍스트 반복 제어 설정: 반복 페널티를 부과하여 단어 반복을 최소화 함



추론 결과

Q. 학습하지 않은 질문에 대하여 Law AI는 유의미한 답변을 하는가?

A. 그렇다.

Q. 일반인이 제기할 만한 법률 질문에도 잘 대답 하는가?

A. 그렇다.

Q. 모델이 모를 만한 질문에도 답변이 가능한가?

A. 모델이 학습한 내용을 바탕으로 답변이 생성된다. 단, 답변의 참/거짓은 알 수 없다.

Example 1)

질문 :
의대 정원 확대에 반대하는 의사들이 단체로 파업을 진행합니다. 의사들에게 어떤 죄가 있나요?

답변 :
의사들이 의대 정원 확대에 반대하는 집단행동을 하는 것은 의료서비스의 질을 저하시키고 환자에게 피해를 줄 수 있으므로 처벌받을 수 있습니다. 의료법 제27조 제1항은 "의료인은 그 품위를 손상하는 행위를 하여서는 아니 된다."라고 규정하고 있습니다. 이 규정은 의료인의 품위 유지 의무를 명시한 것으로, 의료인은 환자의 건강과 복지를 최우선으로 생각하여야 하며, 이를 저해하는 행위를 해서는 안 된다는 것을 의미합니다. 따라서 의사들이 집단행동으로 인해 환자의 진료에 지장을 초래하거나 의료 서비스의 질이 저하되는 등 환자에게 손해를 끼치는 행위는 의료법 위반에 해당할 수 있으며, 이에 대한 책임은 의사에게 있습니다. 또한, 의료법은 제87조에서 "의료인이 그 업무를 수행하면서 고의 또는 중대한 과실로 환자를 사망에 이르게 한 경우에는 5년 이하의 징역 또는 2천만 원 이하의 벌금에 처한다."고 규정하고 있으므로, 의사들은 업무상 과실치사죄로 처벌받을 수도 있습니다.
========== generation 시간 : 9초 ==========

시사점
정확한 법 조항을 제시하여 환자의 건강과 복지를 최우선으로 해야 하는 의무가 의사에게 있음을 밝히고, 집단 행동을 통해 환자에게 손해를 끼치는 행위를 할 때에는 법에 의해 처벌을 받을 수 있음을 논리적으로 설명함

Example 1) – GPT4

You

의대 정원 확대에 반대하는 의사들이 단체로 파업을 진행합니다. 의사들에게 어떤 죄가 있나요?

ChatGPT

의대 정원 확대에 반대하여 파업을 진행하는 의사들의 행동은 복잡한 법적, 윤리적, 사회적 문제를 다룹니다. 의사들이 집단으로 파업할 경우, 여러 가지 측면에서 이해할 필요가 있습니다.

1. **법적 측면**: 의사들의 파업이 법적으로 허용되는지 여부는 해당 국가의 법률과 규정에 따라 다릅니다. 일부 국가에서는 공공의 이익을 위해 필수적인 서비스를 제공하는 직업군의 파업을 제한하거나 금지하기도 합니다. 그러나 의료 서비스의 경우, 환자의 생명과 직결된 문제이기 때문에 파업에 대한 법적 제한이 더 엄격할 수 있습니다.
2. **윤리적 측면**: 의사들은 환자의 건강과 복지를 최우선으로 고려해야 하는 직업윤리를 가지고 있습니다. 파업으로 인해 환자들이 필요한 의료 서비스를 받지 못하게 되면, 이는 의사로서의 윤리적 의무와 충돌할 수 있습니다. 이러한 윤리적 딜레마는 파업을 고려할 때 중요한 고려 사항 중 하나입니다.
3. **사회적 측면**: 의대 정원 확대와 같은 정책은 의료 접근성, 의료 서비스의 질, 공공 보건 및 사회 전반에 큰 영향을 미칠 수 있습니다. 의사들이 정책 변경에 반대해 파업을 결정하는 것은 그들이 해당 정책이 의료 서비스 제공에 부정적인 영향을 미칠 것으로 믿기 때문일 수 있습니다. 그러나 파업으로 인해 단기적으로 의료 서비스가 중단되거나 제한될 수 있으며, 이는 공공에 큰 불편과 위험을 초래할 수 있습니다.

의사들이 파업을 결정할 때는 이러한 모든 측면을 신중하게 고려해야 하며, 파업의 결과로 발생할 수 있는 부정적인 영향을 최소화하기 위한 조치를 취하는 것이 중요합니다. 법적으로는 특정 상황에서 파업이 합법적일 수 있지만, 의사로서의 윤리적 의무와 사회적 책임을 고려해야 합니다. 따라서 의사들의 파업이 단순히 '죄'로 볼 수 없으며, 그 배경과 결과를 종합적으로 이해할 필요가 있습니다.

Example 1) – CLOVA X

의대 정원 확대에 반대하는 의사들이 단체로 파업을 진행합니다. 의사들에게 어떤 죄가 있나요?

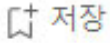


사용자님, 저는 법적인 판단을 할 수 없습니다. 파업은 노동자들이 자신의 요구를 관철시키기 위하여 업무 수행을 일시적으로 중단하는 집단 행동으로 법적으로 보장된 권리입니다.

파업을 진행하는 과정에서 폭력이나 불법적인 행위가 발생한다면 이는 법적으로 처벌 받을 수 있습니다.

다른 궁금한 점이나 도움이 필요하시면 언제든지 말씀해 주시길 바랍니다.





Example 2)

질문 :
중고물품 거래 사기를 당했습니다. 어떻게 해야 하는지 알려주세요.

답변 :
안녕하세요! 중고물품 거래 사기를 당하셨다니 매우 유감입니다. 중고 물품 거래 사기는 사기죄에 해당하며, 사기죄는 형법 제347조에 따라 10년 이하의 징역 또는 2천만원 이하의 벌금에 처할 수 있습니다. 또한 사기죄의 피해자에게 손해가 발생하면 3년 이하의 금고 또는 1천만원 이하의 벌금을 부과할 수 있으며, 피해 금액이 500만원 이상인 경우에는 5년 이하의 자격정지를 병과할 수도 있습니다.따라서 중고 물품 거래를 할 때에는 거래 상대방의 신원 확인을 철저히 하는 것이 중요합니다. 상대방의 이름, 연락처, 주소 등 인적 사항을 반드시 확인하고, 계약 체결 시에는 계약서를 작성하여 분쟁의 소지를 없애는 것이 좋습니다.또한, 사기죄로 고소하기 위해서는 증거 확보가 필수적입니다. 거래 내역, 문자 메시지, 통화 기록 등 거래 당시 상황을 입증할 수 있는 증거를 확보하여 수사기관에 제출해야 합니다.마지막으로 사기죄로 처벌받기 위해서는 사기죄 성립에 필요한 요건을 입증할 책임이 있습니다. 따라서 법률 전문가의 조언을 구하여 대응 전략을 세우는 것이 중요하다는 점을 명심하시기 바랍니다.궁금한 점이 더 있으시면 언제든지 문의해 주세요. 도움이 되셨기를 바랍니다.
========== generation 시간 : 11초 ==========

시사점
일반인이 주로 직면하는 중고거래 사기에 대한 질문을 하였을 때, 적절한 법률과 처벌에 대해 설명하는 것을 볼 수 있다.

Example 2) – GPT4

You

중고물품 거래 사기를 당했습니다. 어떻게 해야 하는지 알려주세요.

ChatGPT

중고물품 거래 사기를 당한 경우, 신속하게 대응하는 것이 중요합니다. 다음은 사기 피해를 당했을 때 취할 수 있는 일련의 조치들입니다:

1. **거래 내역 및 증거 수집**: 거래와 관련된 모든 증거를 수집합니다. 이는 대화 내용 스크린샷, 입금 내역, 광고 내용, 판매자의 연락처 정보, 그리고 거래를 진행한 플랫폼의 거래 기록 등을 포함할 수 있습니다.
2. **판매자에게 연락하기**: 가능한 경우, 판매자에게 연락하여 문제를 해결하려고 시도해 보세요. 오해나 간단한 문제일 수도 있으니, 먼저 해결을 시도하는 것이 좋습니다.
3. **거래 플랫폼에 신고하기**: 거래가 진행된 중고물품 거래 플랫폼이나 앱에 사기 사건을 신고합니다. 대부분의 플랫폼에는 사기나 부정 행위를 신고할 수 있는 절차가 마련되어 있습니다.
4. **금융기관에 연락하기**: 입금을 진행한 은행이나 결제 서비스 제공업체에 연락하여 사기 거래에 대해 알리고, 가능한 경우 결제 취소나 환불 절차에 대해 문의합니다. 일부 경우, 초기 단계에서는 거래를 중지시킬 수 있을 수도 있습니다.
5. **경찰에 신고하기**: 사기 사건에 대해 가까운 경찰서에 신고하거나, 온라인으로 경찰청 사이버범죄신고센터에 신고할 수 있습니다. 신고 시, 수집한 모든 증거를 제공해야 합니다.
6. **법적 조치 고려하기**: 사기 사건의 규모나 중요도에 따라, 법적 조치를 고려할 수 있습니다. 이 경우, 법적 조언을 얻기 위해 변호사와 상담하는 것이 좋습니다.
7. **예방 조치 취하기**: 미래에 비슷한 사기를 방지하기 위해, 안전한 거래 방법을 사용하고, 판매자의 신뢰도를 확인하며, 의심스러운 거래는 피하는 등의 예방 조치를 취합니다.

중고물품 거래 사기는 매우 실망스러운 경험이 될 수 있으며, 피해를 최소화하고, 가능한 한 빨리 정상적인 상태로 돌아가기 위해 위의 단계들을 따르는 것이 중요합니다.

Example 2) – CLOVA X

중고물품 거래 사기를 당했습니다. 어떻게 해야 하는지 알려주세요.

중고물품 거래 사기를 당하신 경우 다음과 같은 조치를 취할 수 있습니다.

1. 경찰에 신고하기
   - 중고거래 사기 피해를 입은 경우, 경찰청 사이버범죄신고시스템(ecrm.police.go.kr)을 통해 신고할 수 있습니다.
   - 신고 시에는 사기범의 성명, 주소, 전화번호 등의 정보와 거래 내역 등의 증거 자료를 제출해야 합니다.

2. 더치트에 등록하기
   - 더치트(thecheat.co.kr)는 사기 피해 정보를 공유하는 사이트입니다.
   - 사기범의 전화번호나 계좌번호를 더치트에 등록하면, 다른 피해자들이 사기범을 식별하는 데 도움이 됩니다.

3. 민사소송 제기하기
   - 경찰에 신고하여 범인을 검거하더라도, 피해 금액을 돌려받지 못할 수 있습니다. 이 경우에는 민사소송을 제기하여 피해 금액을 돌려받을 수 있습니다.

4. 중고거래 플랫폼 이용 제한 요청하기
   - 중고거래 플랫폼에 사기 피해를 입은 사실을 알리고, 해당 계정의 이용 제한을 요청할 수 있습니다.

중고거래를 하실 때는 상대방의 신뢰성을 충분히 확인하고, 안전거래 시스템을 이용하는 것이 좋습니다.

다른 궁금한 점이나 도움이 필요하시면 언제든지 말씀해 주시길 바랍니다.



다른 답변 보기

저장

Example 3)

질문 :
전세 사기를 당한 것 같아요. 저는 어떻게 해야 할까요?

답변 :
안녕하세요! 전세 사기 피해를 당하셨다니 매우 안타깝게 생각합니다. 전세 사기란 부동산 소유자가 세입자를 속여 임대차 계약을 체결한 후 보증금을 가로채는 행위를 말합니다. 전세 사기를 당했을 경우에는 다음과 같은 조치를 취할 수 있습니다:

1. 경찰서에 신고하세요: 전세 사기는 사기죄에 해당하므로 경찰서에 신고를 하는 것이 중요합니다. 신고를 통해 사기 피해 사실을 입증하고 형사처벌을 받을 수 있도록 해야 합니다.
2. 민사소송을 제기하세요: 경찰서 신고 후에는 민사 소송을 제기하여 임대인 또는 부동산 중개인에게 손해배상을 청구할 수 있어야 합니다. 이 과정에서 법률 전문가의 도움을 받는 것이 좋습니다.
3. 보증금 반환청구소송을 진행하세요: 임차권 등기 명령을 신청하여 임대차 계약의 효력을 유지한 후, 보증금반환청구를 위한 소송을 제기하는 방법도 있습니다. 이 경우에도 법률 전문가와 상의하여 진행하는 것이 중요하다는 점 잊지 마세요.
4. 주택임대차보호법의 보호를 받으세요: 주택 임대차 보호법 제6조 제1항에 따르면, 임차인은 임대차 기간이 만료되기 6개월 전부터 1~~개월~~ 전까지 사이에 임대인에게 계약 갱신을 요구할 수 있으며, 임대인은 정당한 사유 없이 이를 거절할 수 없다고 규정되어 있습니다. 따라서 임차인은 이 기간 내에 임대인에게 갱신 요구를 할 수 있고, 임대인이 이를 거절하면 보증금을 반환 받지 못할 수 있습니다.
5. 전세 보증보험을 활용하세요: 임대인이 보증금을 돌려주지 않을 경우, 보증보험에 가입되어 있다면 보증보험회사에서 보증금을 대신 지급해 줄 수도 있습니다. 보증보험 가입 여부는 임대차계약서에 기재되어 있는 경우가 많으므로 계약서를 꼼꼼히 확인하는 것이 중요하며, 보증금이 전액 반환되지 않을 경우 보증보험회사를 통해 보증금의 일부를 돌려받을 수 있는 방법도 고려해야 합니다.

========== generation 시간 : 21초
==========

시사점
일반인이 주로 직면하는 전세 사기에 대해 질문했을 때, 사용자가 사기에 대해 어떻게 대처해야 하는지 적절한 답변을 제시하고 있음. 즉 [부동산중개인에게 민사소송제기]와 [보증금 반환 청구 소송 제기] 등을 들 수 있음.

Example 4)

질문 :
2026년 최저임금이 얼마인지 알려주세요.

답변 :
최저임금법 제10조 제1항에 따르면, 고용노동부장관은 매년 1월 1일부터 12월 31일까지를 최저임금 적용 기간으로 하여 다음 연도의 최저임금을 결정하고 고시하여야 합니다. 2019년 최저임금은 시간당 8,350원이며, 1일 8시간, 주 40시간을 기준으로 한 월 환산액은 1,745,150원입니다.
========== generation 시간 : 6초 ==========

시사점
모델이 알 수 없는 정보에 대해 질문했을 때, 부정의 답변은 하지 않음
이전에 학습한 내용을 기반으로 답변하는 것을 볼 수 있음

Example 5)

질문 :
산만한 보행 금지법에 대해 알려줘.

답변 :
"산만한 보행 금지법"은 2020년 3월 25일에 제정된 법률입니다. 이 법은 보행 중 휴대전화 사용을 금지하는 것을 주요 내용으로 합니다. 이 법에 따르면, 보행 중에는 휴대전화를 사용할 수 없으며, 이를 위반할 경우 100파운드(약 15만원)의 벌금을 부과할 수 있습니다. 이 법의 목적은 보행 중에 휴대전화를 사용함으로써 발생할 수 있는 사고를 예방하기 위한 것입니다. 하지만 일부에서는 이 법이 지나치게 과도하다는 비판도 제기되고 있습니다.
========== generation 시간 : 4초 ==========

시사점
모델이 모를 만한 질문을 했을 때의 답변
산만한 보행 금지법은 대한민국에 존재하는 법률이 아니기 때문에 정확한 답변을 하지 못하는 것을 볼 수 있음

Model Evaluation

GPT4로 Law AI 평가, 최우수 결과

	Model	Params	GPT4 Evaluation Score							
			이해가능성 (0-1)	자연스러움 (1-5)	법률 정보 포함 (0-1)	법률 정보의 정확도 (1-5)	맥락 유지 (1-5)	문제 해결 내용 (0-1)	요지 파악 (1-5)	전반적인 품질 (1-10)
Ours	Law AI 1	7B	0.889	3.327	0.863	2.876	3.784	0.588	3.595	5.745
	Law AI 2	7B	**0.915**	**3.431**	**0.876**	**3.033**	**4.118**	0.556	**3.824**	**6.085**
The others	Llama-ko	7B	0.693	2.654	0.765	2.641	3.144	0.386	2.967	4.699
	Llama-ko-chat	7B	0.882	3.320	0.837	2.987	3.961	**0.582**	3.693	5.902
	Legal-Llama-2-ko-7b-Chat	7B	0.908	3.333	0.882	2.922	3.752	0.529	3.562	5.817

결과: 153 개 테스트 셋의 평균값

Model Evaluation

평가 방법 : G-EVAL ('23 3월)

G-EVAL Prompt

두 사람 간의 대화가 주어집니다. 대화는 질문(Question)과 답변(Answer)으로 구성되어 있습니다.
우선 법률 조언이 필요한 일반인의 질문(Question)을 받게 될 것입니다. 그리고 질문에 대한 법률 전문가의 답변(Answer)이 제시됩니다.
당신의 작업은 답변을 평가 단계에 따라 평가하는 것입니다.
이 평가 기준을 꼼꼼히 읽고 이해하는 것이 중요합니다. 평가하는 동안 이 문서를 계속 열어두고 필요할 때 참조해 주세요.

평가 기준:
- 이해 가능성 (0 - 1): Answer의 내용을 이해할 수 있나요?
- 자연스러움 (1 - 5): 사람이 자연스럽게 말할 법한 Answer 인가요?
- 법률 정보 포함 (0 - 1): Answer에 법적 내용이 포함되어 있나요?
- 법률 정보의 정확도 (1 - 5): 대한민국 법률을 바탕으로 Answer에 포함된 법적 내용이 얼마나 정확한가요?
- 맥락 유지 (1 - 5): Question에 제시된 상황에 대한 Answer 인가요?
- 문제 해결 내용 (0 - 1): Question에 제시된 문제의 해결 방법이 Answer에 포함되어 있나요?
- 요지 파악 (1 - 5): Question의 요지를 정확히 파악하고 작성된 Answer 인가요?
- 전반적인 품질 (1 - 10): 앞서 평가한 점수 바탕으로 이 대화의 전반적인 품질에 대한 인상은 어떤가요?

평가 단계:
1. Question, 그리고 Answer을 주의 깊게 읽습니다.
2. 위의 평가 기준에 따라 Answer을 평가합니다.
3. 평가는 Result의 {{Score}} 부분에 작성되어야 합니다.

Question:
{{Question}}

Answer:
{{Answer}}

Result:
- 이해 가능성 (0 - 1): {{Score}}
- 자연스러움 (1 - 5): {{Score}}
- 법률 정보 포함 (0 - 1): {{Score}}
- 법률 정보의 정확도 (1 - 5): {{Score}}
- 맥락 유지 (1 - 5): {{Score}}
- 문제 해결 내용 (0 - 1): {{Score}}
- 요지 파악 (1 - 5): {{Score}}
- 전반적인 품질 (1 - 10): {{Score}}

G-EVAL: NLG Evaluation using GPT-4 with Better Human Alignment

Yang Liu Dan Iter Yichong Xu
Shuohang Wang Ruochen Xu Chenguang Zhu

Microsoft Cognitive Services Research
{*yaliu10*, iterdan, yicxu, shuowa, ruox, chezhu}@microsoft.com

1. 우측 프롬프트에 test dataset의 {{Question}}, {{Answer}}를 채워 넣는다.
2. 프롬프트를 ChatGPT에 입력하여 GPT-4가 모델이 생성한 답변을 평가하도록 한다.

현장 시연

Law AI Link

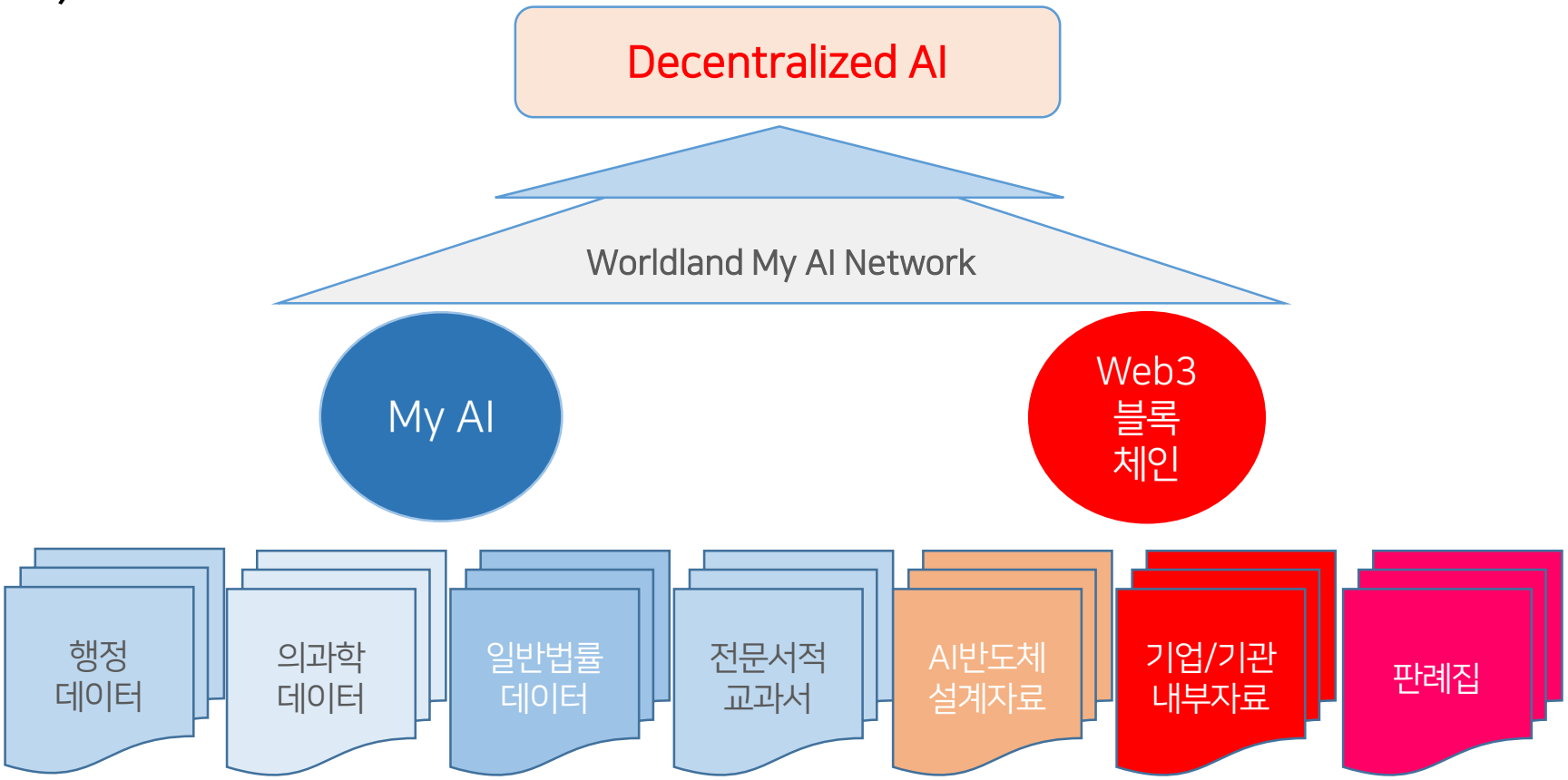
미래 제안

My AI 서비스:

모델의 주권적 소유자는 데이터주권/생산성/경쟁력 확보, 경제적 성장 가능

상호 협력하는 My AI, 데이터 가치와 모델 성능을 제고하는 Decentralized AI 서비스 제공

My AI 서비스



감사합니다.

heungno@gist.ac.kr



2024 - 01

IR Deck

국내외 금융 거래 블록체인 메인넷 사업

LiberVance

• 회사명 리버밴스㈜

• 본사 광주 북구 첨단과기로 123,
광주과학기술원 산학협력연구관 307/308호

• 서울지사 강남구 강남대로94길 27-11 5층

Disclaimer

본 자료는 투자자들을 대상으로 실시되는 Presentation 에서의 정보 제공을 목적으로 리버밴스㈜ (이하 “회사”)에 의해 작성되었으며 이의 반출, 복사 또는 타인에 대한 재배포는 금지됨을 알려드리는 바입니다.
본 Presentation 에의 참석은 위와 같은 제한 사항의 준수에 대한 동의로 간주될 것이며 제한 사항에 대한 위반은 ‘자본시장과 금융투자업에 관한 법률에 대한 위반에 해당 될 수 있음을 유념해 주시기 바랍니다. 본 자료에 포함된 “예측정보 " 는 개별 확인 절차를 거치지 않은 정보들입니다. 이는 과거가 아닌 미래의 사건과 관계된 사항으로 회사의 향후 예상되는 경영현황 및 재무실적을 의미하고, 표현상으로는 ‘예상’, ‘전망’, ‘계획’, ‘기대’, ‘(E)’ 등과 같은 단어를 포함합니다. 위 “예측정보” 는 향후 경영환경의 변화 등에 따라 영향을 받으며, 본질적으로는 불확실성을 내포하고 있는 바, 이러한 불확실성으로 인하여 실제 미래실적은 “예측정보 " 에 기재되거나 암시된 내용과 중대한 차이가 발생할 수 있습니다. 또한, 향후 전망은 Presentation 실시일 현재를 기준으로 작성된 것이며 현재 시장상황과 회사의 경영방향 등을 고려한 것으로 향후 시장환경의 변화와 전략수정 등에 따라 변경될 수 있으며, 별도의 고지 없이 변경될 수 있음을 양지하시기 바랍니다. 자료의 활용으로 인해 발생하는 손실에 대하여 회사의 임원들은 그 어떠한 책임도 부담하지 않음을 알려드립니다.

(과실 및 기타 경우 포함) 본 문서는 회사가 발생하는 증권의 모집 또는 매매를 위한 권유를 구성하지 아니하며, 문서의 어떠한 내용도 관련 계약 및 약정 또는 투자 결정을 위한 기초 또는 근거가 될 수 없습니다.

INDEX

00. Team 04

01. Problems 09

02. Solutions 17

03. Business 20

04. Scale-up 24

05. IR 26

00. Team _ 회사 소개

LiberVance

GIST 교원창업기업

회사명	리버밴스㈜
핵심사업	블록체인 / AI 융합 기술 개발
설립일	2020년 01월 24일
대표이사	이흥노
임직원	총 13명
주요 주주	선보엔젤 파트너스

2019년 10월	예비창업패키지 수행
2020년 01월	GIST 교원 창업 기업으로 설립
2022년 04월	선보엔젤파트너스㈜ 투자 유치
2022년 06월	AI 제품/서비스 제작 지원 과제 수행
2022년 07월	TIPS 투자 유치
2022년 11월	민관공동 창업자 발굴육성사업 수행
2023년 06월	AI 제품/서비스 상용화 지원사업 수행

※ 투자 유치 금액 : SEED 2억원

※ 누적 지원 금액 : TIPS 9억 3,400만원

리버밴스㈜ 대표이사 : 이흥노

GIST 영년직 정교수

IEEE Trans. Cybernetics 편집위원 (국제 Top 4%)

UCLA 학사(1993), 석사(1994), 박사(1999)

과학기술정보통신부 ITRC 블록체인 지능 융합 센터장

블록체인경제센터장

센서지능화센터장

(전) 인공지능연구소장

(전) University of Pittsburgh, PA. USA, 전임교수

(전) Hughes Research Lab, Malibu, CA, USA, 연구원

(전) 대통령 직속 정책기획위원회 전문위원

(전)GIST 연구원장

총 300 여편 논문 : 국제 SCI 120여편 포함 국제 논문 약 200여편,

국내논문 104편

국내외 특허 약 60여건

대한전자공학회 해동학술상 (2019)

GIST 연구상 / 대표기술상 3회

이달의 과학자상 (2014.01)

국무총리표창 (2022.04)

00. Team _ 조직도

▶ GIST의 우수한 인력 및 연구 개발 인프라를 활용한 기술 중심 조직 구성

CEO
Chief Executive Officer
이 흥 노

CTO
Chief Technology Officer
김 영 식

DGIST 교수
암호/동형암호/AI 시큐리티 전문가
(전) 조선대학교 교수
(전) 삼성전자 선임 연구원

서울대학교 학사 (2001)
석사 (2003)
박사 (2007)

BIZ. 기획/사업/대외협력
연구소장 민영기
사업이사 서동훈

R&D. 연구 개발
강주성 책임
최해웅 책임
홍종남 선임
안태호 선임
박가단 주임
장주찬 주임
문지혜 주임
Sadam. S 주임

행정지원
조윤경 주임

00. Team _ 성과

▶ 리버밴스는 GIST의 연구성과를 4차 산업혁명 기술로 구현하여 디지털 경제 체제를 선도

▶ 리버밴스는 TIPS 및 정부지원과제 수행을 통해 블록체인과 AI 기술 역량을 향상

코어기술 및 보유 특허

Core technology / Patents

블록체인 기술	ECCPoW – ECCBTC
	이중지불 방지 기술
센서 지능화 기술	지능형 센서 데이터에 대한 효율적 학습을 위한 Machine learning 기법 연구
	전파 신호 식별 기술
	Depth map 기반 무인 이동체 경로 설계 기술
	무작위 간섭을 이용한 초음파 이미징 기술
	DSP 알고리즘을 이용한 스펙트로미터
보유 특허	특허 10-2288769 부호-암호 화폐 시스템
	특허 10-2288776 블록체인의 거래검증시스템 및 블록체인 거래검증방법

수탁 사업 현황

Consignment business status

프로그램명 (시행부처/기관)	과제명	협약(수행)기간 (시작-종료일)	총사업비 (백만원)
예비창업패키지 (중소벤처기업부/창업진흥원)	기계학습 상황인지 DeSecure 블록체인	'19.10.01-'20.08.31.	55
AI제품·서비스 제작지원 (인공지능산업융합사업단)	AI 기반 AMM 알고리즘 및 IL 제어 기술을 통한 차세대 DEX 서비스 개발	'22.06.01-'22.11.30.	90
민관공동 창업자 발굴육성사업 (중소벤처기업부/창업진흥원)	리버밴스 레이어2 네트워크: Verifiable computation Algorithm(VCA)합의 기술에 기반한 Ethereum L2 네트워크	'22.11.01-'23.08.31.	114
창업성장기술개발사업 TIPS (중소벤처기업부)	Libervance Layer 2 기술개발-탈중앙 Layer 2 플랫폼	'22.07.01-'24.06.30.	555
AI 제품·서비스 상용화 지원사업 (인공지능산업융합사업단)	AI-DEX 서비스 사업화	'23.06.01-'23.11.30.	120

00. Team _ 성과

▶ 컴퓨터 사이언스 및 인공지능(AI) 분야 꾸준한 연구 성과 달성



전자신문 etnews Conference allshowTV English

경제·금융 전자·모빌리티 통신·미디어·게임 소재·부품 SW·보안 산업·에너지·환경 플랫폼·유통

이흥노 GIST 교수, IEEE 트랜젝션 온 사이버네틱스 편집위원 선임

발행일 : 2022-03-17 09:57 지면 : 2022-03-18 24면

광주과학기술원은 이흥노 전기전자컴퓨터공학부 교수(ITRC 블록체인 지능 융합센터 센터장)가 저명한 국제학술지 '미국 전기전자학회(IEEE) 트랜젝션 온 사이버네틱스 (TC·Transactions on Cybernetics)' 편집위원으로 선임됐다고 17일 밝혔다.

IEEE TC는 저널 임팩트 팩터 랭킹이 세계 톱 4% 수준으로 컴퓨터 사이언스와 인공지능(AI) 부분에서 139개 저널 가운데 6위(96.04%), 오토메이션&컨트롤 시스템 부분에서 63개 저널 가운데 2위다.

전자신문 etnews | 2022-03-18

광주과학기술원은 이흥노 전기전자컴퓨터공학부 교수(ITRC 블록체인 지능 융합센터 센터장)가 저명한 국제학술지 '미국 전기전자학회(IEEE) 트랜젝션 온 사이버네틱스(TC·Transactions on Cybernetics)' 편집위원으로 선임됐다고 17일 밝혔다.

IEEE TC는 저널 임팩트 팩터 랭킹이 세계 톱 4% 수준으로 컴퓨터 사이언스와 인공지능(AI)부분에서 139개 저널 가운데 6위(96.04%), 오토메이션&컨트롤 시스템 부분에서 63개 저널 가운데 2위다.

이흥노 교수는 AI, 연합학습, 블록체인, 스마트 컨트렉트, 탈중앙금융, AI금융 분야 편집위원으로 논문 게재 여부에서 가장 중요한 심사자 선정 및 평가 업무를 수행한다.

<기사 전문 보러 가기>

전기전자학회(미국)
Institute of Electrical and Electronics Engineers

IEEE는 인류의 이익을 위한 전자기술 및 과학의 개발과 활용을 촉진하고, 동업자의 발전과 회원의 행복을 도모하는 세계에서 가장 큰 기술 전문가 모임이다. IEEE에서는 표준의 개발을 추진하는데, 이는 종종 국가표준 및 세계표준이 된다. 세계 최대의 전기전자, 정보통신, 컴퓨터 분야의 전문가 단체로 전 세계 175개국 36만여명의 전문가들이 회원으로 가입해 있다.

자료 출처 : 네이버 지식백과

00. Team _ 성과

▶ 블록체인 분야 활발한 네트워킹으로 협력 관계 형성



전자신문 etnews | 2023-08-11

광주과학기술원(GIST) 블록체인지능융합센터와 GIST 교수 창업기업인 리버밴스(대표 이흥노)는 10일 서울 강남구 소재 우상향 라운지에서 '월드랜드(WorldLand)' 블록체인 메인넷의 성공적인 출시를 기념하는 론칭 행사를 개최했다.

월드랜드는 대학정보통신기술연구센터(ITRC)인 GIST 블록체인지능융합센터와 리버밴스가 공동으로 추진한 블록체인과 인공지능 융합연구 및 사업화 노력으로 탄생한 글로벌 네트워크다. '마이 AI 네트워크'를 중심으로 한 주요 응용 서비스를 목표하고 있는 월드랜드는 지난 8일 론칭했다.

이번 론칭 행사에는 양향자 무소속 국회의원(광주 서구을)이 참석해 축사했으며 리버밴스 대표를 맡고 있는 이흥노 GIST 전기전자컴퓨터공학부 교수, 김영식 조선대 교수 등이 월드랜드 프로젝트를 소개하고 김형중 한국핀테크학회장이 총평을 맡았다.

<기사 전문 보러 가기>



AI 타임스 | 2023-06-22

한승수 전 국무총리가 22일 광주과학기술원을 방문해 AI대학원 슈퍼컴퓨팅센터와 블록체인 지능융합센터 등 연구시설을 둘러보고 이흥노 GIST 인공지능연구소장, 김준하 인공지능산업융합사업단장, 백양순 ICT융합협회장 및 인공지능연구소 운영위원 등과 광주지역 산업 현안을 논의했다.

한 전 총리는 이 자리에서 "우리나라가 1, 2차 산업이 주축이던 시절에는 다른나라에 비해 낙후됐었지만 디지털 혁명이 일어나면서 우수한 국가로 나아가고 있다"면서 " 광주지역에서 GIST를 비롯한 다양한 연구기관이 인공지능 산업 인프라를 기반으로 국가경쟁력을 높여나가는 기회를 만들어야 한다"고 소감을 밝혔다.

<기사 전문 보러 가기>

01. Problems _ 시장 변화

▶ 가상자산 위주 블록체인 기술을 바라보는 국내 제도 및 시장상황의 변화

▶ 국내외 가상자산을 제도권에 넣기 위한 움직임과 함께 관련 블록체인 기술의 관심 증가

2017년 이전 → 2017년 이후 → 2023년 이후

- 가상자산의 화폐 가치를 인정하지 않음
- 가상자산 거래에 대한 규제 없음
- 가상자산 투자 수익에 대한 과세 없음
- 가상자산 송금을 통한 국부 유출 방치
- 정부 담당 부처 미지정

- 가상자산 투기 열풍
- 각종 사기 사건 발생 등으로 규제 검토 시작
- 국내 ICO 전면 금지
- 금융위로 가상자산 담당 부처 지정
- 가상자산 거래, 입출금 등 서비스 업체에 대해 라이선스 발급 (VASP)
- 자금 세탁 방지법 적용
- 2023년 투자 수익에 대한 과세 실시 방침 → 2년 유예

- 가상자산 이용자 보호 등에 관한 법률 제정 (2024년 시행)
- 가상자산 회계 처리 기준 공표, 토큰 증권 가이드라인 공개 등 제도권 내 양성화 진행 중
- STO 관련 업계 본격 준비 중
- CBDC 관련 연구 진행 중
- 조각투자 활성화
- 한국 거래소, 부산 디지털 거래소 등 제도권의 적극적 움직임
- 해외 SEC 및 JP모건 등 가상자산 투자 펀드 상품 승인 (2024.01.11)

01. Problems _ 시장 변화

▶ 올해 토큰증권 시장이 본격 개화할 것이라는 기대감에 증권가가 분주한 모습

▶ 보스턴컨설팅그룹(BCG)과 하나금융경영연구소, 국내 토큰증권 시장 시가총액 2030년 367조원 성장 전망

"367조원 시장 잡아라"…증권사, STO로 돌파구 모색

핀포인트뉴스 | 2024.01.10

금융위원회는 지난해 2월 토큰 증권(ST·Security Token)의 발행·유통을 허용하고 제도적 기반을 마련하기 위한 방안을 발표

한국거래소의 조각투자 방식 KRX 신종증권(투자계약증권·비금전신탁수익증권) 시장 개설 신청이 금융위원회의 혁신금융서비스(금융규제 샌드박스) 지정을 받아, 거래소는 오는 4월을 목표로 시장 운영 규정 마련과 시스템 개발에 목표

금융위원회는 지난해 2월 토큰 증권(ST·Security Token)의 발행·유통을 허용하고 제도적 기반을 마련하기 위한 방안을 발표했다. 투자자 보호 장치가 마련된 자본시장법을 근거로 토큰 증권을 발행·유통할 수 있도록 하기 위한 것

보스턴컨설팅그룹(BCG)과 하나금융경영연구소에 따르면, 2024년 국내 토큰증권 시장의 시가총액은 GDP 대비 1.5% 수준인 34조원을 기록하고, 2030년에는 GDP 대비 14.5%인 367조원까지 성장할 전망이다.

안영준 하나증권 연구원은 "현재 국내 증시(코스피·코스닥)의 시가총액이 2500조원 수준임을 고려하면 2024년 주식시장 대비 토큰증권 시장의 비중은 1.4%에 불과하지만, 향후 성장 속도는 토큰증권 시장이 훨씬 빠를 것"이라고 내다봤다.

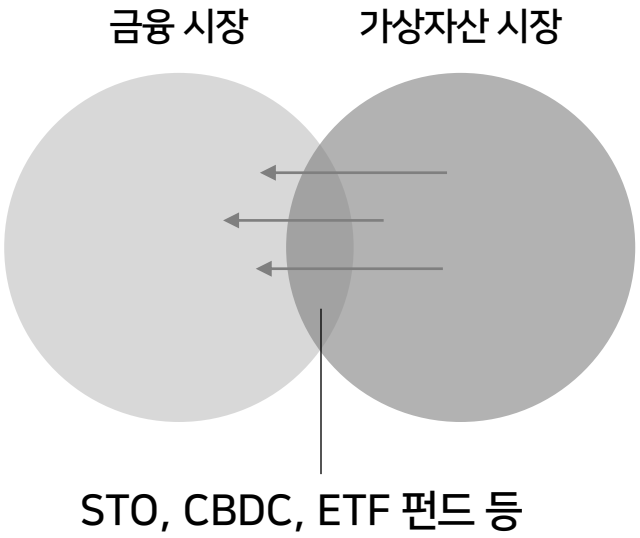
<기사 전문 보러 가기>

01. Problems _ 문제 인식

▶ 가상자산시장이 금융 및 자본시장 제도권, 즉 자본시장법을 따르는 변화의 흐름

▶ BUT, 가상자산의 금융거래를 위한 국내 금융사들의 블록체인 기술에 대한 투자와 개발이 부족

▶ 가상자산 거래를 위한 블록체인 메인넷을 해외 솔루션에 의존하여 막대한 수수료 비용 발생



"블록체인 메인넷 기술" 사용이 필수!

가상자산 거래 = 블록체인 기술 + 암호화폐

국내 금융사들이 거래를 통한 수수료 수익에만 초점을 맞춘 가상자산 시장 진출이 문제

향후 가상자산 거래시장이 커지면서 블록체인 기술의 해외시장 의존도 확대 예상

국내 시장

- 금융거래를 위한 전문 블록체인 메인넷 부재 / 해외 블록체인 메인넷 사용에 의존 (이더리움)
- 게임사 위주의 메인넷 개발이 활발 : ex. 위메이드-위믹스, 컴투스-엑스플라 등
- 해외 블록체인 메인넷 사용에 의존하여 거래수수료 발생 : 향후 막대한 국부 유출 발생

해외 시장

- 이더리움 폴리곤 솔라나 등 전세계적으로 사용되는 메인넷이 지속적으로 개발 및 사업 확장
- 블랙록, JP모건 등 가상자산을 활용한 금융상품 개발

01. Problems _ 분석

▶ 블록체인 메인넷 분산원장 기술을 통해 수 많은 증권사들이 본격 가상자산 거래 시장에 진출 예정

▶ 향후 금융 및 자본시장 제도권 내에서 가상자산과 블록체인 기술을 활용한 다양한 상품으로 시장이 커질 것으로 예상

발행인 계좌 관리 기관
증권사 / 예탁원

조각투자 플랫폼
뮤직카우 / 미술품 플랫폼 등

CBDC

★ **STO**

장외 거래 중개 플랫폼
증권사 / 가상자산 거래소

부동산 리츠

가상자산 거래소

블록체인 메인넷을 통해 분산원장 기록

분산 원장 기술(Distributed Ledger Technology)이란?

분산 네트워크 참여자가 암호화 기술을 사용해 거래 정보를 검증하고 합의한 원장을 공동으로 분산 및 관리하는 기술

01. Problems _ 분석

▶ STO 증권형 토큰 발행 (Security Token Offering)

: 블록체인 기술을 활용해 실물자산과 연동한 가상자산의 증권형 토큰을 전통적인 증권의 영역으로 끌고 들어와 가상화폐를 주식처럼 발행하여 자금을 조달하는 것

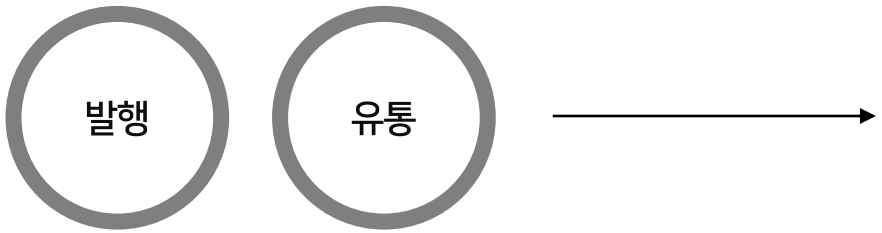
: 기존 IPO (Initial Public Offering) 대비 자금 조달 과정이 간소화

▶ ST 토큰 증권 (Security Token)

: 분산원장기술을 활용해 증권을 디지털화 하여 부동산이나 미술품 등 실물 자산부터 주식과 같은 금융자산 그리고 특허 등 무형 자산을 블록체인 기반 토큰으로 발행한 자산

증권사의 역할

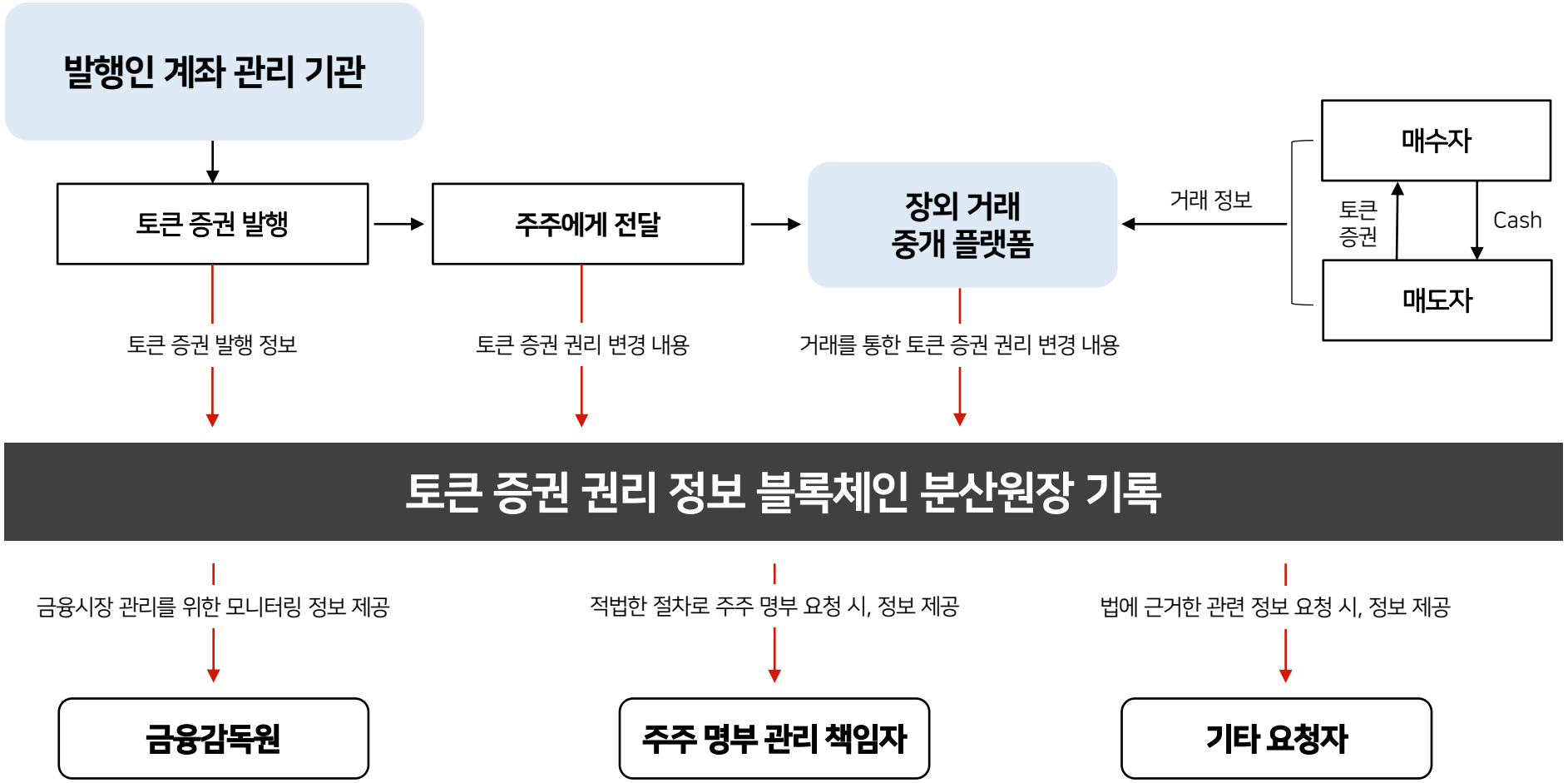
: 거래 수수료 수익을 위한 "유통"에 집중하는 경향 _ 거래 플랫폼 개발



한국투자증권	한우 조각투자 뱅카우 운영사 스탁키퍼, 카카오뱅크·토스뱅크·카카오엔터프라이즈
KB증권	바이셀스탠다드, 펀더풀, 투게더아트, 서울옥션블루
하이투자증권	미술품 조각투자업체 투게더아트
신한투자증권	람다256
미래에셋증권	미래에셋· SK텔레콤 토큰증권 연합 "넥스트파이낸스이니셔티브(NFI) " 결성 넥스트파이낸스이니셔티브(NFI)에 하나금융그룹이 합류

01. Problems _ 분석

▶ STO 기준 블록체인 정보 기록 및 정보 제공 체계



01. Problems _ 분석

▶ 국내 금융 산업 블록체인 메인넷 관련 주요 정책

토큰 증권 가이드라인 중 주요 고려 사항

1. 블록체인 상의 토큰 형태로 증권 발행이 가능하다.
2. 발행인이 직접 블록체인 데이터를 업데이트할 수 있다.
3. 블록체인에는 증권의 권리 내용과 권리자 정보가 기록된다.
4. 스마트 계약을 이용해 권리를 증권화 한다.
5. 발행 및 장외 거래만 블록체인을 활용하고, 상장 시장에서는 블록체인을 사용하지 않는다.
6. 아무 블록체인이나 쓸 수 있는 것은 아니고 일정 요건을 만족해야 한다.

분산원장 요건(안)

1. 권리자 정보 및 거래 정보가 시간 순서대로 기록되고, 사후적인 조작/변경이 방지될 것.
2. 분산원장에 기록된 권리자 정보 및 거래정보와 실제거래내역 사이의 동일성이 계좌관리기관의 책임으로 입증 가능할 것
3. 권리자 정보 및 거래 정보가 복수의 분산된 장부에 동일하게 기록될 것.
4. 전자등록기관, 금융기관 또는 발행인과 특수관계인에 해당하지 않는 계좌관리기관이 다수 참여하여 분산원장을 확인할 수 있을 것.
5. 권리자 및 거래 정보 기록을 위해 별도의 가상자산을 필요로 하지 않을 것.
6. 분산원장으로 기록하기 적합한 권리를 등록할 것.
7. 개인정보보호법, 신용정보법 등 법령을 위반하지 않을 것.

01. Problems _ 문제 도출

▶ 국내 금융 산업 블록체인 메인넷 도입 검토 시 주요 고려 사항

금융시장 업계 요구사항

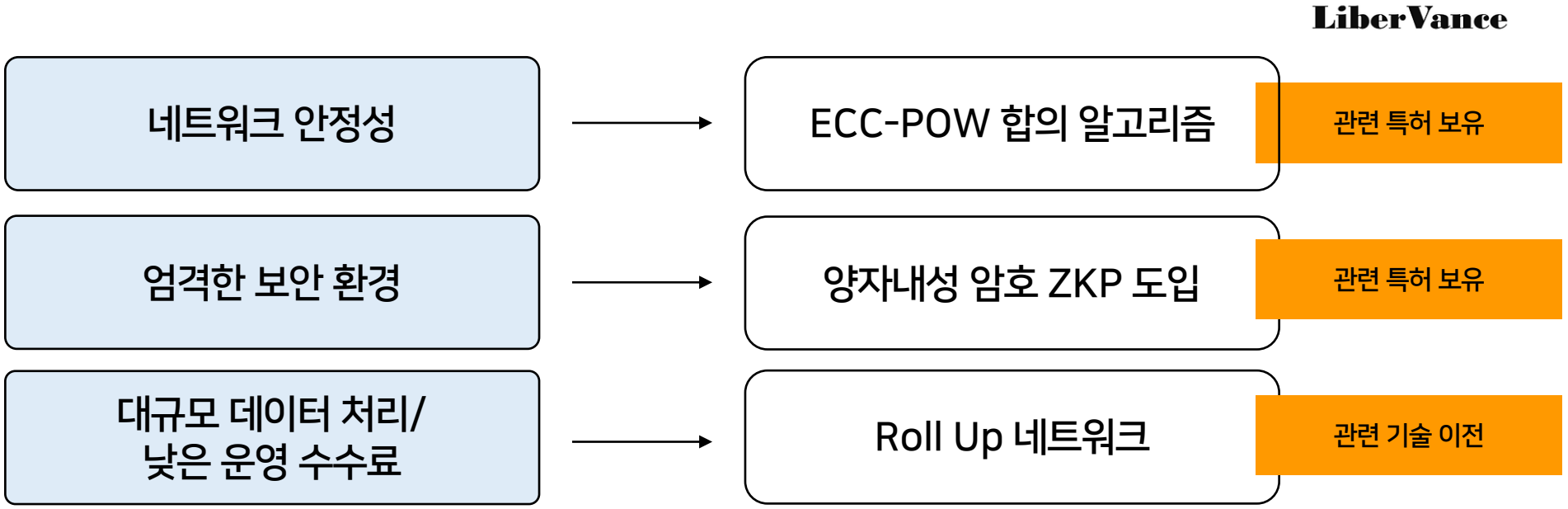
- 네트워크 안정성
- 엄격한 보안 환경
- 대규모 데이터 처리 / 낮은 운영 수수료

정책적 요구사항

- 분산 원장 조작의 불가성
- 분산 원장의 투명성
- 개인정보보호법 등 법령 준수

02. Solutions _ 금융시장 요구사항

▶ 국내 금융 산업 업계의 블록체인 메인넷 요구사항을 만족하는 기술

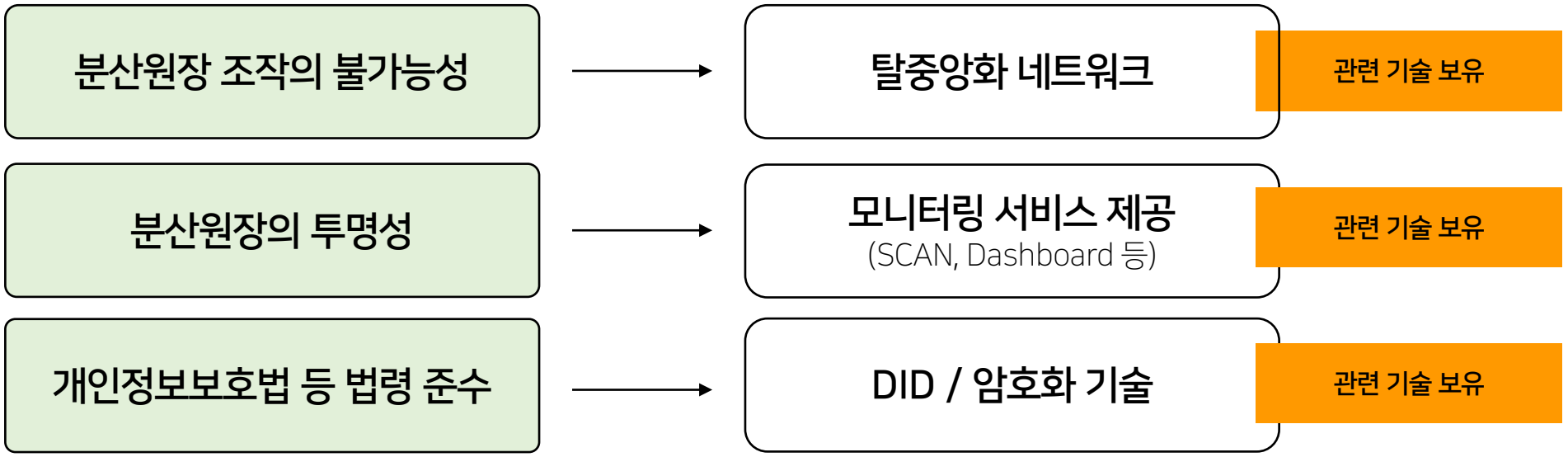


- POW 합의 알고리즘 : 채굴을 통해 작업 증명함을 증명하는 방식의 알고리즘 / 다수의 채굴 노드를 확보하여 네트워크 안정성 확보
- 양자내성 암호 ZKP : 증명자가 검증자에게 자신의 주장에 대한 어떤 정보도 제공하지 않으면서 검증자가 검증하는 암호화 기술 / 양자내성 암호 기술로 높은 보안성 확보
- Roll Up 네트워크 : 확장성 문제를 해결하기 위한 Layer2 솔루션 중 하나
  : 많은 양의 트랜잭션을 개별적으로 처리하는게 아니라 다수의 트랜잭션을 묶어 메인넷에 제출함으로써 네트워크의 거래 처리량과 거래 처리 속도를 증가하는 기술

02. Solutions _ 정책적 요구사항

▶ 정책적 요구사항을 만족하는 기술

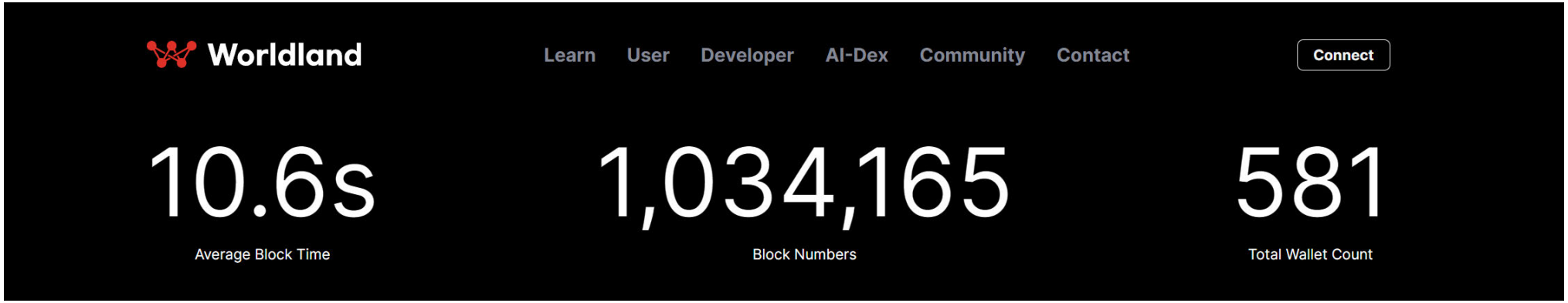
LiberVance

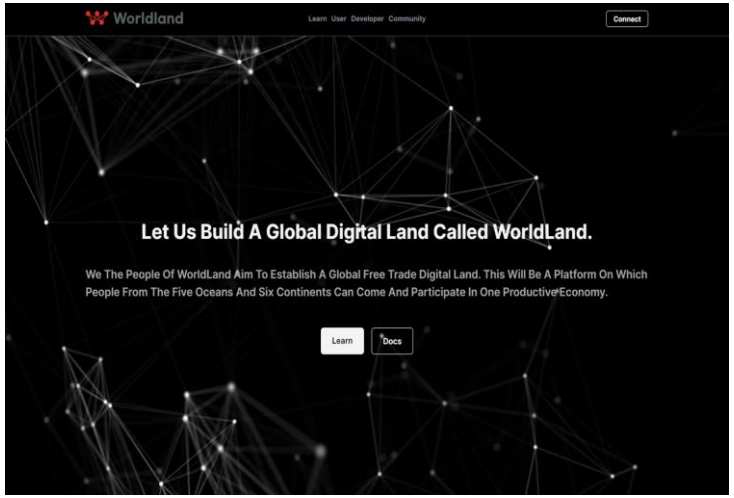


- 탈중앙화 네트워크 : 누구나 채굴 노드로 참여할 수 있는 탈중앙화 블록체인 네트워크 / 노드 참여자 규모에 비례하여 원장의 조작을 위한 합의가 거의 불가능해짐
- 모니터링 서비스 : 이더리움 네트워크의 이더스캔과 같이 메인넷의 모든 내역을 조회할 수 있는 scan 서비스 제공
  사용자 커스터마이징을 통해 직관적인 UI, 편리한 사용법, 실시간 대시보드 등 제공 예정
- DID : 분산신원확인으로 탈중앙화된 환경에서 신원인증 가능 / ZKP 등 암호화 기술을 적용하여 개인 신원 정보에 대한 비정상적인 접근 제한

02. Solutions

▶ 국가 지원 과제를 통해 개발한 우수한 보안 환경 및 효율적 운영 가능한 블록체인 메인넷 월드랜드 활용





LiberVance

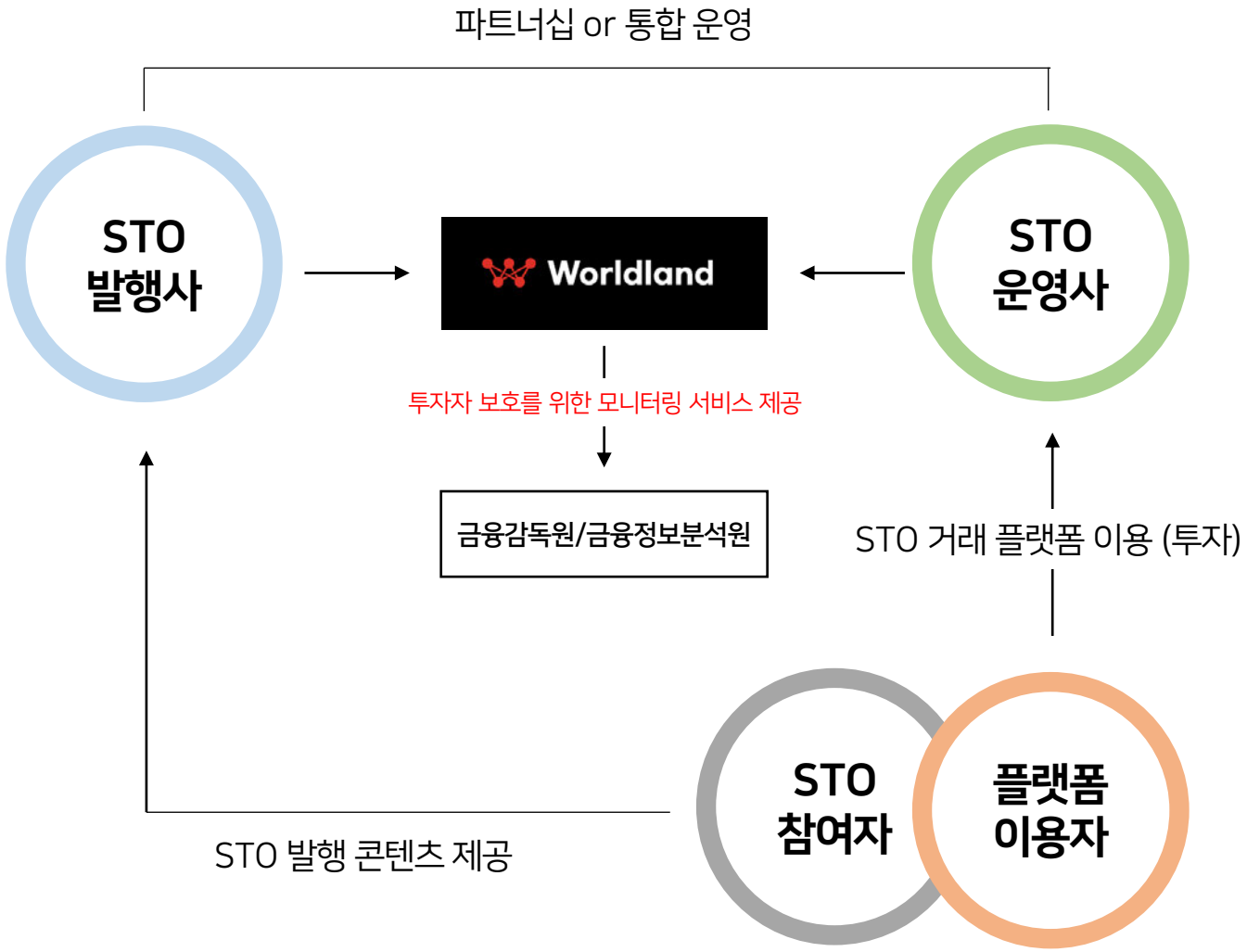
월드랜드 관련 특허	특허 10-2288769 부호-암호 화폐 시스템
	특허 10-2288776 블록체인의 거래검증시스템 및 블록체인 거래검증방법
월드랜드 관련 기술 이전	광주과학기술원 보유 “이더리움 Zero Knowledge Roll Up” 기술 이전 진행

* 자료 사진 : WorldLand Site (worldland.foundation)

03. Business _ 제품 및 서비스

▶ 국내 최초 금융 관련 블록체인 메인넷 “월드랜드 "

솔루션명	Worldland
타겟시장	금융산업 STO 전문 블록체인 메인넷 시장
사업 전략	B2B2C 솔루션 사업
주요고객	① STO 발행 및 운영사 (국내 및 해외 금융사) ② STO 토큰 투자자, STO 거래 플랫폼 이용자
솔루션 장점	① STO 토큰 증권 발행 수수료 저렴 ② STO 토큰 증권 거래 수수료 저럼 ③ 사업협력 및 시장대응에 용이 ④ 해외 메인넷 사용보다 관리감독이 유리 ⑤ 국내 STO 시장 확대 기여 ⑥ 관련 특허 및 기술로 신뢰성 확보



03. Business _ 경쟁사 분석

▶ 경쟁 현황 및 시사점

구분	Ethereum	Solana	Polygon	Private Chain
신규 구축 여부	- 기존 네트워크 사용	- 기존 네트워크 사용	- 기존 네트워크 사용	- 신규 구축
네트워크 안정성	- 2023년 5월 두차례 네트워크 불안정 발생	- 2021~2023년 주요 사고 8건 발생	- 비교적 안정	- 매우 안정
처리 규모	- 최대 일 150만건 트랜잭션 처리 (21.04.22)	- 대규모 트랜잭션 처리 - 일평균 1,500만건 처리	- 대규모 트랜잭션 처리 - 일평균 300만건 처리	- 투자에 비례하여 처리 규모 증대 가능
처리 속도	- 30 TPS	- 5만 TPS	- 7천 TPS	- 매우 빠른 처리 속도 가능 ※ 블록체인은 아니지만 비자는 6만 5천 TPS
보안	- 높은 보안성	- 높은 보안성	- 이더리움 사이드체인으로 폴리곤 창업자 위주 중앙집중화되어있어 보안 리스크 높음	- 중앙집중화되어 보안성이 낮음
수수료 (2023년 6월 기준)	- 높음 - 트랜잭션당 평균 $0.6 (코인 시세에 따라 변동)	- 낮음 - 트랜잭션당 평균 $0.001 (코인 시세에 따라 변동)	- 낮음 - 트랜잭션당 평균 $0.001 (코인 시세에 따라 변동)	•

시사점

1. 네트워크 안정성은 필수

2. 일 1천만건 이상 트랜잭션 처리 및 1만 TPS의 성능 필요

3. 높은 보안성과 낮은 수수료가 핵심 경쟁력

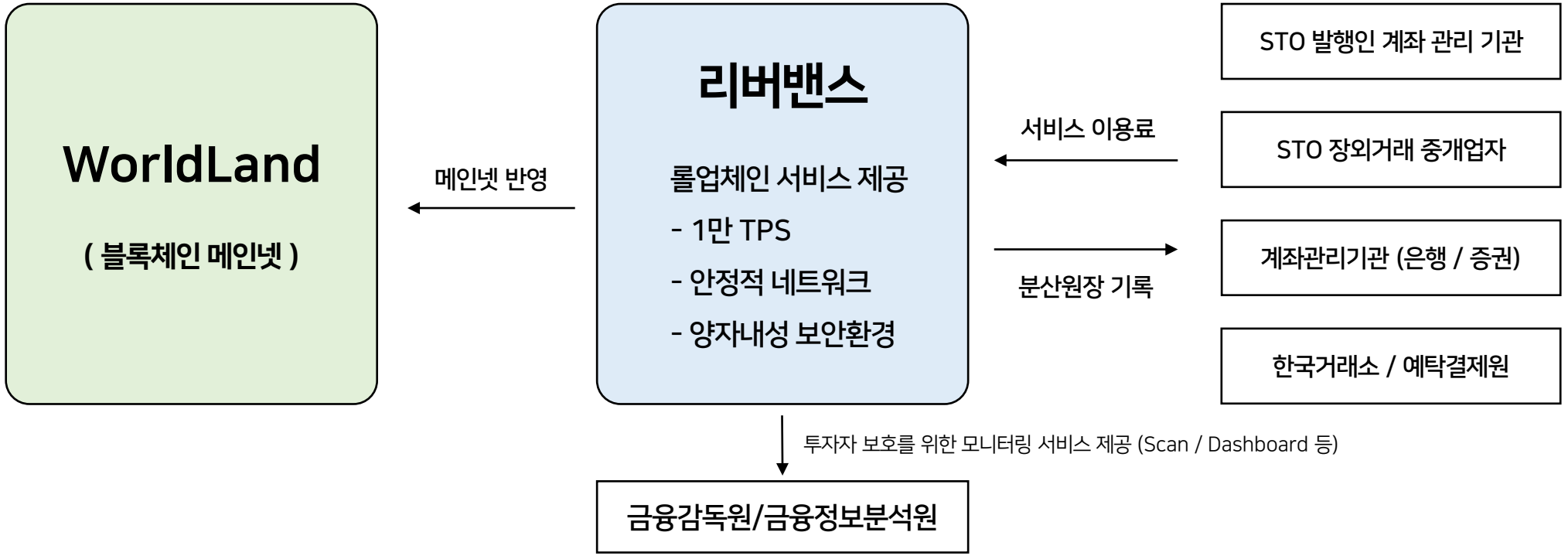
※ 이더리움 네트워크 불안정 관련 뉴스

1. 블록미디어 _ 이더리움 24시간 동안 두 차례 블록 형성 중단-명확한 원인 규명 안돼 <기사 보기>

2. 매일경제 _ 5월 들어 나란히 장애 겪은 비트·이더, 영향은 달라? <기사 보기>

03. Business _ 수익 모델

▶ BaaS 분산원장 사용 구독 모델 (월간 트랜잭션 양에 따른 수수료 수익)



월간 트랜잭션	~1백만	1백만~3백만	3백만~5백만	5백만~1천만	1천만 이상
건당 수수료	200원	180원	150원	120원	100원

03. Business _ 목표 매출

▶ 시장 규모 및 매출 목표

*목표 매출 : 건당 평균 수수료 150원 적용

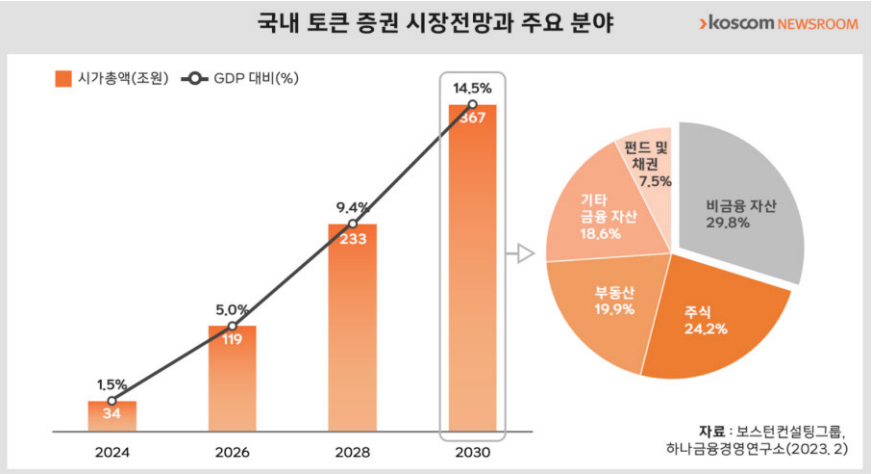
구분	2024년	2025년	2026년	2027년	2028년	2029년	2030년
STO 시가총액	34조원	77조원	119조원	176조원	233조원	300조원	367조원
STO 거래금액	77조원	172조원	268조원	396조원	525조원	675조원	826조원
트랜잭션 수	5,800만 건	1억 3,100만 건	2억 400만 건	3억 200만 건	4억 건	5억 1,500만 건	6억 3,100만 건
목표 시장 점유율	30%	40%	50%	50%	50%	50%	50%
목표 매출	26억원	79억원	153억원	227억원	300억원	386억원	473억원

※ 2022년 국내 유가 증권 시장 _ 자료 출처 : 코스콤 뉴스룸

- 시가총액 2,082조원
- 거래금액 4,687조원 : 시가 총액 대비 거래금액 비율 225%
- 거래건수 36억 건 : 거래금액 대비 건수 0.00008%

※ STO 참고자료 및 추정 근거

보스턴컨설팅그룹, 하나금융경영연구소 (2023. 02)



04. Scale-up

▶ 2030년 글로벌 시장 규모는 최대 4조 달러 (약 5,200조 원)로 추정

citibank

Citi: Tokenized securities market could reach $4 trillion by 2030

씨티은행 자체 보고서 내용 요약

- 2030년 STO 시장규모를 4조달러로 예측
- 가치 있는 거의 모든 것은 토큰화할 수 있음
- CBDC 도입과 게임 콘텐츠 내 자산 및 소셜미디어의 결제가 블록체인에 의해 주도될 것
- 블록체인 기술의 대중화(10억명 이상의 사용자 확보)에는 6~8년 정도 기간이 필요할 것으로 예상
- 토큰화의 장점을 검토한 대형 기관 및 기업들은 기술 도입에 긍정적으로 변화하는 중

글로벌 주요 동향

미국
- 2020년 43건의 STO 진행
- STO 거래 플랫폼 tZERO 2022년 11월 거래량 180만달러 <기사 전문 보러 가기>

일본
- 2019년 노무라, 다이와 등 6개 증권사 STO 협회 설립
- 2020년 일본 금융청 증권형 토큰 발행 허용
- 2021년 회사채 토큰증권 등장
- 채권 및 부동산을 기초자산으로 하는 비상장 STO 위주

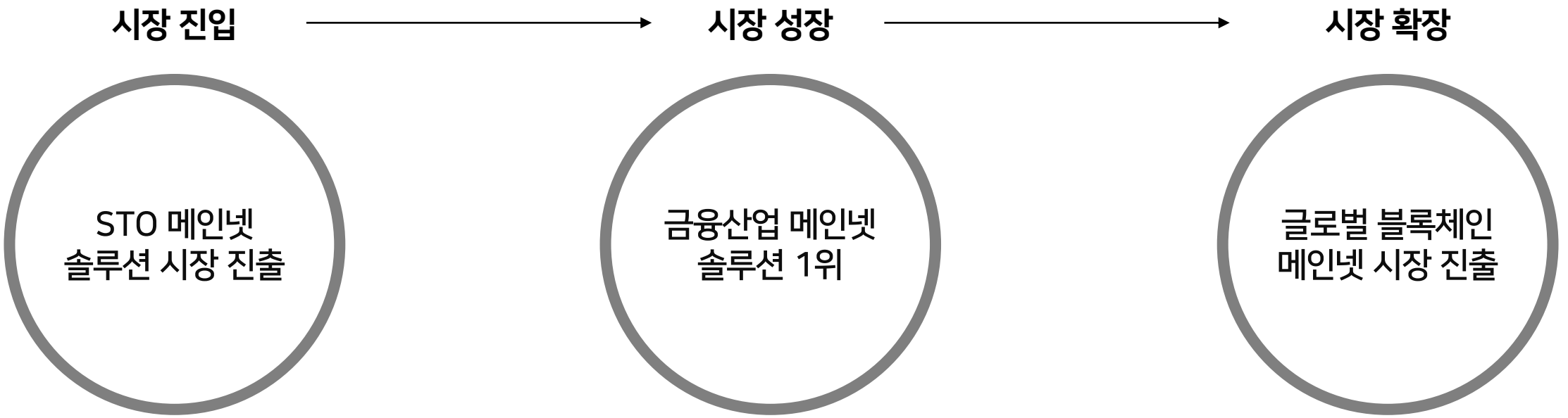
싱가포르
- 2017년 STO 발행 가이드라인 발표
- STO 발행 및 유통 플랫폼, 싱가포르 거래소, 싱가포르 개발은행 등 금융권 참여

홍콩
- 2023년 2월 그린본드 토큰 증권 발행

04. Scale-up

▶ 국내 시장 점유율 1위 달성 후 제도적으로 STO를 허용하고 있는 싱가포르, 일본 등 아시아권 주요 거점 해외 진출

▶ 이후 미국, 유럽 등 글로벌 시장 전반 진출



- 2024년 STO 시장 서비스 오픈
- 메인넷 모니터링 서비스 오픈
- 국내 기업의 메인넷 수수료 해외 유출 방어

- 2025년 국내 금융산업
  블록체인 메인넷 솔루션 점유율 1위
- 조각 투자, CBDC 등 금융 관련
  메인넷 솔루션 확대 제공

- 2026년 글로벌 시장 진출 시작
- 일본, 싱가폴 1차 타겟

05. IR _ 투자 유치

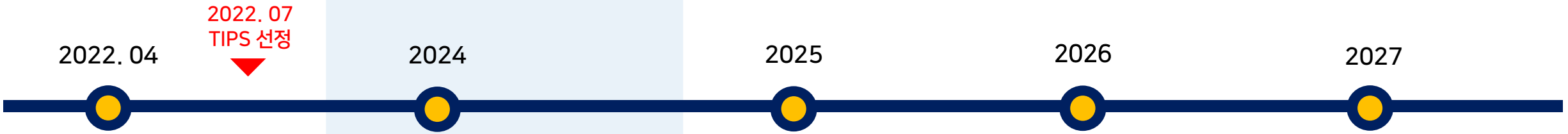
▶ 투자 라운드 개요

Pre A

IR라운드 일정	2024.02.16 까지	주식 발행	신주 발행
목표 기업가치	200억원	인수주식 수	453
목표 투자금액	20억원	투자 단가	4,411,116 원
희망 투자자	금융관련 기업	할당 지분	10%
주요 주주	이흥노 79%	투자금 사용용도	연구소 설립, 네트워킹 인력확보, 사업역량 확보
	선보엔젤파트너스 12%		
회사 자본금	2,200 만원 (2022년 기준)		

05. IR _ 자금유치 및 엑시트 전략

▶ 투자 라운드 플랜



SEED

기업가치 17억원
투자금액 2억원
투자자 VC 선보엔젤파트너스

주요 활동
- 블록체인 메인넷 개발 및 운영
- 연구 인재 채용
- MY AI 서비스 개발

Pre-A

목표 기업가치 200억원
목표 투자금액 20억원
희망 투자자 금융관련 기업

주요 활동
- 국내 금융산업 전용
  블록체인 메인넷 솔루션 개발
- STO, 조각투자 등 관련 업체 대상 영업
- 정부 정책 친화적 서비스 구축 대관 업무

Series A

목표 기업가치 500억원
목표 투자금액 50억원
희망 투자자 국내 은행/증권사

주요 활동
- 국내 블록체인 메인넷 솔루션
  점유율 1위 달성
- 증권사, 은행, 거래소 등 금융산업
  전반에 대한 영업 범위 확대
- 정부 정책 대관 업무 강화
- 글로벌 진출 전략 수립

Series B

목표 기업가치 1,000억원
목표 투자금액 100억원
희망 투자자 글로벌 VC

주요 활동
- 국내 블록체인 메인넷 솔루션
  점유율 1위 유지
- 글로벌 거점 진출 (일본, 싱가폴 등)
- 서비스의 글로벌 표준화
- 현지화 개발 병렬 진행
- IPO 준비

IPO & M&A

목표 기업가치 3,000억원
목표 투자금액 300억원

주요 활동
- IPO 주관 증권사 선정 및 IPO 추진
- 글로벌 거점 확대 (미국, EU, 홍콩 등)
- M&A 추진

05. IR _ 투자금 활용 계획

▶ Pre-A 투자유치금 활용 방안



연구소 설립

- 국내 금융권 요구사항에 대응하는 성능의 메인넷 개발을 위한 연구소 설립
- 연구소장 / 개발 인력 충원
- 블록체인 운영 노드 인프라 구축



네트워킹 인력 확보

- 금융권, STO, 조각투자 플랫폼 등 관련 경력자 출신 임원 영입
- 금융감독원 등 기관 대상 대관 업무 경력자 영입
- 네트워킹 전담 부서 설립



사업 역량 확보

- 국내외 사업 역량 확보를 위한 사업 부서 설립
- 블록체인 마케팅 경력 CMO 영입
- 마케팅/홍보 예산 수립 및 집행

05. IR _ 투자 유치

▶ 투자 경쟁력

1. 블록체인 메인넷 국내 솔루션 분야 최고 권위의 CEO
2. GIST 교원창업 기업으로서, 사업 관련 최적의 연구개발인력 보유
3. 사업활동에 필요한 기술력 관련 특허를 보유
4. SEED 투자 단계에서 국내 최고 수준의 VC 투자유치로 향후 투자 라운드 진행 유리
5. TIPS를 통해 정부에서 인정하는 기술력을 인증
6. 블록체인 가상자산 시장에서의 시장 니즈가 명확
7. 문제해결을 위한 명확한 솔루션 포인트를 인지하여 사업활동
8. 국가적 차원에서 금융거래 관련 국내 메인넷 기업이 필요
9. 금융기관의 향후 가상자산 시장 진입에 필요한 블록체인 기술 확보 및 협력 용이



2024 - 01

IR Deck

THANK YOU

LiberVance

• 회사명 리버밴스㈜

• 본사 광주 북구 첨단과기로 123,
광주과학기술원 산학협력연구관 307/308호

• 서울지사 강남구 강남대로94길 27-11 5층



Worldland

A Global Digital Land

November 2023

01. Market status

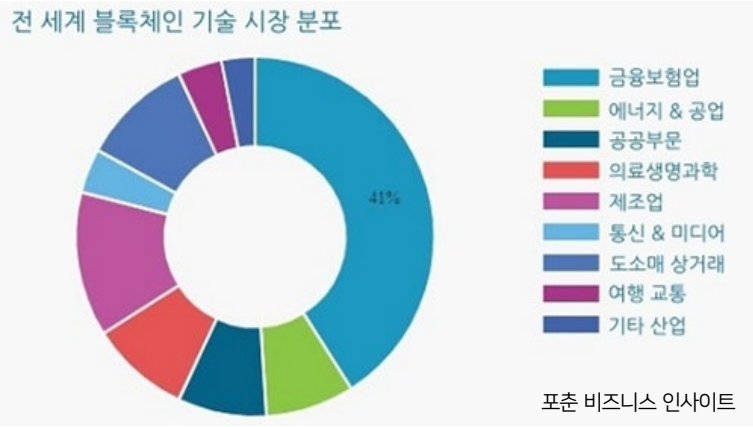
가상자산 관련 규제 강화에도 불구하고 블록체인 기술의 활용 지속적으로 성장 기대

데이터 분산화 및 저장, 소유권 통제 가능 차세대 웹 환경인 Web3.0의 탈중앙화된 플랫폼은 모두 '블록체인' 기술 기반

블록체인 시장은 2026년까지 연간 68.4% 성장 예상되며, 가상자산 전체 시가총액은 약 1,826조 규모에 달함 (23.11.20 기준)

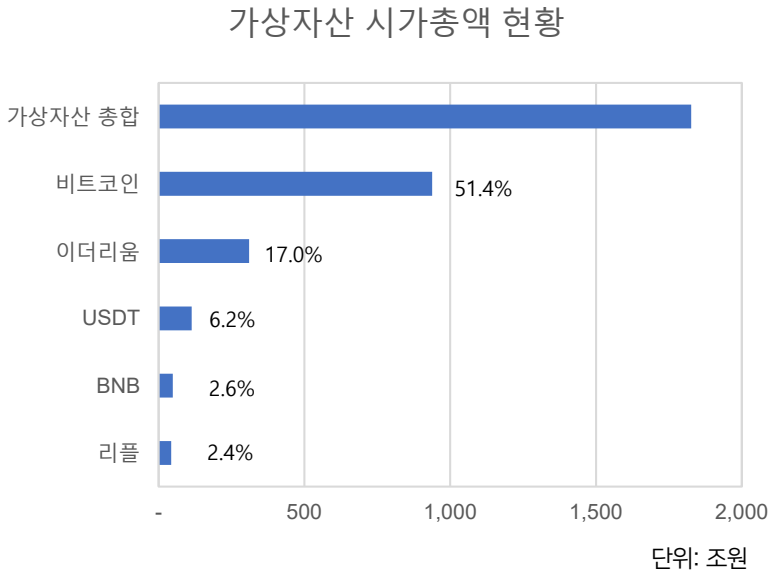
지속적으로 증가하는 블록체인 시장

- *블록체인은 다양한 플랫폼 (NFT, P2E게임, DeFi, 메타버스, 시티코인, 비트코인 등) 의 기반기술*
- 블록체인 기술에 대한 투자 증가, 은행 및 사이버 보안 분야의 블록체인 솔루션 사용 확대 등이 시장 성장 주도할 것으로 예상
- 공공 및 정부의 인지도 향상, 법제도화로 이어져 블록체인 산업 성장이 점차 가속



가상자산 시장 규모 ('23.11.20 기준)

- **비트코인은 시가총액 938조원으로 전체 가상자산 시가 총액 중 51.4% 비중**
- 시가총액 기준 이더리움 310조원, USDT113조원, BNB 48조원, 리플 43조원 순



*알트코인 : 비트코인을 제외한 모든 가상화폐

01. Market status

확장성 트릴레마 (Scalability Trilemma) 및 해결책

확장성, 탈중앙성, 보안성이 모두 우수한 블록체인을 만드는 것은 현재 블록체인 기술에서 가장 풀기 어려운 문제
해결책은 중앙 집중형 합의 스케줄러를 추구하는 Top Down 방식 vs 시장 수요에 의한 자연발생을 추구하는 Bottom Up 방식

블록체인의 성능지표: 확장성, 탈중앙성, 보안성

- **확장성** (최대 초당 거래수 (TPS) 가 얼마나 높은가?)
  - ✔DS Attack 저항위해서, 블록 용량과 BGT(블록생성속도)를 제한함
  - ✔PS 감소: 21년 5월, NFT 공급량 초과로 거래수수료 평균 건당 68달러 도달

- **탈중앙성** (특정 개인 또는 집단이 채굴을 독점할 수 있는가?)
  - ✔각종 공격에 대한 보안 유지를 위한 충분한 분산 노드가 필요
  - ✔채굴 ASIC의 등장으로 인해 새로운 노드의 진입이 어려운 문제 발생
  - ✔누구나 고르게 참여 가능한 합의 알고리즘 필요

- **보안성** (각종 공격으로부터 얼마나 안전한가?)
  - ✔확장성 개선을 위한 Layer-2 솔루션 혹은 사이드체인의 보안 문제
  - ✔사이드체인의 분산성 결여, 각종 공격에 취약
  - ✔다양한 공격 대비를 위한 암호학적 보안 솔루션 필요

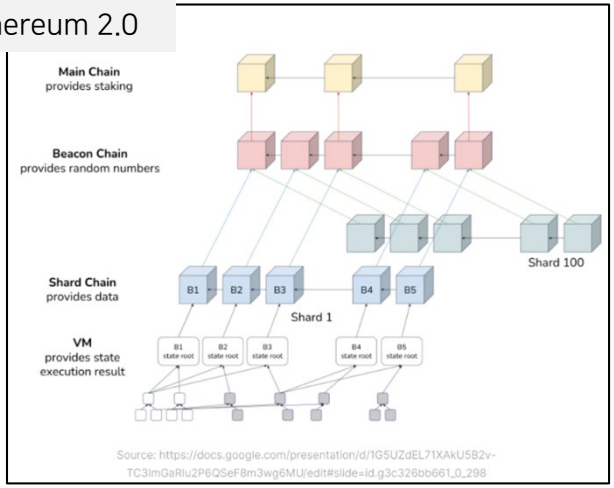
Top Down vs Bottom Up

Bottom Up 방식
- 자연 발생적
- 자연 선택적 방식
- ETH는 글로벌 Layer 로 존재
- 병렬적 독립 체인이 연결되며 확장성 문제 해결 가능

- 중앙 집중 형
- Ethereum 2.0
- Beacon + 64 shards

Top Down 방식

Ethereum 2.0



01. Market status

Layer-1 vs Layer-2

서로 독립적이고, 높은 트렌젝션 수를 감당하지 못하는 Layer-1 스케일링 솔루션
사용성 위주의 자유로운 설계가 가능한 Layer-2 스케일링 솔루션

Layer-1 문제점

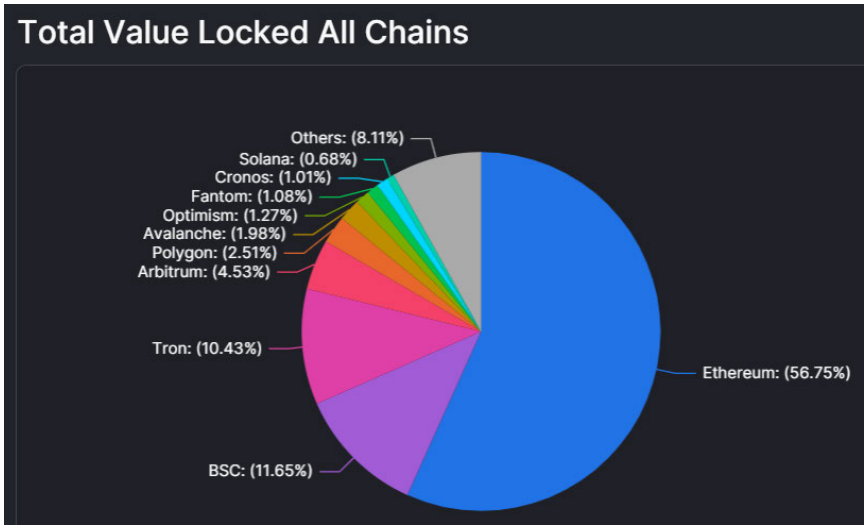
- L1 코인들은 서로 교류 어려움, 각자의 영역에서만 거래 가능
- 주로 이더리움 네트워크 기반으로 사용하고 있어 거래 비용이 높고 Transaction을 감당하지 못함

왜 Layer-2 로 가야 하는가?

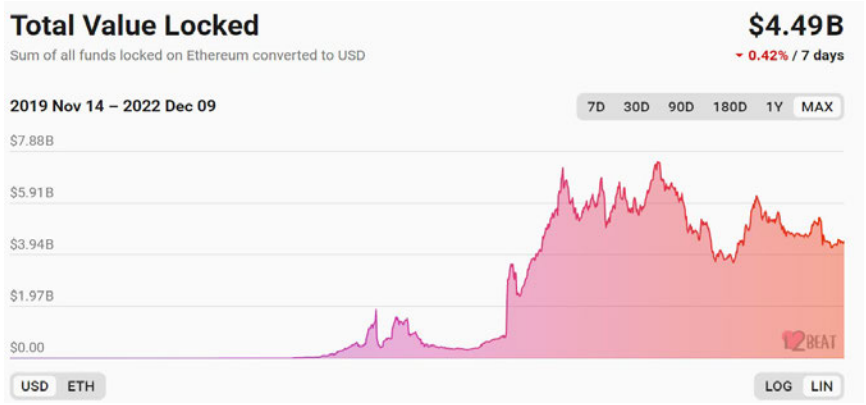
- 성장 기회가 큰 새로운 네트워크인 L2
- 낮은 거래 수수료 및 빠른 거래
- 속도와 사용성 위주의 자유로운 설계 가능
- 이더리움과 결합되어 실행되면 L1보다 L2를 더 많이 사용할 수 있을 것으로 전망 (The Block 리서치, 21.12)
- 최근 *dYdX의 두드러진 성장, L2 매우 초기 단계

*dYdX : 이더리움 블록체인의 스마트 계약을 활용한 탈중앙화 거래소(DEX)

알트코인 TVL 포션 비교('22.12 기준)



Layer 2 TVL 추이('22.12 기준)



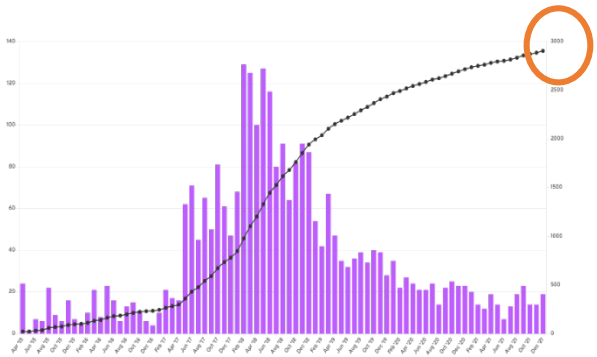
01. Market status

가장 많이 이용하는 블록체인 네트워크 ' 이더리움 '

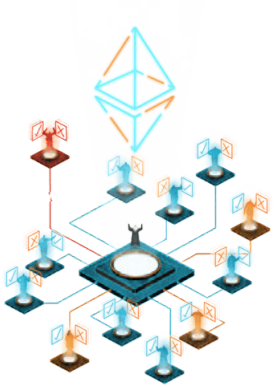
전체 탈 중앙화 서비스 (dApp) 중 75%가 이더리움 기반 서비스를 제공
시장 성장에 따른 혼잡성, 수수료 증가 등의 문제 존재, 이를 해결하기 위한 Layer-2 솔루션 제시
2년 내 약 80%까지 Layer 2 로 이동할 것으로 예상

하루 사용자 약 1억 명, 거래 2억 4천여 건

- 계속하여 증가하는 사용자 수와 거래량으로 인해 거래완료에 걸리는 시간, 혼잡성, 수수료 증가 등의 문제 야기
- 그러나 22년1월 솔라나 다운으로 인한 디도스 공격 취약성 등 안전성 문제, 非탈중앙성 등의 문제 지속적 대두



[이더리움 앱 수 및 신규
추가현황 (21.12월 기준)]



문제 해결 위한 Layer-2 솔루션

- 스마트 계약을 통해 Layer-1 네트워크 위에서 유지
- 주요 네트워크와 상호작용
- 기본 프로토콜 수정에 의존하지 않음
- Layer-2 프로젝트의 유동성 인센티브가 등장하며 점점 성장하고 있다고 진단 (The Block 리서치)



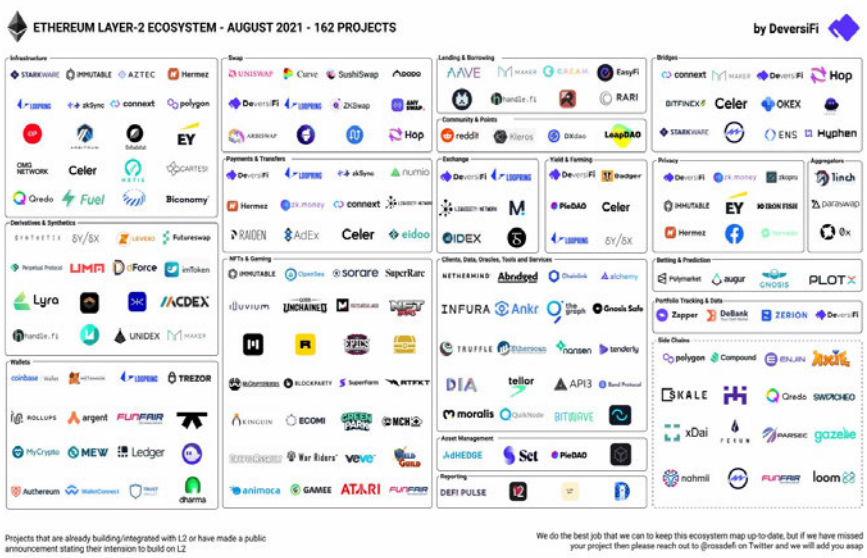
01. Market status

Layer-2 솔루션 생태계

폴리곤 활성 이용자 수 50만 건 이상으로, 이더리움 활성 이용자를 능가 (21.10 기준)
Layer 2 시장의 성장은 뚜렷하며, 최대 50배까지 성장할 것으로 예상

대표적 사례 - 폴리곤(Polygon)

- 2021년 급성장한 *L2는 사이드체인, 플라즈마로 시작
- 2021년 5월 이더리움 네트워크 체증 문제 증대로 폴리곤(Polygon) 확장 솔루션 수요 폭증
- 이더리움을 능가하는 활성이용자 보유



* By DeversiFi, 21.8월 기준

Layer-2의 시장 가능성

- 크립토 경제는 이미 무시할 수 없는 규모로 성장
- 경제시장의 원활한 작동을 위해 확장성과 상호운용성 솔루션 제공 기본 인프라 구축은 필수 불가결
- L2의 성장은 뚜렷하며 향후 20~50배 성장 전망 (21.11, Ryan Berckmans)

*L2 : Layer-2

02. Problem

기존 Layer 2 솔루션의 문제점: 보안 문제, 탈중앙성 이슈

보안성이 최대 장점인 블록체인 플랫폼에서 보안문제 발생, 꾸준히 제기되는 탈중앙성 이슈

→ TPS 제고 위해 탈중앙성이나 보안성이 취약한 합의 알고리즘과 체인 연결 구조 선택

(문제점 1) 기존 L2 솔루션의 합의 알고리즘 상 문제점

- 기존 L2 솔루션은 PoS/PoA 기반 합의 방식 사용
  - 블록 생성 노드, 검증자 노드의 숫자가 적음
    - ☐ DoS 공격, 권한 탈취 공격에 취약해지는 보안문제 발생
  - 탈중앙성 훼손
    - ☐ 블록 생성 노드 운영자에 대한 신뢰 문제 발생

(문제점 2) 기존 L2 솔루션의 체인 연결 기술상 문제점

- 브릿지 구조상 보안성 문제가 발생하기 쉬움
- 문제 발생 시, 연결 체인 사용자 전체가 피해를 입을 수 있음
- 체인 연결 구조에 따른 사용상의 문제 ex:
  - Plasma, Optimistic rollup : exit을 위해 challenge 기간 동안 기다려야 함
  - zk rollup : 네트워크의 초기 신뢰성 문제, 계산량 문제, 양자 컴퓨팅으로 인한 해킹 우려
  - Cosmos : 각 체인들이 보안성을 각자 확보해야 함
  - Polkadot : 검증자를 다시 감시하는 fisherman 이 별도로 필요

1) **합의 알고리즘**과 2) **체인 연결 구조**를 개선하여 이러한 문제들을 해결할 수 있음

02. Problem

합의 알고리즘상 문제점 – PoA 권한 해킹 사례

액시 인피니티 암호화폐 유출 사태 – 로닌 사이드 체인

- 22년 03월, 이더리움과 액시 인피니티의 유틸리티 토큰을 이어주는 사이드체인 '로닌'에서 보안 침해사고 발생
- 권위 증명 (Proof of Authority, PoA) 합의 알고리즘 사용
    - 사이드 체인 내 거래를 검증하는 벨리데이터(Validator) 풀이 매우 소규모임을 의미 ***(탈 중앙성 훼손)***
    - 액시 인피니티의 경우 9개 유효성 검사기 중 최소 서명에 필요한 개수인 5개 권한을 해킹당했으며, 임의 Transaction을 발생시켜 암호화폐를 탈취함.
- 피해 금액은 17만 3,600 ETH 및 2,550만 USDC 로, 약 7,500억원이며, 이는 역대 최대 규모임.
- 해킹 후 6일만에 발견되었으며, '로닌'의 거버넌스 토큰인 RON 은 20% 가까이 하락함.





02. Problem

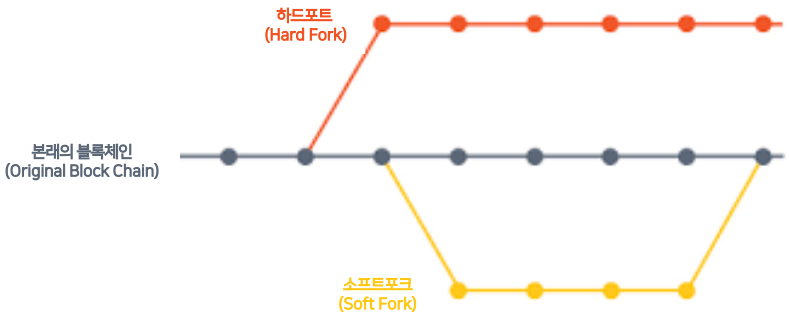
합의 알고리즘상 문제점 – 탈중앙성 훼손 논란 사례

PoS 합의 알고리즘의 탈중앙성 훼손

폴리곤 비공개 하드포크 진행으로 제기된 탈중앙성 훼손 논란

- 21년 12월, 메이저 L2 솔루션 중 하나인 폴리곤 네트워크가 일반 사용자들에게 공지 하지 않고 하드포크를 진행함.
- 블록체인 하드포크 (BlockChain Hard Fork)
    - 체인의 기본 동작에 해단하는, 기본 코드 (Base Code)를 업데이트 하는 일
- 폴리곤 COO인 샌딥 나이왈은 '하드포크를 알리지 않고 추진한것은, PoS 보안 취약점이 악용될 수 있었기 때문이며, 최선이었다' 해명함.
- Vitalik Buterin 의 정치적 탈중앙화 (Political Decentralization)
    - 얼마나 많은 수의 개인 혹은 조직이 시스템을 구성하는 컴퓨터를 제어하는가?
- 참여 노드들에게 알리지 않고, 신속하게 하드포크가 가능했다는 점에서, 정치적 탈 중앙화 측면에서의 의혹이 제기됨.
- 많은 수의 L2 솔루션들이 탈중앙성에 있어 문제가 있음이 드러남.

polygon



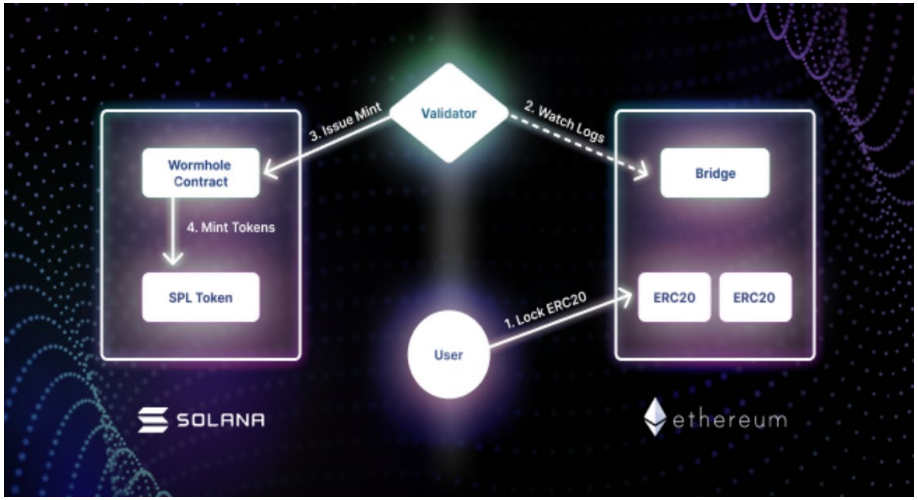
02. Problem

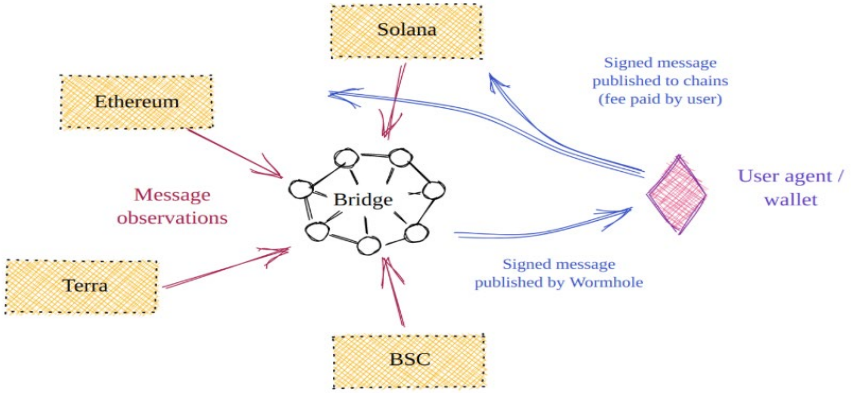
체인 연결 기술상 문제점 – Bridge 해킹 사례

체인 연결 컨트랙트의 취약점을 이용한 해킹

Solana – Wormhole 해킹 사건

- 22년 02월, 이더리움과 솔라나 블록체인을 연결하는 크로스체인 '웜홀'에서 보안 침해사고 발생
- 웜홀 서명검증 과정 (Wormhole Contract 코드) 상에서의 보안 취약점 이용 ***(논리 구조 취약)***
  - Transaction 위조, 12만개의 WETH 임의 발행
  - 브릿지 솔루션을 활용, ETH 로 바꿔 인출함.
- 피해 금액은 약 3800억원 어치로, 역대 4번째 규모임.
- 이번 해킹 사건으로 인해, Solana 와 Ethereum 모두, 10% 이상의 가치 하락이 발생함.
- 브릿지 구조에서의 보안성이 문제가 된 사례





WorldLand 정의

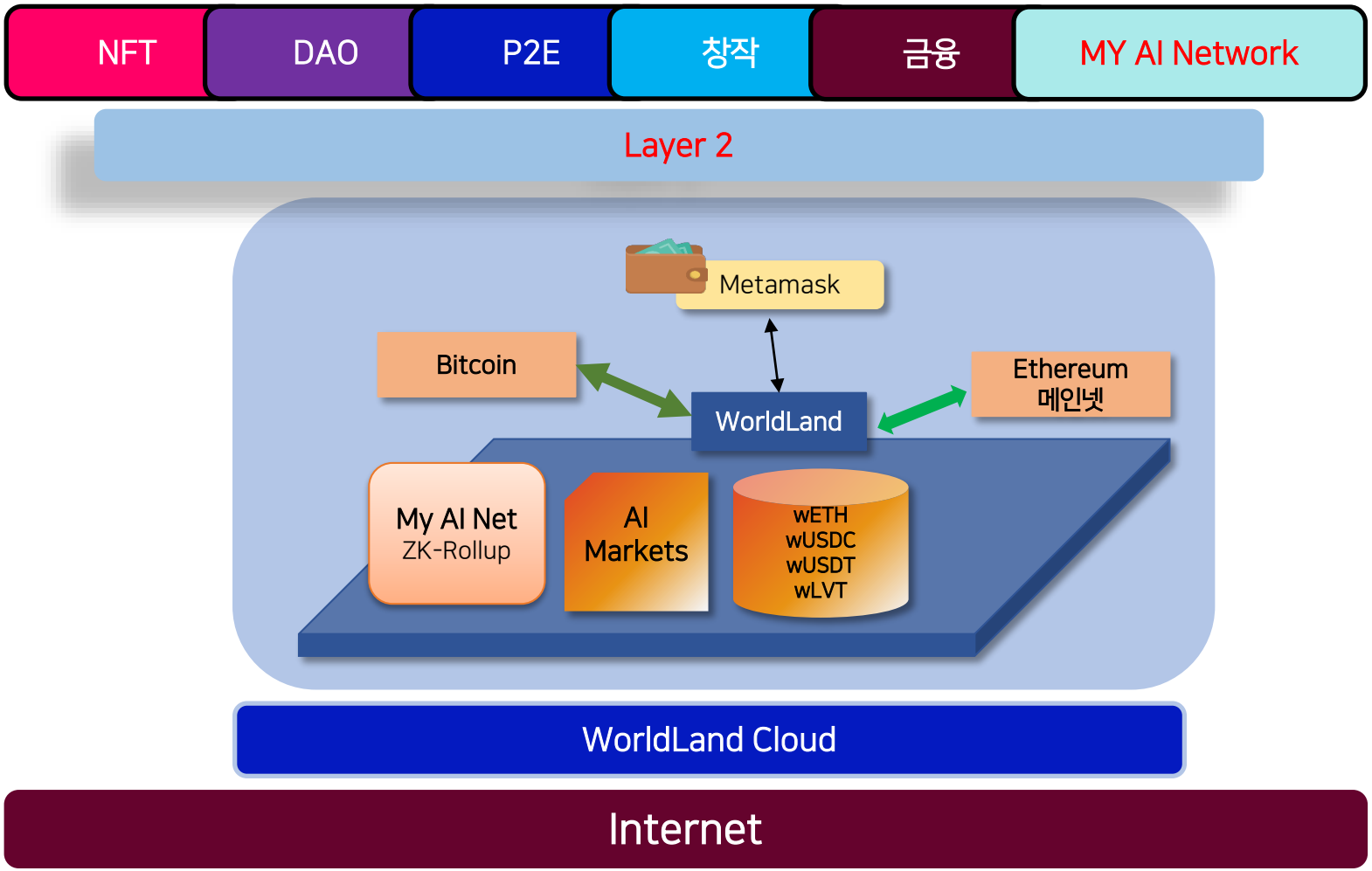
WorldLand는

- 세계인이 교류하는 디지털 랜드다.
- 반감기를 적용한 화폐 WLC를 발행한다.
- 참여 노드가 WLC 주조권을 갖는다.
- 노드는 Green VCA로 합의한다.
  - 양자내성암호를 사용한다.
  - 동전던지기함수와 오류정정부호POW로 합의한다.
  - EVM을 사용한다.
- 양자내성 영지식 증명을 활용하여 확장성을 제고한다.
- 킬러 dApp My AI Net 플랫폼을 제공한다.
- My AI Net는 개인에 귀속된 AI가 상호 협력하여 교사-학생 학습, 전이학습, 연합학습 등으로 **다양한 사회 구성원과 함께 진화하는 네트워크를 의미한다.**

03. Solution

WorldLand는 글로벌 메인넷이며 다양한 MY AI Network 등 다양한 dApp 서비스를 제공

MY AI Network는 다양한 형태의 맞춤형 AI 서비스 구현

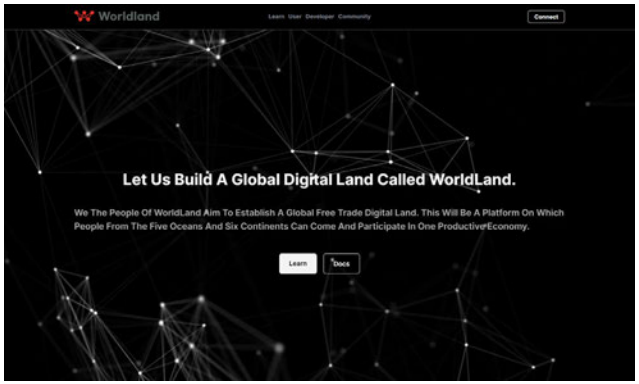


WorldLand 개요

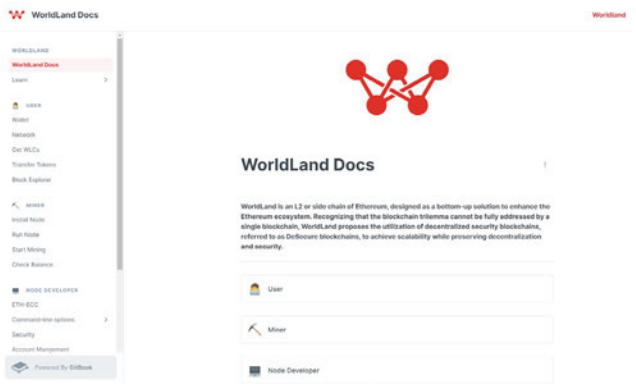
구분	내용		특이점
네트워크	코인(WLC)을 발행하는 메인넷		Ethereum Sidechain
용도	Payment currency		ZKP L2 My AI Network
메인넷특징	EVM호환, Ethereum 사이드체인		양자내성암호 서명함수
주 개발사	리버밴스㈜		GIST교수창업기업
혁신기술	합의알고리즘	VCA + Green	블록생성주기 10 초 블록당 WLC 4 개 반감기 2년
	코인발행스케쥴	4번의 반감기 이후 매년 4%씩 증가	
	암호체계	양자내성암호 (합의, 서명, VCT, ZKP)	
거버넌스	DAO: 5년 이내 탈 중앙 거버넌스 완료		
D-Date 론칭	D-Date: 2023년 3분기 메인넷 글로벌 론칭		
프로젝트 사이트	https://worldland.foundaiton		

WorldLand Seoul

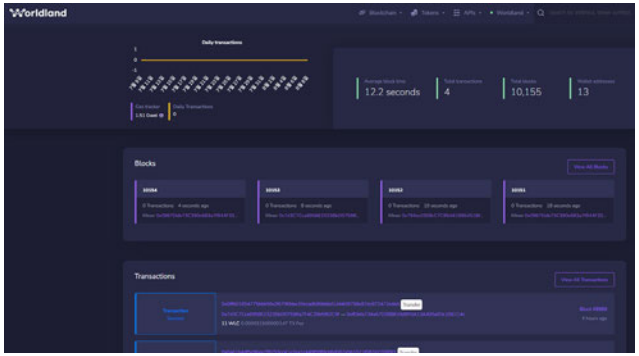
2023년 8월 8일, 메인넷 WorldLand Seoul 런칭



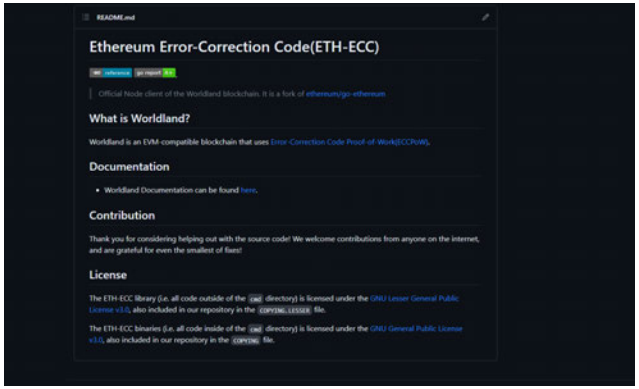
WorldLand Site (worldland.foundation)



WorldLand Docs (docs.worldland.foundation)



WorldLand Scan (scan.worldland.foundation)



WorldLand Github (github.com/cryptoecc/ETH-ECC)

WorldLand 네트워크 개요 (1).

	내용		비고
네트워크	코인(WLC)을 발행하는 메인넷		
코인의 용도	Payment		
메인넷특징	EVM호환, Ethereum 사이드체인		
주 개발사	리버밴스㈜		GIST교수창업기업
혁신기술	합의알고리즘	VCA	블록생성주기 12 초
	코인발행스케쥴	4번의 반감기 이후 매년4%씩 증가	
	암호체계	양자내성암호 (합의, 서명, VCT, ZKP)	

WorldLand 사업 개요 (2).

	내용	비고
하드포크	Ethereum 하드포크: Block Height @15,000,000	
거버넌스	DAO 5년 이내 탈 중앙 거버넌스 완료	
토큰세일	DAO 토큰 WLT 발행 메인넷 제네시스 블록에서 WLT를 WLC로 교환	
프로젝트 사이트	https://worldland.foundation	

03. Solution

WorldLand 화폐 발행 스케쥴은 어떤 것 인가요?

4번의 반감기를 거쳐 안정화 단계에 이르고 그 후로는 매년 4%씩 추가 발행합니다.

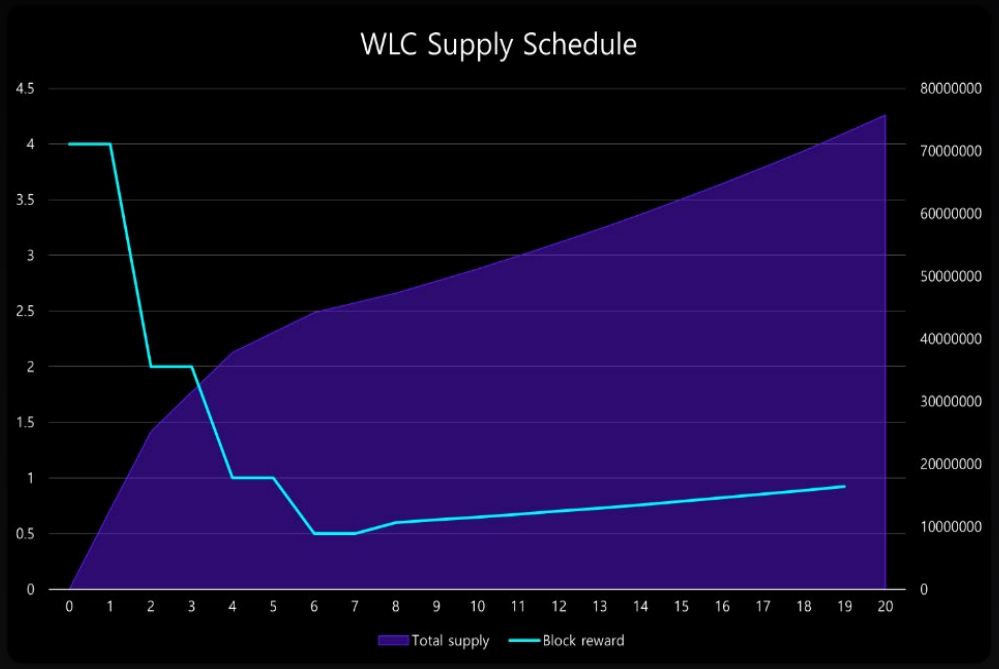
❑ WorldLand경제 규모 매년 4% 예측 가능하게 성장

❑ 과도한 호들링(Hodling) 방지

❑ 코인 사용 및 투자 촉진

WorldLand Coinomics

Coinomics of WorldLand involves the issuance and utilization of the mainnet coin WLC to support the development ecosystem of dApps and the operation of the mainnet. Here are the key aspects of WorldLand's coinomics:



VCA 합의 프로토콜 – 개발 요소기술 1

VCA합의알고리즘은 오류정정부호에 기반함. 고치고자하는 오류의 갯수로 Puzzle 문제의 난이도를 결정함.
SHA함수와 달리, 수많은 오류정정부호를 만들 수 있음. 따라서 ASIC 칩의 출현을 방지.
탈중앙성과보안성을 제고하게 되며, 전력 소모를 최소화 할 수 있게됨.

VCA 합의 프로토콜 : 오류정정부호에 기반한 새로운 합의 기술

- 탈중앙성과 보안성이 높은 블록체인 플랫폼 개발

기존 합의 기술상 문제점

- PoS/PoA 방식: 탈중앙성이 훼손되기 쉬움, 해킹에 취약
- PoW 방식: 에너지 소모가 큼, ASIC 채굴장비로 인해 탈중앙성 훼손

VCA 합의 프로토콜의 차별성

- VCA 합의 프로토콜 기반 블록체인 플랫폼에서는
  - 대규모 p2p 노드가 블록 생성에 참여함으로써 탈중앙성 확보
  - Time-varying crypto puzzle 방식으로 이중 지불 공격 (DS attack) 에 대한 높은 보안성
  - ASIC 저항성으로 인해 탈중앙성 강화 및 에너지 소모 감소



VCA 합의 프로토콜 – 개발 요소기술 1

선행 연구 개발 결과: 논문 게재, 테스트 서버 가동, Github 코드 공개 완료

✔ 선행 논문: ieeexplore: Heung-No Lee eccpow

✔ Github: https://github.com/cryptoecc/ETH-ECC

VCA 합의 프로토콜의 정의

- 모든 노드가 동일한 절차를 독립적으로 수행함으로 합의에 이름

Verifiable Computation Algorithm (VCA) 과정

Step 1: 참여하고자 원하는 노드는 누구나 Stake하고 참여

Step 2: 각 노드는 VRF 기반 동전던지기 함수 실행하고 블록 생성 노드로 스스로 선출

Step 3: 블록 생성 노드로 선출(동전던지기에 당첨)된 노드는 Puzzle Solving (PS) 수행

Step 4: PS에 성공한 노드는 노드 선출 증거 및 Puzzle 해답 증거를 첨부하여 생성된 블록 배포

VCA 합의 프로토콜 규칙

- 가장 많은 VC Energy가 저장된 체인이 올바른 체인
- 각 동전의 앞면성공확률을 조절하여 에너지 소모량 조절

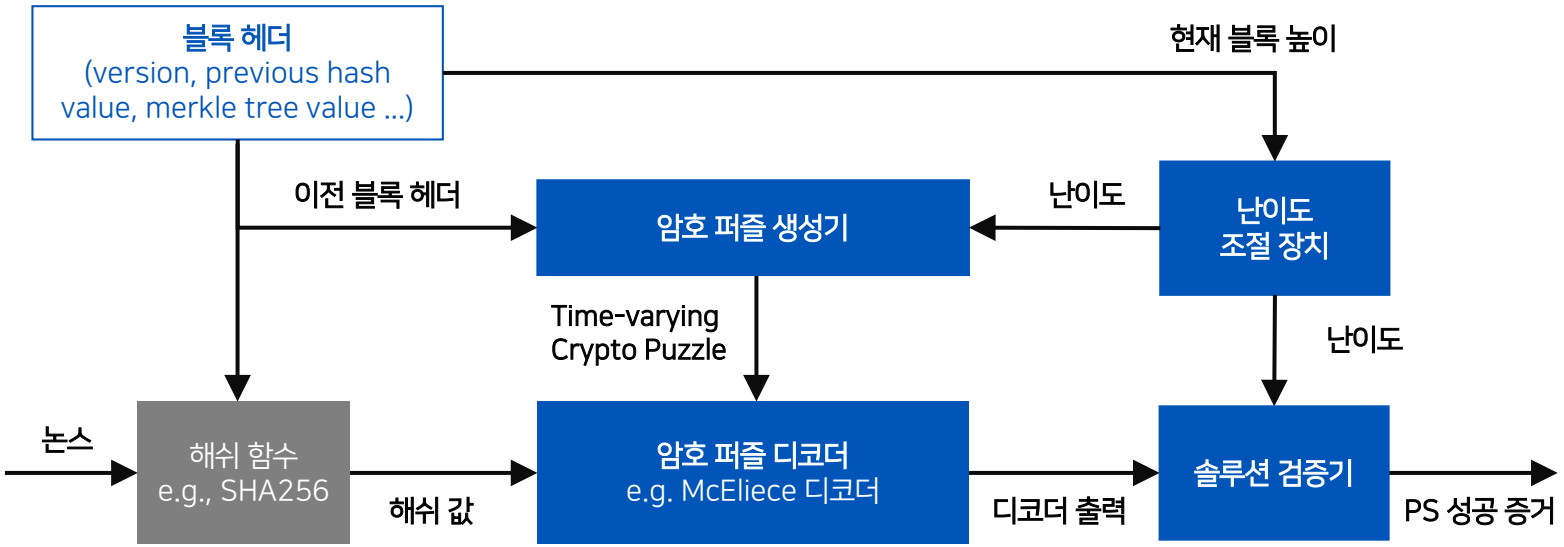
VCA 합의 프로토콜 – 개발 요소기술 1

알고리즘의 특징은 Time-Varying Puzzle 생성 가능 ASIC 저항성 확보
오류정정부호 기반으로 양자 내성 확보 가능.
난이도 조정력을 내포하여 Energy 저장 조절 가능. (예: L1 메인넷은 높은 난이도, L2 넷은 낮은 난이도)

Puzzle Solving 알고리즘의 특징

- 기존 블록체인 (ex: Bitcoin, Ethereum) 의 PoW 는 해쉬함수(hash function) 기반
- **오류정정부호**를 사용하여 Time-Varying Puzzle 생성
- **ASIC 저항성**을 확보 탈중앙성 강화
- **양자 내성 확보** 가능*

* McEliece, MDPC, Concatenated RM-RS 코드에 기반한 여러 알고리즘들이 양자 내성 공개키 암호 방식 표준으로 고려되고 있음

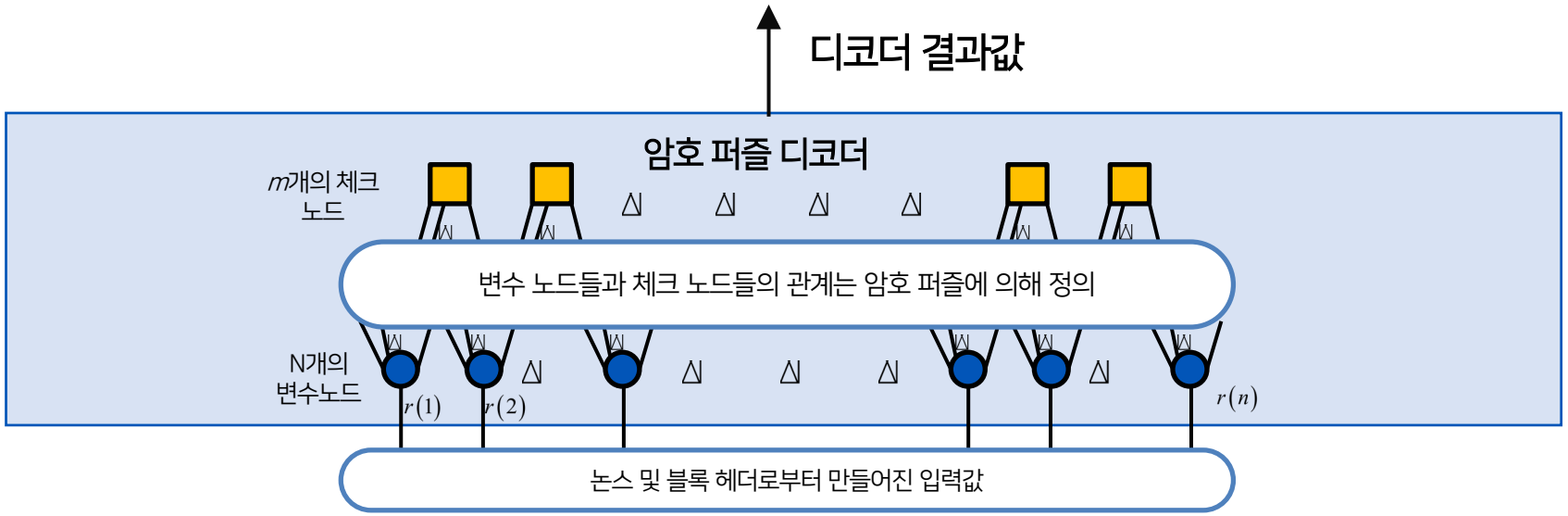


VCA 합의 프로토콜 – 개발 요소기술 1

어떻게 Time-Varying Puzzle 문제를 출제할 수 있는가?
어떻게 Verifiable Computation을 이루는가?
어떻게 수많은 노드의 합의 참여가 가능하게 되는 가?

Time-Varying Puzzle Solving 알고리즘의 구조 설명 (왜 보안성과 탈 중앙성이 제고되는가?)

- **필요한 총 연산량 계산**을 통해 필요한 **노드의 숫자 및 PS 난이도 계산 가능, 디코더 Puzzle 문제 제시**
- 모든 노드는 제시된 디코더 PS 알고리즘 수행, 그 중 성공한 노드가 **디코딩 성공 결과값** (PS성공증거) **제시**
- 디코더의 결과값이 제시되면, 다른 노드는 PS성공 증거 검증, **검증하는 것은 매우 쉬움**
- 각 노드는 ***독립적으로 PS를 수행***하는 것으로 **합의가 완료**, 따라서 **수많은 노드의 합의 참여가 매우 쉽게 됨.**
- 노드간 빈번한 통신이 필요한 BFT 계열 합의 알고리즘보다 **신뢰성과 보안성이 높음**

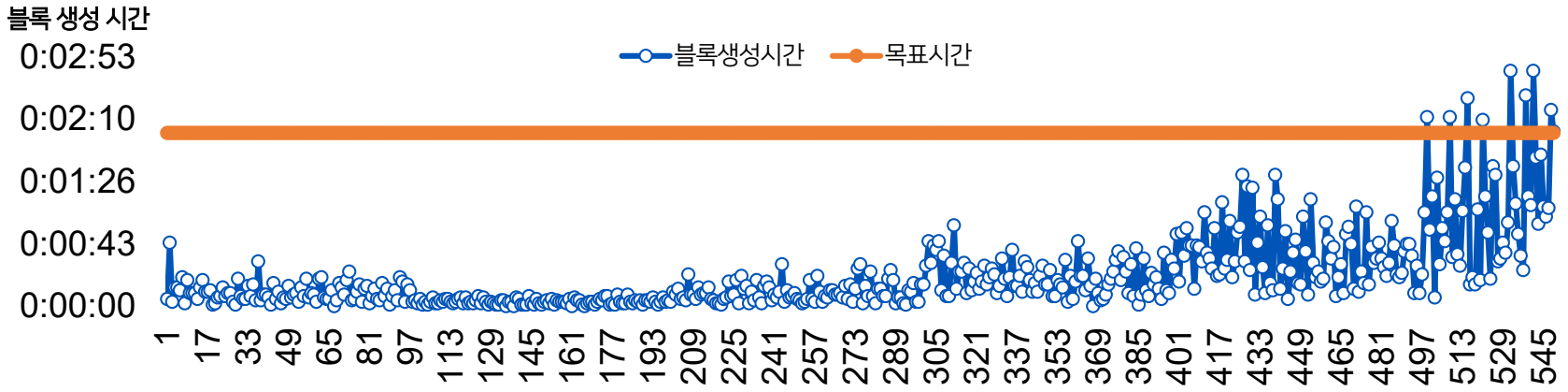


VCA 합의 프로토콜 – 개발 요소기술 1

블록 생성 시간 안정화

VCA 합의 프로토콜 연구 개발 목표 1: 블록 생성 시간 안정화 연구

- 연구 필요성:
  - 빠르고 안정적인 거래 종결성(finality) 확보를 위해 **블록 생성 시간 안정화** 필요
  - 노드의 컴퓨팅 파워의 변동, 매 블록 바뀌는 암호 퍼즐등의 문제로 블록 생성 시간을 일정하게 유지할 방법 제시 필요
- 연구 세부 내용
  - 암호 퍼즐 파라메터에 따른 난이도 변화 분석 연구
  - 난이도 조절 알고리즘 고도화 연구
  - 컴퓨팅 파워 변동 대응 알고리즘 연구



<블록생성시간-블록높이 그래프(블록 생성 목표시간 2분)>

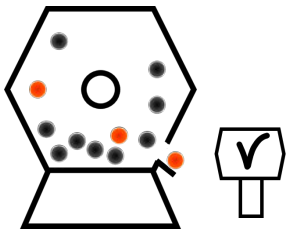
VCA 합의 프로토콜 – 개발 요소기술 1

Verifiable Coin-Toss Function 만들어 빠르고 공정한 블록 생성 노드 선출 방식 연구 및 개발

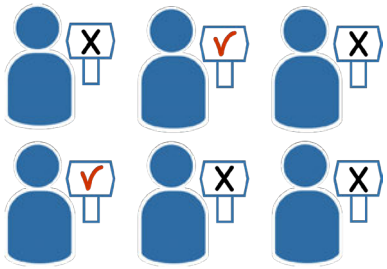
VCA 합의 프로토콜 연구 개발 목표 2: 빠르고 공정한 블록 생성 노드 선출 방식 연구

- 연구 필요성
  - VCA 합의 프로토콜 기반 블록체인은 ASIC 채굴이 불가능하므로 PoW 에 비해 에너지 소모가 낮음
  - 이에 더해 매 블록, 블록 생성 노드를 선출하여 이 노드들이 VCA 합의를 진행토록 하면 에너지 소모를 더욱 낮출 수 있음
  - 하지만 블록 생성 노드 선출이 비효율적이거나 불공정하면 블록체인의 성능 및 보안성에 문제가 생김
  - 빠르고 공정한 블록 생성 노드 선출 방식 연구가 필요함

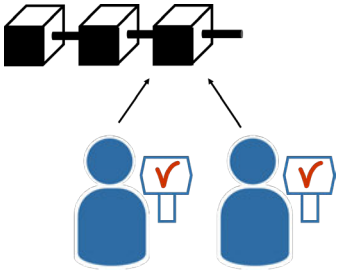
- 연구 세부 내용
  - VRF 기반의 블록 생성 노드 선출 기술 개발
  - VRF 의 보안성, 양자내성 강화 연구
  - 블록 생성 노드 선출 프로토콜을 적용하여 VCA 합의 기술 고도화 연구



VRF 기반 노드 선출



노드 스스로 선출&증거 첨부



선출된 노드 VCA 합의 진행

DeSecure 복합체인 연결 기술 – 개발 요소기술 2

WorldLand 넷은 각각 De-Secure블록체인임, 안전한 WL Net을 용도별 특화하여 각각 개발하고 상호 연결함
나아가 WL 넷을 타 L1 체인과도 연결하여 web3 생태계 확장에 기여

DeSecure 복합체인 연결 기술

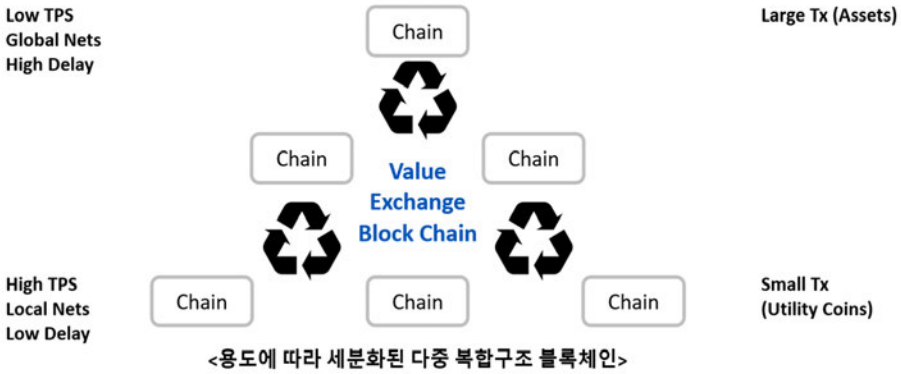
- VCA 합의 프로토콜 기반의 개별 체인 간 및 Ethereum 과 같은 다른 L1 과 연결

기존 L2 솔루션 체인 연결 기술의 문제점

- 브릿지 구조상의 보안성 문제
- 체인간 복잡한 통신 필요, 보안성 미흡

DeSecure 복합체인 연결 기술의 차별성

- VCA 기반의 체인간 연결 효율화
  - 같은 보안 레벨 체인간의 빠른 데이터 교환
  - 다른 보안 레벨 체인간 연결은 높은 보안 레벨에 맞춰서
- Ethereum 과 같은 다른 L1과 연결하여 보안성 강화



<용도에 따라 세분화된 다중 복합구조 블록체인>

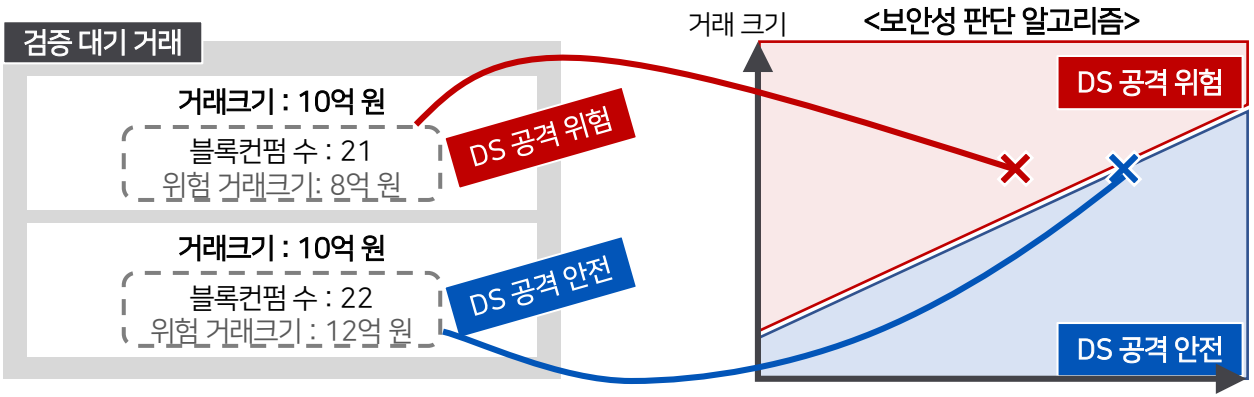
참조발표: https://www.youtube.com/watch?v=VuRgKBDW4Lk&list=PLWq4e5y2b_S9b9AHMPpLyLqF41xBRUd4C&index=2

DeSecure 복합체인 연결 기술 – 개발 요소기술 2

연구 개발 목표: 체인 연결 구조 설계에 있어서 보안 해킹 사고를 저감시키기 위하여
거래 크기에 따라 보안 등급 조절 필요.
Profitable Double Spending Attack* 연구 결과 활용 가능.

DeSecure 연구 개발 목표 : 체인 연결 구조 설계

- 연구 필요성
  - 다양한 체인 연결 구조들이 제안 되었지만 보안성 문제, 프로토콜 복잡성 문제등이 존재함
  - 기존 체인 연결 구조의 장단점에 대한 분석 연구가 필요함
  - 용도별로 보안 파라메터가 다른 체인들을 효율적으로 연결할 수 있도록 하는 새로운 체인 연결 구조의 설계가 필요함
- 연구 세부 내용
  - 체인 보안 레벨 판단 알고리즘 연구
  - 연결된 두 체인의 보안 레벨에 따른 안전한 체인 연결 구조 설계
- Key Idea: 거래 크기에 따라 보안 등급 조절 필요. Profitable Double Spending Attack 연구 결과 활용 가능.



* 참조논문: Heung-No Lee, 2020. https://www.mdpi.com/2076-3417/10/23/8477

04. Technology

경쟁사비교

Zk-STARK 활용 및 새로운 자체 개발 VCA 합의에 기반한 탈 중앙성, 보안성 제고
De-Secure 브릿지 연결, 체인간 보안 레벨 차이에 따른 안전한 연결 구조 설계

Projects	Goal	합의노드 선출방식	합의 알고리즘	합의 알고리즘 Anti-ASIC	확장성 솔루션 제공 여부	DeFi 솔루션 제공 여부	양자내성 지원 여부*	최대 TPS
Polygon	확장성 트릴레마 문제 해결	PoW	GHOST	X	Polygon	O	X (zk-STARKs)	Scalable
Algorand	확장성 트릴레마 문제해결	VRF	Pure PoS BFT	N/A	X	X	X (EdDSA)	45,000 (2021목표)
Stacks	Bitcoin확장성 솔루션	Proof-of-transfer	Verified random function	N/A	X	O	X	~7
Cosmos	*체인간 연결 프로토콜*	PoS	Tender-mint (BFT)	N/A	인터체인 통신	O	X (ECDSA)	Scalable
Solana	*Layer 1*	PoH	BFT	N/A	X	O	X (EdDSA)	Maximum: 710,000
Polkadot	*체인간 연결 프로토콜*	PoS	BFT	N/A	인터체인 통신	예정	X (ECDSA, schnorr, EdDSA)	Scalable
Liber-Vance	**확장성 트릴레마 문제 해결**	**VRF**	**VCA**	**O**	**DeSecure 복합체인 연결기술**	**예정**	**오류정정부호 기반 양자내성 암호화**	**Scalable**

*** RSA, Elliptic Curve 방식의 경우, 양자 내성 확보 불가**
(NIST 양자 내성 암호 기술 표준화, 정보통신기획평가원 주간기술동향 1976호, 20.12.09.)

기술개발 내용 및 목표

세계적인 Ethereum L2 Solution 대비 성능 평가 수행
평가 환경에 따른, 공인 평가원 평가를 통한 성능 달성치 제시

성능지표

평가 항목 (주요성능)	단 위	전체 항목에서 차지하는 비중(%)	세계 최고수준 보유국/보유기관 성능수준	연구개발 목표치	목표 설정 근거
블록 생성 시간 (BGT) 표준편차	%	20 %	1.7 % (Polygon, 4월 기준)	< 2 %	선행연구결과: 90%
최소 조정 가능한 BGT 비율	%	10 %	100 % (Bitcoin)	< 100 %	
난이도 변동시 BGT 안정 시간	블록	20 %	2016 (Bitcoin)	< 7	Bitcoin 의 블록 컨펌 숫자 이내로 안정화를 목표로 함
VRF 기반 노드 선출 공정성	P-value	20 %	0.9 (Algorand)	> 0.9	균등분포 검정 (Anderson-darling test*) 점수 기준
Profitable DS 공격** 가능성	확률	30 %	1.0 (Bitcoin, Ethereum)		현존 블록체인은 해시파워 대여 서비스를 통한 DS공격에 취약함

평가방법

순번	평가항목(성능지표)	평가방법	평가환경
1	BGT 표준 편차	한국시험인증원 평가	네트워크 해시 파워 고정
2	최소 조정 가능한 BGT 비율	한국시험인증원 평가	노드 수, 해시 파워 고정
3	난이도 변동시 BGT 안정 시간	한국시험인증원 평가	현재 해시 파워에서 10% 변동 시
4	VRF 기반 노드 선출 공정성	한국시험인증원 평가	Anderson-darling 테스트, 샘플 수 고정
5	Profitable DS 공격 가능성	한국시험인증원 평가	거래당 최대 크기 기준치를 $단위로 설정

*Anderson-darling test: 데이터가 균등 분포를 따르는지 확인하는 통계적 테스트 방법 (1: 균등, 0: 비균등), 이흥노교수 논문 링크
** Heung-No Lee, 2020. https://www.mdpi.com/2076-3417/10/23/8477

기술개발 현황 및 추진 방향

이흥노 교수 랩 연구 개발 성과로 확보한 핵심 원천기술 (TRL 2~5단계)
TIPS R&D 사업으로 TRL 단계 고도화 및 사업화 추진
서비스 개시를 위해 UI/UX가 제고된 포털 개발, 지갑서비스, 사용자 접근성 제고 등 추진 필요

핵심 기술 개발 현황

기술 분류	개발 현황	개발 목표	세부 개발 목표
VCA 합의 프로토콜	논문 게재 및 특허 출원 완료 핵심 기능 설계 완료 및 시제품 소프트웨어 개발 완료 소규모 테스트 노드 운영중 대규모 테스트 및 부가기능 구현, 기능 고도화 필요 **(TRL 2~ 5단계)**	**블록 생성 시간 안정화**	암호 퍼즐 파라메터에 따른 난이도 변화 분석
			난이도 조절 알고리즘 고도화
			컴퓨팅 파워 변동 대응 알고리즘 구현
		블록 생성 노드 선출 기술	VRF 기반의 블록 생성 노드 선출 기술 개발
			VRF 의 보안성, 양자내성 강화
			블록 생성 노드 선출 프로토콜을 적용하여 VCA 합의 기술 고도화
DeSecure 복합체인 연결기술	핵심 알고리즘에 설계, 실험, 최적화 완료 시스템 전체에 대한 설계 및 테스트 필요 **(TRL 3단계)**	**체인 연결 구조 설계**	체인 보안 레벨 판단 알고리즘 고도화
			체인의 보안 레벨에 따른 안전한 체인 연결 구조 설계

04. Technology

지식재산권

개발 기술 원형의 특허 등록을 통한 지재권 보호 완료
개발 기술을 LV로 이전하여 최고 기술 경쟁력 확보 예정

LiberVance 지식 재산권

- 개발 기술 특허 등록 및 기술이전을 통한 기술 실시권 확보
- LiberVance
  - ✔국내 특허: 등록 2건

INFONET 연구실 지식 재산권

- INFONET 연구실 (https://heungno.net/)
  - ✔국내외 논문 300+편 (SCI 107+ 편)
  - ✔국제 특허: 출원 9건, 등록 7건
  - ✔국내 특허: 출원 7건, 등록 22 건
  - ✔기술이전: 3건 (2억 2천만원)

순번	구분	지식재산권 출원인 (실시권자)	출원국	진행상황	출원 및 등록번호	지식재산권(특허)명	비고
1	보유 특허	이흥노 (리버밴스㈜)	한국	등록	10-2288769	**부호-암호 화폐 시스템**	ECCPoW
2	보유 특허	이흥노 (리버밴스㈜)	한국	등록	10-2288776	블록체인의 거래검증시스템, 및 블록체인의 거래검증방법	ECCPoW
3	보유 특허 (기술이전 예정)	이흥노 (INFONET)	한국	출원	10-2020-0171424	블록체인 전자투표시스템, 그 시스템의 운용방법	DeSecure
4	보유 특허 (기술이전 예정)	이흥노 (INFONET)	한국	출원	10-2022-0018393	**스택 머신 프로그램을 위한 범용 영지식 스나크(zk-snark) 증명 시스템 및 방법**	DeSecure
5	보유 특허 (기술이전 예정)	이흥노 (INFONET)	미국	출원	62/699751	새로운 블록체인 및 암호화폐 작업증명 생성, 증명, 검증 시스템	ECCPoW

05. Scale-up

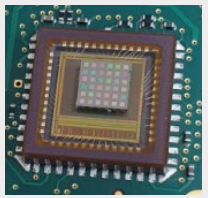
목표시장 및 경쟁사 현황 -요약

경쟁사 대비 자체 플랫폼, 복합체인 연결기술, Anti-ASIC 및 양자내성 특성 등, 기술적 우위 확보

기술적 우위에 의한 시가총액 추월 목표

	Worldland	WORLDCOIN FOUNDATION	Stacks	SOLANA	COSMOS ATOM	Polkadot.	polygon
상호	WorldLand	Worldcoin	Stacks	Solana	Cosmos	Polkadot	Polygon
창립일	2023 년	2019 년	2013 년	2017 년	2017 년	2017 년	2017 년
시가 총액	N/A	약 0.3조 원	약 1.2 조 원	약 32.5 조 원	약 4.4 조 원	약 8.8 조 원	약 10.1 조 원
주요 서비스	My AI Net 초분광카메라 DID 	홍채 인식 DID, UBI 	플랫폼 (NFT, DID, Web3)	플랫폼 (NFT, DeFi)	상호운용 플랫폼	상호운용 플랫폼	오프체인 거래 지원 Layer-2 솔루션
경쟁사 대비 자사의 비교우위점	자체 플랫폼	자체 플랫폼	멀티체인 구현 x		합의 알고리즘 Anti-ASIC 및 양자 내성을 통한 신뢰 검증 기능 부족		

- Green VCA 합의
- Proof-of-Personhood
- My AI Net
- MAC-ZKP



이홍노 GIST 교수팀, 대량 생산 가능한 소형 분광기 개발

발행일 : 2022-03-30 09:12

적은 수의 필터로 고해상도 실현…필터 배열 대량 생산으로 비용 절감

펀딩 및 론칭 스케쥴

구분	준비	1st Qtr D-date	2nd Qtr	3rd Qtr	4th Qtr	2nd yr Qtr-1	2nd yr Qtr-2	2nd yr Qtr-3
White Paper								
External Testnet								
Ethereum Wallet Airdrop								
WorldLand Mainnet		G						
Tech Roadmap		Lattice VCT	MY AI Net Design	MY AI Net Inference Service	MY AI Net Training Service	Bitcoin Bridge	My AI Net WorldLand	Lattice MAC-ZKP

LTE – Long Term Endowment
WLC은 비증권성 검토 의견서 (법무법인 디라이트 조원희 변호사)

05. Scale-up

목표시장 및 경쟁사 현황 – 기술 가치 평가 $1.2M

Polkadot, Polygon 대비 자체 합의 알고리즘 보유, 확장성 솔루션 제공 및 양자 내성 특성 지원

진보된 기술력 기반, 시가총액 추월 목표

Projects	Goal	합의노드 선출방식	*독창적 합의 알고리즘*	합의 알고리즘 Anti-ASIC	*확장성 솔루션 제공 여부*	DeFi 솔루션 제공 여부	*양자내성 지원 여부**	최대 TPS (transactions per second)
Polygon	확장성 트릴레마 문제 해결	PoW (Ethereum)	***GHOST***	**X**	***Polygon (Layer-2)***	O	*X*	scalable
Cosmos/ Polkadot	*체인간 연결 프로토콜*	PoS	***BFT***	**X**	*인터체인 통신*	O	*X*	scalable
LiberVance	확장성 트릴레마 문제 해결	VRF	***VCA***	**O**	***DeSecure 복합체인 연결 기술***	예정	*오류정정부호 기반 양자내성 암호화*	scalable

<Cosmos, Polkadot 평가가치>

Cosmos 시가총액 약 4.4 조원
Polkadot 시가총액 약 8.8 조원
(23.11.20.coinmarketcap 기준)



<Polygon 평가가치>

Polygon 시가총액 약 10.1 조원
(23.11.20.coinmarketcap 기준)



< LiberVance 평가가치 >

Cosmos, Polkadot 그리고 Polygon 대비 진보된 기술력

1. 합의 프로토콜에서의 비교: VCA 합의 프로토콜의 이점
   - PoS 방식인 Cosmos, Polkadot 대비 높은 보안성
   - ASIC 저항성 기술 및 Time-varying 퍼즐 생성 기술을 적용하여, Polygon 대비 더 탈중앙적
   - BFT 방식인 Cosmos, Polkadot 대비 간단한 구성
2. 체인 연결 기술 측면에서의 비교: DeSecure 기술의 이점
   - 체인 보안 레벨 판단 기술 적용으로 Polygon, Cosmos, Polkadot 대비 높은 보안성

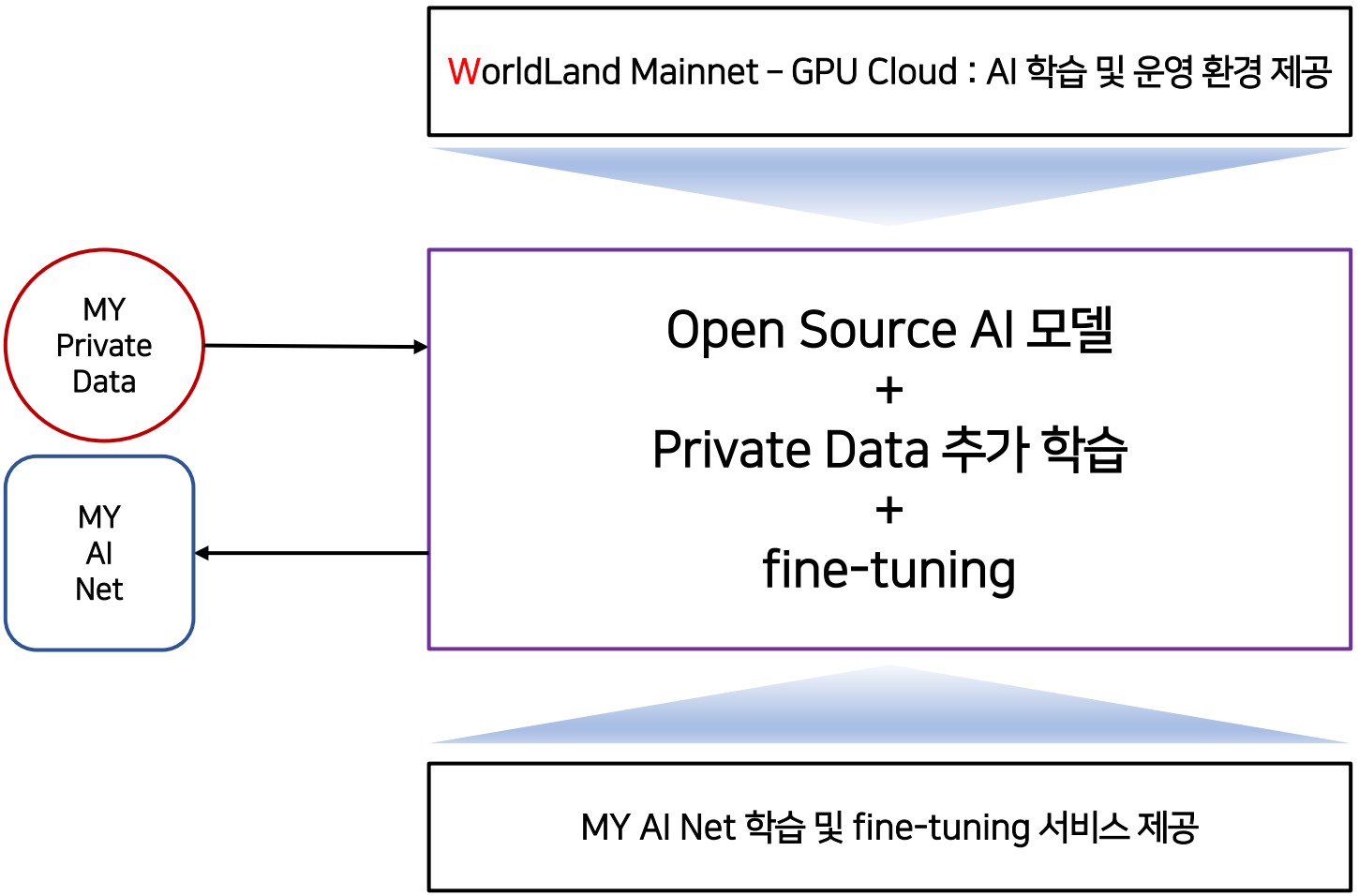
05. Scale-up

목표 펀딩규모 22,000 ETH ($26.4M)
Series A-1, Series A-2, Series B, Public Sale

Funding		$26.40	M$							
Spec	Round	ETH@WLT		Ether	Implied US$		WLT 개수		Lock Up (yr)	$WLT
Private Sales	Seed	0.00002	1%	1,000	$1,200,000.00	4.35%	42,000,000	56.99%	20% TGE 80% 4Yr Vesting	$0.03
	A-1	0.00010	5%	1,000	$1,200,000.00	4.35%	10,000,000	13.57%	2	$0.12
	A-2	0.00029	14%	2,000	$2,400,000.00	8.70%	7,000,000	9.50%	1	$0.34
	B	0.00055	28%	4,000	$4,800,000.00	17.39%	7,200,000	9.77%	1	$0.67
Public Sale		0.002	100%	15,000	$18,000,000.00	65.22%	7,500,000	10.18%	N / A	$2.40
							73,700,000	100.00%		$0.37

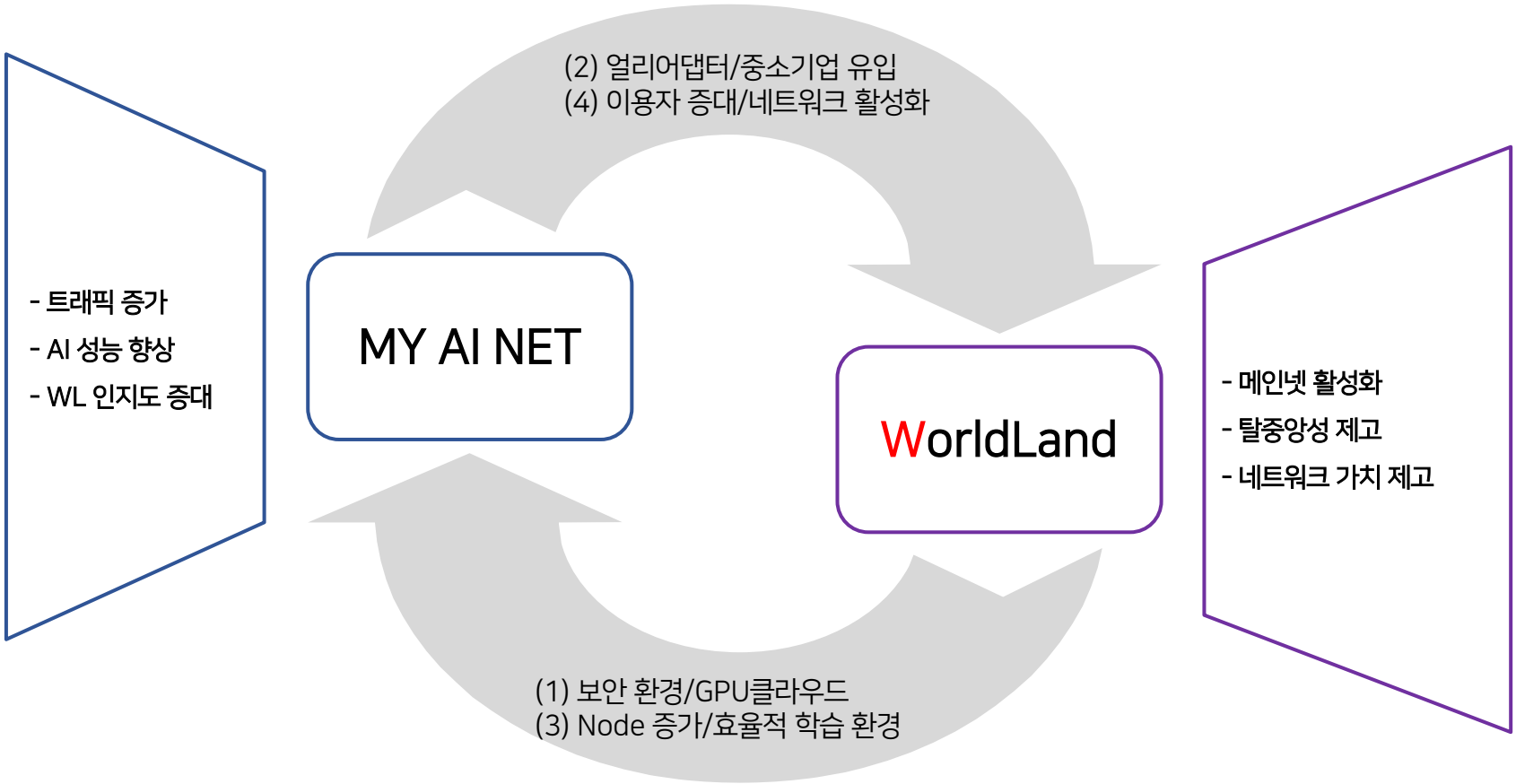
05. Scale-up : MY AI Network

WorldLand (+ Web3), MY AI Network 인프라 제공, 초거대 AI기업 독점 문제 해결
데이터 보안 / 학습 비용 등의 문제, 블록체인 및 보안 기술 통해 해결



05. Scale-up : MY AI Network

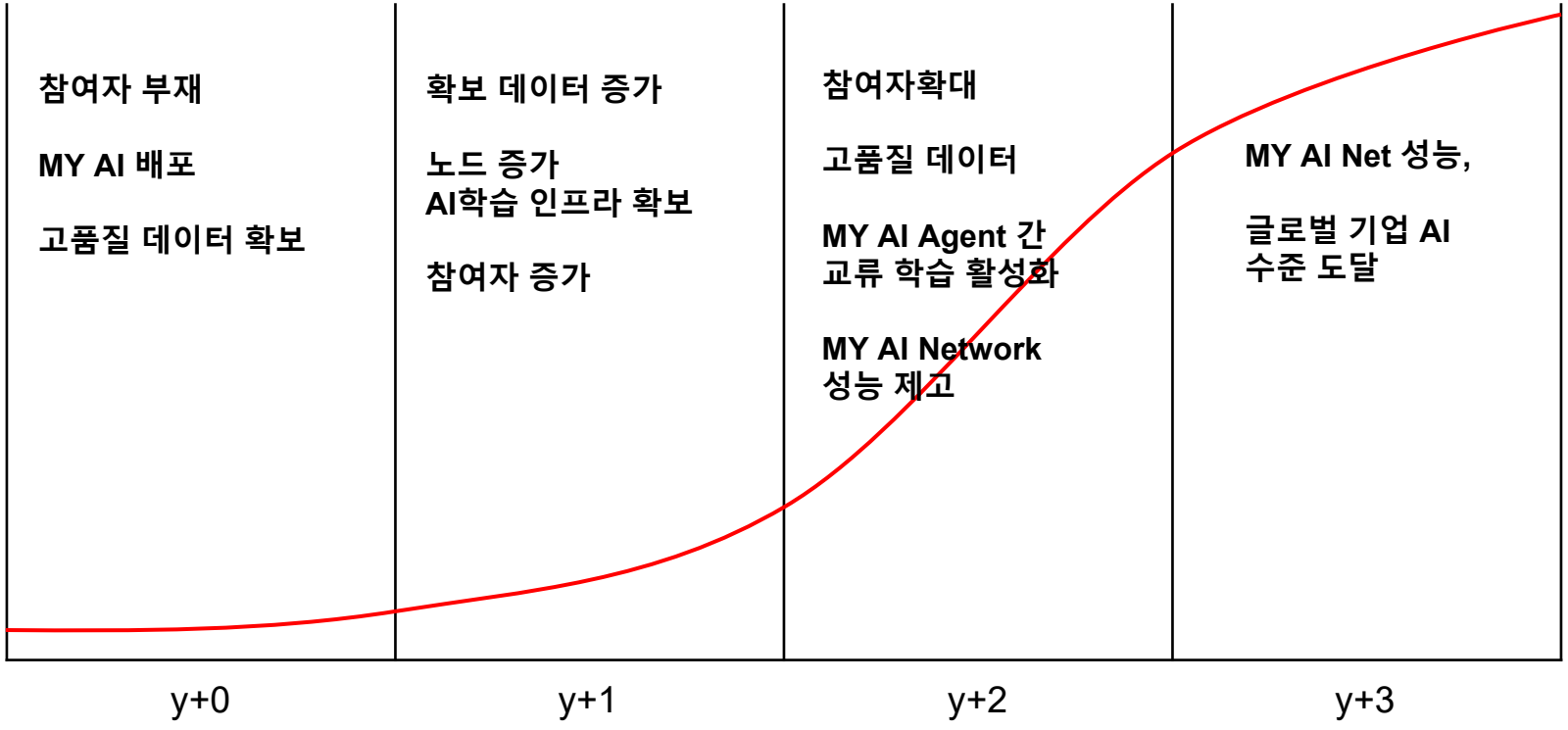
WorldLand 인프라 활용, MY AI Net 이용자 증대, 네트워크 활성화
활성화된 WorldLand, Node(GPU 자원) 증가, MY AI Net 서비스 활성화, 선순환 구조 구축



05. Scale-up : MY AI Network

개인 고급 데이터 학습 완료한 My AI Agent 수 증가, 교류 학습 기하급수적

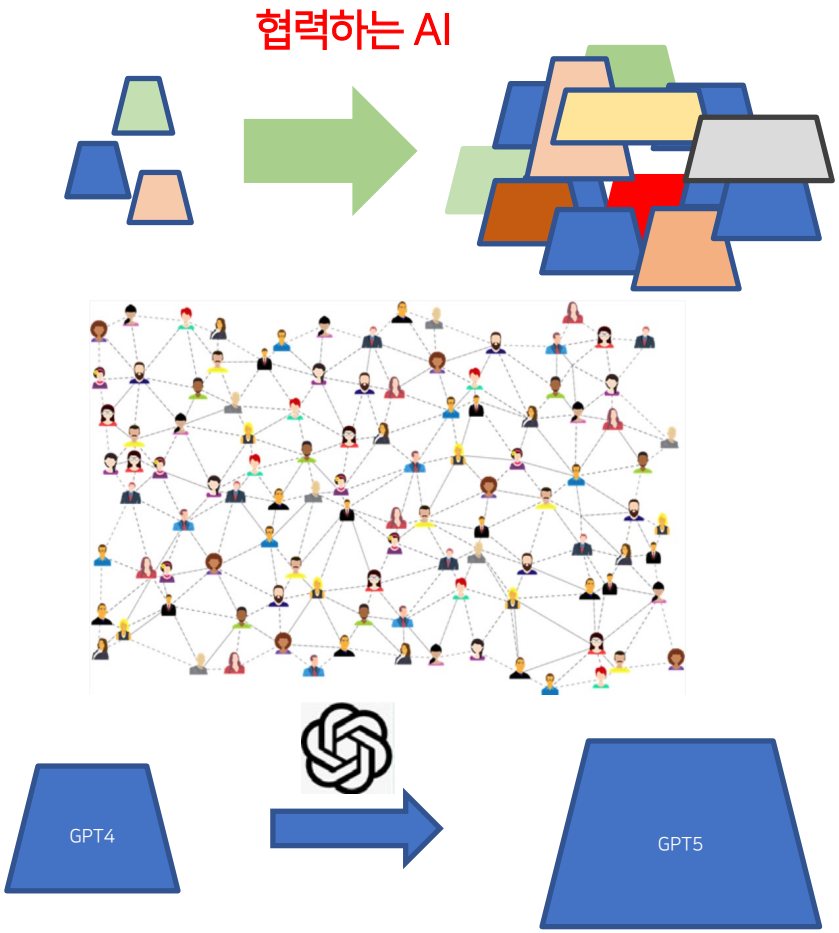
[MY AI Network 발전 과정]



WorldLand

05. Scale-up : MY AI Network

WorldLand MY AI Net 과 독점적 AI API 서비스, 누가 더 바르게, 누가 더 주도적으로 성장할까?



☐ 협력하는 AI

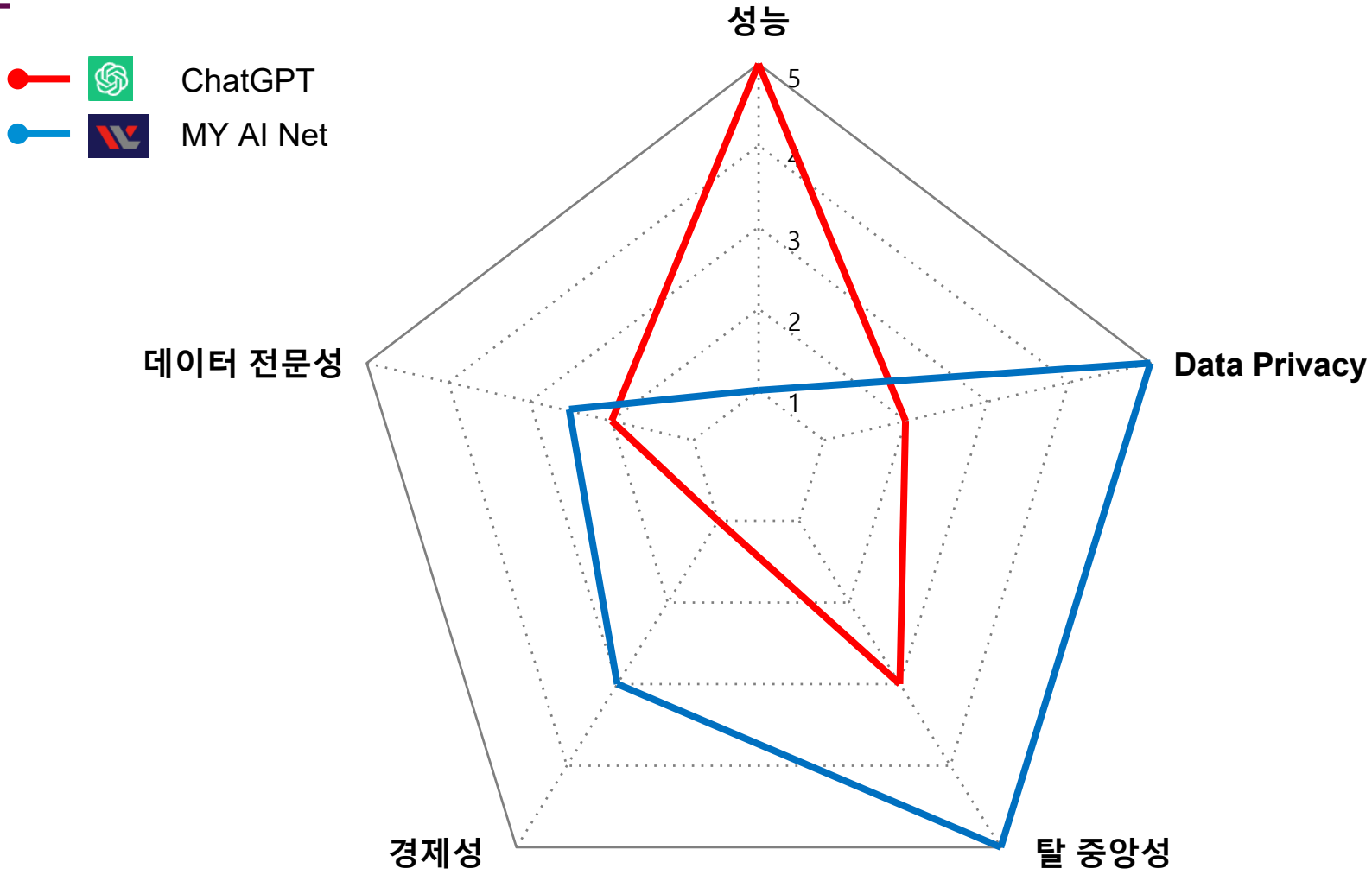
■ 다양한 참여자가 자신들의 AI를 가지고 직접 참여
■ 자신들의 프라이빗 데이터를 바탕으로 AI 학습
■ AI 모델 간 교류 학습을 통해 진화
■ 교류를 통해 생성된 경제적 결과물은 공평하게 분배

☐ 거대한 중앙집중 독점적 AI

■ 천문학적 비용을 바탕으로 획득한 데이터와 GPU 자원으로 경쟁력있는 AI 지속 개발
■ 경제적 효과는 독점

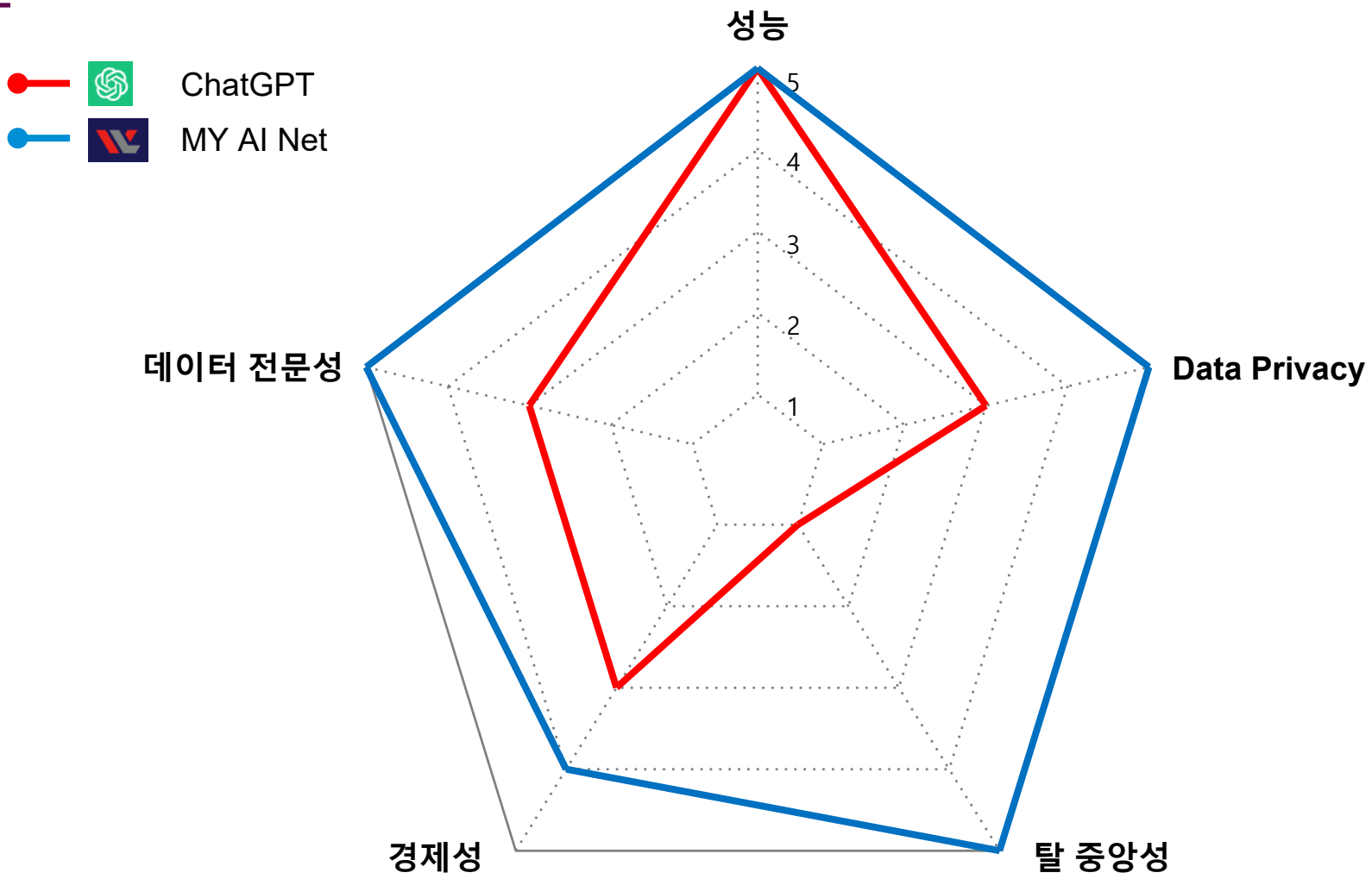
05. Scale-up : MY AI Network

MY AI Net vs. 챗GPT (현재)



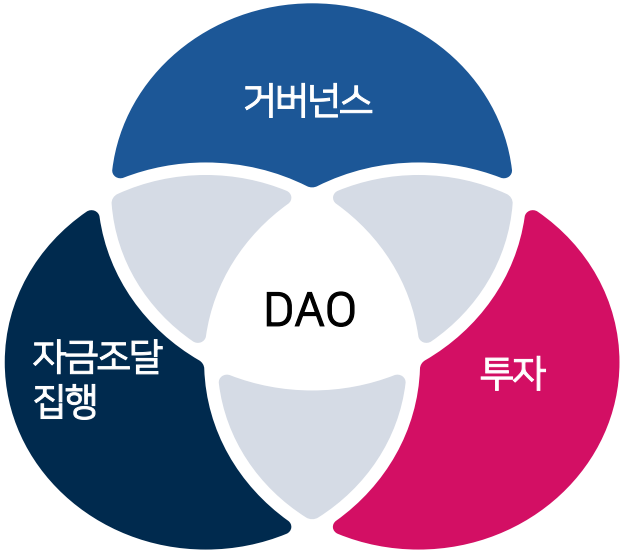
05. Scale-up : MY AI Network

MY AI Net vs. 챗GPT (3년후)



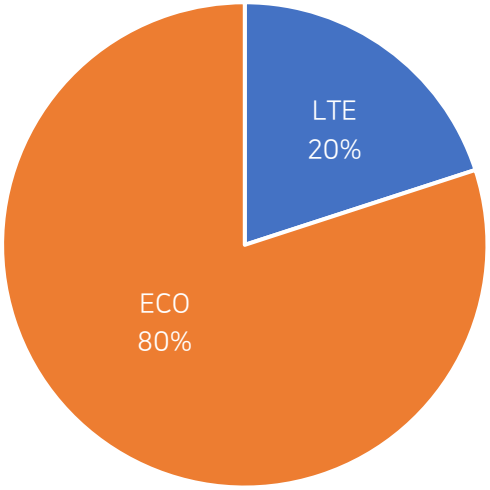
WorldLand DAO

1. WorldLand DAO는 재단 DAO 형태를 따른다.
2. 코인 지분 증명으로 중요한 의사를 결정한다.
3. WorldLand Ecosystem Growth Foundation 역할을 수행한다.
   - G직후 설립, 5년간 ECO생성 자금 집행.

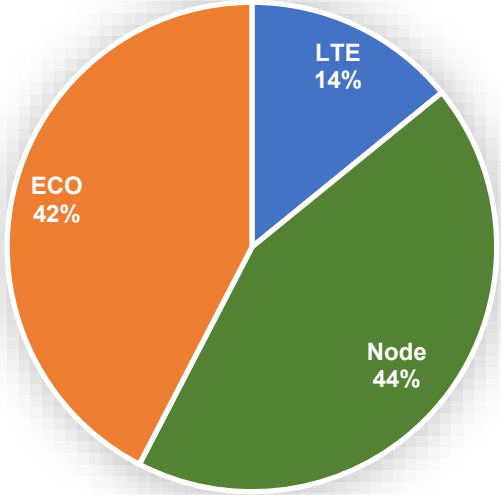


재단의 탈중앙화 전략 : ECO 육성 정책

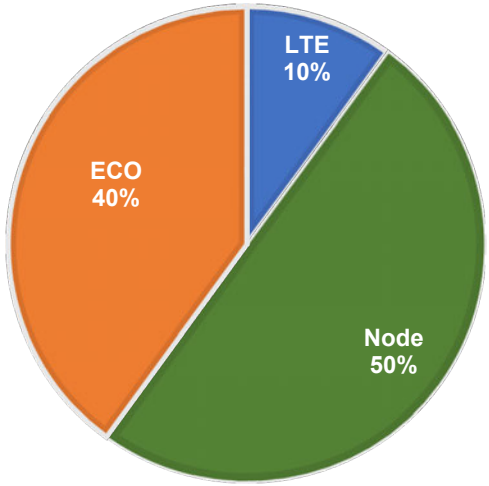
DAO 지분@제네시스



DAO지분@YR2

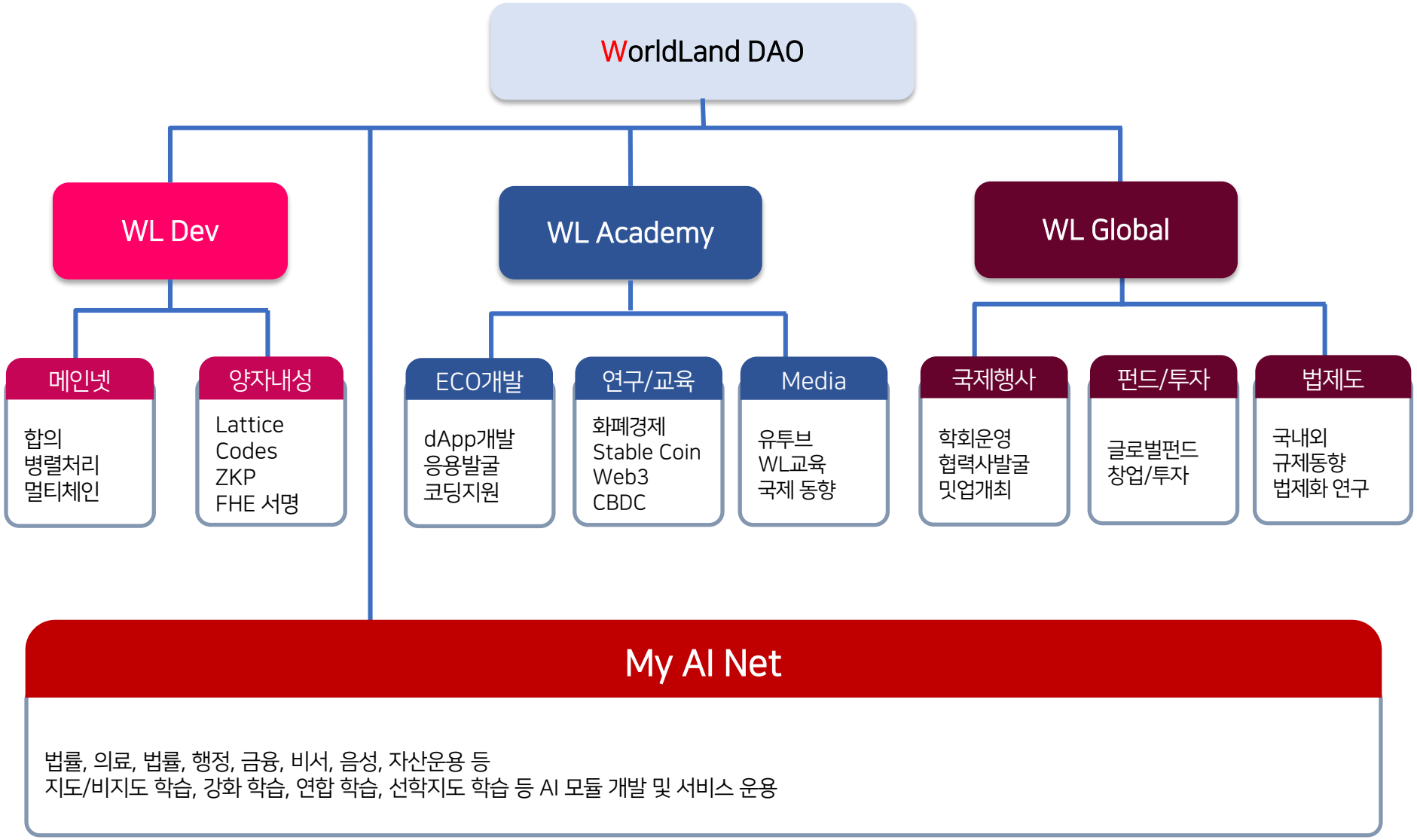


DAO지분@YR5



Long-Term Endowment (LTE)
ECO 육성 대상 사업: Wallets, NFT/게임/투표/Bridge, My AI Network, 메인넷, 연구, 교육, 투자 등

05. Scale-up


WorldLand DAO
WL Dev
WL Academy
WL Global
메인넷
합의
병렬처리
멀티체인
양자내성
Lattice
Codes
ZKP
FHE 서명
ECO개발
dApp개발
응용발굴
코딩지원
연구/교육
화폐경제
Stable Coin
Web3
CBDC
Media
유투브
WL교육
국제 동향
국제행사
학회운영
협력사발굴
밋업개최
펀드/투자
글로벌펀드
창업/투자
법제도
국내외
규제동향
법제화 연구
My AI Net
법률, 의료, 법률, 행정, 금융, 비서, 음성, 자산운용 등
지도/비지도 학습, 강화 학습, 연합 학습, 선학지도 학습 등 AI 모듈 개발 및 서비스 운용

WorldLand ECO 주요 이해 당사자들
WorldLand 생태계 內 플레이어들의 역할

Players in WorldLand



DAO 재단
- 토큰 발행
- 토큰 상장



투자자
- 토큰과 투자금 스왑
- 토큰 상장 시 EXIT가능



AI Net 개발사
- AI Net 프로토콜 완성 및 공개
- AI 프로토콜 설치 노드사업자에 제공



노드 사업자
- 노드 구매 or 직접설치 운영
- 전기사용 노드 운영
- 노드운영 수익
  - ✔ 주조차익, gas fee, AI서비스 fee



사용자
- 지갑
- dApp
- AI 학습 및 추론 서비스 사용자



dApp 사업자
- Smart Contract
  - ✔ My AI Net
  - ✔ 스왑서비스
  - ✔ Soul Bound Tokens
  - ✔ 데이터사업

05. Scale-up

WorldLand My AI Net ECO 재창출
다양한 AI dApp 생태계 內 탈 중앙 서비스 제공을 통한 생태계 재창출
다양한 dApp 사업자들이 오픈 WorldLand에 AI dApp 서비스 스마트 컨트랙트를 배포 (법률, 의료, 금융, 이상 징후 파악 등)

WorldLand Layer 2 네트워크 생태계 재창출 구조

My AI Net 응용 부분별 사업

Players in a Single MY AI dApp

My AI dApp 재단
- 토큰 발행 및 상장
- TVL 상승

AI Net 개발사
- 재단 개발비 지원
- AI 프로토콜 완성/공개/업그레이드 등 새로운 서비스 지속 개발
- Smart Contract 기반 새로운 AI 솔루션 제시

My AI dApp 사용자
- A토큰가치 = 가중치 가치
- A 토큰 매도/매수

↓

보다 똑똑한 My AI 사용

편익 증가

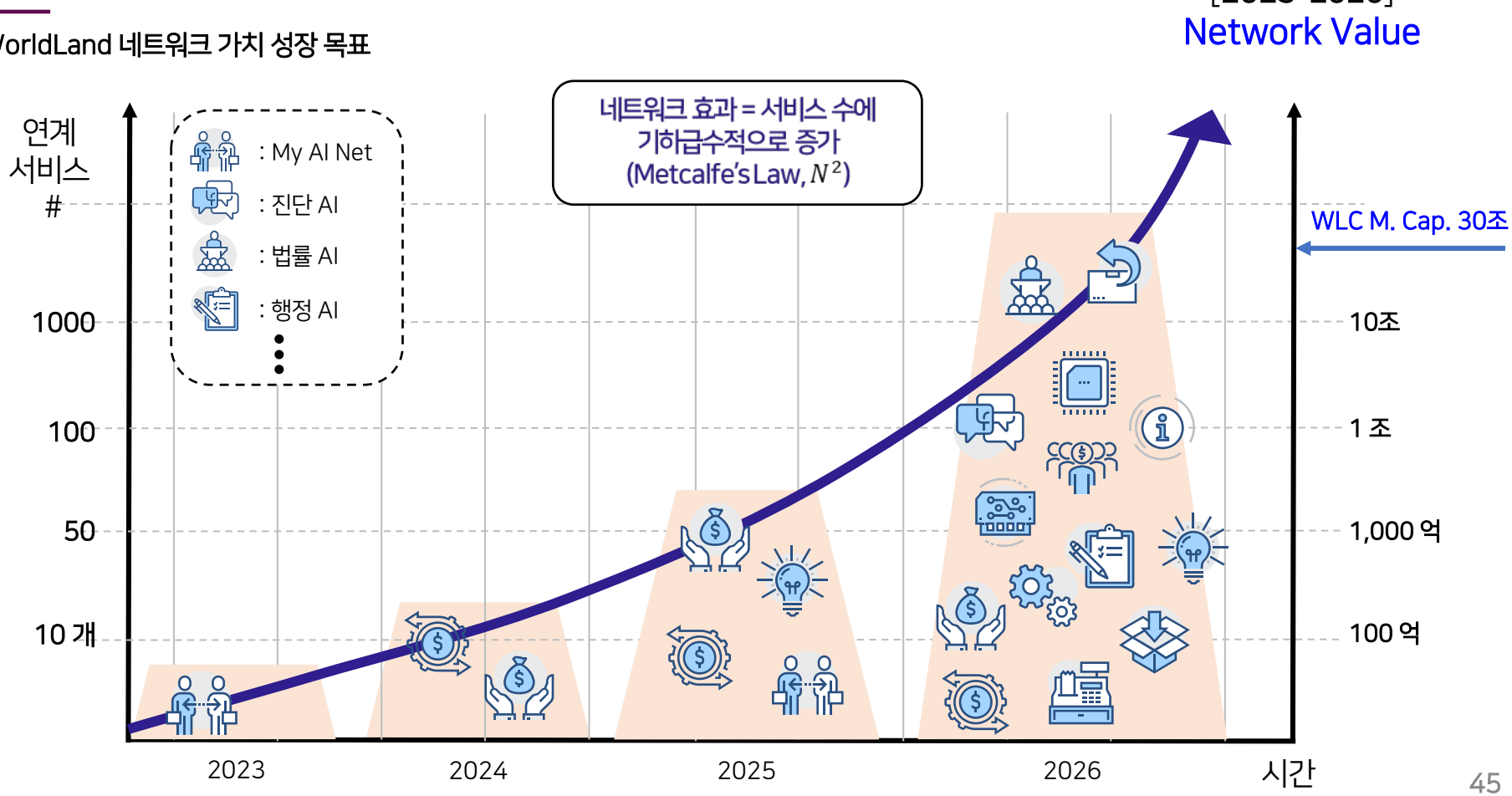
노드 사업자
- AI Agent 상호 교류
- TX fee, Gas fee
- AI 서비스 fee
- 노드 운영 수입 증가
- 주조차익

AI솔루션 공급자
- AI개발사, 대학/기업/기관 등 연구기관
- 다양한 AI모델, 데이터, 훈련, 자동화 솔루션 공급
- **법률, 의료, 금융, 게임, 자산운용 등**

05. Scale-up

론칭 후 재단의 ECO 생성 지원 자금 집행 결과는?
3년 후 시가총액 50조 가능

WorldLand 네트워크 가치 성장 목표



오픈 후 3년간 WLC 가격 및 시가 총액 시뮬레이션 (VS. Polygon)

Polygon 대비 기술적 우위를 바탕으로 공격적인 ECO 생태계를 구축하여 서비스 오픈 시점 기준 3년 경과 후 시가총액 역전

구분	Y+0	Y+1	Y+2	Y+3
WLC 가격	98원	224원	1,963원	689,343원
WLC 시가 총액	40억원	120억원	1,300억원	50조원
Polygon 가격	16원	16원	31원	2,125원
Polygon 시가 총액	1,612억원	1,612억원	3,120억원	21조 2,550억원

※ WLC 2023년~2026년 / Polygon 2019년~2022년
※ Polygon 관련 자료 출처 : Investing.com

06. Team

LiberVance

회사명	리버밴스㈜
소재지	광주 북구 첨단과기로 123, 광주과학기술원 산학협력연구관 307/308호
핵심사업	AI / 블록체인 융합 기술 개발
설립일	2020년 1월 24일
대표이사	이흥노
임직원	총 13명

주요 연혁

2019년 10월	예비창업패키지 수행
2020년 01월	GIST 교원 창업 기업으로 설립
2022년 04월	선보엔젤파트너스㈜ 투자 유치
2022년 06월	AI 제품/서비스 제작 지원 과제 수행
2022년 07월	TIPS 투자 유치
2022년 11월	민관공동 창업자 발굴육성사업 수행
2023년 06월	AI 제품/서비스 상용화 지원사업 수행

※ 투자 유치 금액 : 약 200백만원
※ 누적 지원 금액 : 약 934백만원



대표이사 : 이흥노

- 리버밴스㈜ 대표이사
- IEEE Trans. Cybernetics 편집위원 (국제 Top 4%)
- GIST 교수 / 인공지능연구소장
- 과기부 ITRC 블록체인 지능 융합 센터장
- (전) University of Pittsburgh, PA. USA, 전임교수
- (전) Hughes Research Lab, Malibu, CA, USA, 연구원
- 대통령 직속 정책기획위원회
- 블록체인경제센터장
- 센서지능화센터장
- (전)GIST 연구원장
- 총 285편 논문 : SCI 100여편 포함 국제 논문 약 200여편
  국내논문 104편
- 국내외 특허 약 60여건
- 해동학술상 (2019)
- GIST 연구상 / 대표기술상 3회
- 이달의 과학자상 (2014.01)
- 국무총리표창 (2022.04)

06. Team

○ 리버밴스는 GIST의 연구성과를 4차 산업혁명 기술로 구현하여 디지털 경제 체제를 선도하고자 합니다.

○ 리버밴스는 TIPS 및 정부지원과제 수행을 통해 AI와 블록체인 기술 역량을 향상 시키고 있습니다.

코어기술 / 보유 특허

블록체인 기술	ECCPoW – ECCBTC
	이중지불 방지 기술
센서 지능화 기술	지능형 센서 데이터에 대한 효율적 학습을 위한 Machine learning 기법 연구
	전파 신호 식별 기술
	Depth map 기반 무인 이동체 경로 설계 기술
	무작위 간섭을 이용한 초음파 이미징 기술
	DSP 알고리즘을 이용한 스펙트로미터
보유 특허	특허 10-2288769 부호-암호 화폐 시스템
	특허 10-2288776 블록체인의 거래검증시스템 및 블록체인 거래검증방법



전자신문 etnews Conference allshowTV English

경제·금융 전자·모빌리티 통신·미디어·게임 소재·부품 SW·보안 산업·에너지·환경 플랫폼·유통

이흥노 GIST 교수, IEEE 트랜젝션 온 사이버네틱스 편집위원 선임

발행일 : 2022-03-17 09:57 지면 : 2022-03-18 24면

광주과학기술원은 이흥노 전기전자컴퓨터공학부 교수(ITRC 블록체인 지능 융합센터 센터장)가 저명한 국제학술지 '미국 전기전자학회(IEEE) 트랜젝션 온 사이버네틱스 (TC·Transactions on Cybernetics)' 편집위원으로 선임됐다고 17일 밝혔다.

IEEE TC는 저널 임팩트 팩터 랭킹이 세계 톱 4% 수준으로 컴퓨터 사이언스와 인공지능(AI) 부분에서 139개 저널 가운데 6위(96.04%), 오토메이션&컨트롤 시스템 부분에서 63개 저널 가운데 2위다.

주요 지원 사업 현황

프로그램명 (시행부처/기관)	과제명	협약(수행)기간 (시작-종료일)	총사업비 (백만원)
예비창업패키지 (중소벤처기업부/창업진흥원)	기계학습 상황인지 DeSecure 블록체인	'19.10.01-'20.08.31.	55
AI제품·서비스 제작지원 (인공지능산업융합사업단)	AI 기반 AMM 알고리즘 및 IL 제어 기술을 통한 차세대 DEX 서비스 개발	'22.06.01-'22.11.30.	90
민관공동 창업자 발굴육성사업 (중소벤처기업부/창업진흥원)	리버밴스 레이어2 네트워크: Verifiable computation Algorithm(VCA)합의 기술에 기반한 Ethereum L2 네트워크	'22.11.01-'23.08.31.	114
창업성장기술개발사업 TIPS (중소벤처기업부)	Libervance Layer 2 기술개발-탈중앙 Layer 2 플랫폼	'22.07.01-'24.06.30.	555
AI 제품·서비스 상용화 지원사업 (인공지능산업융합사업단)	AI-DEX 서비스 사업화	'23.06.01- '23.11.30.	120

06. Team

○ GIST의 우수한 인력 및 연구 개발 인프라를 활용한 기술 중심 조직 구성

○ INFONET 연구실 및 GIST와 연계한 원천기술 확보 / 연구개발 인력 교육을 통한 핵심 원천기술 개발 역량 확보

대표이사

기획/사업/대외협력

송상훈 연구소장
서동훈 이사

연구개발

강주성 책임
최해웅 책임
홍종남 선임
안태호 선임
박가단 주임
장주찬 주임
문지혜 주임
박선우 주임
박준서 주임
Sadam. S 주임

행정지원

조윤경 주임

협력 기관/업체

광주과학기술원 Gwangju Institute of Science and Technology

AICA 인공지능산업융합사업단 Artificial Intelligence Industry Cluster Agency

광주광역시

Web 3.0, AI on Blockchain

Thank You

[문의]

이흥노 대표: ceo.libervance@gmail.com

서동훈 이사: dhseo.libervance@gmail.com

My AI 한국어 Foundation 모델 개발

- Decentralized My AI 서비스 -

연구팀 | INFONET 연구실

개발팀 | 리버밴스㈜

발표 장소: 마음 AI 사옥
날자: 2023년 12월 27일

CONTENTS

1. LiberVance 연구 개발 목표
2. 법률 AI란?
3. 연구 개발 추진 체계
4. 데이터셋 (Corpus) 확보 방안
5. My AI Development
    1. Pre-training / Fine-tuning
    2. RAG development
6. Cowork Process
7. Q&A

마음 AI 기술 세미나

WorldLand My AI Network 개요

발표자 | 이흥노 대표

CONTENTS

1. LiberVance 연구 개발 목표
2. 법률 AI란?
3. 연구 개발 추진 체계
4. 데이터셋 (Corpus) 확보 방안
5. My AI Development
    1. Pre-training / Fine-tuning
    2. RAG development
6. Cowork Process
7. Q&A

01. LiberVance 연구 개발 목표

Worldland My AI Network

왜 My AI 소유가 필요한가?

- 모델의 주권적 소유자는 데이터주권/생산성/경쟁력 확보, 경제적 성장 가능

My AI Network란 무엇인가? 왜 필요한가?

- 개별 My AI는 글로벌 중앙집중형 AI에 비해 성능이 열악할 수 있으나
- 상호 협력하는 My AI, 데이터 가치와 모델 성능을 제고하는 **Decentralized AI**

왜 Web3 블록체인 적용이 필요한가?

- **원천 데이터의** 기원/접근권한/지재권 **투명관리 및 보안성 제고**
- 스마트계약과 토큰에 의한 생성 **디지털 콘텐츠 자산화** 및 경제성 제고
- AI모델 훈련 및 운영 필수재 **GPU 클라우드 확보**
- AI시스템 결정에 대한 **신뢰성 제고**
- 분산자율조직(DAOs)에서 AI활용 거버넌스 프로세스 최적화

01. LiberVance 연구 개발 목표

AI를 Web3 블록체인에서 해야 하는 이유는 무엇인가?

- Privacy를 높여준다구? 클라우드에서 하는게 더 안전하고 믿을만한 Private AI 아냐?
- Private (사내)서버에서 사내데이터 학습, 가장 private 하지 않아?
- Cloud Server (아마존,구글)에서 학습하고 거기 서버를 사용해도 되지 않아?
- chatGPT 믿고 맡기면 쉽고 빠른데 뭐하러 블록체인에서 해?
  - chatGPT에 사내 data 올려서 학습시키고, 내 AI API를 공개해서 소득을 만들면 되잖어?
- 도대체 블록체인을 이용하면 My AI의 데이터 프라이버시가 뭐가 어떻게 더 좋아진다는 거야?
- LLaMa, Falcon, FlanT 등 많은 LLM이 존재하고 Open Source 코드로 공유되고 있는데, My AI Network이 무슨 강점이 있어?

01. LiberVance 연구 개발 목표

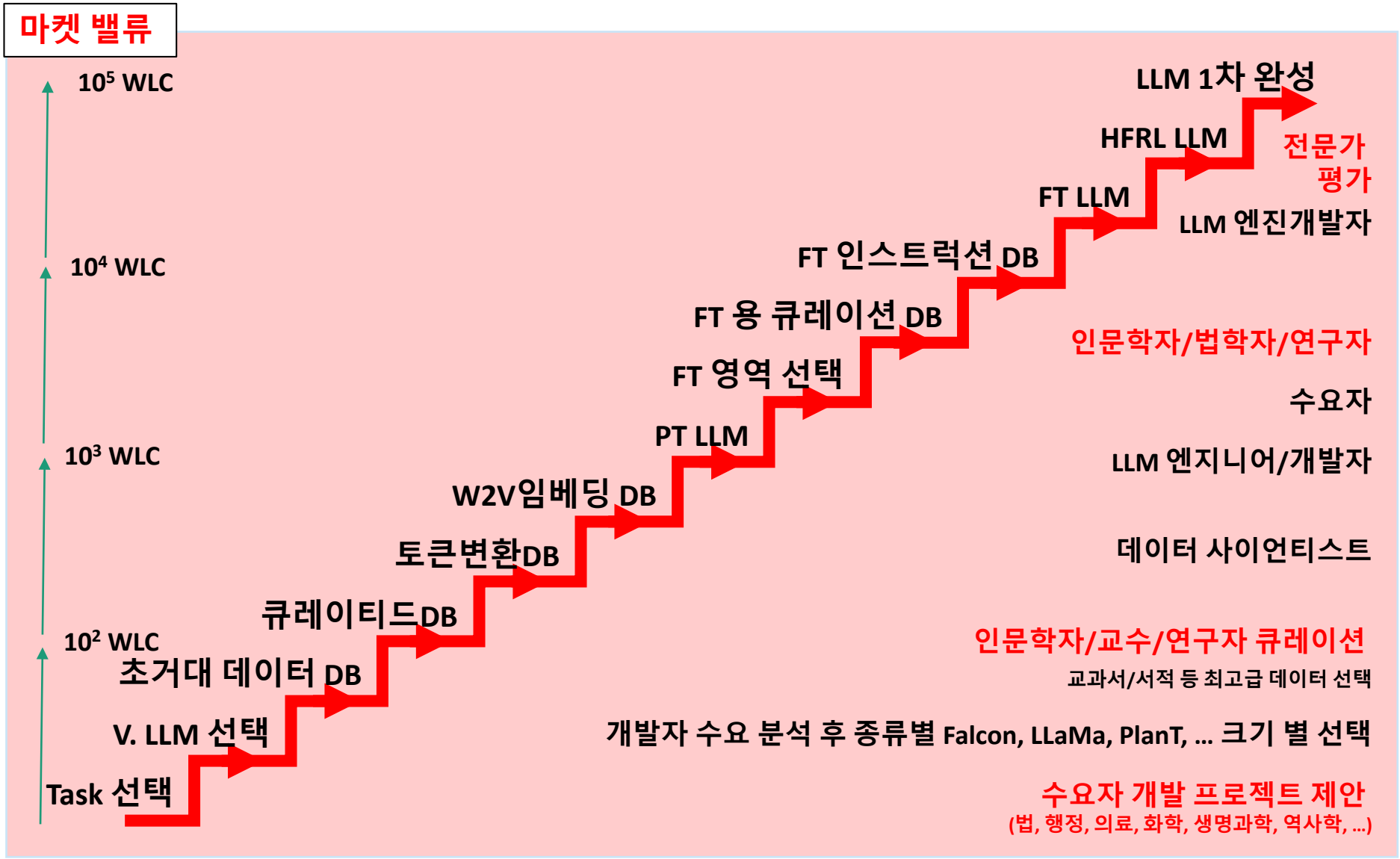
생성형 AI 개발 과정

- chatGPT는 최강 개발자가 사용자가 편하게 사용할 수 있도록 모든 과정을 자동화한 시스템을 구축 해 놓았음.
  - 사용자는 단순히 마이 데이터를 올리면 My AI Agent를 갖게됨
  - 전세계 사람들은 데이터 올리고 사용하는 사용자로 전락

- 생성형 AI 개발 과정과 구조를 해부해 보면 매우 복잡한 시스템으로 구성됨
  1. Task 정의
  2. Select LLM 모델
  3. 데이터를 끌어모아 Task에 맞게 큐레이션
  4. Curated 데이터 DB
  5. 토큰 데이터 변환 DB
  6. W2V임베딩 통해 벡터 DB
  7. Selected LLM에 프리트레이닝 --> PT LLM
  8. PT LLM 파인튜닝 → FT LLM
  9. 사람/chatGPT 검수 → HFRL or GFRL
  10. Task에 알맞은 LLM 완성
  11. 많은 사람들이 사용해서 추가 학습 DB 구축
  12. GPU 클라우드에서 주기적으로 추가 학습

01. LiberVance 연구 개발 목표

특화 생성형 AI 개발 과정과 마켓 밸류 상승 과정



01. LiberVance 연구 개발 목표

Worldland My AI Network 탈중앙경제 시스템

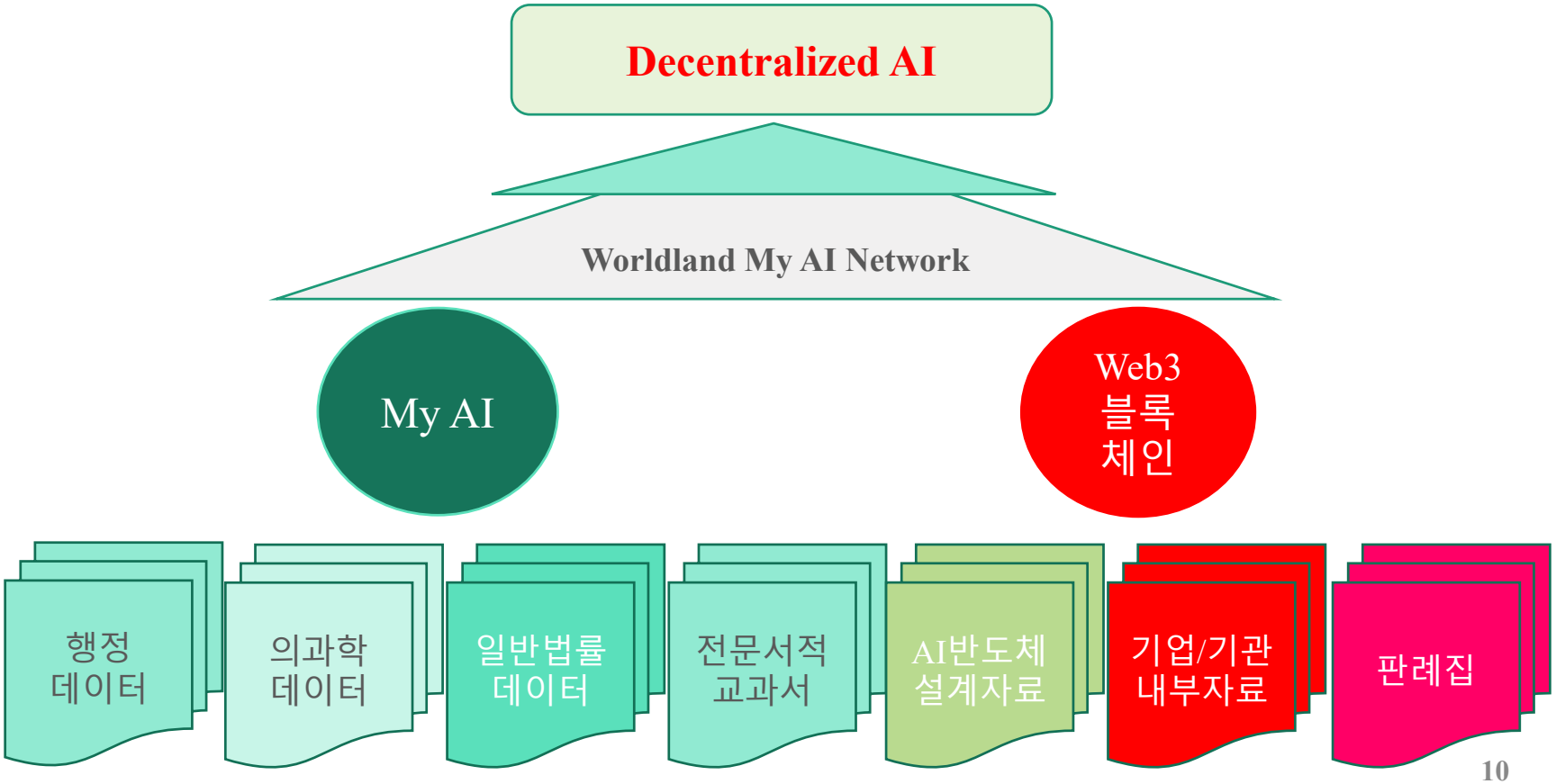
- 각 단계 결과물에 마켓 벨류를 만드려면 필요한 것은?
  - 각 단계 디지털 결과물을 시장에 내어 놓고 팔 수 있어야 함
  - 디지털 결과물을 토큰화하여 판매하는 플랫폼 필요 → 월드랜드 AI Market
  - 판매하려는 자와 수요자가 만날수 있는 마켓플레이스 필요

- 월드랜드 DAO의 역할
  - 다양한 참여자가 경제 활동 할 수 있도록 기반을 구축을 지원 함.
  - 각 Task별로 LLM(V. LLM, PT LLM, FT LLM, ...)과 데이터 매매 플랫폼 개발
  - LLM 및 데이터 마켓플레이스 개발 지원 필요
  - PT과 FT에는 GPU가 필요
  - GPU 자원을 확보하여 제공하는 노드운영자 지원 필요
  - My AI 앱 개발 전과정에 전문성을 제공하고 전문가 지원 필요
    - 데이터 큐레이터 및 DB 구축자 (인문학자, 데이터 사이언티스트)
    - PT/FT전문가 (AI 엔지니어, 개발자)
    - FB제공 전문가 (변호사, 의사, 교수, ... 등)
    - My AI 앱 사업가 (법률, 행정, 의료 문답 서비스로 사업을 영위)

02. My 법률 AI 서비스란?

My AI Network 서비스:
AI 모델의 주권적 소유자는 데이터주권/생산성/경쟁력 확보를 통한 경제적 성장 가능
상호 협력하는 My AI, 데이터 가치와 모델 성능을 제고하는 **Decentralized AI** 서비스 제공

Worldland My AI Network 서비스



02. My 법률 AI 서비스란?

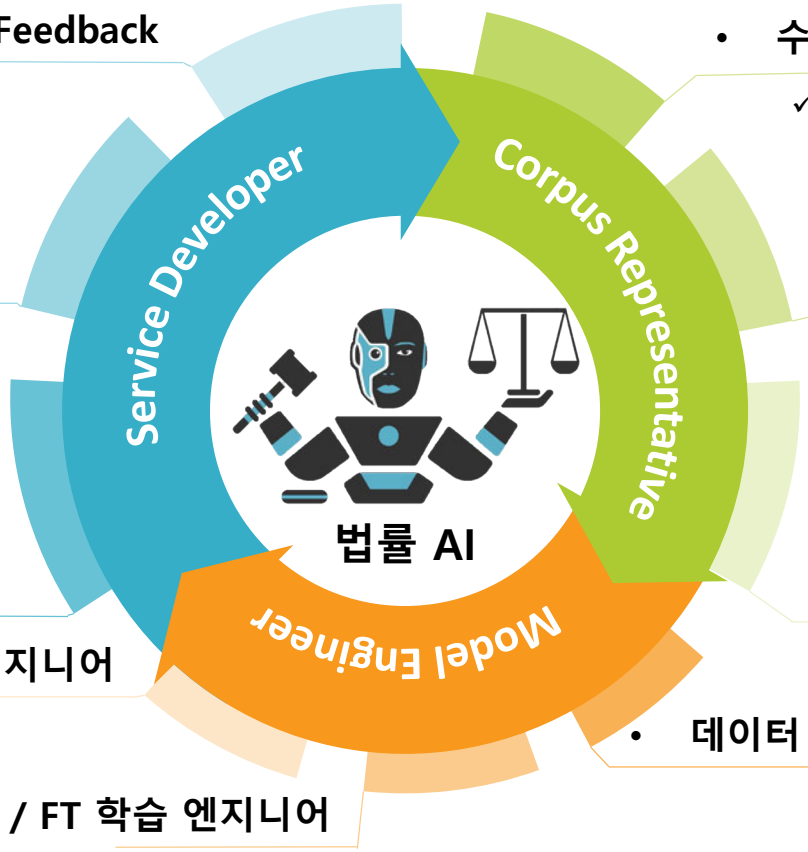
기술 스택 관점 LiberVance 차별성:
기반기술, 특화기술 및 응용기술 기반 **Decentralized AI** 플랫폼 구현



02. My 법률 AI 서비스란?

My AI 네트워크 서비스 – My '법률' AI 서비스

법률 AI 모델을 위한 다양한 참가자 : **Corpus 관계자, Model Engineer, Service Developer** 등
참가자 간 유기적 협력을 통한 **My 법률 AI 서비스** 제공

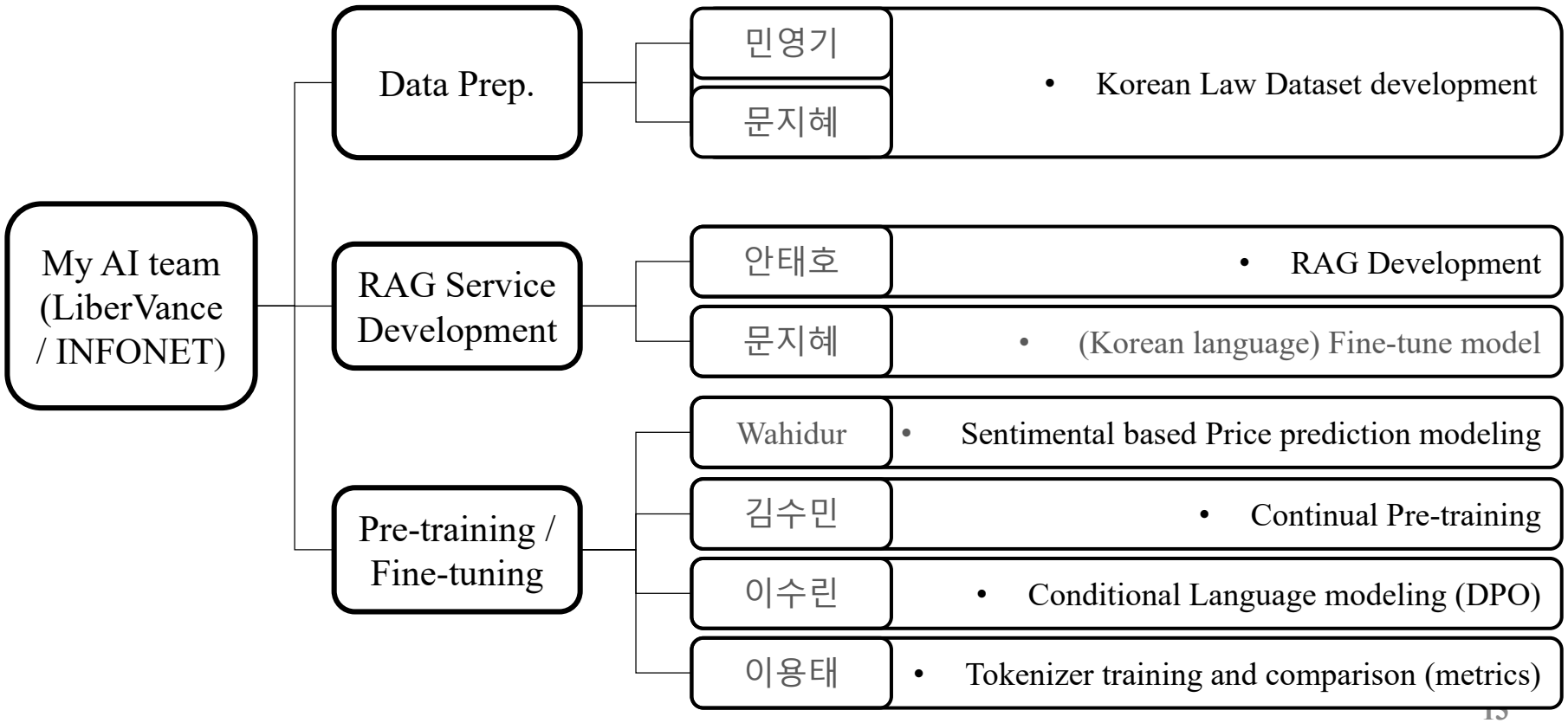


- 전문가 평가 및 Feedback
- LangChain 프레임워크
- PlugIn 엔지니어
- RAG 엔지니어
- PT / FT 학습 엔지니어
- 수요 기반 프로젝트 제안자
  - ✓ 의료, 법, 행정, 교육, 연구, 수학, 전기공학, 화학, 생명과학, 역사학
- 데이터 제공자
  - ✓ 출판사, 신문사, 잡지사
- 인문학/교수 데이터 큐레이션
- 데이터 사이언티스트

03. 연구개발 추진체계

My AI Development Team:
Market Price 예측, 법률 분석, My AI 모델 학습 및 평가 전담팀으로 구성하여 진행 중

My AI Development Team:

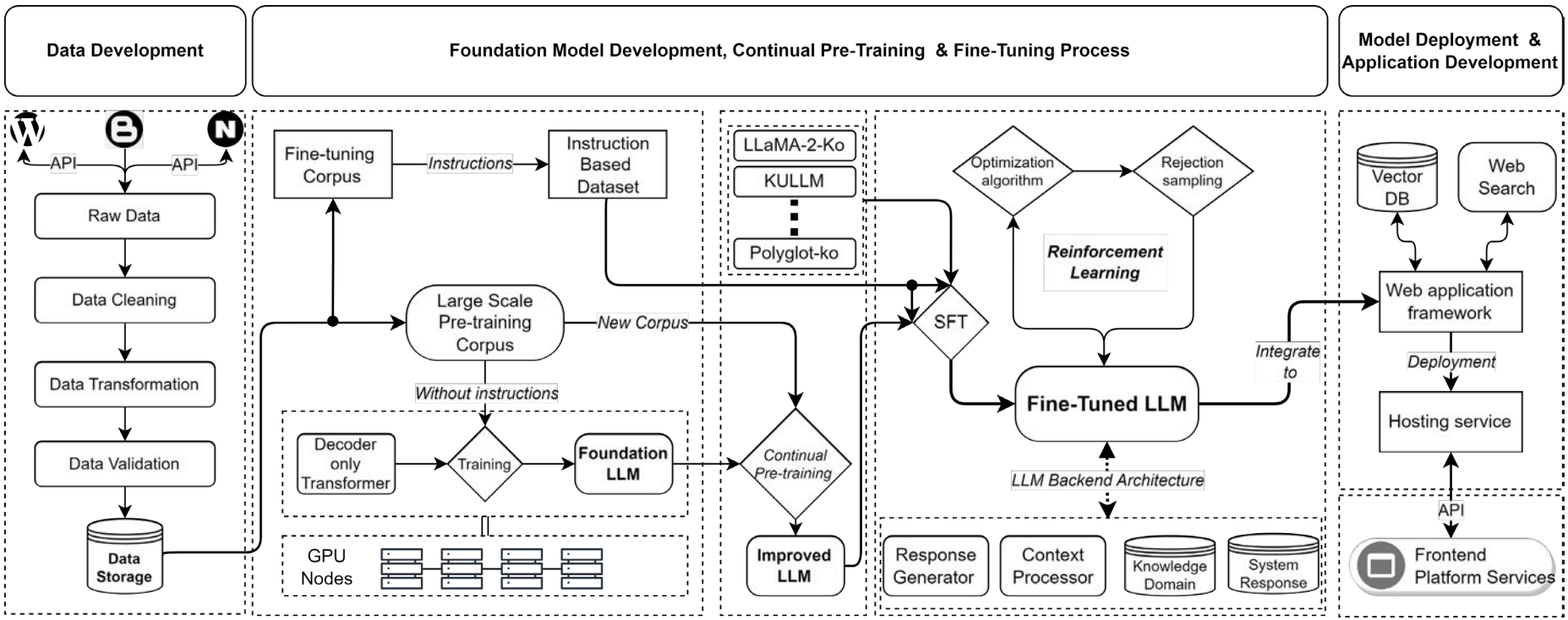


03. 연구개발 추진체계

My AI Development Work Flow:
각 팀별 유기적 연계를 통한 My Foundation AI model 개발

My-AI Work Flow:

- 국문 법률 Dataset 확보
- 법률 AI 모델 Pre-training / Fine-tuning 기법 연구
- RAG 기반 신뢰성 확보 기법 개발

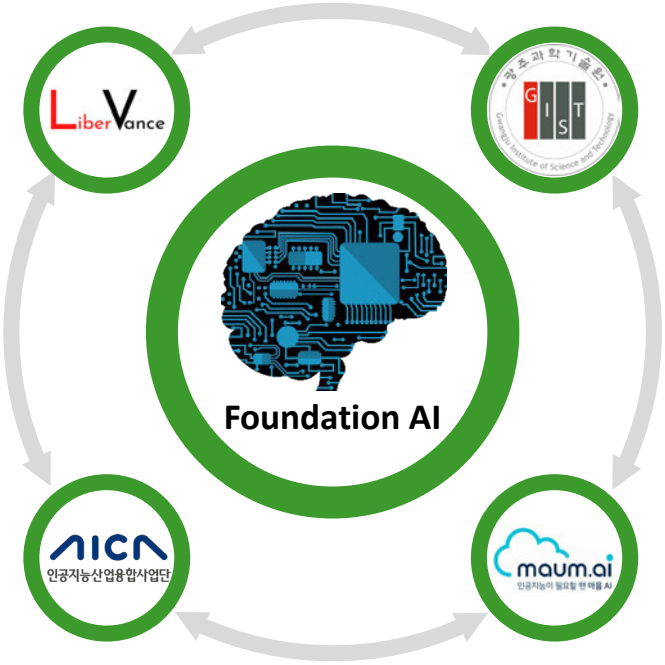


03. 연구개발 추진체계

Hardware resource based ML DebOps:
AICA, 마음 AI, Libervance 간 연계를 통한 ML DebOps 실행

Foundation model 개발을 위한 4자 MoU :

- "리버밴스 – AICA – GIST – 마음 AI" 4자간 MoU
- A100 기반 Hardware resource 제공 예정





전자신문 etnews Conference allshowTV English

경제·금융 전자·모빌리티 통신·미디어·게임 소재·부품 SW·보안 산업·에너지·환경 플랫폼

리버밴스-AICA-GIST-마음AI, 파운데이션 모델 공동 개발 4자 업무협약 체결

발행일 : 2023-12-11 09:10

<이홍노 리버밴스 대표와 김준하 인공지능산업융합사업단장, 유태준 마음AI 대표(왼쪽부터)가 파운데이션 모델 공동 개발 4자 업무협약을 체결하고 있다.>

블록체인 메인넷과 지능형 센서 융복합 기업 리버밴스(대표 이흥노)는 인공지능산업융합사업단(AICA), 광주과학기술원(GIST), AI 기업 마음AI와 리버밴스 서울사무소에서 파운데이션 모델(대규모 데이터를 기반으로 학습한 초거대AI) 공동 개발을 위한 업무협약을 체결했다고 11일 밝혔다.

이번 4자 협약은 △파운데이션 모델 개발을 위한 AI 학습 △AI 학습 환경 제공 △향후 파운데이션 모델 개발 협의체 운영 활성화 등을 목표로 한다.

04. 데이터셋 (Corpus) 확보 방안

Dataset (Corpus) 확보 방안:
법률 정보 크롤링 기반 국문 Dataset (Corpus) 구축

말뭉치 corpus:

- 현행 법령
  - 건수 : 총 210,455건
  - 출처: 국가법령정보 공동활용
  - 형식: CSV(172MB), JSON(진행중)
- 판례
  - 건수 : 총 85,830건
  - 출처: 국가법령정보 공동활용
  - 형식: CSV(897MB), JSON(진행중)
- 헌법 재판소
  - 건수 : 총 648건
  - 출처: 헌법재판소 분야별 판례
  - 형식: JSON(5.47MB)

질문-답변 corpus :

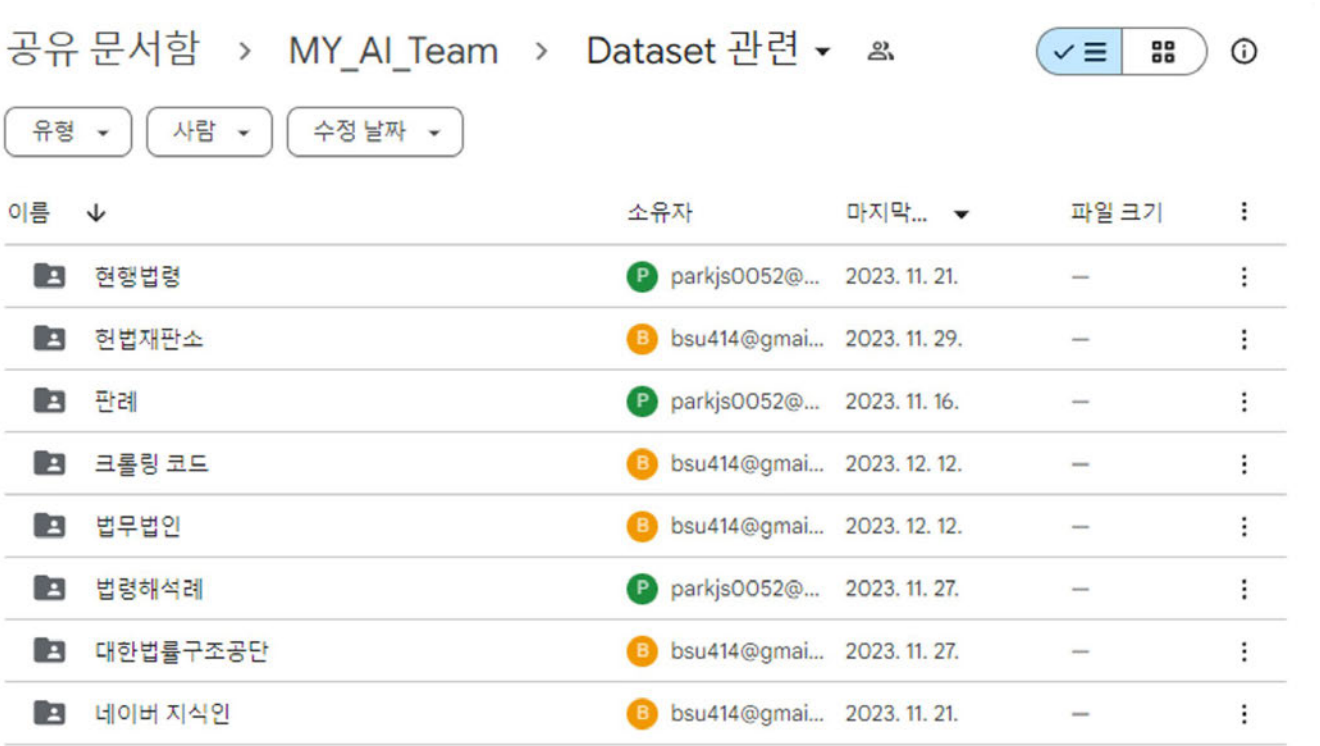
- 대한법률구조공단 – 법률상담사례
  - 건수 : 약 1만건
  - 출처: 대한법률구조공단 법률상담사례
  - 형식: JSON(22.2MB)
- 대한법률구조공단 – 사이버상담
  - 건수 : 총 2,639건
  - 출처: 대한법률구조공단 사이버상담
  - 형식: JSON(9.12MB)
- 법무법인 로시컴(lawsee), 로톡(lawtalk)
  - 건수 : lawsee(약4.1만건), lawtalk (진행중)
  - 출처: lawsee, lawtalk
  - 형식: JSON (lawsee-111MB, lawtalk - 진행중)
- 네이버 지식인 약 6만건, 법령해석례 총 7,672건 확보

04. 데이터셋 (Corpus) 확보 방안

Dataset (Corpus) 확보 방안:
법률 정보 크롤링 기반 국문 Dataset (Corpus) 구축

말뭉치 및 질문-답변 corpus:

- 법률 데이터셋 크롤링 및 Google Drive DB화 진행 중



04. 데이터셋 (Corpus) 확보 방안

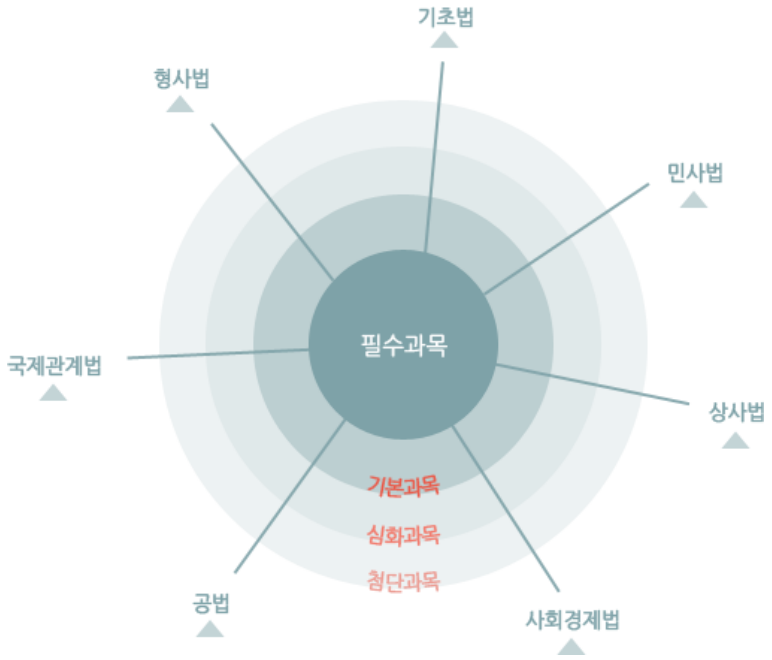
Dataset (Corpus) 확보 방안:
법률 정보 크롤링 기반 국문 Dataset (Corpus) 구축

법학 대학/대학원 커리큘럼 기반 Corpus set

- 법학계열 커리큘럼 확보 및 DB화 진행 예정

- 법학대학 커리큘럼 *
  - 법학개론, 법률문장론, 민법총칙
  - 헌법, 채권각론, 채권총론, 형법총론, 상법총론, 국제법1, 협업각론, 법사회학
  - 물권법, 친족법, 회사법, 행정법, 법철학, 노동법, 민사소송법, 세법, 재산법특강, 회사법특강, 입법학, 법과정치, 인권법,
  - 상속법, 행정구제법, 산업재산권법, 법의학, 지방자치법, 환경법, 국제사법, 증권거래법,

- 법학대학원 커리큘럼 *
  - 기초법전공, 헌법전공, 민사법전공, 상사법전공, 형사법전공, 행사법전공, 국제법전공 등



[*] 법학대학 커리큘럼, https://law.snu.ac.kr/page/Univ_faculty_curriculum.php

마음 AI 기술 세미나

AI Model Development team
- Pre-training & Fine-tuning -

발표자 | 강주성 책임

CONTENTS

1. LiberVance 연구 개발 목표
2. 법률 AI란?
3. 연구 개발 추진 체계
4. 데이터셋 (Corpus) 확보 방안
5. My AI Development
    1. Pre-training / Fine-tuning
    2. RAG development
6. Cowork Process
7. Q&A

05-01. My AI Model Development

My AI Model Development

Challenge 0

Sentimental Analysis

SNS based semantic analysis, targeting for crypto. price prediction.

Challenge 1

Korean Tokenizer

Specialized Tokenizer development for Korean Language

Challenge 2

Continual DAP

Domain transfer w.o. Catastrophic Forgetting

Challenge 3

Preference Learning (w. DPO)

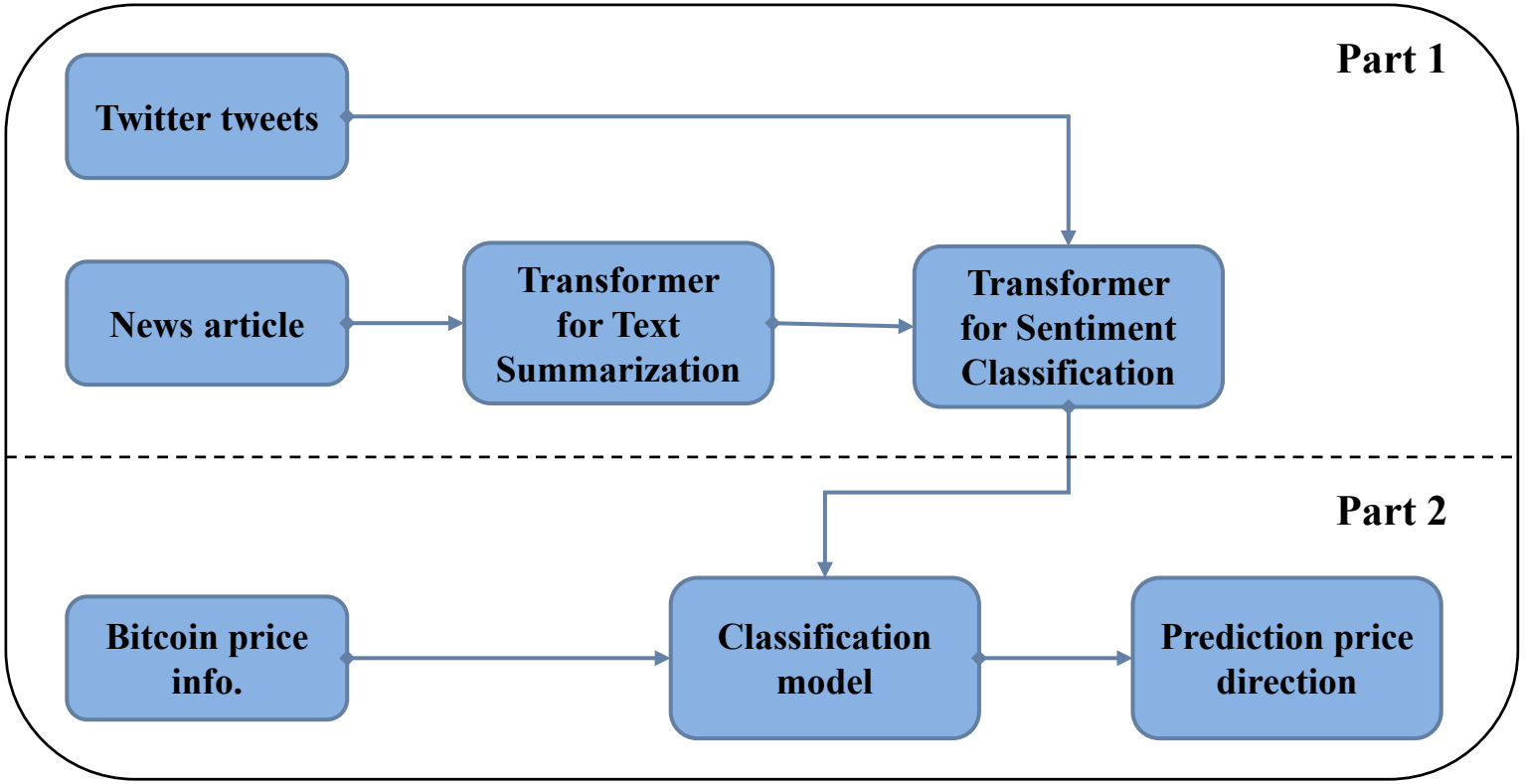
Enhancing model from human preference

Starting from our experience of the semantic analysis in human SNS corpus, we plan to develop our My AI model into a Korean Law-based Model by addressing these challenges, i.e. **1) Korean Tokenizer, 2) Continual Pre-training, and 3) Human Preference Learning.**

05-01. Sentimental based Price prediction modeling [1/4]

Goal of Sentimental Analysis :

My AI project – User semantic analysis based on SNS data.



Sentimental Analysis topology

05-01. Sentimental based Price prediction modeling [1/4]

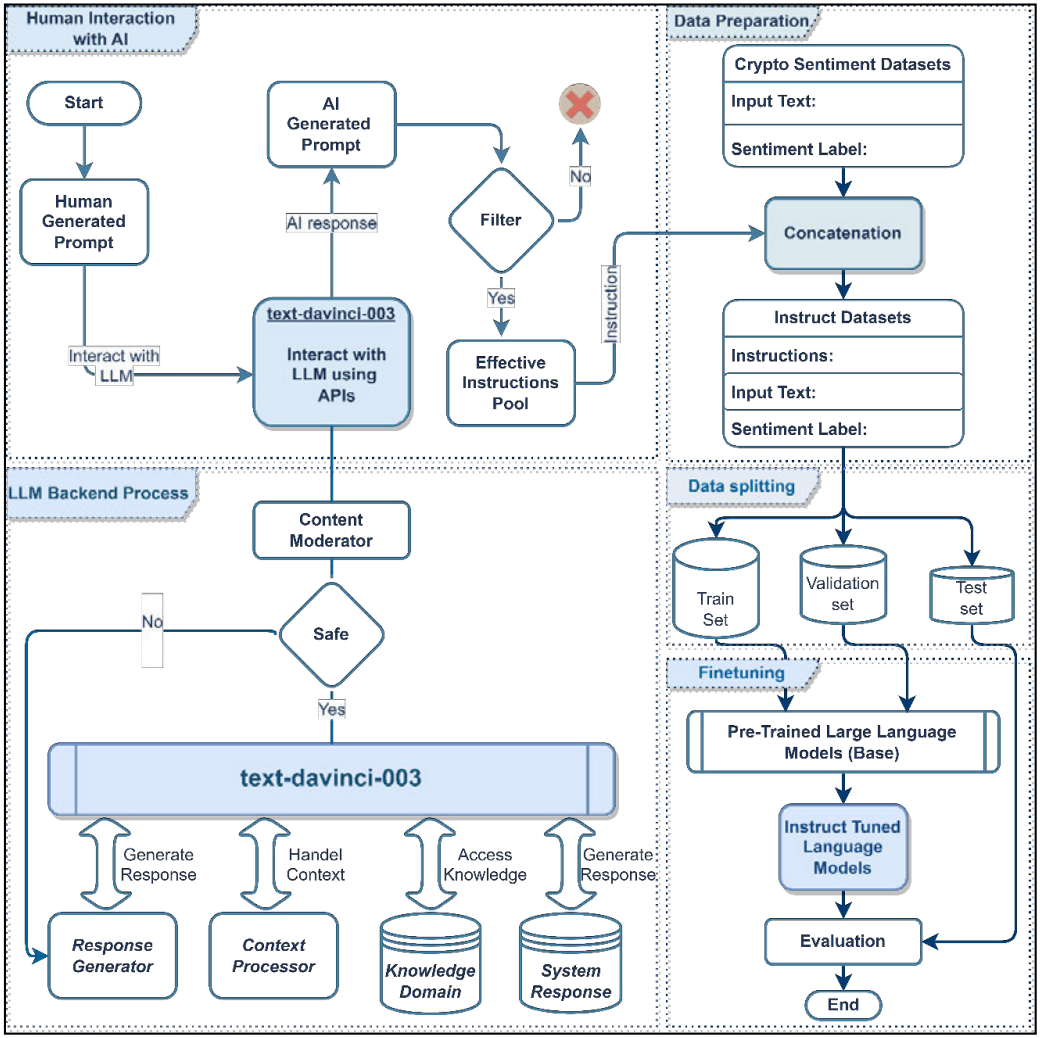
Graphical Abstract Figure:

- Instruction tuning based fine-tuning approach

Research questions:

1. Effectiveness of fine-tuning
2. Effect of prompt quality
3. Relation between instruction tuning and model size
4. Effect of training corpus size

Type	Prompt
Shot and simple	*Please detect the sentiment.*
	Detect the sentiment of the text.
	Please detect the sentiment of the given text.
Long and complex	*Classify the sentiment of the provided cryptocurrency related social media posts or messages.*
	Determine the emotional tone of the given text, which primarily revolves around cryptocurrencies and their associated concepts.
	Categorize the sentiment expressed in the provided dataset consisting of the text snippets related to cryptocurrency and computer science, focusing on capturing positive or negative sentiments.



05-01. Sentimental based Price prediction modeling [1/4]

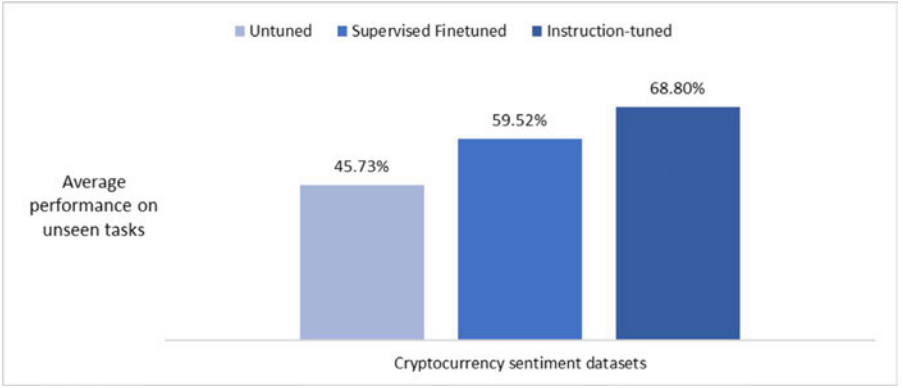
Result 1&2

– Effectiveness of fine-tuning :

– Effect of prompt quality :

Effects of fine-tuning

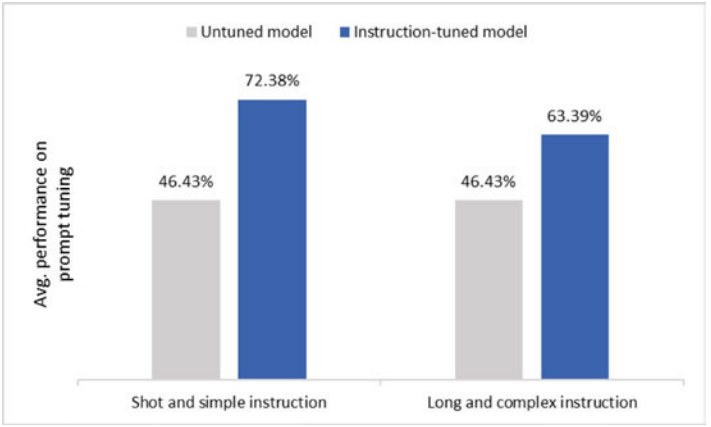
- 40% accuracy score improves after fine-tuning.
- IT (instruction-based) fine-tuned models outperformed SFT (Supervised Fine-tuning) by 16%



Zero-shot performance analysis

Effect of prompt quality

- The vanilla model's performance is consistent for both classes.
- The fine-tuned model performs 12% better with short and clear instructions.



Performance analysis across various prompts.

05-01. Sentimental based Price prediction modeling [1/4]

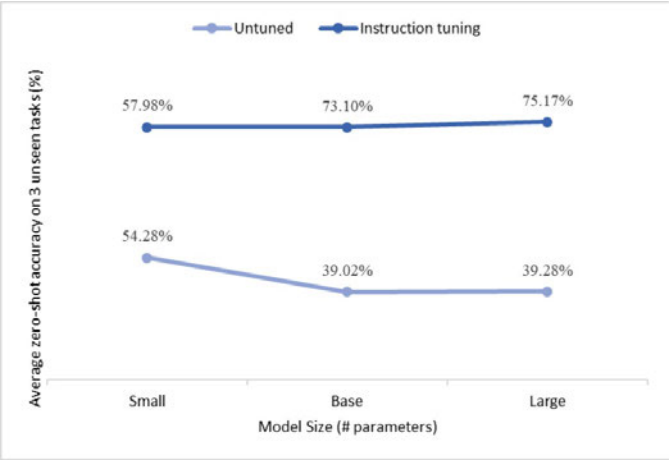
Result 3 & 4

– Relation between instruction tuning and model size :

– Effect of training corpus size :

Relation between instruction tuning and model size :

- Vanilla model having inconsistent performance.
- Fine-tuned model's performance is consistent, and effectiveness grows with model sizes.
- Larger fine-tuned models perform better than small fine-tuned models.



Instruction tuning efficacy

Effect of training corpus size :

- The optimal fine-tuning corpus size is 6,000 data points.

Sample size	DistilBert	MiniLM	FLAN-T5	Average	Best Score
2K	54.58%	63.03%	58.51%	58.71%	63.03%
4K	56.54%	69.08%	60.27%	61.93%	69.08%
6K	**66.64%**	**69.67%**	**61.16%**	**65.82%**	**69.67%**
8K	57.83%	66.82%	60.89%	61.85%	66.82%
10K	55.93%	54.58%	61.04%	57.19%	61.04%
12K	58.26%	59.72%	60.93%	59.64%	60.93%

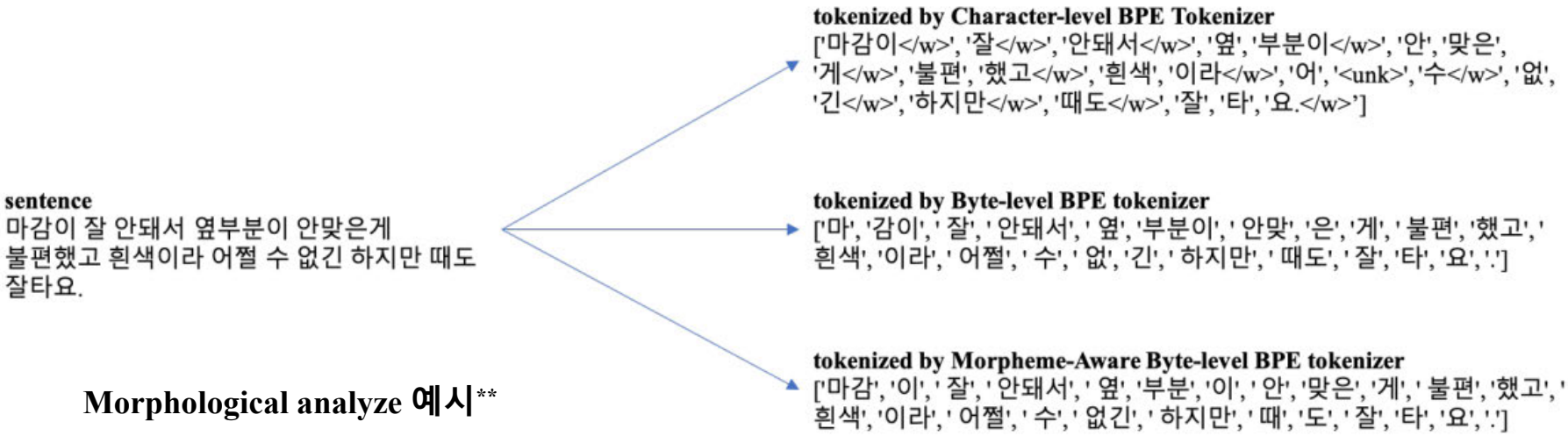
Zero-shot accuracy

05-01. Tokenization with Linguistic awareness [2/4]

Goal of Tokenization :
Development of Korean efficient tokenizer.

Morpheme(형태소)-aware sub-word Byte Pair Encoding (BPE)*

- BPE + linguistic awareness of Korean language
  - Apply morphological analyzer first and then BPE on it
    - E.g. ‘ㅎ’, ‘하’, ‘한’ are all considered different Unicode bytes
      => conventional BPE for English LLM utilization cannot handle Korean language proficiently.



[*] Kyubyong Park, Joohong Lee, Seongbo Jang, Dawoon Jung. 2020. An Empirical Study of Tokenization Strategies for Various Korean NLP Tasks *arXiv:2010.02534*

[**] Kim et al., 2021. What Changes Can Large-scale Language Models Bring? Intensive Study on HyperCLOVA: Billions-scale *Korean Generative Pretrained Transformers*

05-01. Tokenization with Linguistic awareness [2/4]

Korean Tokenization :

Efficient morphological analyzer is required

Morpheme(형태소)-aware sub-word Byte Pair Encoding (BPE)*

- Korean LLMs - Polyglot-Ko, KULLM, HyperCLOVA
- Morphological analyzer for whole training dataset
  - Train inefficiency => Tokenization strategy needs further improvement.

EleutherAI /polyglot-ko

KULLM

HyperCLOVA

Tokenization	Vocab Size	KorQuAD Dev (EM/F1)	KorNLI Dev	KorNLI Test	KorSTS Dev	KorSTS Test	NSMC Dev	NSMC Test	PAWS-X Dev	PAWS-X Test
CV	166	59.66 / 73.91	70.60	71.20	77.22	71.47	87.97	87.89	58.00	55.20
Syllable	2K	69.10 / 83.29	73.98	73.47	82.70	75.86	88.94	89.07	68.65	67.20
Morpheme	32K	68.05 / 83.82	74.86	74.37	82.37	76.83	87.87	88.04	69.30	67.20
	64K	70.68 / 85.25	75.06	75.69	83.21	77.38	88.72	88.88	73.40	68.65
Subword	4K	71.48 / 83.11	74.38	74.03	83.37	76.80	89.08	89.30	72.00	69.60
	8K	72.91 / 85.11	74.18	74.65	83.23	76.42	89.08	89.19	73.45	69.00
	16K	73.42 / 85.75	74.46	75.15	83.30	76.41	88.89	88.88	73.40	70.70
	32K	**74.04** / 86.30	74.74	74.29	83.02	77.01	89.39	89.38	74.05	70.95
	64K	**74.04 / 86.66**	73.73	74.55	83.52	77.47	88.80	89.19	75.85	72.10
Morpheme-aware Subword	4K	67.53 / 81.93	73.53	73.45	83.34	76.03	88.93	89.32	69.75	67.45
	8K	70.90 / 84.57	74.14	73.95	83.71	76.07	89.37	89.29	73.40	71.30
	16K	69.47 / 83.36	75.02	74.99	83.22	76.59	89.33	89.41	75.05	71.70
	32K	72.65 / 86.35	74.10	75.13	83.65	**78.11**	89.53	89.65	74.60	71.60
	64K	69.48 / 83.73	**76.39**	**76.61**	**84.29**	76.78	**89.82**	**89.66**	**76.15**	**74.00**
Word	64K	~~1.54 / 8.86~~	64.06	65.83	69.00	60.41	70.10	70.58	58.25	55.30

[*] Kyubyong Park, Joohong Lee, Seongbo Jang, Dawoon Jung. 2020. An Empirical Study of Tokenization Strategies for Various Korean NLP Tasks *arXiv:2010.02534*

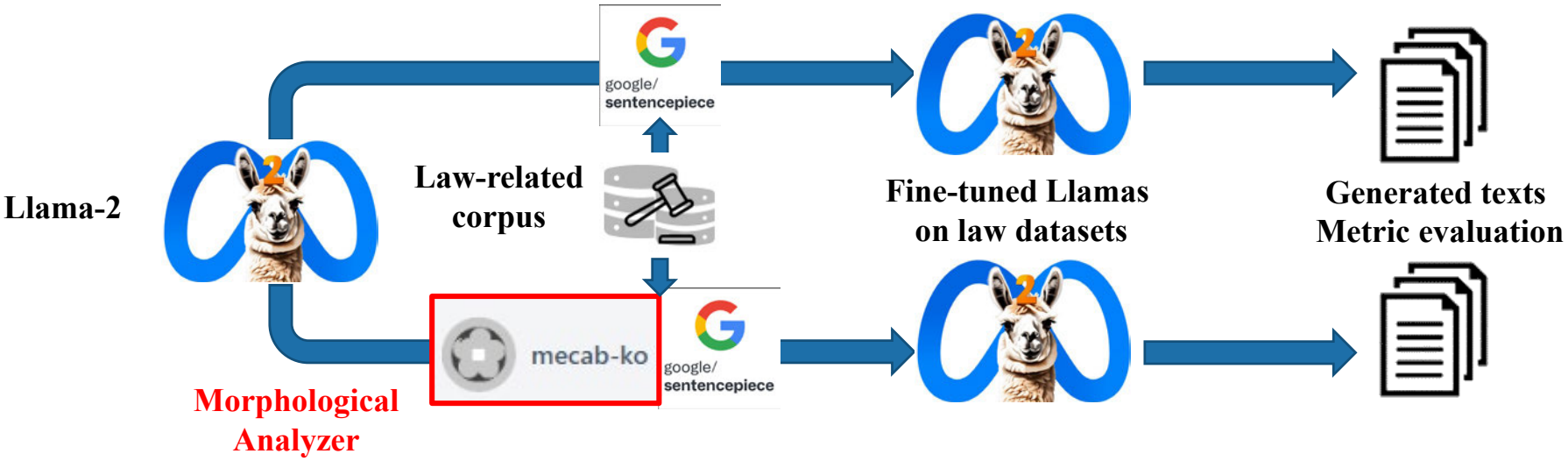
05-01. Tokenization with Linguistic awareness [2/4]

Current work :

Effect evaluation of the Morphological analyzer

Current work – Morphological analyzer evaluation

- Exchange Tokenizer from LLMs (e.g. Llama2)
  - Sentencepiece + Fine-tuned LLMs
  - Korean Morphological analyzer (e.g. Mecab-ko) + Sentencepiece + Fine-tuned LLMs

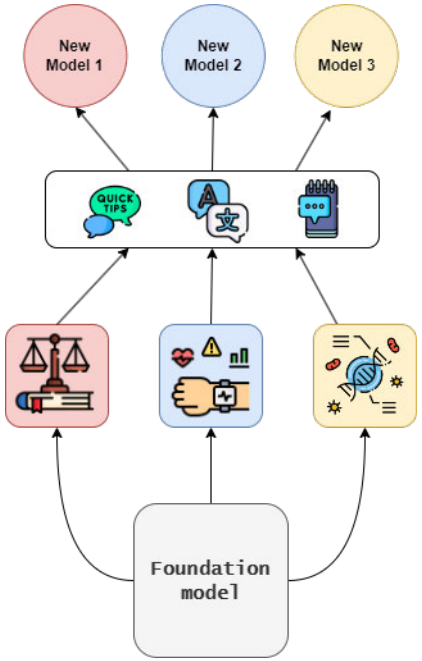


05-01. Continual Domain Adaptive Pre-training [3/4]

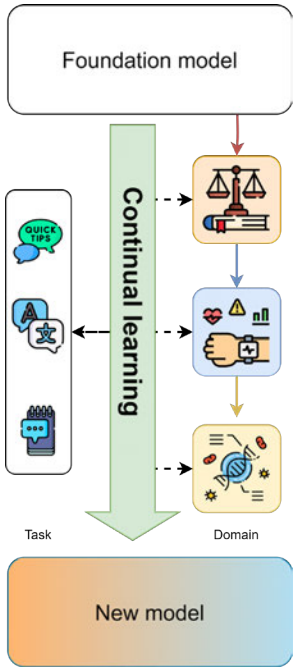
Goal of Continual Domain Adaptive Pre-training (Continual DAP) :
다양한 도메인에서의 모델 학습을 위한, **Catastrophic forgetting 최소화 메커니즘** 개발

Continual DAP

- Knowledge transfer for domain specific fine-tuned task across all domains.



Domain specific learning



Continual learning

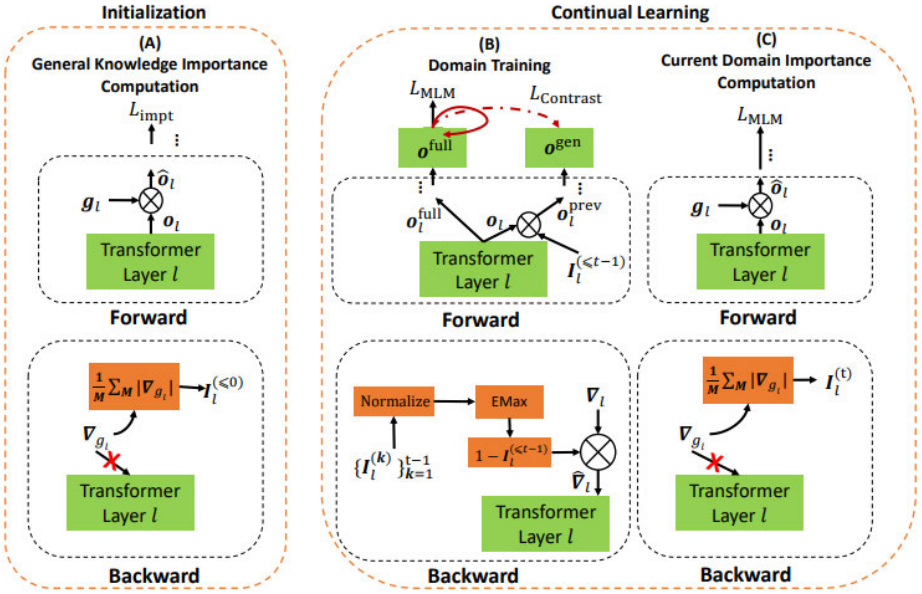
[*] K. Arumae, Q. Sun, and P. Bhatia, "An empirical investigation towards efficient multi-domain language model pre-training," in Proc. EMNLP, vol. 1, 2020, pp. 4854–4864.

05-01. Continual Domain Adaptive Pre-training [3/4]

Related works :
Soft-masking 시 Domain importance 값을 고려

Continual Pre-training *

- **Soft-masking & Contrastive loss**
  - Soft masking 기반 Domain transfer learning
- **Initialization:**
  - Previous domain 으로부터의 general knowledge importance 계산
  - Current domain 에서의 Soft-masking 에 활용
  - Previous model 의 relevant knowledge 유지에 유리



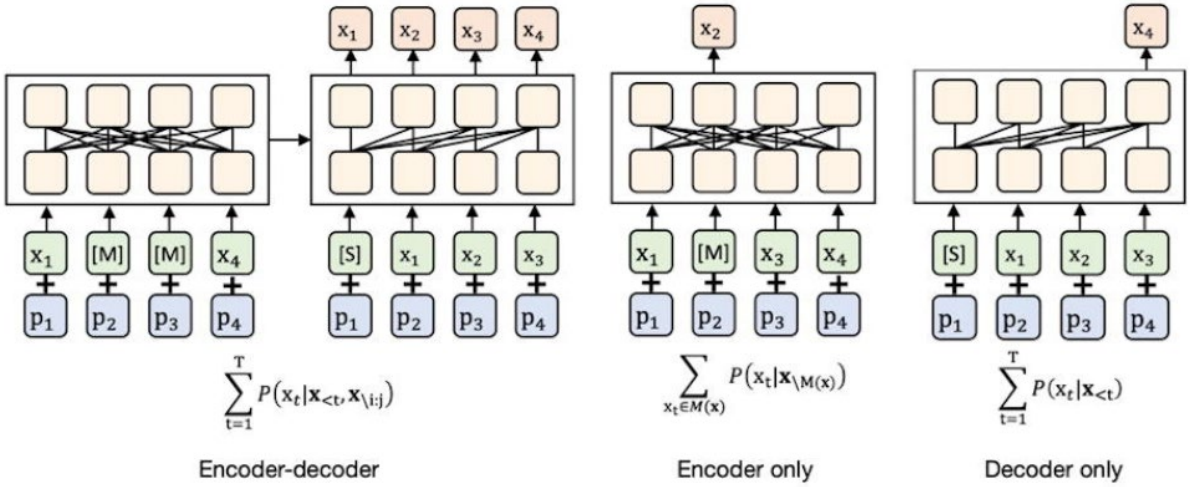
[*] Ke, Zixuan, et al. "Continual Pre-training of Language Models." *The 11th ICLR*, 2022.

05-01. Continual Domain Adaptive Pre-training [3/4]

Challenges :
Soft-masking 방식은 Decoder only model 구조에 효율적이지 못함.

Challenges - Decoder-only LLM model

- Sentence generation 에 특화된 Decoder-only model*
  - Soft-masking 방식의 경우 Decoder-only model에 효율적이지 못함.
    - Decoder 구조상 이전의 추론 단어에 기반하여 다음 단어를 추론함.
  - 모델 구조를 고려한 효율적인 접근이 필요함.



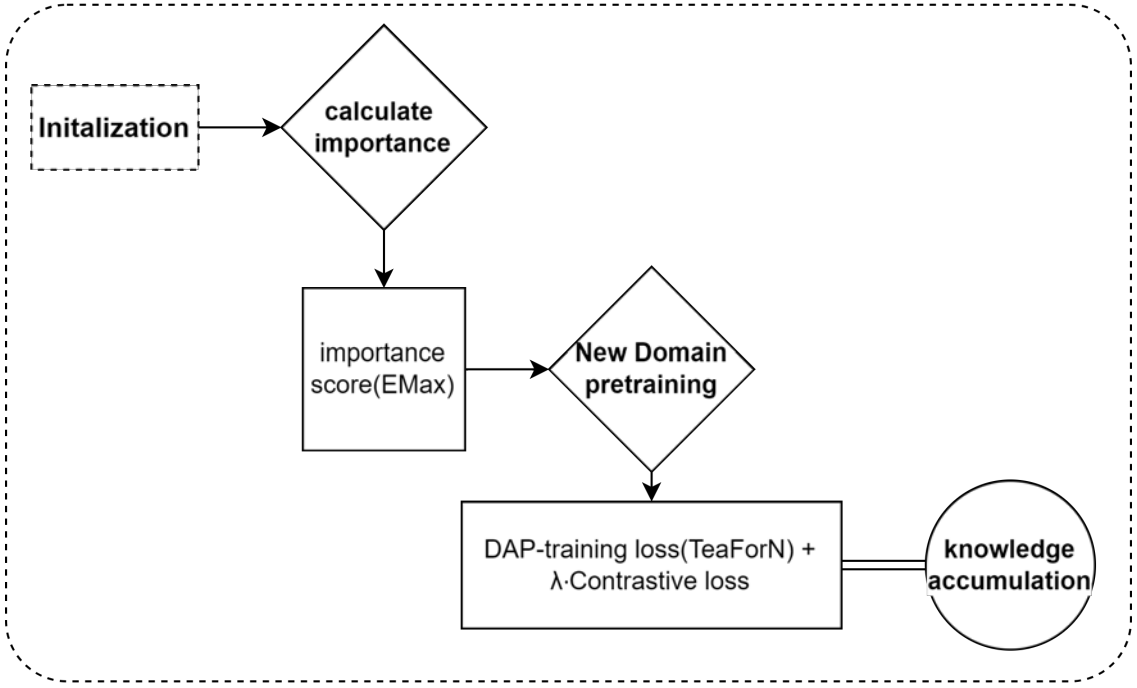
[*] H. Wang, J. Li, H. Wu, E. Hovy, and Y. Sun, ‘‘Pre-trained language models and their applications,’’ Engineering, vol. 25, pp. 51–65, Jun. 2023.

05-01. Continual Domain Adaptive Pre-training [3/4]

Current work - Proposed method:
Decoder only model 을 고려한 **TeaForN**방법 *,** 기반 importance score 반영

Proposed method – Contrastive learning + computing every unit's importance + TeaForN*,**

- Decoder only model 구조 특성을 고려한 Teacher forcing 방법; **TeaforN(Teacher-Forcing with N-grams)** *,**
- Importance score 에 기반한 TeaforN 방법 적용



[*] Ke, Zixuan, et al. "Continual Pre-training of Language Models." *The Eleventh International Conference on Learning Representations*. 2022.
[**] Goodman, Sebastian, Nan Ding, and Radu Soricut. "Teaforn: Teacher-forcing with n-grams." *arXiv preprint arXiv:2010.03494* (2020).

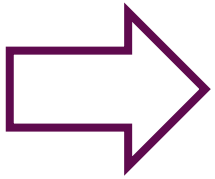
05-01. Preference Learning (DPO) [4/4]

Related works[*,**] **:**
Reinforcement Learning from Human Feedback (RLHF) 기반 fine-tuning 방법에서의 문제점

Reinforcement Learning from Human Feedback (RLHF)[*,**]

RLHF pipeline

1. A supervised fine-tuning (SFT)
2. The process of annotating data with preference labels
3. Training a reward model on the preference data
4. And the RL optimization step



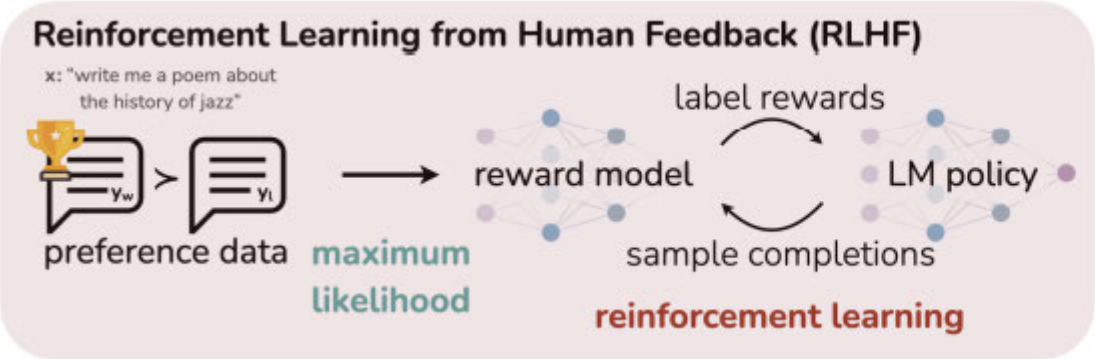
- Complex
- High GPU resources
- Unstable procedure
- Mode Collapse

[*] Santacroce, Michael, et al. "Efficient RLHF: Reducing the Memory Usage of PPO." *arXiv:* 2309.00754 (2023)
[**] Casper, Stephen, et al. "Open Problems and Fundamental Limitations of Reinforcement Learning from Human Feedback." *arXiv*:2307.15217 (2023)

05-01. Preference learning (DPO) [4/4]

Related works* :
RLHF에서의 복잡성을 피하면서도 사용자 선호도를 직접 학습하는 방법
우수한 구현 편의성, 학습 효율성

Direct Preference Optimization (DPO)*





RLHF
Training a reward model on the preference data And the RL optimization step

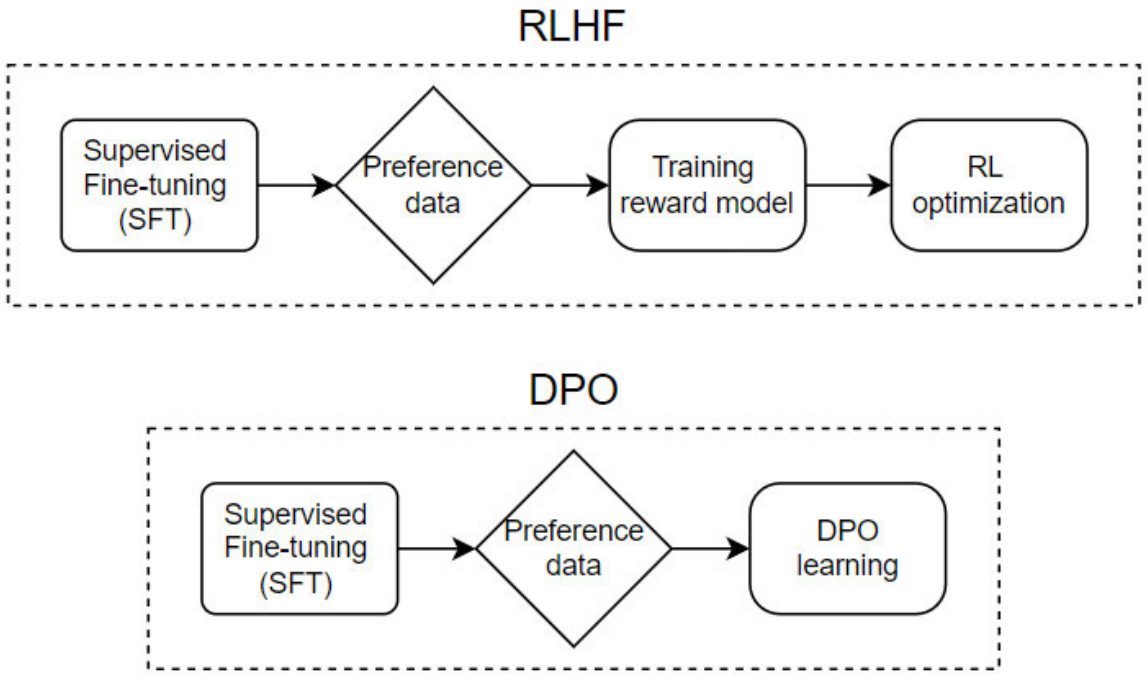
DPO
DPO training directly optimizes the DPO object on preference annotated data

[*] Rafailov, R., Sharma, A., Mitchell, E., Ermon, S., Manning, C.D., and Finn, C. 'Direct preference optimization: Your language model is secretly a reward model' arXiv: 2305.18290, 2023.

05-01. Preference learning (DPO) [4/4]

Goal of Preference Learning:
Direct Preference Optimization (DPO) 기반 My Legal AI model fine-tuning 진행

동기 및 배경



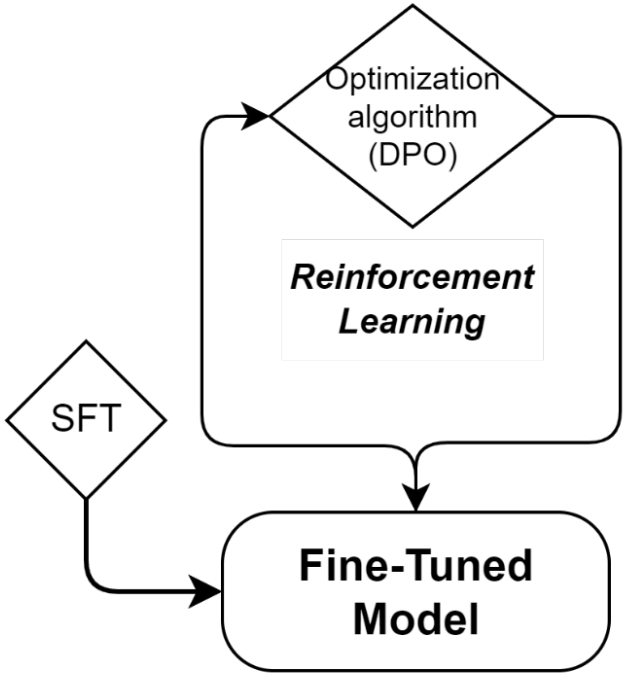
- DPO는 RLHF의 복잡성을 피하면서도 사용자의 선호도를 직접적으로 최적화
- 이를 통해 높은 품질의 법률 AI 모델을 보다 간편하게 Fine-tuning

05-01. Preference learning (DPO) [4/4]

Goal of Preference Learning:

Direct Preference Optimization (DPO) 기반 My Legal AI model fine-tuning 진행

연구 목표



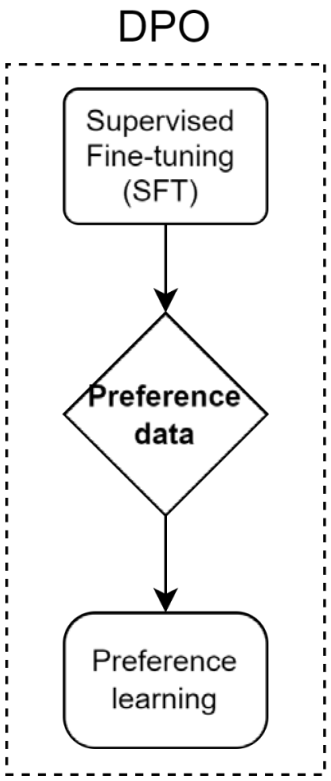
1. 정확성 및 전문성 향상
   - 법적 용어 및 문맥 정확이 이해하고 처리, 전문성이 요구되는 법률 서비스에서 높은 품질의 텍스트 생성

2. 특화된 대응
   - 특정한 법률적 요구사항 및 선호에 따라 대응,특화된 대화 능력 향상

05-01. Preference learning (DPO) [4/4]

Current work - Proposed method:
Preference dataset 구성 및 DPO 기반 fine-tuning 수행

연구 방법: Creating Preference Data




```
DPO_dataset_dict = {
    "prompt": [
        "상속 관련 법률 상담이 필요합니다.",
        "회사 설립 절차에 대해 알려주세요.",
        "임대차 계약서 작성법에 대해 설명해주세요.",
        # 추가적인 prompt들...
    ],
    "chosen": [
        "상속 절차에 대해 자세히 알려드리겠습니다.",
        "회사 설립을 위한 절차와 주의사항에 대해 안내해드릴게요.",
        "임대차 계약서 작성법과 주요 포인트를 안내합니다.",
        # 추가적인 선택된 답변들...
    ],
    "rejected": [
        "미안한데 다른 주제로 질문해주시겠어요?",
        "회사 설립 외에 다른 주제로 도움을 드릴 수 있을까요?",
        "임대차 계약서 작성 이외에 궁금한 내용이 있을까요?",
        # 추가적인 비선호 답변들...
    ]
}
```

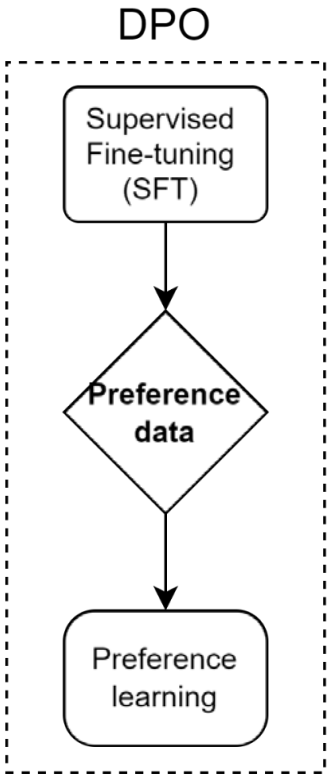
Preference data sample

1. **Prompt**
   - 사용자의 질문
2. **Chosen**
   - 모델이 생성한 선호하는 답변
3. **Rejected**
   - 모델이 생성한 비선호하는 답변

05-01. Preference learning (DPO) [4/4]

Current work - Proposed method:
Preference dataset 구성 및 DPO 기반 fine-tuning 수행

연구 방법: DPO learning (preference learning)



1. Preference learning pipeline 설정
: 모델 답변 수집 → 사용자 평가 → 점수 기록 및 학습 데이터 구성

2. DPO 알고리즘 적용 및 Fine-tuning
: DPO update를 수행하여 모델을 fine-tuning, 모델이 선호하는 답변을 더 잘 생성하도록 업데이트됨

3. 성능 평가

4. 반복 및 개선

5. 사용자 평가 및 피드백 수집

마음 AI 기술 세미나

AI Service Development team
- RAG Service development -

발표자 | 안태호 팀장

CONTENTS

1. LiberVance 연구 개발 목표
2. 법률 AI란?
3. 연구 개발 추진 체계
4. 데이터셋 (Corpus) 확보 방안
5. My AI Development
    1. Pre-training / Fine-tuning
    2. RAG development
6. Cowork Process
7. Q&A

05-02. RAG Development [1/2]

LLM Application 서비스를 위한 웹앱 개발과 배포

업무 요약:

① 실 사용자들이 마주할 서비스의 웹페이지 구현

② 언어모델이 추론 결과를 내놓는 모델 서빙 서버 구현

③ 적절한 응답을 내놓기 위한 로직 구현 (w/ Prompt Engineering, RAG)

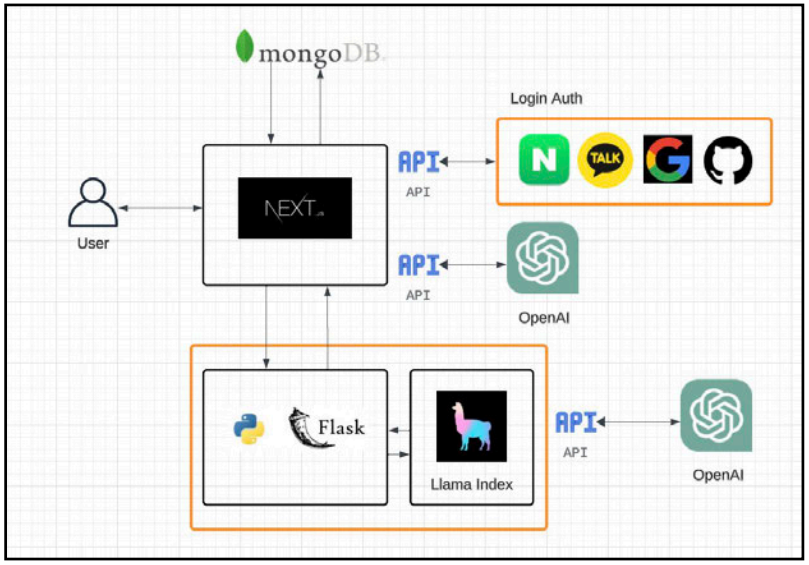
④ 개발된 서비스의 버전 관리와 배포 (w/ Git, AWS)

05-02. RAG Development [1/2]

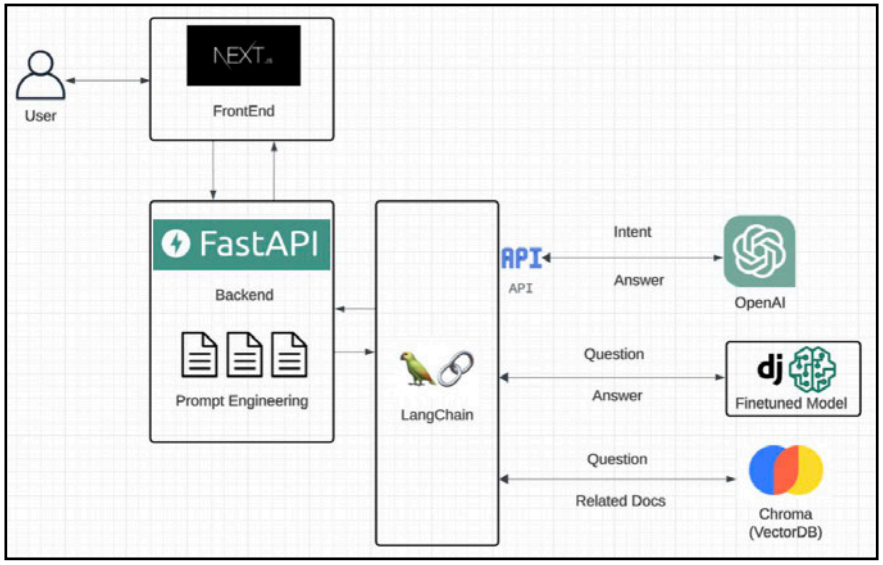
LLM Application 서비스를 위한 웹앱 개발과 배포

개요 그림:

① 업로드한 PDF 기반 질의응답



② 법률 지식 기반 질의응답



05-02. RAG Development [1/2]

LLM Application 서비스를 위한 웹앱 개발과 배포

연구 내용 및 결과:

LLM을 활용한 서비스를 구현하기 전,
당시의 LLM 현황과 LLM Application의 case study 진행

① Llama, Alpaca, KoAlpaca 그리고 한국어 사전학습 모델
② 서비스 형태별 LLM Application 정리
③ 법률 데이터와 Finetuning API 활용해 생성한 모델의 응답 성능 비교

05-02. RAG Development [1/2]

LLM Application 서비스를 위한 웹앱 개발과 배포

연구 내용 및 결과 : ① **Llama, Alpaca, KoAlpaca 그리고 한국어 사전학습 모델**

Llama 공개 후,

3. 오픈소스 LLM 활용해 finetuning (ex. LlaMa)

장점 : chatgpt급은 아니어도 어느정도 나오는 성능

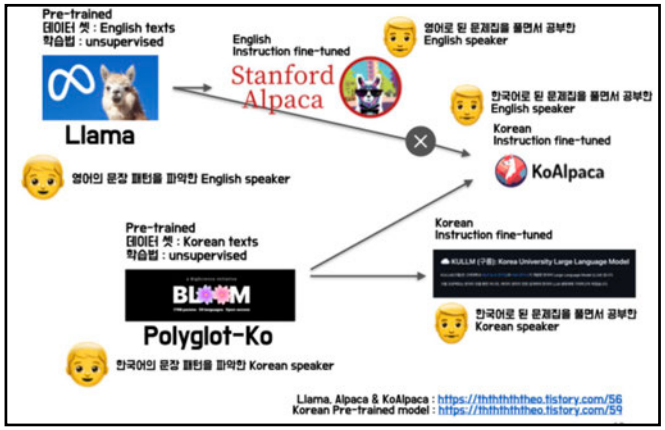
서비스 업체가 원하는 도메인 지식 추가 학습 가능

단점 : 100B 이상의 LLM 보다 떨어지는 성능

도메인 관련 데이터셋 구축 필요

AI 엔지니어 필요

서비스 운영 시, 모델 서빙에도 GPU 필수



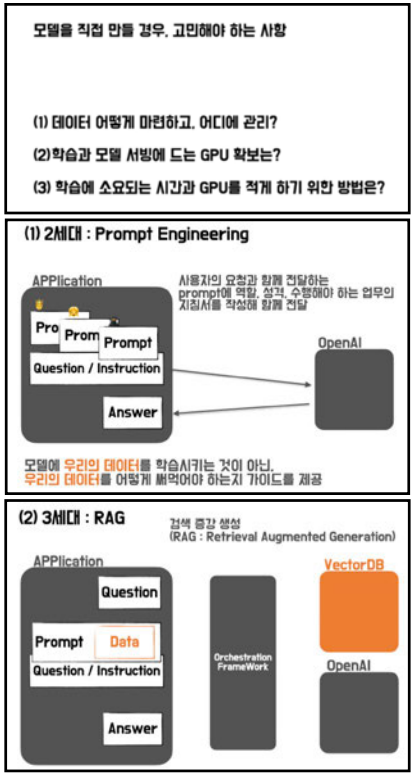
Llama, Alpaca 그리고 KoAlpaca의 생성 과정 파악
➞ Finetuning을 통한 LLM 이용 방법 확인
➞ 한국어로 된 사전학습 모델의 필요성 확인

한국어로 사전학습된 대규모 모델들 파악
➞ Polyglot-Ko와 KULLM

출처 : Llama, Alpaca, KoAlpaca 보고용 리포트, LLM Application 보고용 리포트

LLM Application 서비스를 위한 웹앱 개발과 배포

연구 내용 및 결과 : ② 서비스 형태별 **LLM Application** 정리



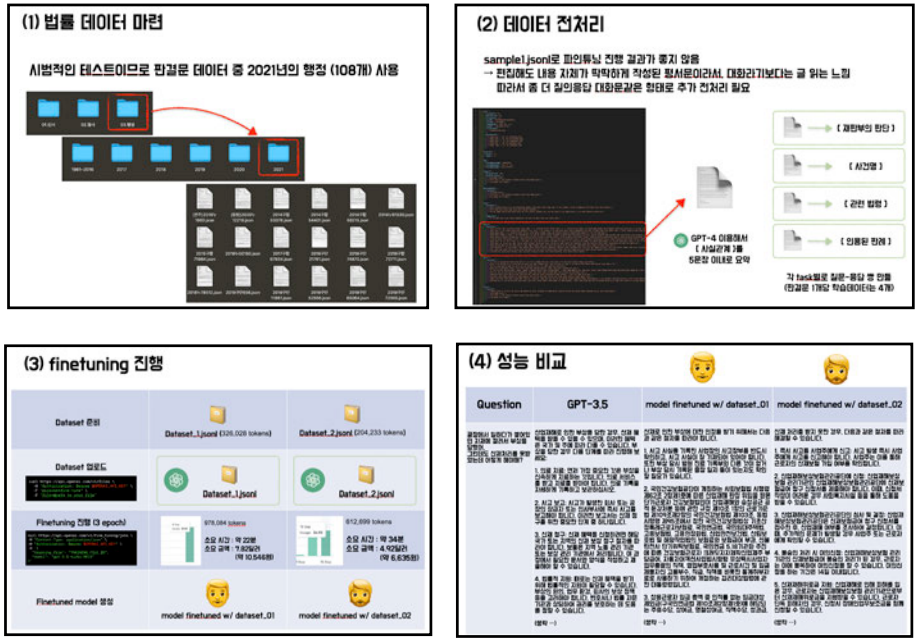
LLM을 활용하는 서비스에는 다양한 형태로 존재

방향 1 : 성능이 좋은 LLM을 만들기 위한 노력
방향 2 : API 활용 + 프롬프트 엔지니어링, RAG

출처 : LLM Application 보고용 PPT

LLM Application 서비스를 위한 웹앱 개발과 배포

연구 내용 및 결과: ③ 법률 데이터와 Finetuning API 활용해 생성한 모델의 성능 비교



OpenAI Finetuning API 출시
이를 활용해 모델 생성 후, 성능 확인

gpt-3.5보다는 법률 전문가 페르소나가 부여된 느낌
프롬프트 엔지니어링에 대한 수고나 토큰 절약

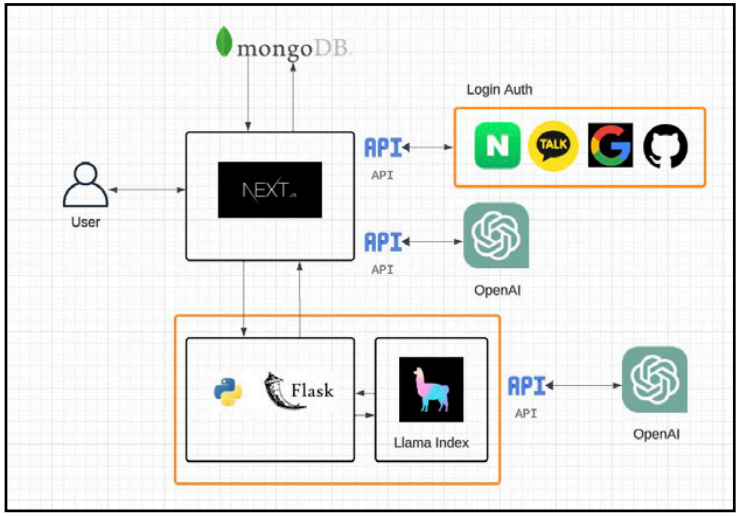
하지만 특정 내용에 대해 묻거나 찾아주는 질문의 경우,
RAG를 이용해 얻은 응답이 더 뛰어남

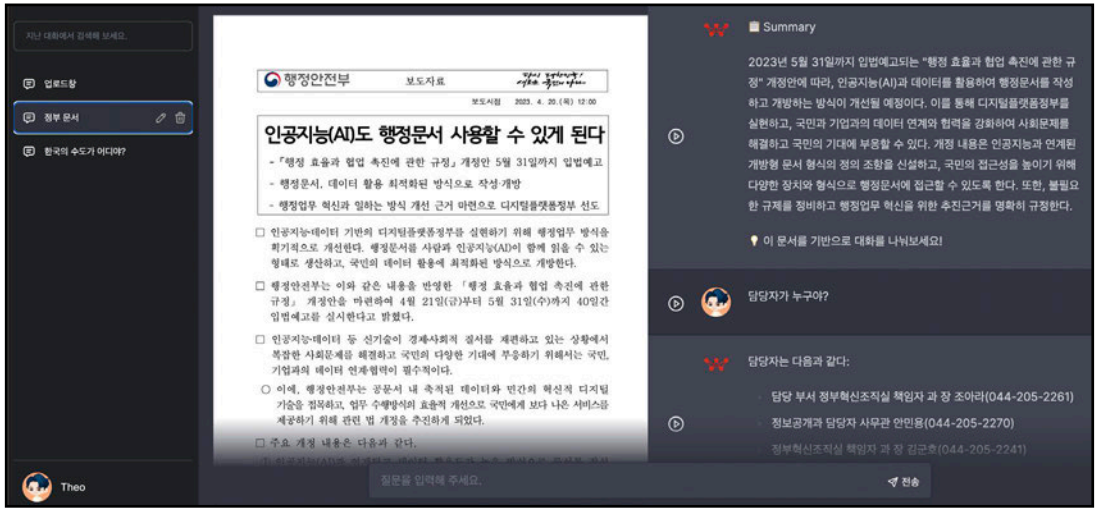
출처 : 법률 도메인 응답 생성 성능 비교 보고용 PPT

05-02. RAG Development [1/2]

LLM Application 서비스를 위한 웹앱 개발과 배포

주요 개발 Task 및 개발 현황: ① 업로드한 PDF 기반 질의응답



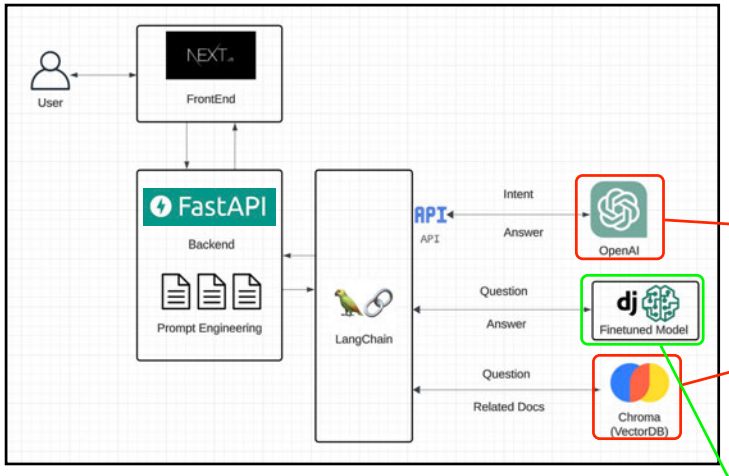


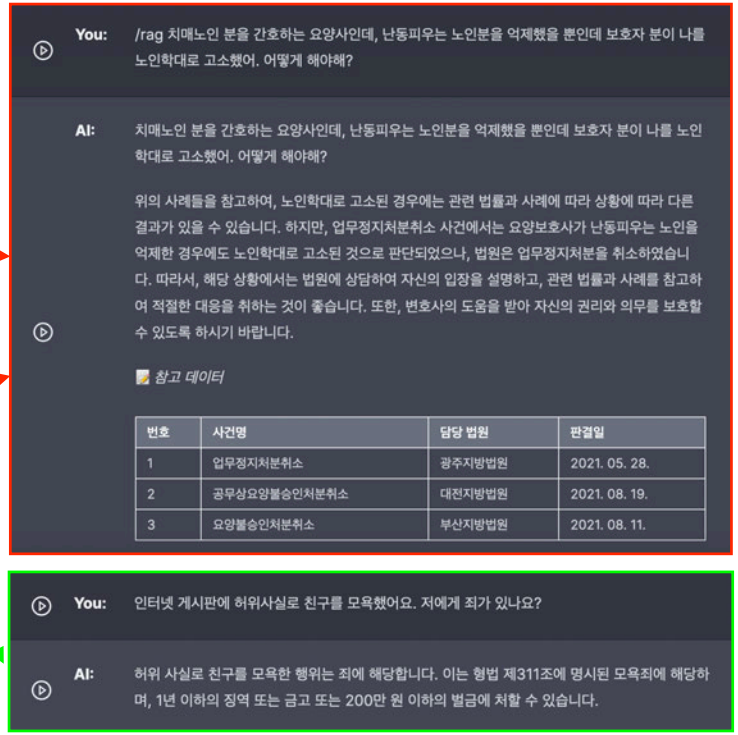
개발 완료 후, AWS EC2 인스턴스 서버에 배포

https://ai.worldland.foundation/ (해당 도메인으로 접속 가능)

LLM Application 서비스를 위한 웹앱 개발과 배포

주요 개발 Task 및 개발 현황: ② 법률 지식 기반 질의응답



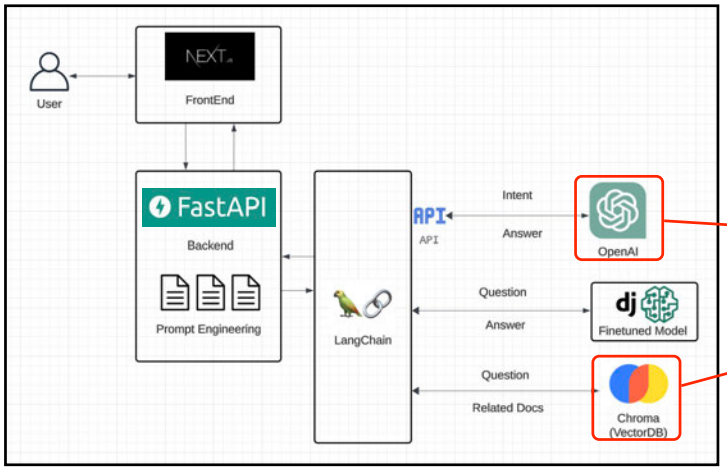


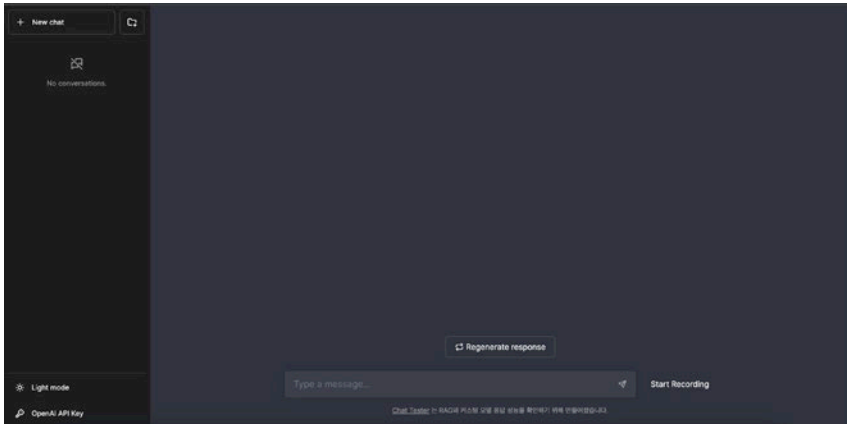
커스텀 모델의 추론을 위한 서버 구축과 추론 테스트 완료
파인튜닝 완료된 모델 완성 후, 서버 새로 구축 예정

05-02. RAG Development [1/2]

LLM Application 서비스를 위한 웹앱 개발과 배포

주요 개발 Task 및 개발 현황: ② 법률 지식 기반 질의응답

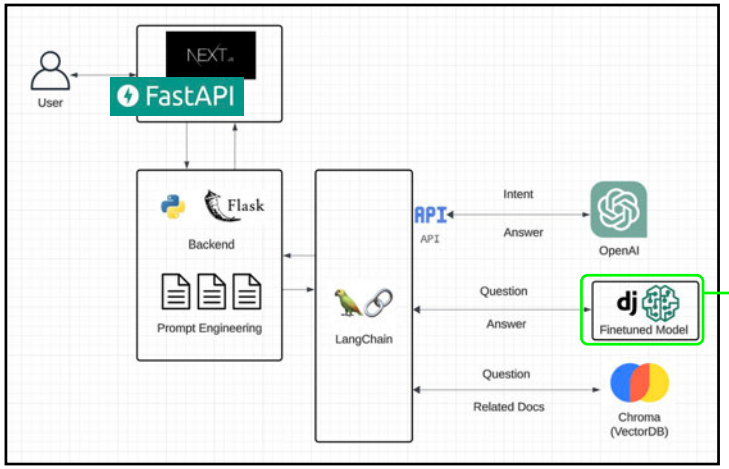


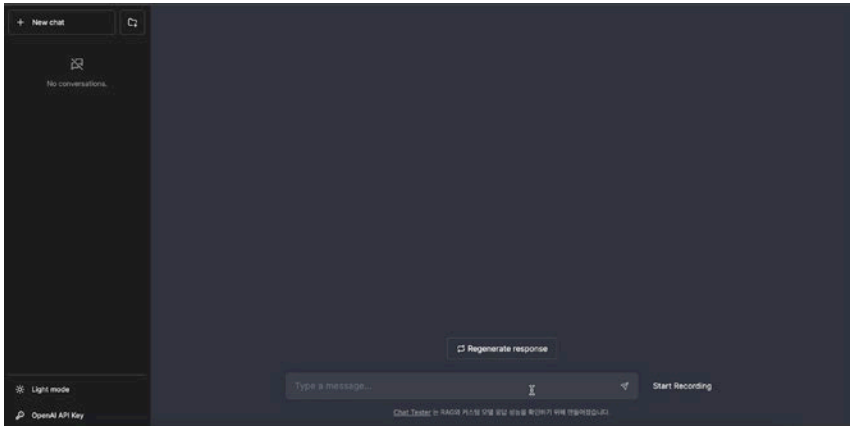


05-02. RAG Development [1/2]

LLM Application 서비스를 위한 웹앱 개발과 배포

주요 개발 Task 및 개발 현황: ② 법률 지식 기반 질의응답

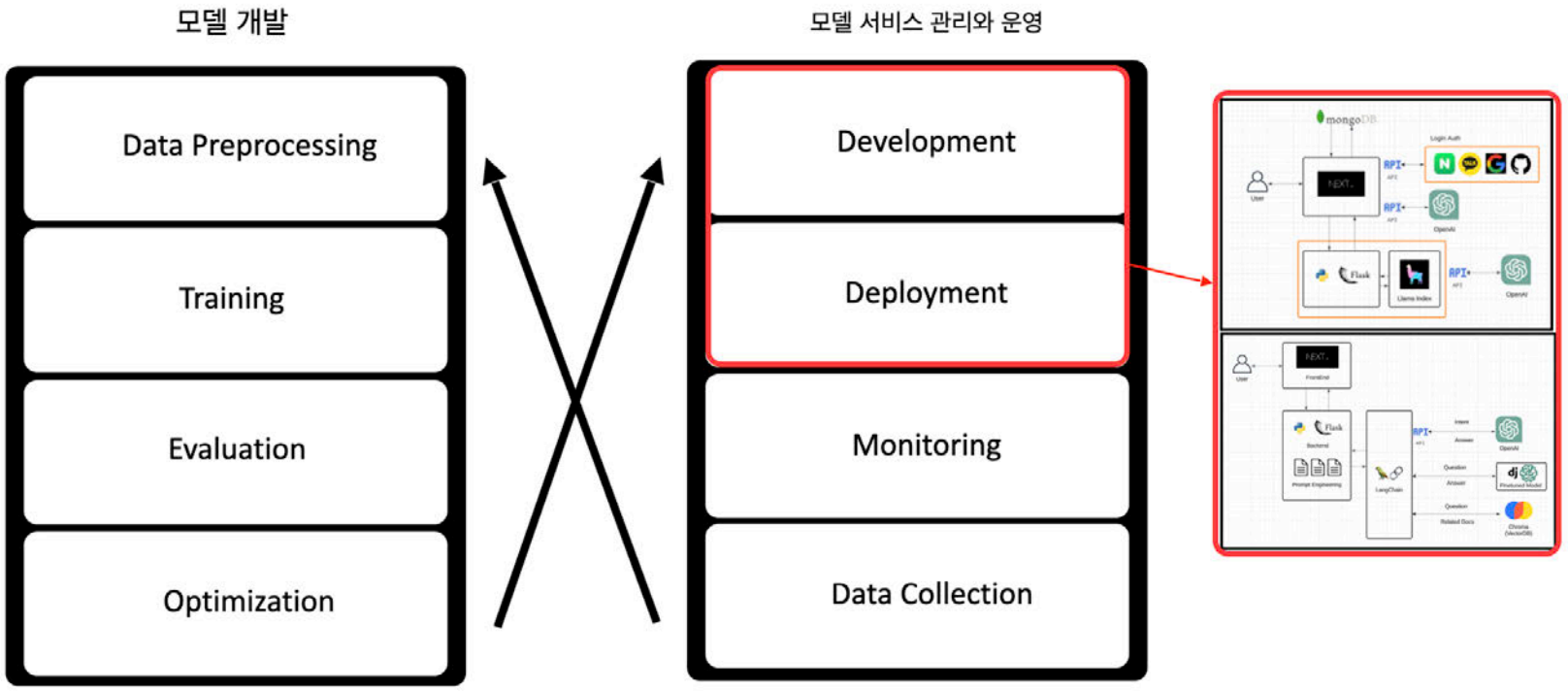




05-02. RAG Development [1/2]

LLM Application 서비스를 위한 웹앱 개발과 배포

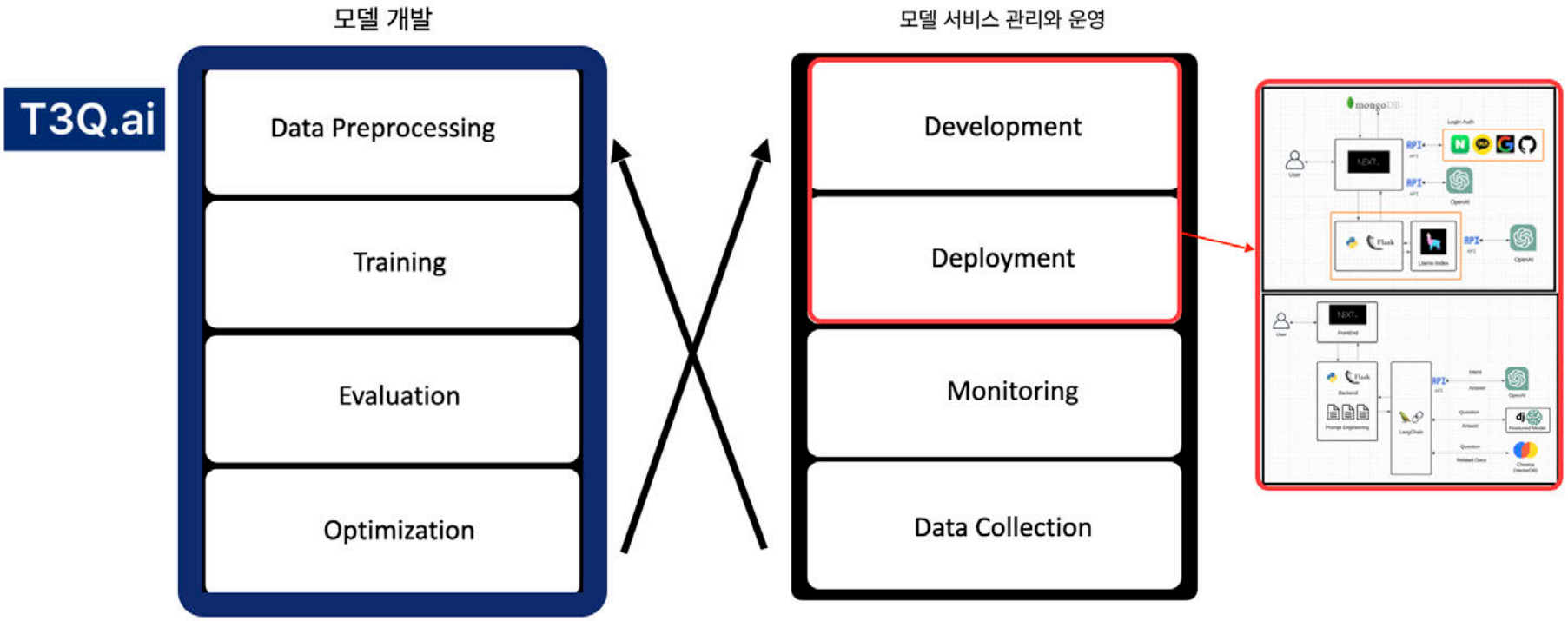
기대 효과 및 향후 추진 방향: **MLOps**를 *위한 파이프라인 구축과 AICA 공용풀*



AI 서비스의 운영에는 해당 모델의 업데이트와 버전 관리가 필수적
이러한 지속적인 통합과 배포가 이루어지도록 자체적 파이프라인 구축

LLM Application 서비스를 위한 웹앱 개발과 배포

기대 효과 및 향후 추진 방향: **MLOps**를 위한 파이프라인 구축과 ***AICA*** 공용풀



모델 개발 과정을 AICA의 공용풀에서 진행 가능할 것으로 사료됨

05-02. (Korean) Fine-Tuned Model Service [2/2]

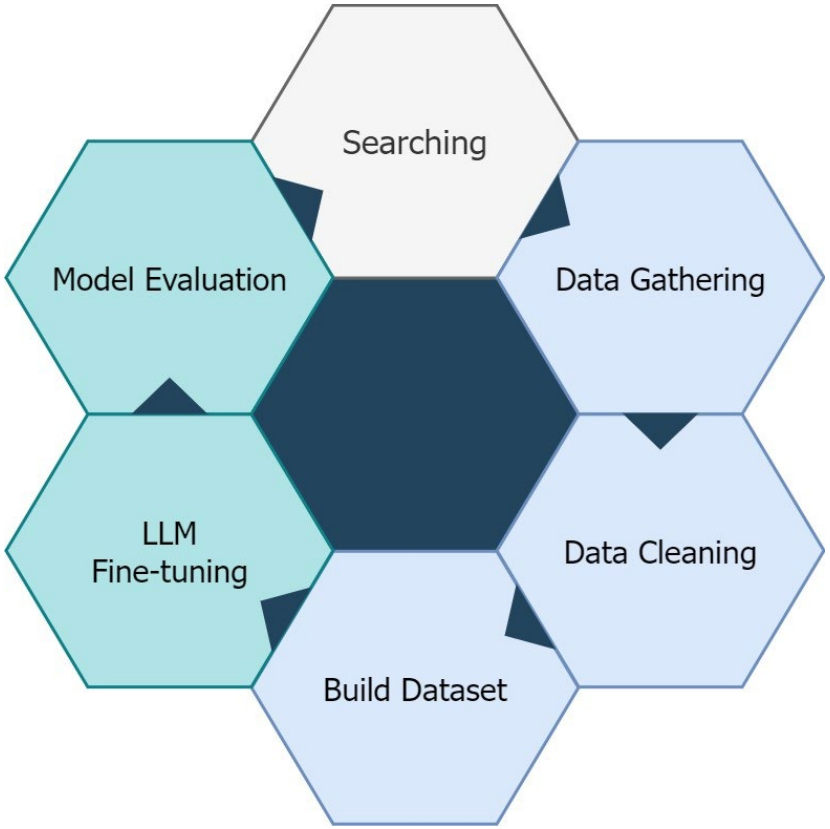
법률 도메인에 특화된 AI 모델을 학습하여, 법률 조언을 제공하는 챗봇 개발

업무 요약:

- 법률 QA dataset 구축 (완료)
  - 크롤링
  - 데이터 합성
  - 데이터 전처리
- 한국어 LLM Fine-tuning (진행중)
  - LLaMA-2, KULLM, Polyglot-ko
- 모델 평가 (진행 전)
  - Human Evaluation
  - G-EVAL

05-02. (Korean) Fine-Tuned Model Service [2/2]

개요



✔ Human Evaluation
✔ G-EVAL

✔ 찾기쉬운생활법령정보
✔ 대한법률구조공단

✔ LLaMA-2-ko 모델 학습
✔ Kullm-polyglot 모델 학습

✔ 특수문자 제거
✔ 공단 안내문구 제거
✔ 쿠션어 제거

✔ 합성 데이터 생성
✔ 데이터 병합

05-02. (Korean) Fine-Tuned Model Service [2/2]

연구 내용 및 결과

1. 모델 선정 : LLaMA vs. Polyglot-ko

(내용) 한국어로 Pre-train 된 LLM 후보를 선정하고, 법률 QA 샘플 데이터

(결과) LLaMA 계열의 모델들이 법적 내용이 포함된 답변을 더 잘 생성하였음

LLM 후보

- ✓ KULLM-12.8B
- ✓ KULLM-5.8B
- ✓ Polyglot-ko-5.8B

⋮

- ✓ LLaMA-2-ko-7B
- ✓ LLaMA-2-ko-chat-7B

⋮

Hugging Face

Datasets: jihye-moon / LawQA-Ko



LLM 선정

- ✓ **LLaMA-2-ko-chat-7B**

05-02. (Korean) Fine-Tuned Model Service [2/2]

연구 내용 및 결과

2. Prompt Engineering

(내용) 모델이 수행해야 하는 작업에 대한 설명과 질문을 Prompt에 적용

(결과) 백본 모델의 프롬프트와 유사하면서, “법률”, “법적”이라는 단어가 포함된 프롬프트가 학습이 가장 잘 되었음

✓ prompt_with_precedent template

“아래는 상황을 설명하는 사용자의 질문과 이와 관련한 **법률** 문서를 제공하는 입력이 짝을 이루는 예제입니다. 사용자의 질문에 대해 적절한 **법적** 답변이 포함된 응답을 작성하세요.

질문:
{question}

법률 문서:
{precedent}

응답:
”

✓ prompt_without_precedent template

“아래는 상황을 설명하는 사용자의 질문입니다. 사용자의 질문에 대해 적절한 **법적** 답변이 포함된 응답을 작성하세요.

질문:
{question}

응답:
”

05-02. (Korean) Fine-Tuned Model Service [2/2]

Build to Question – Answer Dataset of Legal domain

1. Data Gathering

- 대한법률구조공단 법률상담사례 질문-답변 10,037건
- 대한법률구조공단 사이버상담 질문-답변 2,579건
- 찾기쉬운생활법령정보 백문백답 질문-답변 2,195건

2. Synthetic data using GPT-4

- 데이터셋을 보완하기 위해 OpenAI API를 사용하여 Synthetic Data 생성
- GPT-4으로 생성한 법률 질문-답변 960건 (비용 약 $ 20)

3. 데이터 전처리 및 질문-답변 데이터셋 구축

- 학습에 불필요한 텍스트(공단 안내 문구, 격려의 말 등) 제거
- 특수문자, url, 전화번호 등 제거
- 총 15,771건 질문-답변 데이터셋 구축 완료

https://huggingface.co/datasets/jihye-moon/LawQA-Ko

05-02. (Korean) Fine-Tuned Model Service [2/2]

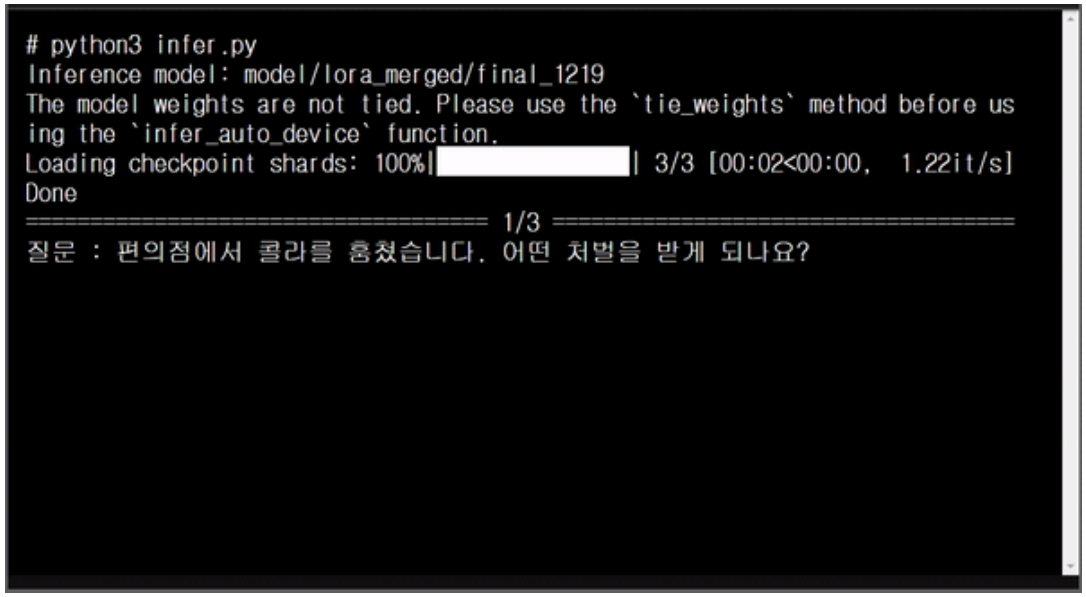
Fine-tune Korean Large Language Model with legal QA dataset and Evaluation

1. 후보 모델 선정 및 실험 (완료)

- 한국어 LLM 후보 조사 (LLaMA-2-ko, KULLM, Polyglot-ko 등)
- Sample Dataset 으로 후보 모델 학습 후, 서비스 모델 선정

2. Fine-tune LLaMA-2-ko-7B series of models (진행중)

- LLaMA-2-ko-7B 계열의 모델을 법률 QA 데이터셋으로 Fine-tuning



05-02. (Korean) Fine-Tuned Model Service [2/2]

기대 효과 및 향후 추진 방향

기대효과 1) 편리하고 쉬운 법률 서비스

법적 조언이 필요한 사람들이 해당 서비스를 활용하여 편리하고 쉽게 법적 조언을 얻을 수 있다.

기대효과 2) 전문가와 일반인의 정보 격차 해소

법에 대한 지식과 정보가 부족한 일반인들이 서비스를 사용하며 법률 전문가와의 정보 격차를 해소할 수 있다.

추진방향) MLOps 관점에서의 서비스 설계

한 번 만들어진 모델로 계속 서비스를 제공하는 것이 아니라, 사용자 대화 데이터를 지속적으로 학습하여 점진적으로 발전하는 인공지능 서비스가 되어야 한다.

마음 AI 기술 세미나

Foundation 모델 개발을 위한 양사 협업 전략

발표자 | 이흥노 대표

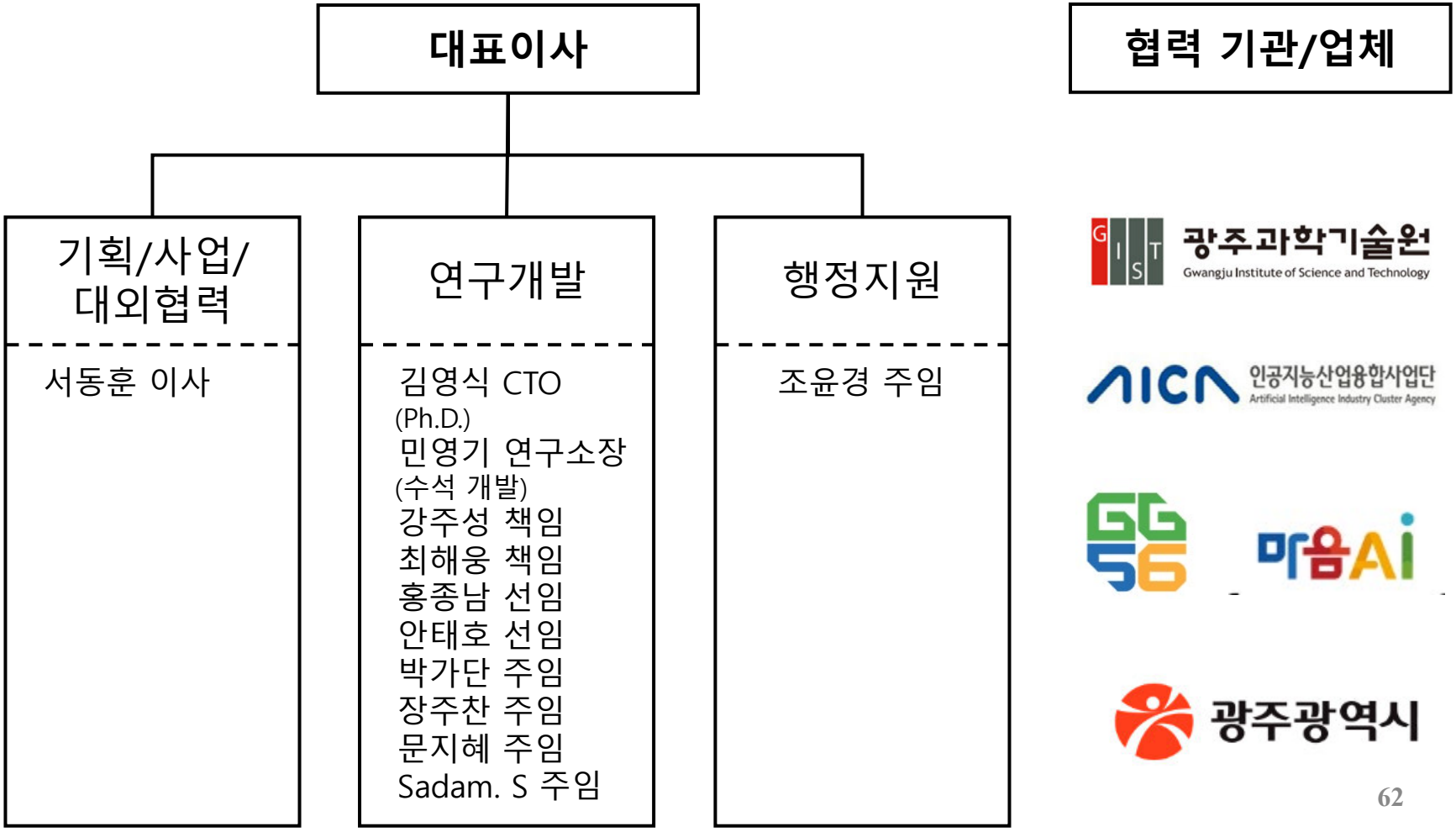
CONTENTS

1. LiberVance 연구 개발 목표
2. 법률 AI란?
3. 연구 개발 추진 체계
4. 데이터셋 (Corpus) 확보 방안
5. My AI Development
    1. Pre-training / Fine-tuning
    2. RAG development
6. Cowork Process
7. Q&A

06. Cowork Process

LiberVance – 마음 AI 협업 전략

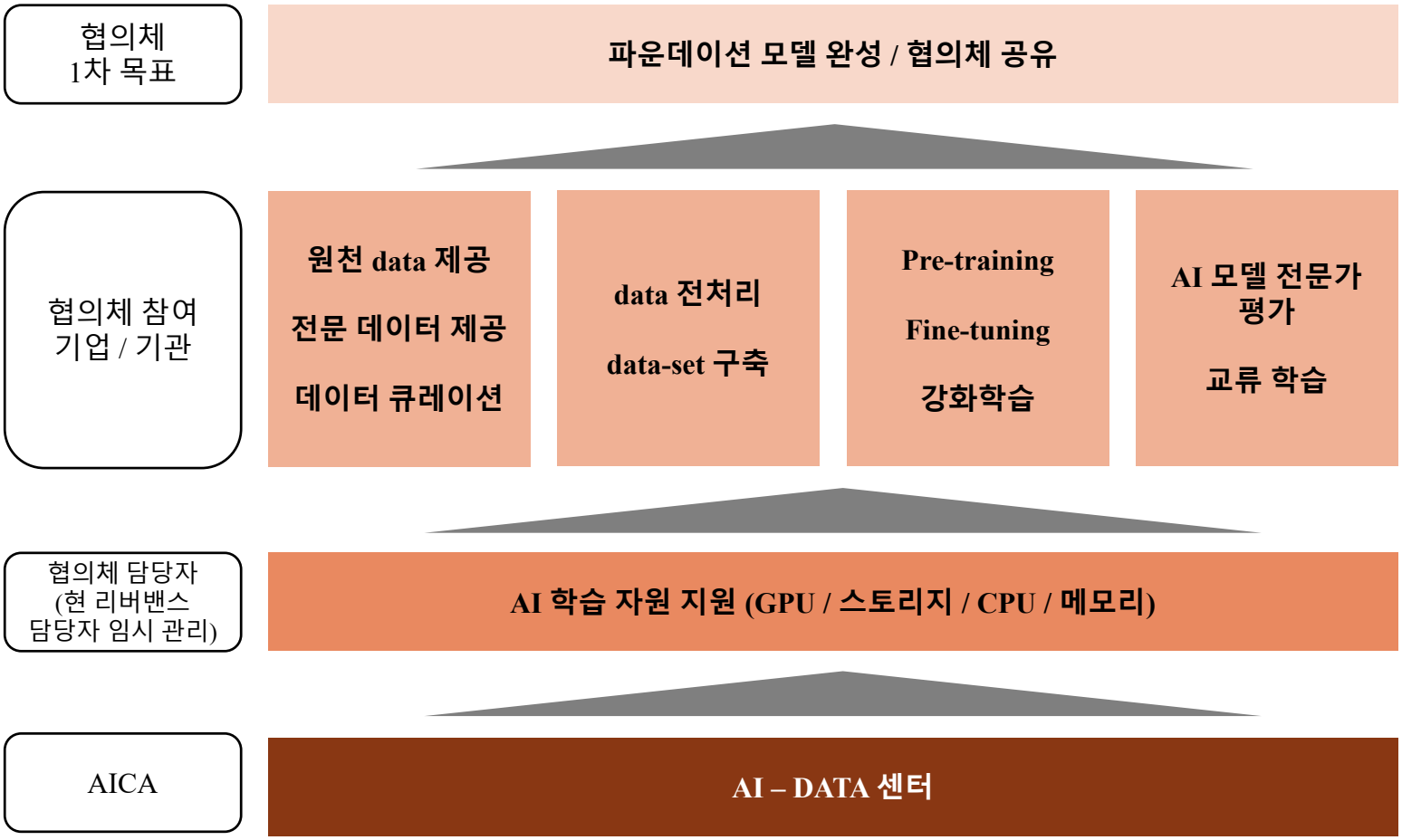
GIST 우수한 인력 기반 AI 기술 중심의 LiberVance,
AI 모델 서비스 개발 노하우 기반의 AI 개발력 중심의 마음 AI
상호 연계한 원천기술 확보 및 서비스 개발을 통한 법률 Foundation AI 모델 확보



06. Cowork Process

협의체 협업체계 구축

파운데이션 모델 개발을 목표로 협의체 운영
참여 업체별 기여 방안 수립 및 유기적 협업 체계 구축을 통해 목표 달성



Q&A

Web 3, AI on Blockchain

Thank You for listening

[문의]

- 이흥노 대표: ceo.libervance@gmail.com
- 서동훈 이사: dhseo.libervance@gmail.com

2023년 AI 제품·서비스 상용화 지원사업

AI-DEX 서비스 사업화

- 최종 평가 발표용 -

리버밴스㈜

CONTENTS

01. 과제 개요
02. 과제 추진현황
03. 성과현황
04. 향후 추진일정
05. 사업비 사용내역

1. 과제 개요

1-1. 과제 개요

1.2. 배경 및 필요성, 기능 및 기술, 시장동향

1-1. 개요

AI-DEX 사업 현황

- 사업 개요: DEX의 고질적 문제 비영구적 손실(IL)을 AI-엔진을 활용해 제거하여 유동성 공급자와 소비자를 확보함.
- 목표: 2023년, 유동성 공급자 모집 진행 중, AI-DEX 사업 런칭
- WTA독점 경제에서 탈 중앙 경제로 전환:
  - Swap거래자: 최소 TX 수수료 및 LP수수료 지불하고 교환 가능
  - 유동성 공급자: 공급 자본 손실 최소화, 수익 최대화
  - AI-DEX 플랫폼 제공자: 거버넌스 토큰 수입
  - 자체 메인넷 WorldLand (Ethereum L2)를 AI-DEX전용으로 사용
- 임펙트: Uniswap DEX 및 CEX거래소를 능가하는 AI-DEX 서비스 제공 가능. 다수의 유동성 공급자가 금융 소득을 안정적으로 얻을 수 있게 됨. 소비자는 유동성이 풍부한 시장에서 안전하게 환전할 수 있게 됨.
- 현황: 예산 100%소진, 고용 100%, 기술고도화 100%, 서비스런칭 완료

1-2. 배경 및 필요성

DEX의 정의

DEX는 탈중앙거래소입니다. 즉, 이더리움에서 암호 자산 교환 서비스를 제공하는 Smart Contract를 의미합니다.

DEX 란?

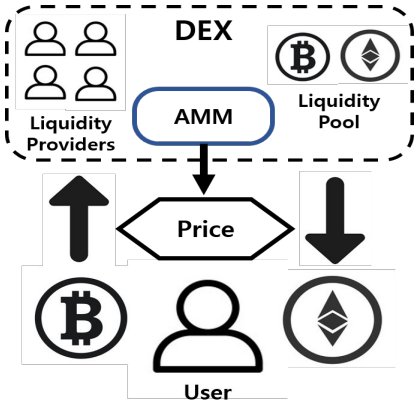
- 탈중앙거래소(decentralized exchange; DEX)는 디지털자산 교환(decentralized application; dApp)하는 알고리즘 임.
- DEX를 이용하면 **중앙화거래소의 불투명한 시스템을** 신뢰할 필요가 없음.
- DEX는 지갑을 갖고 있는 누구나 암호자산을 거래할 수 있음(즉, KYC 필요치 않음)

BINANCE UPbit

Centralized Ex (CEX)
Binance, Upbit

- 수수료 독점
- 거래내역열람불가
- 자전거래 의심
- 거래순서조정 의심





Uniswap

- 오픈소스코드
- 수수료수입투명 배분
- 거래투명성
- 비허가성
- 다양한서비스활용 (스테이킹, 렌딩, 풀)

중앙화 거래소 (좌) 및 탈중앙 거래소 (우) 개념도

1-2. 배경 및 필요성

기존 DEX의 문제점

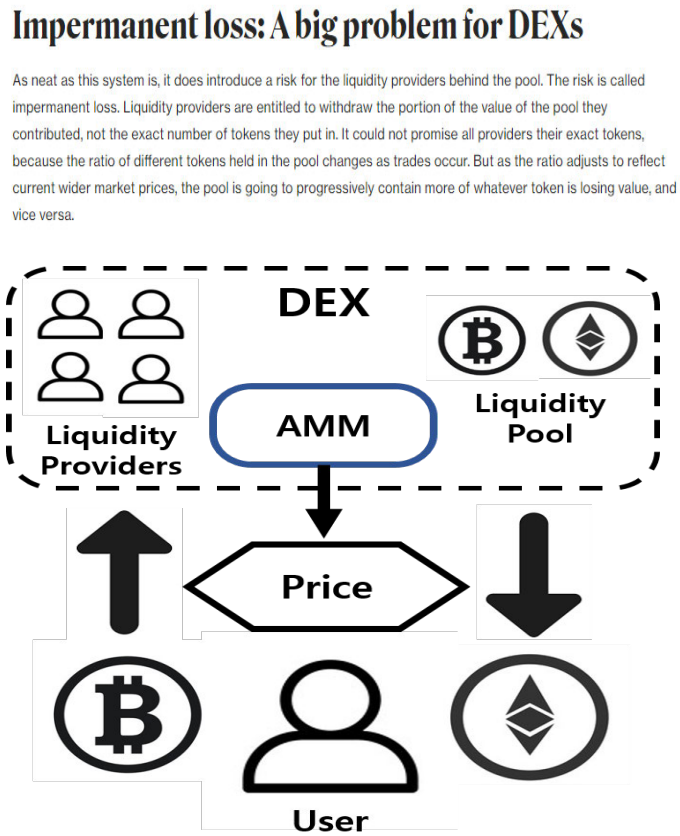
비영구적 손실 (impermanent loss; IL) 에 대한 리스크가 DEX 사용 및 성장을 저해함

기존 DEX의 문제점

- 대부분의 DEX는 빠른 거래 체결을 위해 AMM(automated market maker) 알고리즘에 기반함.
- AMM 기반 DEX 에서의 Key 플레이어 중 하나가 바로 **유동성 공급자(liquidity provider; LP)**임.
- LP는 암호자산 교환에 필요한 유동성을 제공하여 교환 거래가 즉시 체결되도록 하고, 그 보상으로써 거래액에 비례한 수수료를 얻음.
- 하지만 **LP들에게는 비영구적 손실 (IL) 이라는 금융적 리스크가 있음**. 비영구적 손실은 **LP가 유동성 공급에 참여하기 꺼리는 원인이** 됨.
- DEX의 유동성이 부족해지면 일반 사용자들은 거래지연 또는 과도하게 큰 슬리피지 같은 문제를 겪을 수 있으며 이는 **DEX의 성장을 방해**함.
- 따라서 **비영구적 손실을 제거하여 사용자 경험을 개선할 수 있는 새로운 DEX에 대한 요구**가 늘어나고 있음.

AI 예측모델을 이용하여 문제 해결

- **AI 예측모델**을 사용하여 **향후 수십블록 동안의 시장가격을 예측**하고, **이를 batch 입력** 한다면 상기 문제 해결 가능
- 그 결과 **현실적인 수준의 비용으로도 IL를 제거**한 DEX 운영 가능



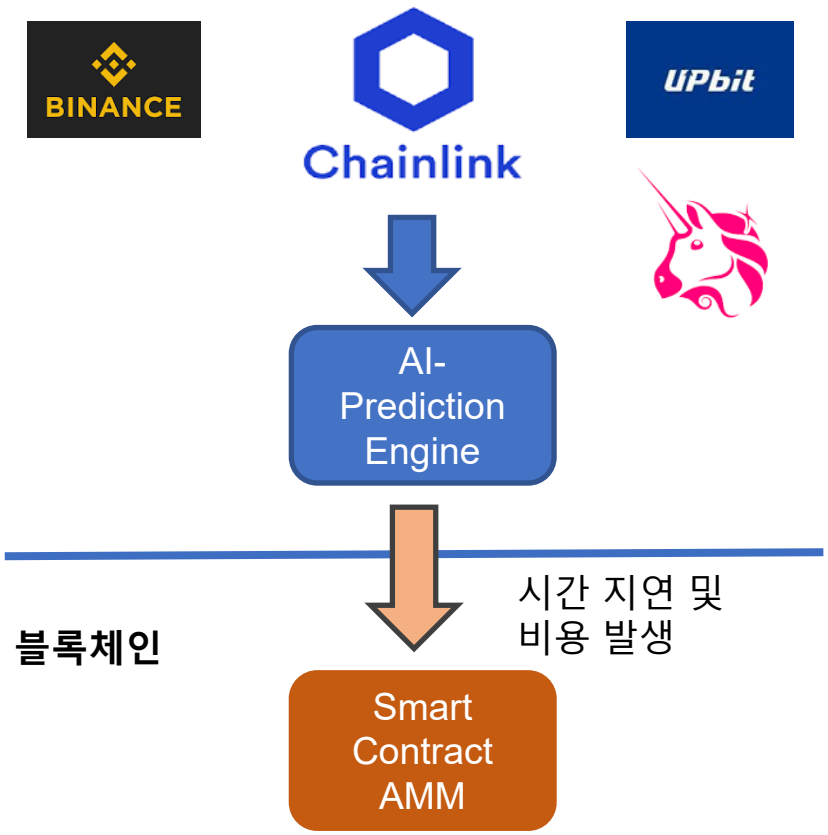
탈중앙 거래소 (DEX) 개요도

1-2. 배경 및 필요성

기술개발의 방향

AI-엔진의 역할, 자산싯가를 예측, 블록체인으로 데이터피드 하는 것

- 비영구적 손실의 원인은 차익거래임.
- **시장값으로 자산을 거래하면 차익거래는 사라짐.**
- AI-엔진은 자산의 시장값을 전세계 자산시장에서 불러와 기초 데이터를 확보함
- 평균값과 거래량 등을 계산하여 요약데이터를 만듬.
- AI-엔진은 확보한 데이터를 활용, **자산의 근미래 싯가를 예측하고, 블록체인에 저장함.**
- AI-DEX Smart Contract는 저장된 자산 예측값을 활용하여 거래 시점 자산쌍의 Swap 거래값으로 제시함.



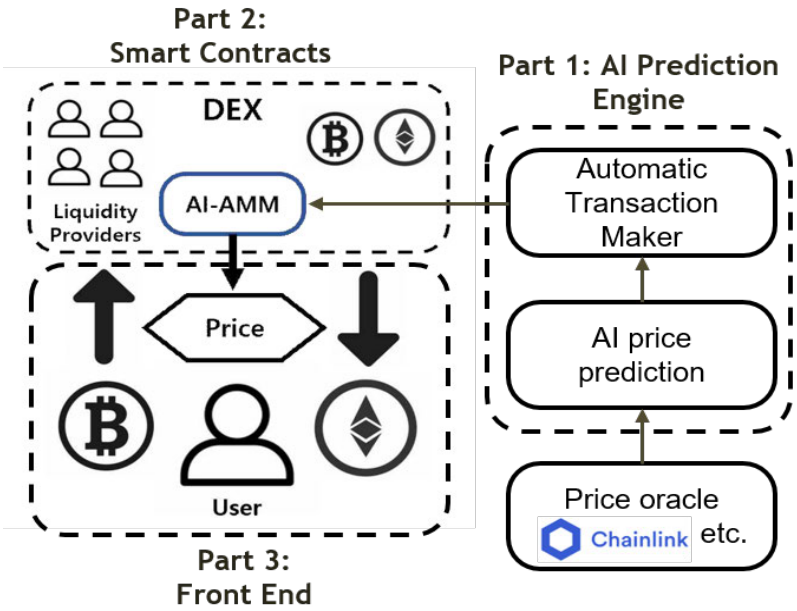
1-2. 배경 및 필요성

기 개발 AI-DEX 요약

외부 오라클로부터 암호화폐 시장 가격 수집하며, AI 예측 모델을 활용한 시장 가격 예측
시장가격을 DEX Smart Contract 에 주기적 입력하며, DEX가 싯가에 교환 가격을 산정하여 거래를 진행하도록 함

기 개발 AI-DEX 의 동작 방식

1. 외부 오라클로부터 암호화폐 시장 가격 수집
2. AI 예측모델을 사용하여 향후 수 블록 동안의 시장 가격 예측
3. 예측된 시장가격을 DEX Smart Contract 에 주기적 입력 (순간 기울기가 시장값이 되도록 Curve Control 함)
4. DEX Smart Contract (AI-AMM) 는 이 시장가격을 기반으로 현실의 시장가와 비슷한 교환 가격을 산정하여 거래 진행



1-2. 배경 및 필요성

기 개발 AI-DEX 의 성능

싯가를 Dynamic Curve 제어 알고리즘에 적용하여, 비 영구적 손실의 제거가 가능함을 보임.
운용비용 현실화, 기존 운용비용 대비 1/10, 또는 예측 모델의 성능에 따라 그 이상으로 운용비용 절감

기 개발 완료된 AI-DEX의 성능 실험 결과

1. IL 제거
- 이론적 한계점 (실시간 시장가를 사용한 DCPMM)에 근접할 정도로 IL이 낮음을 확인 가능
- AI 예측모델의 예측 정확도 개선을 통해 IL 제거 성능 향상 가능

2. 운용비용 절감
- 실시간 시장가를 사용할 때의 운용비용 대비 1/10 로 절감
- AI 예측모델의 예측기간 개선을 통해 그 이하로 절감 가능

H. Park, "Dynamic Constant Product Market Maker Using Cryptocurrency Price Prediction Based on Deep Learning," 2022, GIST 석사학위논문 (지도교수 이흥노)

- **[AI-AMM]** 'DCPMM' 대비 **일부 성능 열화를 확인**할 수 있으나, **그 폭이 크지 않음**을 확인함.
  - ✓ DCFMM 기술의 *실시간 price oracle 문제를 개선*하면서도 *IL 제어가 가능함을 확인*함.

AMM algorithms	IL Values (단위: DAI)	IL in percentages [IL values / Total Values]
CPMM	12842.30	0.022 %
DCSMM (w. AI prediction)	585909.17	1.021 %
DCPMM (w. Real price)	-28.188	-0.000049 %
AI-AMM (w. AI-prediction)	3263.58	0.0057 %
[비고] Dex 참여 안할 시 Pool 내 총 자산 가치	57392001.96	-

AI-AMM 동작에 따른 IL 제어 효과 실험 결과

*2022년도 AI (시)제품 제작 지원사업 결과 중 발췌.
https://heungno.net/ai-dex/

2. 과제 추진현황

2-1. 과제 요약

2-2. 추진현황

2-3. 목표달성 현황

2-4. 중간점검 결과 지적사항 이행 및 개선사항

2-5. 최종결과물

2-1. 과제 요약

구분	목표	실적	달성도(%)
과제 결과물	AI-DEX 기술 고도화	1) AI 예측 모듈 고도화 (완료) 2) Smart-Contract (SC) 모듈 고도화 (완료) 3) 프론트엔드 (FE) 모듈 고도화 (완료) 4) 기능 모듈간 연동 고도화 (완료)	100
	서비스 배포 및 초기 유동성 공급자 유치	1) 배포 대상 네트워크 선정 (완료) 2) 초기 유동성 공급자 및 사용자 유치(진행중) - 다양한 채널을 통한 홍보 활동 - 컨소시엄 연계 활용, 파트너사와의 협력 강화 - 베타테스트 운영	75
	AI 품질 확보	1) 성능 테스트 실시	100

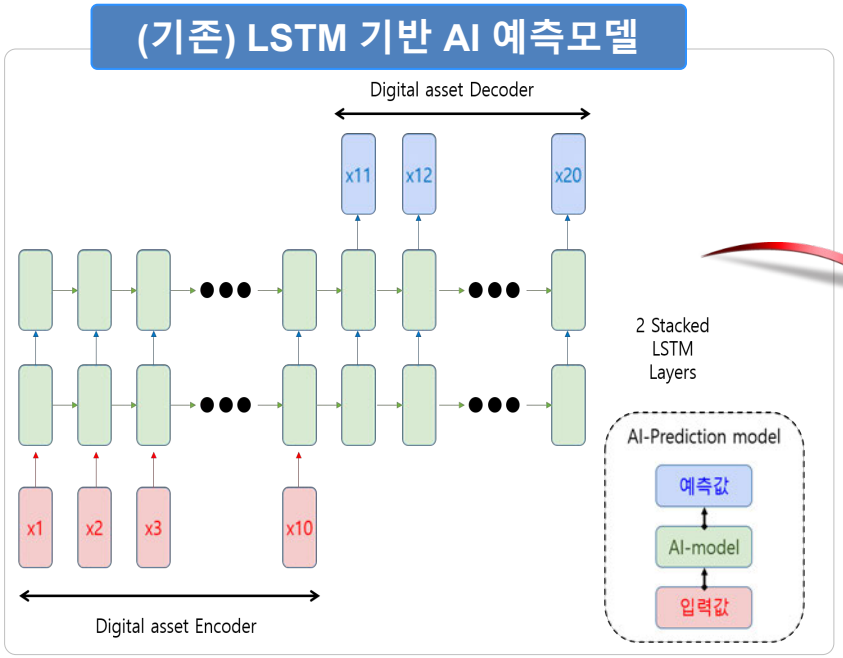
2-2. 추진현황

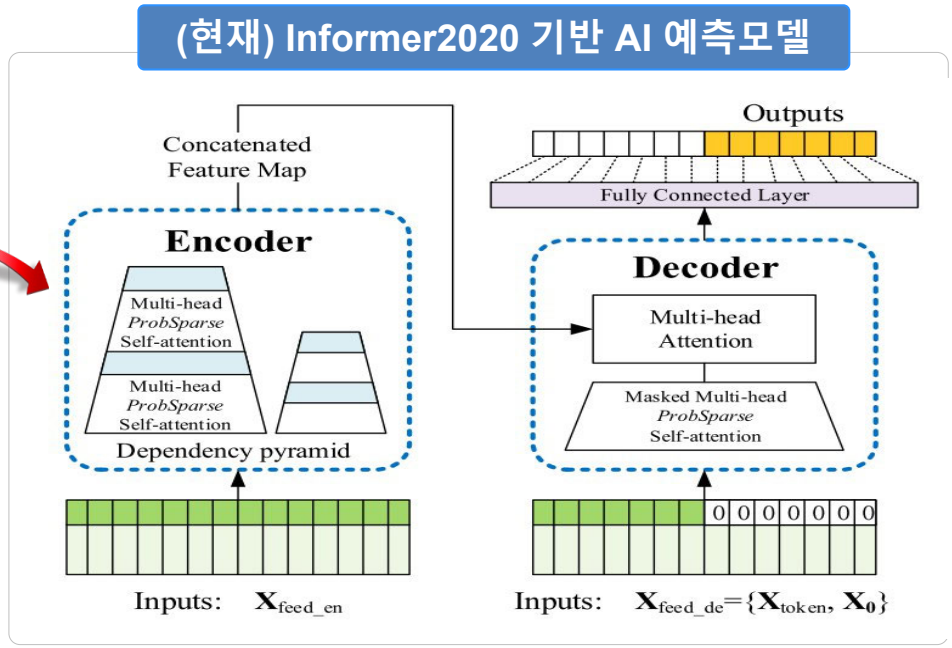
1) AI 제품·서비스(AI-DEX) 개선 및 실증

- 시제품 레벨로, 각 기능 모듈(SC, AI, FE)의 정상동작이 가능하고 기능 모듈 간 제한적인 연동이 완료된 상태임
- 본 과제를 통해 각 기능 모듈의 고도화, 기능 모듈 간 연동 고도화를 수행하여 서비스 가능 수준의 AI-DEX 개발함

AI 예측 모듈 고도화

- 예측 모델 고도화(LSTM → Informer)를 통해 기 개발된 예측 모델의 정확도 향상 완료



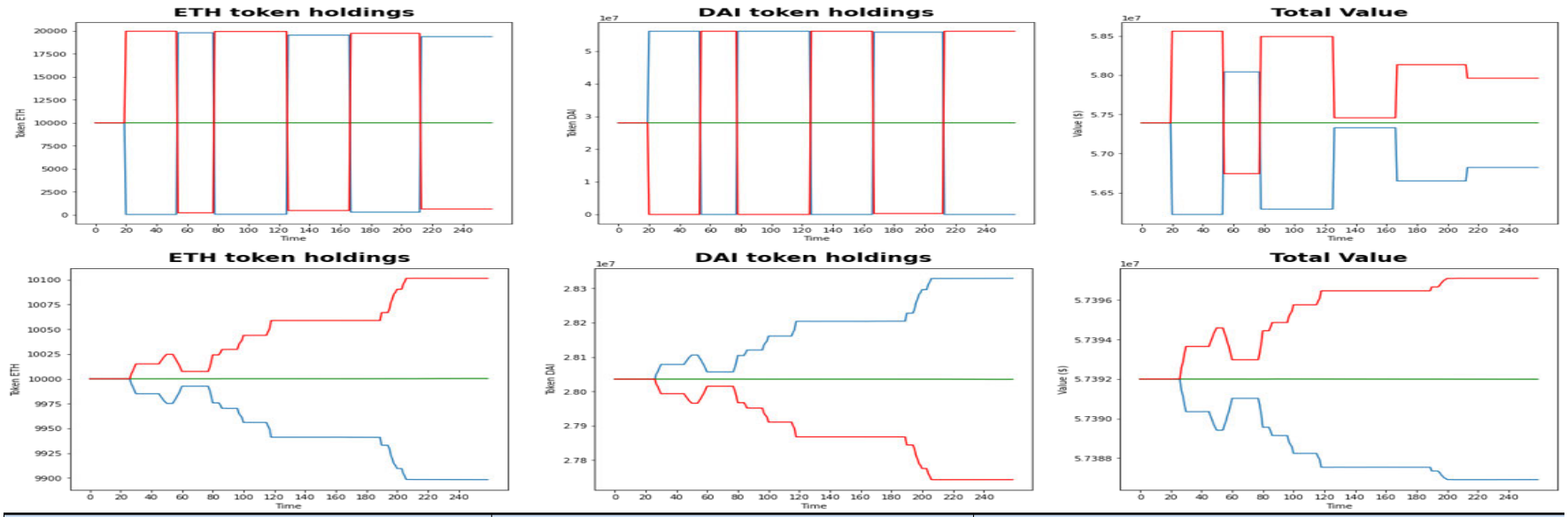


2-2. 추진현황

Smart-Contract (SC) 모듈 고도화

- (가격-블록 높이) 쌍 형태로 저장하여 시간에 따른 추가기능 구현이 쉽도록 변경 완료
- AMM 알고리즘 고도화 (DCSMM → DCPMM)를 통해 유동성 고갈 문제 해결 완료

DCSMM 및 DCPMM 알고리즘을 사용할 때의 차익거래 시뮬레이션 결과



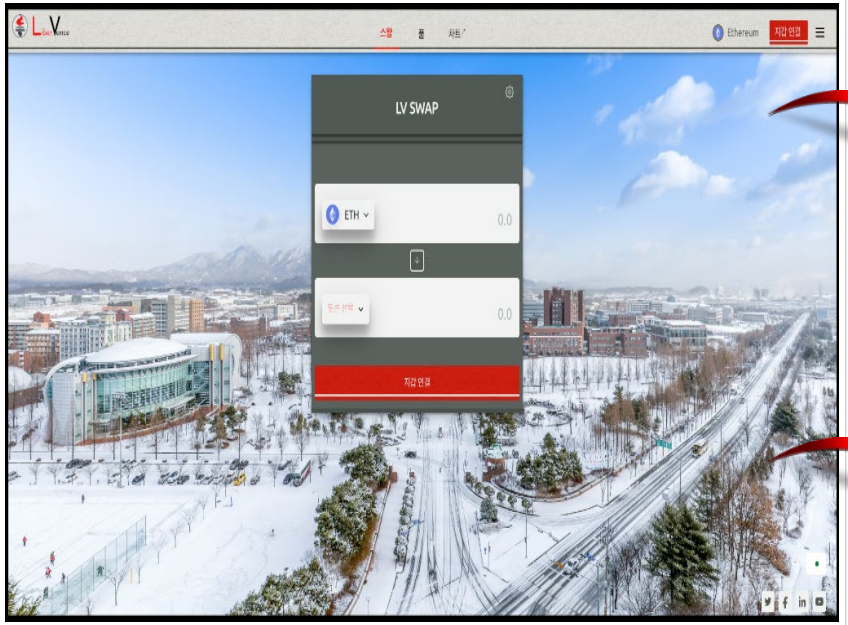
AMM algorithms	IL Values(단위: DAI)	IL in percentages [IL values / Total Values]
DCSMM (w. AI prediction)	585909.17	1.021 %
DCPMM (w. AI prediction)	4908.81	0.0086 %

2-2. 추진현황

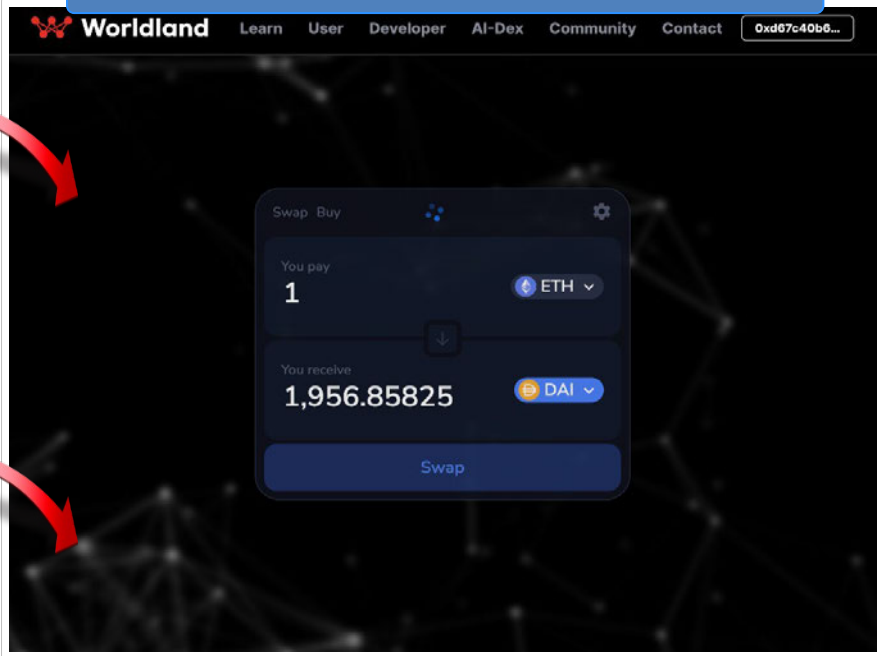
프론트앤드(FE) 모듈 고도화

- UI/UX 개선 완료, 모바일을 포함한 다양한 기기에서의 사용성 개선
- (실시간 크롤러) 실시간으로 Price oracle 데이터를 수집·검증하여 AI 예측모델에 입력
- (자동 트랜잭션 발생기) 가격 예측 값을 주기적으로 SC에 입력, 예외상황 발생 시 처리
- (FE 연동) AI 모델의 예측결과 가격, SC에 입력된 가격 및 두 가격의 싱크로 상태를 파악하여 FE에 표시

(AS- IS) FE



(TO-BE)현재 개발 완료한 FE

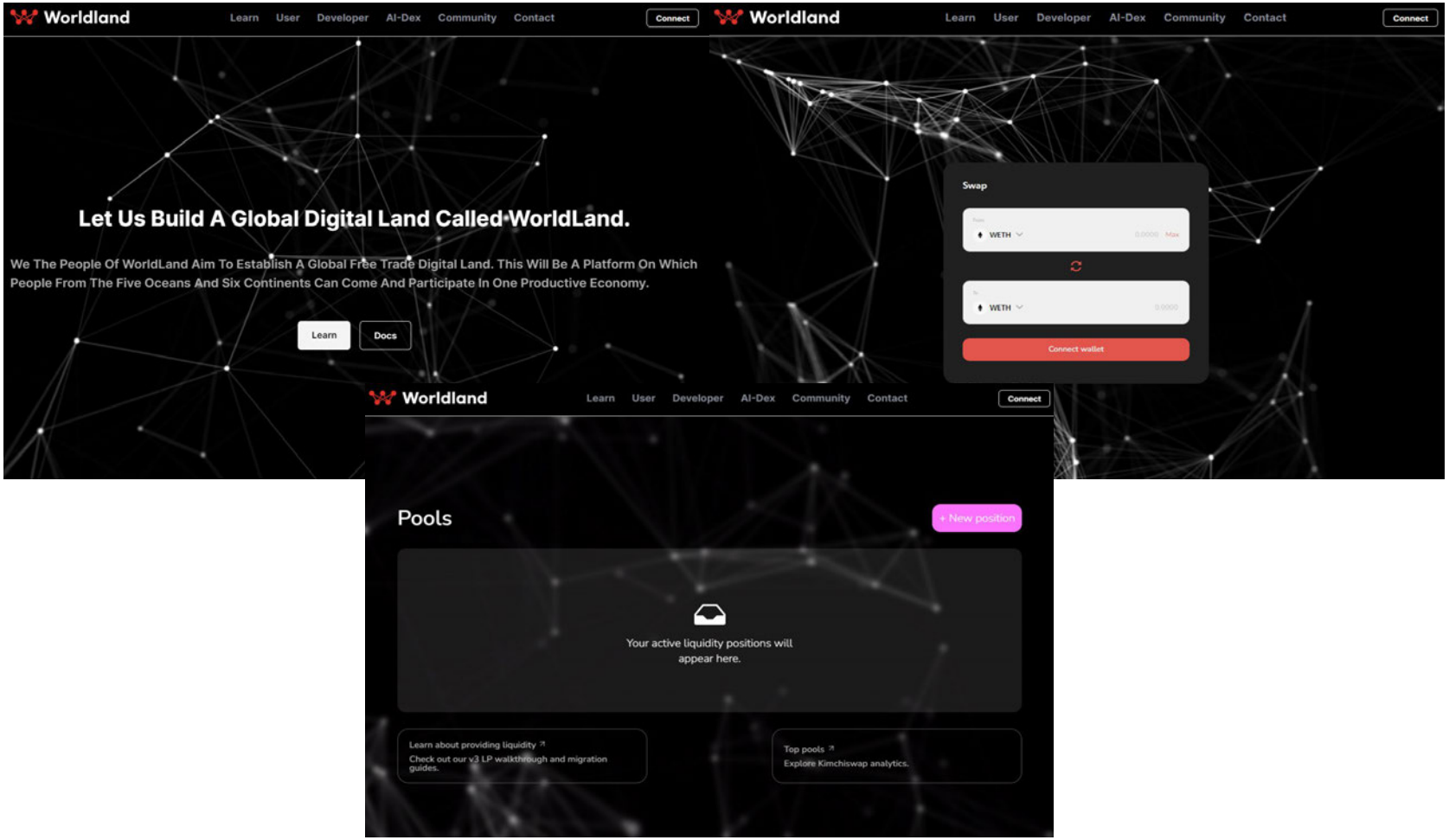


2-2. 추진현황

AI 제품·서비스 베타테스트 중

양산 및 판매 서비스 수행 계획

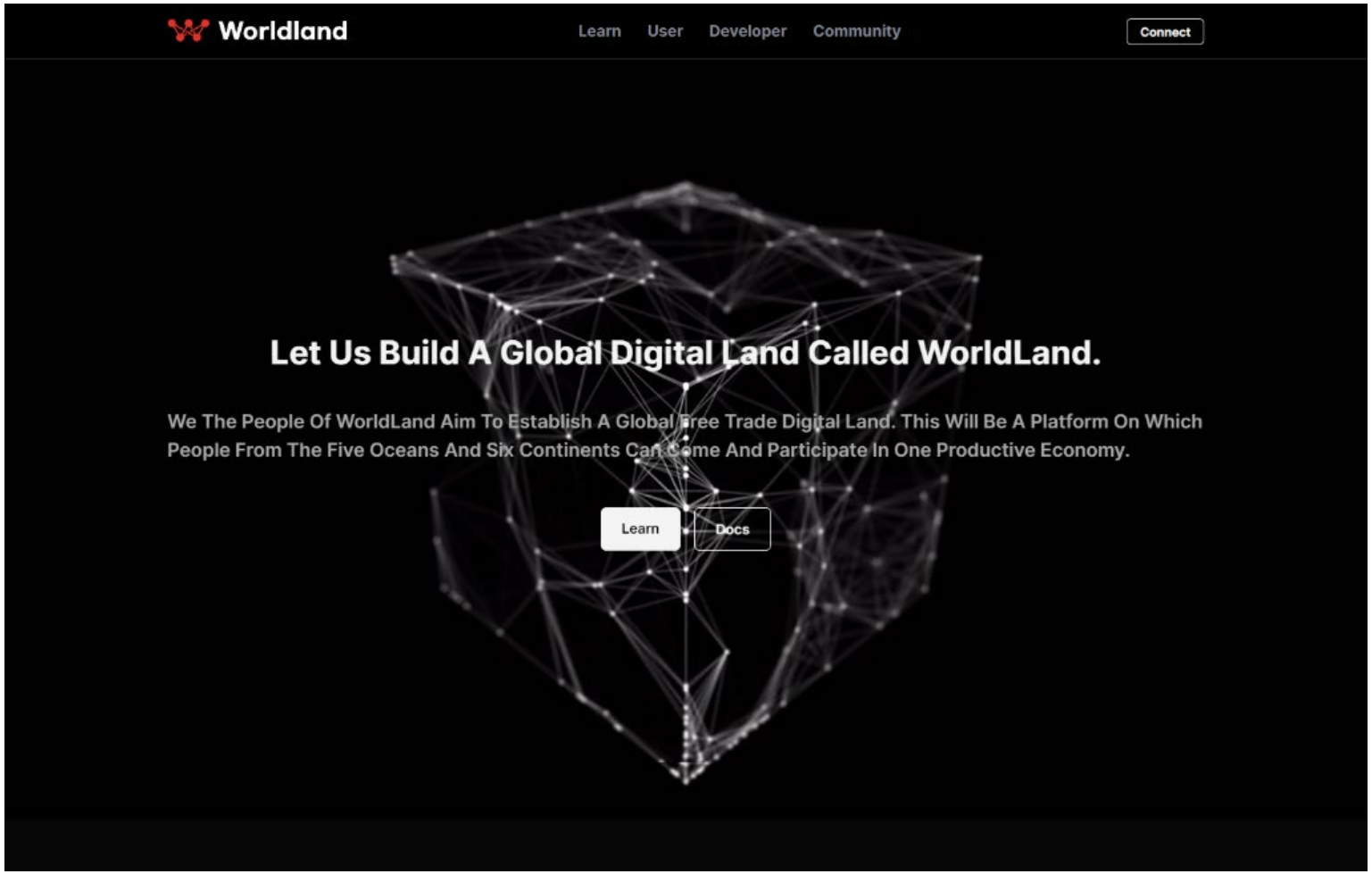
- 상용 네트워크에 AI-DEX를 배포하고, DEX 서비스를 런칭함



2-2. 추진현황

배포대상 네트워크 선정

- AI-DEX 서비스 런칭할 대상 네트워크를 선정
- 높은 보안성(양자내성)을 가지면서, 사고 발생 시 빠른 대처가 가능한 자체 개발 네트워크에 우선 배포를 결정함



2-2. 추진현황

초기 유동성 공급자 및 사용자 유치를 위한 홍보 활동

- 월드랜드 메인넷 런칭(2023.8.10.)
- AI-TECH 2023 행사 참여 및 홍보



WorldLand 메인넷 런칭 행사
서울특별시 강남구 우상향 라운지



AI-TECH2023 행사 참여 및 서비스 홍보

2-2. 추진현황

초기 유동성 공자 및 사용자 유치를 위한 홍보 활동

- KBW 2023 참가(2023.09.05 ~ 06)
- WorldLand network 및 AI-DEX 보도자료 발표



KBW 2023. 서울 신라호텔



202.07.04. 인공지능 신문

2-2. 추진현황

초기 유동성 공자 및 사용자 유치를 위한 홍보 활동

- WorldLand network 및 AI-DEX 보도자료 발표

GIST 교원창업기업 리버밴스, 8일 '월드랜드 블록체인 메인넷' 공식 출시

발행일 : 2023-08-07 09:28 지면 : 2023-08-08 20면

검증 가능한 VCA 적용 양자내성 높은 보안성 갖춰
에너지 소비율 90% 감축 환경친화 블록체인 구현

Worldland

<'월드랜드 메인넷' 로고.>

광주과학기술원(GIST) 교원창업기업 리버밴스(대표 이흥노·전기전자컴퓨터공학부 교수)가 검증 가능한 계산 알고리즘(VCA)을 적용해 차세대 블록체인 서비스를 제공하는 '월드랜드(WorldLand) 블록체인 메인넷'을 8일 공식 출시한다.

월드랜드는 인공지능(AI)과 블록체인 등 글로벌 변화와 발전 속도에 선제적으로 대응하기 위한 우리나라의 융합기술 발전과 이를 접목한 사업화 노력의 결과물이다.

리버밴스는 과학기술정보통신부·정보통신기획평가원(IITP) 대학정보통신기술(ICT)육성지원사업(ITRC) 과제와 중소벤처기업부 주관 민간투자주도형 기술창업지원(TIPS) 프로그램 지원으로 개발했다.

월드랜드에 적용한 VCA 합의 프로토콜은 투명하고 견고하며 진보된 솔루션 제공을 목표로 한다. VCA의 독자적 기술을 통해 높은 탈중앙성을 갖추면서도 고도화된 암호화 기술로 양자 내성은 물론 일반 집적회로와 달리 특정한 용도에 따라 제작된 주문형 반도체(Anti-ASIC) 특성을 갖춰 높은 보안성을 보인다.

월드랜드는 가상자산업계가 맞닥뜨리고 있는 환경문제에 대한 해결책으로서의 블록체인 기술이 집약돼 있다.

비트코인으로 대표되는 가상자산 채굴은 에너지 집약적 작업이다. 채굴자의 채굴과정(해싱 프로세스)에서 대규모 전력소모로 기후변화 촉진에 영향을 준다. 하지만 월드랜드는 현재 비트코인 네트워크 에너지 소비율의 90% 이상을 줄일 수 있는 이른바 '그린 VCA 기술'을 적용해 작업증명(PoW)에 기반함에도 불구, 획기적인 환경친화적 블록체인을 구현했다.

이흥노 리버밴스 대표는 "블록체인과 그에 따른 가상자산 시장은 코로나 19에 따른 비대면 기술 수요에 편승해 폭발적 성장을 이뤄냈다"면서 "그러나 탈중앙성이 악용돼 신뢰할 수 없거나 기술력을 갖추지 못한 법인 혹은 재단의 무분별한 가상화폐 발행이 투자자 및 시장참여자에게 피해를 주는 사례 또한 빈번히 발생했다"고 말했다.

2023.08.07. 전자신문

GIST 블록체인지능융합센터·리버밴스, '월드랜드' 블록체인 메인넷 론칭 행사 개최

발행일 : 2023-08-11 09:09



<GIST 블록체인지능융합센터와 GIST 교수 창업기업인 리버밴스가 10일 서울 강남구 소재 우상향 라운지에서 개최한 '월드랜드(WorldLand)' 블록체인 메인넷의 성공적인 출시를 기념하는 론칭 행사에서 참석자들이 기념촬영하고 있다.>

광주과학기술원(GIST) 블록체인지능융합센터와 GIST 교수 창업기업인 리버밴스(대표 이흥노)는 10일 서울 강남구 소재 우상향 라운지에서 '월드랜드(WorldLand)' 블록체인 메인넷의 성공적인 출시를 기념하는 론칭 행사를 개최했다.

월드랜드는 대학정보통신기술연구센터(ITRC)인 GIST 블록체인지능융합센터와 리버밴스가 공동으로 추진한 블록체인과 인공지능 융합연구 및 사업화 노력으로 탄생한 글로벌 네트워크다. '마이 AI 네트워크'를 중심으로 한 주요 응용 서비스를 목표하고 있는 월드랜드는 지난 8일 론칭했다.

마이 AI 네트워크는 이용자 개별적인 요구에 부응하도록 설계된 AI 네트워크로 개인이나 기업, 기관 등이 보유한 고유 데이터로 각각의 마이(MY) AI 에이전트를 훈련시킬 수 있다. 훈련된 마이 AI 에이전트들은 이용자의 서비스 요구에 프라이버시를 지키며 대응할 수 있도록 설계됐다. 필요에 따라 법률, 의료, 행정, 금융 등 다양한 분야의 전문 에이전트로 성장할 수 있다. 마이 AI 에이전트들은 마이 AI 네트워크'내에서 협력학습과 강화학습을 통해 사람들이 마주하는 다양한 문제를 공동으로 해결해 나갈 수 있다.

2023.08.11. 전자신문

2-2. 추진현황

초기 유동성 공자 및 사용자 유치를 위한 홍보 활동

- WorldLand network 및 AI-DEX 보도자료 발표

[GIST 기술융합혁신]블록체인과 AI 융합 '혁신의 길'을 닦다

발행일 : 2023-08-28 14:00 지면 : 2023-08-29 14면

Worldland

<GIST 블록체인지능융합센터와 교원창업기업 리버밴스가 최근 론칭한 월드랜드 메인넷 로고.>

인공지능(AI)과 블록체인 두 기술은 전 세계적으로 저성장이 뉴노멀로 자리잡고 있는 요즘, 새로운 성장동력을 제공할 수 있는 핵심 범용 기술로 자리 잡고 있다. 하지만 두 기술이 미래 성장을 견인할 핵심기술로 인정을 받고 있음에도, 개별 기술을 들여다보면 극복해야 할 한계점 역시 존재한다.

AI가 가진 편향성과 윤리성 훼손에 대한 우려, 오류문제, 학습 데이터 확보방안 문제 등은 많은 연구자들이 지적한 AI의 해결과제로 꼽힌다. 블록체인 역시 한계가 존재한다. 탈중앙성과 보안성, 확장성 등 기술적 제한 등 일명 트릴레마로 꼽히는 블록체인의 대표적 난제는 차세대 블록체인 혁신을 위한 극복 과제로 여전히 놓여있다.

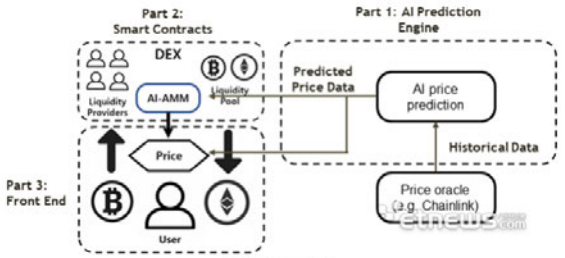
특히 독립적인 블록체인 메인넷 개발은 블록체인 개발에서 넘어야 할 큰 산으로 꼽힌다. 메인넷은 블록체인 네트워크의 핵심인 제네시스(최초) 블록부터 운영하는 블록들이 기록되는 공간이자 블록체인 프로젝트를 수행할 수 있는 독립적인 플랫폼이다.

광주과학기술원(GIST) 블록체인지능융합센터와 GIST 교원창업기업 리버밴스는 이러한 두 기술을 상호보완한 융합연구 노력을 기울이고 있다. 리버밴스는 블록체인을 통해 AI와 같은 기술의 성장으로 인해 발생하는 불평등을 완화하고 포용성, 투명성, 분산화를 촉진하며 모든 시장 참여자의 성장을 목표로 하는 블록체인 및 AI 기술 전문 기업이다. 최근 센터와 공동 연구로 탄생한 월드랜드 메인넷이 이러한 노력의 대표적 성과이다.

2023.08.28. 전자신문

[GIST 기술융합혁신]리버밴스, 'AI-DEX'로 DEX의 위험을 해결한다

발행일 : 2023-08-28 14:00 지면 : 2023-08-29 14면



<AI-DEX의 작동방식>

탈중앙 거래소(DEX)는 중앙 관리자 없이 암호화폐를 교환할 수 있는 탈중앙 애플리케이션(dApp)의 일종으로 빠른 거래 체결을 위해 자동화된 마켓 메이커(AMM) 알고리즘에 기반해 운영된다. 알고리즘에 기반한 DEX를 이용하면 중앙화 거래소의 불투명한 시스템을 신뢰할 필요가 없이 사용자 간 자유롭게 가상자산을 거래할 수 있다.

DEX 운영에서 가장 중요한 핵심 중 하나는 유동성 공급자(LP)인데, LP는 가상자산 교환에 필요한 유동성을 제공해 교환 거래가 즉시 체결되도록 하고 그 보상으로 거래액에 비례한 수수료를 얻을 수 있다. 하지만 LP들에겐 비영구적 손실로 대표되는 금융 리스크가 있어 이로 인해 유동성 공급에 참여하기를 꺼려한다.

이러한 리스크로 인해 DEX의 유동성이 부족해지면 일반 사용자들은 거래지연 또는 주문 가격과 체결 가격의 차이가 발생하는 슬리피지 문제를 겪을 수 있다. 비영구적 손실을 제거해 사용자 경험을 개선할 수 있는 새로운 DEX에 대한 요구가 늘어가고 있다.

리버밴스가 개발한 인공지능(AI)-DEX는 비영구적 손실을 해결하기 위해 AI 기술을 사용해 자산 쌍의 미래 가격을 예측하고 거래가 이뤄질 때 자산의 가치 변동을 고려해 거래를 처리한다. 이를 통해 AI-DEX의 LP에게 강요되는 비영구적 손실로 인한 리스크를 줄이고 사용자들에게는 안정적인 거래 경험을 제공할 수 있는 장점이 있다.

AI-DEX의 기술적 차별화와 경쟁력은 AMM 알고리즘과 프라이스 오라클 기술을 통해 제공한다. AI-DEX의 AMM 알고리즘은 온-체인 스마트 컨트랙트(SC)에 시장값을 알려주는 프라이스 오라클 문제를 해결해 비영구적 손실을 없애는 것은 물론, 시장값 오차 문제와 슬리피지, 가격영향 문제를 개선할 수 있다. 이를 위해 AI-DEX는 다양한 마켓의 거래 데이터를 종합한 프라이스 오라클 데이터를 수집해 향후 수(數) 블록 동안의 시장 가격을 AI 모델로 예측해 온-체인 SC에 입력시킨다.

2023.08.28. 전자신문

2-2. 추진현황

초기 유동성 공자 및 사용자 유치를 위한 홍보 활동

- WorldLand network 및 AI-DEX 보도자료 발표

[GIST 기술융합혁신]이흥노 블록체인융합센터장·리버밴스 대표 "모두를 위한 기술 혁신 꿈꾼다"

발행일 : 2023-08-28 14:00 지면 : 2023-08-29 14면



<이흥노 GIST 블록체인융합센터장·리버밴스 대표.>

이흥노 광주과학기술원(GIST) 전기전자컴퓨터공학부 교수가 센터장을 맡고 있는 GIST 블록체인지능융합센터는 2021년 7월 설립 이후 GIST, 고려대, 서울대, 아주대 연구진이 블록체인과 AI, 사물인터넷(IoT) 등의 기술융합 연구를 수행하고 있다. 지난해 12월 기준 과학기술논문인용색인(SCI)급 논문 40여 건을 포함한 100여 건의 논문 발간과 같은 우수한 연구성과를 냈다.

140여 명의 석·박사생들이 센터 연구에 참여하며 슈퍼블록, 자일로랩스, 우락부락빌리지 등 블록체인과 AI 분야에 대한 창업기업도 탄생했다. 블록체인지능융합센터의 기반 사업인 대학정보통신기술(ICT)연구센터육성지원사업(ITRC)은 센터 운영기간(2021년-2028년) 동안 총 8건의 창업성과를 예상하고 있다.

리버밴스 또한 교원창업기업으로 이흥노 교수가 GIST에서 쌓아온 연구·개발(R&D) 기술을 사업화하기 위한 2020년 설립했으며 이후 3년 여 간 연구와 기술력을 발전시켰다.

이 대표는 "리버밴스가 가진 연구성과와 기술력은 세계적 수준이라 자부하지만 기술의 사업화는 또 다른 영역이었다"며 "GIST의 연구실에서 탄생한 블록체인과 AI 기반 기술성과가 월드랜드, 분광기, AI-DEX 등 시장에서 인정받는 사업모델로 변환되기까지 많은 시도와 고민이 수반됐다"고 말했다.

"지난 2021년 대한민국의 R&D 투자비중은 경제협력개발기구(OECD) 국가 중 2위를 차지할 정도로 높고, 올해 정부의 R&D 예산은 사상 처음으로 30조원을 돌파한 31조2000억원 규모로 편성됐습니다. 하지만 이러한 투자를 통해 얻은 R&D 성과에 대한 기술사업화는 여전히 연구자들의 인식과 관련 정책, 체계 등의 부족으로 미미한 수준입니다".

그럼에도 리버밴스가 지속적으로 R&D 역량을 사업화하려 했던 동력은 기술이 혁신을 창출할 수 있다는 믿음 때문이다. 블록체인과 AI를 융합해 성장위주의 발전이 아닌 시장 참여자 모두의 성장을 위한 포용의 기술이 될 수 있도록 발전시키기 위한 연구를 계속하고 있다.

그는 "리버밴스의 사업 비전은 글로벌 탈중앙 디지털 네트워크를 구축해 전 세계인들이 천부인권과 개인의 존엄성을 존중하며 일대일 거래가 가능한 투명한 플랫폼을 만드는 것"이라며 "월드랜드라는 플랫폼을 기반으로 실현되는 리버밴스의 AI 프로젝트는 중앙집중형 AI와 차별성을 지닌다. 우리의 AI 프로젝트는 분산적이며 투명하고, 효율적인 모델 학습이 가능하다"고 설명했다.

이어 "MY AI Net을 통해 모인 네트워크 진화형 집단지능은 네트워크 참여자가 직면한 다양한 상황과 영역에 적용 가능한 혁신적인 결과를 도출할 수 있다"면서 "월드랜드와 MY AI Net이 인류가 서로 돕고 함께 성장하는 글로벌 사회의 기반이 되기를 바란다"고 덧붙였다.

리버밴스의 노력은 창업 2년 후 2022년 벤처캐피탈(VC) 엔젤투자를 확보하면서 결실을 맺기 시작했다. 엔젤투자 유치에 성공하자 민간투자주도형 기술창업지원(TIPS)를 비롯한 외부 투자를 연이어 성공시킬 수 있었다.

이 대표는 "그동안 고민했던 기술의 사업화 노력이 이제 빛을 보게 된 것 같아 기쁘다"며 "지식과 시간이 창출한 가치를 기술을 통해 모두와 공유할 수 있도 록 보유기술에 대한 연구와 개발을 지속하겠다"고 강조했다.

2023.08.28. 전자신문

2-2. 추진현황

초기 유동성 공급자 유치계획 수립

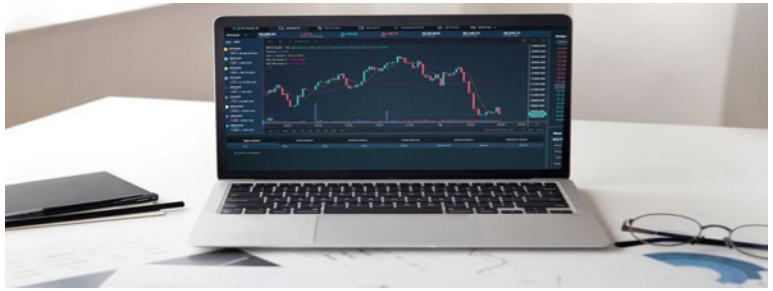
유동성 공급자 참여의사 재확인

- AI-DEX 서비스 런칭 시 유동성 공급자 참여 의향서 제출한 파트너

의향서 제출 기업	액수	참여 의사
㈜그린바이오	10억원	○
(주)블록체인랩	10억원	○
㈜블록체인팩토리	5억원	○
OwlMembers Co., Ltd	10억원	○
㈜엑스포체인	10억원	○
최재경	10억원	○

세부 추진 계획

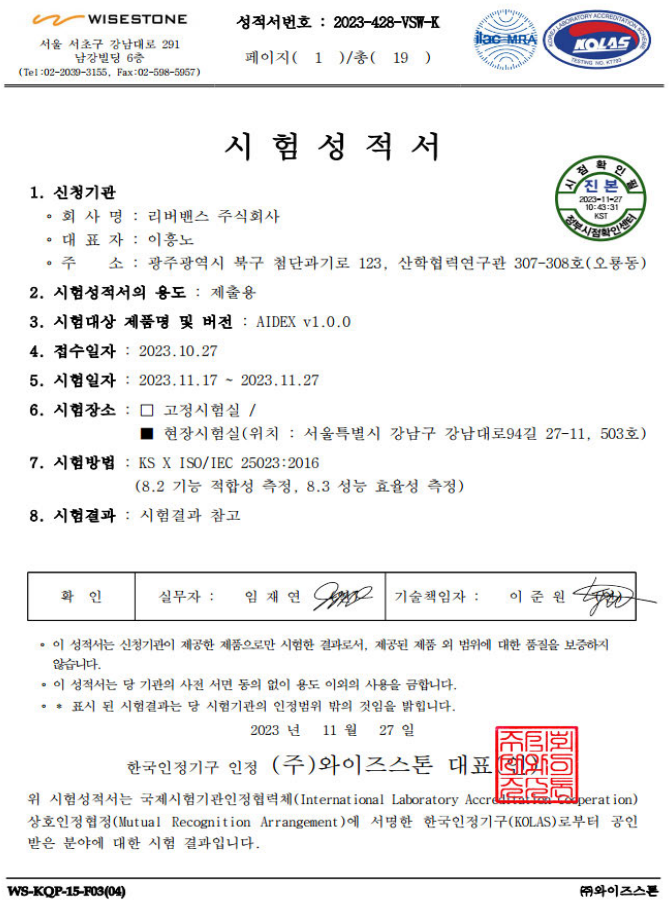
- 제공한 AI-DEX 서비스의 안정화와 트래픽 해결
- 안정적인 서비스 운용과 위험 요소 예측 및 예방
  → 안전한 거래 환경 제공을 통한 거래 활성화
  → 유동성 공급자 참여
- 적극적인 프로모션 계획수립을 통한 파트너 모집 → 활성화를 위한 혜택 계획 수립 (보상을 유동성 풀에 재투자하는 자동 복리 방식 도입)
- 시장의 가격, 규정 준수, 재무 건전도, 평판 등의 기준을 고려한 공급 계획 수립



2-2. 추진현황

AI 품질 확보

- AI 서비스 성능 테스트 및 학습용 데이터 세트 품질검증을 전문 기업·기관을 활용하여 수행하고 해당 항목의 테스트 결과서, 시험 성적서를 발행함



WISESTONE
서울 서초구 강남대로 291
남강빌딩 6층
(Tel:02-2039-3155, Fax:02-598-5957)

성적서번호 : 2023-428-VSW-K

페이지(1)/총(19)

시 험 성 적 서

1. **신청기관**
   - 회 사 명 : 리버벤스 주식회사
   - 대 표 자 : 이홍노
   - 주 소 : 광주광역시 북구 첨단과기로 123, 산학협력연구관 307-308호(오룡동)
2. **시험성적서의 용도** : 제출용
3. **시험대상 제품명 및 버전** : AIDEX v1.0.0
4. **접수일자** : 2023.10.27
5. **시험일자** : 2023.11.17 ~ 2023.11.27
6. **시험장소** : □ 고정시험실 /
   ■ 현장시험실(위치 : 서울특별시 강남구 강남대로94길 27-11, 503호)
7. **시험방법** : KS X ISO/IEC 25023:2016
   (8.2 기능 적합성 측정, 8.3 성능 효율성 측정)
8. **시험결과** : 시험결과 참고

확 인	실무자 : 임 재 연	기술책임자 : 이 준 원

- 이 성적서는 신청기관이 제공한 제품으로만 시험한 결과로서, 제공된 제품 외 범위에 대한 품질을 보증하지 않습니다.
- 이 성적서는 당 기관의 사전 서면 동의 없이 용도 이외의 사용을 금합니다.
- * 표시 된 시험결과는 당 시험기관의 인정범위 밖의 것임을 밝힙니다.

2023 년 11 월 27 일

한국인정기구 인정 (주)와이즈스톤 대표

위 시험성적서는 국제시험기관인정협력체(International Laboratory Accreditation Cooperation) 상호인정협정(Mutual Recognition Arrangement)에 서명한 한국인정기구(KOLAS)로부터 공인 받은 분야에 대한 시험 결과입니다.

WS-KQP-15-F03(04) ㈜와이즈스톤



WISESTONE
서울 서초구 강남대로 291
남강빌딩 6층
(Tel:02-2039-3155, Fax:02-598-5957)

성적서번호 : 2023-428-VSW-K

페이지(2)/총(19)

시 험 결 과

번호	신청기관 제시 항목			KS X ISO/IEC 25023:2016 측정 지표	
	시험 항목	시험 기준	시험 결과	측정 지표	결과
1	AI 예측 엔진 오차율 (MSE)	0.001 이하	**기준 만족** (0.000 5)	8.2.2 기능 적합성 측정 - 기능 정확성 측정 - 기능 정확성	1.0
2	AI 예측 엔진 오차율 (MAPE)	0.005% 이하	**기준 만족** (0.001 1%)	8.2.2 기능 적합성 측정 - 기능 정확성 측정 - 기능 정확성	1.0
3	Swap 트랜잭션 처리 시간	20 s 이하	**기준 만족** (19.6 s)	8.3.1 성능 효율성 측정 - 시간 반응성 측정 - 평균 반환 시간	19.6 s
4	Swap 트랜잭션 효율 (gas 사용량)	160 000 이하	**기준 만족** (108 526)	8.2.2 기능 적합성 측정 - 기능 정확성 측정 - 기능 정확성	1.0
5	Swap 트랜잭션 오류율	1회 이하	**기준 만족** (0회)	8.2.2 기능 적합성 측정 - 기능 정확성 측정 - 기능 정확성	1.0
6	비허가 데이터 변조 가능성	불가능	**기준 만족** (변조 불가능함을 확인함)	8.2.2 기능 적합성 측정 - 기능 정확성 측정 - 기능 정확성	1.0

※ 본 성적서에 명시된 결과는 성적서에 명시된 시험 대상, 시험 환경, 시험 방법에 국한된 결과임

WS-KQP-15-F04(04) ㈜와이즈스톤

2-3. 목표 달성 현황

목표대비 달성도

구분	목표	실적	달성도(%)
과제목표	AI-DEX 기술 고도화	1) 개별 모듈 고도화 완료 2) 모듈간 연동 고도화 완료	100
	서비스 배포 및 초기유동성 공급자 유치	1) 배포 네트워크 선정 완료 2) 초기 유동성 공급자 유치 진행 중	75
실적 목표	특허출원 2건, 상표2건	1) 특허 출원 1건 진행 중, 1건 진행 예정 2) 상표 출원/등록완료	75
	신규고용창출 3인	신규 고용 창출 3명(서동훈, 송상훈, 사담)	100
	서비스 성능 테스트 결과서	시험 성적서, 결과서	100

목표 추진 현황

과업 내용		달성도(%)
AI-DEX 기술 고도화	1-1) 각 기능 모듈 성능 및 품질 고도화	100
	1-2) 기능 모듈간 연동 고도화	100
AI-DEX 서비스 사업화	2-1) 배포 대상 네트워크 선정	100
	2-2) 초기 유동성 공급자 및 사용자 유치 등의 구체적인 계획	50
성과목표	AI-DEX 서비스 공개	80
	지식재산권 확보(특허출원)	75
	신규고용창출	100

2-4. 중간 점검 결과 지적사항 및 이행 개선 사항

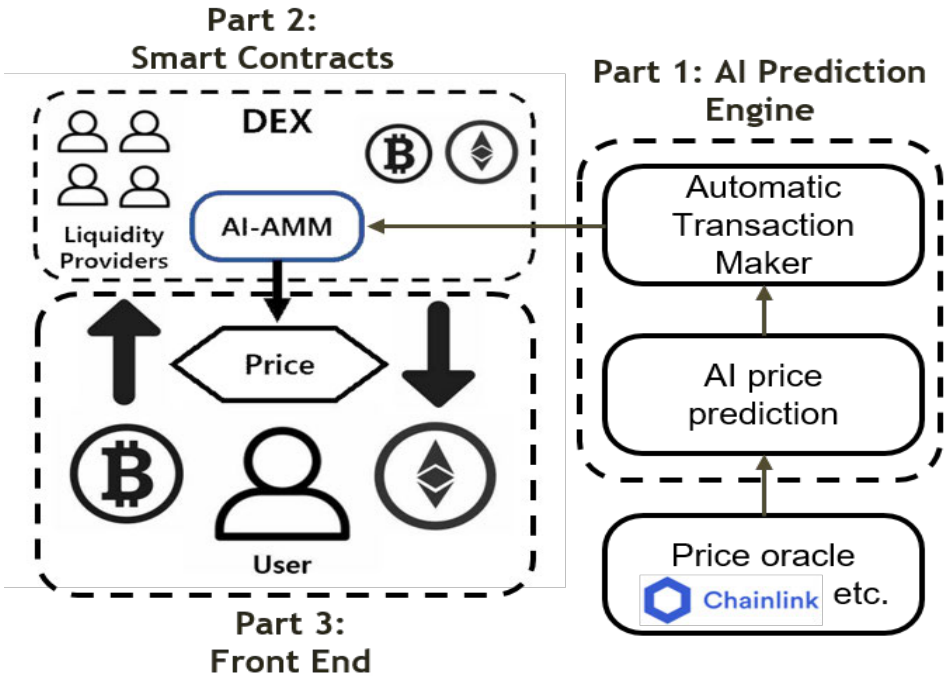
중간 점검 결과 지적사항 및 이행 개선 사항

구분	중간점검 결과	이행 및 개선 사항
의견 1	◦ 상용화 및 사업화 가능성 분석과 예측 ◦ 과제 실적 목표(특허출원, 고용창출)	◦ 개별 모듈 고도화 완료 ◦ 모듈간 연동 고도화를 통해 사업화 가능 성을 높임 - AI 예측 모듈 고도화, Smart-Contract (SC) 모듈 고도화 - 프론트엔드(FE) 모듈 고도화 ◦ 특허출원 진행중(1건), 진행 예정(1건), 상표출원(2건) ◦ 고용창출 3명
의견 2	◦ 유동성 공급자 유치 등의 구체적인 계획	◦ 배포 네트워크(WorldLand) 선정 완료 ◦ 유동성 공급자 유치를 위한 홍보행사(월드랜드 메인넷 런칭, 2023.08.10.서울 강남구 우상향) 진행과 언론 홍보 자료 발표

2-5. 최종결과물

AI-DEX 서비스 사업화

- 서비스 가능 레벨의 AI-DEX S/W
- AI-DEX 서비스 배포 및 공개
- 서비스 운영을 위한 유동성 공급자 유치



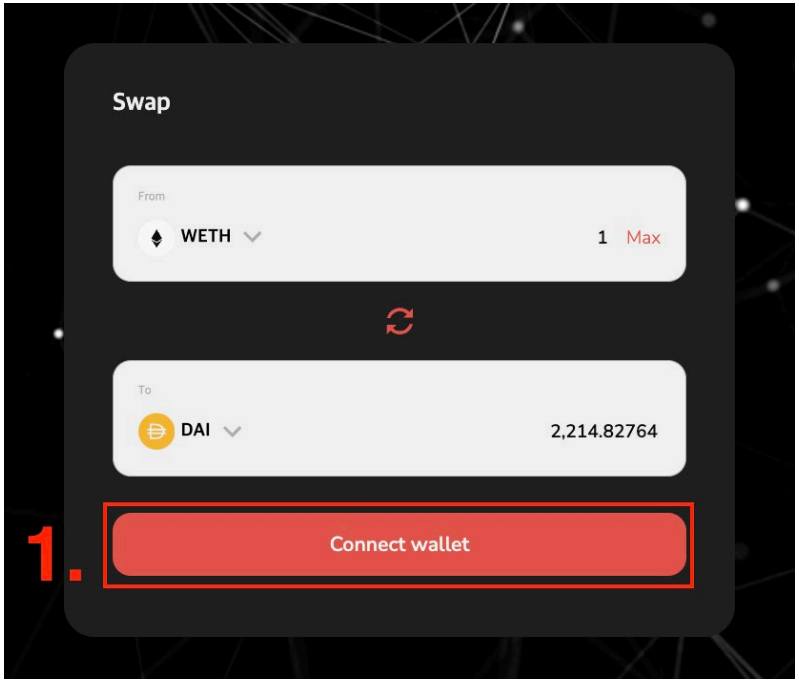
AI-DEX 의 구체적인 기술

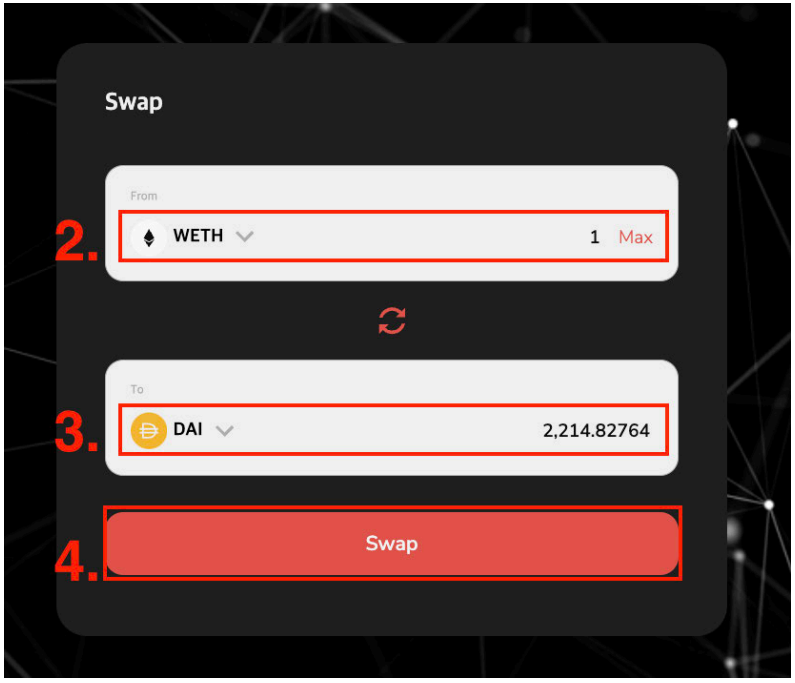
- (Part 1) AI Prediction Engine
- (Part 2) Smart Contracts
- (Part 3) Front End

2-5. 최종결과물(AI-DEX 서비스 배포 및 공개)

AI-DEX 서비스 사용자 설명

1. (프론트엔드 인터페이스) 사용자는 Connect 버튼을 통하여 메타마스크 연결작업을 진행
2. (스왑 진행) 현재 사용자가 보유중인 WETH에서 스왑하고자하는 수량 입력
3. (스왑 진행) 사용자가 입력한 WETH 수량에 따른 DAI 토큰 수량 확인
4. (스왑 진행) 스왑버튼을 클릭하여 트랜잭션 발생

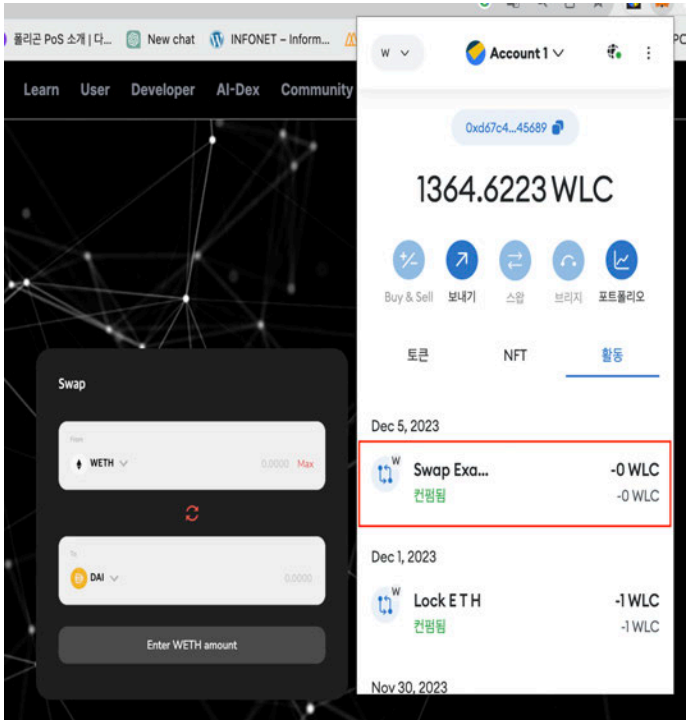




2-5. 최종결과물(AI-DEX 서비스 배포 및 공개)

AI-DEX 서비스 사용자 설명

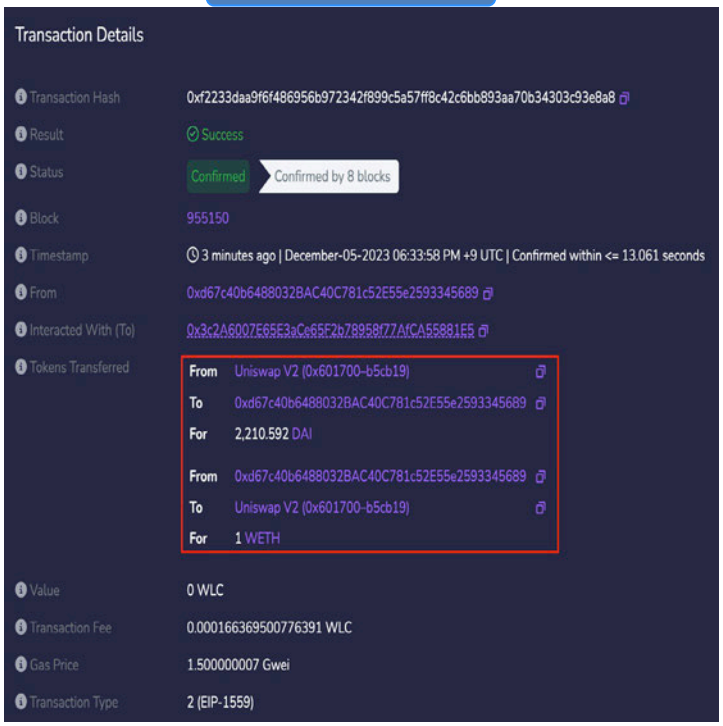
Swap 컨펌



Tx Hash



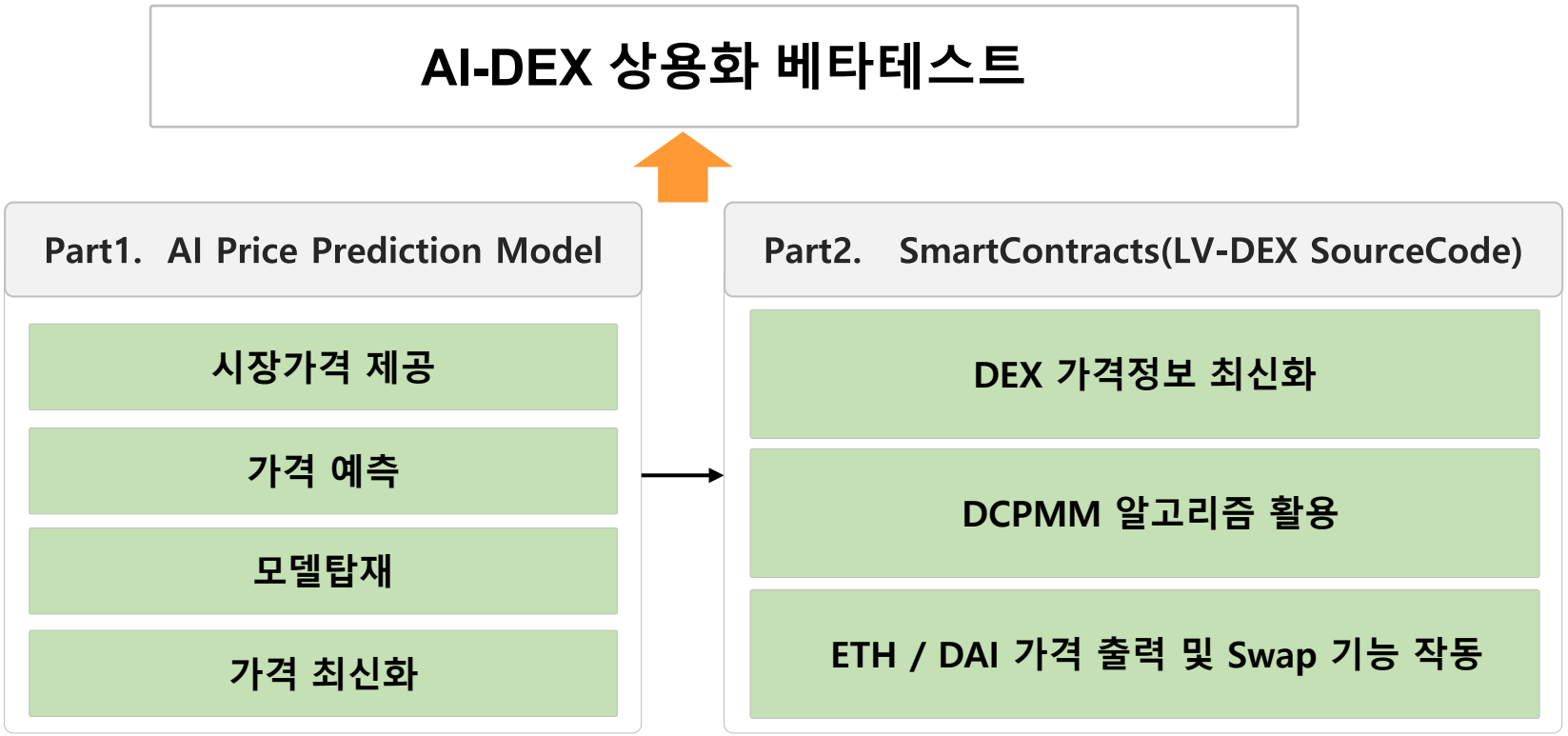
Tx Details



2-5. 최종결과물(AI-DEX 상용화 베타테스트)

AI-DEX 상용화 베타테스트 실시(Part 1 AI Price Prediction Model)

- 현재 시장 가격 제공받음 → AI 가격 예측 모델을 활용하여 10개 블록(2분간의 시간)가격 예측
- 예측모델은 재 학습 및 유지보수성을 고려하여 LSTM 기반 모델을 탑재하였음
- 서버는 매 30초 마다 LV-DEX Contract Router의 `setMarketPriceAtPool` 함수를 호출하여 예측된 가격을 담아 트랜잭션을 전송하여 시장 가격을 최신화함



2-5. 최종결과물(AI-DEX 상용화 베타테스트)

AI-DEX 상용화 베타테스트 실시(Part 2. SmartContracts(LV-DEX SourceCode)

- `setMarketPricesAtPool` 함수의 인자값으로 token0, token1, blockNumber, prices를 받아 반복 실행하여 DEX의 가격정보를 최신화함
- `setMarketPricesAtPool` 함수는 PriceSetter로 지정된 계정만 호출할 수 있음

Implemented SPEC

1. Each `Pair` contract stores a mapping from block number to corresponding marketPrice, i.e.,

```
mapping(uint => uint) public marketPrice;
```

2. A `PRICESETTER` (default: deployer of `Router02` contract) can set `marketPrice` (array input) of each `Pair` contract, by calling `setMarketPricesAtPool` function on `Router02` contract.

```
function setMarketPricesAtPool(
        address token0,
        address token1,
        uint blockNumber,
        uint[] calldata prices
    ) external;
```

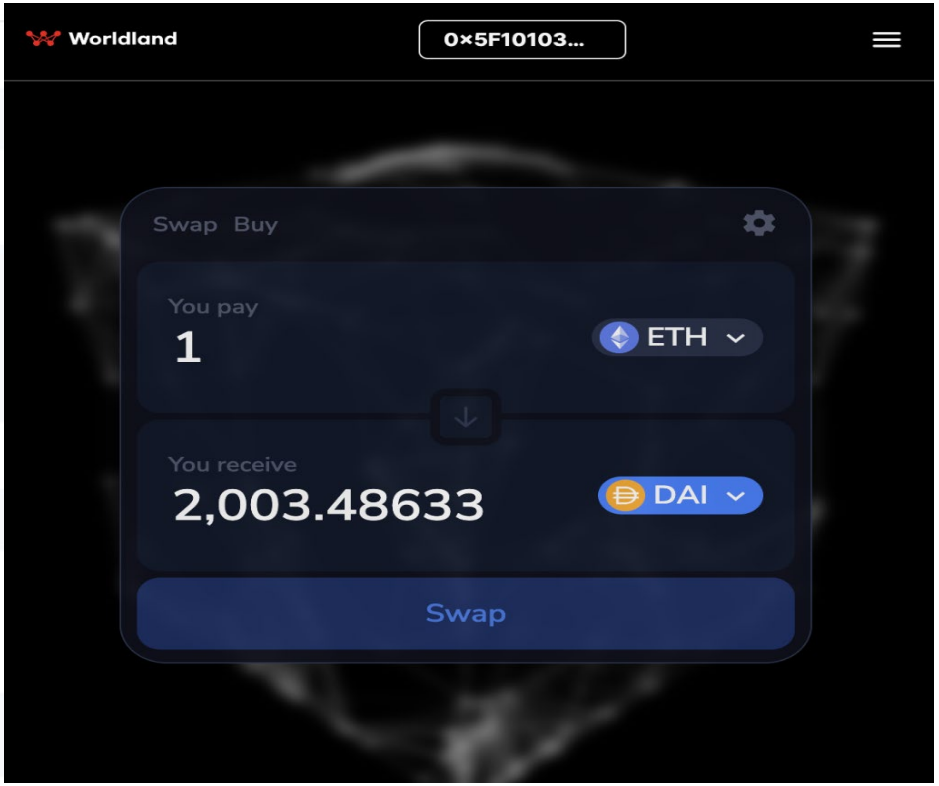
Or, by calling `setMarketPrice` function on `Pair` contract directly.

```
function setMarketPrice(uint newMarketPrice, uint index) external;
```

- encoding for marketPrice: uint with wei unit.
  ex: to set marketPrice = 1.0, marketPrice[block.number] = uint(1.0 * 10**18);

3. The trade ratio for a `Pair` is calculated by following formula (DCPMM with fee rate 0.3%):

```
OutputAmount1 = 0.997 * inputAmount0 * reserve1 / (0.997 * inputAmount0 + (reserve1 / price))
OutputAmount0 = 0.997 * inputAmount1 * reserve1 / ((0.997 * inputAmount1 + reserve1) * price)
```



2-5. 최종결과물(AI-DEX 상용화 베타테스트)

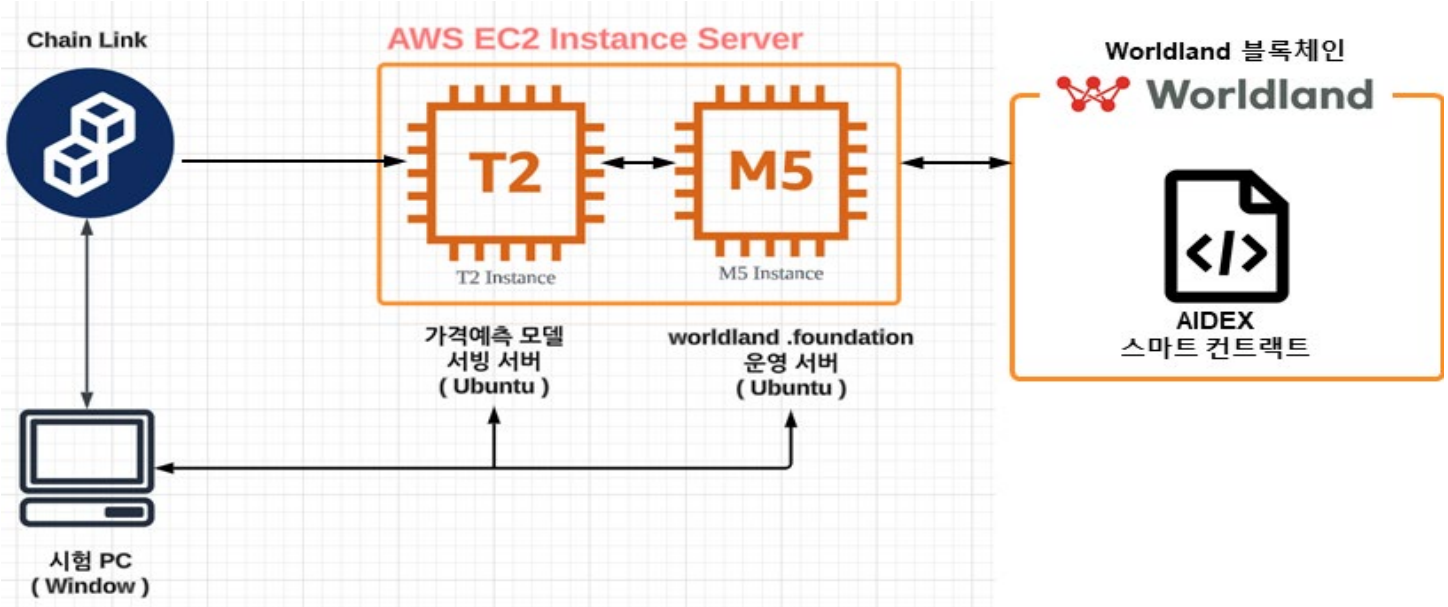
AI-DEX 상용화 베타테스트 실시 (WorldLand Network 테스트 과정)

・(프론트엔드 인터페이스) 사용자는 프론트엔드 인터페이스를 통해 1ETH당 예측된 DAI 수량으로 스왑할 수 있는 옵션을 확인하고 선택함

・(스왑 진행) 사용자가 1 ETH 을 선택하고 스왑 버튼을 클릭하면, 시스템은 블록체인과 상호작용하여 스왑 트랜잭션을 생성함

・(트랜잭션 확인) 사용자는 "WorldLand Scan"에서 스왑 트랜잭션의 해시를 확인하여 트랜잭션 상세 정보를 조회할 수 있음

・(메타마스크 확인) 사용자는 메타마스크를 열어 실제 스왑된 수량을 확인함

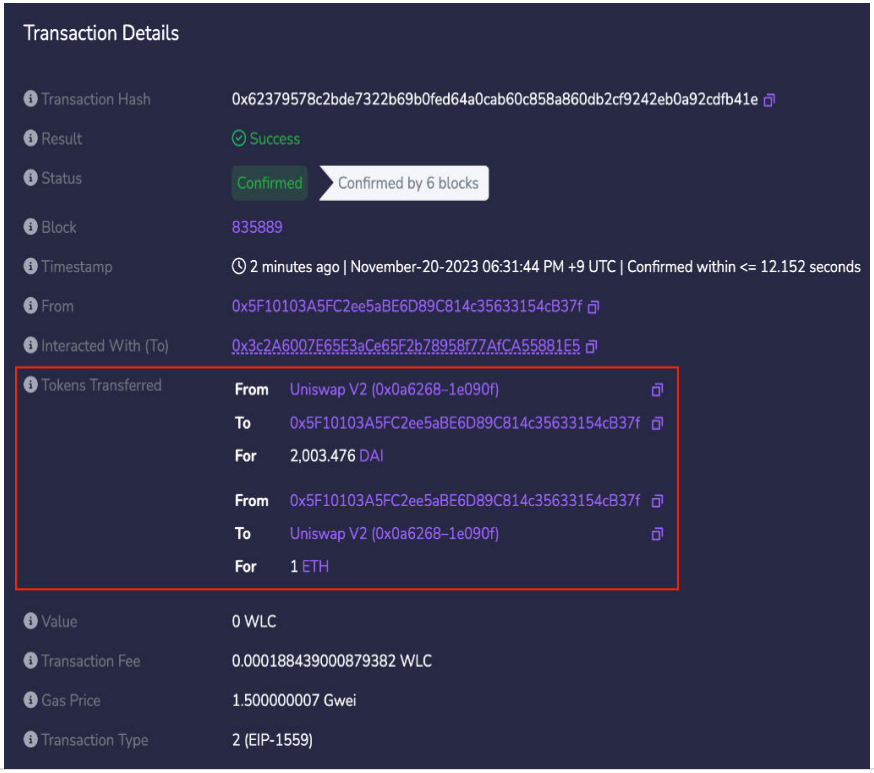


2-5. 최종결과물(AI-DEX 상용화 베타테스트)

AI-DEX 상용화 베타테스트 실시(테스트 결과 시각화 및 확인)

Scan Transaction Details

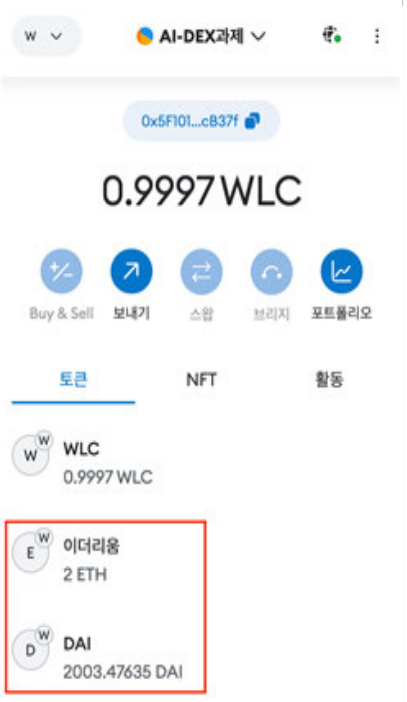
- 테스트 중 생성된 트랜잭션의 상세 정보 및 스왑 진행 내역을 확인할 수 있음



메타마스크 수량 확인

- Swap 후 메타마스크의 수량을 확인하며, 3 ETH의 수량이 2 ETH로 변경되었으며 0 DAI의 수량이 2003 DAI로 변경되었음





2-5. 최종결과물

서비스 성능 테스트 결과서(시험성적서)

WISESTONE
서울 서초구 강남대로 291
남강빌딩 6층
(Tel:02-2039-3155, Fax:02-598-5957)

시 험 성 적 서

시험확인 진본 2023-11-27 10:43:31 KST 정보시험확인센터

1. **신청기관**
   - 회 사 명 : 리버벤스 주식회사
   - 대 표 자 : 이홍노
   - 주 소 : 광주광역시 북구 첨단과기로 123, 산학협력연구관 307-308호(오룡동)
2. **시험성적서의 용도** : 제출용
3. **시험대상 제품명 및 버전** : AIDEX v1.0.0
4. **접수일자** : 2023.10.27
5. **시험일자** : 2023.11.17 ~ 2023.11.27
6. **시험장소** : □ 고정시험실 /
   ■ 현장시험실(위치 : 서울특별시 강남구 강남대로94길 27-11, 503호)
7. **시험방법** : KS X ISO/IEC 25023:2016
   (8.2 기능 적합성 측정, 8.3 성능 효율성 측정)
8. **시험결과** : 시험결과 참고

확 인	실무자 : 임 재 연	기술책임자 : 이 준 원

- 이 성적서는 신청기관이 제공한 제품으로만 시험한 결과로서, 제공된 제품 외 범위에 대한 품질을 보증하지 않습니다.
- 이 성적서는 당 기관의 사전 서면 동의 없이 용도 이외의 사용을 금합니다.
- * 표시 된 시험결과는 당 시험기관의 인정범위 밖의 것임을 밝힙니다.

2023 년 11 월 27 일

한국인정기구 인정 (주)와이즈스톤 대표이사

위 시험성적서는 국제시험기관인정협력체(International Laboratory Accreditation Cooperation) 상호인정협정(Mutual Recognition Arrangement)에 서명한 한국인정기구(KOLAS)로부터 공인 받은 분야에 대한 시험 결과입니다.

WISESTONE
서울 서초구 강남대로 291
남강빌딩 6층
(Tel:02-2039-3155, Fax:02-598-5957)

시 험 결 과

번호	신청기관 제시 항목			KS X ISO/IEC 25023:2016 측정 지표	
	시험 항목	시험 기준	시험 결과	측정 지표	결과
1	AI 예측 엔진 오차율 (MSE)	0.001 이하	**기준 만족** (0.000 5)	8.2.2 기능 적합성 측정 - 기능 정확성 측정 - 기능 정확성	1.0
2	AI 예측 엔진 오차율 (MAPE)	0.005% 이하	**기준 만족** (0.001 1%)	8.2.2 기능 적합성 측정 - 기능 정확성 측정 - 기능 정확성	1.0
3	Swap 트랜잭션 처리 시간	20 s 이하	**기준 만족** (19.6 s)	8.3.1 성능 효율성 측정 - 시간 반응성 측정 - 평균 반환 시간	19.6 s
4	Swap 트랜잭션 효율 (gas 사용량)	160 000 이하	**기준 만족** (108 526)	8.2.2 기능 적합성 측정 - 기능 정확성 측정 - 기능 정확성	1.0
5	Swap 트랜잭션 오류율	1회 이하	**기준 만족** (0회)	8.2.2 기능 적합성 측정 - 기능 정확성 측정 - 기능 정확성	1.0
6	비허가 데이터 변조 가능성	불가능	**기준 만족** (변조 불가능함을 확인함)	8.2.2 기능 적합성 측정 - 기능 정확성 측정 - 기능 정확성	1.0

※ 본 성적서에 명시된 결과는 성적서에 명시된 시험 대상, 시험 환경, 시험 방법에 국한된 결과임

3. 성과현황 및 향후 추진 일정

3-1. 성과현황

3-2. 향후 추진 일정

3.1. 성과 현황

- **(AI-DEX 고도화 S/W)** 고도화된 AI-DEX S/W 1건
- **(지식재산권 출원)** 특허 출원 1건 진행 중, 1건 추가 진행 예정. 상표권 2건 등록 완료.
- **(신규 고용 창출)** 신규 고용 창출 3명 수행
- **(성능 테스트)** ISO/IEC 25023 표준 기반 AI-DEX 시험 평가서 발행
- **(투자 유치 홍보 실적)** 투자 IR 피칭 3회 수행
- **(정부 지원사업 연계 실적)** NIPA 2023 블록체인 컨설팅 지원사업 참여

3.1. 성과 현황

고용창출(3명, 서동훈, 송상훈, 사담)

출력일시 : 2023.01.18 09:50

4대 사회보험
사업장 가입자 명부

발급번호	20230118640321	발급일시	2023-01-18 09:50	사업장 관리번호	83888015090

구분	국민연금	건강보험	산재보험	고용보험
사업자등록번호	838-88-01509	838-88-01509	838-88-01509	838-88-01509
사업장 명칭	리버밴스 주식회사	리버밴스 주식회사	리버밴스 주식회사	리버밴스 주식회사

■ 가입 내역(발급일자 현재기준) 1 / 2

연번	주민(외국인) 등록번호	성명	자격취득일			
			국민연금	건강보험	산재보험	고용보험
1	040414-3******	박선우	2023.01.16	2023.01.16	2023.01.16	2023.01.16
2	040614-3******	박준서	2023.01.16	2023.01.16	2023.01.16	2023.01.16
3	691231-1******	김항진	2022.10.01	2022.10.01	2022.10.01	2022.10.01
4	790101-2******	최경심	2022.09.01	2022.09.01	2022.09.01	2022.09.01
5	870217-1******	김광우	2022.11.01	2022.11.01	2022.11.01	2022.11.01
6	880614-1******	강주성	2021.09.01	2021.09.01	2021.09.01	2021.09.01
7	940309-1******	홍종남	2022.11.16	2022.11.16	2022.11.16	2022.11.16
8	940811-2******	김유진	2023.01.09	2023.01.09	2023.01.09	2023.01.09

▷ 위 사업장 가입자 명부는 4대사회보험 정보연계시스템이 국민연금공단, 국민건강보험공단 근로복지공단의 가입자 정보를 실시간 연계받아 제공하는 것이며, 발급사실 여부는 발급일로부터 90일까지 4대사회보험 포털사이트(www.4insure.or.kr)의 [발급사실확인] 메뉴에서 확인 가능합니다.
"편리한 정보연계서비스, 4대 사회보험이 함께 합니다."

출력일시 : 2023.11.01 11:06

4대 사회보험
사업장 가입자 명부

발급번호	20231101449677	발급일시	2023-11-01 11:06	사업장 관리번호	83888015090

구분	국민연금	건강보험	산재보험	고용보험
사업자등록번호	838-88-01509	838-88-01509	838-88-01509	838-88-01509
사업장 명칭	리버밴스 주식회사	리버밴스 주식회사	리버밴스 주식회사	리버밴스 주식회사

■ 가입 내역(발급일자 현재기준) 1 / 2

연번	주민(외국인) 등록번호	성명	자격취득일			
			국민연금	건강보험	산재보험	고용보험
1	040414-3******	박선우	2023.01.16	2023.01.16	2023.01.16	2023.01.16
2	040614-3******	박준서	2023.01.16	2023.01.16	2023.01.16	2023.01.16
3	761103-1******	서동훈	2023.06.12	2023.06.12	2023.06.12	2023.06.12
4	781115-1******	송상훈	2023.06.01	2023.06.01	2023.06.01	2023.06.01
5	861105-1******	안태호	2023.04.20	2023.04.20	2023.04.20	2023.04.20
6	880614-1******	강주성	2021.09.01	2021.09.01	2021.09.01	2021.09.01
7	901230-1******	최혁동	2023.03.01	2023.03.01	2023.03.01	2023.03.01
8	940309-1******	홍종남	2022.11.16	2022.11.16	2022.11.16	2022.11.16
9	940723-2******	박가단	2023.05.01	2023.05.01	2023.05.01	2023.05.01
10	950708-2******	조윤경	2023.05.23	2023.05.23	2023.05.23	2023.05.23
11	951121-2******	문지혜	2023.10.01	2023.10.01	2023.10.01	2023.10.01
12	960401-1******	장주찬	2023.04.20	2023.04.20	2023.04.20	2023.04.20
13	980727-5******	사담	2023.10.01			
14	980727-5******	사담백알로키디노보		2023.10.01	2023.10.01	2023.10.26

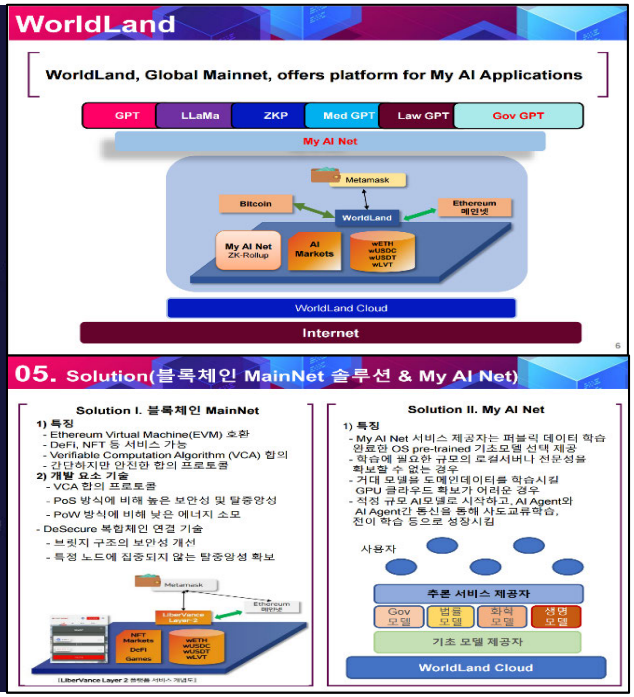
▷ 위 사업장 가입자 명부는 4대사회보험 정보연계시스템이 국민연금공단, 국민건강보험공단 근로복지공단의 가입자 정보를 실시간 연계받아 제공하는 것이며, 발급사실 여부는 발급일로부터 90일까지 4대사회보험 포털사이트(www.4insure.or.kr)의 [발급사실확인] 메뉴에서 확인 가능합니다.
"편리한 정보연계서비스, 4대 사회보험이 함께 합니다."

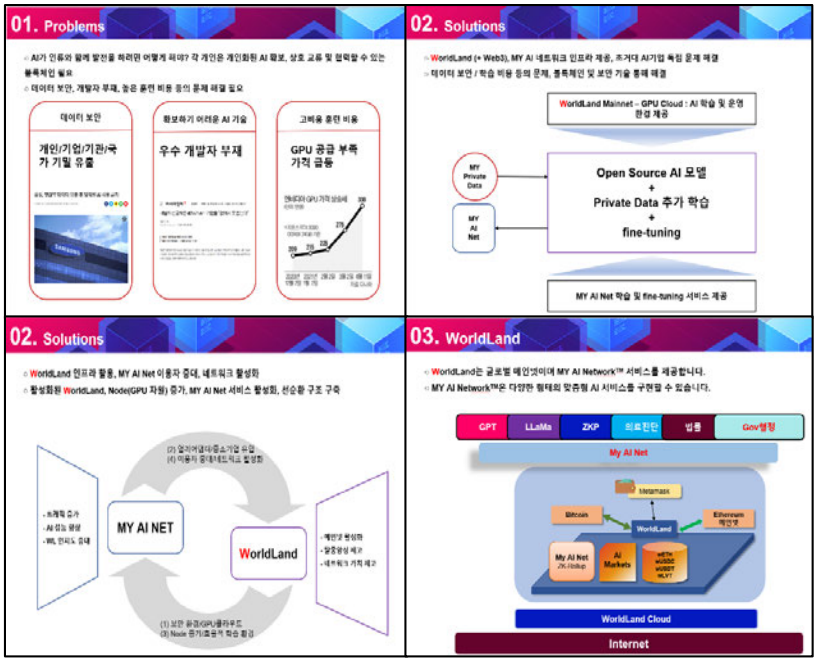
3.1. 성과 현황

투자유치(홍보) 실적

투자 유치명	투자유치 일정	비고
뻔뻔(FUN FUND)한 투자유치 IR	2023.06.13.	광주광역시 광주창조경제혁신센터
제2핀테크랩 투자사 IR(1)	2023.08.25.	서울특별시
제2핀테크랩 투자사 IR(2)	2023.09.01.	서울특별시



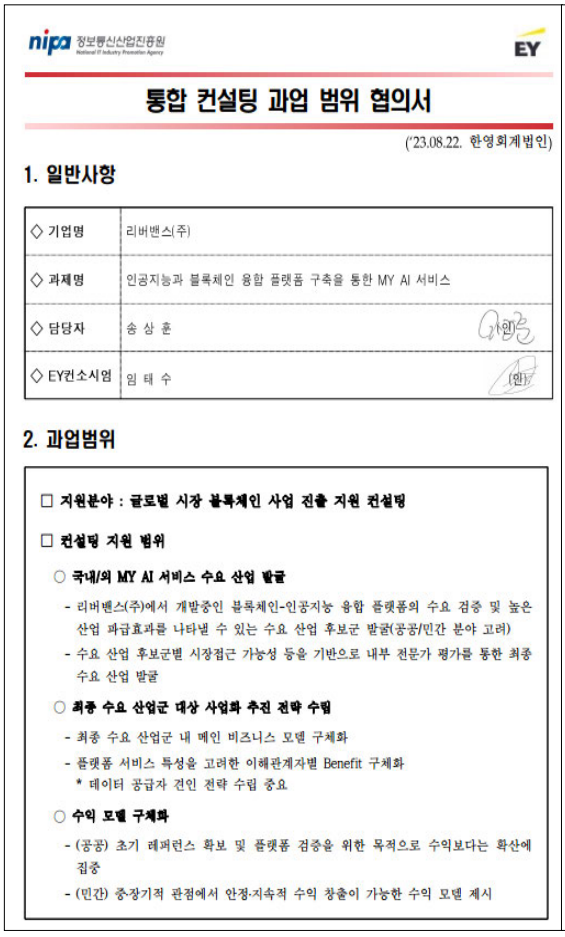




3.1. 성과 현황

정부 지원사업 연계 실적

사업명	선정 기관	비고
2023 블록체인 전문컨설팅 지원사업	NIPA(정보통신산업진흥원)	블록체인 전문 컨설팅(원포인트)



nipa 정보통신산업진흥원 EY

통합 컨설팅 과업 범위 협의서

('23.08.22. 한영회계법인)

1. 일반사항

◇ 기업명	리버밴스(주)
◇ 과제명	인공지능과 블록체인 융합 플랫폼 구축을 통한 MY AI 서비스
◇ 담당자	송 삼 훈
◇ EY컨소시엄	임 태 수

2. 과업범위

☐ **지원분야 : 글로벌 시장 블록체인 사업 진출 지원 컨설팅**

☐ **컨설팅 지원 범위**

○ **국내/외 MY AI 서비스 수요 산업 발굴**

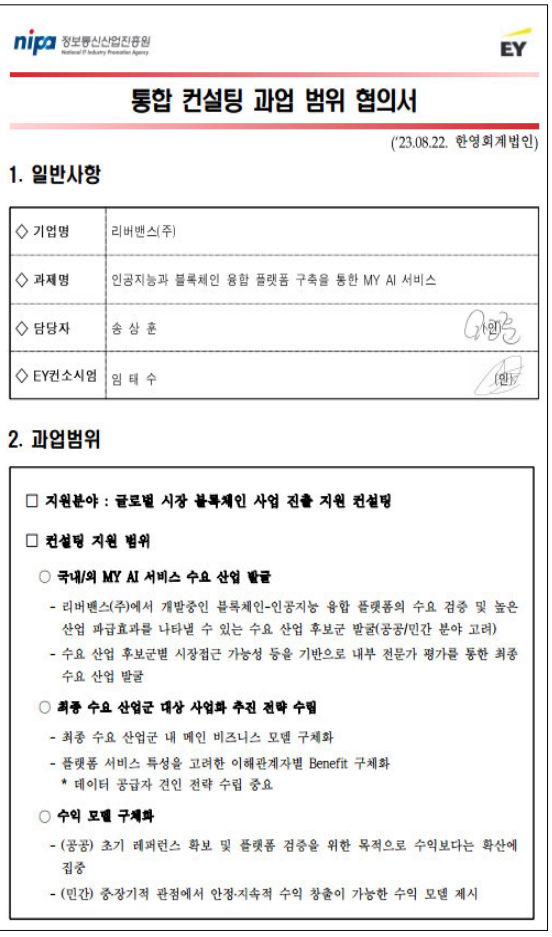
- 리버밴스(주)에서 개발중인 블록체인-인공지능 융합 플랫폼의 수요 검증 및 높은 산업 파급효과를 나타낼 수 있는 수요 산업 후보군 발굴(공공/민간 분야 고려)
- 수요 산업 후보군별 시장접근 가능성 등을 기반으로 내부 전문가 평가를 통한 최종 수요 산업 발굴

○ **최종 수요 산업군 대상 사업화 추진 전략 수립**

- 최종 수요 산업군 내 메인 비즈니스 모델 구체화
- 플랫폼 서비스 특성을 고려한 이해관계자별 Benefit 구체화
  * 데이터 공급자 견인 전략 수립 중요

○ **수익 모델 구체화**

- (공공) 초기 레퍼런스 확보 및 플랫폼 검증을 위한 목적으로 수익보다는 확산에 집중
- (민간) 중장기적 관점에서 안정·지속적 수익 창출이 가능한 수익 모델 제시



nipa 정보통신산업진흥원 EY

통합 컨설팅 과업 범위 협의서

('23.08.22. 한영회계법인)

1. 일반사항

◇ 기업명	리버밴스(주)
◇ 과제명	인공지능과 블록체인 융합 플랫폼 구축을 통한 MY AI 서비스
◇ 담당자	송 삼 훈
◇ EY컨소시엄	임 태 수

2. 과업범위

☐ **지원분야 : 글로벌 시장 블록체인 사업 진출 지원 컨설팅**

☐ **컨설팅 지원 범위**

○ **국내/외 MY AI 서비스 수요 산업 발굴**

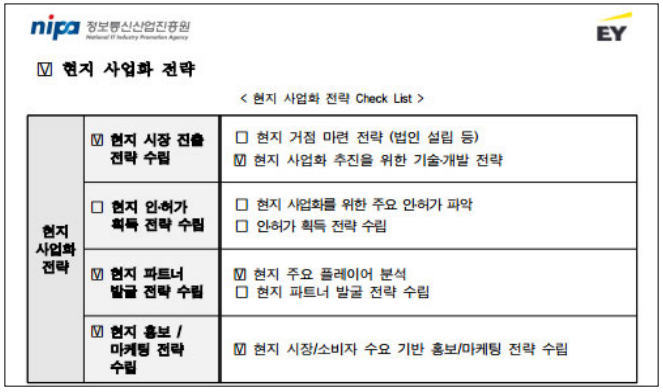
- 리버밴스(주)에서 개발중인 블록체인-인공지능 융합 플랫폼의 수요 검증 및 높은 산업 파급효과를 나타낼 수 있는 수요 산업 후보군 발굴(공공/민간 분야 고려)
- 수요 산업 후보군별 시장접근 가능성 등을 기반으로 내부 전문가 평가를 통한 최종 수요 산업 발굴

○ **최종 수요 산업군 대상 사업화 추진 전략 수립**

- 최종 수요 산업군 내 메인 비즈니스 모델 구체화
- 플랫폼 서비스 특성을 고려한 이해관계자별 Benefit 구체화
  * 데이터 공급자 견인 전략 수립 중요

○ **수익 모델 구체화**

- (공공) 초기 레퍼런스 확보 및 플랫폼 검증을 위한 목적으로 수익보다는 확산에 집중
- (민간) 중장기적 관점에서 안정·지속적 수익 창출이 가능한 수익 모델 제시

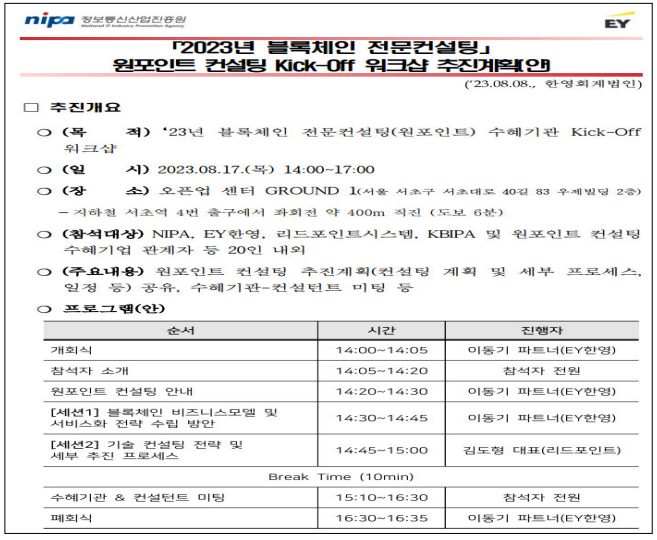


nipa 정보통신산업진흥원 EY

Ⅳ 현지 사업화 전략

< 현지 사업화 전략 Check List >

현지 사업화 전략	Ⅳ 현지 시장 진출 전략 수립	☐ 현지 거점 마련 전략 (법인 설립 등) Ⅳ 현지 사업화 추진을 위한 기술·개발 전략
	☐ 현지 인·허가 획득 전략 수립	☐ 현지 사업화를 위한 주요 인허가 파악 ☐ 인허가 획득 전략 수립
	Ⅳ 현지 파트너 발굴 전략 수립	Ⅳ 현지 주요 플레이어 분석 ☐ 현지 파트너 발굴 전략 수립
	Ⅳ 현지 홍보 / 마케팅 전략 수립	Ⅳ 현지 시장/소비자 수요 기반 홍보/마케팅 전략 수립



nipa 정보통신산업진흥원 EY

「2023년 블록체인 전문컨설팅」
원포인트 컨설팅 Kick-Off 워크샵 추진계획(안)

('23.08.08., 한영회계법인)

☐ **추진개요**

○ **(목 적)** '23년 블록체인 전문컨설팅(원포인트) 수혜기관 Kick-Off 워크샵

○ **(일 시)** 2023.08.17.(목) 14:00~17:00

○ **(장 소)** 오픈업 센터 GROUND 1(서울 서초구 서초대로 40길 83 우세빌딩 2층)
- 지하철 서초역 4번 출구에서 좌회전 약 400m 직진 (도보 6분)

○ **(참석대상)** NIPA, EY한영, 리드포인트시스템, KBIPA 및 원포인트 컨설팅 수혜기업 관계자 등 20인 내외

○ **(주요내용)** 원포인트 컨설팅 추진계획(컨설팅 계획 및 세부 프로세스, 일정 등) 공유, 수혜기관-컨설턴트 미팅 등

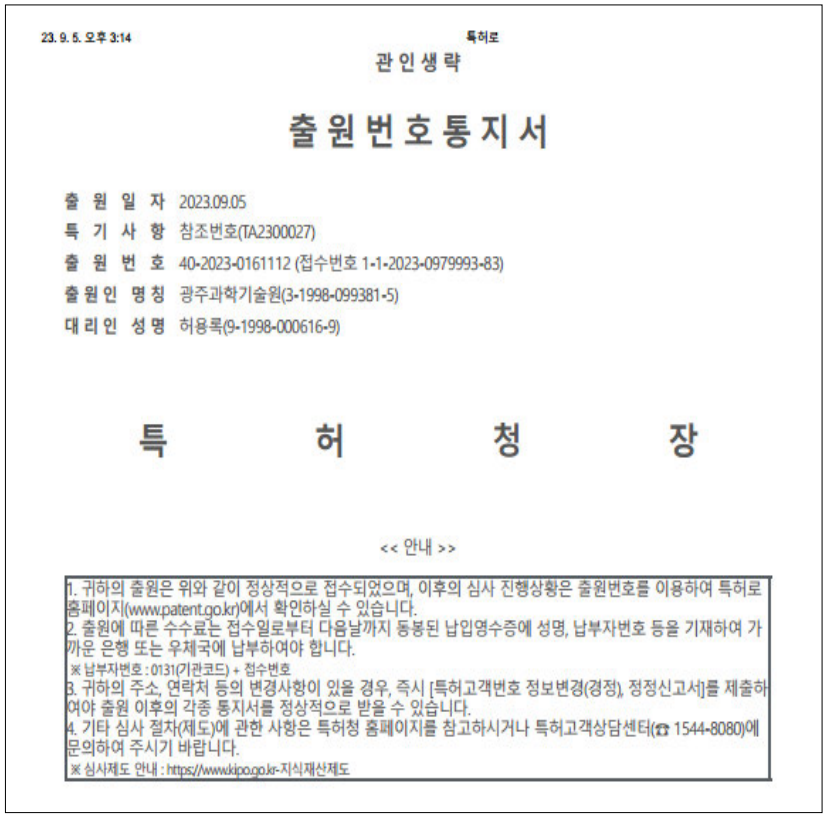
○ **프로그램(안)**

순서	시간	진행자
개회식	14:00~14:05	이동기 파트너(EY한영)
참석자 소개	14:05~14:20	참석자 전원
원포인트 컨설팅 안내	14:20~14:30	이동기 파트너(EY한영)
[세션1] 블록체인 비즈니스모델 및 서비스화 전략 수립 방안	14:30~14:45	이동기 파트너(EY한영)
[세션2] 기술 컨설팅 전략 및 세부 추진 프로세스	14:45~15:00	김도형 대표(리드포인트)
Break Time (10min)		
수혜기관 & 컨설턴트 미팅	15:10~16:30	참석자 전원
폐회식	16:30~16:35	이동기 파트너(EY한영)

3.1. 성과 현황

지식재산권 실적

지재권 종류	지재권 명	출원번호	출원일	비고
상표	월드랜드	40-2023-0161112	2023.09.05.	
	My AI Network	40-2023-0161133	2023.09.05.	



23. 9. 5. 오후 3:14 특허로

관인생략

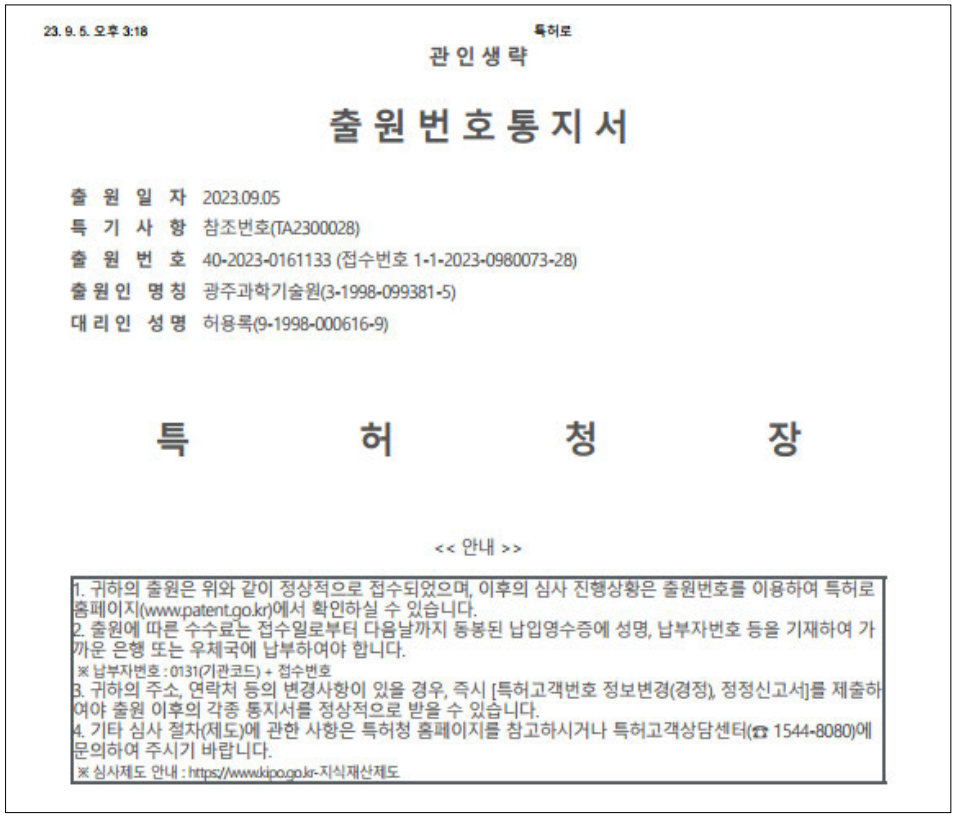
출원번호통지서

출 원 일 자 2023.09.05
특 기 사 항 참조번호(TA2300027)
출 원 번 호 40-2023-0161112 (접수번호 1-1-2023-0979993-83)
출원인 명칭 광주과학기술원(3-1998-099381-5)
대리인 성명 허용록(9-1998-000616-9)

특 허 청 장

<< 안내 >>

1. 귀하의 출원은 위와 같이 정상적으로 접수되었으며, 이후의 심사 진행상황은 출원번호를 이용하여 특허로 홈페이지(www.patent.go.kr)에서 확인하실 수 있습니다.
2. 출원에 따른 수수료는 접수일로부터 다음날까지 동봉된 납입영수증에 성명, 납부자번호 등을 기재하여 가까운 은행 또는 우체국에 납부하여야 합니다.
※ 납부자번호 : 0131(기관코드) + 접수번호
3. 귀하의 주소, 연락처 등의 변경사항이 있을 경우, 즉시 [특허고객번호 정보변경(경정), 정정신고서]를 제출하여야 출원 이후의 각종 통지서를 정상적으로 받을 수 있습니다.
4. 기타 심사 절차(제도)에 관한 사항은 특허청 홈페이지를 참고하시거나 특허고객상담센터(☎ 1544-8080)에 문의하여 주시기 바랍니다.
※ 심사제도 안내 : https://www.kipo.go.kr-지식재산제도



23. 9. 5. 오후 3:18 특허로

관인생략

출원번호통지서

출 원 일 자 2023.09.05
특 기 사 항 참조번호(TA2300028)
출 원 번 호 40-2023-0161133 (접수번호 1-1-2023-0980073-28)
출원인 명칭 광주과학기술원(3-1998-099381-5)
대리인 성명 허용록(9-1998-000616-9)

특 허 청 장

<< 안내 >>

1. 귀하의 출원은 위와 같이 정상적으로 접수되었으며, 이후의 심사 진행상황은 출원번호를 이용하여 특허로 홈페이지(www.patent.go.kr)에서 확인하실 수 있습니다.
2. 출원에 따른 수수료는 접수일로부터 다음날까지 동봉된 납입영수증에 성명, 납부자번호 등을 기재하여 가까운 은행 또는 우체국에 납부하여야 합니다.
※ 납부자번호 : 0131(기관코드) + 접수번호
3. 귀하의 주소, 연락처 등의 변경사항이 있을 경우, 즉시 [특허고객번호 정보변경(경정), 정정신고서]를 제출하여야 출원 이후의 각종 통지서를 정상적으로 받을 수 있습니다.
4. 기타 심사 절차(제도)에 관한 사항은 특허청 홈페이지를 참고하시거나 특허고객상담센터(☎ 1544-8080)에 문의하여 주시기 바랍니다.
※ 심사제도 안내 : https://www.kipo.go.kr-지식재산제도

3.2. 향후 추진 일정

향후 추진 방향 및 추진 계획

- 사용자 피드백을 반영한 서비스 개선을 위한 추가 개발
- 프로모션을 위해 필요한 개발(홍보 사이트 개발, 프로모션 어플 개발)
- (파트너십 확보) 기존 dApp 서비스 업체와 네트워킹, dApp 서비스를 준비/구상 중인 업체에 기술지원 및 사업 협력을 통해 파트너십 체결
- (유동성 공급자 유치) 안정적인 거래 환경 제공과 거래 활성화를 통한 유동성 공급자 참여 유도

사업내용	세부추진내용	12월				'24년 이후
		1주	2주	3주	4주	
지식재산권 출원	특허출원(1) 비영구적 손실 저감을 위한 온디멘드, 자체업로딩 (ODSU) 오라클 기반의 탈중앙거래소					
	특허출원(2)					
시제품 양산	유저 피드백 분석 및 개선 필요 사항 도출					
	유저 피드백을 반영하여 서비스 개선					
	제품 출시 및 서비스 공급제품 출시 및 서비스 공급					
마케팅	홍보 사이트 개발, 프로모션용 앱 개발					

감사합니다



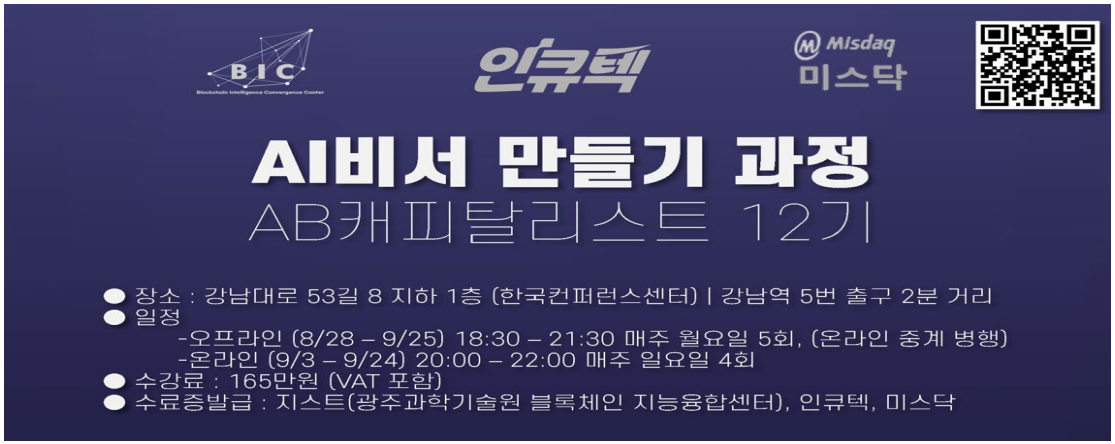
AI 역사와 GPT 모델

2023.08.28 | GIST/리버밴스㈜

우상향, 서울 강남구 강남대로 374 13층

TIPS 투자(22년 5월), R&D(22’ 6월), 해외사업(22’11월)

https://heungno.net



정규 과정 강의시간 : 월요일 18:30 – 21:30, 일요일 20:00 – 22:00

날짜	강의 주제	강사
AI의 역사와 GPT 모델 8월 28일(월)	AI : 어디까지 왔나 : 인공지능의 진화와 LLM의 등장	GIST 이흥노 교수
	챗GPT란 무엇인가 : LLM의 원리와 생성AI의 미래	
	입학식, 개강 파티 참가자 자기 소개 및 원우회 구성	
AI비서 만들기 환경 세팅 9월 3일(일)	Introduction : 챗GPT, 바드, 빙챗, Llama 등 소개 및 계정 생성	고려대학교 송인규 겸임교수
	오픈AI API 키 생성 및 구글 드라이브 사용법 기초	인큐텍 송경섭 실장
AI프롬프트 원리와 응용 9월 4일(월)	AI 어디까지 왔나 : 다양한 AI의 종류와 활용사례 개괄	POST AI 최규문 대표
	챗GPT 프롬프트 기초 사용법 : 기본 원리와 대화법	
	챗GPT4 개인 맞춤화(커스텀 인스트럭션) 설정 요령	

01. 리버밴스 대표

이흥노 교수

- GIST교수/인공지능연구소장
- 과기부 ITRC 블록체인 지능융합 센터장
- 대통령직속 정책기획위원회
- 블록체인경제센터장
- 센서지능화센터장
- (전)GIST연구원장
- 총 285편의 논문(국제 200여편, **SCI 100여편 포함**), 국내논문 104편,국내외 특허 60여건
- 해동학술상(2019)
- GIST연구상/대표기술상(3회)
- 이달의과학자상(2014.1)
- 국무총리표창 (2022.4)

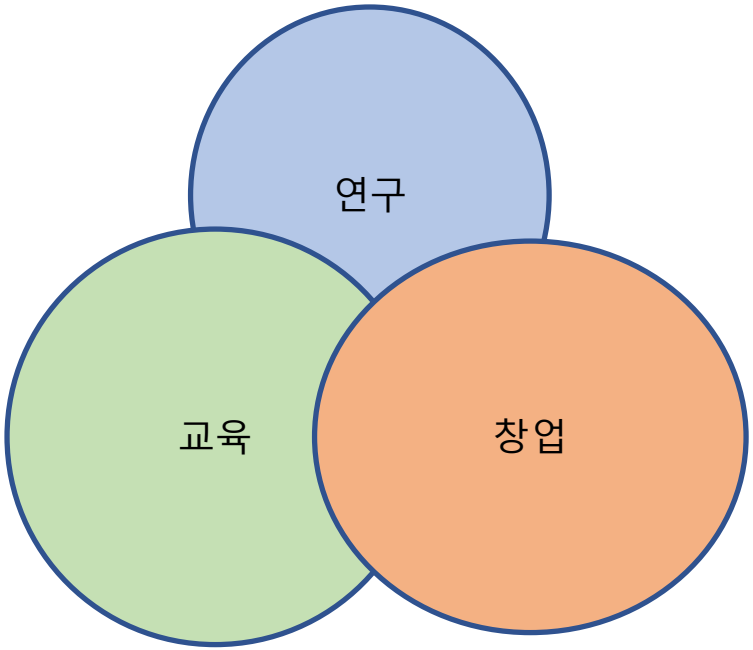
http://gist-itrc.org



2022년 4월 버전, 상세 이력 @ https://heungno.net

주요 이력

- IEEE Trans. Cybernetics 편집위원 (국제 Top 4%)
- ITRC 블록체인지능융합 센터장 (20’ – 28’)
- (전)GIST 인공지능연구소장
- GIST Infonet 연구실 지도교수
- (전) University of Pittsburgh, PA, USA, 전임 교수
- (전) Hughes Research Lab 연구원, Malibu, CA, USA.



지능

두뇌와 인류문명

45억년 전 지구의 탄생

20만년 전 Homo Sapiens(지혜로운 인간) 출현

BC 3,000년 문명의 시작

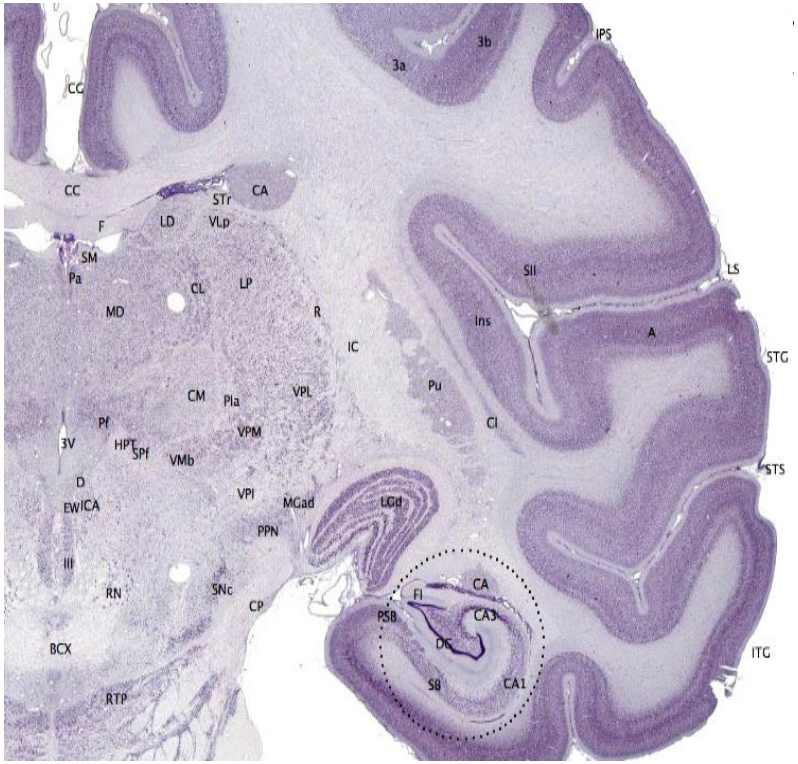
AD 1세기 지구상의 인구는 1억, 평균 수명 20세

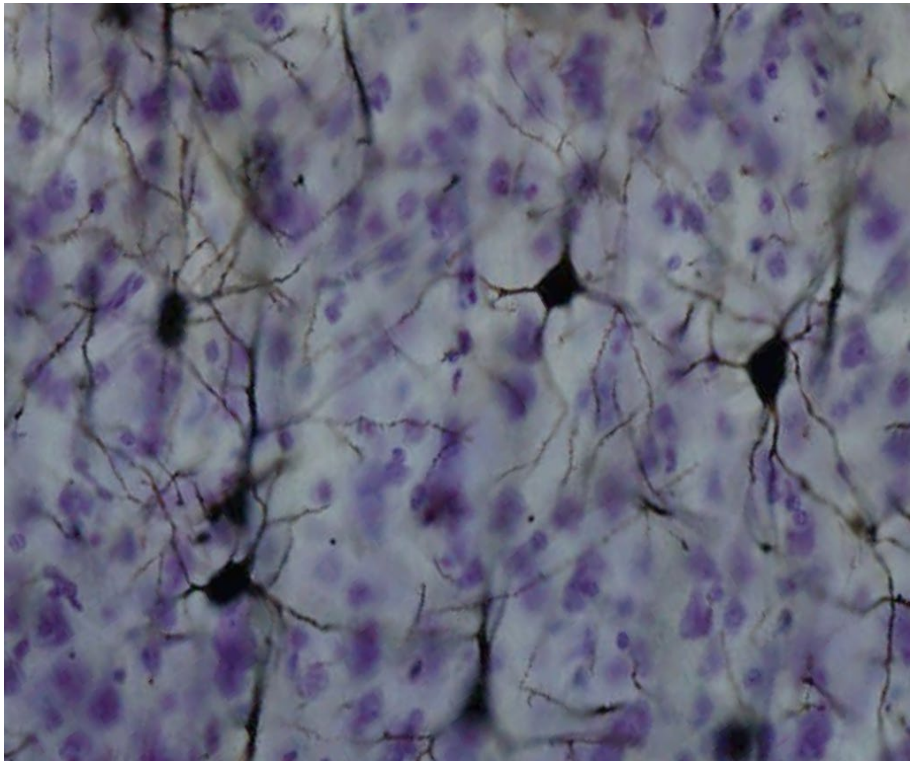
두뇌와 협력으로 이루어낸 기적!



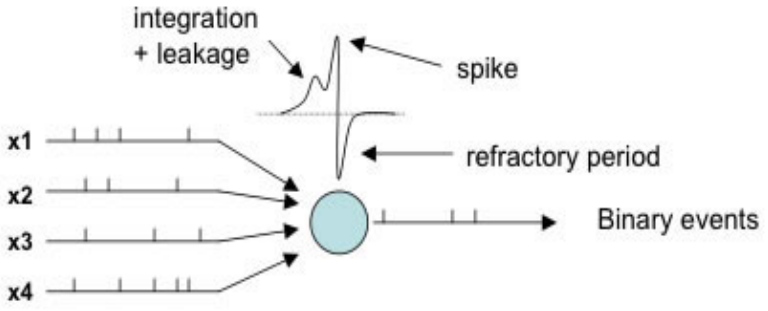


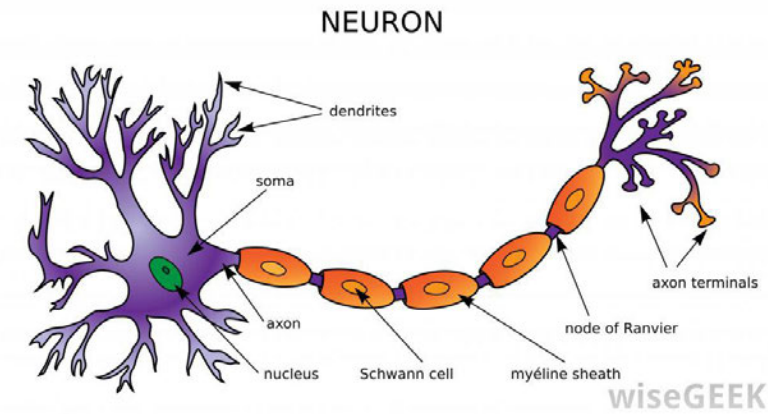
인간의 두뇌 – network of cells

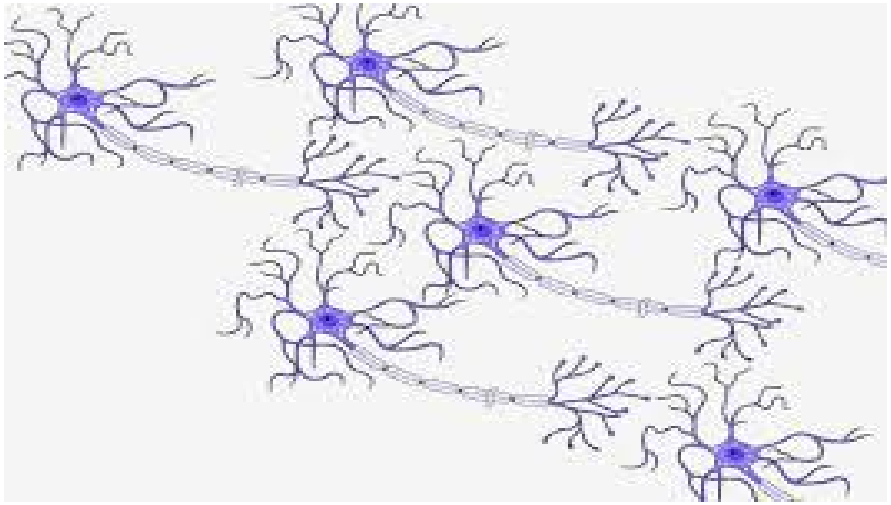




Neuron : 입출력 장치







02. 두뇌

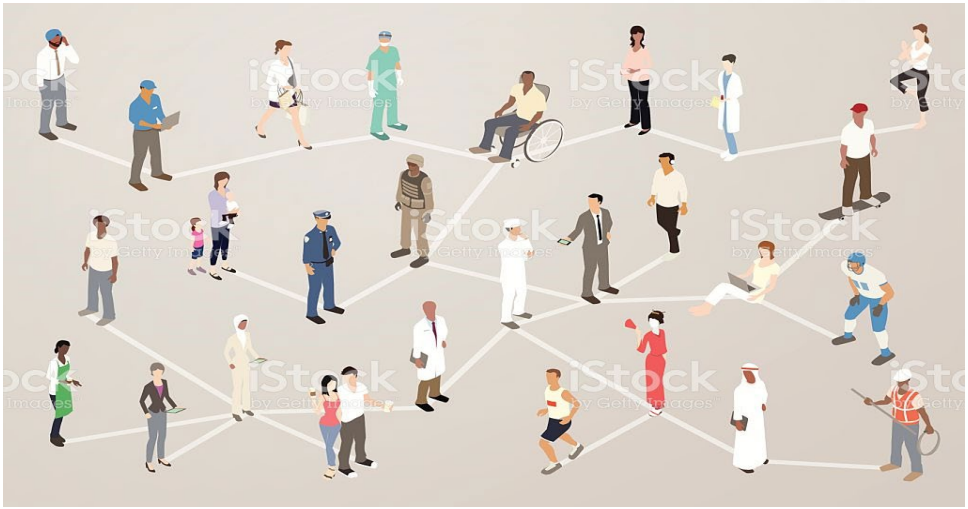
두뇌와 두뇌는 연결된 사회속에서 성장한다.

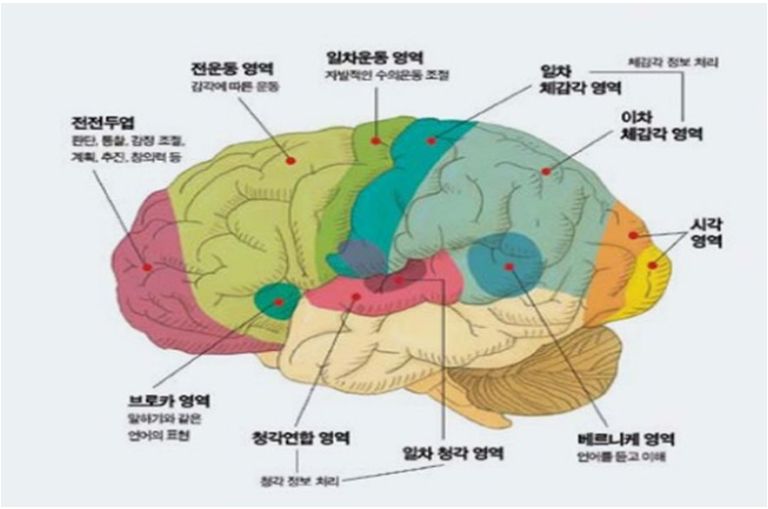
사람의 지능은 두뇌에 있다.

사회에 속해있다.

사람은 생명체다.



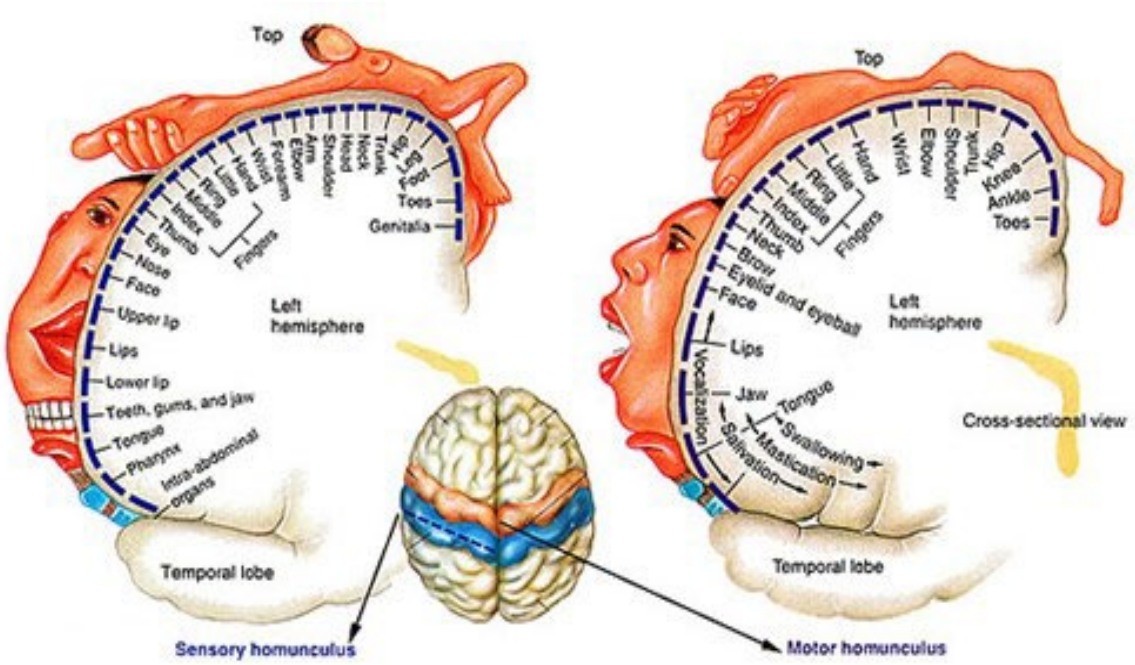




Sensing 과 Action

손가락, 손바닥, 혀, 입술 등에 뇌의 영역이 집중된 것을 시각화한 그림





02. 두뇌

우리의 두뇌는 사회 활동의 결과물이다.

Sensors
신호습득

보고, 듣고, 냄새맡고,
만져보고, 느끼고

부모에게 물어보고

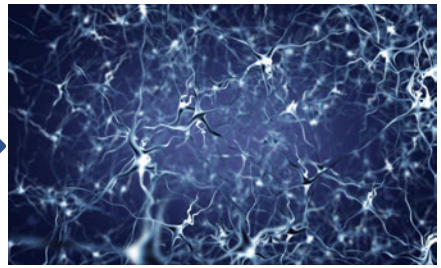
선생님과 대화해 보고

책을 읽고, TV를 보고

SNS 등을 통한

센싱을 통해

Situational Awareness
(신호 + 경험)





Actions
결정

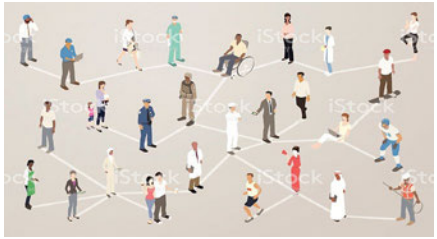
이동하는 것
먹는 것
입는 것
공부하는 것
노는 것

정치/경제/사회

협력/공유/신뢰/책임

결정한다

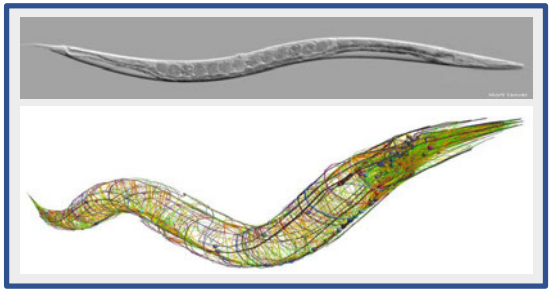




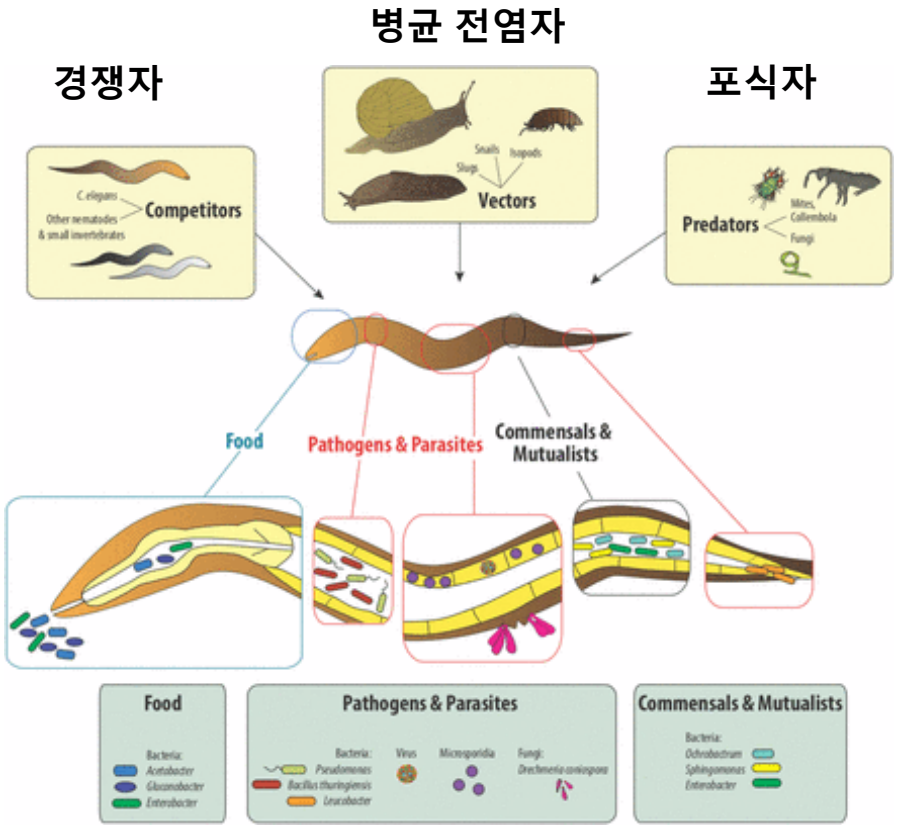


02. 두뇌

꼬마선충(C elegans), 302개의 뉴론으로 늪지에서 산다.







02. 두뇌

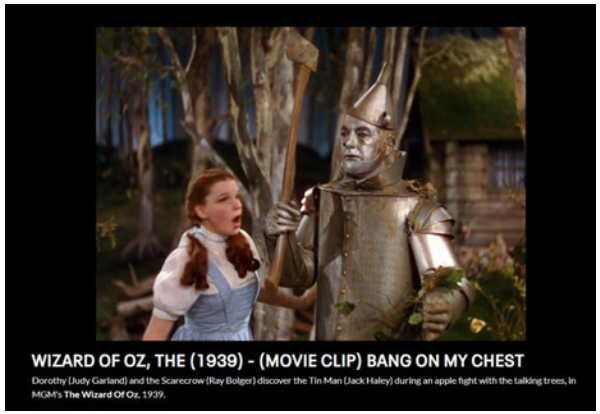
지능이란?

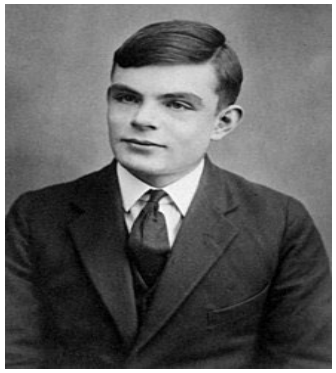
- **인간(뇌)은 센싱을 통해, 상황을 판단하고, 행동의사를 결정한다.**
- **세상은 인간의 결정에 반응한다.**
- **인간은 세상의 반응을 통하여 학습하고 의사결정능력을 고도화 한다.**
- **시의 적절하게 *잘 내릴 수 있는* 의사결정능력을 지능이라 한다.**

인공지능

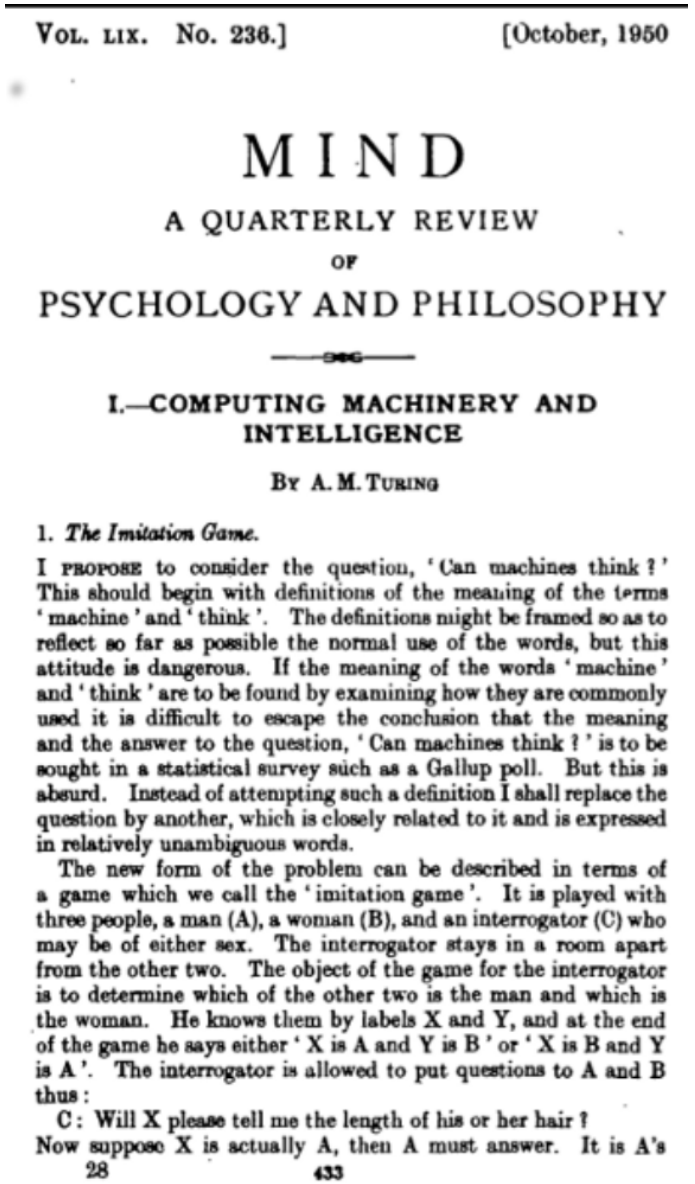
03. 인공지능

Start: Can machines think?





- **1950**: Alan Turing proposes the Turing Test as a measure of machine intelligence in his paper "Computing Machinery and Intelligence."



VOL. LIX. No. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the 'imitation game'. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either 'X is A and Y is B' or 'X is B and Y is A'. The interrogator is allowed to put questions to A and B thus:

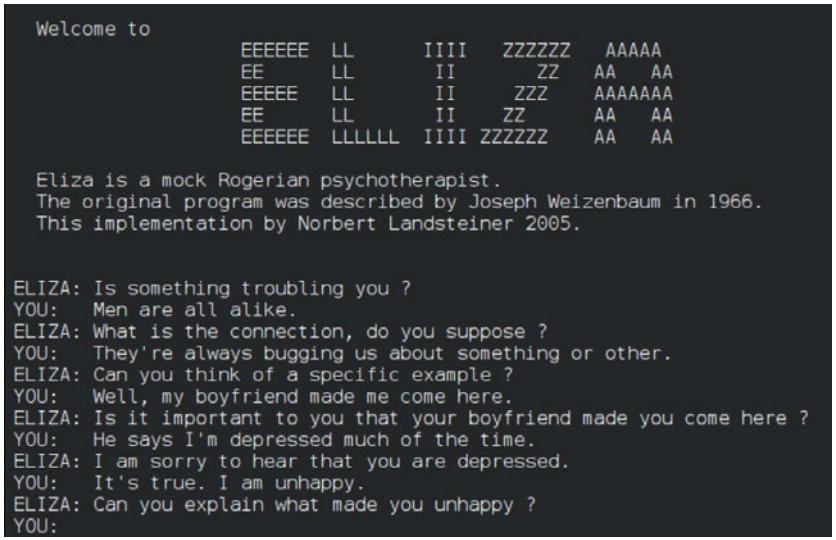
C: Will X please tell me the length of his or her hair?

Now suppose X is actually A, then A must answer. It is A's

28 433

ELIZA

- The 1960s saw the development of early AI programs such as:
- **ELIZA**: Created by Joseph Weizenbaum, simulates a Rogerian psychotherapist.



- 로제이안 심리치료: 환자 자아 존중, 환자 스스로 자신의 문제에 해답을 찾도록 도와주는 방법론

SHRDLU 1968

- **SHRDLU**: Developed by Terry Winograd, demonstrates natural language understanding in a blocks world.

Person: Pick up a big red block.
Computer: OK.
Person: Grasp the pyramid.
Computer: I DON'T UNDERSTAND WHICH PYRAMID YOU MEAN.
Person (changing their mind): Find a block which is taller than the one you are holding and put it into the box.
Computer: BY "IT", I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER THAN THE ONE I AM HOLDING.
Computer: OK.
Person: What does the box contain?
Computer: THE BLUE PYRAMID AND THE BLUE BLOCK.
Person: What is the pyramid supported by?
Computer: THE BOX.
Person: How many blocks are not in the box?
Computer: FOUR OF THEM.

Lighthill Report

- **Early 1970s**: The limitations of early AI systems become apparent, leading to reduced funding and the first "AI winter.“
- **1973**: The "Lighthill Report" in the UK criticizes the lack of practical progress in AI, leading to reduced funding.

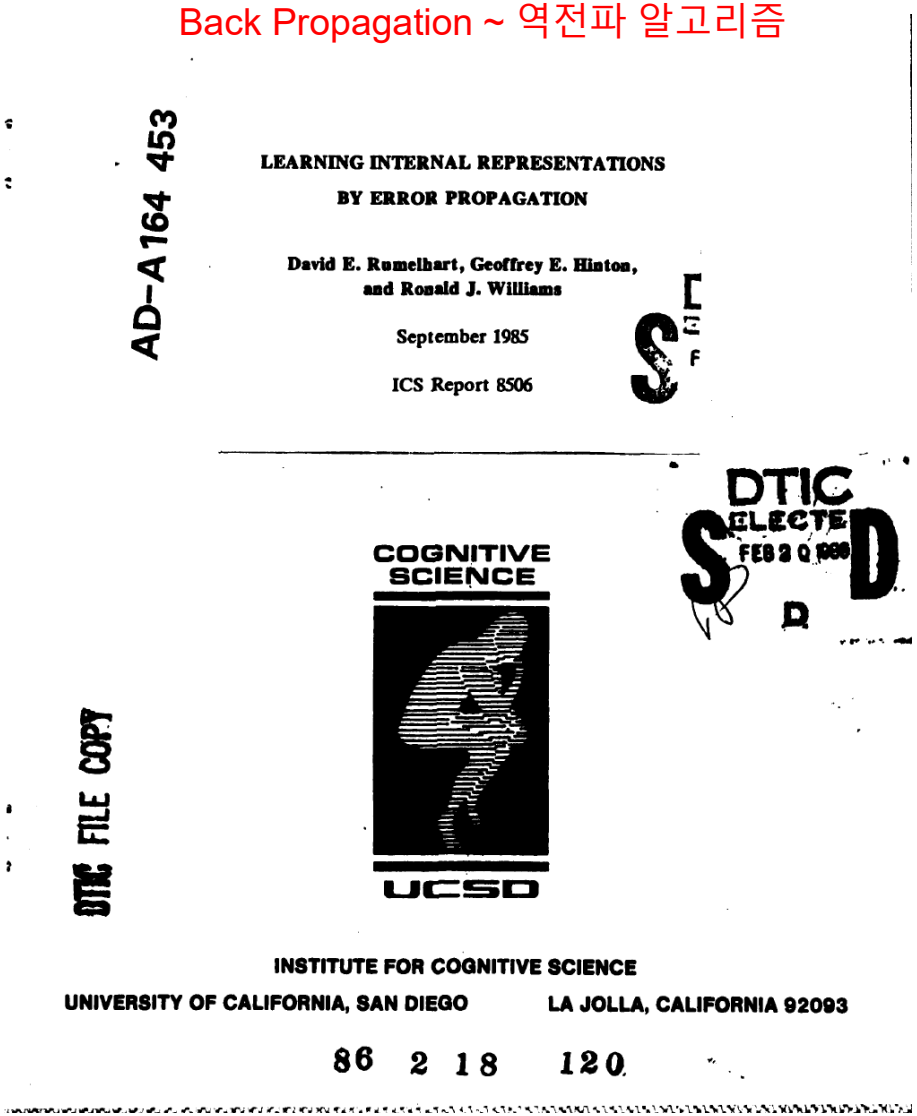
03. 인공지능

80s/90s



https://en.wikipedia.org/wiki/Expert_system

Back Propagation ~ 역전파 알고리즘



AD-A164 453

LEARNING INTERNAL REPRESENTATIONS
BY ERROR PROPAGATION

David E. Rumelhart, Geoffrey E. Hinton,
and Ronald J. Williams

September 1985

ICS Report 8506

DTIC
ELECTE
FEB 2 0 1986
D

COGNITIVE
SCIENCE

UCSD

DTIC FILE COPY

INSTITUTE FOR COGNITIVE SCIENCE
UNIVERSITY OF CALIFORNIA, SAN DIEGO LA JOLLA, CALIFORNIA 92093

86 2 18 120

2000s

- **2006**: Geoffrey Hinton ~ “deep learning” to fix the “vanishing gradient” problem.



A fast learning algorithm for deep belief nets *

Geoffrey E. Hinton and **Simon Osindero**
Department of Computer Science University of Toronto
10 Kings College Road
Toronto, Canada M5S 3G4
{hinton, osindero}@cs.toronto.edu

Yee-Whye Teh
Department of Computer Science
National University of Singapore
3 Science Drive 3, Singapore, 117543
tehyw@comp.nus.edu.sg

Abstract

We show how to use “complementary priors” to eliminate the explaining away effects that make inference difficult in densely-connected belief nets that have many hidden layers. Using complementary priors, we derive a fast, greedy algorithm that can learn deep, directed belief networks one layer at a time, provided the top two layers form an undirected associative memory. The fast, greedy algorithm is used to initialize a slower learning procedure that fine-tunes the weights using a contrastive version of the wake-sleep algorithm. After fine-tuning, a network with three hidden layers forms a very good generative model of the joint distribution of handwritten digit images and their labels. This generative model gives better digit classification than the best discriminative learning algorithms. The low-dimensional manifolds on which the digits lie are modelled by long ravines in the free-energy landscape of the top-level associative memory and it is easy to explore these ravines by using the directed connections to display what the associative memory has in mind.

1 Introduction

Learning is difficult in densely-connected, directed belief nets that have many hidden layers because it is difficult to infer the conditional distribution of the hidden activities when given a data vector. Variational methods use simple approximations to the true conditional distribution, but the approximations may be poor, especially at the deepest hidden layer where the prior assumes independence. Also, variational learning still requires all of the parameters to be learned together and makes the learning time scale poorly as the number of parameters increases.

We describe a model in which the top two hidden layers form an *undirected* associative memory (see figure 1) and the remaining hidden layers form a directed acyclic graph that converts the representations in the associative memory into observable variables such as the pixels of an image. This hybrid model has some attractive features:

1. There is a fast, greedy learning algorithm that can find a fairly good set of parameters quickly, even in deep networks with millions of parameters and many hidden layers.
2. The learning algorithm is unsupervised but can be applied to labeled data by learning a model that generates both the label and the data.
3. There is a fine-tuning algorithm that learns an excellent generative model which outperforms discriminative methods on the MNIST database of hand-written digits.
4. The generative model makes it easy to interpret the distributed representations in the deep hidden layers.
5. The inference required for forming a percept is both fast and accurate.
6. The learning algorithm is local: adjustments to a synapse strength depend only on the states of the presynaptic and post-synaptic neuron.
7. The communication is simple: neurons only need to communicate their stochastic binary states.

Section 2 introduces the idea of a “complementary” prior which exactly cancels the “explaining away” phenomenon that makes inference difficult in directed models. An example of a directed belief network with complementary priors is presented. Section 3 shows the equivalence between restricted Boltzmann machines and infinite directed networks with tied weights.

Section 4 introduces a fast, greedy learning algorithm for constructing multi-layer directed networks one layer at a time. Using a variational bound it shows that as each new layer is added, the overall generative model improves. The greedy algorithm bears some resemblance to boosting in its repeated use of the same “weak” learner, but instead of reweighting each data-vector to ensure that the next step learns something new, it re-represents it. The “weak” learner that

*To appear in Neural Computation 2006

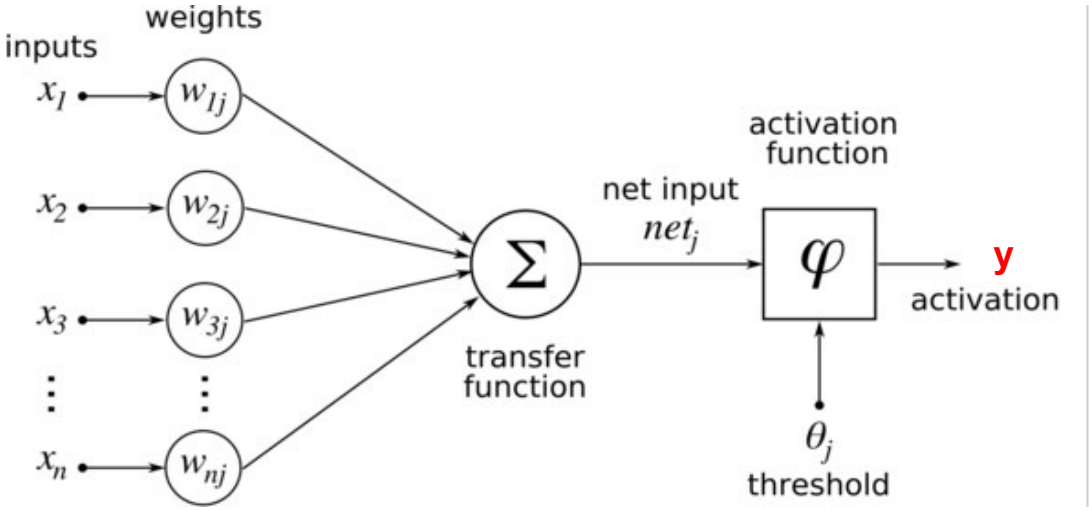
03. 인공지능

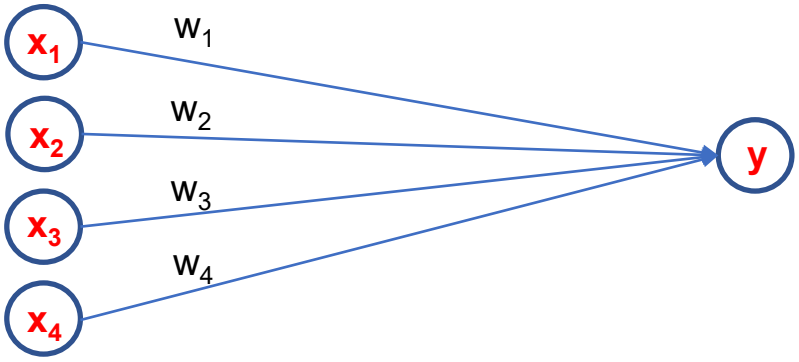
2010s:

- **2011**: IBM's Watson wins the quiz show Jeopardy!, defeating human champions.
- **2012**: AlexNet, a deep neural network, achieves a breakthrough in image recognition accuracy in the ImageNet competition.
- **2014**: Chatbot Eugene Goostman reportedly passes the Turing Test, though this claim is widely debated.
- **2015**: Google DeepMind's AlphaGo defeats a human professional Go player.
- **2018**: OpenAI's GPT-2 shows that large language models can generate coherent and contextually appropriate text.

뉴론 Neuron

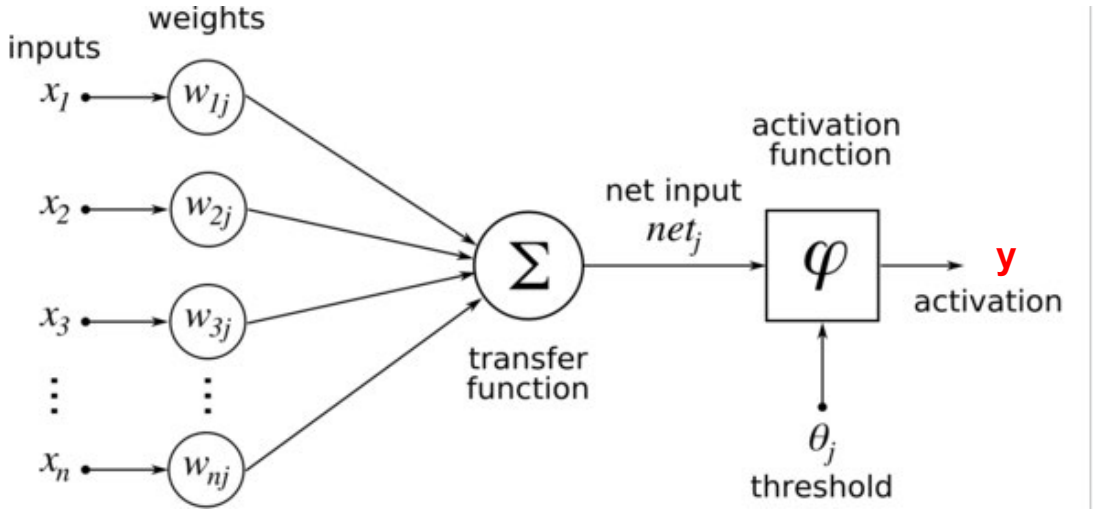
- 입력 x
  - 가중치 w
  - 편향
  - 합계
  - 활성함수
- 출력 y



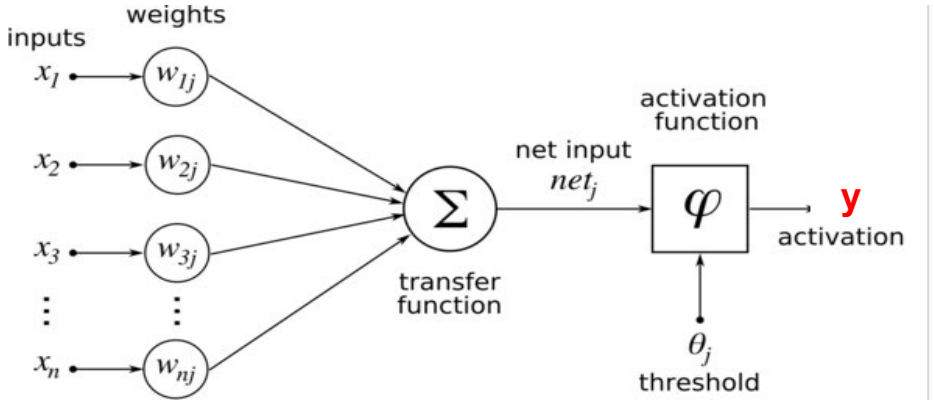


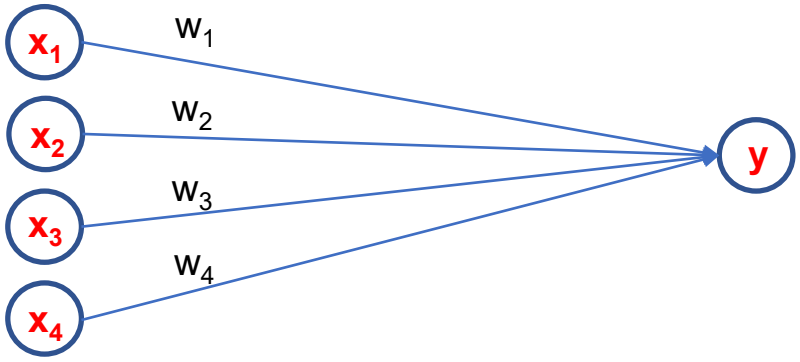
학습 Training

- 랜덤 변수로 시작
- 에러 e = y – y_d
- e 최소화 변수 찾기
  - Backpropagation
- 다수의 개와 고양이 사진 제시

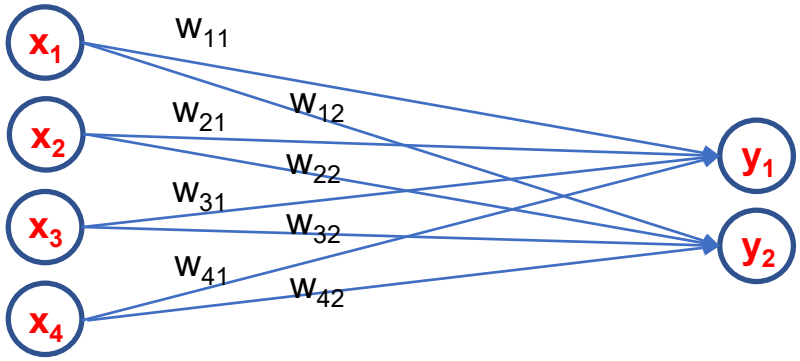


간단한 표현법


weights
inputs
x_1
x_2
x_3
x_n
w_{1j}
w_{2j}
w_{3j}
w_{nj}
Σ
transfer function
net input
net_j
activation function
φ
y
activation
θ_j
threshold

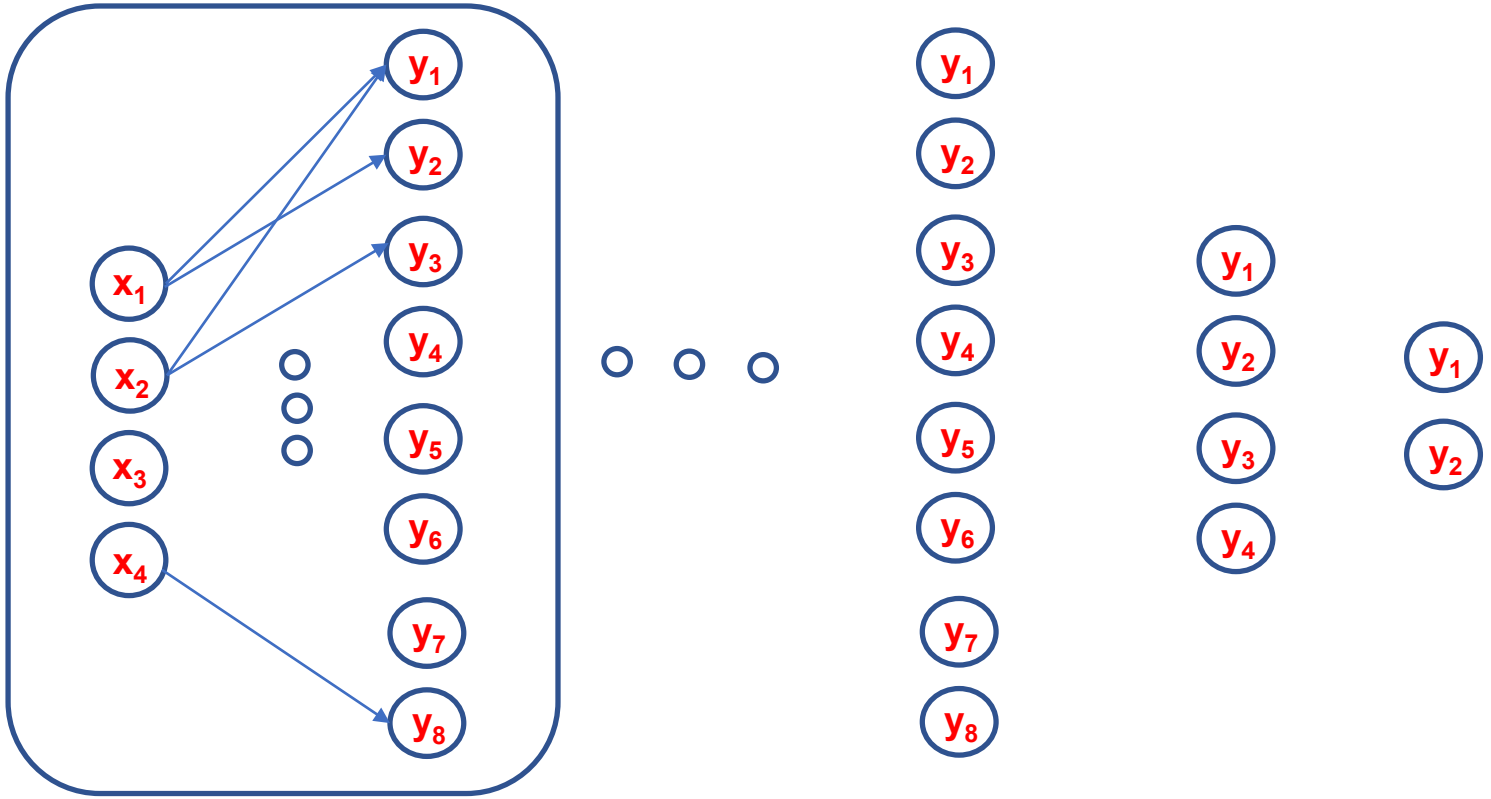

x_1
x_2
x_3
x_4
w_1
w_2
w_3
w_4
y

간단한 표현법 – 2 outputs



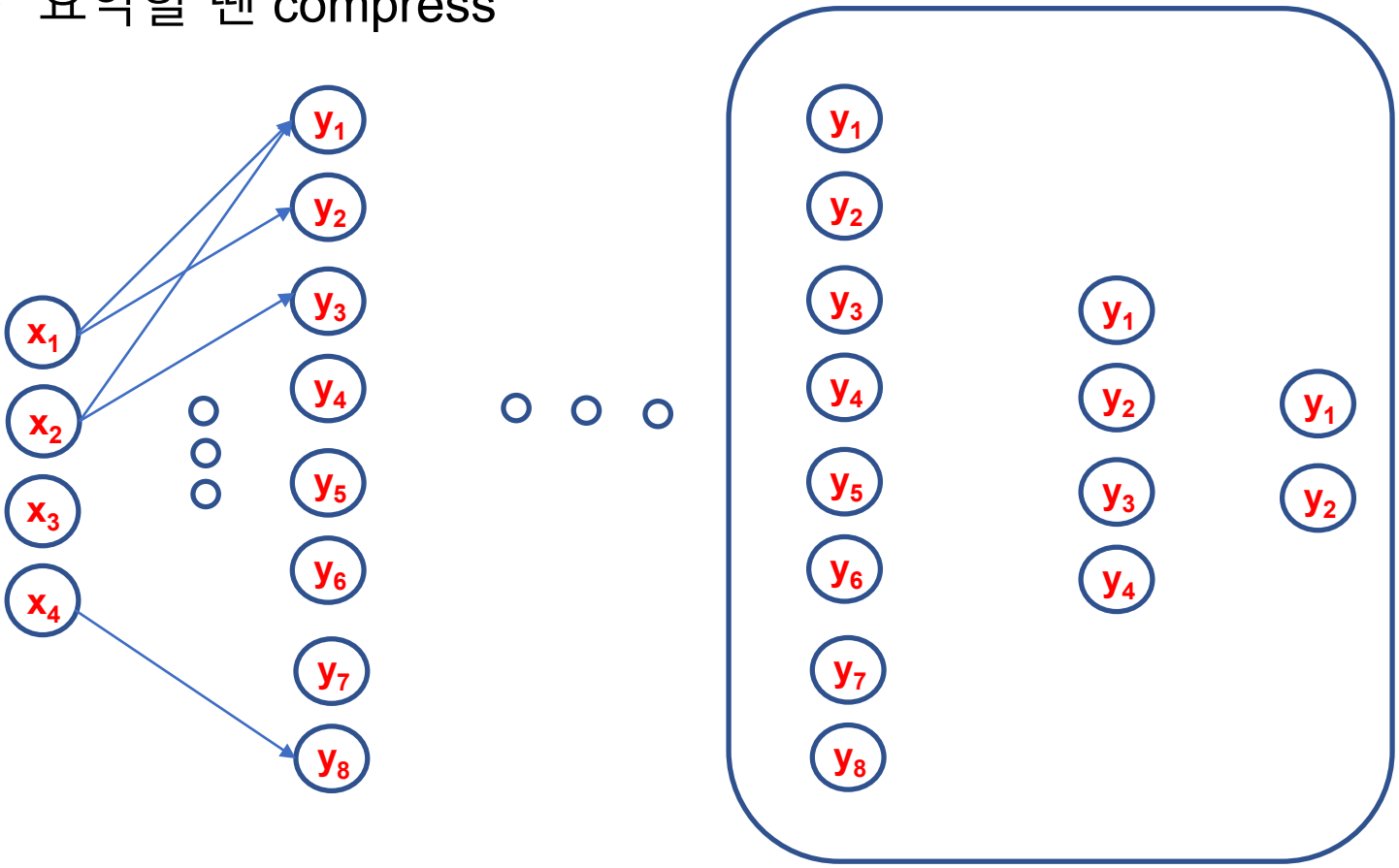
분석과 요약

- 분석할 땐 expand

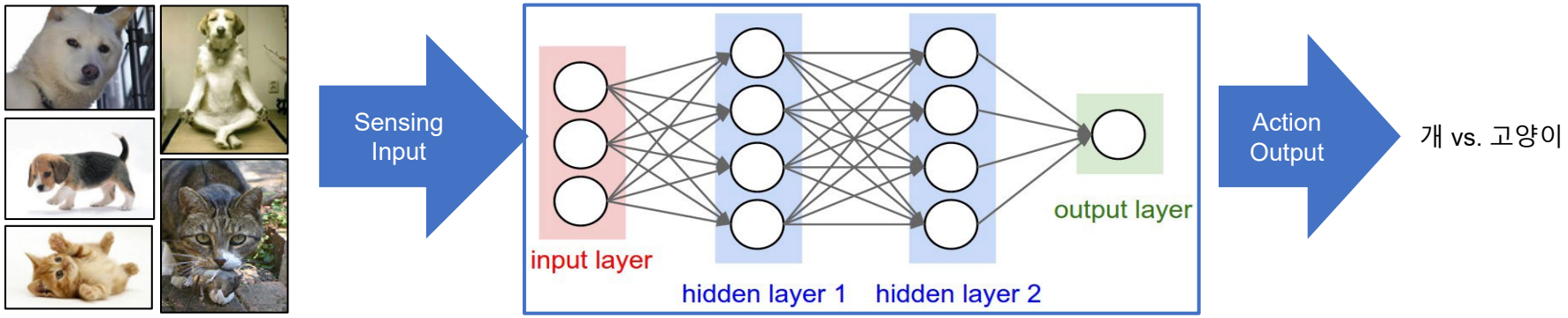


분석과 요약

- 요약할 땐 compress



심층신경망 훈련은 데이터와 GPU인프라가 필요하다.



대규모 데이터 확보, 딥러닝 모델을 학습시키고, 학습 완료된 모델을 사용하는것

학습 : 원하는 IN-OUT 관계를 위해 네트워크 가중치와 매개 변수를 조정하는것

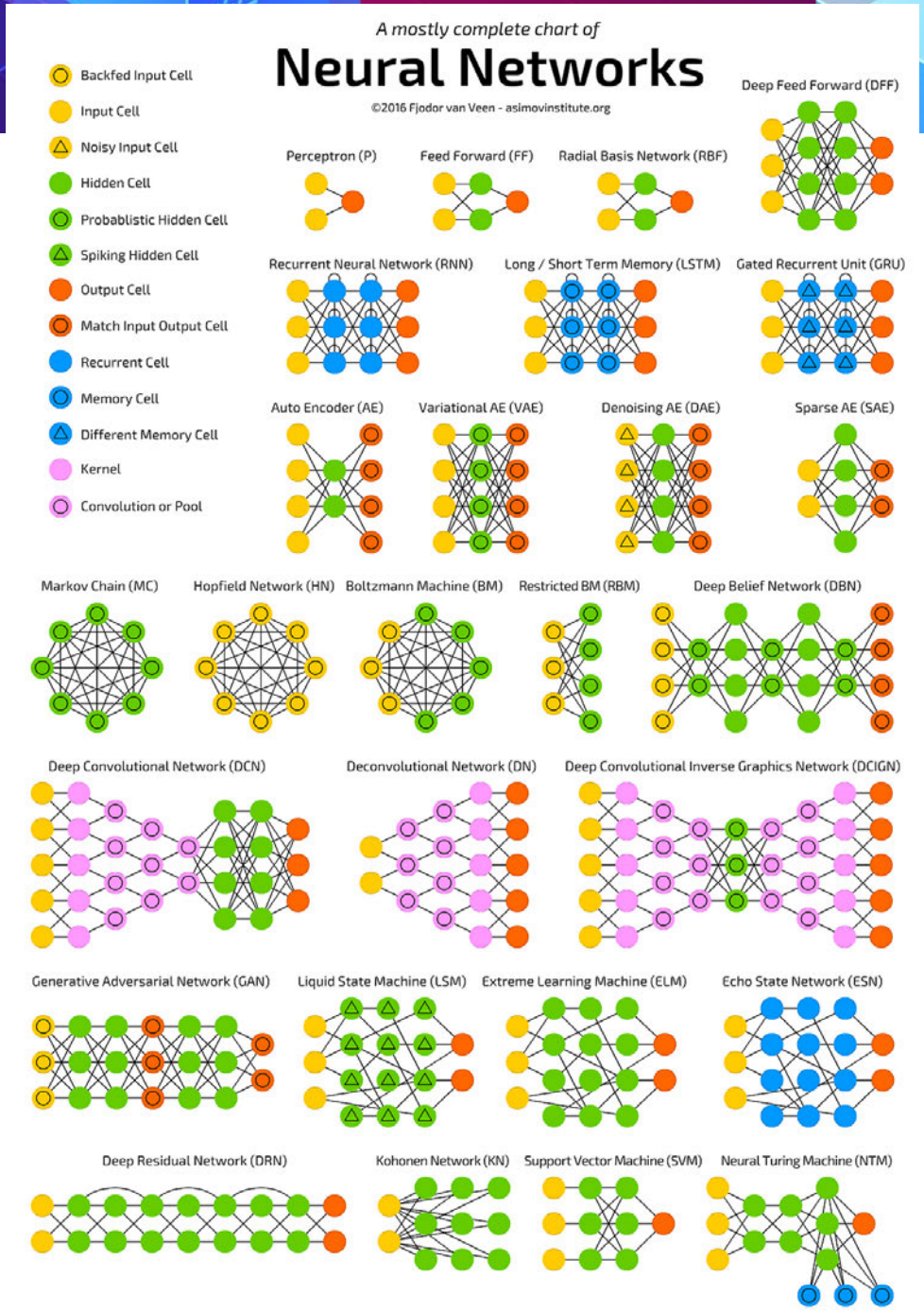
추론 : 학습완료된 뉴럴넷을 새로운 인풋에 대해 예측, 분류, 생성 등 판단에 사용

주요 기술: 심층신경망 (deep neural network), 역전파 (back propagation),
확률경사하강법 stochastic gradient descent (SDG)

03. 인공지능

Deep Neural Networks 의 종류

20 ~ 30여 개의 구조



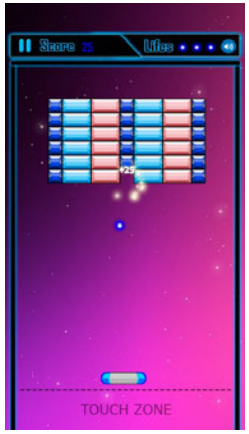
Machine Learning Methodologies

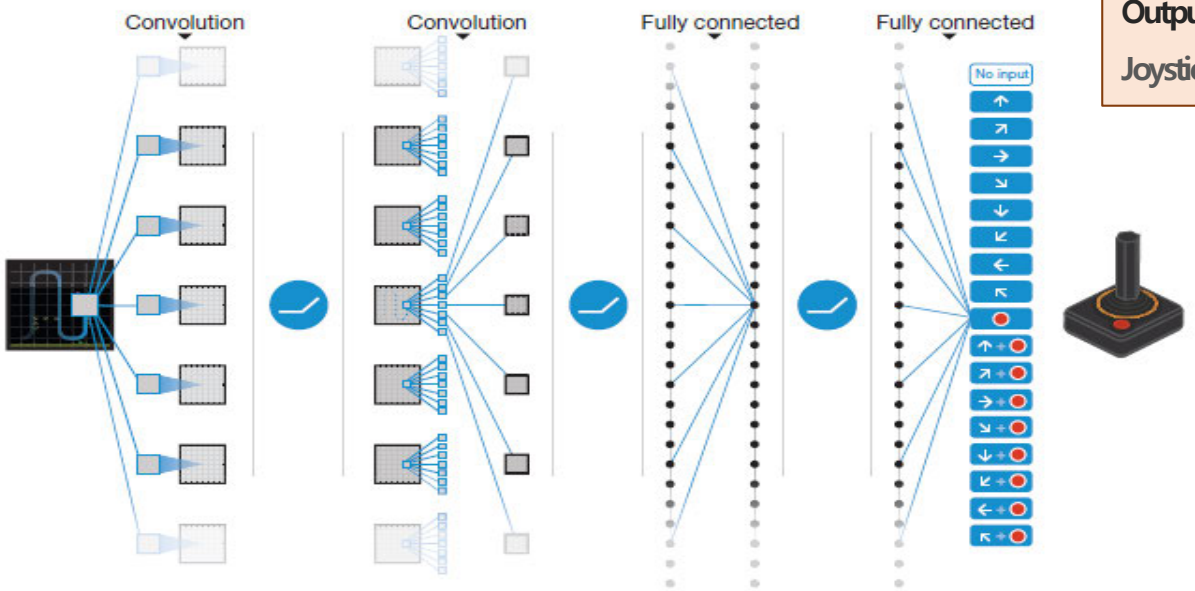
- Unsupervised learning
- Supervised learning
- Reinforcement learning
- Multi Agent Reinforcement Learning

AlphaGo의 Deep Neural Network

Sensing 상황인식 **Action**

Input : Screen



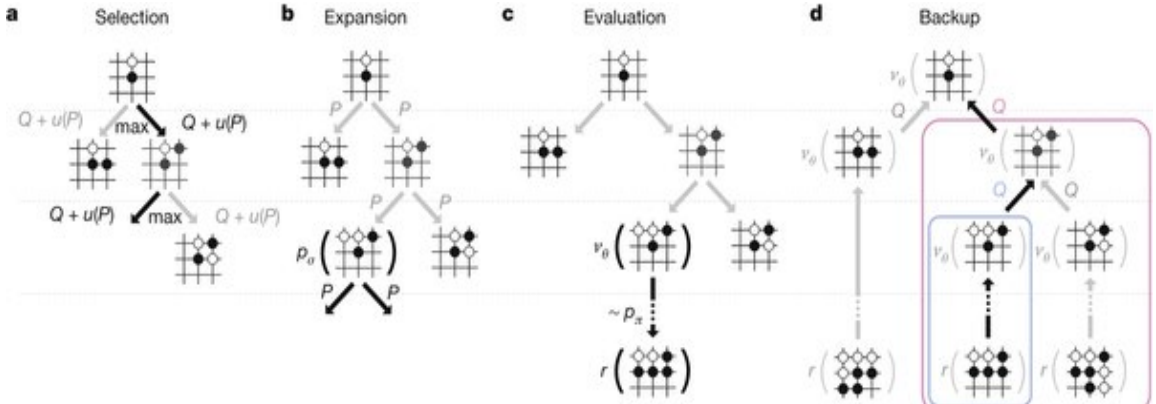


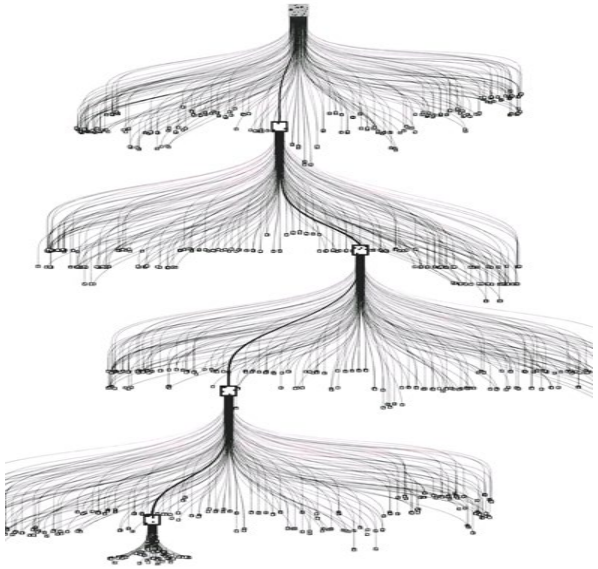
Output :
Joystick Action

인공지능

AlphaGo: Tree Search

Cloud Computers를 활용해 넓고, 깊게, 게임을 이기는 수를 찾는 바둑 알고리즘





03. 인공지능

QUIZ

이 컴퓨터의 이름은?



IBM에서 개발

자연어 형식으로 된 질문들에 답할 수 있는 인공지능 컴퓨터 시스템

'11년 미국 ABC방송 Jeopardy 퀴즈쇼에서 우승

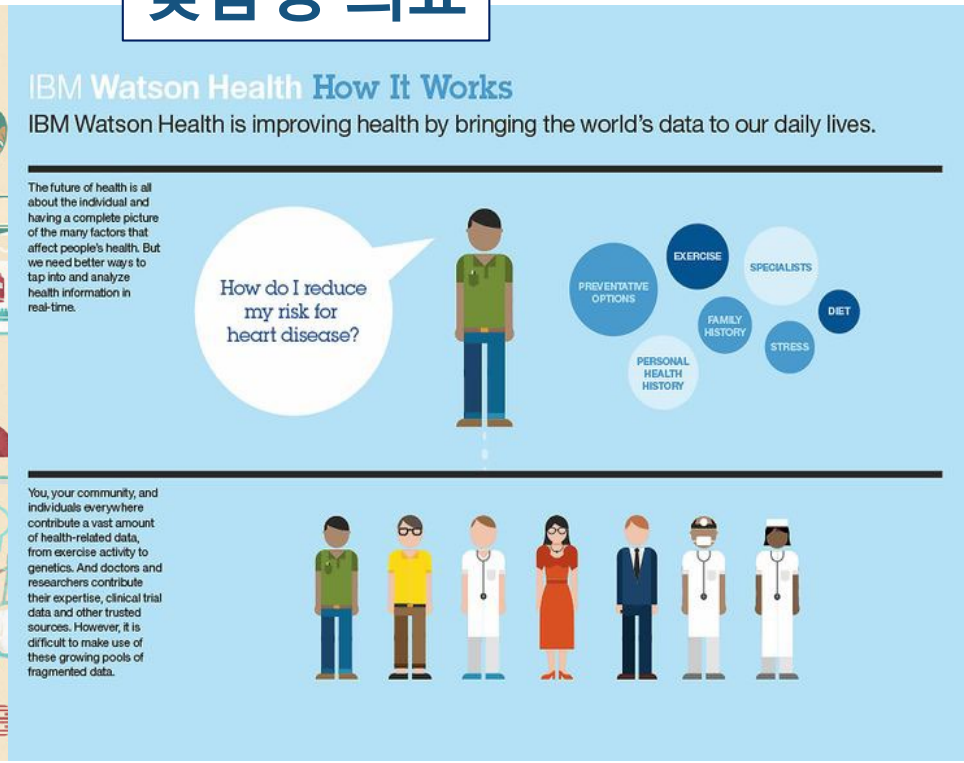
03. 인공지능

IBM Watson Capabilities

맞춤형 교육



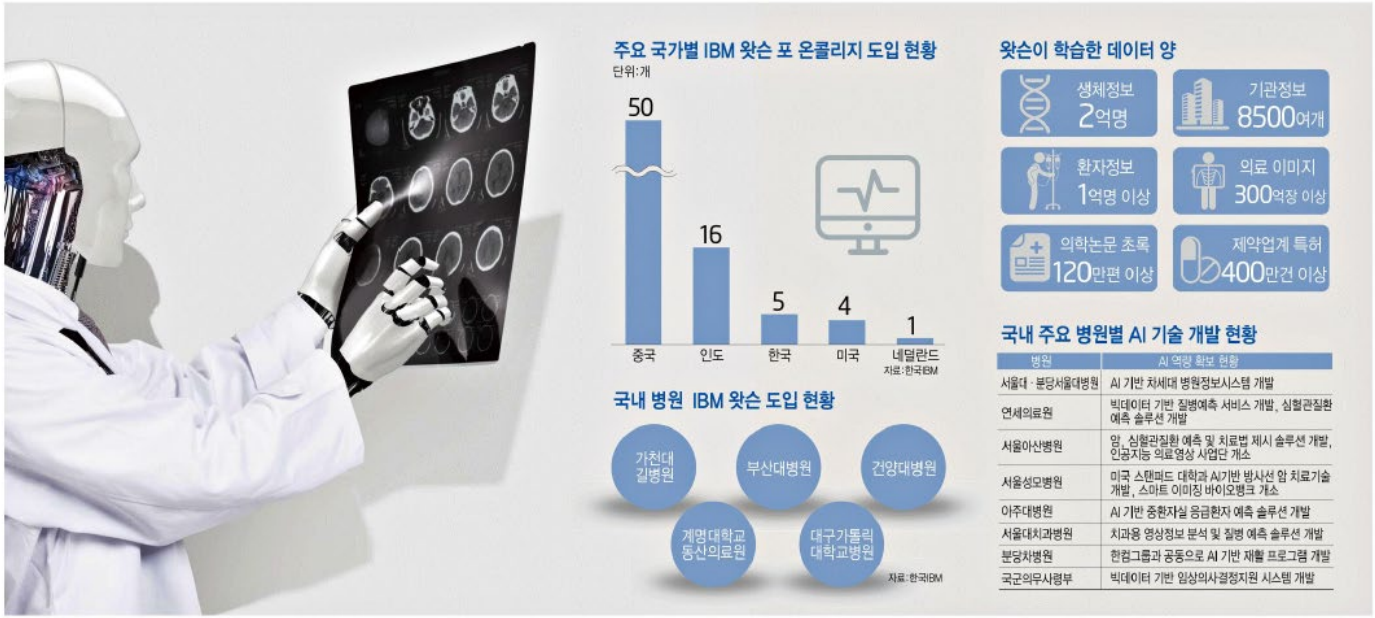
맞춤형 의료



2017년 4월 11일자 전자신문

전자신문

2017년 04월 11일 화요일 004면 종합



AI 주치의 등장…의료혁신 脈 제대로 짚을까

◇세계 톱3 'AI 도입국', 의료 혁신 시작됐다

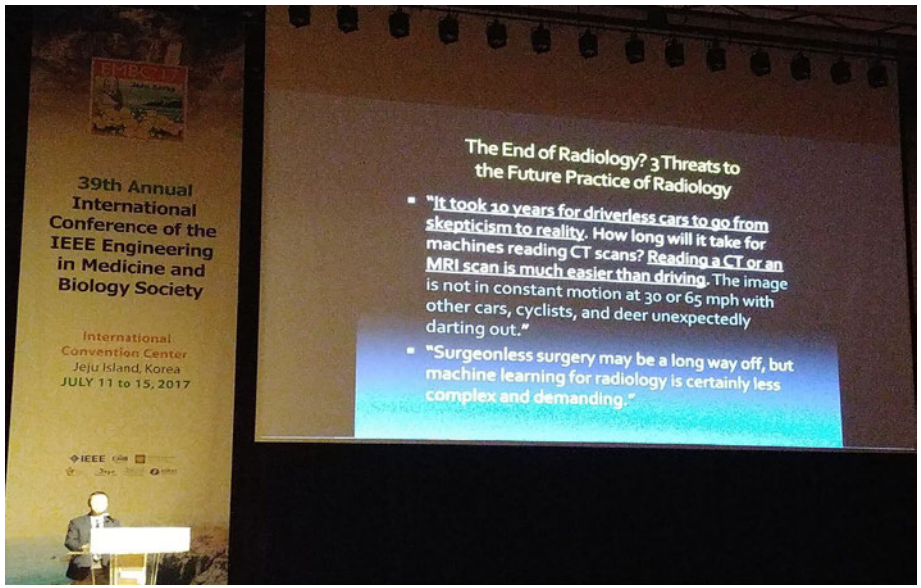
알파고가 불러온 AI 신드롬은 병원에서 IBM '왓슨'이 바통을 이어 받았다. 코그니티브(인지) 컴퓨팅 솔루션 왓슨은 빅데이터를 분석해서 자연어로 된 질문을 이해하고 답을 제시한다. 매일 쏟아지는 300여종의 의학저널, 200여종의 의학 교과서, 1500만쪽에 달하는 의료 정보를 학습해서 최적의 치료법을 제시한다. 암 진단·치료에 도움을 주는 '온콜로지' △유전자 분석에 초점을 맞춘 '지노믹스' △임상 시험을 돕는 '클리니컬 트라이얼 매칭' △연구개발(R&D)용 '라이프 사이언스' 등이 대표 솔루션이다.

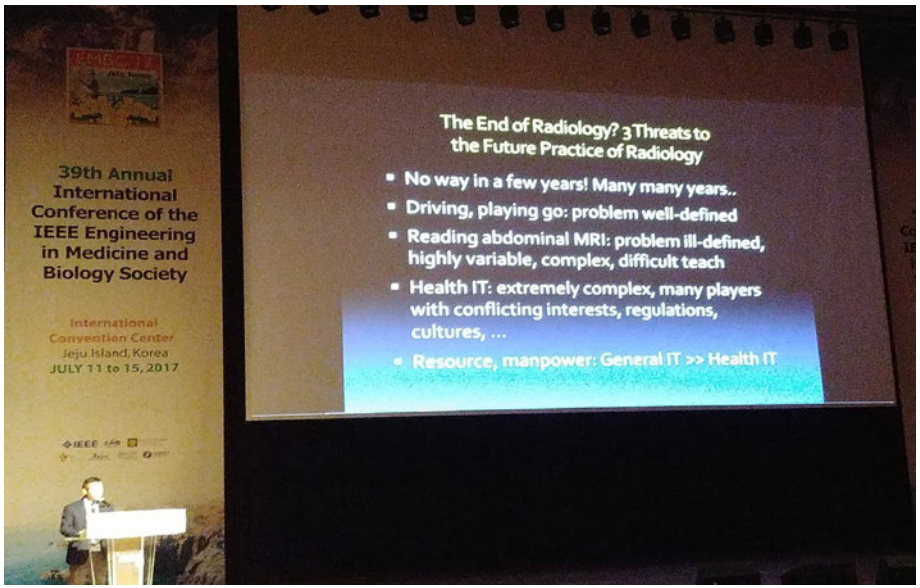
2015년 국내에 첫선을 보인 왓슨 포 온콜로지는 지난해 9월 가천대 길병원을 시작으로 부산대병원, 건양대병원, 계명대 동산의료원, 대구 가톨릭대학병원, 중앙보훈병원 등 6개 병원이 도입했거나 도입할 예정이다. 세계 각국과 비교해서 도입 비율이 높다. 왓슨 포 온콜리지를 도입한 병원은 중국이 50곳으로 가장 많다. 인도가 마니팔 병원 그룹 내 16곳이 도입해 뒤를 이었다. 우리나라(5곳)는 3위다. 미국이 4개 병원, 태국·네팔·네덜란드가 각 1곳이다.

03. 인공지능

A Radiologist I met at IEEE EMBS 2017

The end of Radiology?
He says No!
Why?
Because the job is very complex for AI to take over.
He says there are lots of hypes about AI.





Large Language Models

심층신경망 자연어 처리 NLP 발전 주요 이력

	paper	비고
2013.01	Word2vec: “Efficient estimation of word representations in vector space,” Google **단어를 고차원 벡터로 표현. 대규모데이터로 학습. 유사한 연관 단어 벡터는 서로 가깝게 표현.**	BoW: 연관단어 → 타겟단어 Skip-gram: 타겟단어 → 연관단어
2014.02	“Glove: Global vectors for word representation,” by Pennington et.al., Stanford University	전역적인 count-base 통계 모델 및 예측 모델 활용
2016.07	“Enriching word vectors with subword information,” Bojanowski et. al. 단어를 하위어 단어로 분해하고 임베딩, 형태론적 변형과 유사성 포착	단어를 Subword로 쪼개어 형태론적 변형 포착
2017.06	Transformer. “*Attention is all you need*,” Vaswani et.al. **주의집중 메카니즘 활용, 다중 계층 인코더로 문맥 파악, 다중 계층 디코더로 문맥화된 단어 생성**	**주의 집중 메카니즘**
2018.02	Elmo. “Deep contextualized word representations,” Peters et.al. 양방향 LSTM사용, 좌우 문맥을 고려, 각 단어의 문맥화된 표현 생성	양방향 LSTM문맥
2018.10	BERT: bidirectional encoder representation from Transformers, Devlin et.al. Transformer를 활용 단어 문맥 양방향 고려, 단어임베딩	양방향 Transformer문맥

Transformer

- Attention is all you need! (NIPS2017)
  - GER-ENG 번역기
  - Input 문장을 받아 Output 문장을 생성
  - IN 문장에서 단어들간의 관계 및 주목할 단어 파악, OUT에서 한 단어씩 생성
- 수십 층의 심층 신경망
  - 학습 매개 변수

$$O(layers \times d_{\text{model}}{}^2) \sim \text{ 10s million}$$

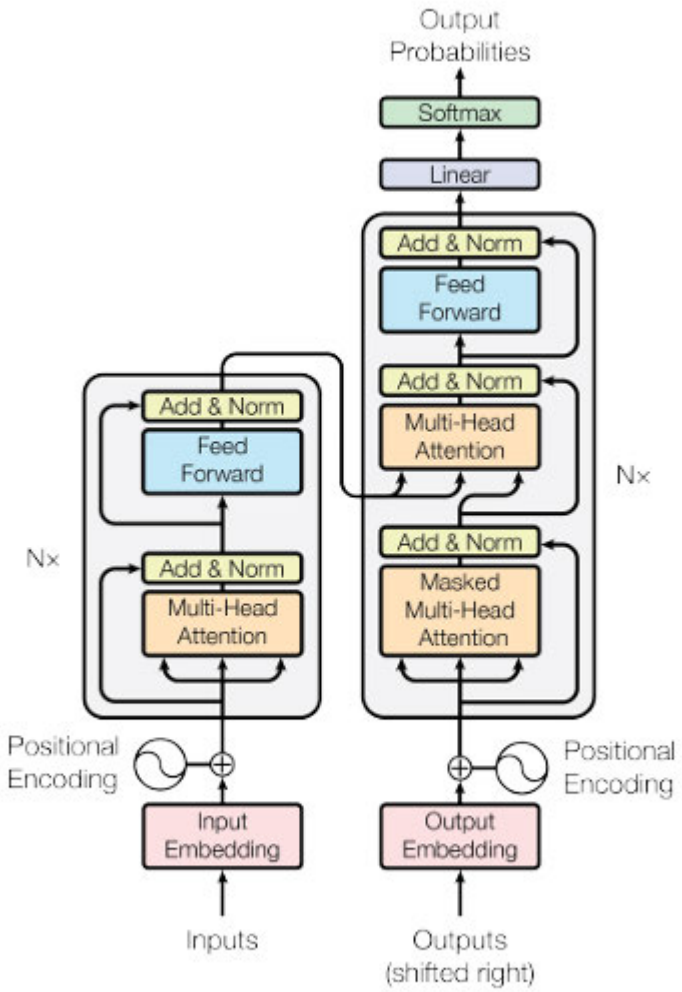


Figure 1: The Transformer - model architecture.

04. Large Language Models

Transformer는 언어 변환 모델로 시작하였다.



04. Large Language Models

OpenAI는 Transformer를 맥락 인지 문장 생성기로 발전시킴.

- GPT-3, July 2020
- GPT-3.5, March 2022
  - 1750억 개 매개변수
  - 800 GB storage
- Emergent Abilities:
  - Few-shot learning
  - Comm. language
- Computer codes, poetry, prose, "Terrifying ability," NYT 2020.

Table 2.2: Datasets used to train GPT-3.

Dataset	Quantity (tokens)	Weight in training mix	Epochs elapsed when training for 300B tokens
Common Crawl (filtered)	410 billion	60%	0.44
WebText2	19 billion	22%	2.9
Books1	12 billion	8%	1.9
Books2	55 billion	8%	0.43
Wikipedia	3 billion	3%	3.4

1. **500 billion 토큰 사전, 비지도 학습**
2. **휴먼 피드백 세부 튜닝 학습**

04. Large Language Models

LLMs 은 few-shot 학습자다.

- **언어 모델의 규모 확장**, 소수 데이터 학습(few-shot learning) 성능 크게 향상.
- 번역, 질의 응답, 빈칸 채우기 작업, 3자리 산술 연산 등 즉석 추론이나 도메인 적응 필요 작업도 처리 가능.
- GPT-3가 뉴스 기사 작성, **인간이 쓴 기사와 구별하기 어려움**.

arXiv:2005.14165v4 [cs.CL] 22 Jul 2020

Language Models are Few-Shot Learners

Tom B. Brown* Benjamin Mann* Nick Ryder* Melanie Subbiah*
Jared Kaplan† Prafulla Dhariwal Arvind Neelakantan Pranav Shyam Girish Sastry
Amanda Askell Sandhini Agarwal Ariel Herbert-Voss Gretchen Krueger Tom Henighan
Rewon Child Aditya Ramesh Daniel M. Ziegler Jeffrey Wu Clemens Winter
Christopher Hesse Mark Chen Eric Sigler Mateusz Litwin Scott Gray
Benjamin Chess Jack Clark Christopher Berner
Sam McCandlish Alec Radford Ilya Sutskever Dario Amodei
OpenAI

Abstract

Recent work has demonstrated substantial gains on many NLP tasks and benchmarks by pre-training on a large corpus of text followed by fine-tuning on a specific task. While typically task-agnostic in architecture, this method still requires task-specific fine-tuning datasets of thousands or tens of thousands of examples. By contrast, humans can generally perform a new language task from only a few examples or from simple instructions – something which current NLP systems still largely struggle to do. Here we show that scaling up language models greatly improves task-agnostic, few-shot performance, sometimes even reaching competitiveness with prior state-of-the-art fine-tuning approaches. Specifically, we train GPT-3, an autoregressive language model with 175 billion parameters, 10x more than any previous non-sparse language model, and test its performance in the few-shot setting. For all tasks, GPT-3 is applied without any gradient updates or fine-tuning, with tasks and few-shot demonstrations specified purely via text interaction with the model. GPT-3 achieves strong performance on many NLP datasets, including translation, question-answering, and cloze tasks, as well as several tasks that require on-the-fly reasoning or domain adaptation, such as unscrambling words, using a novel word in a sentence, or performing 3-digit arithmetic. At the same time, we also identify some datasets where GPT-3's few-shot learning still struggles, as well as some datasets where GPT-3 faces methodological issues related to training on large web corpora. Finally, we find that GPT-3 can generate samples of news articles which human evaluators have difficulty distinguishing from articles written by humans. We discuss broader societal impacts of this finding and of GPT-3 in general.

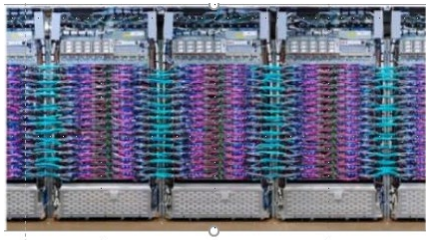
*Equal contribution
†Johns Hopkins University, OpenAI

Author contributions listed at end of paper.

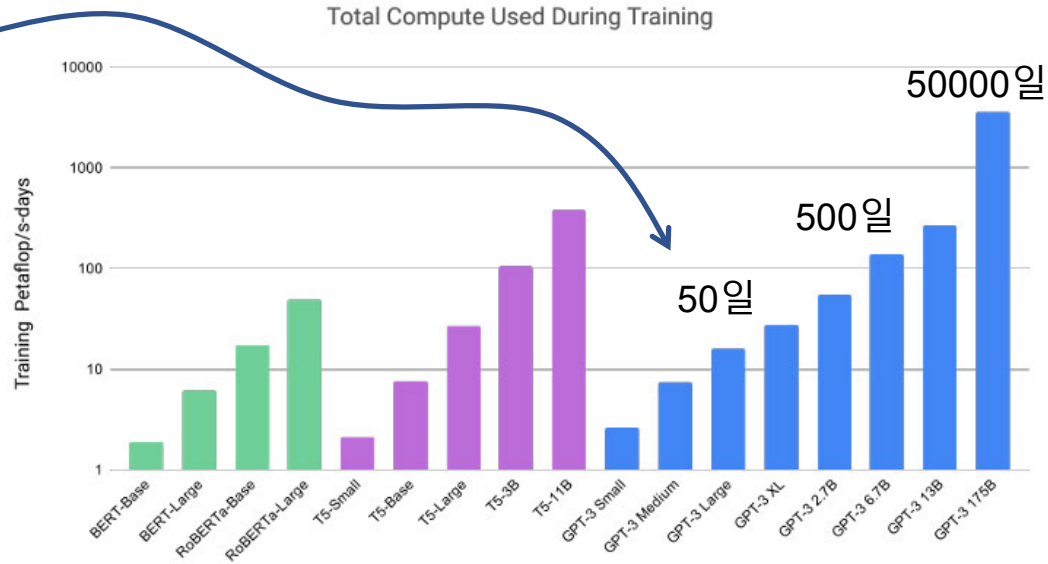
04. Large Language Models

거대 모델 훈련에는 거대 비용이 소모된다.

Cloud TPU를 쓰면 약 4주 만에 학습 완료, 2020 GPT-3 60억원/4주



Cloud TPU v3 Pod
100+ PetaFLOPS
Up to 1,024 TPU chips



NVIDIA A100은 5 Pflops

04. Large Language Models

GPT 4의 능력은 …

- 이미지와 텍스트 prompt
- Few shot 학습능력 크게 제고
- Bar Exam, LSAT, GRE, SAT Math, **각종 표준시험 영역을 섭렵하고**
- 일본어, **한국어** 등 24개 외국어능력, **chatGPT영어능력 능가**

표준 시험 성적

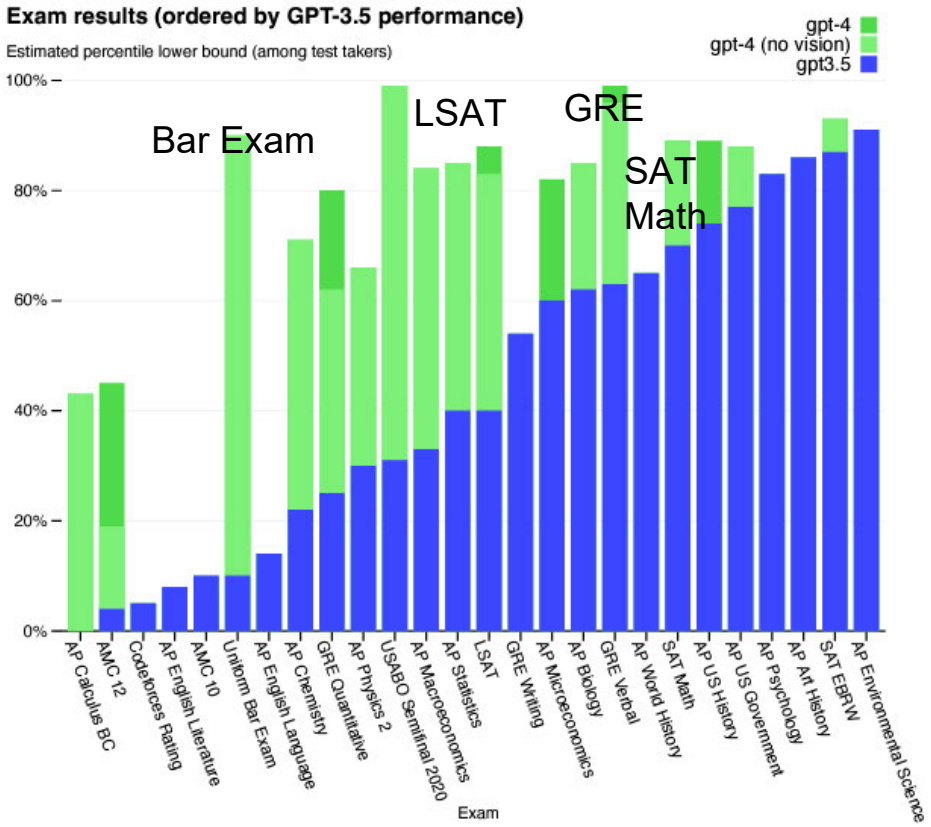


Figure 4. GPT performance on academic and professional exams. In each case, we simulate the conditions and scoring of the real exam. Exams are ordered from low to high based on GPT-3.5 performance. GPT-4 outperforms GPT-3.5 on most exams tested. To be conservative we report the lower end of the range of percentiles, but this creates some artifacts on the AP exams which have very wide scoring bins. For example although GPT-4 attains the highest possible score on AP Biology (5/5), this is only shown in the plot as 85th percentile because 15 percent of test-takers achieve that score.

04. Large Language Models

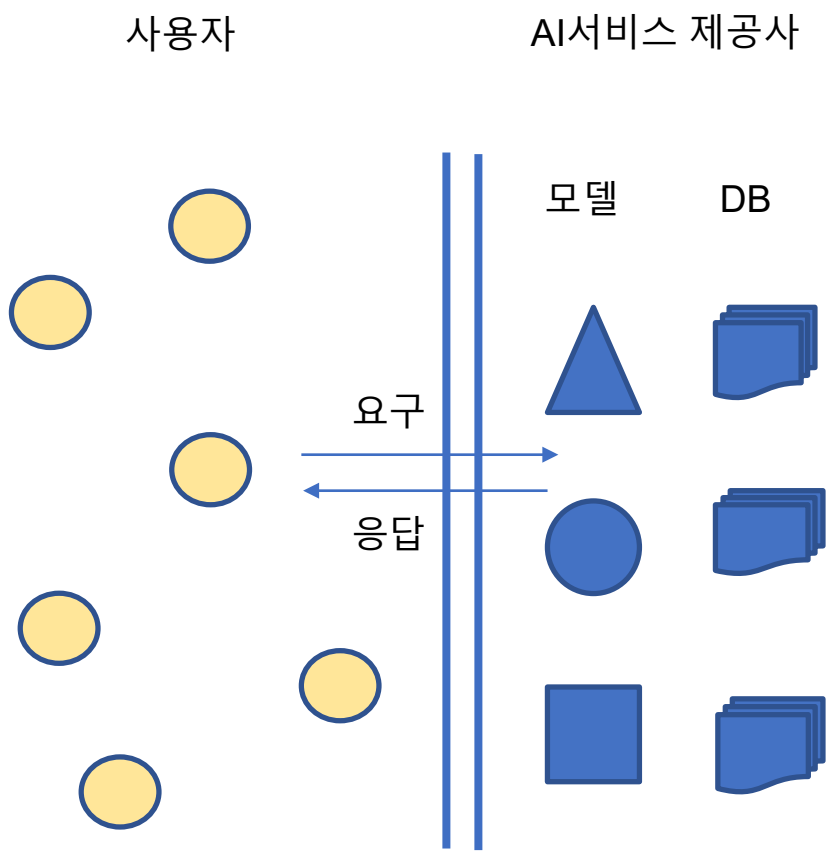
초 거대 기계학습 결과물의 등장, 혁신성, 문제점

- chatGPT (GPT-4)
  - 초 거대 모델/데이터셋 + 비지도/지도/강화 학습의 결과물
  - 문답을 통해 누구나 원하는 결과물을 신속하게 얻는게 가능
  - 요구사항 반영한 writing, poetry, 코드, …
  - AI 서비스 제품 출시 경쟁 촉발

- 문제점
  - 중앙집중성
  - 조 단위 학습 비용
  - 데이터소유권
  - 데이터 패권 (학습 가중치 왜곡)
  - 생성 데이터 만연 우려
  - AI Divide(전문가 vs 비전문가)

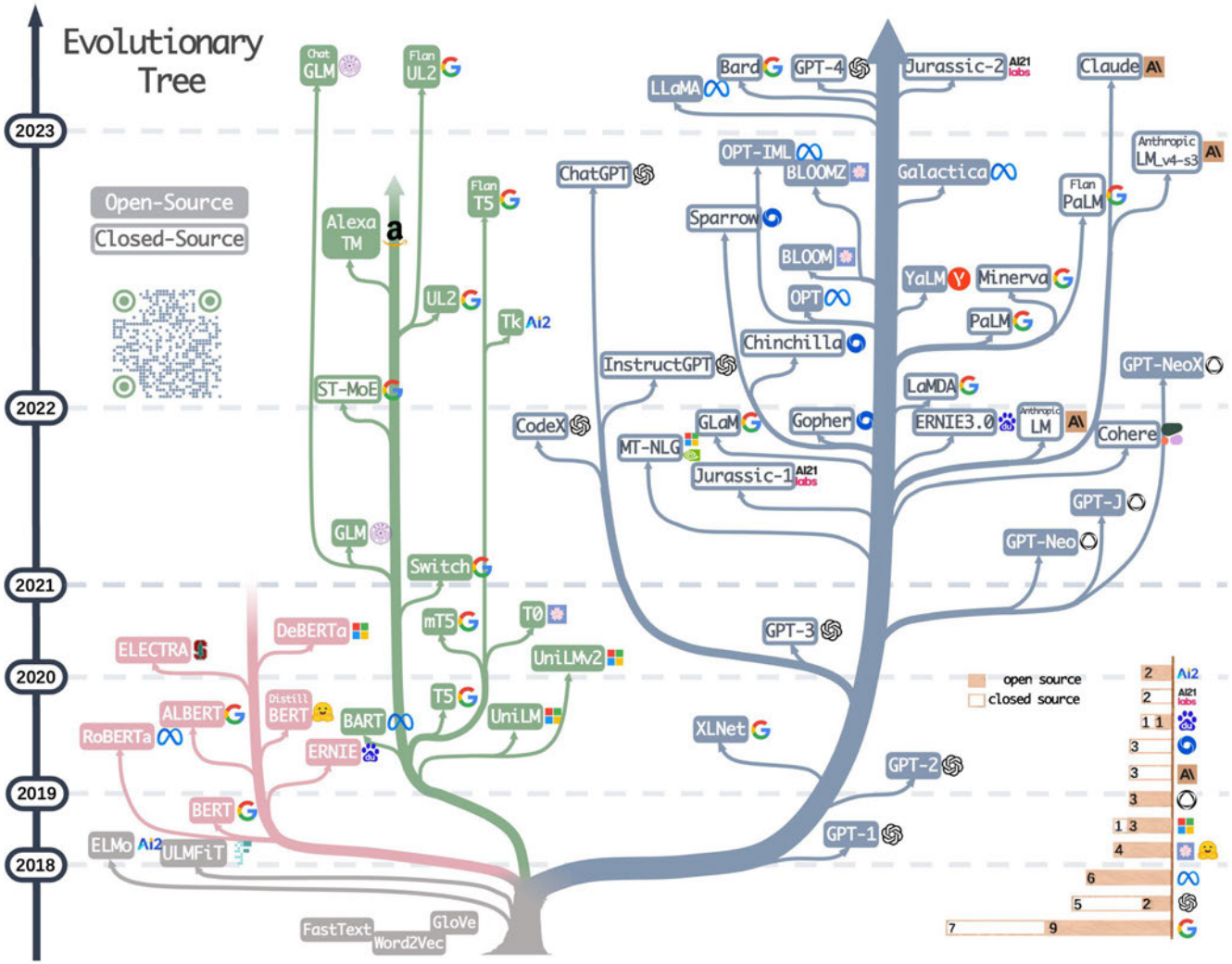
초거대 AI 모델을 초거대 AI기업이 독점할 때 생기는 문제

- AI 산업, 패권국에 종속
  - Who will control the future of this amazing technology? *
- EU 개인정보보호이사회
  - GDPR 위반 가능성 지적
- 비투명성
  - 학습 데이터 사용자 제공 데이터 반영 학습
- 모델 무결성
  - 차등 모델, 차등 서비스 요금
  - 사용자 간, 동일룰 적용 신용점수결정
- 개인정보 이슈
  - 비밀/기밀 데이터



* MIT Technology Review https://www.technologyreview.com/2023/03/25/1070275/chatgpt-revolutionize-economy-decide-what-looks-like/

04. Large Language Models



04. Large Language Models

Compact LLMs

- LLaMa
- Alpaca
- Ko-Alpaca
- LLaMa2
- AutoGPT

My AI Network on Blockchain

독점 AI에 대항하는 우리의 AI Network

https://worldland.foundation
https://twitter.com/Worldland_space

05. Worldland

Worldland Mainnet 론칭 행사 2023년 8월 8일08시08분



https://drive.google.com/file/d/1jgrbjElvpO9cqQyVAhmlcNgOUh22VjHx/view?usp=sharing 발표 자료 URL

https://www.youtube.com/watch?v=G_VuFSeMpRc&t=1917s YouTube 동영상

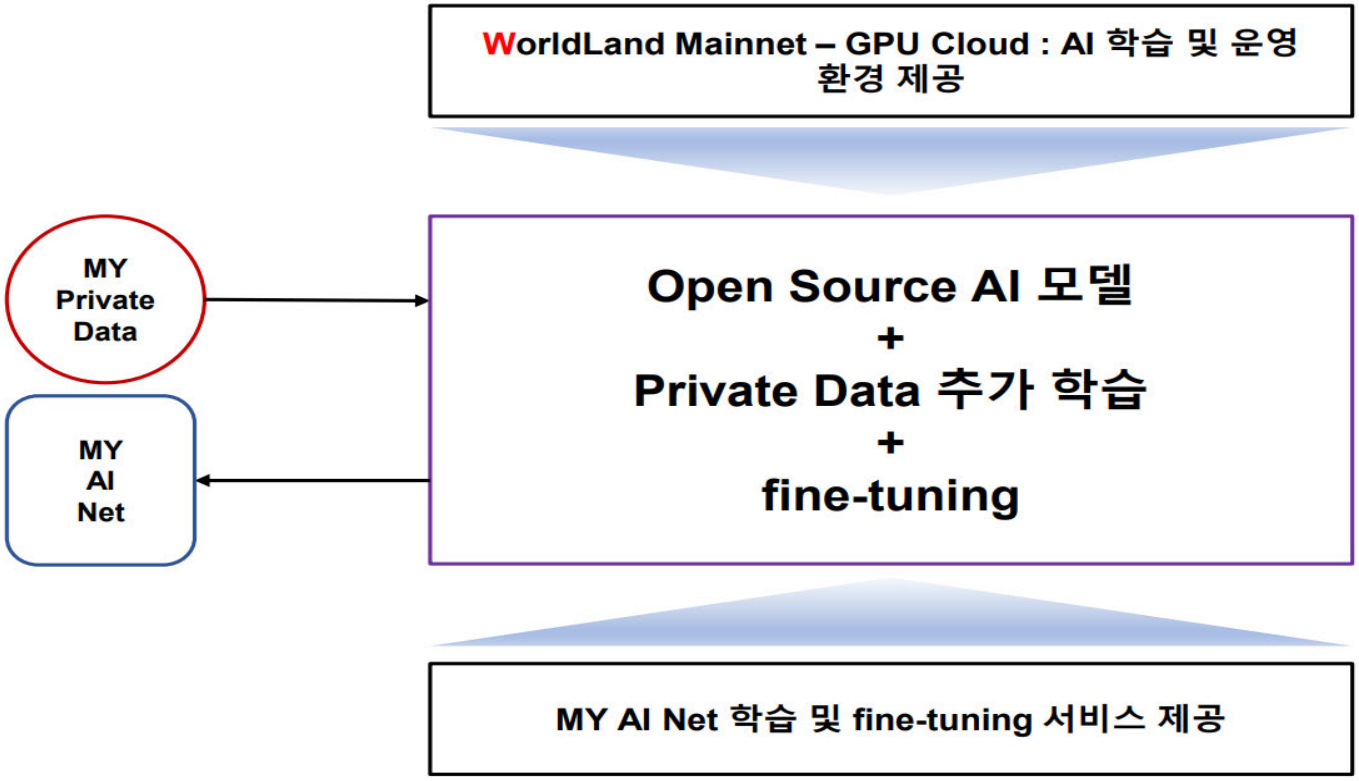
05. Worldland

지갑 생성하고, 주소 보내주소, WLC 200개 Aidrop 받기

- Worldland Mainnet 런칭 행사 비데오 참조
- 1. 핸폰에서 메타마스크 설치 및 지갑을 생성한다.
- 2. https://worldland.foundation/ 에 접속하여 지갑을 Connect한다.
- 3. 지갑 주소, 성함, 전화번호를 작성해 주십시오. https://docs.google.com/spreadsheets/d/1XRXz9FaIaP4f2kCOrCH0mWcIGnhBYZ6eCvlY08jzEvg/edit?pli=1#gid=0

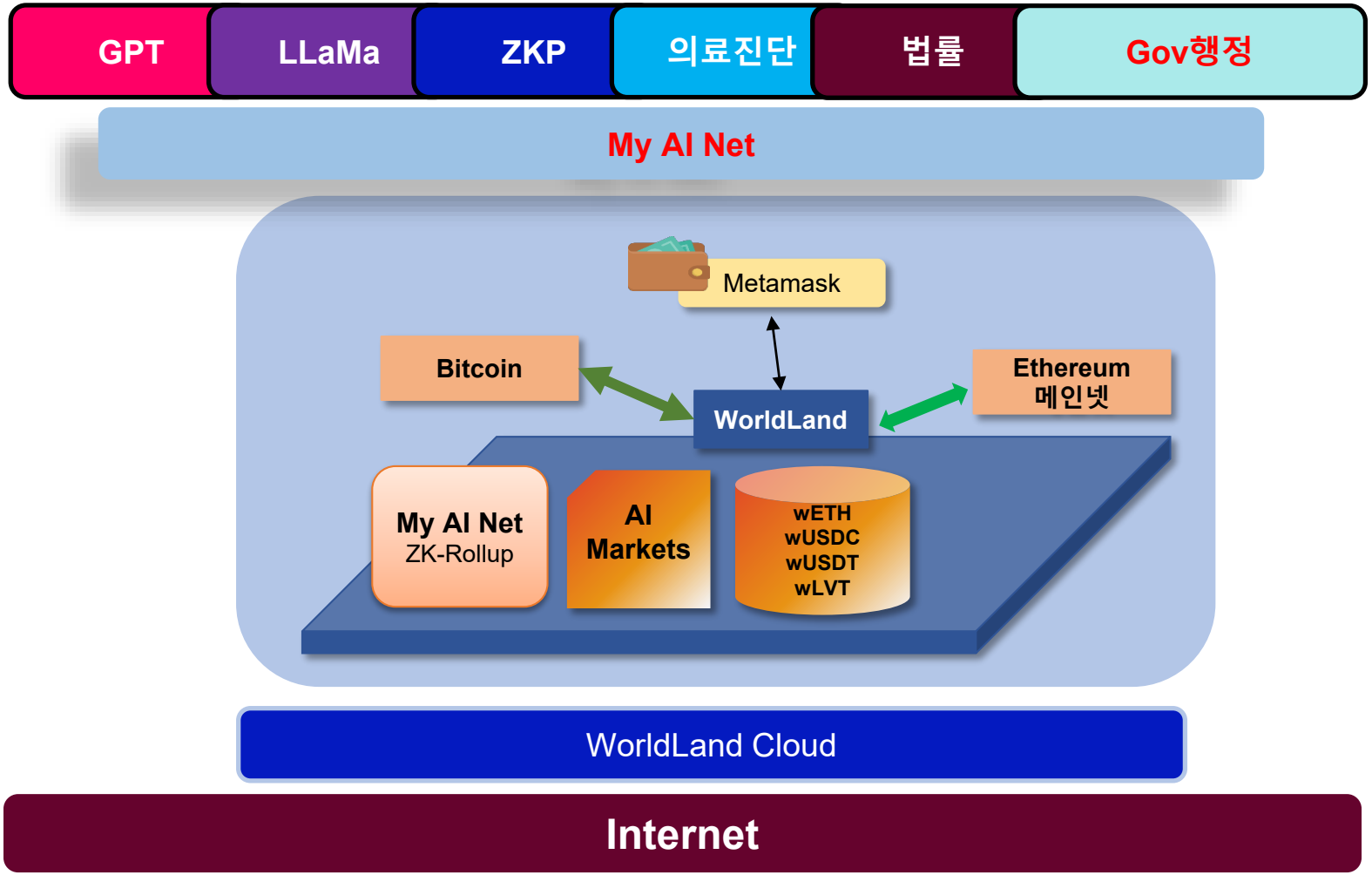
My AI란 무엇? 왜 필요한가? 왜 블록체인 인가?

- WorldLand (+ Web3), MY AI 네트워크 인프라 제공, 초거대 AI기업 독점 문제 해결
- 데이터 보안 / 학습 비용 등의 문제, 블록체인 및 보안 기술 통해 해결



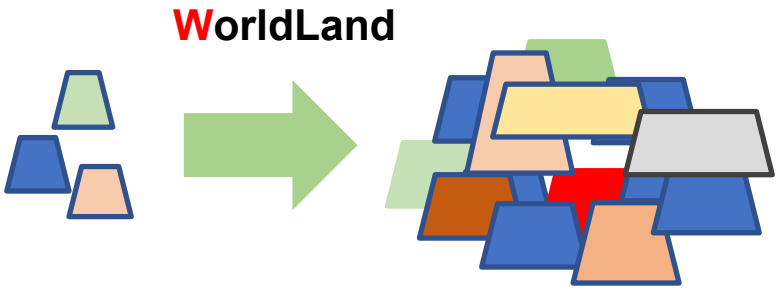
05. WorldLand

WorldLand는 글로벌 Mainnet 이며 My AI Net 서비스를 제공한다.

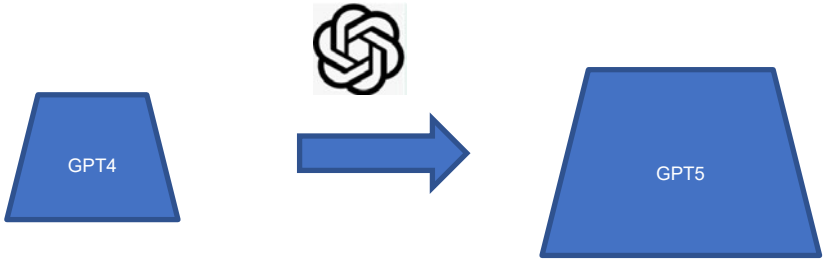


05. WorldLand

My AI Net과 OpenAI GPT, 누가 더 빠르게 성장할까?







- WorldLand My AI Net
  - 다양한 참여자가 다양한 지분을 갖고 직접 참여
  - 진짜데이터로 AI모델 학습
  - 학습결과물을 공유하고
  - 경제적 효과를 나누어 가짐

참여하고 협력하여 주도할 것인가?

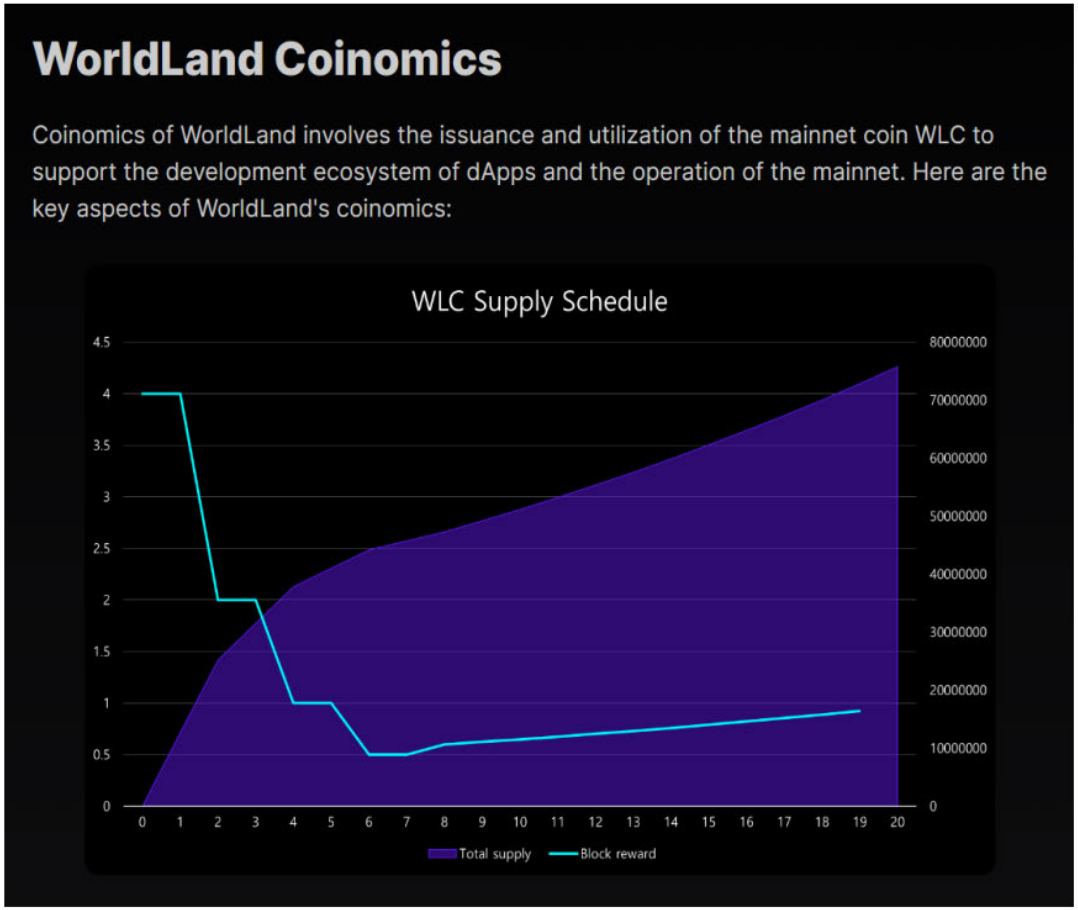
- OpenAI
  - 천문학적 규모의 돈으로
  - 생성된 데이터로 GPT학습
  - 경제적 효과는 독점

○ **WorldLand 화폐 발행 스케쥴은**
- 4번의 반감기 거쳐 안정화 단층에 이르고, 그 후로는 매년 4%씩 추가 발행

❑ **WorldLand 경제 규모 매년 4% 예측 가능하게 성장**

❑ **과도 호들링(Hodling) 방지**

❑ **화폐 사용 및 투자 촉진**



노드 설치하고 WLC 채굴에 참여하기

인터넷에 연결된 Windows PC를 켠다.

채굴설명 링크 : https://docs.worldland.foundation/miner/install-node 를 참조한다.

1. https://github.com/cryptoecc/ETH-ECC/releases

위 링크에서 worldland-windows-amd64-1.0.0-unstable-f47693c2.exe 파일 다운로드 및 실행

2. 설치 진행 : 보안 관련 경고가 뜨는 경우 확인을 통해 설치 진행
3. 프로그램 파일 추가 확인
4. 명령 프롬프트 cmd 실행
5. Cmd 창에 worldland console 입력 / 엔터
6. Cmd창에 miner.setEtherbase(" 자기 지갑 주소 ") 입력/엔터
7. True 메시지 확인
8. Cmd창에 miner.start(1) 입력 / 엔터

9 채굴 실행 완료

10. https://worldland.foundation 에 지갑 주소로 연결하면, 채굴 진행 상황 실시간 확인 가능

요약

- 인공지능은 Artificial General Intelligence로 발전중
- AGI 독점은 비극, 누구나 AGI개발에 능동적으로 참여하고, 동반 성장할때 인류에 큰 도움
- **WorldLand** 는 Green VCA기반 탈중앙 블록체인 네트워크
- **WorldLand**는 My AI Net 에 특화된 형태로 성장하고자 목표

09. 참고 자료

인공지능 관련 이흥노 교수 최근 3년간 논문 참조 자료

Full list at https://heungno.net/publications/

- **118. Manjit Kaur, Dilbag Singh, Vijay Kumar, Heung-No Lee**, "MLNet: Metaheuristics-based Lightweight Deep Learning Network for Cervical Cancer Diagnosis", IEEE Journal of Biomedical and Health Informatics, (Impact factor: 7.021, Category Mathematical & Computational Biology, JCR Quartile: Q1: JIF Percentile: 93.86). *(Link)*
- **117. Dilbag Singh, Manjit Kaur, Jazem Mutared Alanazi, Ahmad Ali AlZubi, Heung-No Lee**, "Efficient evolving deep ensemble medical image captioning model", IEEE Journal of Biomedical and Health Informatics, (Impact factor: 7.021, Category Mathematical & Computational Biology, JCR Quartile: Q1: JIF Percentile: 93.86). *(Link)*
- **116. Varsha Himthani, Vijaypal Singh Dhaka, Manjit Kaur, Geeta Rani, Meet Oza, and Heung-No Lee,** Comparative Performance Assessment of Deep Learning based Image Steganography Techniques" *Scientific Reports*, (Impact factor: 4.996, Category Multidisciplinary Sciences, JCR Quartile: Q2: JIF Percentile: 74.66).
- **115. Jian Su; Zhenlong Liao, Zhengguo Sheng, Alex X. Liu, Dilbag Singh, and Heung-No Lee,** " Human Activity Recognition Using Self-powered Sensors Based on Multilayer Bi-directional Long Short-Term Memory Networks" *IEEE Sensors Journal*, (Impact factor: 4.325, Category Instruments & Instrumentation, JCR Quartile: Q1: JIF Percentile: 78.91, Do-Yak project). *(Link)*
- **114. Rohini Raina, Naveen Kumar Gondhi, Chaahat, Dilbag Singh, Manjit Kaur, and Heung-No Lee,** "A Systematic Review on Acute Leukemia Detection Using Deep Learning Techniques" Archives of Computational Methods in Engineering, Springer, (Impact factor: 7.302, Category: Mathematics, Interdisciplinary Applications, JCR Quartile: Q1: JCR Percentile: 99.54, IITP & Do-Yak project). *(Link)*
- **108. Shashank S. Singh, Vishal Srivastava, Ajay Kumar. Shailendra Tiwari, Dilbag Singh, and Heung-No Lee**, "Social Network Analysis: A Survey on Measure, Structure, Language Information Analysis, Privacy, and Applications" Accepted in ACM Transactions on Asian and Low-Resource Language Information Processing, (Impact factor: 3.316, Category: Computer Science, Artificial Intelligence, JCR Quartile: Q4: JCR Percentile: 18.35, IITP & Do-Yak project). *(PDF)*
- **93. Cheolsun Kim, Pavel Ni, Kang Ryeol Lee, and Heung-No Lee***, "Mass production-enabled computational spectrometers based on multilayer thin films", Accepted in Scientific Reports, (Impact Factor: 4.380, Category: Multidisciplinary Sciences, JCI Quartile: Q1: JCI Percentile: 85.55, Do-Yak project).
- **82. Mohamed Yaseen Jabarulla and Heung-No Lee***, "Blockchain and Artificial Intelligence-based Patient-Centric Healthcare System for Combating COVID-19 Pandemic: Opportunities and Applications," *Healthcare,* 2021, 9(8), 1019. doi: https://doi.org/10.3390/healthcare9081019 (Impact Factor: 2.645, Do-Yak project). Paper: (Journal Page) (Cover Story) (Cover Image).

감사합니다!

비즈니스 협력 연락처

이흥노 대표: ceo.libervance@gmail.com

WorldLand

AI Nexus

2023.05.14 | 리버밴스㈜

TIPS 투자(22년 5월), R&D(22’ 6월), 해외사업(22’11월)

https://heungno.net

01. 리버밴스 대표

이흥노 교수

- GIST교수/인공지능연구소장
- 과기부 ITRC 블록체인 지능융합 센터장
- 대통령직속 정책기획위원회
- 블록체인경제센터장
- 센서지능화센터장
- (전)GIST연구원장
- 총 285편의 논문(국제 200여편, **SCI 100여편 포함**), 국내논문 104편,국내외 특허 60여건
- 해동학술상(2019)
- GIST연구상/대표기술상(3회)
- 이달의과학자상(2014.1)
- 국무총리표창 (2022.4)

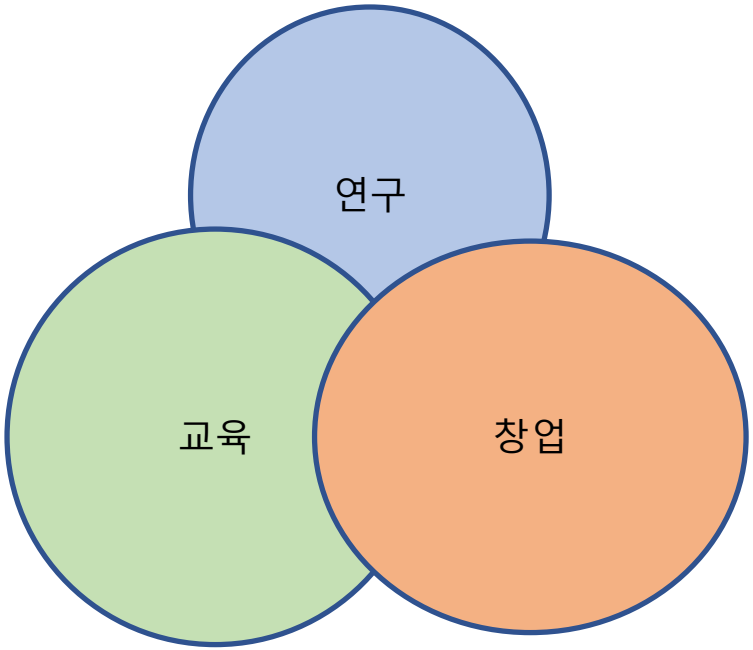
http://gist-itrc.org



2022년 4월 버전, 상세 이력 @ https://heungno.net

주요 이력

- IEEE Trans. Cybernetics 편집위원 (국제 Top 4%)
- ITRC 블록체인지능융합 센터장 (20’ – 28’)
- GIST 인공지능연구소장
- GIST Infonet 연구실 지도교수
- (전) University of Pittsburgh, PA, USA, 전임 교수
- (전) Hughes Research Lab 연구원, Malibu, CA, USA.



Executive Summary

월드랜드 사업은 인공지능과 블록체인의 장점을 결합한 사업입니다. 월드랜드는 GPU 클라우드가 될 것입니다. 왜냐하면 월드랜드 합의알고리즘 Green VCA는 ASIC을 허용하지 않기 때문입니다. 전기에너지의 과도한 사용이 필요치 않으면서도 탈중앙성을 극대화합니다. 그런데 GPU는 AI학습에 반드시 필요한 두 가지 자원중 하나입니다. 또다른 필수 자원은 기업과 기관이 보유하고 있는 공유 불가능한 내부 기밀 자료입니다. 월드랜드는 기업이 기밀을 지키며 AI모델을 고도화할 수 있는 환경을 제공합니다. 나아가 기업이 기밀을 지키며 모델을 연합학습할 수 있는 기술적 솔루션을 제공합니다. 전 세계의 기업과 기관이 참여할 수 있고 그 역량을 월드랜드에서 하나로 묶어 집중할 수 있습니다. 그러므로 월드랜드는 중앙집중형 글로벌 AI 기업을 능가할 수 있습니다. 유능한 리더가 이끄는 집단지능은 다양성과 의외성을 확보할 수 있으며 지속으로 급변하는 상황속에서 살아남을 수 있는 지혜를 찾을 수 있습니다. 수익 극대화라는 유일한 목표를 일관되게 추구할 수 밖에 없는 중앙집중형 시스템을 늘 이기는 이유입니다. 바로 자연이 가르쳐주는 이치입니다.

02. WorldLand 비전

오늘날 인류가 직면한 여러 문제들로 인해 전 세계 사람들은 많은 어려움을 겪고 있습니다. 이러한 문제들은 정치 경제적 불확실성, 글로벌 경제 둔화, 높은 부채 수준, 인플레이션, 양극화, 인간 가치의 저하 등 다양한 양상으로 나타납니다. 이러한 어려움에 대응하기 위해 **WorldLand** 사업을 제안하게 되었습니다.

WorldLand 사업의 비전은 글로벌 탈중앙 디지털 네트워크를 구축하여 전 세계 시민들이 천부 인권과 개인의 존엄성을 존중하며 일대 일 거래가 가능한 투명한 플랫폼을 만드는 것입니다. 이를 통해 혁신 지식 생산 활동을 촉진하고 탈세계화가 진행되어도 전쟁을 억제하고 평화를 정착할 수 있는 환경을 조성하며, 인류가 서로 돕고 함께 성장하는 글로벌 사회를 만들고자 합니다.

WorldLand 헌장은 다음과 같은 목표를 세우고 있습니다:
1.기본 인권에 대한 신념을 고수하고 인간의 존엄성과 가치를 존중합니다.
2.남녀, 큰 국가와 작은 국가의 평등한 권리를 증진합니다.
3.더 큰 자유 속에서 사회 발전과 삶의 질 향상을 도모합니다.
4.좋은 이웃으로서 평화롭게 함께 살아갈 수 있도록 합니다.
5.국제 평화와 안전을 유지하기 위해 힘을 결집합니다.
6.인류의 경제 및 사회 발전을 촉진하기 위해 **인공지능**과 **블록체인**과 같은 디지털 기술을 활용합니다.

이러한 목표를 달성하기 위해, **WorldLand** 는 전 세계 디지털 네트워크를 구축하고자 합니다. 이 플랫폼을 통해 사람들은 서로 돕고 협력하여 인류가 직면한 여러 문제들을 해결하고, 보다 나은 글로벌 사회를 만드는데 기여할 수 있습니다.

03. WorldLand 정의

WorldLand 는

- 세계인이 교류하는 디지털 랜드다.
- 반감기를 적용한 화폐 WLC를 발행한다.
- 참여 노드가 WLC 주조권을 갖는다.
- 노드는 Green VCA로 합의한다.
  - 양자내성암호를 사용한다.
  - 동전던지기함수와 오류정정부호POW로 합의한다.
  - EVM을 사용한다.
- 양자내성 영지식 증명을 활용하여 확장성을 제고한다.
- 킬러 dApp AI Nexus 플랫폼을 제공한다.
- AI Nexus는 AI가 교사-학생 학습, 전이학습, 연합학습 등으로 다양한 사회 구성원과 함께 진화하는 네트워크를 의미한다.

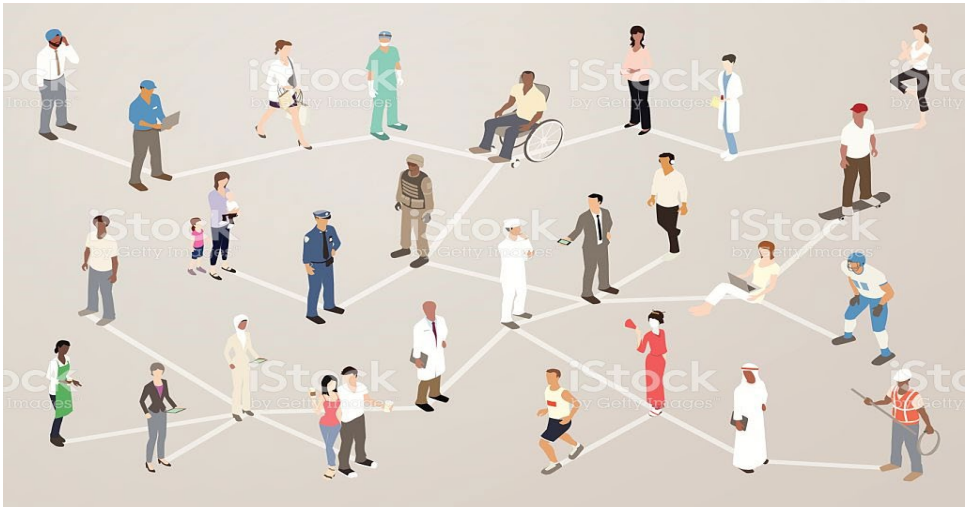
두뇌와 두뇌는 연결된 사회속에서 성장한다.

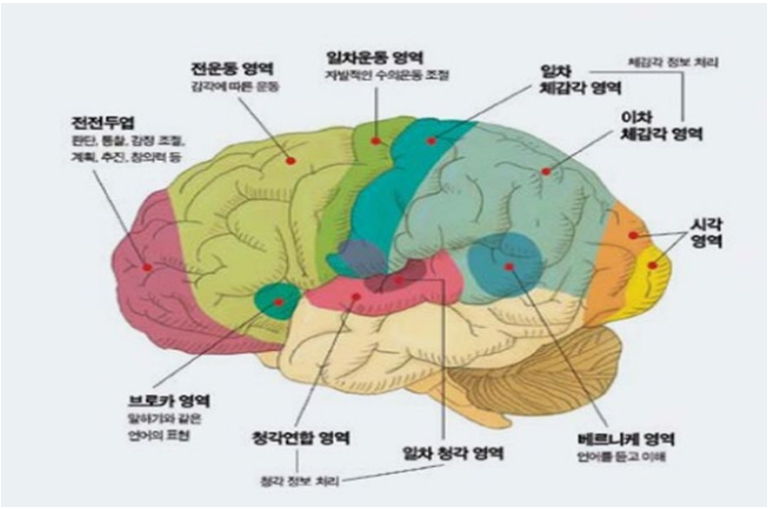
사람의 지능은 두뇌에 있다.

사회에 속해있다.

사람은 생명체다.







우리의 두뇌는 사회 활동의 결과물이다.

Sensors
신호습득

보고, 듣고, 냄새맡고,
만져보고, 느끼고

부모에게 물어보고

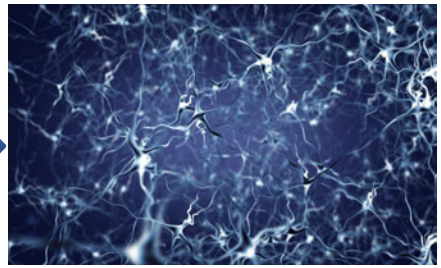
선생님과 대화해 보고

책을 읽고, TV를 보고

SNS 등을 통한

센싱을 통해

Situational Awareness
(신호 + 경험)





Actions
결정

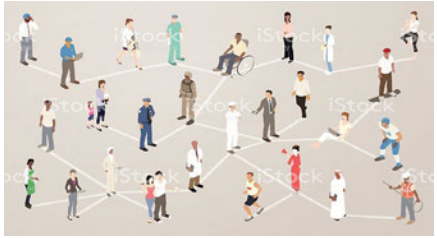
이동하는 것
먹는 것
입는 것
공부하는 것
노는 것

정치/경제/사회

협력/공유/신뢰/책임

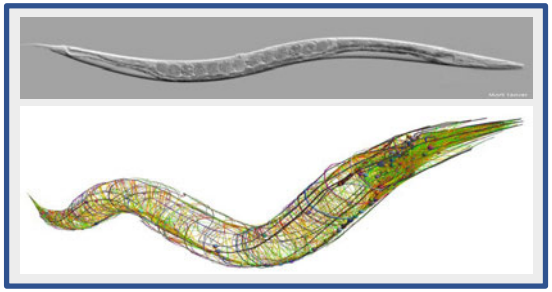
결정한다



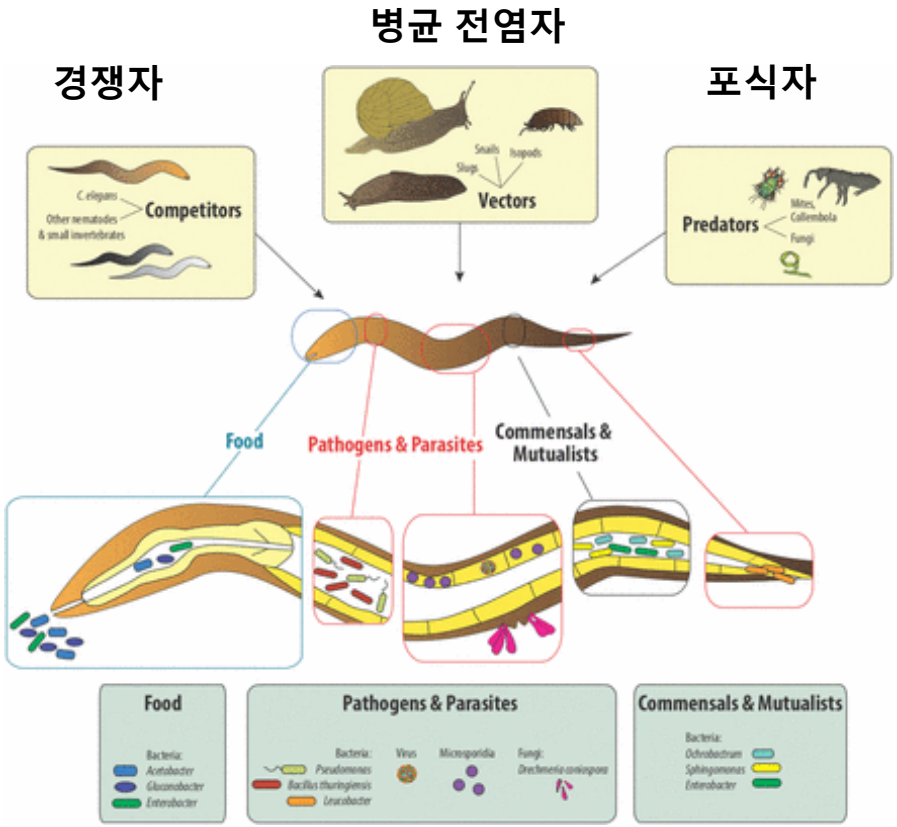




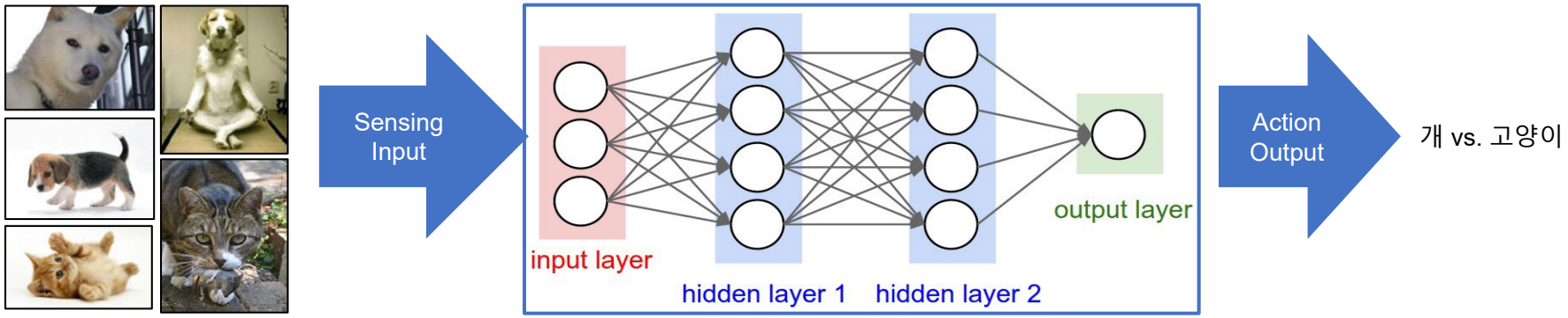
꼬마선충(C elegans), 302개의 뉴론으로 늪지에서 산다.







심층신경망 훈련은 데이터와 GPU인프라가 필요하다.



대규모 데이터 확보, 뉴럴넷 모델을 학습시키고, 학습 완료된 모델을 사용하는것

학습 : 네트워크 가중치와 매개 변수를 조정하여 IN –OUT 관계 파악하고 예측, 분류, 생성 값을 개선하는 것을 의미

추론 : 학습된 뉴럴넷을 새로운 Input에 대해 예측, 분류, 생성 등 판단에 사용

주요 기술: 심층신경망 (deep neural network), 역전파 (back propagation), 확률경사하강법 stochastic gradient descent (SDG)

심층신경망 자연어 처리 NLP 발전 주요 이력

	paper	비고
2013.01	Word2vec: “Efficient estimation of word representations in vector space,” Google **단어를 고차원 벡터로 표현. 대규모데이터로 학습. 유사한 연관 단어 벡터는 서로 가깝게 표현.**	BoW: 연관단어 → 타겟단어 Skip-gram: 타겟단어 → 연관단어
2014.02	“Glove: Global vectors for word representation,” by Pennington et.al., Stanford University	전역적인 count-base 통계 모델 및 예측 모델 활용
2016.07	“Enriching word vectors with subword information,” Bojanowski et. al. 단어를 하위어 단어로 분해하고 임베딩, 형태론적 변형과 유사성 포착	단어를 Subword로 쪼개어 형태론적 변형 포착
2017.06	Transformer. “Attention is all you need,” Vaswani et.al. **주의집중 메카니즘 활용, 다중 계층 인코더로 문맥 파악, 다중 계층 디코더로 문맥화된 단어 생성**	**주의 집중 메카니즘**
2018.02	Elmo. “Deep contextualized word representations,” Peters et.al. 양방향 LSTM사용, 좌우 문맥을 고려, 각 단어의 문맥화된 표현 생성	양방향 LSTM문맥
2018.10	BERT: bidirectional encoder representation from Transformers, Devlin et.al. Transformer를 활용 단어 문맥 양방향 고려, 단어임베딩	양방향 Transformer문맥

Transformer

- Attention is all you need! (NIPS2017)
  - GER-ENG 번역기
  - Input 문장을 받아 Output 문장을 생성
  - IN 문장에서 단어들간의 관계 및 주목할 단어 파악, OUT에서 한 단어씩 생성
- 수십 층의 심층 신경망
  - 학습 매개 변수

$$O(layers \times d_{\text{model}}{}^2) \sim \text{ 10s million}$$

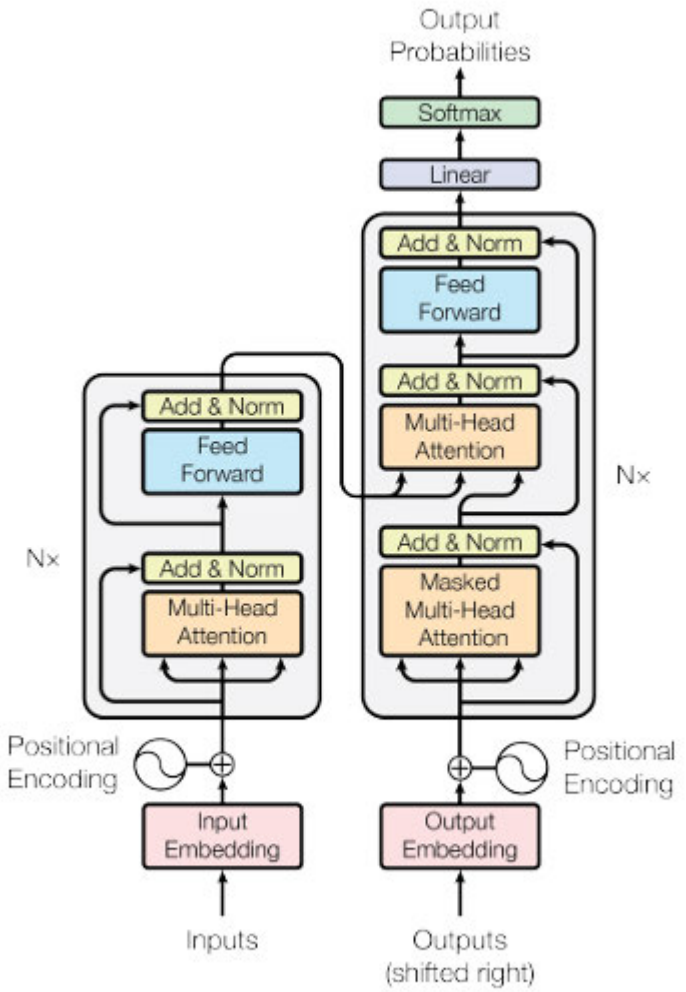


Figure 1: The Transformer - model architecture.

Transformer는 언어 변환 모델로 시작하였다.



OpenAI는 Transformer를 맥락 인지 문장 생성기로 발전시킴.

- GPT-3, July 2020
- GPT-3.5, March 2022
  - 1750억 개 매개변수
  - 800 GB storage
- Emergent Abilities:
  - Few-shot learning
  - Comm. language
- Computer codes, poetry, prose, "Terrifying ability," NYT 2020.

Table 2.2: Datasets used to train GPT-3.

Dataset	Quantity (tokens)	Weight in training mix	Epochs elapsed when training for 300B tokens
Common Crawl (filtered)	410 billion	60%	0.44
WebText2	19 billion	22%	2.9
Books1	12 billion	8%	1.9
Books2	55 billion	8%	0.43
Wikipedia	3 billion	3%	3.4

1. **500 billion 토큰 사전, 비지도 학습**
2. **휴먼 피드백 세부 튜닝 학습**

LLMs 은 few-shot 학습자다.

- **언어 모델의 규모 확장**, 소수 데이터 학습(few-shot learning) 성능 크게 향상.
- 번역, 질의 응답, 빈칸 채우기 작업, 3자리 산술 연산 등 즉석 추론이나 도메인 적응 필요 작업도 처리 가능.
- GPT-3가 뉴스 기사 작성, **인간이 쓴 기사와 구별하기 어려움**.

arXiv:2005.14165v4 [cs.CL] 22 Jul 2020

Language Models are Few-Shot Learners

Tom B. Brown* Benjamin Mann* Nick Ryder* Melanie Subbiah*
Jared Kaplan† Prafulla Dhariwal Arvind Neelakantan Pranav Shyam Girish Sastry
Amanda Askell Sandhini Agarwal Ariel Herbert-Voss Gretchen Krueger Tom Henighan
Rewon Child Aditya Ramesh Daniel M. Ziegler Jeffrey Wu Clemens Winter
Christopher Hesse Mark Chen Eric Sigler Mateusz Litwin Scott Gray
Benjamin Chess Jack Clark Christopher Berner
Sam McCandlish Alec Radford Ilya Sutskever Dario Amodei

OpenAI

Abstract

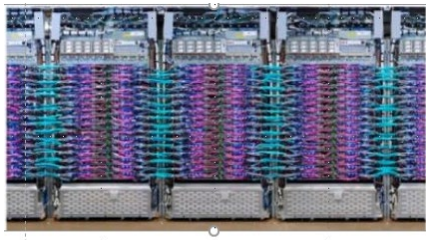
Recent work has demonstrated substantial gains on many NLP tasks and benchmarks by pre-training on a large corpus of text followed by fine-tuning on a specific task. While typically task-agnostic in architecture, this method still requires task-specific fine-tuning datasets of thousands or tens of thousands of examples. By contrast, humans can generally perform a new language task from only a few examples or from simple instructions – something which current NLP systems still largely struggle to do. Here we show that scaling up language models greatly improves task-agnostic, few-shot performance, sometimes even reaching competitiveness with prior state-of-the-art fine-tuning approaches. Specifically, we train GPT-3, an autoregressive language model with 175 billion parameters, 10x more than any previous non-sparse language model, and test its performance in the few-shot setting. For all tasks, GPT-3 is applied without any gradient updates or fine-tuning, with tasks and few-shot demonstrations specified purely via text interaction with the model. GPT-3 achieves strong performance on many NLP datasets, including translation, question-answering, and cloze tasks, as well as several tasks that require on-the-fly reasoning or domain adaptation, such as unscrambling words, using a novel word in a sentence, or performing 3-digit arithmetic. At the same time, we also identify some datasets where GPT-3's few-shot learning still struggles, as well as some datasets where GPT-3 faces methodological issues related to training on large web corpora. Finally, we find that GPT-3 can generate samples of news articles which human evaluators have difficulty distinguishing from articles written by humans. We discuss broader societal impacts of this finding and of GPT-3 in general.

*Equal contribution
†Johns Hopkins University, OpenAI

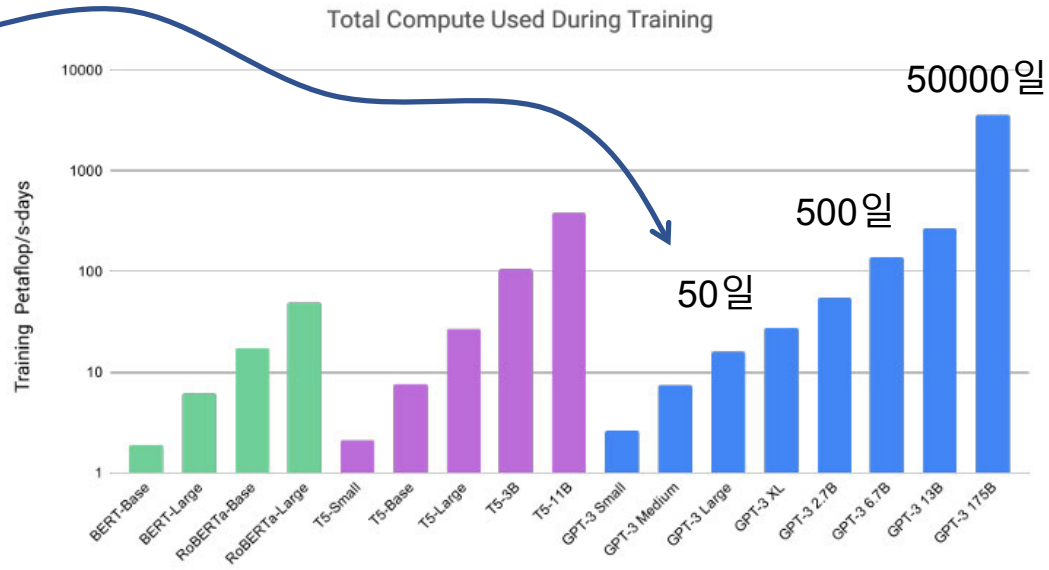
Author contributions listed at end of paper.

거대 모델 훈련에는 거대 비용이 소모된다.

Cloud TPU를 쓰면 약 4주 만에 학습 완료, 2020 GPT-3 60억원/4주



Cloud TPU v3 Pod
100+ PetaFLOPS
Up to 1,024 TPU chips



NVIDIA A100은 초당 9.7 Pflops

GPT 4의 능력은 …

- 이미지와 텍스트 prompt
- Few shot 학습능력 크게 제고
- Bar Exam, LSAT, GRE, SAT Math, **각종 표준시험 영역을 섭렵하고**
- 일본어, **한국어** 등 24개 외국어능력, **chatGPT영어능력 능가**

표준 시험 성적

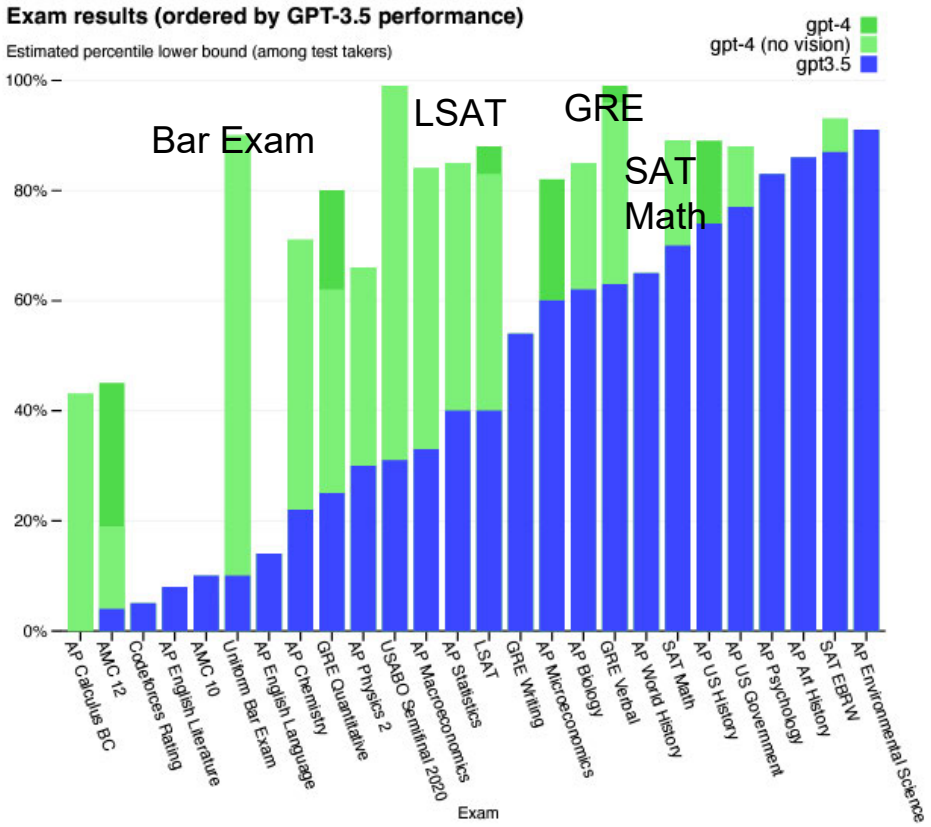


Figure 4. GPT performance on academic and professional exams. In each case, we simulate the conditions and scoring of the real exam. Exams are ordered from low to high based on GPT-3.5 performance. GPT-4 outperforms GPT-3.5 on most exams tested. To be conservative we report the lower end of the range of percentiles, but this creates some artifacts on the AP exams which have very wide scoring bins. For example although GPT-4 attains the highest possible score on AP Biology (5/5), this is only shown in the plot as 85th percentile because 15 percent of test-takers achieve that score.

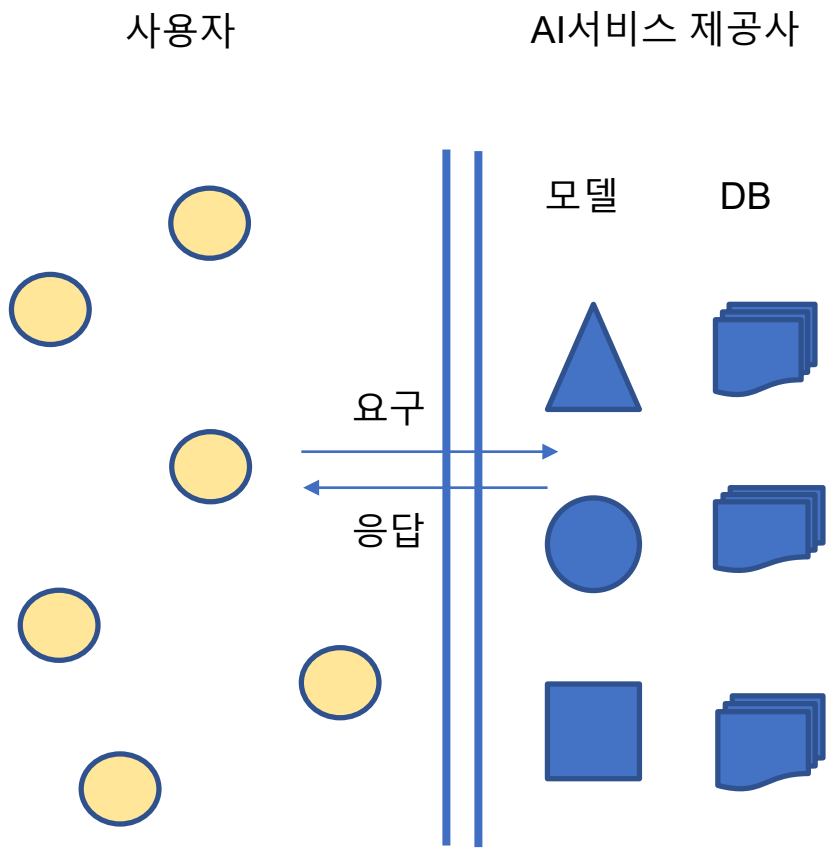
초 거대 기계학습 결과물의 등장, 혁신성, 문제점

- chatGPT (GPT-4)
  - 초 거대 데이터셋 + 비지도/지도/강화 학습의 결과물
  - 문답을 통해 누구나 원하는 결과물을 신속하게 얻는게 가능
  - 요구사항 반영한 writing, poetry, 코드, …
  - AI 서비스 제품 출시 경쟁 촉발

- 문제점
  - 중앙집중성
  - 조 단위 학습 비용
  - 데이터소유권
  - 데이터 패권 (학습 가중치 왜곡)
  - 생성 데이터 만연 우려
  - AI Divide(전문가 vs 비전문가)

초거대 AI 모델을 초거대 AI기업이 독점할 때 생기는 문제

- AI 산업, 패권국에 종속
  - Who will control the future of this amazing technology? *
- EU 개인정보보호이사회
  - GDPR 위반 가능성 지적
- 비투명성
  - 학습 데이터 사용자 제공 데이터 반영 학습
- 모델 무결성
  - 차등 모델, 차등 서비스 요금
  - 사용자 간, 동일룰 적용 신용점수결정
- 개인정보 이슈
  - 비밀/기밀 데이터

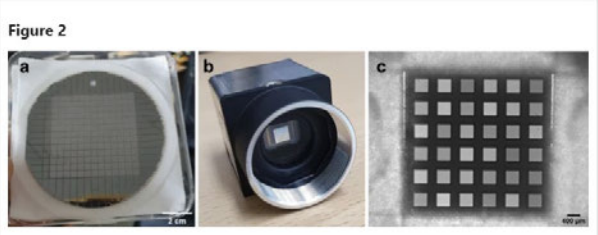


* MIT Technology Review https://www.technologyreview.com/2023/03/25/1070275/chatgpt-revolutionize-economy-decide-what-looks-like/

상기한 문제점들은 web3.0과

- 연합학습, 비투명성 제고
  - 개인데이터 반영 모델 학습
  - 학습모델공유
- 영지식증명, 모델 진위성/무결성
  - 차등 모델, 차등 서비스 요금
  - 사용자 간, 동일룰 적용
  - 신용점수결정
- Green VCA
  - VCT + VRF
  - ECCPoW
- 영지식센싱, 개인정보비유출
  - 사용자 인증시 (홍채) 비저장

해결 가능하다.



Mass production-enabled computational spectrometers based on multilayer thin films | Scientific Reports (nature.com)

Article | Open Access | Published: 08 March 2022

Mass production-enabled computational spectrometers based on multilayer thin films

Cheolsun Kim, Pavel Ni, Kang Ryeol Lee & Heung-No Lee

Scientific Reports **12**, Article number: 4053 (2022) | Cite this article

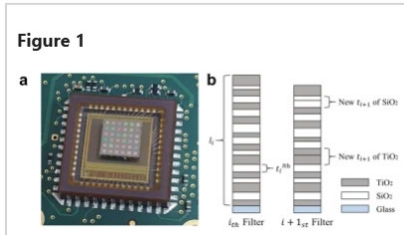
1467 Accesses | **2** Citations | Metrics

Abstract

Multilayer thin film (MTF) filter arrays for computational spectroscopy are fabricated using stencil lithography. The MTF filter array is a 6 × 6 square grid, and 169 identical

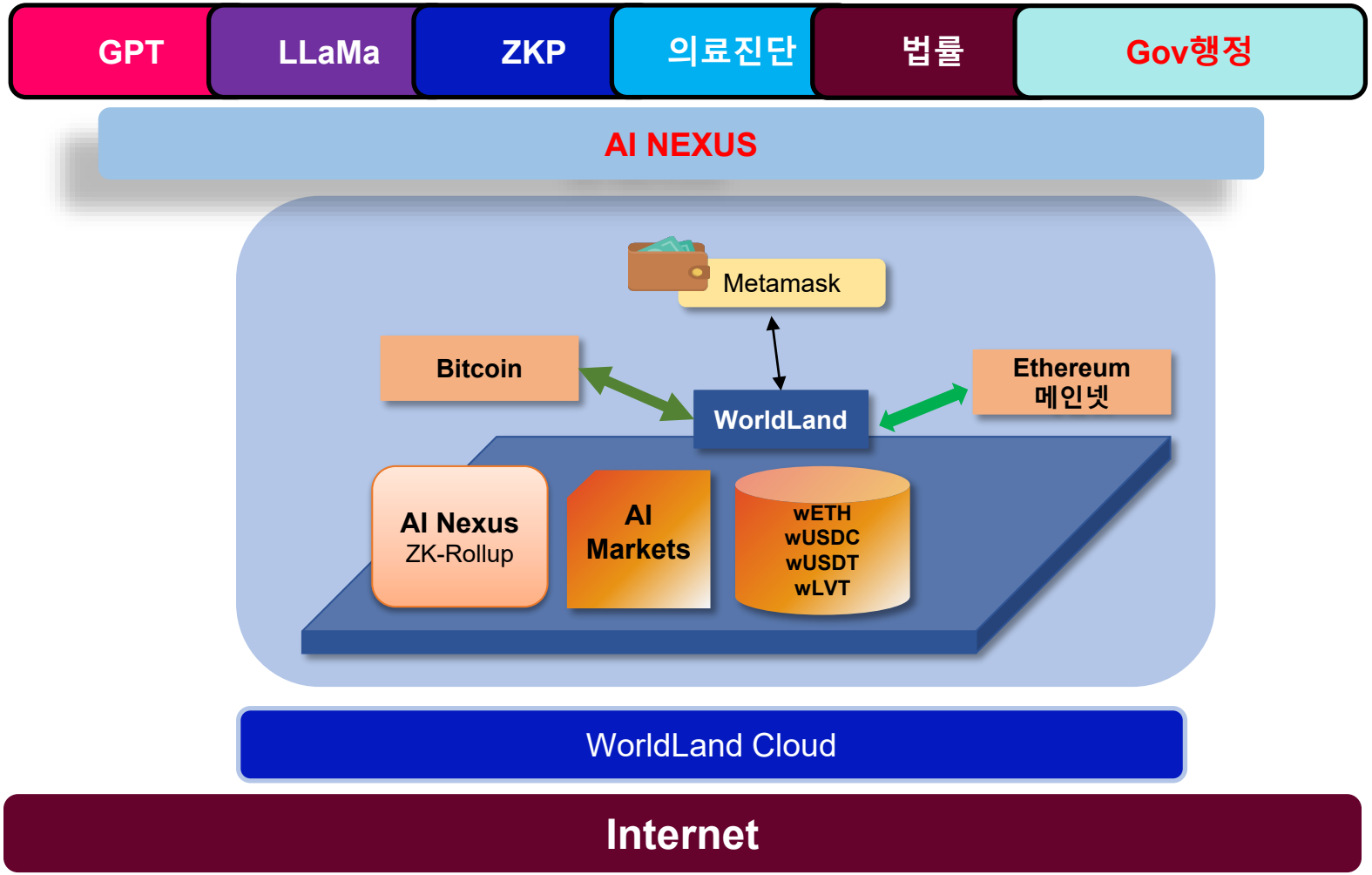
대량 생산 가능한,
초소형 초분광카메라로
신원 인증 가능

images and reconstructed 350 monochromatic images in the wavelength range of 500 to 849 nm, with a spacing of 1 nm. These computational spectrometers could be useful for various applications that require compact size, high resolution, and wide working range.



03. WorldLand

WorldLand는 글로벌 Mainnet 이며 AI Nexus 서비스를 제공한다.



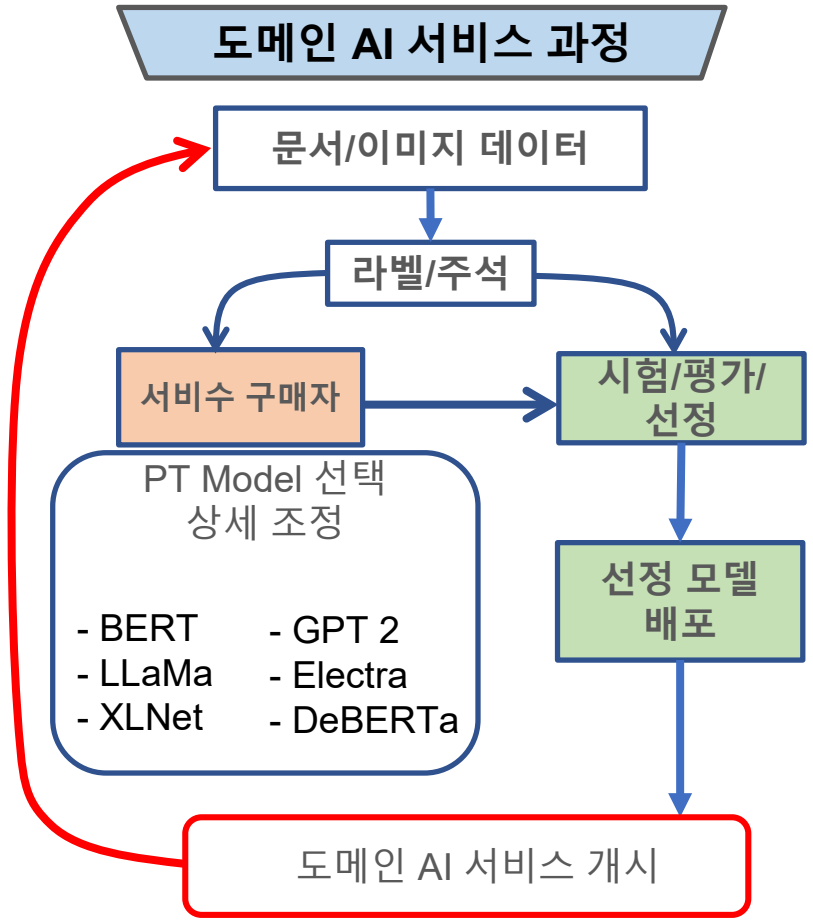
03. WorldLand

AI Nexus는 탈중앙 데이터, AI 학습 및 추론 서비스 네트워크

- NEXUS: “연결된 것”
- AI Nexus는 연결된 AI Agent 넷을 의미
- **AI Nexus Nexus는 WorldLand의 dApp**
- AI Nexus 생태계
  - 데이터 제공 참여자에게 Soul Bound Token (SBT)지급
  - 특수 도메인 데이터로 맞춤형 훈련 원하는 기업/기관
  - GPU 및 AI Nexus 서비스 제공자
    - 개인 맞춤형 AI 훈련 서비스
    - 훈련 완료된 AI 추론 모델 서비스
- 법률, 의료진단, 교육, 행정 등 기관AI, 기업AI, GovAI 제공 가능

사용자(기업/연구소/정부)가 보유한 도메인 데이터로 안전하게 완성형 AI모델을 확보하는 방법은 없을까?

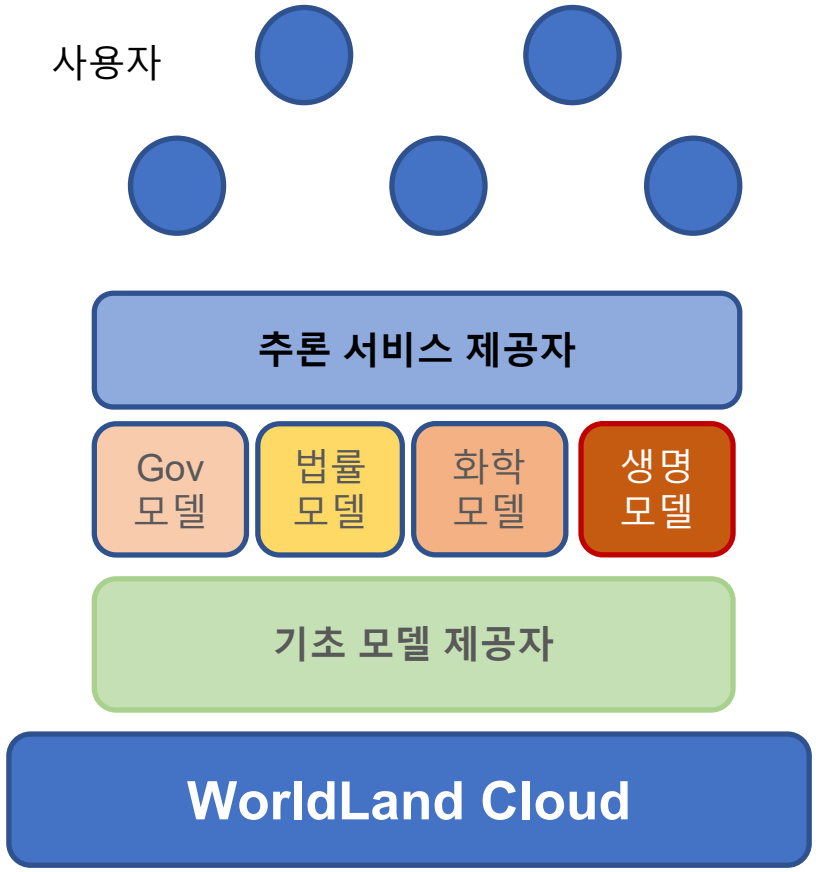
- AI-Nexus서비스 제공자는 퍼블릭 데이터 학습 완료한 OS pre-trained 기초모델 선택 제공
- 서비스 구매자는 도메인 데이터로 로칼서버에서 기초모델을 fine-tuning 한 후 완성 AI모델 확보
- 사용자는 서비스 제공자에게 WLC 지급



03. WorldLand

사용자(기업/연구소/정부)가 보유한 도메인 데이터로
안전하게 완성형 AI모델을 확보하는 방법은 없을까?

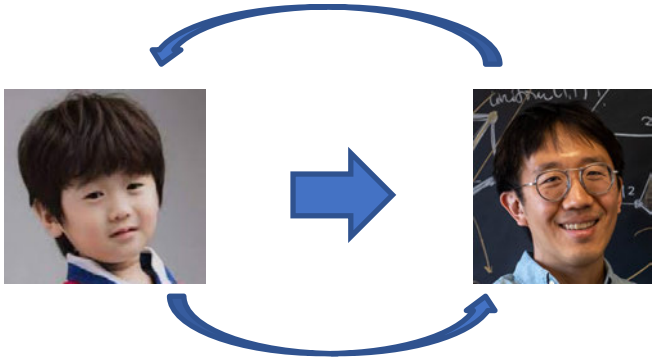
- 학습에 필요한 규모의 로컬 서버나 전문성을 확보할 수 없는 경우
  - 즉, 거대 모델을
  - 도메인데이터를 학습시킬
  - GPU 클라우드 확보가 어려운 경우
- 적정 규모 AI모델로 시작하고, AI Agent와 AI Agent간 통신을 통해 사도교류학습, 전이 학습 등으로 성장시킴



사람이 학습하는데 걸리는 시간

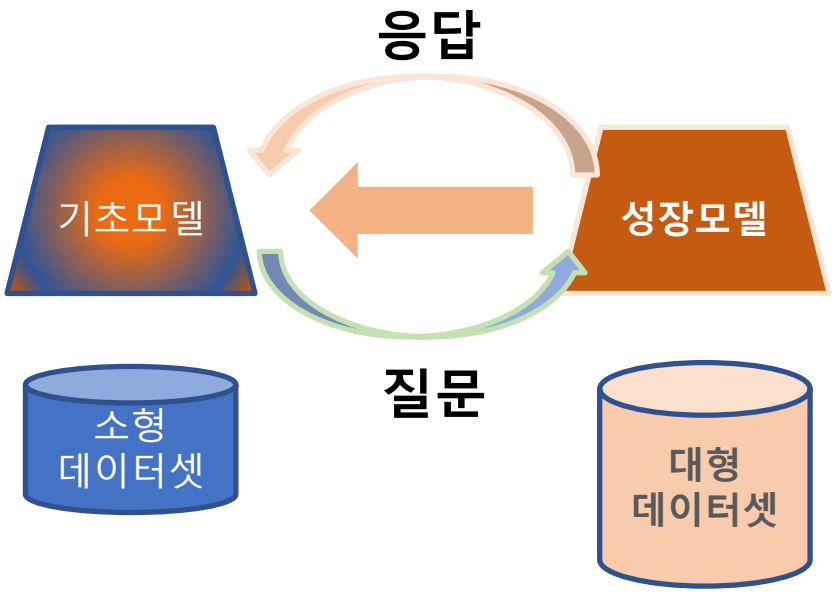
- 사람의 신경망은 서로 상이함
- 초등학생이 대학생으로 성장하는데 걸리는 시간 12년
- 박사로 성장하려면?

- AI Agent 학습의 경우는얼마나 걸릴까?



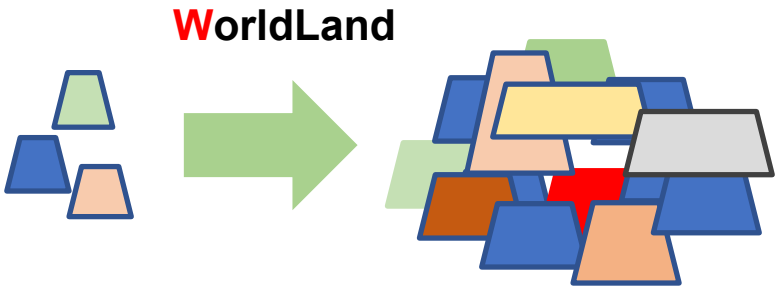
AI Agent 가 학습하는데 걸리는 시간은?

- 모든 디지털 두뇌 즉 신경망 구조는 유사함
- 선생/제자 사도교류 학습, 전이 학습, 연합학습 등은 속도는 매우 빠름
  - Emergent 통신
  - Emergent 언어 발현

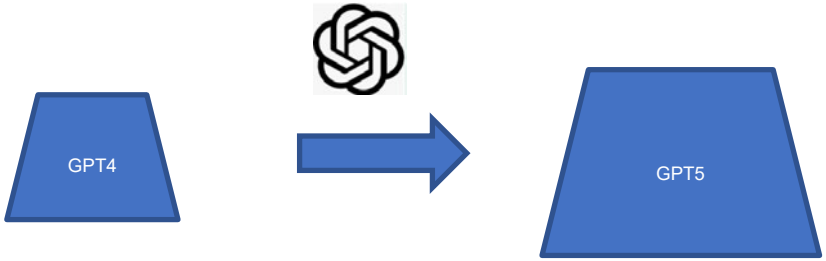


03. WorldLand

AI Nexus 와 OpenAI GPT, 누가 더 바르게 성장할까?







- WorldLand AI Nexus
  - 다양한 참여자가 다양한 지분을 갖고 직접 참여
  - 진짜데이터로 AI모델 학습
  - 학습결과물을 공유하고
  - 경제적 효과를 나누어 가짐

참여하고 협력하여 주도할 것인가?

- OpenAI
  - 천문학적 규모의 돈으로
  - 생성된 데이터로 GPT학습
  - 경제적 효과는 독점

AI Nexus 개요

- 상기 사업은 분산 연구 및 개발을 위한 Web3.0 활성화 거대 **블록체인-인공지능 융합 플랫폼**을 구축하는 것을 목표로 합니다.
- 초 거대 데이터셋을 수집하고, 비지도 학습, 지도 학습, 연합 학습, 강화 학습을 통해 AI 모듈을 학습시킵니다. 또한 **인간 라벨러가 라벨링 피드백을 제공합니다**.
- 본 플랫폼은 WorldLand **블록체인과 암호화폐 WLC를 기반으로하여** 다양한 참여자가 데이터의 분산 저장과 AI 솔루션을 제공하고, 플랫폼 통하여 연구 및 서비스 **개발 결과를 공유하도록 인센티브를 부여**합니다.
- 나아가, **사업자가 학습 완료된 AI 모듈 서비스를 대중에게 제공**하여 글로벌 사업을 영위할 수 있는 플랫폼을 제공합니다.
- AI Nexus 플랫폼은 인공지능 연구 및 개발에 중점을 두고 있지만, Web3.0 기술을 활용하여 법률, 의료진단 산업 등 다른 분야에도 적용될 수 있습니다.
- 전반적으로 Nexus AI 프로젝트는 분산적이고 투명하며 효율적인 인공지능 모델 학습과 연구개발 결과 공유를 위한 플랫폼을 제공하여 다양한 산업에 적용 가능한 혁신적인 결과를 도출하고자 합니다.

AI Nexus 적용 가능 분야

- Nexus AI 플랫폼은 다음과 같은 분야에서 적용될 수 있습니다:
- **인공지능 연구 및 개발**: Nexus AI는 인공지능 모델 학습 및 연구 결과 공유를 위한 플랫폼으로, 인공지능 연구 및 개발에 큰 도움이 될 것입니다.
- **의학**: Nexus AI는 질병 진단 및 치료 방법 개발, 신약 발굴 등 의학 분야에서도 활용될 수 있습니다.
- **화학**: Nexus AI는 새로운 물질 및 화합물 개발을 위한 분자 시뮬레이션 및 모델링에 활용될 수 있습니다.
- **게임 산업**: Nexus AI는 분산 게임 플랫폼과 플레이어 소유 게임 자산을 구현하는 데에 활용될 수 있습니다.
- **보안**: Nexus AI는 네트워크 보안 및 공격 차단에도 적용될 수 있습니다.
- 이 외에도, Nexus AI는 다양한 분야에서 적용 가능하며, 연구 결과에 따라 새로운 분야의 적용 가능성이 계속해서 발견될 수 있습니다.

법률GPT

- **법률 연구 및 분석**: AI Nexus는 대량의 법적 문서와 사례를 수집하여 분석하고, 법률 연구에 활용할 수 있는 지식 그래프나 예측 모델을 개발할 수 있습니다.
- **전자 검색 및 문서 분류**: AI Nexus를 사용하여 법적 문서의 효율적인 검색과 분류를 수행할 수 있으며, 속도와 정확성을 향상시킬 수 있습니다.
- **계약 분석 및 검증**: AI Nexus는 계약 문서를 분석하여 중요한 조항이나 위험 요소를 식별하고, 계약의 유효성과 규정 준수 여부를 검증할 수 있습니다.
- **판례 예측**: AI Nexus는 과거 판례와 법률 데이터를 기반으로 한 예측 모델을 개발하여, 특정 사건에 대한 판례 결과를 예측하고 법률 전략 수립에 도움을 줄 수 있습니다.
- **개인 정보 보호**: AI Nexus는 개인 정보 보호 및 데이터 처리 관련 법규를 분석하고, 기업이나 개인이 법적 요구 사항을 준수할 수 있는 지원을 제공할 수 있습니다.

➢ 이러한 방식으로 AI Nexus를 법률 분야에 적용함으로써, 변화하는 법적 환경에서 더 빠르고 정확한 결정을 내리고, 효율적인 리서치와 문서 처리를 수행할 수 있게 될 것입니다.

03. WorldLand

WorldLand 개요 (1).

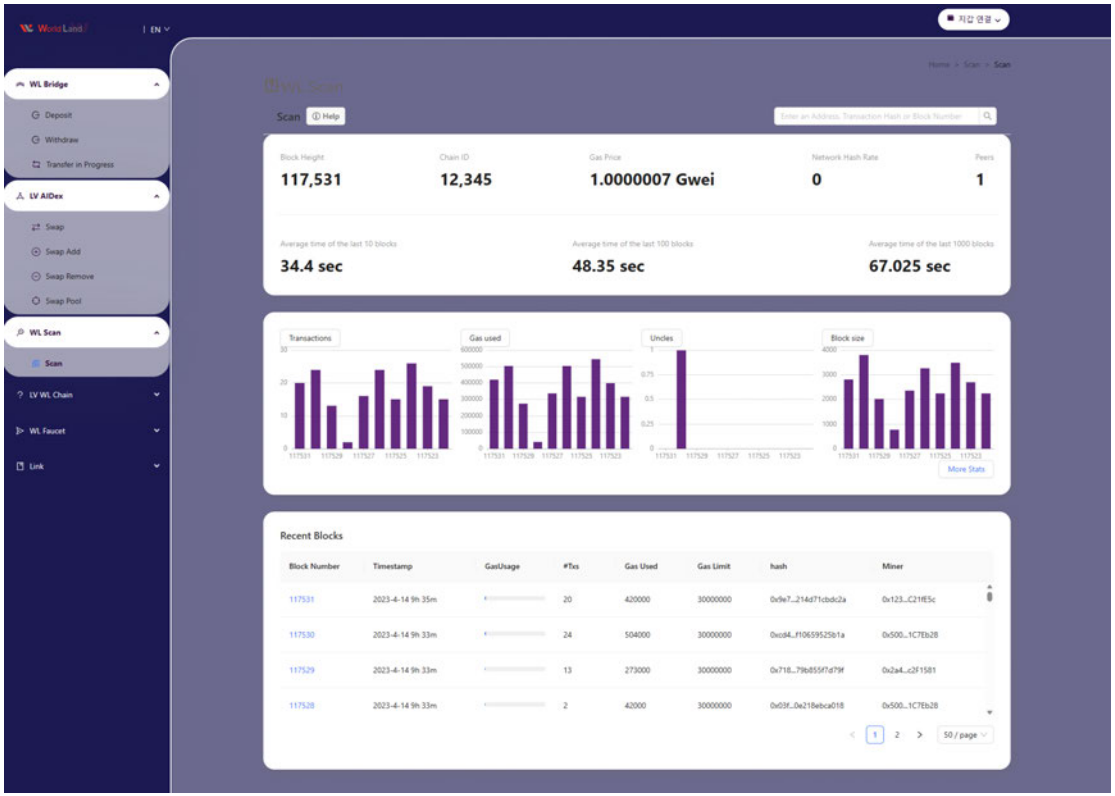
	내용		특이점
네트워크	코인(WLC)을 발행하는 메인넷		
용도	Payment currency		ZKP L2 AI Nexus
메인넷특징	EVM호환, Ethereum 사이드체인		양자내성암호 서명함수
주 개발사	리버밴스㈜		GIST교수창업기업
혁신기술	합의알고리즘	VCA + Green	블록생성주기 12 초 블록당 WLC 4 개 반감기 2년
	코인발행스케쥴	4번의 반감기 이후 매년 4%씩 증가	
	암호체계	양자내성암호 (합의, 서명, VCT, ZKP)	

WorldLand 개요 (2).

	내용	특이점
거버넌스	DAO 3년 이내 탈 중앙 거버넌스 완료	
토큰세일	DAO 토큰 WLT 발행 메인넷 제네시스 블록에서 WLT만큼 WLC를 발행 지급	WLC는 DAO의 투표권으로 활요
D-Date 론칭 및 상장 계획	D-Date: 2023년 3분기 1분기: Testnet 1 2분기: Testnet 2 3분기: 메인넷 글로벌 론칭 3분기: CEX/DEX상장	Testnet 운영중 (’19)
프로젝트 사이트	https://heungno.net	

WorldLand 메인넷 개발 현황 (TRL 9 사업화 단계)

- WorldLand Testnet
- WorldLand Bridget
- WorldLand Scan
- WorldLand Faucet
- AI-DEX
- Visit home page
  - LiberVance.net
  - https://heungno.net/worldland



WorldLand AI Nexus 노드

	분류	응용분야	수익구조
WorldLand AI Nexus 노드	EVM 서비스	DeFi, NFT, ERC-20, Bridges	생태계 Token 지급
		Soul Bound Tokens	
		L2: Zero-Knowledge Proof (검증부분)	
		AI-Nexus (Contract 부분) 서비스 요청 및 수수료 계약	
	비EVM 서비스	Consensus Green VCA	WLC 주조권
		Zero-Knowledge Proof (연산 증명 부분)	TX 수수료
		AI Nexus (컴퓨팅 부분)	계약 수익금

WorldLand 클라우드 서버 레이어 운용 방안

1. 참가 노드 등록 (Cloud Portal)

각 참가자(그래픽 처리 장치(GPU) 컴퓨터 소유자)는 네트워크에 자신의 장치를 등록. 등록 과정에서 시스템은 작업 스케줄링 및 로드 밸런싱을 위해 GPU 사양을 수집: GPU 모델 (예: NVIDIA A100, H100, GeForce, RTX), GPU 메모리, GPU 코어 수, 네트워크 속도, 사용 가능한 저장 공간 등 등록.

2. 작업 제출 (Smart Contract)

AI 모델을 훈련시키려는 사용자는 웹 포털이나 API를 통해 작업을 제출. 작업 제출에는 다음이 포함: 훈련시킬 AI 모델, 사용할 데이터셋, 훈련 매개변수, 원하는 GPU 사양(있는 경우), 작업 우선 순위 조정.

3. 작업 스케줄링 및 로드 밸런싱: (노드)

작업 스케줄러는 작업 제출을 받고, GPU 사양, 현재 로드, 작업 우선 순위에 따라 작업을 어디에 할당할지 결정.

4. 작업 실행: (노드)

작업 스케줄러가 GPU에 작업을 할당하면, 참가자의 GPU 컴퓨터에 설치된 작업자 에이전트는 작업 세부 정보를 받고 작업을 시작. 작업자 에이전트는 또한 작업 상태와 GPU 사용량에 대한 정기적인 업데이트를 작업 스케줄러에게 보냄.

5. 작업 완료 및 결과 검색: (Cloud)

작업이 완료되면, 결과는 패키지화되어 사용자에게 돌려짐. 결과는 웹 포털을 통해 다운로드 가능하게 하거나, API를 통해 검색할 수 있음.

6. 청구 및 지불: (Smart Contract)

청구 시스템은 GPU 사용량과 작업 기간에 따라 작업의 비용을 계산합니다. 사용자는 이에 따라 청구되며, 사용된 GPU의 소유자는 지불의 일부를 받습니다.

7. 보안 (계속): (WorldLand 보호 노드)

사용자와의 모든 통신은 암호화됩니다. 또한, 작업자 에이전트는 무단 접근으로부터 사용자의 데이터를 보호하기 위해 격리된 환경에서 실행.

8. 모니터링 및 유지 보수: (Cloud Monitoring)

모니터링 시스템은 참가자의 GPU와 작업 스케줄러의 상태를 지속적으로 확인. GPU가 오프라인 상태가 되거나 작업이 실행되지 않는 등의 문제가 발견되면, 시스템은 자동으로 시스템 관리자에게 알림.

9. 확장성:

시스템이 많은 수의 참가자와 작업을 처리할 수 있도록, 작업 스케줄러, 데이터베이스, 웹 포털/API는 수평 확장이 가능하게 설계. 가령 AWS EC2와 같은 컴퓨팅 리소스, AWS RDS와 같은 데이터베이스, AWS Elastic Load Balancer와 같은 로드 밸런싱을 사용함으로써 달성 가능.

10. 사용자 인터페이스:

작업 제출자와 GPU 소유자 모두를 위한 사용자 친화적인 인터페이스를 설계. 이 인터페이스는 작업 상태, GPU 사용량, 청구 정보, 시스템 상태에 대한 정보를 제공.

Proof of Useful Work를 하는 노드는 SBT 지급
SBT 양에 따라 동전던지기 함수에서 선택될 가능성을 제고함.

03. WorldLand

목표시장 및 경쟁사 현황 -요약

경쟁사 대비 자체 플랫폼, 복합체인 연결기술, Anti-ASIC 및 양자내성 특성 등, 기술적 우위 확보

기술적 우위에 의한 시가총액 추월 목표

	WorldLand	Decentraland	Stacks	SOLANA	COSMOS ATOM	Polkadot.	polygon
상호	WorldLand	Decentraland	Stacks	Solana	Cosmos	Polkadot	**Polygon**
창립일	2023 년	2017 년	2013 년	2017 년	2017 년	2017 년	**2017 년**
시가 총액	N/A	약 6.5 조 원	약 3.2 조 원	약 55.6 조 원	약 10.7 조 원	약 31.6 조 원	**약 20.2 조 원**
주요 서비스	AI NEXUS	메타버스, NFT	플랫폼 (NFT, DID, Web3)	플랫폼 (NFT, DeFi)	상호운용 플랫폼	상호운용 플랫폼	**오프체인 거래 지원 Layer-2 솔루션**
경쟁사 대비 자사의 비교우위 점	자체 플랫폼	자체 플랫폼 없음	멀티체인 구현 x		합의 알고리즘 Anti-ASIC 및 양자 내성을 통한 신뢰 검증 기능 부족		

- **Green VCA 합의**
- **AI NEXUS**
- **MAC-ZKP**

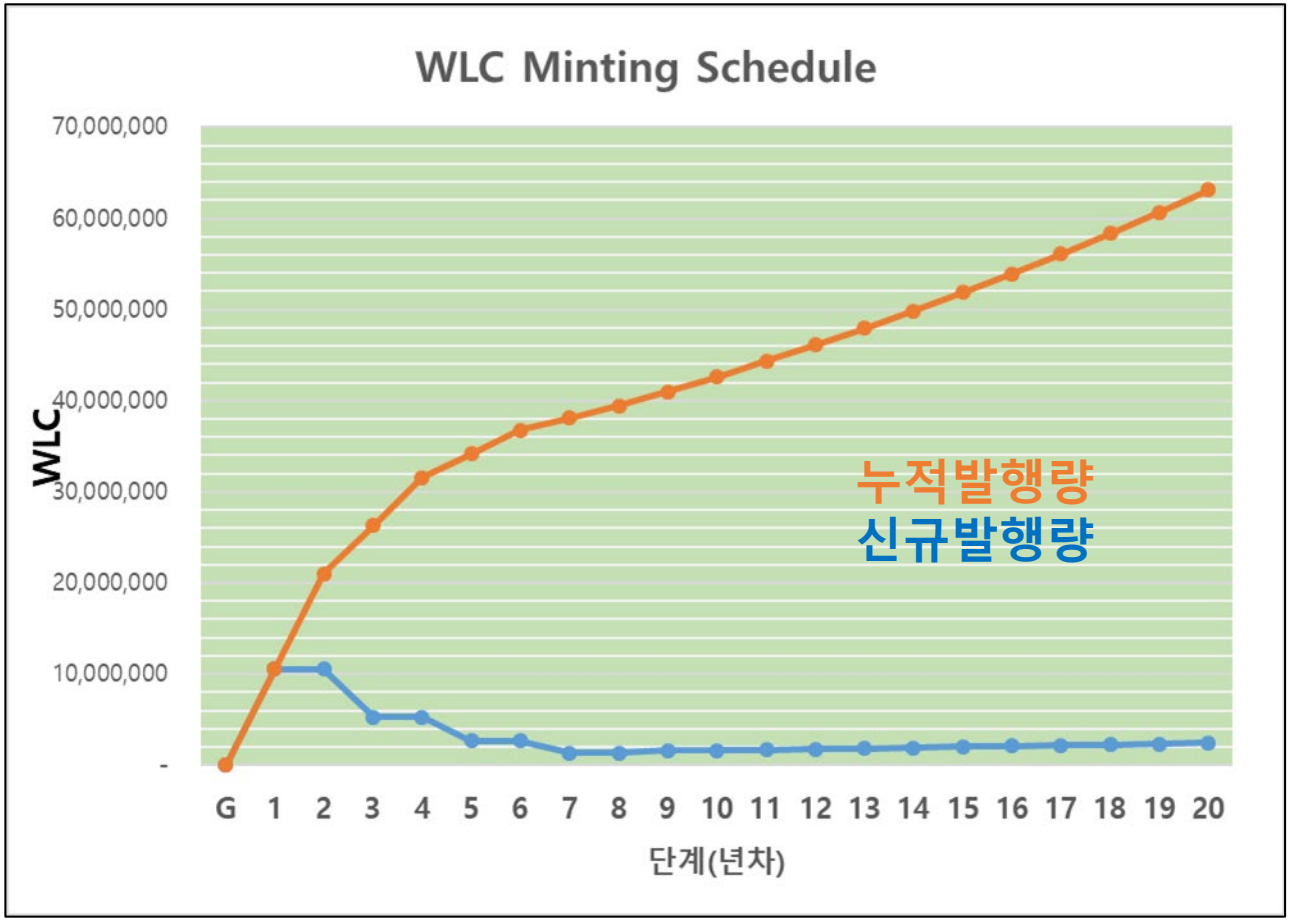
03. WorldLand

WorldLand 화폐 발행 스케쥴은 어떤 것인가요?
4번의 반감기를 거쳐 안정화 단층에 이르고, 그 후로는 매년 4%씩 추가 발행합니다.

❑ **WorldLand 경제 규모 매년 4% 예측 가능하게 성장**

❑ **과도한 호들링(Hodling) 방지**

❑ **화폐 사용 및 투자 촉진**



04. 양자 내성 암호

WorldLand는 양자 내성 암호를 사용합니다.

- Elliptic curve cryptography, RSA
- 양자 컴퓨터의 특성
  - 1995 Shor: **소인수분해 문제**와 **이산대수문제**를 다항식 시간 안에 풀어줄 수 있는 양자 알고리즘을 제안

    https://arxiv.org/abs/quant-ph/9508027

    [Submitted on 30 Aug 1995 (v1), last revised 25 Jan 1996 (this version, v2)]

    Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer

    Peter W. Shor (AT&T Research)

  - 양자컴퓨터(1000 큐빗 이상)가 개발시 현재 사용하고 있는 공개키 암호 시스템은 붕괴.

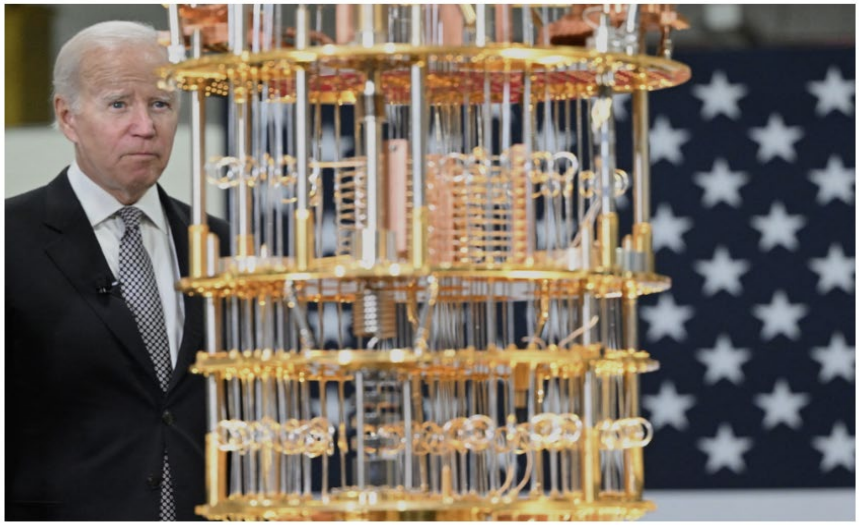
이산대수: Discrete logarithms
소인수분해: Prime Factorization

ARGUMENT *An expert's point of view on a current event.*

The U.S. Wants to Make Sure China Can't Catch Up on Quantum Computing

Washington is likely to impose new controls in the race for a key technology.

By Kevin Klyman, a technology researcher at Harvard's Belfer Center.

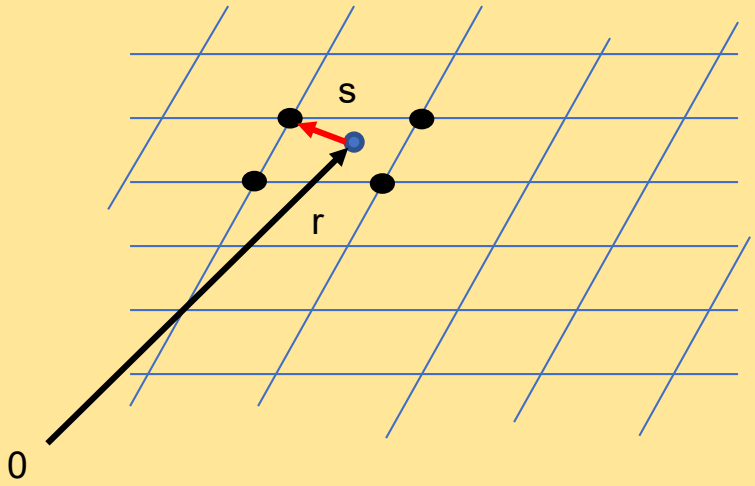


U.S. President Joe Biden looks at a quantum computer as he tours an IBM facility in Poughkeepsie, New York on Oct. 6, 2022. MANDEL NGAN/AFP VIA GETTY IMAGES

양자 내성 암호 개발 동향

- NIST Post-Quantum Cryptography Contests
  - Falcon, Dillithium, Sphincs
    https://csrc.nist.gov/News/2022/pqc-candidates-to-be-standardized-and-round-4
- 격자 부호 (lattice code)
- 랜덤 오류 정정 부호 (error-correction code)
  - Random keys, **AB** = 0.
  - Find a short vector **As** = **Hash(m)** around r.
  - Shortest Integer Solution

서명: m → (r, s)
확인: r+s가 격자점인가?
s가 짧은벡터인가?



04. 양자내성암호: MAC-ZKP

WorldLand 의 합의알고리즘 ECCPoW과 영지식증명 MAC-ZKP 는 양자내성 암호입니다.

- **WorldLand** ECCPoW와 MAC-ZKP는 격자부호
- 가장 근접한 랜덤 쇼트벡터 혹은 격자부호점을 찾는 문제
- 양자 내성 확보

ECCPoW

$$\mathbf{Ae} = H(BH)$$

합의알고리즘: 수 백만대의 컴퓨터가 경쟁적으로 매우 어려운 양자내성암호를 푸는 문제

MAC-ZKP

$$\mathbf{A}(y + r) = H(m)$$

ZKP: 비밀을 알고 있는 하나의 컴퓨터가 어려운 양자내성암호를 푸는 문제

04. 양자 내성 암호: 합의알고리즘

Green VCA
합의알고리즘 논문

Vol. 1, No. 1, October 2022 THE JOURNAL OF DIGITAL ASSETS 33

Green Bitcoin: Global Sound Money

Heung-No Lee, Young-Sik Kim, Dilbag Singh, and Manjit Kaur

Abstract

Modern societies have adopted government-issued fiat currencies many of which exist today mainly in the form of digits in credit and bank accounts. Fiat currencies are controlled by central banks for economic stimulation and stabilization. Boom-and-bust cycles are created. The volatility of the cycle has become increasingly extreme. Social inequality due to the concentration of wealth is prevalent worldwide. As such, restoring sound money, which provides stored value over time, has become a pressing issue. Currently, cryptocurrencies such as Bitcoin are in their infancy and may someday qualify as sound money. Bitcoin today is considered as a digital asset for storing value. But Bitcoin has problems. The first issue of the current Bitcoin network is its high energy consumption consensus mechanism. The second is the cryptographic primitives which are unsafe against post-quantum (PQ) attacks. We aim to propose Green Bitcoin which addresses both issues. To save energy in consensus mechanism, we introduce a post-quantum secure (self-election) verifiable coin-toss function and novel PQ secure proof-of-computation primitives. It is expected to reduce the rate of energy consumption more than 90 percent of the current Bitcoin network. The elliptic curve cryptography will be replaced with PQ-safe versions. The Green Bitcoin protocol will help Bitcoin evolve into a post-quantum secure network. In addition, it improves the properties of Bitcoin's hash PoW while addressing environmental concerns.

Keywords: Bitcoin, energy consumption, error-correction codes, post-quantum security, sound money, verifiable random function

I. INTRODUCTION

WE the global citizens not by our choice live in a world of boom-and-bust cycles created by the Federal Reserve Board (FED) of the United States. In the boom phase of a cycle, the FED supplies debt-monetized US dollars to the world, and supplied USDs are used to purchase real items, goods, and services from

Time-Variant Proof-of-Work Using Error-Correction Codes

Sangjun Park, Haeung Choi, and *Heung-No Lee, *Senior Member, IEEE*

***Abstract*— The protocol for cryptocurrencies can be divided into three parts, namely consensus, wallet, and networking overlay. The aim of the consensus part is to bring trustless rational peer-to-peer nodes to an agreement to the current status of the blockchain. The status must be updated through valid transactions. A proof-of-work (PoW) based consensus mechanism has been proven to be secure and robust owing to its simple rule and has served as a firm foundation for cryptocurrencies such as Bitcoin and Ethereum. Specialized mining devices have emerged, as rational miners aim to maximize profit, and caused two problems: *i*) the re-centralization of a mining market and *ii*) the huge energy spending in mining. In this paper, we aim to propose a new PoW called Error-Correction Codes PoW (ECCPoW) where the error-correction codes and their decoder can be utilized for PoW. In ECCPoW, puzzles can be intentionally generated to vary from block to block, leading to a time-variant puzzle generation mechanism. This mechanism is useful in repressing the emergence of the specialized mining devices. It can serve as a solution to the two problems of recentralization and energy spending.**

***Index Terms*— Consensus, Cryptocurrency, Blockchain, Proof-**

If a node was re-forging all the blocks alone, it could spend the total amount of works done to all the mined blocks.

Without PoW, anybody with a computer can alter the content of the blockchain, implying unauthorized changes in any mined blocks can be possible. If PoW is attached to each mined block, attackers cannot make any unauthorized modifications without redoing all the works. No node can alone alter any mined block, meaning an immutability property.

In Bitcoin, miners make rational decisions to maximize their profits by following a two stage process in which *i*) the miners select a blockchain whose length is the longest and *ii*) they extend this longest one by adding a newly mined block. Suppose there are two blockchains where one is longer than the other one in terms of the length. Since the longer chain has the more accumulated works, altering it is more difficult. This longer chain shall be treated the more trustable and preferable by the miners. Thus, they select the longer chain. Making such a selection is rational for the sake of keeping the mining re-

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 25, 2021, accepted September 7, 2021, date of publication September 16, 2021, date of current version October 11, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3113522

Error-Correction Code Proof-of-Work on Ethereum

HYOUNGSUNG KIM[1], JEHYUK JANG[1], SANGJUN PARK[2], AND HEUNG-NO LEE[1], (Senior Member, IEEE)

[1]School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea
[2]Electronics and Telecommunications Research Institute (ETRI), Gwangju 500-712, South Korea

Corresponding author: Heung-No Lee (heungno@gist.ac.kr)

This research was supported in part by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2021-0-01835) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation), in the part by the IITP Grant through Korean Government MSIT under Grant 2020-0-00958, and in part by the National Research Foundation of Korea (NRF) Grant through Korean Government MSIP under Grant NRF-2021R1A2B5B03002118.

ABSTRACT The error-correction code proof-of-work (ECCPoW) algorithm is based on a low-density parity-check (LDPC) code. ECCPoW can impede the advent of mining application-specific integrated circuits (ASICs) with its time-varying puzzle generation capability. Previous research studies on ECCPoW algorithm have presented its theory and implementation on Bitcoin. In this study, we have not only designed ECCPoW for Ethereum, called ETH-ECC, but have also implemented, simulated, and validated it. In the implementation, we have explained how ECCPoW algorithm has been integrated into Ethereum 1.0 as a

05. 양자 내성 암호: 영지식 증명

영지식 증명 시스템

- **검증자**는 증명자에게 질문/문제/명제를 제시하고 답과 증명을 요구
- **증명자**는 답과 증명을 검증, 명제 수행 완전성을 검증하는 시스템
- 가령 “증명자는 비공개 입력값 w과, 주어진 입력 값 m, 주어진 함수 f에 대하여 결과값 y = f(m, w)를 정직하게 계산하였다. ” 와 같은 명제이다.
- 대표적인 예는 개인인증 시스템임
  - w는 개인 정보
  - y 는 결과
  - f는 프로그램
- **증명자는 주어진 문제를 푼 후 결과값과 증거를 생성**, 검증자에게 전달
- **검증자는 증거를 검증함으로써 명제의 옳고 그름을 판단할 수 있음**
- 단, **비밀정보는 검증 과정에서 절대로 누출되지 않아야 함**.

05. 양자 내성 암호: 영지식 증명

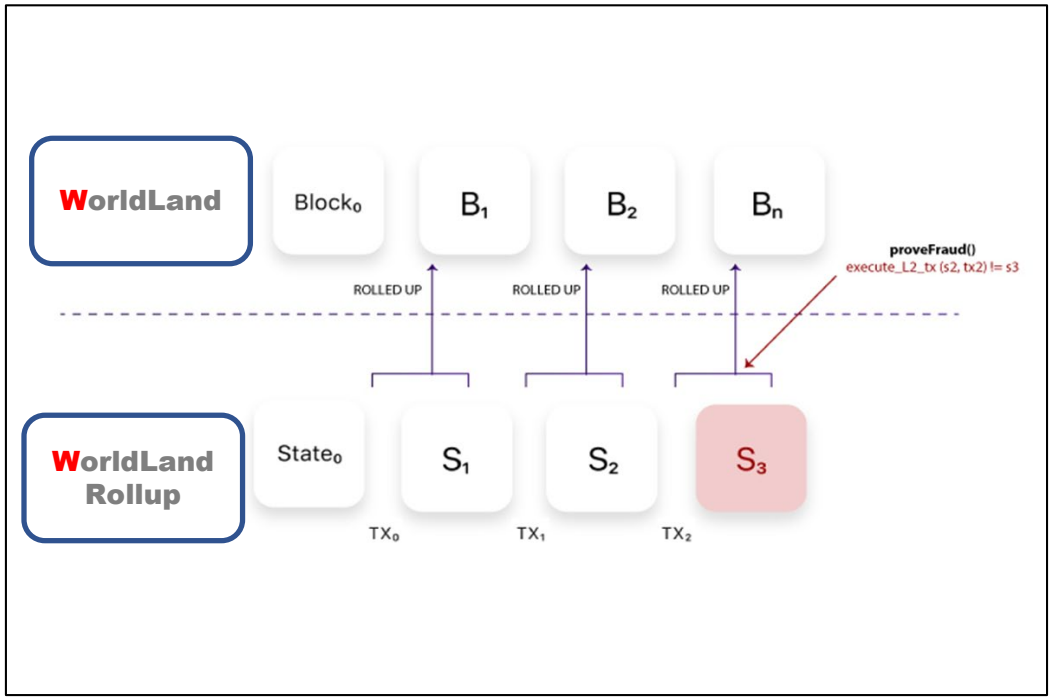
ZKP Use Cases

- 나이를 밝히지 않고 내가 미성년자가 아님 증명하기
- 온라인 무기명 투표
  - 투표자: 내 투표가 집계에 정확히 반영되었는지 확인하고자 함
  - 선관위: 모든 투표지를 정확히 반영하여 집계하였음을 증명
  - 선관위의 증명이 집계정확성 외외에 다른 어떤 정보 (투표자가 어느 후보자에게 투표했는지)도 유출하지 않음
- DNA Profile Match Problem
  - Andy: 마약 복용 의심 받는 대통령 후보자
  - FBI: 마약 복용자 DNA DB 보유, Andy's DNA 수신
  - 국민: Andy's DNA가 FBI DB에 존재하는가?
  - FBI DB와 Andy' DNA 등의 개인 정보는 국민에게 비누출

05. 양자 내성 암호: 영지식 증명

WorldLand & ZK-Rollup

- ZK-Rollup
  - L2 스케일링
  - Faster blocks
  - 무합의
  - y = f(m, w)
    - m = (S_2, TX_2)
    - w = priKey



05. 양자 내성 암호: 영지식 증명

영지식 증명은 증명자는 연산을 많이하고, 검증자는 쉽게 검증하는 시스템이다.

- Statement: y = f(x, w)
- Prover: 증명자
- Verifier: 검증자

- Completeness: Pr{정직증명 & +검증} <= 1 – e_c.
- Soundness: Pr{거짓증명 & +검증} <= e_s.
- Zero-Knowledge: Pr{witness revealed}<= e_{zk}.

05. 양자 내성 암호: 영지식 증명

우수한 ZKP 시스템이란 무엇인가?

- T_c : 연산 수행 소요 시간
- T_p : 연산 수행과 증거 생성 소요 시간, $T_p > T_c$.
- T_v : 연산결과와 증거를 대조 검증 소요 시간, $T_v << T_c$.
- n : ZKP 시스템 차원
  - 예를 들면 데이터베이스의 n 개의 프로파일 엔트리.
  - Zk-SNARK [Groth17] 경우 O(n) multiplication circuits.

05. 양자 내성 암호: 영지식 증명

영지식 증명 연구 동향

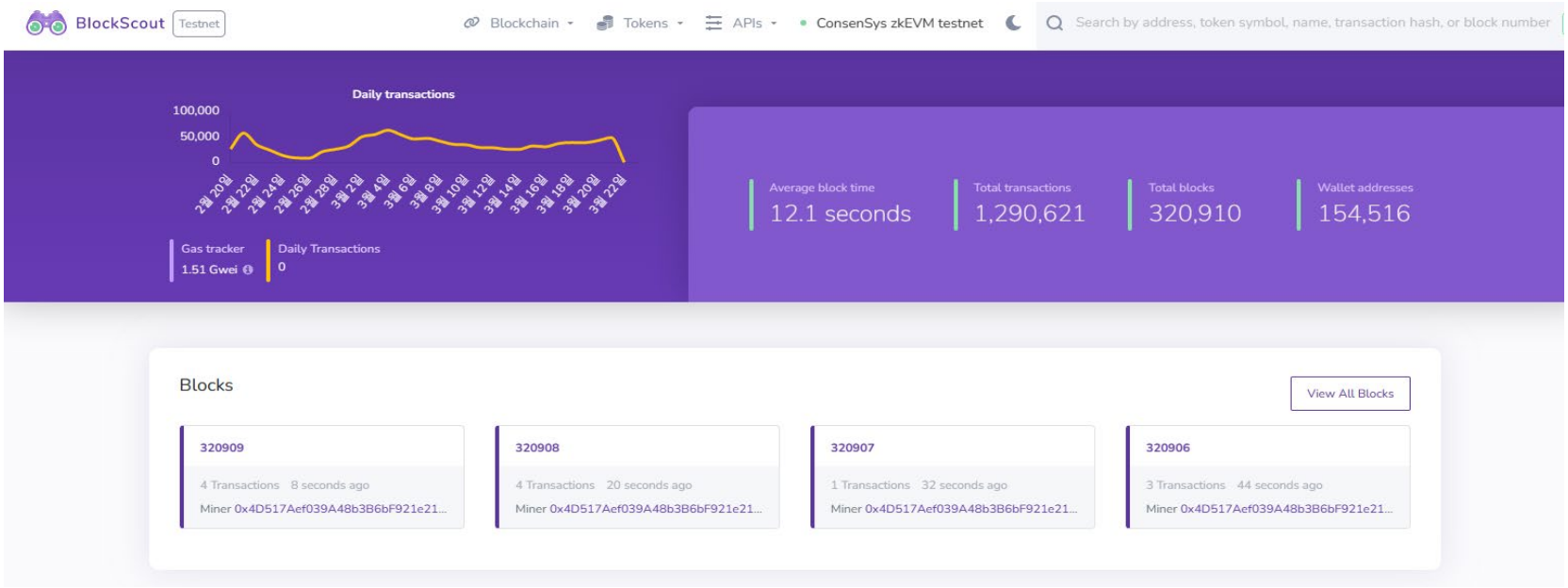
- 대화형 (1985)
- 비대화형 (1993)
- ZK PCPs (2001)
  - 복잡한 연산 검증 개념 추가
- Zk-SNARK (2012)
  - Zokrates (Remix tool)
- Zcash (2016)
  - Zk-SNARK, private거래에 사용
- ZK-STARK/Orion/lattice codes (2018 ~)
  - 증거 크기가 작고, 증명과 검증이 빠르며, 견고한 시스템

Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof-systems. Proceedings of 17th Symposium on the Theory of Computation, Providence, Rhode Island. 1985. Draft. Extended abstract

최신 연구 동향은
연산 속도를 최적화 위해
그래프 기반 랜덤 코드로
옮겨가고 있음

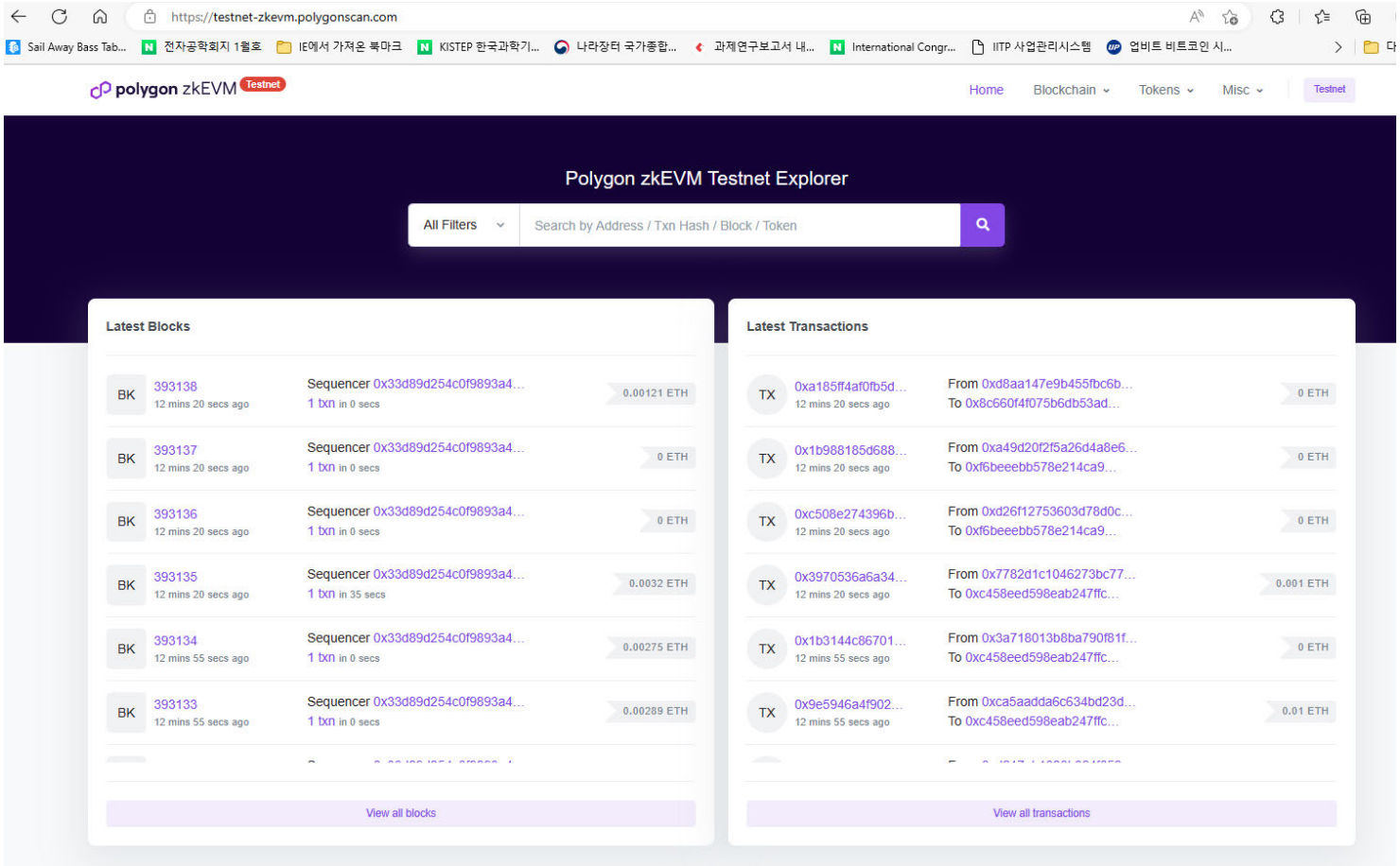
ConsenSys zk EVM

- ConsenSys will open its EVM-equivalent zkEVM rollup to public testnet on March 28th, 2023.
- Our innovative lattice-based prover offers leading performance in proof generation, resulting in cheap transaction fees. ConsenSys zkEVM offers fast finality, high throughput, and the security of Ethereum settlement.



05. 영지식 증명 동향

Polygon zkEVM (23년 2월 오픈)



Zk-EVM 베타 출시 경쟁, 확장성 제고 능력은 아직 검증 중

- **폴리곤, ‘zkEVM’ 메인넷 베타 버전 출시**

  2023/03/29글쓴이: W, Lynn

- 이더리움 레이어2 확장 솔루션 폴리곤(MATIC)이 영지식 이더리움 가상머신(zkEVM) 베타 메인넷 버전을 27일 오전 출시했다. zkEVM은 영지식증명 기술을 사용하는 새로운 확장성 기술로 메인넷 출시 시 이더리움의 거래 비용과 속도를 개선, 궁극적인 온체인 프라이버시 구현도 가능하다.
- **폴리곤, zkEVM 메인넷 베타 버전 공개**
- 코인데스크에 따르면 폴리곤의 zkEVM 출시는 경쟁사인 매터랩스가 자체 zkEVM인 zk싱크 에라를 출시한 지 며칠 만에 이뤄졌다. 폴리곤 측은 “개발자들이 코드를 연구하고 공유하여 영지식 연구에 기여할 수 있도록 zkEVM 관련 기술을 오픈 소스로 공개하고 있다”고 밝혔다.
- 폴리곤 공동 창립자 샌디프 네일월은 코인데스크와 인터뷰에서 “우리는 계속해서 **웹3 정신을 추구할** 것”이라며 “많은 도전을 이어 나갈가겠다”고 말했다. 그는 “zkEVM이 공개됐지만 새로운 기술이기 때문에 ‘베타 메인넷’이라고 부른다”며 “투자자들은 기술을 이해한 뒤 우리에게 투자해도 늦지 않는다”고 강조했다.
- 폴리곤 재단은 이는 **오픈소스로 배포한다고** 밝혔다. 향후 소스코드가 수정된 버전도 오픈소스로 배포되도록 라이선스를 규정해 **특정 주체의 독점적인 사용을 방지했다**.

장점: 오픈 소스로 공유

단점: TPS제고 미흡

시사점: zk-SNARK기술 기반, TPS제고 위해 추가 연구 필요 할 듯

새로운 ZKP 연구개발 추진 필요

06. WorldLand MAC-ZKP

Novel MAC-ZKP

- 주어진 문제 y = f(m, w)
  - y_m = f(m, 0) 문제와 y_w = f(0, w)문제로 분리
  - 즉, 2 사용자 MAC 통신 채널 문제로 모델링
  - 두 개의 polynomial 혹은 graph code로 모델링
  - 두 개의 코드 구조를 parity-check으로 결합함
- 핵심 새로운 아이디어
  - 증거 부분은 code rate을 capacity 보다 크게
  - 메시지 부분은 code rate을 capacity 보다 매우 작게
  - 거짓증거를 사용하면 문제가 전혀풀수없고, 참 증거를 사용했을 때에만 문제가 풀리게
  - 문제가 풀렸을때 거짓 메세지를 사용하면 Verifier가 간파할 수 있게 설계한다.
- 목표:
  - 정보이론 활용 매우 간단하고 명확한 이론
  - Zk-STARK 장점 계승, 증명과 검증이 빠르고(10배) 알고리즘

06. WorldLand MAC-ZKP

Spatially Concatenated Channel-Network Code for Underwater Wireless Sensor Networks

Zafar Iqbal and Heung-No Lee, *Senior Member, IEEE*

***Abstract*— Underwater environment monitoring is an important application of wireless sensor networks (WSNs). However, WSNs face challenges, such as erroneous communication, ensuring the lifetime and robustness of the network, and cost constraints. The underwater acoustic channel (UAC) is highly frequency-selective, and the channel response changes over time because of variations in the channel conditions. Therefore, designing a cooperative coded orthogonal frequency division multiplexing (COFDM) system that is suitable for the doubly selective UAC and has reduced power consumption is very challenging. We propose a cooperative spatial-domain coding scheme combined with the low-density parity-check-coded OFDM system, called spatially concatenated channel-network code, for underwater acoustic WSNs. The designed underwater acoustic WSN exhibits a significant advantage regarding the required number of sensors, bit error rate (BER), and power consumption over the non-cooperative COFDM communication system. We also analyze sensor deployment schemes and find out the area in which our proposed scheme can be beneficial in terms of reduced power consumption and enhanced BER.**

***Index Terms*— Underwater acoustic communication, OFDM, LDPC, multipath fading, network coding, deployment, WSN.**

at the transmitter and/or receiver. Furthermore, the multipath delay spread due to the reflections at the sea surface and bottom causes inter-symbol interference (ISI) and frequency-selective fading. Hence, these factors degrade the system performance [4], [5].

To overcome such performance falloffs, coded orthogonal frequency division multiplexing (COFDM) systems have been proposed that employ low-density parity-check (LDPC) codes [6]–[8], Reed-Solomon codes [9], and adaptive modulation and coding (AMC) [10] in OFDM systems for UACs. LDPC-coded and OFDM-based underwater acoustic communication has been well investigated in [6]–[8] and [11]–[13], respectively. However, we observe that a non-cooperative LDPC-COFDM system exhibits significant performance degradation in the presence of random fading. Also, because point-to-point systems are vulnerable to long-term deep fading, Doppler spread, and shadow zones [1], [14], [15], the interest in designing cooperative communication systems with network coding, has recently increased.

Z. Iqbal and H.-N. Lee, "Spatially concatenated channel-network code for underwater wireless sensor networks," IEEE Transactions on Communications, Vol. 64, No. 9, pp. 3901-3914, Sep. 2016.
https://heungno.net/wp-content/uploads/2012/08/07518674_SCCNC_published.pdf

Novel MAC-ZKP: 그래프 코드의 모습

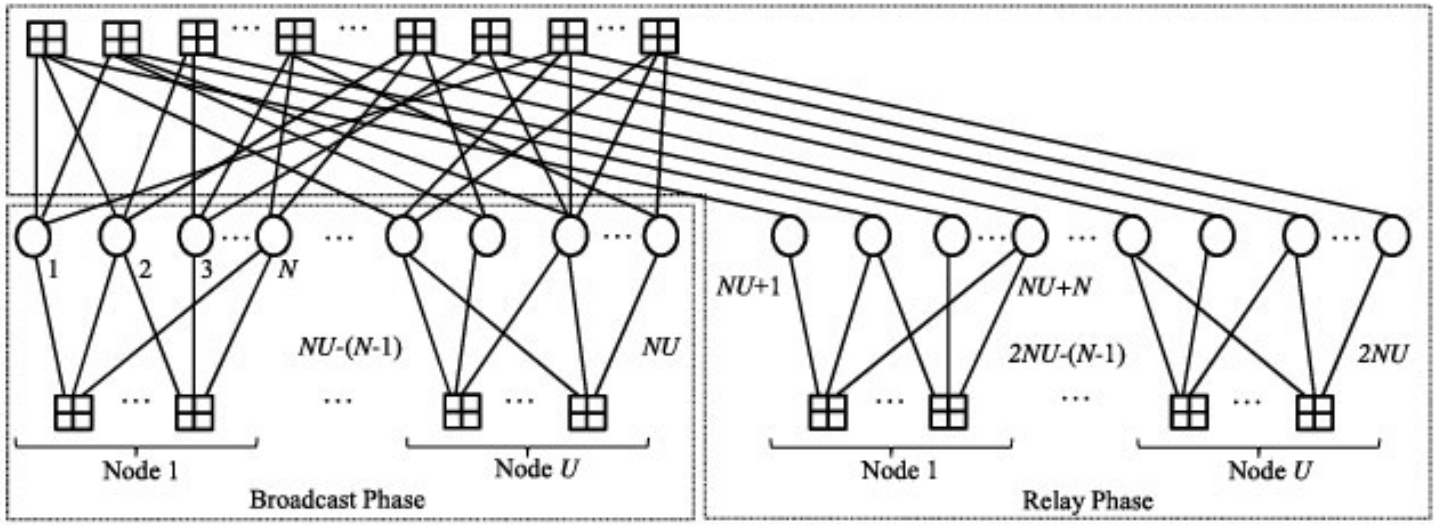
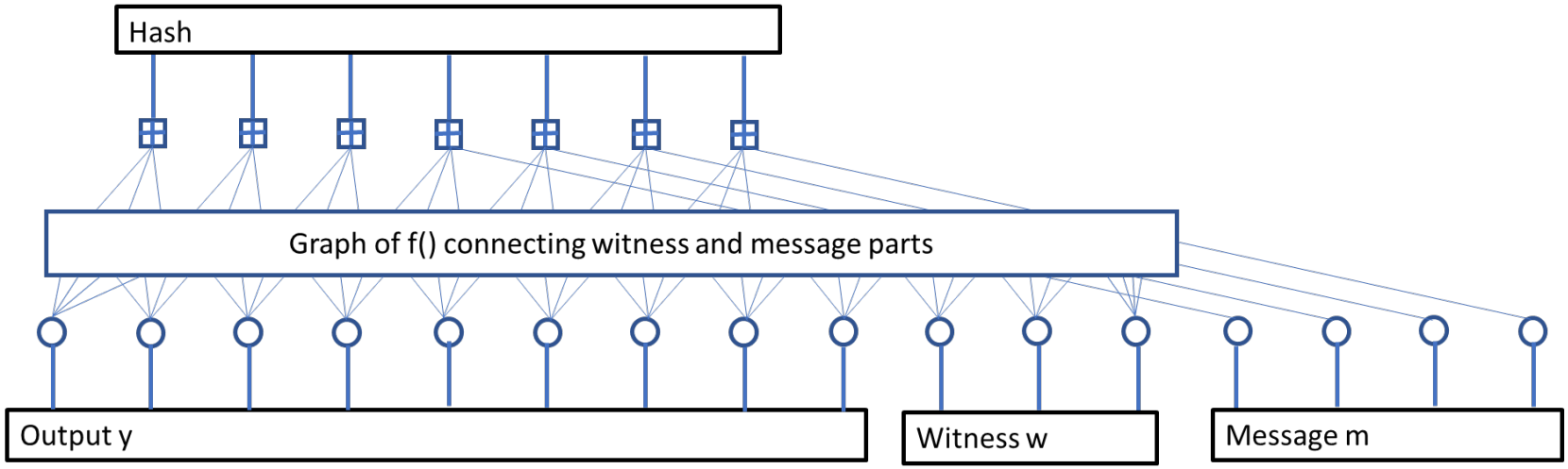


Fig. 6. Example of the SCCNC scheme: each sensor sends an LDPC codeword in the first phase. The spatial-domain checksums are computed and sent during the second phase, which are the LDPC coded symbols received during the first phase.

Z. Iqbal and H.-N. Lee, "Spatially concatenated channel-network code for underwater wireless sensor networks," IEEE Transactions on Communications, Vol. 64, No. 9, pp. 3901-3914, Sep. 2016.
https://heungno.net/wp-content/uploads/2012/08/07518674_SCCNC_published.pdf

MAC-ZKP 개요도



f(w, m)=y의 graph적 표현

y ~ 상태변화값

w ~ secret (비밀정보)

m ~ public state0, TXs

06. WorldLand MAC-ZKP

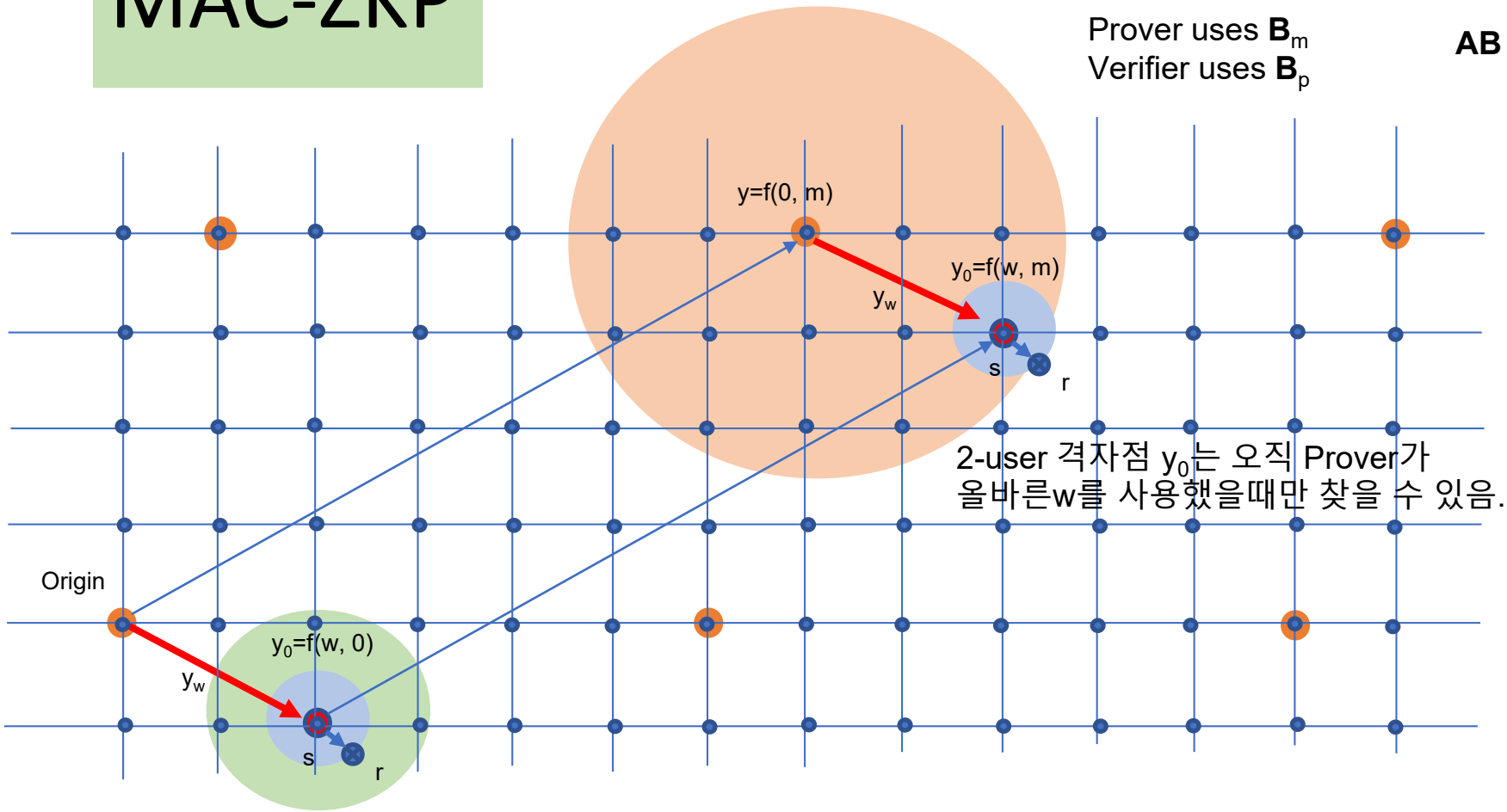
MAC-ZKP

f(0, m) 코드워드

f(w, m) 코드워드

Prover uses $\mathbf{B}_m$
Verifier uses $\mathbf{B}_p$

AB = 0



2-user 격자점 y_0는 오직 Prover가 올바른w를 사용했을때만 찾을 수 있음.

w를 모르면 r로 부터 정답격자점을 찾을 수 없음

Summary of MAC-ZKP Benefits

WorldLand에서는 세계 최초의 MAC-ZKP 알고리즘을 개발하고 Rollup에 적용하는 것을 목표로 합니다. 한 가지 시작 아이디어는 다중 센서 공간 주파수 변조(SF) 코딩 방식[RM1]을 사용하는 것입니다. 다른 접근 방식은 Reed-Solomon 코딩 체계를 사용하는 것입니다. 두 방식 모두 격자 부호 이론을 사용합니다. 따라서 MAC-ZKP는 격자 기반의 ZK 시스템으로 분류할 수 있습니다.

다음은 SF 코딩 체계가 제공하는 몇 가지 주요 속성입니다.

- 효율성: SF 코딩 방식은 계산 및 검증 시간 측면에서 매우 효율적으로 설계되었습니다. 따라서 계산 리소스가 제한된 EVM에서 사용하기에 적합합니다.
- 프라이버시: SF 코딩 체계는 영지식 증명입니다. 즉, 계산에 사용된 입력 또는 중간 값에 대한 정보를 공개하지 않고 증명자가 계산의 정확성을 입증할 수 있습니다.
- 보안: SF 코딩 체계는 SIS(Short Integer Solution) 문제의 경도를 포함하도록 설정할 수 있으며, 이는 양자 컴퓨터의 공격에 대해서도 안전하게 됩니다. 이는 장기적인 보안을 위한 유망한 옵션입니다.
- 확장성: SF 코딩 체계는 확장 가능하도록 설계되었습니다. 즉, 비실용적으로 증명 및 검증 속도가 느리거나 저장 공간 등의 주요 리소스를 많이 사용하지 않고 대규모 계산을 처리할 수 있습니다.

시사점

- AI사업요약
  - 각 응용 분야별 뉴럴넷 방식 선택, 확보한 데이터로 학습, 서비스 제공
- 향후 AI기업/정부/기관/연구 그룹이 고민하는 부분
  - 도메인 거대 학습데이터 확보
  - 학습데이터를 저장하고, 훈련된 서비스를 운용할 수 있는 GPU클라우드
  - 개인 프라이버시를 제고할 수 있는 기술력
  - Open source 경쟁에 나설 AI 학습 및 서비스 인프라 확보 방안
- WorldLand AI Nexus
  - Cloud
  - 데이터 제공 인센티브
  - Privacy
  - 참여 공유 형 생태계

07. Scale-up

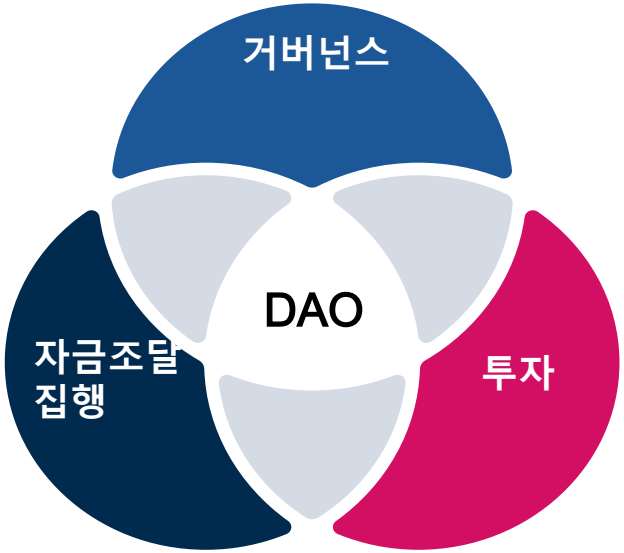
펀딩 및 론칭 스케쥴: 펀딩으로 확보한 자금, ECO생성 활용

	준비	1st Qtr D-date	2nd Qtr	3rd Qtr	4th Qtr	2nd yr Qtr-1	2nd yr Qtr-2	2nd yr Qtr-3
Testnet 자체검증	→							
Testnet 외부		→						
토큰 발행 및 세일 A, B, P		▲	▲	▲				
Ethereum 지갑계승 이벤트				▲	Genesis			
WorldLand Mainnet			공식테스트넷1	공식테스트넷 2	→	→	→	→
상장				DEX ▲	CEX ▲			
기술로드맵			AI Nexus 개발 및 검증	AI Nexus 추론 서비스 개발	AI Nexus 학습 서비스 개발	AI Nexus 론칭	격자부호 VCT	격자부호 ZKP

WLT, WLC은 비증권임 (디라이트 조원희 변호사 법률검토의견서)

WorldLand DAO

1. WorldLand DAO는 재단 DAO 형태를 따른다.
2. 글로벌 투자금 유치와 집행은 Smart Contract (Ethereum)를 통해 진행한다.
3. 토큰 지분으로 자금 집행 등 중요한 의사를 결정한다.
4. WorldLand Ecosystem Growth Foundation 역할을 수행한다.
5. G직후 설립, 5년간 개발 및 ECO생성 자금을 집행한다.

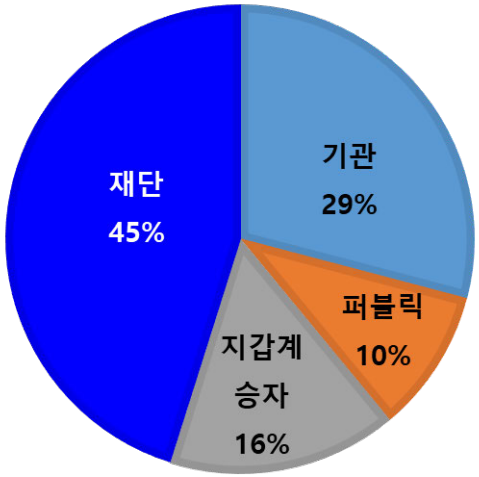


07. Scale-up

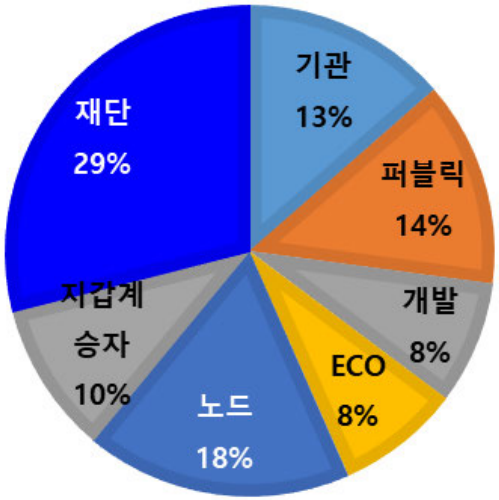
재단의 탈중앙화 전략 = ECO 구축 정책

5개년 간 개발 및 ECO생성 자금 집행후 스스로 탈 중앙화 함

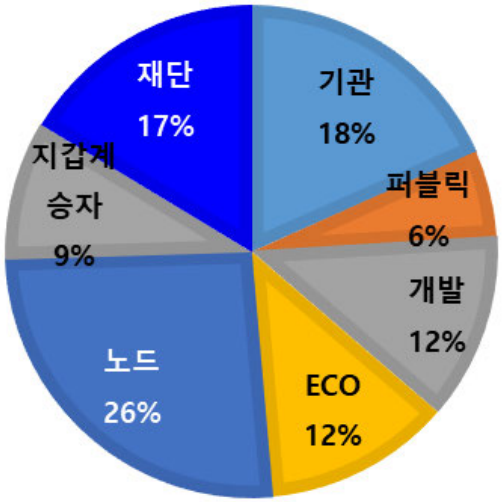
DAO 지분 @제네시스



DAO지분@YR2

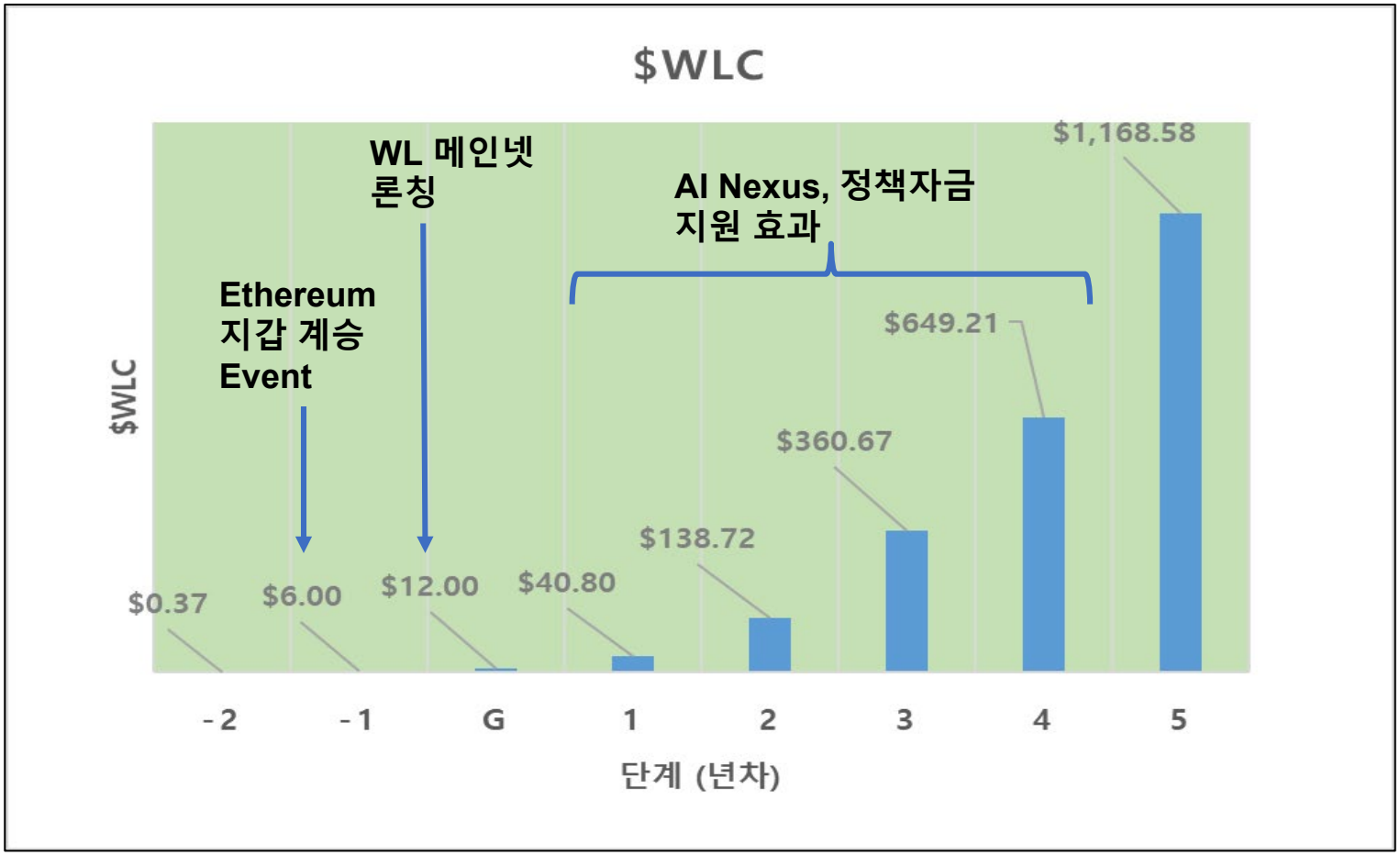


DAO지분@YR5



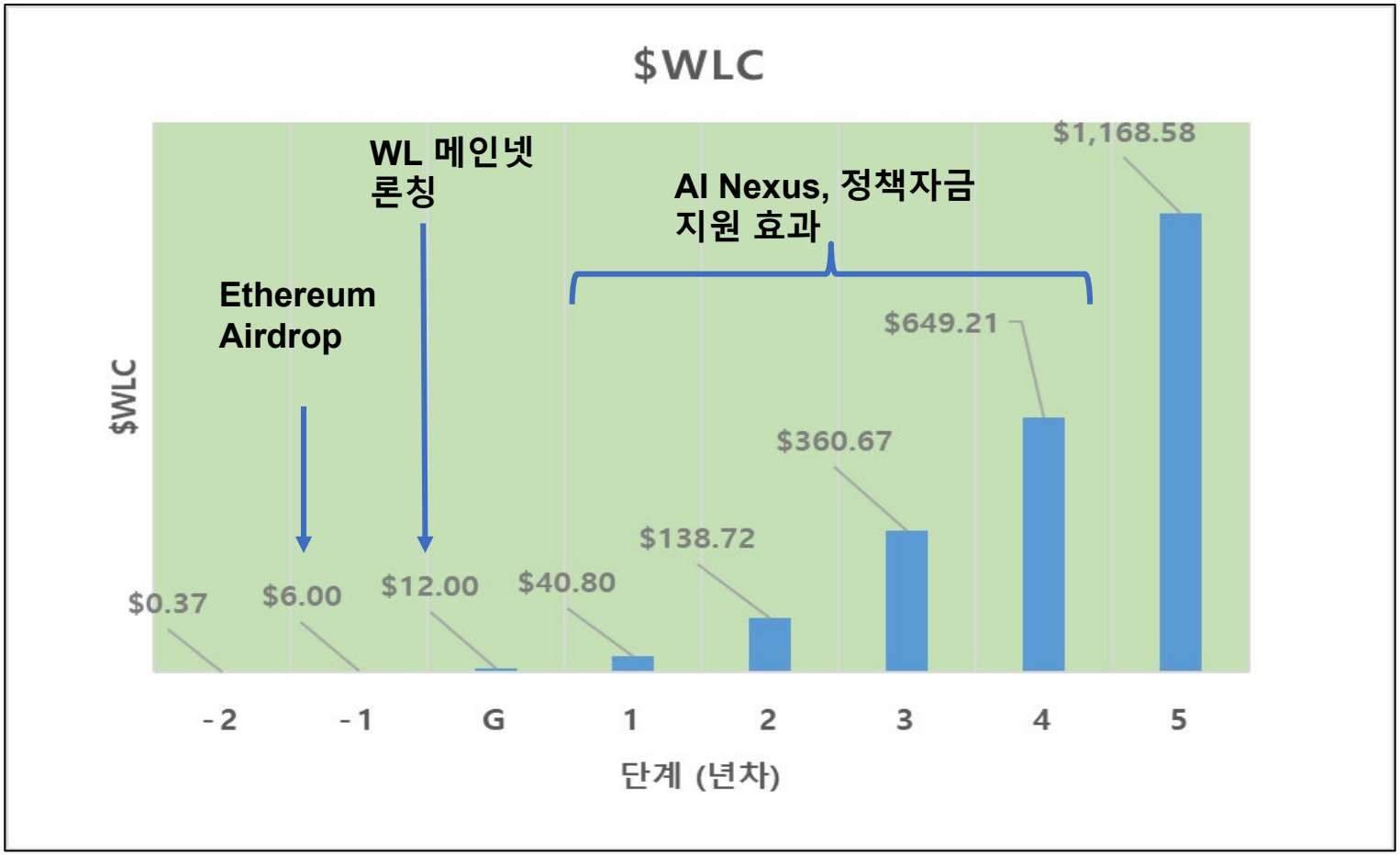
07. Scale-up

투자금



07. Scale-up

$WLC Simulation



07. Scale-up

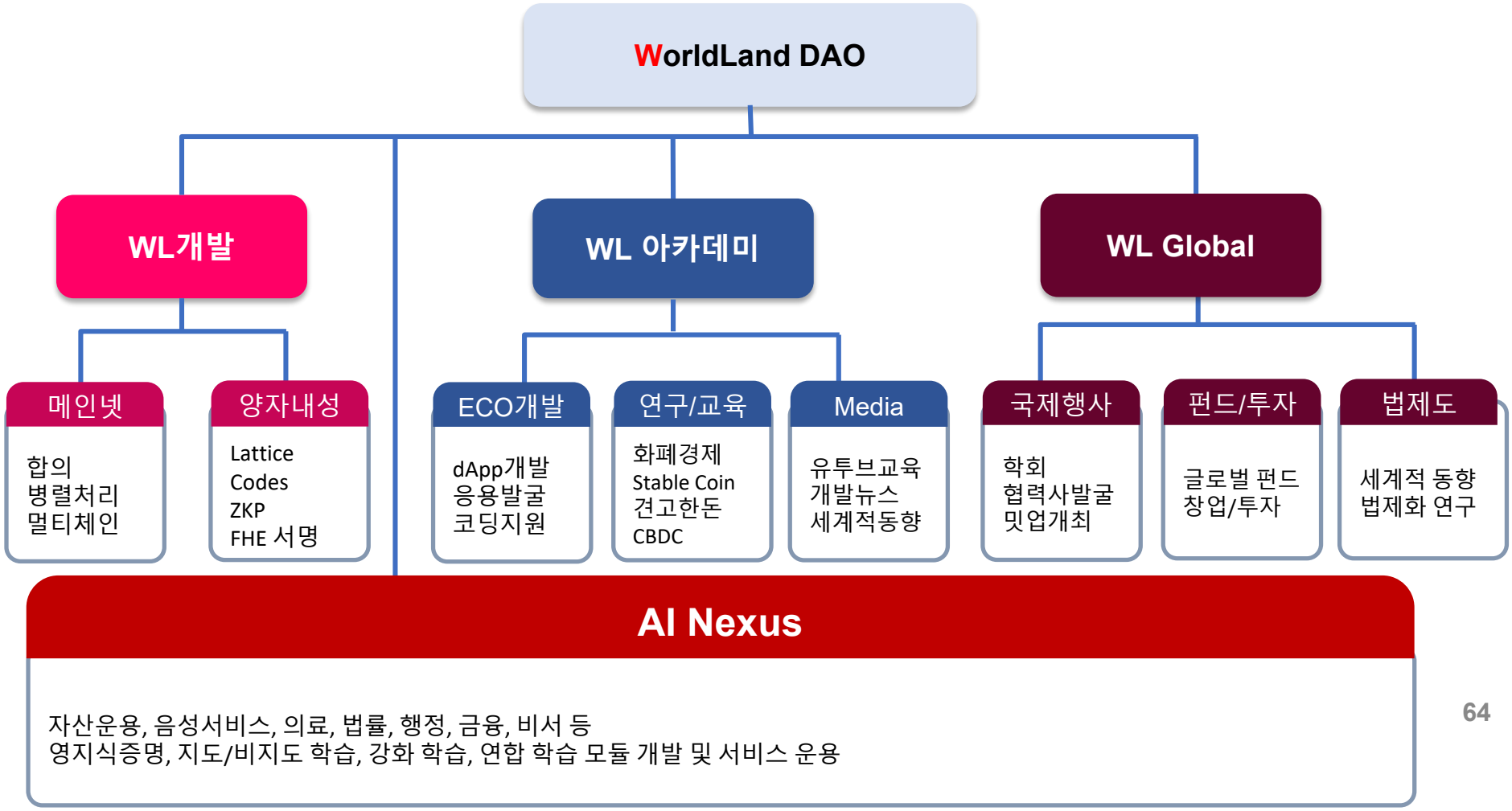
Market Cap Simulation

3년 차에 $50B 목표, 네트워크 효과(활성 지갑, 노드, 댑 등).



07. Scale-up

WL DAO



07. Scale-up

WorldLand ECO 주요이해당사자들
WorldLand 생태계 內 플레이어들의 역할
기하 급수적 성장 가능

Players in WorldLand

DAO 재단
- 토큰 발행
- 토큰 상장

투자자
- 토큰과 투자금 스왑
- 토큰 상장 시 EXIT가능

AI블록체인 개발사
- AI Nesux 프로토콜 완성 및 공개
- AI 프로토콜 설치 노드를 VC사업자에 판매

노드 사업자
- 노드 구매 or 직접설치 운영
- 전기사용 노드 운영
- 노드운영 수익
  - ✔ 주조차익, Gas Fee

사용자
- 지갑
- dApp
- AI 학습 및 추론 서비스 사용자

dApp 사업자
- Smart Contract
  - ✔ AI Nexus
  - ✔ 스왑서비스
  - ✔ Soul Bound Tokens
  - ✔ 데이터사업

07. Scale-up

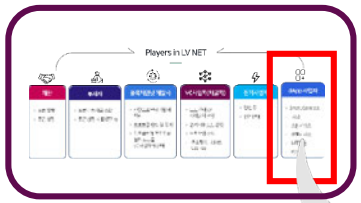
WorldLand AI Nexus ECO 재창출

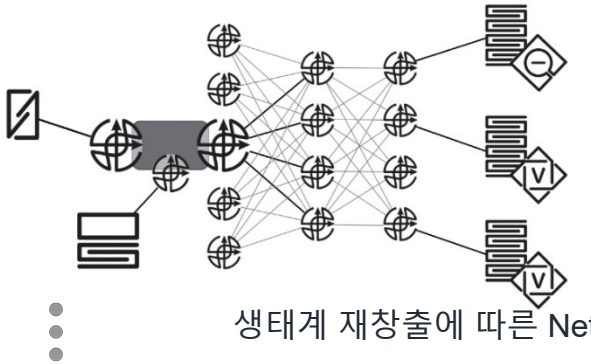
다양한 AI dApp 생태계 內 탈 중앙 서비스 제공을 통한 생태계 재창출

다양한 dApp 사업자들이 오픈 WorldLand에 AI dApp 서비스 스마트 컨트랙트를 배포

(법률, 의료, 금융, 이상징후 파악 등)

WorldLand Layer 2 네트워크 생태계 재창출 구조





생태계 재창출에 따른 Network effect ↑

AI Nexus 응용 부분별 사업

dApp 재단

- 토큰 발행 및 상장
- TVL 상승->토큰값 상승

dApp 개발사

- 재단으로부터 개발비 지원
- AI 프로토콜 완성/공개/업그레이드 등 새로운 서비스 지속 개발
- V1->V2->V3
- Smart Contract 보완 새로운 솔루션 제시

dApp 사용자

- A토큰을 Pool에 제공
- B토큰을 구매

↓

유동성 공급자에게 수수료 지불

노드 사업자

- TX fee, Gas fee →LV Net의 노드 사업자
- LP fee →LP 공급자
- Swap서비스 개발비→Swap 서비스 개발자

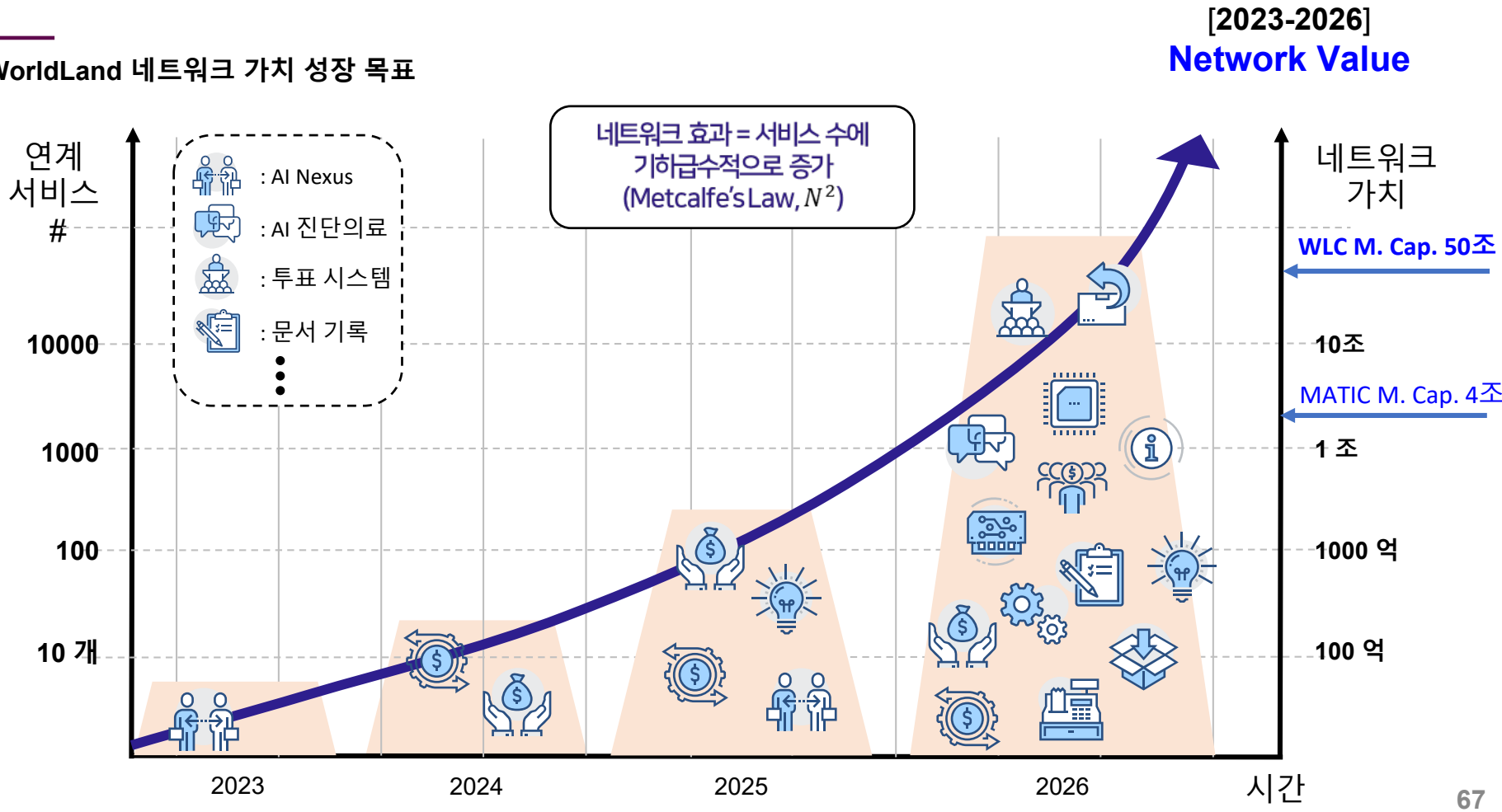
유동성 공급자

- 자본을 투자하여 Swap SV 사용자가 Swap 할 수 있도록 유동성 공급
- 자산pair로 Pool 형성
- (ex)Uniswap의 Pool 개수: **50개**

07. Scale-up

론칭 직후 재단의 5개년 ECO 생성 지원 자금 집행의 결과는?
3년 안에 50조 가능

WorldLand 네트워크 가치 성장 목표



07. Scale-up: 기반 구축 전략

전략적 제휴 추진 현황

- Megazone Cloud
- 씨이랩 (XIIIlab)
- 인공지능융합사업단

광주 AI데이터센터



□ 추진경과

- ㅇ 인공지능산업융합 사업단 발족 및 사업단장 임명 : '20. 1월
- ㅇ 과기정통부-정보통신산업진흥원 간 사업 협약 체결 : '20. 1월
- ㅇ **인공지능산업융합사업단 설립** : '20. 1월
- ㅇ **사업 수행계획 수립 및 NIPA적정성 검토** : '20. 3월
- ㅇ **NIPA-광주광역시-사업단 간 사업 협약 체결** : '20. 3월
- ㅇ 사업 추진방향 및 '20년 추진계획 운영위원회 심의 : '20. 3월
- ㅇ **사업단-참여기관 간 세부사업 협약체결** : '20. 4월

2023년 4월 11일 88PF 규모

광주 AI데이터센터에 국내 최대 규모 슈퍼컴 인프라 들어선다

나호정 기자 입력 2023.04.11 18:36 수정 2023.04.12 18:35 댓글 1 좋아요 0

인공지능산업융합사업단, 10월 67PF 용량 'H100' GPU 도입 계획...
총 88.5PF 규모로 10월 오픈 목표...글로벌 수준 초거대 AI 모델 개발 가능해져



광주 첨단 3지구에 건설 중인 인공지능 데이터센터 (사진=인공지능산업융합사업단)

07. Scale-up

글로벌 진출전략

글로벌 그룹과의 MOU 체결을 통한 시장 진출 전략

핵심 팀 멤버 이흥노 대표 – (국제학술활동) 톱저널 편집위원, CVA 멤버, Consensus 2019, ConsenSys MOU

글로벌 진출을 위한 커뮤니티 참석 성과 및 계획

행사명	참여 형태	일시, 장소
Crypto Valley Conference on Blockchain Technology (CVCBT) 2018	ECCPoW 홍보 및 개발자 커뮤니티 교류	18.06.20 ~ 18.06.22
Crypto Valley Association (CVA) New York Consensus 2019	DeSecure Blockchain 연구 프로젝트 ***홍보 부스 개설***	19.05.13 ~ 19.05.16
ICML 2022	**완전동형암호 Bootstrapping *기술 논문 오랄 발표***	**22.05.30 ~ 22.06.03, Baltimore, MD**
Consensus 2023	**WorldLand 네트워크 홍보 및 개발자 커뮤니티 교류 참여**	**23.04.26 ~ 22.04.28, Austin, Texas**
Future Blockchain Summit 2023	**WorldLand 네트워크 기술 홍보 및 체험 부스 개설**	**23. 5.10 ~ 22.5.12, Dubai, UAE**

* 외 AsiaCrypt 2022 (Taiwan), Crupto 2022 (USA), DeFi Conference 2023 (Ireland) 등 국제 개발자 커뮤니티 적극 참여 예정



글로벌 ConsenSys와 MOU



Consensus 2019, NY, USA



IEEE Trans. 편집위원

08. Team

주 개발사 리버밴스

주요 임원



CEO 이흥노

GIST 전기전자컴퓨터공학부 정교수
(전)대통령 직속 정책기획위원회 전문위원
ITRC 블록체인 지능 융합 센터장
국내외 논문 400여편 및 특허 60여건 보유
Associate Editor of IEEE Transactions on Cybernetics (세계 톱4% SCI Journal) 등 다수 국제 저널의 편집 위원 수행 중



법률 고문 조원희 변호사/변리사

법무법인 디라이트 설립자/대표변호사
과학기술정보통신부 고문변호사
(전) 법무법인 태평양 변호사



CTO 김영식

조선대학교 정보통신공학부 정교수
양자내성암호연구단(KpqC) 기술자문위원
IEEE Globcom2022 SAC Cloud, Technical Program Committee Members
동형암호 관련 EUROCRYPT 2022/2021 게재 외 국내외 논문 160여편



CMO / CFO 이동건

리버밴스㈜ CMO/CFO
(전) GIST 기업지원센터 본부장
(전) 광양보건대 교수
(전) 한국정보전략학회 부회장 및 상임이사

요약

- **WorldLand** 는 Green VCA기반 탈중앙 블록체인 네트워크
  - ECCPoW + Green VRF
  - 테스트넷 가동중
  - MAC-ZKP에 기반한 새로운 ZK-Rollup
  - AI Nexus
- **WorldLand**는 AI Nexus 에 특화된 형태로 성장하고자 목표

09. 참고 자료

블록체인/영지식 관련 이흥노 교수의 핵심 참조 자료

Full list at https://heungno.net/publications/

- [RM1] Z. Iqbal and H.-N. Lee, “Spatially concatenated channel-network code for underwater wireless sensor networks,” IEEE Transactions on Communications, Vol. 64, No. 9, pp. 3901-3914, Sep. 2016. https://heungno.net/wp-content/uploads/2012/08/07518674_SCCNC_published.pdf
- [RB1] J. Jang and H-.N. Lee, “Universal Zero-Knowledge Succinct Non-Interactive Argument Of Knowledge Proof System For Stack-Based Virtual Machine Program 스택기반 가상머신 프로그램을 위한 효율적인 유니버셜 영지식 증명 시스템,” US patent 18/070,497, 출원일 29-Nov-2022.
- [RB2] H. Kim, J. Jang, S. Park, and H.-N. Lee, “Error-Correction Code Proof-of-Work on Ethereum”, IEEE Access, Vol. 9, pp. 135942-135952, Sep 2021. doi: https://doi.org/10.1109/ACCESS.2021.3113522.
- [RB3] J. Jang and H.-N. Lee, “Profitable Double-Spending Attacks,” Applied Sciences, 10, 8477, Nov. 2020. doi: https://doi.org/10.3390/app10238477. MATLAB evaluation script: https://codeocean.com/capsule/2308305/tree.
- [RB4] H. Jung, H.-N. Lee, “ECCPoW: Error-Correction Code Based Proof-of- Work for ASIC Resistance”, Symmetry, June. 2020, 12(6), 988.
- [RB5] S. Park, H. Choi and H.-N. Lee, “Time-Variant Proof-of-Work using Error-Correction Codes”, Submitted to IEEE Trans. on Information Forensics and Security. (https://arxiv.org/abs/2006.12306)
- [RB10] Heung-No Lee, Young-Sik Kim, Dilbag Singh, and Manjit Kaur, “Green Bitcoin: Global Sound Money,” The Journal of Digital Assets, Vol.1, No. 1, pg. 33-47, October, 2022. https://arxiv.org/abs/2212.13986

연합학습, 초분광뉴럴넷 등 인공지능 관련 자료는 https://heungno.net/publication 에서 확인 바람

09. 참고 자료

영지식 증명 관련 참조 자료

[Martin22] Martin, et al., "Lattice-based SNARKs: Publicly Verifiable, Preprocessing, and Recursively Composable," Cryptology ePrint Archive, Paper 2022/941, year 2022, https://eprint.iacr.org/2022/941.

[Sasson22] Eli Ben-Sasson, Iddo Bentove, Yinon Horesh, M. Riabzev, "Scalable, transparent, and post-quantum secure computational integrity," https://starkware.co/wp-content/uploads/2022/05/STARK-paper.pdf.

[Groth17] Jens Groth and Mary Maller. Snarky signatures: Minimal signatures of knowledge from simulation-extractable snarks. In Advances in Cryptology - CRYPTO 2017 - 37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part II, pages 581–612, 2017.

[Orion22] Tiancheng Xie, Yupeng Zhang, and Dawn Song, "Orion: Zero Knowledge Proof with Linear Prover Time," Cryptology ePrint Archive, Paper 2022/1010, Year 2022, https://eprint.iacr.org/2022/1010.

[Boneh20] Boneh, D., Drake, J., Fisch, B., Gabizon, A.: Halo infinite: Recursive zk-SNARKs from any additive polynomial commitment scheme. Cryptology ePrint Archive, Report 2020/1536 (2020).

[Sasson19] Ben-Sasson, E., Bentov, I., Horesh, Y., Riabzev, M.: Scalable zero knowledge with no trusted setup. In: CRYPTO (2019).

양자내성 암호 관련 참조 자료

[MacEliece78] McEliece, R.: A public key cryptosystem based on algebraic coding theory. DSN progress report, 42-44:114–116 (1978).

[Sidelnikov92] Sidelnikov, Vladimir Michilovich, and Sergey O. Shestakov. "On insecurity of cryptosystems based on generalized Reed-Solomon codes." (1992): 439-444.

[Janwa96] Janwa, H. and Moreno, O.: McEliece public key cryptosystems using algebraic geometric codes. Designes, Codes and Cryptography, 8:293–307 (1996).

[Dillithium17] Lyubashevsky, Vadim, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Peter Schwabe, Gregor Seiler, Damien Stehlé, and Shi Bai. "Crystals-dilithium." Submission to the NIST Post-Quantum Cryptography Standardization [NIS] (2017).

[Falcon20] Pierre-Alain Fouque Jeffrey Hoffstein Paul Kirchner, Vadim Lyubashevsky Thomas Pornin Thomas Prest Thomas Ricosset, Gregor Seiler William Whyte Zhenfei Zhang, “Falcon: Fast-Fourier Lattice-based Compact Signatures over NTRU,” Specification v1.2 — 01/10/2020, https://falcon-sign.info/falcon.pdf. Visit https://falcon-sign.info/ for complete package.

[Durandal19] Aragon, N., Blazy, O., Gaborit, P., Hauteville, A., Zémor, G. (2019). Durandal: A Rank Metric Based Signature Scheme. In: Ishai, Y., Rijmen, V. (eds) Advances in Cryptology – EUROCRYPT 2019. EUROCRYPT 2019. Lecture Notes in Computer Science, vol 11478. Springer, Cham. https://doi.org/10.1007/978-3-030-17659-4_25.

[Cover06] Thomas M. Cover and Joy A. Thomas, Elements of Information Theory, 2nd Edition, Wiley-Interscience, 2016. http://staff.ustc.edu.cn/~cgong821/Wiley.Interscience.Elements.of.Information.Theory.Jul.2006.eBook-DDU.pdf.

[Gallage63] Robert Gallager, Low-Density Parity-Check Codes, MIT, 1963. https://web.stanford.edu/class/ee388/papers/ldpc.pdf.

감사합니다!

비즈니스 협력 연락처

이흥노 대표: ceo.libervance@gmail.com

팁스프로그램(TIPS) 기술창업기업 사업계획서 Part2

LiberVance 네트워크 기술 개발

- Layer2 플랫폼

운영사 | 선보엔젤파트너스㈜

창업팀 | 리버밴스㈜

CONTENTS

01. Market status

02. Problem

03. Solution

04. Technology

05. Scale-up

06. Team

07. 마무리

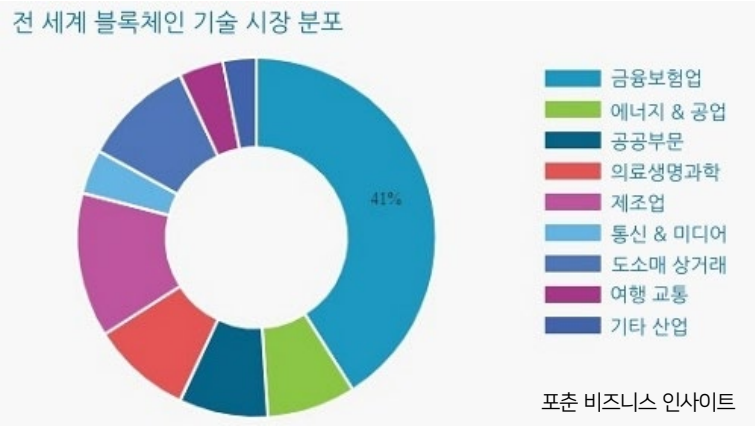
01. Market status

블록체인 기술의 활용 및 지속적으로 성장하는 시장규모

데이터의 분산화 및 저장, 소유권을 가져 통제할 수 있는 차세대 웹 환경인 Web3.0의 탈중앙화된 플랫폼은 모두 '블록체인' 기술 기반
블록체인 시장은 2026년까지 연간 68.4% 성장 예상되며, 가상자산 전체 시가총액은 약 2,634조 규모로 2021초 대비 약 200% 증가

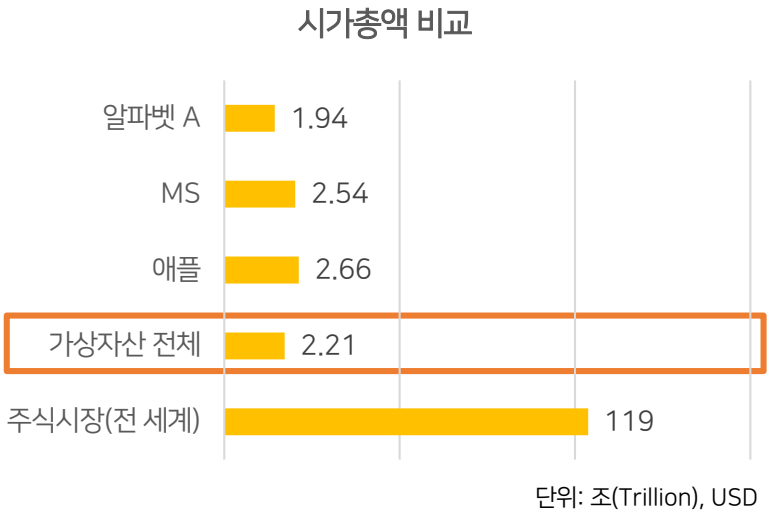
지속적으로 증가하는 블록체인 시장

- *블록체인은 다양한 플랫폼 (NFT, P2E게임, DeFi, 메타버스, 시티코인, 비트코인 등) 의 기반기술*
- 블록체인 기술에 대한 투자 증가, 은행 및 사이버 보안 분야의 블록체인 솔루션 사용 확대 등이 시장 성장 주도할 것으로 예상
- 공공 및 정부의 인지도 향상, 규제완화로 이어져 블록체인 산업 성장이 점차 가속



가상자산 시장 규모

- 가상자산 전체 시가총액은 2021년 초 대박 약 200% 증가
- 가상자산 시장 확대 주력은 ' *알트코인 ' 으로 1.4T 달러, 6배 이상 증가



*알트코인 : 비트코인을 제외한 모든 가상화폐

01. Market status

확장성 트릴레마 (Scalability Trilemma) 및 해결책

확장성, 탈중앙성, 보안성이 모두 우수한 블록체인을 만드는 것은 현재 블록체인 기술에서 가장 풀기 어려운 문제
해결책은 중앙 집중형 합의 스케줄러를 추구하는 Top Down 방식 vs 시장 수요에 의한 자연발생을 추구하는 Bottom Up 방식

블록체인의 성능지표: 확장성, 탈중앙성, 보안성

- **확장성** (최대 초당 거래수 (TPS) 가 얼마나 높은가?)
  - ✓DS Attack 저항위해서, 블록 용량과 BGT(블록생성속도)를 제한함
  - ✓PS 감소: 21년 5월, NFT 공급량 초과로 거래수수료 평균 건당 68달러 도달
- **탈중앙성** (특정 개인 또는 집단이 채굴을 독점할 수 있는가?)
  - ✓각종 공격에 대한 보안 유지를 위한 충분한 분산 노드가 필요
  - ✓채굴 ASIC의 등장으로 인해 새로운 노드의 진입이 어려운 문제 발생
  - ✓누구나 고르게 참여 가능한 합의 알고리즘 필요
- **보안성** (각종 공격으로부터 얼마나 안전한가?)
  - ✓확장성 개선을 위한 Layer-2 솔루션 혹은 사이드체인의 보안 문제
  - ✓사이드체인의 분산성 결여, 각종 공격에 취약
  - ✓다양한 공격 대비를 위한 암호학적 보안 솔루션 필요

Top Down vs Bottom Up

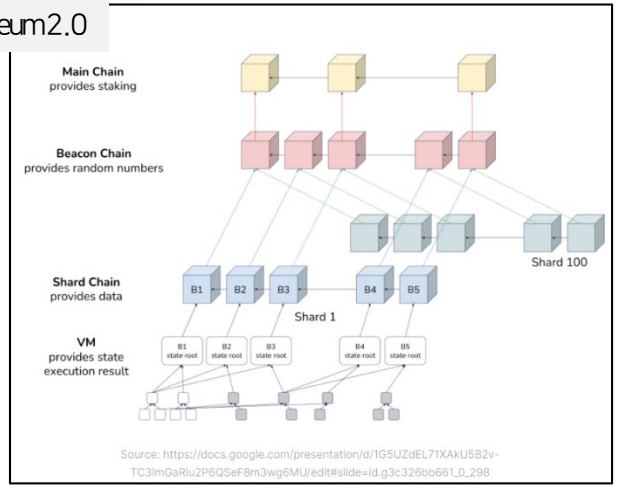
Bottom Up 방식

- 자연발생적,
- 자연선택적 방식
- ETH, BTC는 글로벌 상위 Layer로 존재
- 병렬적 L2의 확장성 문제 해결 가능

- 중앙집중형
- Ethereum 2.0
- Beacon + 64 shards

Top Down 방식

Ethereum 2.0



01. Market status

Layer-1 vs Layer-2

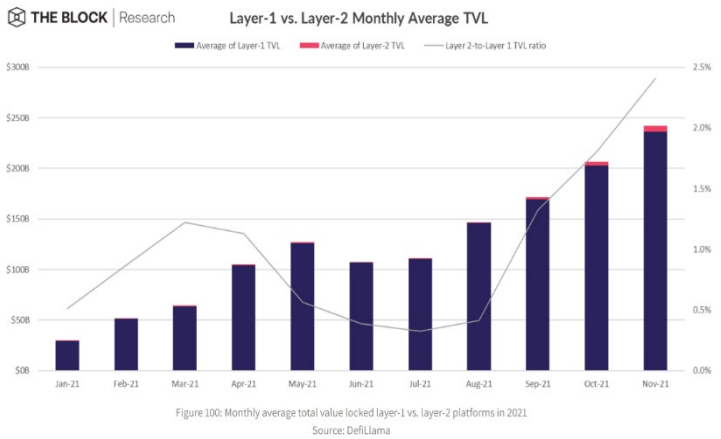
서로 독립적이고, 높은 트렌젝션 수를 감당하지 못하는 Layer-1 스케일링 솔루션
사용성 위주의 자유로운 설계가 가능한 Layer-2 스케일링 솔루션

Layer-1 문제점

- L1 코인들은 서로 교류 불가, 각자의 영역에서만 거래 가능
- 주로 이더리움 네트워크 기반으로 사용하고 있어 거래 비용이 높고 Transaction을 감당하지 못함

왜 Layer-2 로 가야 하는가?

- 성장 기회가 큰 새로운 네트워크인 L2
- 낮은 거래 수수료 및 빠른 거래
- 속도와 사용성 위주의 자유로운 설계 가능
- 이더리움과 결합되어 실행되면 L1보다 L2를 더 많이 사용할 수 있을 것으로 전망 (The Block 리서치, 21.12)
- 최근 *dYdX의 두드러진 성장, L2 매우 초기 단계



[2021년 월평균 TVL 변화]

polygon
omisego
LOOPRING
ARBITRUM
SYNTHETIX
Bancor

*dYdX : 이더리움 블록체인의 스마트 계약을 활용한 탈중앙화 거래소(DEX)

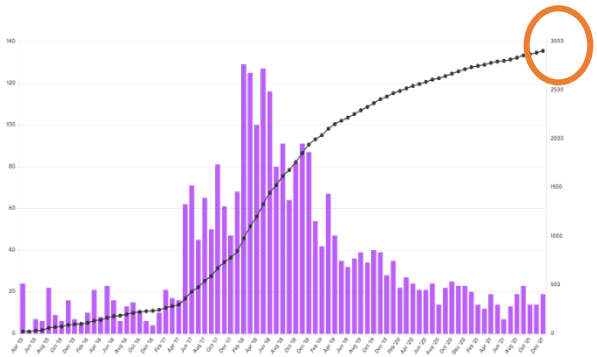
01. Market status

가장 많이 이용하는 블록체인 네트워크 ' 이더리움 '

전체 탈 중앙화 서비스 (dApp) 중 75%가 이더리움 기반 서비스를 제공
시장 성장에 따른 혼잡성, 수수료 증가 등의 문제 존재, 이를 해결하기 위한 Layer-2 솔루션 제시
2년 내 약 80%까지 Layer 2 로 이동할 것으로 예상

하루 사용자 약 1억 명, 거래 2억 4천여 건

- 계속하여 증가하는 사용자 수와 거래량으로 인해 거래완료에 걸리는 시간, 혼잡성, 수수료 증가 등의 문제 야기
- 그러나 최근(22년1월) 솔라나 다운으로 인한 디도스 공격 취약성 등 안전성 문제, 非탈중앙성 등의 문제 지속적 대두



[이더리움 앱 수 및 신규 추가현황 (21.12월 기준)]



문제 해결 위한 Layer-2 솔루션

- 스마트 계약을 통해 Layer-1 네트워크 위에서 유지
- 주요 네트워크와 상호작용
- 기본 프로토콜 수정에 의존하지 않음
- Layer-2 프로젝트의 유동성 인센티브가 등장하며 점점 성장하고 있다고 진단 (The Block 리서치)



01. Market status

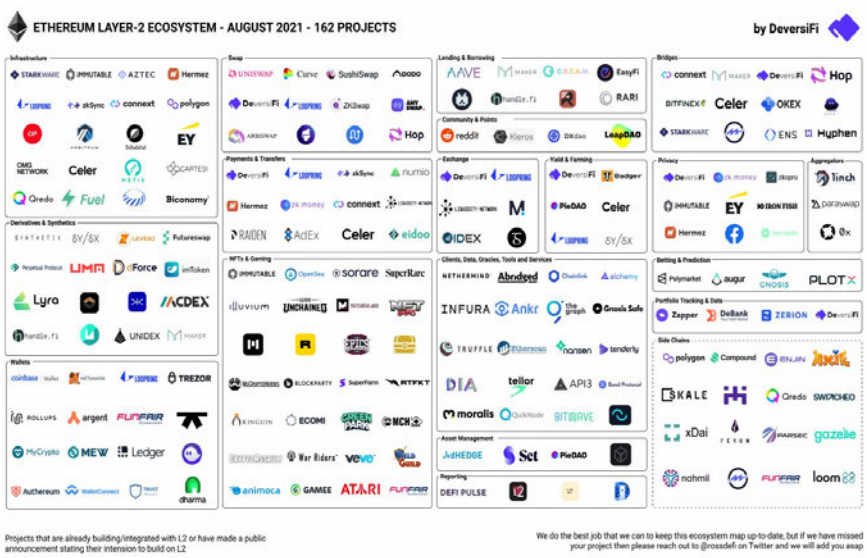
Layer-2 솔루션 생태계

폴리곤 활성 이용자 수 50만 건 이상으로, 이더리움 활성 이용자를 능가 (21.10 기준)

Layer 2 시장의 성장은 뚜렷하며, 최대 50배까지 성장할 것으로 예상

대표적 사례 - 폴리곤(Polygon)

- 2021년 급성장한 *L2는 사이드체인, 플라즈마로 시작
- 2021년 5월 이더리움 네트워크 체증 문제 증대로 폴리곤(Polygon) 확장 솔루션 수요 폭증
- 이더리움을 능가하는 활성이용자 보유



* By DeversiFi, 21.8월 기준

Layer-2의 시장 가능성

- 크립토 경제는 이미 무시할 수 없는 규모로 성장
- 경제시장의 원활할 작동을 위해 확장성과 상호운용성 솔루션 제공 기본 인프라 구축은 필수 불가결
- L2의 성장은 뚜렷하며 향후 20~50배 성장 전망 (21.11, Ryan Berckmans)

*L2 : Layer-2

02. Problem

기존 Layer 2 솔루션의 문제점: 보안 문제, 탈중앙성 이슈

보안성이 최대 장점인 블록체인 플랫폼에서 보안문제 발생, 꾸준히 제기되는 탈중앙성 이슈

→ TPS 제고 위해 탈중앙성이나 보안성이 취약한 합의 알고리즘과 체인 연결 구조 선택

(문제점 1) 기존 L2 솔루션의 합의 알고리즘 상 문제점

- 기존 L2 솔루션은 PoS/PoA 기반 합의 방식 사용
  - 블록 생성 노드, 검증자 노드의 숫자가 적음
    → DoS 공격, 권한 탈취 공격에 취약해지는 보안문제 발생
  - 탈중앙성 훼손
    → 블록 생성 노드 운영자에 대한 신뢰 문제 발생

(문제점 2) 기존 L2 솔루션의 체인 연결 기술상 문제점

- 브릿지 구조상 보안성 문제가 발생하기 쉬움
- 문제 발생 시, 연결 체인 사용자 전체가 피해를 입을 수 있음
- 체인 연결 구조에 따른 사용상의 문제 ex:
  - Plasma, Optimistic rollup : exit을 위해 challenge 기간 동안 기다려야 함
  - zk rollup : 네트워크의 초기 신뢰성 문제, 계산량 문제, 양자 컴퓨팅으로 인한 해킹 우려
  - Cosmos : 각 체인들이 보안성을 각자 확보해야 함
  - Polkadot : 검증자를 다시 감시하는 fisherman 이 별도로 필요

1) 합의 알고리즘과 2) 체인 연결 구조를 개선하여 이러한 문제들을 해결할 수 있음

02. Problem

합의 알고리즘상 문제점 – PoA 권한 해킹 사례

합의 알고리즘의 취약점을 이용한 해킹

액시 인피니티 암호화폐 유출 사태 – 로닌 사이드 체인

- 22년 03월, 이더리움과 액시 인피니티의 유틸리티 토큰을 이어주는 사이드체인 '로닌'에서 보안 침해사고 발생
- 권위 증명 (Proof of Authority, PoA) 합의 알고리즘 사용
  - 사이드 체인 내 거래를 검증하는 벨리데이터(Validator) 풀이 매우 소규모임을 의미 ***(탈 중앙성 훼손)***
  - 액시 인피니티의 경우 9개 유효성 검사기 중 최소 서명에 필요한 개수인 5개 권한을 해킹당했으며, 임의 Transaction을 발생시켜 암호화폐를 탈취함.
- 피해 금액은 17만 3,600 ETH 및 2,550만 USDC 로, 약 7,500억원이며, 이는 역대 최대 규모임.
- 해킹 후 6일만에 발견되었으며, '로닌'의 거버넌스 토큰인 RON 은 20% 가까이 하락함.





02. Problem

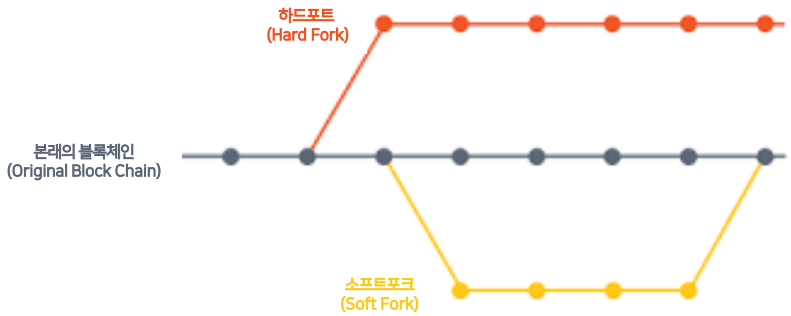
합의 알고리즘상 문제점 – 탈중앙성 훼손 논란 사례

PoS 합의 알고리즘의 탈중앙성 훼손

폴리곤 비공개 하드포크 진행으로 제기된 탈중앙성 훼손 논란

- 21년 12월, 메이저 L2 솔루션 중 하나인 폴리곤 네트워크가 일반 사용자들에게 공지 하지 않고 하드포크를 진행함.
- 블록체인 하드포크 (BlockChain Hard Fork)
  - 체인의 기본 동작에 해단하는, 기본 코드 (Base Code)를 업데이트 하는 일
- 폴리곤 COO인 샌딥 나이왈은 '하드포크를 알리지 않고 추진한것은, PoS 보안 취약점이 악용될 수 있었기 때문이며, 최선이었다' 해명함.
- Vitalik Buterin 의 정치적 탈중앙화 (Political Decentralization)
  - 얼마나 많은 수의 개인 혹은 조직이 시스템을 구성하는 컴퓨터를 제어하는가?
- 참여 노드들에게 알리지 않고, 신속하게 하드포크가 가능했다는 점에서, 정치적 탈 중앙화 측면에서의 의혹이 제기됨.
- 많은 수의 L2 솔루션들이 실제로는 탈중앙성에 있어 논란이 야기됨.

polygon



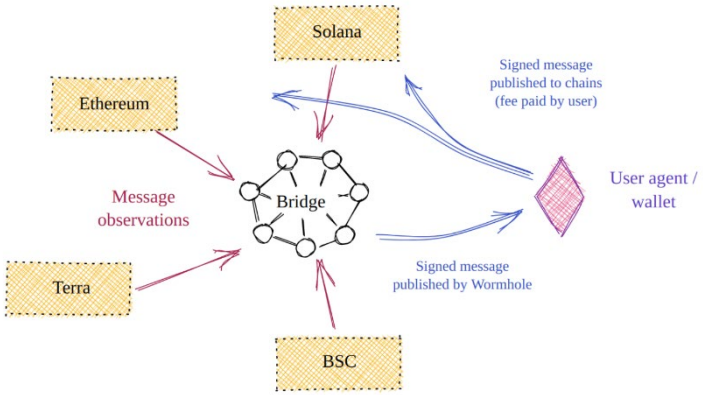
02. Problem

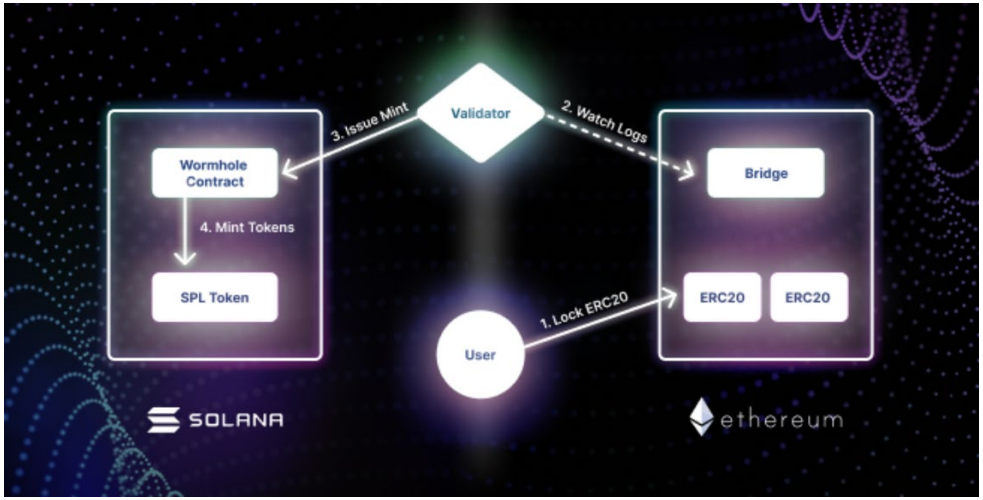
체인 연결 기술상 문제점 – Bridge 해킹 사례

체인 연결 컨트랙트의 취약점을 이용한 해킹

Solana – Wormhole 해킹 사건

- 22년 02월, 이더리움과 솔라나 블록체인을 연결하는 크로스체인 '웜홀'에서 보안 침해사고 발생
- 웜홀 서명검증 과정 (Wormhole Contract 코드) 상에서의 보안 취약점 이용 ***(논리 구조 취약)***
  - Transaction 위조, 12만개의 WETH 임의 발행
  - 브릿지 솔루션을 활용, ETH 로 바꿔 인출함.
- 피해 금액은 약 3800억원 어치로, 역대 4번째 규모임.
- 이번 해킹 사건으로 인해, Solana 와 Ethereum 모두, 10% 이상의 가치 하락이 발생함.
- 브릿지 구조에서의 보안성이 문제가 된 사례





03. Solution

LiberVance 네트워크 – 개발 목표 기술

리버밴스는 VCA 합의 기술을 개발, Ethereum L2인 LV Net 구축을 추진.
Chain 연결기술을 활용하여 Ethereum 생태계를 확장하고자 함.
EVM 사용으로 다양한 NFT, DeFi, P2E 게임 Dapp을 지원할 수 있게 됨.

LiberVance 네트워크

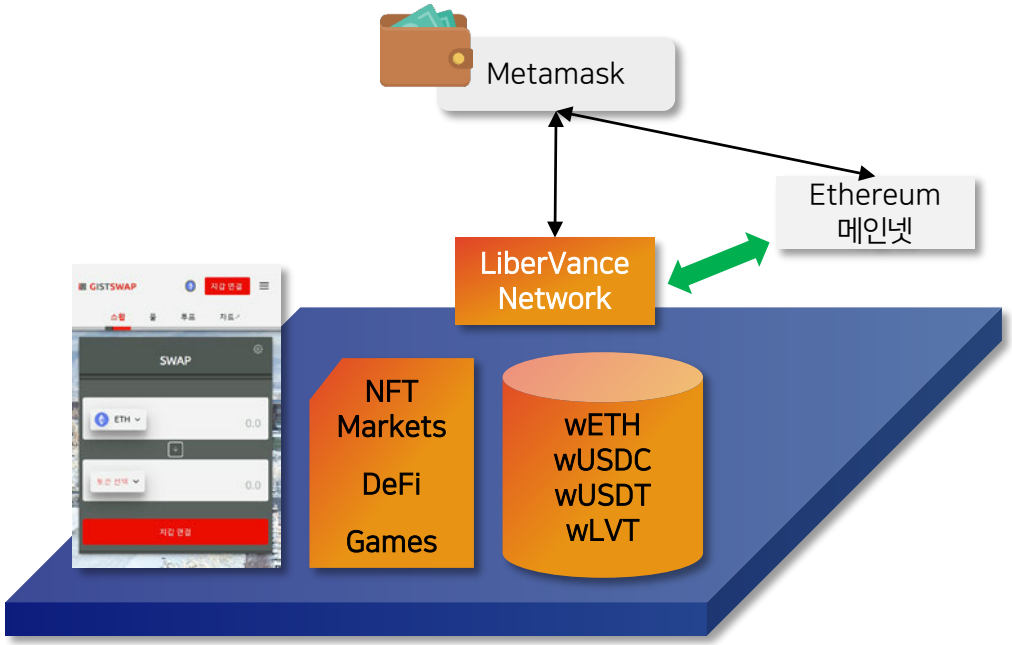
- 새로운 합의 기술에 기반한 Ethereum L2 네트워크

특징

- Ethereum Virtual Machine (EVM) 호환
  - DeFi, NFT 등 서비스 가능
- Verifiable Computation Algorithm (VCA) 합의
  - 간단하지만 안전한 합의 프로토콜

개발 요소 기술

- **VCA 합의 프로토콜**
  - ✓ PoS 방식에 비해 높은 보안성 및 탈중앙성
  - ✓ PoW 방식에 비해 낮은 에너지 소모
- **DeSecure 복합체인 연결 기술**
  - ✓ 브릿지 구조의 보안성 개선
  - ✓ 특정 노드에 집중되지 않는 탈중앙성 확보



[LiberVance Layer 2 플랫폼 서비스 개념도]

VCA 합의 프로토콜 – 개발 요소기술 1

VCA합의알고리즘은 오류정정부호에 기반함. 고치고자하는 오류의 갯수로 Puzzle 문제의 난이도를 결정함.
SHA함수와 달리, 수많은 오류정정부호를 만들 수 있음. 따라서 ASIC 칩의 출현을 방지.
탈중앙성과보안성을 제고하게 되며, 전력 소모를 최소화 할 수 있게됨.

VCA 합의 프로토콜 : 오류정정부호에 기반한 새로운 합의 기술

- 탈중앙성과 보안성이 높은 블록체인 플랫폼 개발

기존 합의 기술상 문제점

- PoS/PoA 방식: 탈중앙성이 훼손되기 쉬움, 해킹에 취약
- PoW 방식: 에너지 소모가 큼, ASIC 채굴장비로 인해 탈중앙성 훼손

VCA 합의 프로토콜의 차별성

- VCA 합의 프로토콜 기반 블록체인 플랫폼에서는
  - 대규모 p2p 노드가 블록 생성에 참여함으로써 탈중앙성 확보
  - Time-varying crypto puzzle 방식으로 이중 지불 공격 (DS attack) 에 대한 높은 보안성
  - ASIC 저항성으로 인해 탈중앙성 강화 및 에너지 소모 감소



VCA 합의 프로토콜 – 개발 요소기술 1

선행 연구 개발 결과: 논문 게재, 테스트 서버 가동, Github 코드 공개 완료

- ✓ 선행 논문: ieeexplore: Heung-No Lee eccpow
- ✓ Github: https://github.com/cryptoecc/ETH-ECC

VCA 합의 프로토콜의 정의

- 모든 노드가 동일한 절차를 독립적으로 수행함으로 합의에 이름

Verifiable Computation Algorithm (VCA) 과정

Step 1: 참여하고자 원하는 노드는 누구나 Stake하고 참여
Step 2: 각 노드는 VRF 기반 동전던지기 함수 실행하고 블록 생성 노드로 스스로 선출
Step 3: 블록 생성 노드로 선출(동전던지기에 당첨)된 노드는 Puzzle Solving (PS) 수행
Step 4: PS에 성공한 노드는 노드 선출 증거 및 Puzzle 해답 증거를 첨부하여 생성된 블록 배포

VCA 합의 프로토콜 규칙

- **가장 많은 VC Energy가 저장된 체인이 올바른 체인**
- 각 동전의 앞면성공확률을 조절하여 에너지 소모량 조절

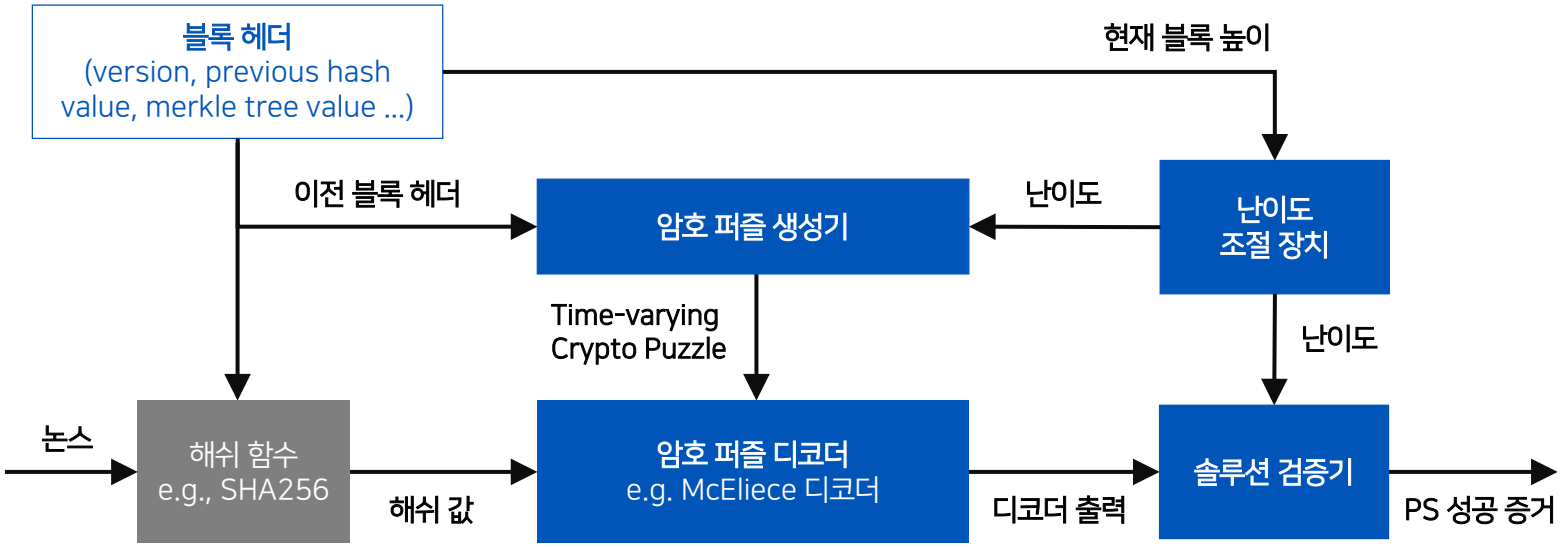
VCA 합의 프로토콜 – 개발 요소기술 1

알고리즘의 특징은 Time-Varying Puzzle 생성 가능 ASIC 저항성 확보
오류정정부호 기반으로 양자 내성 확보 가능.
난이도 조정력을 내포하여 Energy 저장 조절 가능. (예: L1 메인넷은 높은 난이도, L2 넷은 낮은 난이도)

Puzzle Solving 알고리즘의 특징

- 기존 블록체인 (ex: Bitcoin, Ethereum) 의 PoW 는 해쉬함수(hash function) 기반
- **오류정정부호**를 사용하여 Time-Varying Puzzle 생성
- **ASIC 저항성**을 확보 탈중앙성 강화
- **양자 내성 확보** 가능*

* McEliece, MDPC, Concatenated RM-RS 코드에 기반한 여러 알고리즘들이 양자 내성 공개키 암호 방식 표준으로 고려되고 있음

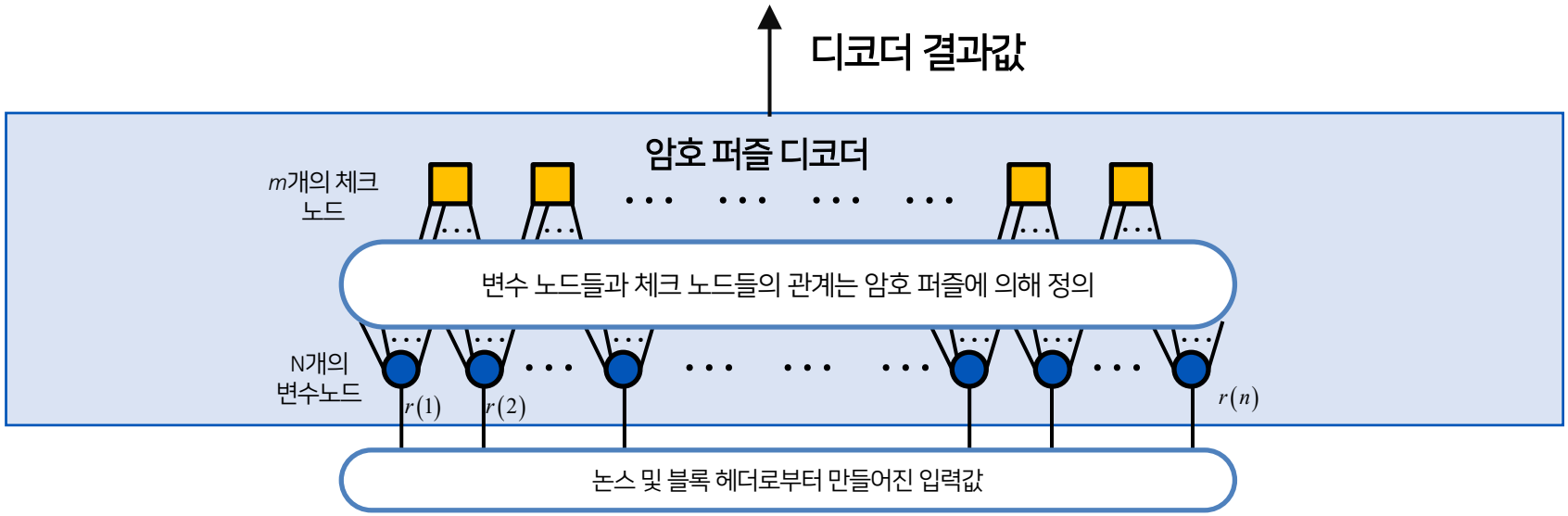


VCA 합의 프로토콜 – 개발 요소기술 1

어떻게 Time-Varying Puzzle 문제를 출제할 수 있는가?
어떻게 Verifiable Computation을 이루는가?
어떻게 수많은 노드의 합의 참여가 가능하게 되는 가?

Time-Varying Puzzle Solving 알고리즘의 구조 설명 (왜 보안성과 탈 중앙성이 제고되는가?)

- **필요한 총 연산량 계산**을 통해 필요한 **노드의 숫자 및 PS 난이도 계산 가능, 디코더 Puzzle 문제 제시**
- 모든 노드는 제시된 디코더 PS 알고리즘 수행, 그 중 성공한 노드가 **디코딩 성공 결과값** (PS성공증거) **제시**
- 디코더의 결과값이 제시되면, 다른 노드는 PS성공 증거 검증, **검증하는 것은 매우 쉬움**
- 각 노드는 ***독립적으로 PS를 수행***하는 것으로 **합의가 완료**, 따라서 **수많은 노드의 합의 참여가 매우 쉽게 됨.**
- 노드간 빈번한 통신이 필요한 BFT 계열 합의 알고리즘보다 **신뢰성과 보안성이 높음**

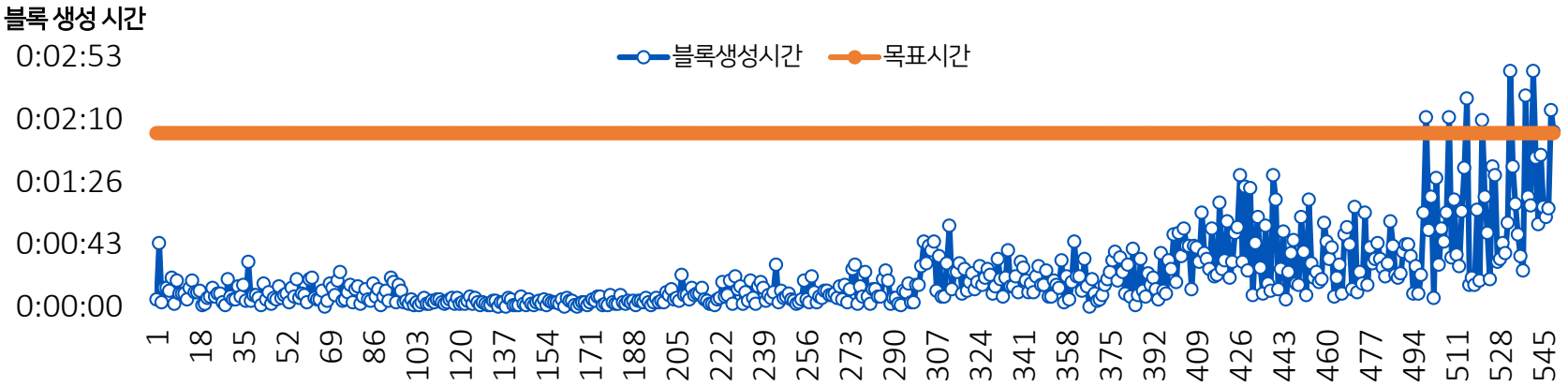


VCA 합의 프로토콜 – 개발 요소기술 1

연구 개발 목표: 블록 생성 시간 안정화 연구

VCA 합의 프로토콜 연구 개발 목표 1: 블록 생성 시간 안정화 연구

- 연구 필요성:
  - 빠르고 안정적인 거래 종결성(finality) 확보를 위해 **블록 생성 시간 안정화** 필요
  - 노드의 컴퓨팅 파워의 변동, 매 블록 바뀌는 암호 퍼즐등의 문제로 블록 생성 시간을 일정하게 유지할 방법 제시 필요
- 연구 세부 내용
  - 암호 퍼즐 파라메터에 따른 난이도 변화 분석 연구
  - 난이도 조절 알고리즘 고도화 연구
  - 컴퓨팅 파워 변동 대응 알고리즘 연구



<블록생성시간-블록높이 그래프(블록 생성 목표시간 2분)>

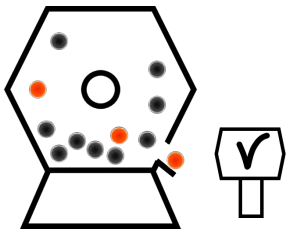
VCA 합의 프로토콜 – 개발 요소기술 1

연구 개발 목표: Verifiable Random Function 만들어 빠르고 공정한 블록 생성 노드 선출 방식 연구 및 개발

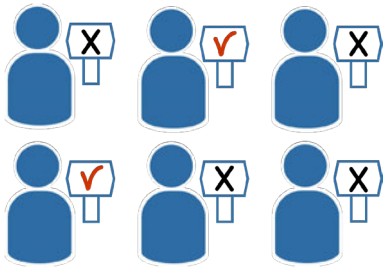
VCA 합의 프로토콜 연구 개발 목표 2: 빠르고 공정한 블록 생성 노드 선출 방식 연구

- 연구 필요성
  - VCA 합의 프로토콜 기반 블록체인은 ASIC 채굴이 불가능하므로 PoW 에 비해 에너지 소모가 낮음
  - 이에 더해 매 블록, 블록 생성 노드를 선출하여 이 노드들이 VCA 합의를 진행토록 하면 에너지 소모를 더욱 낮출 수 있음
  - 하지만 블록 생성 노드 선출이 비효율적이거나 불공정하면 블록체인의 성능 및 보안성에 문제가 생김
  - 빠르고 공정한 블록 생성 노드 선출 방식 연구가 필요함

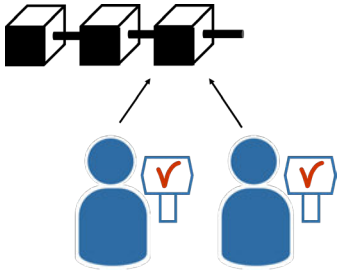
- 연구 세부 내용
  - VRF 기반의 블록 생성 노드 선출 기술 개발
  - VRF 의 보안성, 양자내성 강화 연구
  - 블록 생성 노드 선출 프로토콜을 적용하여 VCA 합의 기술 고도화 연구



VRF 기반 노드 선출



선출된 노드로 합의체 구성



선출된 노드들이 VCA 합의 진행

DeSecure 복합체인 연결 기술 – 개발 요소기술 2

LV 넷은 각각 De-Secure블록체인임, 안전한 LV Net을 용도별 특화하여 각각 개발하고 상호 연결함
나아가 LV 넷을 타 L1 체인과도 연결하여 web3 생태계 확장에 기여

DeSecure 복합체인 연결 기술

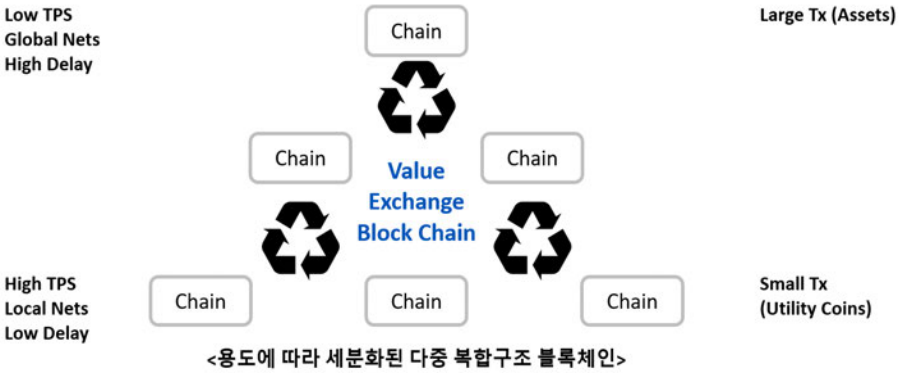
- VCA 합의 프로토콜 기반의 개별 체인 간 및 Ethereum 과 같은 다른 L1 과 연결

기존 L2 솔루션 체인 연결 기술의 문제점

- 브릿지 구조상의 보안성 문제
- 체인간 복잡한 통신 필요, 보안성 미흡

DeSecure 복합체인 연결 기술의 차별성

- VCA 기반의 체인간 연결 효율화
  - 같은 보안 레벨 체인간의 빠른 데이터 교환
  - 다른 보안 레벨 체인간 연결은 높은 보안 레벨에 맞춰서
- Ethereum 과 같은 다른 L1과 연결하여 보안성 강화



<용도에 따라 세분화된 다중 복합구조 블록체인>

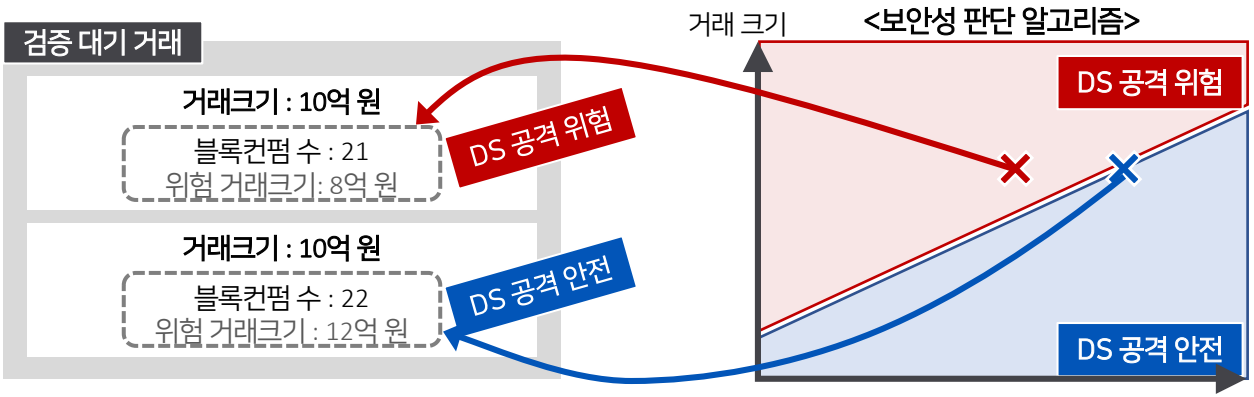
참조발표: https://www.youtube.com/watch?v=VuRgKBDW4Lk&list=PLWq4e5y2b_S9b9AHMPpLyLqF41xBRUd4C&index=2

DeSecure 복합체인 연결 기술 – 개발 요소기술 2

연구 개발 목표: 체인 연결 구조 설계에 있어서 보안 해킹 사고를 저감시키기 위하여
거래 크기에 따라 보안 등급 조절 필요.
Profitable Double Spending Attack* 연구 결과 활용 가능.

DeSecure 연구 개발 목표 : 체인 연결 구조 설계

- 연구 필요성
  - 다양한 체인 연결 구조들이 제안 되었지만 보안성 문제, 프로토콜 복잡성 문제등이 존재함
  - 기존 체인 연결 구조의 장단점에 대한 분석 연구가 필요함
  - 용도별로 보안 파라메터가 다른 체인들을 효율적으로 연결할 수 있도록 하는 새로운 체인 연결 구조의 설계가 필요함
- 연구 세부 내용
  - 체인 보안 레벨 판단 알고리즘 연구
  - 연결된 두 체인의 보안 레벨에 따른 안전한 체인 연결 구조 설계
- Key Idea: 거래 크기에 따라 보안 등급 조절 필요. Profitable Double Spending Attack 연구 결과 활용 가능.



* 참조논문: Heung-No Lee, 2020. https://www.mdpi.com/2076-3417/10/23/8477

04. Technology

경쟁사비교

Zk-STARK 활용 및 새로운 자체 개발 VCA 합의에 기반한 탈 중앙성, 보안성 제고
De-Secure 브릿지 연결, 체인간 보안 레벨 차이에 따른 안전한 연결 구조 설계

Projects	Goal	합의노드 선출방식	합의 알고리즘	합의 알고리즘 Anti-ASIC	확장성 솔루션 제공 여부	DeFi 솔루션 제공 여부	양자내성 지원 여부*	최대 TPS
Polygon	확장성 트릴레마 문제 해결	PoW	GHOST	X	Polygon	O	X (zk-STARKs)	Scalable
Algorand	확장성 트릴레마 문제해결	VRF	Pure PoS BFT	N/A	X	X	X (EdDSA)	45,000 (2021목표)
Stacks	Bitcoin확장성 솔루션	Proof-of-transfer	Verified random function	N/A	X	O	X	~7
Cosmos	*체인간 연결 프로토콜*	PoS	Tender-mint (BFT)	N/A	인터체인 통신	O	X (ECDSA)	Scalable
Solana	*Layer 1*	PoH	BFT	N/A	X	O	X (EdDSA)	Maximum: 710,000
Polkadot	*체인간 연결 프로토콜*	PoS	BFT	N/A	인터체인 통신	예정	X (ECDSA, schnorr, EdDSA)	Scalable
Liber-Vance	**확장성 트릴레마 문제 해결**	**VRF**	**VCA**	**O**	**DeSecure 복합 체인 연결기술**	**예정**	**오류정정부호 기반 양자내성 암호화**	**Scalable**

* RSA, Elliptic Curve 방식의 경우, 양자 내성 확보 불가
(NIST 양자 내성 암호 기술 표준화, 정보통신기획평가원 주간기술동향 1976호, 20.12.09.)

기술개발 내용 및 목표

세계적인 Ethereum L2 Solution 대비 성능 평가 수행
평가 환경에 따른, 공인 평가원 평가를 통한 성능 달성치 제시

성능지표

평가 항목 (주요성능)	단 위	전체 항목에서 차지하는 비중(%)	세계 최고수준 보유국/보유기관 성능수준	연구개발 목표치	목표 설정 근거
블록 생성 시간 (BGT) 표준편차	%	20 %	1.7 % (Polygon, 4월 기준)	< 2 %	선행연구결과: 90%
최소 조정 가능한 BGT 비율	%	10 %	100 % (Bitcoin)	< 100 %	$\frac{\text{현재 } BGT - \text{다음 난이도 } BGT}{\text{현재 } BGT}$
난이도 변동시 BGT 안정 시간	블록	20 %	2016 (Bitcoin)	< 7	Bitcoin 의 블록 컨펌 숫자 이내로 안정화를 목표로 함
VRF 기반 노드 선출 공정성	P-value	20 %	0.9 (Algorand)	> 0.9	균등분포 검정 (Anderson-darling test*) 점수 기준
Profitable DS 공격** 가능성	확률	30 %	1.0 (Bitcoin, Ethereum)	≈ 0.0	현존 블록체인은 해시파워 대여 서비스를 통한 DS공격에 취약함

평가방법

순번	평가항목(성능지표)	평가방법	평가환경
1	BGT 표준 편차	한국시험인증원 평가	네트워크 해시 파워 고정
2	최소 조정 가능한 BGT 비율	한국시험인증원 평가	노드 수, 해시 파워 고정
3	난이도 변동시 BGT 안정 시간	한국시험인증원 평가	현재 해시 파워에서 10% 변동 시
4	VRF 기반 노드 선출 공정성	한국시험인증원 평가	Anderson-darling 테스트, 샘플 수 고정
5	Profitable DS 공격 가능성	한국시험인증원 평가	거래당 최대 크기 기준치를 $단위로 설정

*Anderson-darling test: 데이터가 균등 분포를 따르는지 확인하는 통계적 테스트 방법 (1: 균등, 0: 비균등), 이흥노교수 논문 링크
** Heung-No Lee, 2020. https://www.mdpi.com/2076-3417/10/23/8477

기술개발 현황 및 추진 방향

이흥노 교수 랩 연구 개발 성과로 확보한 핵심 원천기술 (TRL 2~5단계)
TIPS R&D 사업으로 TRL 단계 고도화 및 사업화 추진
서비스 개시를 위해 UI/UX가 제고된 포털 개발, 지갑서비스, 사용자 접근성 제고 등 추진 필요

핵심 기술 개발 현황

기술 분류	개발 현황	개발 목표	세부 개발 목표
VCA 합의 프로토콜	논문 게재 및 특허 출원 완료 핵심 기능 설계 완료 및 시제품 소프트웨어 개발 완료 소규모 테스트 노드 운영중 대규모 테스트 및 부가기능 구현, 기능 고도화 필요 **(TRL 2~ 5단계)**	**블록 생성 시간 안정화**	암호 퍼즐 파라메터에 따른 난이도 변화 분석
			난이도 조절 알고리즘 고도화
			컴퓨팅 파워 변동 대응 알고리즘 구현
		블록 생성 노드 선출 기술	VRF 기반의 블록 생성 노드 선출 기술 개발
			VRF 의 보안성, 양자내성 강화
			블록 생성 노드 선출 프로토콜을 적용하여 VCA 합의 기술 고도화
DeSecure 복합체인 연결기술	핵심 알고리즘에 설계, 실험, 최적화 완료 시스템 전체에 대한 설계 및 테스트 필요 **(TRL 3단계)**	**체인 연결 구조 설계**	체인 보안 레벨 판단 알고리즘 고도화
			체인의 보안 레벨에 따른 안전한 체인 연결 구조 설계

지식재산권

개발 기술 원형의 특허 등록을 통한 지재권 보호 완료
개발 기술을 LV로 이전하여 최고 기술 경쟁력 확보 예정

LiberVance 지식 재산권

- 개발 기술 특허 등록 및 기술이전을 통한 기술 실시권 확보
- LiberVance
  - ✓ 국내 특허: 등록 2건

INFONET 연구실 지식 재산권

- INFONET 연구실
  - ✓ 국내외 논문 300+편 (SCI 107+ 편)
  - ✓ 국제 특허: 출원 9건, 등록 7건
  - ✓ 국내 특허: 출원 7건, 등록 22 건
  - ✓ 기술이전: 3건 (2억 2천만원)

순번	구분	지식재산권 출원인 (실시권자)	출원국	진행상황	출원 및 등록번호	지식재산권(특허)명	비고
1	보유 특허	이흥노 (리버밴스㈜)	한국	등록	10-2288769	**부호-암호 화폐 시스템**	ECCPoW
2	보유 특허	이흥노 (리버밴스㈜)	한국	등록	10-2288776	블록체인의 거래검증시스템, 및 블록체인의 거래검증방법	ECCPoW
3	보유 특허 (기술이전 예정)	이흥노 (INFONET)	한국	출원	10-2020-0171424	블록체인 전자투표시스템, 그 시스템의 운용방법	DeSecure
4	보유 특허 (기술이전 예정)	이흥노 (INFONET)	한국	출원	10-2022-0018393	**스택 머신 프로그램을 위한 범용 영지식 스나크(zk-snark) 증명 시스템 및 방법**	DeSecure
5	보유 특허 (기술이전 예정)	이흥노 (INFONET)	미국	출원	62/699751	새로운 블록체인 및 암호화폐 작업증명 생성, 증명, 검증 시스템	ECCPoW

04. Technology

표준화 전략

Ethereum은 컴퓨팅 블록체인의 세계적 표준 – Ethereum 재단과의 협업 제안서 제출 (WorldLand Proposal) 완료, Ethereum Foundation Ecosystem Project 개발 논의를 위한 Conference Call (6월 중순 부터 말경) 단계 진입

Ethereum 재단의 Academic Grant 제안서 제출 초대 E-mail

---------- 원본 메일 ----------
보낸사람 : Sara Amaral < sara.amaral@ethereum.org>
받는사람 : < heungno@gist.ac.kr>
참조 : George Trotter < george.trotter@ethereum.org>
받은날짜 : 2022/Feb/10(Thu) 22:20:13
제목 : Research Grant opportunity - The Ethereum Foundation

Dear Professor Heung-No Lee
Gwangju Institute of Science and Technology
School of Information and Communications

I am contacting you **regarding a new wave of grants that The Ethereum Foundation will be launching which will target Academics, Research Centers and the like.** We will be funding formal research that aims to create more knowledge about Ethereum, the blockchain technology, and related domains. This round has up to $750.000 in total available funding.

I am contacting you to **share this upcoming grants round opportunity and explore the possibility of sharing this information among your Research Associates, PhD students and all those interested in these fields.**



SHA256 9884192D0D9B0675F4F5CA37264E0E8604A7EC54CE417A22390DEB45203E7B51

WorldLand

Academic grants proposal sent to: academic-grants@ethereum.org.

Submitted by Blockchain Intelligence Convergence center

Director Heung-No Lee

1 Project name: WorldLand

Project cost: $100,000

Project duration: 12 months

Proposed starting date: June 1st, 2022

Submission date: May 8, 2022

Principle investigator: Prof. Heung-No Lee

Contact information: +(82) 10-4946-4710, heungno@gist.ac.kr, heungno@gmail.com, hnlee@ieee.org

Web-page 1: PI's lab home page: https://infonet.gist.ac.kr/

Web-page 2: PI's BIC center home page: http://www.gist-itrc.org/eng/.

Linkedin: https://www.linkedin.com/in/heung-no-lee-4808b38/

Twitter: @fighterkerry

Telegram user name: heungno

The Blockchain Intelligence Convergence (BIC) center was started with national funding by The Ministry of Science and ICT (MSIT, https://www.msit.go.kr/eng/index.do) in 2021 in the area of Blockchain Foundation. MSIT is a primary administrative agency of the Republic of Korea's government that sets, coordinates, and evaluates science and technology policies; and undertakes scientific and technology research, development, collaboration, and promotion. The BIC center's objectives are twofold: research and education. The research goal is to develop innovative new technologies in the converged field of blockchain and artificial intelligence. The center's researchers are working to advance technologies such as Zero-Knowledge Sensing*, Cryptographic Authentication, and Federated Learning. The educational goal is to create new graduate-level courses and programs that will teach advanced masters and doctorate-level scholars in the unique technologies that will be produced at the center.

* Zero-Knowledge Sensing is referred to as a novel technology invented in the center which aims at preserving the physical uniqueness of real-world signals such as human voice, radio waves, and visual feeds while protecting personal information when these signals are converted into data via relevant sensors. Such preserved physical uniqueness can be utilized to authenticate an individual, a material, a location, and a height.

2 Executive summary–Blockchain technology is envisioned to transform the Internet from an information-sharing platform to a metaverse in which citizens worldwide can gather,

1

제출된 제안서 URL
https://ipfs.io/ipfs/QmQ9RCmW7mfxCdqQo4UFd4LS4XwSQNP7oN6j1LVvEdzYa1

ERC-1155 NFT로 민팅된 첫 번째 페이지at https://opensea.io
ID: 63593980875508525246592578297792355851416414116057972066662242316736767459329

기술개발 추진 일정

1차년도 연구 개발을 통한 VCA 합의 프로토콜 고도화 완료.
체인 연결 구조 설계를 통한 LiberVance 플랫폼 서비스 개발 완료, 서비스 시운영 및 개선 기간 추진

수행기관	담당 기술개발 내용	기술개발 비중(%)
리버밴스㈜	LiberVance 플랫폼 기술 개발	100%

분류	개발 기술	기술 세부 내용	추진기간							
			1차년도				2차년도			
			Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
VCA 합의 프로토콜	블록 생성 시간 안정화	암호 퍼즐 파라메터에 따른 난이도 변화 분석								
		난이도 조절 알고리즘 고도화								
		컴퓨팅 파워 변동 대응 알고리즘 구현								
	블록 생성 노드 선출 기술	VRF 기반의 블록 생성 노드 선출 기술 개발								
		VRF 의 보안성, 양자내성 강화								
		블록 생성 노드 선출 프로토콜을 적용하여 VCA 합의 기술 고도화								
DeSecure 복합체인 연결기술	체인 연결 구조 설계	체인 보안 레벨 판단 알고리즘 고도화								
		안전한 체인 연결 구조 설계								
기술 서비스	서비스 런칭 및 개선	LiberVance L2 플랫폼 서비스 시운영 및 개선								

05. Scale-up

목표시장 및 경쟁사 현황 -요약

경쟁사 대비 자체 플랫폼, 복합체인 연결기술, Anti-ASIC 및 양자내성 특성 등, 기술적 우위 확보

기술적 우위에 의한 시가총액 추월 목표

	LiberVance	Decentraland	Stacks	SOLANA	COSMOS ATOM	Polkadot.	polygon
상호	LiberVance	Decentraland	Stacks	Solana	Cosmos	Polkadot	**Polygon**
창립일	2020 년	2017 년	2013 년	2017 년	2017 년	2017 년	**2017 년**
시가 총액	N/A	약 6.5 조 원	약 3.2 조 원	약 55.6 조 원	약 10.7 조 원	약 31.6 조 원	**약 20.2 조 원**
주요 서비스	DeFi, Swap, NFT, 메타버스, 인증	메타버스, NFT	플랫폼 (NFT, DID, Web3)	플랫폼 (NFT, DeFi)	상호운용 플랫폼	상호운용 플랫폼	**오프체인 거래 지원 Layer-2 솔루션**
경쟁사 대비 자사의 비교우위점	-	자체 플랫폼 없음	멀티체인 구현 x		합의 알고리즘 Anti-ASIC 및 양자 내성을 통한 신뢰 검증 기능 부족		

- **자체 플랫폼 (LiberVance L2) 보유**
- **DeSecure 복합체인 연결 기술로 확장성 확보**
- **VCA 합의 알고리즘의 ASIC 내성 및 양자내성을 통한 데이터 신뢰 기능 가능**

05. Scale-up

목표시장 및 경쟁사 현황 -기술 평가 가치

Polkadot, Polygon 대비 자체 합의 알고리즘 보유, 확장성 솔루션 제공 및 양자 내성 특성 지원
진보된 기술력 기반, 시가총액 추월 목표

Projects	Goal	합의노드 선출방식	*독창적 합의 알고리즘*	합의 알고리즘 Anti-ASIC	*확장성 솔루션 제공 여부*	DeFi 솔루션 제공 여부	*양자내성 지원 여부**	최대 TPS (transactions per second)
Polygon	확장성 트릴레마 문제 해결	PoW (Ethereum)	***GHOST***	X	***Polygon (Layer-2)***	O	*X*	scalable
Cosmos/ Polkadot	*체인간 연결 프로토콜*	PoS	***BFT***	X	***인터체인 통신***	O	*X*	scalable
LiberVance	확장성 트릴레마 문제 해결	VRF	***VCA***	O	***DeSecure 복합체인 연결 기술***	예정	*오류정정부호 기반 양자내성 암호화*	scalable

<Cosmos, Polkadot 평가가치>

Cosmos 시가총액 약 10.7 조원
Polkadot 시가총액 약 31.6 조원
(22. 01. 13. xangle 기준)



<Polygon 평가가치>

Polygon 시가총액 약 20.2 조원
(22. 01. 13. xangle 기준)



< LiberVance 평가가치 >

Cosmos, Polkadot 그리고 Polygon 대비 진보된 기술력

1. 합의 프로토콜에서의 비교: VCA 합의 프로토콜의 이점
   - PoS 방식인 Cosmos, Polkadot 대비 높은 보안성
   - ASIC 저항성 기술 및 Time-varying 퍼즐 생성 기술을 적용하여, Polygon 대비 더 탈중앙적
   - BFT 방식인 Cosmos, Polkadot 대비 간단한 구성
2. 체인 연결 기술 측면에서의 비교: DeSecure 기술의 이점
   - 체인 보안 레벨 판단 기술 적용으로 Polygon, Cosmos, Polkadot 대비 높은 보안성

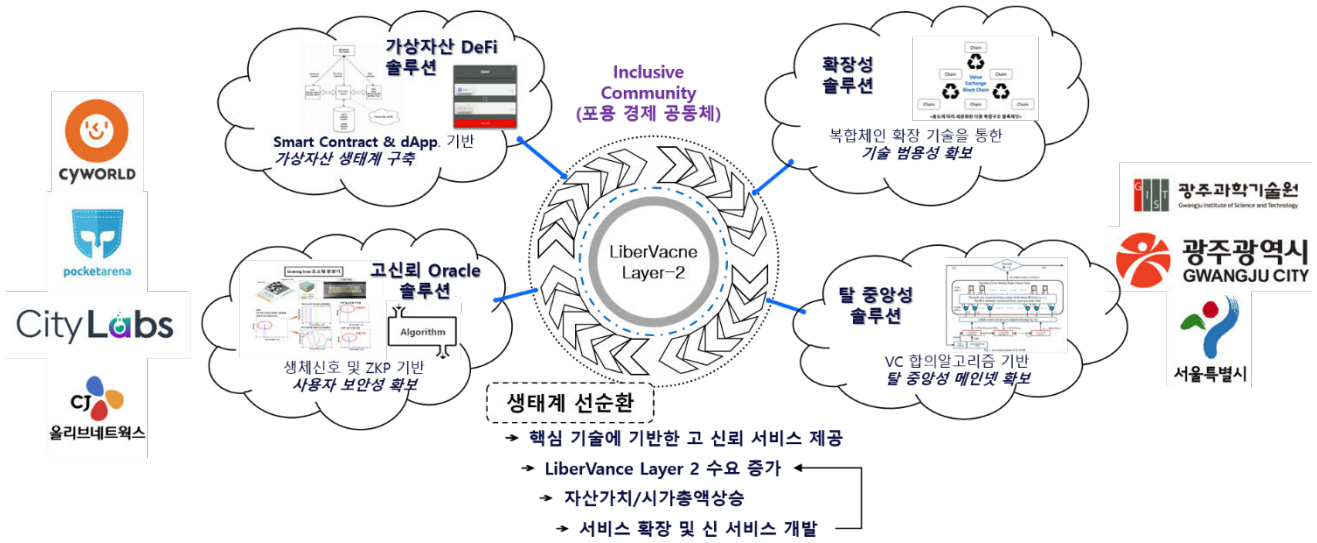
05. Scale-up

사업화 계획 -요약

LiberVance 네트워크 고유의 보안성, 탈중앙성 및 데이터 신뢰성에 따른 안전한 블록체인 플랫폼 서비스 제공
네트워크 사용자 유치를 통한 네트워크 가치 증대 및 포용 경제 공동체 형성

LiberVance 네트워크 포용 경제 공동체 형성

- Layer 2 플랫폼 서비스 기술 제공을 통한, 소요 기업 니즈 충족
- dApp 서비스 개발 과제 수행으로 단기 매출 확보
- 신규 개발 서비스 창출을 통한 LiberVance 네트워크 사용자 유치
- 서비스 증대를 통한 생태계 선순환, 네트워크 가치 제고



05. Scale-up

사업화 계획 (Business model) – LiberVance Net(LV Net)

보다 많은 Key Player에게 탈 중앙 서비스 활용 새로운 Business 기회 Open 가능

디지털 부동산거래, NFT발행, 자산거래소(DEX), 예대마진거래은행, P2E, Stable Coin 등

LiberVance Net Key Players

- LiberVance Net 재단
- LiberVance 개발사
- 투자자 집단
- 노드 사업자 집단
- H/W 인프라 제공 집단 (한전)
- dApp 개발사
- dApp 사용자



LV Swap 서비스 (예시)

- 1st Q
  - AI-DEX 기술개발 및 검증
  - 테스트 컨트랙트 배포(ETH, LVN)
  - 기술 백서 공개
- 2nd Q
  - FE (Swap/Pool/챠트) 오픈
  - LV Swap 런칭
  - 토큰이코노미 백서 공개
  - 글로벌 커뮤니티 구축
    유동성공급자, 사용자, 투자자, 개발자
    Twitter, Telegram, Discord
  - 투자유치프로세스 런칭
- 3rd Q
  - 메인넷 컨트랙트 배포(ETH, LVN, Pol)
  - 유동성 공급사 10개 이상 유치
  - 토큰쌍 100개 이상 생성
  - DAO 오픈
  - Governance 오픈

05. Scale-up

Liber Vance Net(LV Net) Key player

폭증하는 Layer 2 시장 수요 충족을 위한 LiberVance Layer 2 플랫폼 솔루션
리버밴스 생태계 內 플레이어들의 역할

Players in LV NET

재단
- 토큰 발행
- 토큰 상장

투자자
- 토큰과 투자금 스왑
- 토큰 상장 시 EXIT가능

블록체인넷 개발사
- 재단으로부터 개발비 지원
- 프로토콜 완성 및 공개
- 하드웨어에 프로토콜 설치 노드를 VC사업자에 판매

노드 사업자
- 노드 구매 or 직접설치 운영
- 전기사용 노드 운영
- 노드운영 수익
  - ✓ 주조차익, Gas Fee

전기사업자
- 한전 등
- 전기판매

dApp 사업자
- Smart Contract
  - ✓ 지갑
  - ✓ 스왑서비스
  - ✓ 은행서비스
  - ✓ NFT민팅
  - ✓ P2E

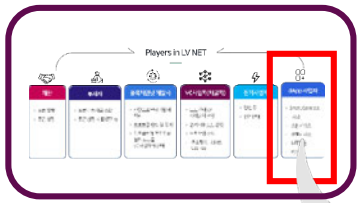
05. Scale-up

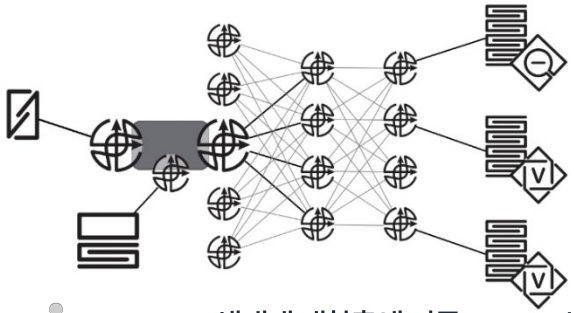
LiberVance Net(LV Net) 생태계 재창출 – 사업 확대

dApp 생태계 內 탈 중앙 서비스 제공을 통한 생태계 재창출

다양한 dApp 사업자들이 오픈 LV Network에 스마트 컨트랙트를 배포 (Wallet, Swap service, Banking, NFT Minting, P2E etc.)

LiberVance Layer 2 네트워크 생태계 재창출 구조





생태계 재창출에 따른 Network effect ↑

dApp(ex. Swap서비스) 생태계

dApp 재단

- 토큰 발행 및 상장
- TVL 상승->토큰값 상승

dApp 개발사

- 재단으로부터 개발비 지원
- 프로토콜 완성/공개/업그레이드 등 새로운 서비스 지속 개발
- V1->V2->V3
- Smart Contract 보완 새로운 솔루션 제시

dApp 사용자

- A토큰을 Pool에 제공
- B토큰을 구매

↓

유동성 공급자에게 수수료 지불

노드 사업자

- TX fee, Gas fee →LV Net의 노드 사업자
- LP fee →LP 공급자
- Swap서비스 개발비→Swap 서비스 개발자

유동성 공급자

- 자본을 투자하여 Swap SV 사용자가 Swap 할 수 있도록 유동성 공급
- 자산pair로 Pool 형성
- (ex)Uniswap의 Pool 개수: 50개

05. Scale-up

타임 라인 별 LiberVance Net(LV Net) 사업화 계획

TIPS 투자 기반 핵심 기술 고도화 개발 완료
네트워크 생태계 확장을 통한 기업가치 성장 동력 확보

타임 라인 별 LiberVance Net(LV Net) 사업화 계획

구분	[단기] 2022 ~ 2023	[중기] 2023 ~ 2024	[장기] 2024 ~
기업 전략 및 방향성	핵심 기술 개발	테스트넷 구축 및 운영	메인넷 (LV-NET) 런칭 및 글로벌 Scale Up 진행
사업 추진 현황 및 계획	• VCA합의기술개발 및 검증 • Ethereum Istanbul HF 및 검증 • 기술 백서 공개	• 테스트넷 런칭 • 토큰이코노미 백서 • 투자유치프로세스 런칭	• 노드운영자 확보 • LV Scan / 지갑 서비스 런칭 • LVNet 스케일업 프로젝트
STAGE 별 사업전략 / 목표	국내외 주요 경쟁 기술 분석 전세계적 수요 기술 파악 VCA 솔루션 기획에 반영 DeSecure 복합체인 기술 확보 LV Net SW Suite완성 양자내성 확보 기술 백서 공개	테스트넷 런칭 토큰이코노미 설계 글로벌 커뮤니티 구축 테스트넷 운영 노드 확보 토큰이코노미 백서 공개 글로벌 커뮤니티 구축	LV 메인넷 런칭 국내외 수요 서비스 발굴 LV-NET 글로벌 스케일 업 LV MainNet 노드운영사 100개 확보 지갑 서비스 사업자 10개사 확보 최소 1개 LV Scan 서비스 개시 멀티체인, dApp, NFT, DeFi, 및 P2E 사업유치

05. Scale-up

사업화 계획 (Business model) – 블록체인 서비스 산업의 Value Chains

블록체인 프로토콜 개발사로서의 오픈소스 비즈니스 모델 - BaaS 서비스 지원

블록체인 dApp 개발사로서의 탈중앙 서비스 개발 솔루션 제공 – 기술용역 서비스 지원

BaaS 서비스 기술용역 서비스

Value Chains	하드웨어 인프라	블록체인 프로토콜 개발		디앱 개발		제품 배포	소비자 (최종 사용자)
	네트워크 하드웨어 제조업체	블록체인 소프트웨어 제조업체	블록체인 프로토콜 운영자	디앱 개발자	디앱 운영 참여자	블록체인 프로토콜/디앱 스토어	서비스 최종 사용자
관련 서비스/ 제품 제공자	• 인터넷 네트워크, IT/모바일 기기 제조업체 • 채굴 하드웨어 제조업체 (CPU, GPU 또는 ASIC)	• 블록체인 프로토콜 개발자	• 채굴자(작업증명) • 합의노드(지분증명, 위임지분증명)	• ***블록체인 디앱 제작/개발업체***	• 서비스 운영에 참여하는 디앱 사용자	• 블록체인 프로토콜/디앱 스토어	• B2B 소비자 • B2C 소비자
수익 모델	• 하드웨어 판매 수익 • 서비스 플랫폼과의 연계 수익 • 채굴 관련 서비스	• 거래 수수료 • 암호화폐 발행에 따른 주조 차익 (시뇨리지) • ***오픈소스 비즈니스 모델***	• 거래 처리 수수료 • 암호화폐 발행에 따른 주조 차익 (시뇨리지)	• 거래 수수료 • 광고 • 프리미엄 기능 • 멤버십 • 구독	• 운영 보상	• 광고 • 스토어 내 결제 서비스료	• 해당 없음
핵심 요소	• 채굴 하드웨어/ 블록체인 서비스에 대한 수요 • 하드웨어 효율성 • 채굴 보상의 가치 • 넓은 사용자 기반 및 lock-in 효과	• 개발 친화적 환경 • 활발한/광범위한 베이스 • 네트워크 효과/가치	• 합의알고리즘의 효율성 및 보안 • 운영 보상	• 활발한/광범위한 사용자 베이스 • 네트우크 효과/가치	• 폭 넓고 적극적인 사용자 기반 • 네트워크 효과/가치 • 운영 보상	• 배포용 스토어의 트래픽 양 • 제공되는 유효 디앱의 수 • 사용자 lock-in 효과	• 제공되는 가치 • 블록체인 서비스에 따라 상이
주변 기능	- 마케팅 판매 - 거버넌스 서비스 - 엑셀러레이팅/컨설팅 - 규정 준수/법률 -감사 -세금						

출처: Block Crafters Research

사업화 계획 (Business model) – LiberVance BM Models

블록체인 기술 인프라 제공자로서의 BM모델
소프트웨어 제조 수익, 프로토콜 운영 수익, 그리고 dApp 제작 개발 수익

LiberVance Layer 2 플랫폼 사업모델

- **본사의 포지션**: 블록체인 소프트웨어 인프라 제공자
- **영업 형태**: 1. 블록체인 인프라 제공, 2.블록체인 서비스 제공, 3. dApp 제작 개발 (용역)

1. L2 네트워크 운영자로서의 수익모델

수익모델 후보	비고
거래 수수료/Block Reward	Ethereum 10조/Polygon 83억

2. 블록체인 소프트웨어 제조업체로서의 수익모델

수익모델 후보	비고
거래 수수료	네트워크 성장에 따라, 노드 사업자 간 경쟁으로 결정
BaaS 서비스	***블록체인 프로젝트 구축에 어려움을 겪는 그룹에 대한 지원 서비스***

3. dApp 제작/개발사로서의 수익모델

수익모델 후보	비고
기술 용역서비스	***블록체인 프로젝트 자체에 대한 dApp 개발 용역 서비스***
프리미엄 기능/멤버쉽/구독	개발 서비스 및 dApp 종류에 따라 제공

사업화 계획 (Business model) – 경쟁사 Value Proposition 분석

Polygon 네트워크 대비 기술 우수성에 기반한 LiberVance 생태계 확장 추구
Layer 2 시장을 주도하는 Polygon과 경쟁을 목표

Polygon 대비 우수한 보안성, 탈중앙성 보유

구분	Polygon	LiberVance
개요	• 이더리움의 인터넷을 표방하는 이더리움 L2솔루션 • L2 중 가장 많은 dApp이 폴리곤에서 출시 (약 600여개) • 이더리움의 속도해결 및 L2시장의 가능성을 보여주는 대표 프로젝트 • Proof of Stake (PoS)	• 리버밴스가 개발한 이더리움 L2솔루션 • 다층 다용도 복합체인 및 연결기술의 블록체인 메인넷 솔루션 • 특화 블록체인 간 연결을 통한 확장성 제고 • 양자내성 VCA 합의 알고리즘
장/단점	• 플라즈마 브릿지 선택 시 출금에 7일이 걸리는 단점 존재 • 최근 코드 임의 하드포크로 인한 중앙화 비판 • 폴리곤 기반 디파이 프로토콜(폴리네트워크) 최대 해킹 피해('21.8월) • 네트워크의 보안 결함 우려 존재	• PDSA 기술 활용, 브릿지 Security강화 • 매 블록마다 새로운 crypto puzzle 제시 탈 중앙성 강화 • 양자내성 보안성과 데이터 신뢰성 제고로 해킹 위험 저감 • 투명한 공개체인 운영, 사용자 신뢰성 제고
사업영역	• 블록체인 플랫폼 사업자/개발사	• 블록체인 플랫폼 사업자/개발사+ 디앱 개발사

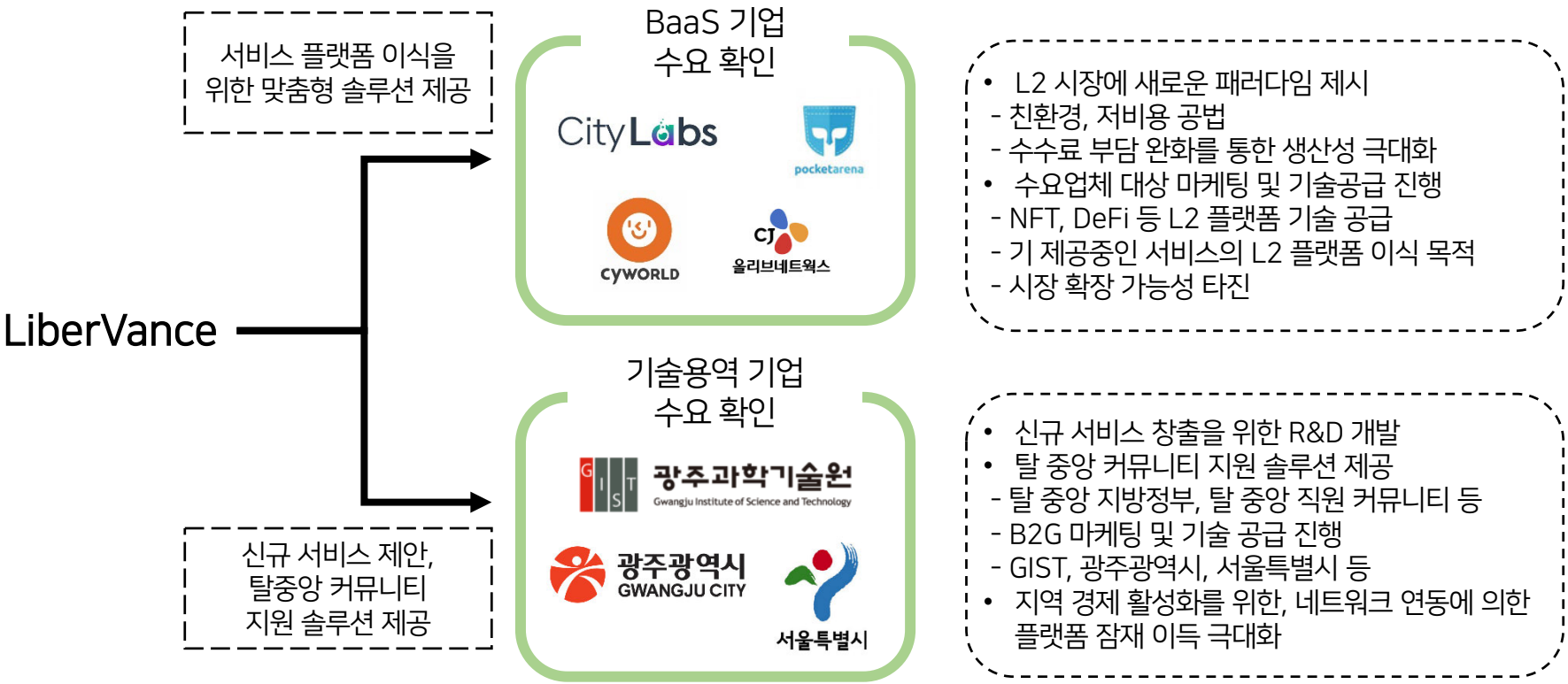
05. Scale-up

사업화 계획 (Business model) – LiberVance 네트워크 Ecosystem 확장 추동

기존 서비스의 Layer 2 플랫폼 이식을 위한 맞춤형 솔루션 제공 – BaaS 서비스 지원

신규 서비스 창출 및 탈중앙 커뮤니티 개발 솔루션 제공 – 기술용역 서비스 지원

비 전공 기술 소요기업/정부/지자체 등 에게 손쉬운 블록체인 솔루션 제공



- L2 시장에 새로운 패러다임 제시
  - 친환경, 저비용 공법
  - 수수료 부담 완화를 통한 생산성 극대화
- 수요업체 대상 마케팅 및 기술공급 진행
  - NFT, DeFi 등 L2 플랫폼 기술 공급
  - 기 제공중인 서비스의 L2 플랫폼 이식 목적
  - 시장 확장 가능성 타진

- 신규 서비스 창출을 위한 R&D 개발
- 탈 중앙 커뮤니티 지원 솔루션 제공
  - 탈 중앙 지방정부, 탈 중앙 직원 커뮤니티 등
  - B2G 마케팅 및 기술 공급 진행
  - GIST, 광주광역시, 서울특별시 등
- 지역 경제 활성화를 위한, 네트워크 연동에 의한 플랫폼 잠재 이득 극대화

05. Scale-up

LV BaaS / 기술용역 사업화 시뮬레이션

고객의 니즈에 따른 Price Tier 선택 전략
선택적 추가 지원에 따른 효율적 서비스 제공

고객이 원하는 기능에 따라 Price Tier 선택:

[실버]
- 사설 네트워크 또는 공개 네트워크 구축 지원, 기본적인 파라메터 (블록생성주기, 코인보상 등) 커스터마이징

[골드]
- 주요 로직 수정 및 새로운 기능 탑재 지원

[플래티넘]
- 고객사가 구축하고자 하는 어플리케이션에 최적화된 특화 프로토콜 설계 및 개발 지원

유지보수 서비스 및 기타 기술 용역 개발 서비스 별도 지원

고객사 자체 개발 대비 장점

- 프로토콜 개발사의 기술 지원으로 자체개발 대비 작업 기간 및 작업 비용 절감
  - 40% ~ 55% 비용 절감
  - 35% ~ 65% 시간 절감
- 고객사의 자체 개발 역량에 따라 Price Tier 선택
- 업데이트, 유지보수 및 운영 책임에서 자유로움

BaaS 서비스 지원	LV public network	Silver	Gold	Platinum
네트워크 구축 지원		O	O	O
프로토콜 커스터마이즈 (basic parameters)		O	O	O
프로토콜 커스터마이즈2 (logics, new features)			O	O
어플리케이션 특화 프로토콜 설계 및 개발				O
LV BaaS 서비스 비용	-	80 백만원 / 2 개월	250 백만원 / 6 개월	500 백만원 / 12 개월
경쟁사/자체개발 추정 비용	-	180 백만원 / 6 개월	470 백만원 / 12 개월	830 백만원 / 18 개월

유지보수 서비스	LV public network	실버	골드	플래티넘
유지보수 지원	Free	9 백만원 /Year	10.8 백만원 /Year	15 백만원 /Year
노드 운영 대행	네트워크 규모에 따라 상이			

기술 용역 개발	비용
지갑 연동	Metamask 연동: 기본 지원
	추가 지갑 연동: 1백만원/건, 작업 기간 10일
	커스텀 지갑 개발 : 28 백만원 / 건 ~
dApp 개발	개발 규모에 따라 상이

05. Scale-up

LiberVance 궁극적 목표는 LiberVance Ecosystem 확장
손쉽고 편리한 블록체인 도입이 가능하도록 LV Network 확장 플랫폼 제공

LV Net, what is it?

LVNET은 손쉬운 블록체인 솔루션 도입을 위한 플랫폼

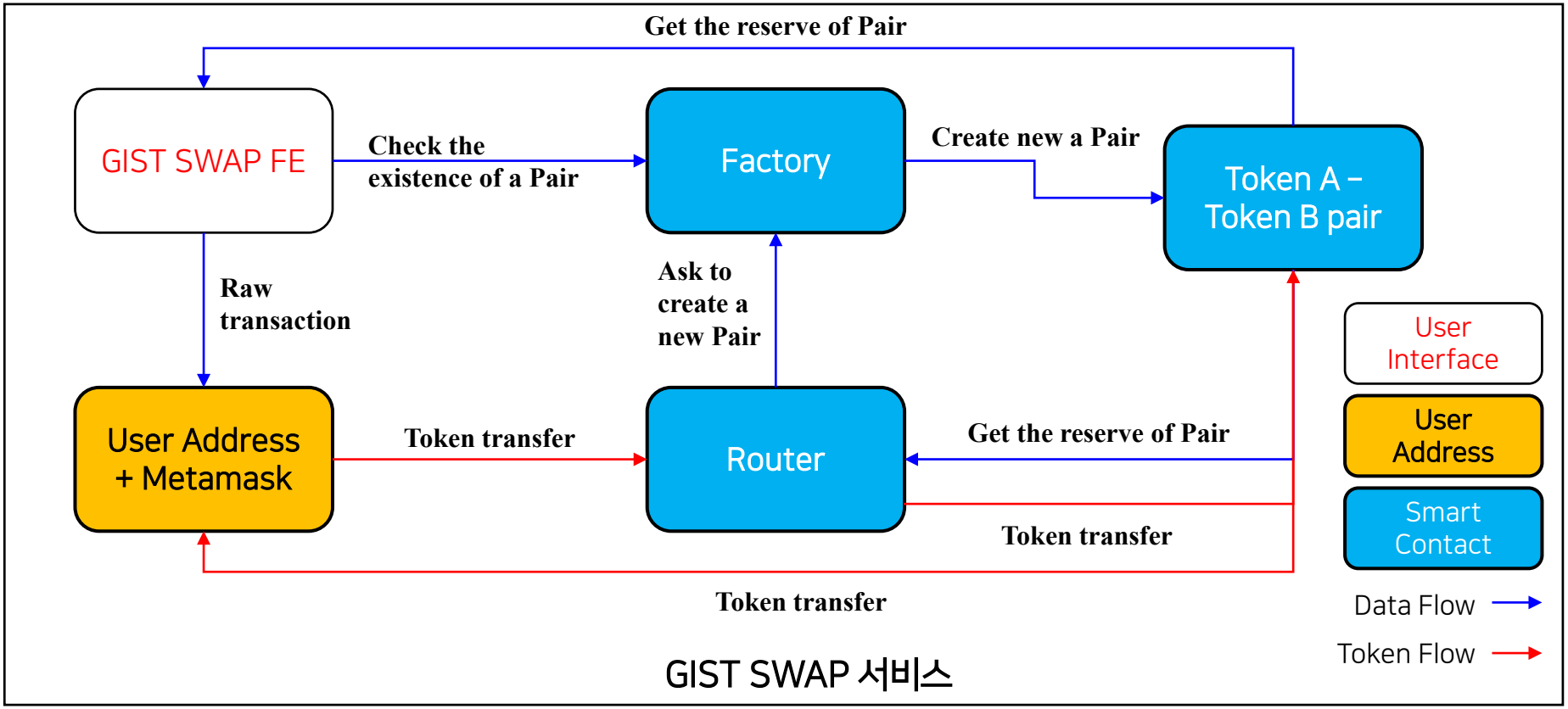
모듈식 프레임워크 제공	옵티미스틱 롤업 체인 / ZK 롤업 체인 / ZKSync 독립형 체인 또는 개발자가 요구하는 기타 인프라 구축
Ethereum VM 호환성	플라즈마 프레임워크와 지분증명 블록체인 아키텍처 / 확장 가능하고 자율적인 스마트 계약의 손쉬운 실행 가능 / 새로운 기능이 기존 증명 기술에 더불어 구축되고 능력을 확대하여 개발자 생태계의 다양한 니즈 충족 / 핵심 기술을 지속적으로 개발하여 더 큰 생태계로 확장
이더리움 멀티 체인 시스템 전환 추동	폴카닷, 코스모스, 아발란체 등 다른 시스템과 연계 이더리움의 보안성, 활기찬 생태계와 개방성의 이점 활용 및 확장

05. Scale-up

dApp 개발 매출 사례 (2021)

Layer 2 기술 기반 신규 dApp 서비스 개발 - SWAP, NFT 서비스

신규 서비스 개발 용역을 통한 소요 기업 니즈 충족



05. Scale-up

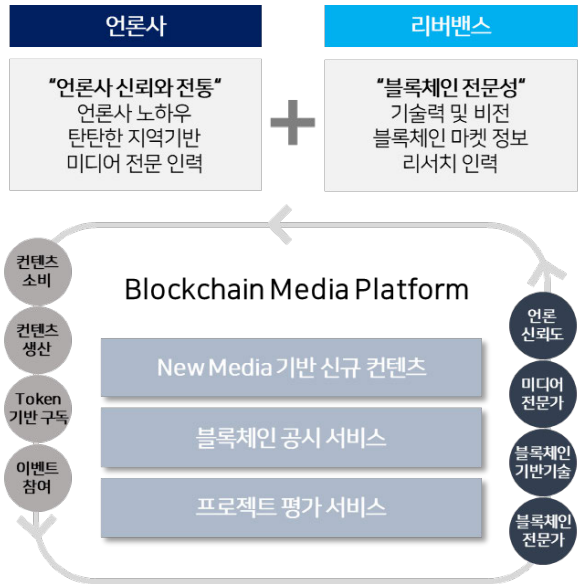
현재 논의 중인 프로젝트

기술력과 개발 경험 보유, 협업 희망하는 기업 컨택 다수

새로운 서비스로 경쟁력 확보 희망하는 기업들의 블록체인 프로젝트, NFT Minting

언론사 미디어 플랫폼 구축

- 지역 언론사와 공동으로 블록체인 미디어 플랫폼 구축



지역 수제맥주 NFT 발행

- 지역 수제맥주 브랜드 NFT Minting
- 부산 갈매기브루잉, 서울 어메이징브루어리 논의중
- 11월 24일 지역 기업가 대상 소규모 발행



B2B 클라우드 서비스 (https://uws.co.kr/)

- UWS (대표 이현주) 사업 추진 중
  - dYdX/Pancake Swap DEX 개발
  - Uniswap AMM 교육
  - AI-DEX 개발
  - 스테이블 코인 개발
  - 개발과 관련된 자문 및 얼라이언스 사업 추진
- UWS는 SK㈜, KB금융그룹, SK텔레콤, LG화학에 클라우드 서비스 제공하는 중견 기업

05. Scale-up

글로벌 진출전략

블록체인 네트워크에서 중요한 커뮤니티 형성을 위해:

- ✓ 글로벌 그룹과의 MOU체결을 통한 해외 Layer 2 시장 진출 목표
- ✓ 국제 커뮤니티 행사 참석 및 적극 홍보를 통한, 글로벌 개발자 커뮤니티 교류 추구

글로벌 진출 목표

추진 목표	내용
해외 블록체인 시장 수요 확보	글로벌 개발사, 사업자 등과 MOU 체결 및 Layer 2 프로젝트 협력
해외 개발자 커뮤니티 구축	개발 기술의 적극적 홍보를 통한 글로벌 개발자 커뮤니티 교류 추구
글로벌 VCA 벨리데이터 네트워크 구축	VCA 노드 운용하는 벨리데이터 모집
글로벌 MOU 및 합작 법인 설립	Validator 확보를 통한 네트워크 생태계 확장 및 해외 시장 고객 니즈 확보

글로벌 진출 추진 일정

추진기간	사업추진 내용(과업)	비고
2022.06 ~ 2022.12	해외 Validator 후보군과의 MOU 체결 / 합작 법인 설립	Europe, Dubai, Taiwan 등
2023.01 ~ 2023.12	해외 Validator 노드 확보	*2곳 이상*
2023.01 ~ 2024.12	Validator 파트너 방문, 사업 설명 및 리크루팅을 통한 국제 개발자 네트워크 협력 추구	
2024.01 ~ 2024.12	해외 시장 니즈 확보(Layer 2 Baas / 기술용역)를 통한 LiberVance Layer 2 네트워크의 생태계 확장	*2건 이상*

05. Scale-up

글로벌 진출전략

글로벌 그룹과의 MOU 체결을 통한 시장 진출 전략

핵심 팀 멤버 이흥노 대표 – (국제학술활동) 톱저널 편집위원, CVA 멤버, Consensus 2019, ConsenSys MOU

글로벌 진출을 위한 커뮤니티 참석 성과 및 계획

행사명	참여 형태	일시, 장소
Crypto Valley Conference on Blockchain Technology (CVCBT) 2018	ECCPoW 홍보 및 개발자 커뮤니티 교류	18.06.20 ~ 18.06.22
Crypto Valley Association (CVA) New York Consensus 2019	DeSecure Blockchain 연구 프로젝트 ***홍보 부스 참석***	19.05.13 ~ 19.05.16
EuroCrypt 2022	**완전동형암호 Bootstrapping *기술 논문 오랄 발표***	**22.05.30 ~ 22.06.03, Norway**
Consensus 2022	**LiberVance 네트워크 기술 홍보 및 개발자 커뮤니티 교류 참여**	**22.06.09 ~ 22.06.12, Texas**
Future Blockchain Summit	**LiberVance 네트워크 기술 홍보 및 체험 부스 개설**	**22.10.10 ~ 22.10.13, Dubai**

* 외 AsiaCrypt 2022 (Taiwan), Crupto 2022 (USA), DeFi Conference 2022 (Ireland) 등 국제 개발자 커뮤니티 적극 참여 예정



글로벌 ConsenSys와 MOU



Consensus 2019, NY, USA

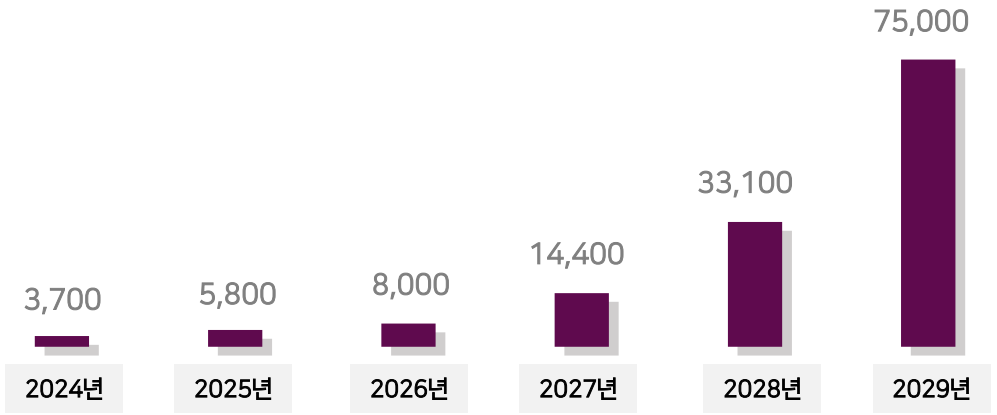


IEEE Trans. 편집위원

05. Scale-up

서비스 매출 부분 고용창출 및 기대효과

신규 서비스 창출을 통한 기업가치 재고 및 고용창출
우수 연구 인력 교육 및 청년 고용 창출



단위 : 백만원,%

매출계획

사업화 성과	세부 성과지표	2024년 (개발종료 해당년)	2025년 (개발종료 후 1년)	2026년 (개발종료 후 2년)	2027년 (개발종료 후 3년)	2028년 (개발종료 후 4년)	2029년 (개발종료 후 5년)
기업 전체 성장	예상 총매출액(A)	3,700	5,800	8,000	14,400	33,100	75,000
개발기술의 사업화 성과	예상 연구개발결과물 제품 매출액(B)	3,700	5,800	8,000	14,400	33,100	75,000
	연구개발결과물 제품 점유비율(C)(C=(B/A)*100)	100	100	100	100	100	100

고용계획

구분	2022년	2023년	2024년	2025년 (개발종료 후 1년)	2026년 (개발종료 후 2년)
신규고용(명)	3	3	5	7	6
상시고용(명)	1	4	7	12	19
계	4	7	12	19	25

사업화 목표 산정 근거 – 매출 추정 근거

BaaS 및 dApp 개발 용역 수행에 기반한 매출 확보
2029년, 영업이익 500억 달성 목표

추정 근거

- 2022년 고용인원 5인 인건비 평균 80백만원 계상, 개발자 판공비 평균 40백만원 계상
- BaaS (초급 1, 중급 1), 기술용역 (초급 1, 중급 1, 고급 1) 계상
- 판관비, 마케팅비, 장비구매비 별도

구분 (단위: 백만원)	제품 구성비	제조 원가	판매가
BaaS	45%	192 (백만원/연,건)	500 (백만원/연,건)
기술 용역	55%	360 (백만원/연,건)	1100 (백만원/연,건)

요약 매출 추정 표

구분 (단위: 백만원)		2022	2023	2024	2025	2026	2027	2028	2029
매출	Baas (건)	500 (1)	1,000 (2)	1,500 (3)	2,500 (5)	2,500 (5)	4,500 (9)	10,000 (20)	20,000 (40)
	기술용역 (건)	1,100 (1)	1,100 (1)	2,200 (2)	3,300 (3)	5,500 (5)	9,900 (9)	23,100 (21)	55,000 (50)
	Total	1,600	2,100	3,700	5,800	8,000	14,400	33,100	75,000
제조 원가		780	972	1,524	2,268	2,988	5,196	11,628	25,908
영업 이익		820	1,128	2,176	3,532	5,012	9,204	21,472	49,092
영업이익율		51%	54%	59%	61%	63%	64%	65%	65%

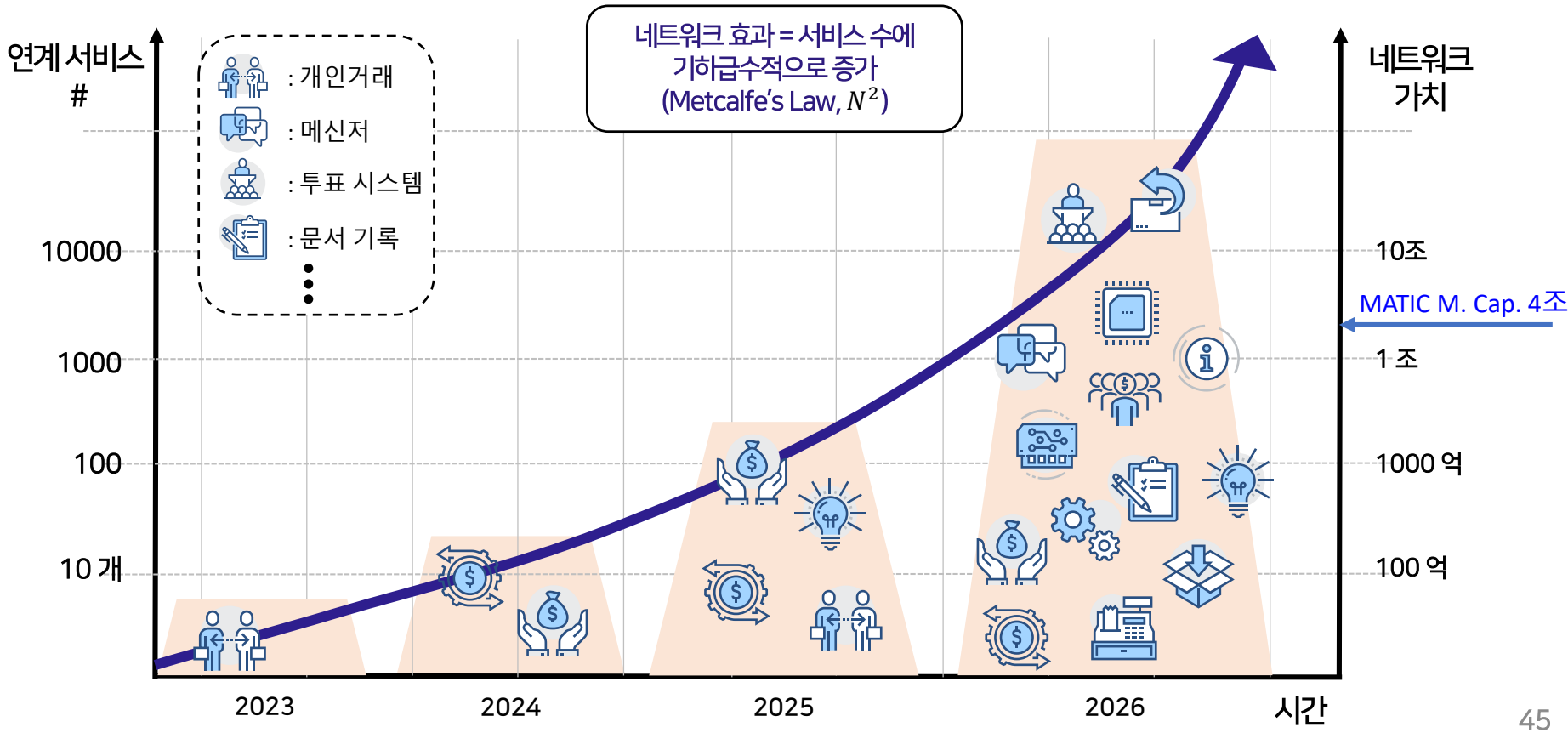
05. Scale-up

플랫폼 부분 사업 목표

전 세계로 탈중앙/보안성이 우수한 LV 플랫폼 제공, 지수적 네트워크 가치 성장을 목표함
2026년, LV 네트워크 가치 10조원 달성 목표

LiberVance 네트워크 가치 성장 목표

[2023-2026]
Network Value



06. Team

광주과학기술원의 우수한 인적/연구/개발 인프라를 활용한 교원 창업

탈 중앙 블록체인 기술 전문가, 양자내성·동형 암호 전문가 를 포함한 최적의 블록체인 암호화 기술 전문가팀

학계를 통해 숙련된 연구 개발자 교육 및 졸업생 리쿠르팅을 통한 핵심 개발 인력 확보



CEO 이흥노

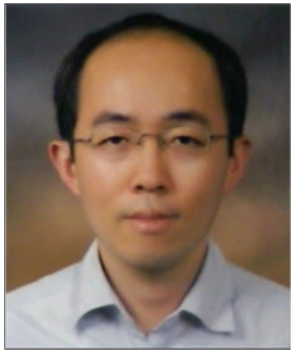
GIST 전기전자컴퓨터공학부 정교수

대통령 직속 정책기획위원회 전문위원

ITRC 블록체인 지능 융합 센터장

국내외 논문 400여편 및 60여건 국내외 특허 보유

Associate Editor of IEEE Transactions on Cybernetics (세계 톱4% SCI Journal) 등 다수 국제 저널의 편집 위원 수행 중



기술이사 김영식(주요주주)

조선대학교 정보통신공학부 정교수

양자내성암호연구단(KpqC) 기술자문위원

IEEE Globcom2022 SAC Cloud, Technical Program Committee Members

동형암호 관련 EUROCRYPT 2022/2021 게재 외 국내외 논문 160여편

기술개발 인력



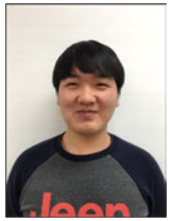
강주성 연구원

- RF 신호 기반 오라클 솔루션
- 복합체인 기술 개발



최해웅 연구원

- 합의 알고리즘 기술
- VCA 합의 프로토콜 개발



김철순 연구원

- 생체신호 기반 오라클 솔루션
- 복합체인 기술 개발

마케팅 및 경영지원 인력



이명은 팀장

- 대외협력
- 기술 마케팅 지원



권효민 연구원

- 마케팅
- 기술 마케팅 지원



정우정 연구원

- 경영지원
- 예산 및 경영 지원

운영사 지원계획

운영사 선보엔젤파트너스 네트워크 및 직접사업을 통한 JV 및 블록체인, NFT 프로젝트 등 연계
해외 지사를 활용한 블록체인 글로벌 커뮤니티 형성 협업 가능

협업가능한 기업 발굴

- 'Web3.0' 관련 포트폴리오 확보를 위한 투자 진행 중
- 다양한 블록체인 기업에 투자를 통해 리버밴스와의 협업을 고려한 포트폴리오 확보 계획

네트워크 활용 콜라보레이션 프로젝트

- 선보엔젤파트너스 관계사인 '인연엔터테인먼트'의 소속 배우 및 가수의 NFT 발행
- 언론사 네트워크 '부산일보', '무등일보'의 블록체인 미디어 플랫폼 콜라보레이션 프로젝트 수행 제안
- 부산지역맥주 '갈매기브루잉' 투자 및 직접사업(맥주NFT 프로젝트) 추진 중에 있어 리버밴스와 공동 런칭 논의 가능

해외 지사 네트워크를 통한 커뮤니티 형성

- 싱가포르 및 독일에 지사를 두고 투자 및 산업네트워크를 확장하고 있는 선보엔젤파트너스를 통해 리버밴스 확장을 주도할 수 있는 글로벌 커뮤니티 형성 계획

GALMEGI BREWING CO.

INYEON ENTERTAINMENT

부산일보

광주·전남 대표 정론지 무등일보

감사합니다!

비즈니스 협력 연락처

대외협력 이명은 팀장: myeongeun@gist.ac.kr
Phone:
이흥노 대표: heungno@gist.ac.kr
Phone: 010-4946-4710

시험성적서

한국시험인증원㈜
서울시 강서구 공항대로219 808호
Tel:02-6929-1033 Fax:02-6929-1037

성적서번호 : KST-19-045

KOTCA 한국시험인증원 Korea Testing Certification Agency

1. 의 뢰 자
 - 회사(기관)명 : 광주과학기술원
 - 주 소 : 광주광역시 북구 첨단과기로 123(오룡동)

2. 시험명 : V&V(Verification & Validation) 시험

3. 시험대상품목/물질/시료명
 - 제품명 : 확장 가능한 탈중앙화 보안성 ECCPoW 블록체인 v1.0
 - 제조사 : 광주과학기술원

4. 시험기간 : 2019년 11월 14일 ~ 2019년 11월 18일

5. 시험방법 : 의뢰자 제시기준

6. 시험결과 : '시험성적서 4. 시험결과' 참조

비고 : 1. 이 성적서는 의뢰자가 제출한 시료에 대한 시험결과이며, 한국시험인증원㈜의 사전 동의 없이 본성적서의 전부 혹은 일부를 복사할 수 없습니다.
2. 이 성적서의 시험결과 중 * 표시된 시험결과는 시험기관의 인정범위 밖의 것임을 밝힙니다.
3. 이 성적서의 진위확인은 기업지원플러스(http://www.g4b.go.kr) 웹페이지에서 G4B조회코드로 확인 가능합니다.

확 인	시험자 성 명 : [redacted] (서명)	기술책임자 성 명 : [redacted] (서명)

2019년 11월 20일

한국시험인증원㈜ 대표이사

TPS-0024-1(01)

G4B(www.g4b.go.kr)진위확인코드 : ZsyxL7847CU=

1. 시험 개요

1.1 시험 기준

- 의뢰자 제시 기준

1.2 시험 장소

- 서울시 강서구 공항대로 219 808호 한국시험인증원 시험실

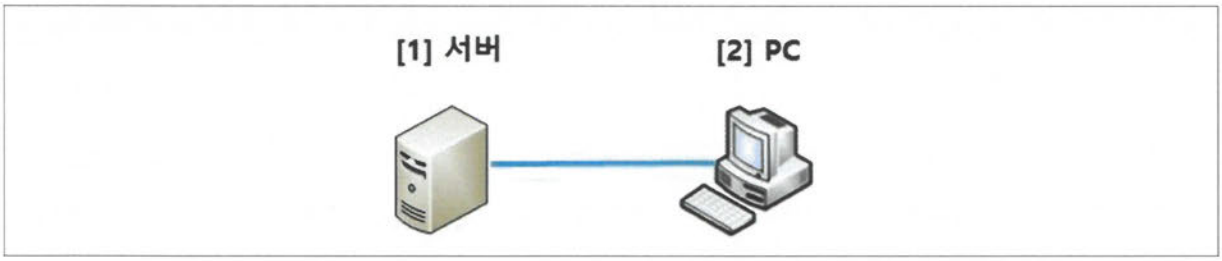
1.3 시험 대상 제품

- 제품명 : 확장 가능한 탈중앙화 보안성 ECCPoW 블록체인 v1.0
- 제조사 : 광주과학기술원

2. 시험 환경

시험환경은 한국시험인증원 시험실에 구축하였으며, 시험환경구성 및 세부사양은 아래와 같다.

<시험환경구성도>



<세부사양>

No	Role	OS	CPU	Memory	HDD/SSD	Pre-Requisite
1	서버	Amazon Web Services EC2				-
2	PC	Ubuntu 18.04 LTS	Intel(R) Core(R) CPU i5-7600U @ 2.60GHz	16GB	SSD 256GB	-

G4B(www.g4b.go.kr)진위확인코드 : ZsyxL7847CU=

3. 시험 목적 및 방법(의뢰자 제시)

3.1 시험항목 1

시험내용(의뢰자 제시)

1. 시험항목명 : 분산성 평가

2. 시험목적

○ ECCPoW의 채굴성공분포도 계산

$$\text{채굴성공분포도}(\%) = \frac{(\text{채굴성공개수의평균})}{\sqrt{(\text{채굴성공개수의분산}) + (\text{채굴성공개수의평균})^2}} \cdot 100$$

3. 시험방법

1) AWS instance 생성

Instance 역할	생성개수	인스턴스종류	vCPU	RAM
Seed Instance	3개	m5.xlarge	2	8GB
Mining Instance	20개	m5.xlarge	2	8GB

2) 블록 생성 목표 시간을 3분으로 설정
3) 난이도 변경 주기는 20 블록으로 설정
4) ECCPoW 채굴을 위한 스크립트 실행
5) 블록 100개 생성 시점에서 채굴 종료
6) 채굴 결과가 저장된 csv파일을 확인하여 채굴 성공 개수 확인

G4B(www.g4b.go.kr)진위확인코드 : ZsyxL7847CU=

3.2 시험항목 2

시험내용(의뢰자 제시)

1. 시험항목명 : 보안성 평가

2. 시험목적

○ Bitcoin과 ECCPoW의 고아 체인 블록 비율 계산

$$\text{고아 체인 블록 비율} = \frac{\text{고아 체인 블록 개수}}{\text{블록체인의 총 높이}} \cdot 100$$

- 고아 체인 블록 개수 : 주 체인으로부터 분기되어 높이가 41 이상인 체인에 속한 블록 개수
- 블록체인의 총 높이 : 주 체인의 높이

3. 시험방법

1) AWS instance 생성

구분	Instance 역할	생성개수	인스턴스종류	vCPU	RAM
Bitcoin	Seed Instance	3개	m5.xlarge	2	8GB
Bitcoin	Mining Instance	20개	m5.xlarge	2	8GB
ECCPoW	Seed Instance	3개	m5.xlarge	2	8GB
ECCPoW	Mining Instance	20개	m5.xlarge	2	8GB

2) 블록 생성 목표 시간을 3분으로 설정
3) 난이도 변경 주기는 20 블록으로 설정
4) Bitcoin과 ECCPoW 채굴을 위한 스크립트 실행
5) 블록 100개 생성 시점에서 채굴 종료
6) Bitcoin과 ECCPoW의 블록체인을 시각화하여 PNG파일로 변환
7) PNG파일에서 높이가 41 이상인 고아 체인의 블록 개수 확인

G4B(www.g4b.go.kr)진위확인코드 : ZsyxL7847CU=

3.3 시험항목 3

시험내용(의뢰자 제시)

1. 시험항목명 : 확장성 평가

2. 시험목적

○ Bitcoin과 ECCPoW의 최근 생성된 10개 블록 TPS 측정

$$TPS = \frac{\text{10개 블록의 총 보유 트랜잭션 수}}{\text{10개 블록의 총 생성 시간}}$$

3. 시험방법

1) AWS instance 생성

구분	Instance 역할	생성개수	인스턴스종류	vCPU	RAM
Bitcoin	Seed Instance	3개	m5.xlarge	2	8GB
Bitcoin	Mining Instance	20개	m5.xlarge	2	8GB
ECCPoW	Seed Instance	3개	m5.xlarge	2	8GB
ECCPoW	Mining Instance	20개	m5.xlarge	2	8GB

2) 블록 생성 목표 시간을 3분으로 설정
3) 난이도 변경 주기는 20 블록으로 설정
4) Bitcoin과 ECCPoW 채굴을 위한 스크립트 실행
5) 블록 100개 생성 시점에서 채굴 종료
6) 최근 발생된 10개 블록의 생성시간과 블록에 포함된 트랜잭션 개수 확인

G4B(www.g4b.go.kr)진위확인코드 : ZsyxL7847CU=

4. 시험 결과

4.1 시험항목 1

시험항목	분산성 평가
시험목적	ECCPoW 채굴성공분포도 계산

시험결과

○ ECCPoW 채굴성공률분포도 : 92.21944%

$$\text{채굴성공분포도}(\%) = \frac{6.2}{\sqrt{6.76 + 38.44}} \cdot 100$$

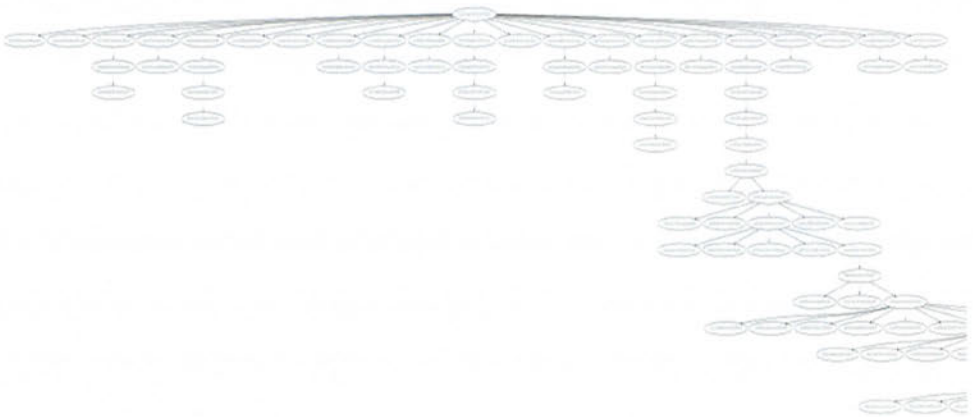
전체 채굴 성공 개수	채굴 성공 개수의 평균	평균 채굴 성공 개수의 제곱	채굴 성공 개수의 분산	채굴성공 분포도
124	6.2	38.44	6.76	92.21944

- 채굴노드 및 채굴 성공 개수

채굴노드	채굴 성공 개수	채굴노드	채굴 성공 개수
1723171	4	172312222	7
17231050	9	172313135	4
172310105	3	172316142	12
172310123	4	172316250	5
172311029	6	172319215	11
172311339	4	1723110197	2
172311355	6	1723112243	7
172311463	6	1723114125	6
172311482	5	1723114253	10
172312158	5	1723115189	8

G4B(www.g4b.go.kr)진위확인코드 : ZsyxL7847CU=

4.2 시험항목 2

시험항목	보안성 평가
시험목적	Bitcoin과 ECCPoW의 고아 체인 블록 비율 계산
시험결과	ㅇ Bitcoin 고아 체인 블록 비율 : 0% - Bitcoin 고아 체인 블록 개수 : 0 ㅇ ECCPoW 고아 체인 블록 비율 : 0% - ECCPoW 고아 체인 블록 개수 : 0 - Bitcoin 고아 체인 블록 화면  - ECCPoW 고아 체인 블록 화면 

G4B(www.g4b.go.kr)진위확인코드 : ZsyxL7847CU=

4.3 시험항목 3

시험항목	확장성 평가
시험목적	Bitcoin과 ECCPoW의 TPS 측정

시험결과

○ Bitcoin TPS : 0.945922747건/초
○ ECCPoW TPS : 0.940334129건/초

- Bitcoin

구분	Height	블록생성 소요시간(초)	트랜잭션수(건)
1	91	719	672
2	92	19	22
3	93	94	88
4	94	144	141
5	95	275	263
6	96	40	40
7	97	313	296
8	98	6	11
9	99	574	534
10	100	146	137
합계		2330	2204
TPS		0.945922747초/건	

G4B(www.g4b.go.kr)진위확인코드 : ZsyxL7847CU=

시험결과

- ECCPoW

구분	Height	블록생성 소요시간(초)	트랜잭션수(건)
1	91	151	140
2	92	192	182
3	93	369	351
4	94	361	331
5	95	156	151
6	96	214	205
7	97	124	120
8	98	267	250
9	99	585	548
10	100	514	480
합계		2933	2758
TPS		0.940334129초/건	

G4B(www.g4b.go.kr)진위확인코드 : ZsyxL7847CU=

시험성적서

한국시험인증원㈜
서울시 강서구 공항대로219 808호
Tel:02-6929-1033 Fax:02-6929-1037

성적서번호 : KST-19-048

KOTCA 한국시험인증원 Korea Testing Certification Agency

1. 의 뢰 자
 - 회사(기관)명 : 광주과학기술원
 - 주 소 : 광주광역시 북구 첨단과기로 123(오룡동)

2. 시험명 : V&V(Verification & Validation) 시험

3. 시험대상품목/물질/시료명
 - 제품명 : 확장 가능한 탈중앙화 보안성 ECCPoW 블록체인 v1.0
 - 제조사 : 광주과학기술원

4. 시험기간 : 2019년 11월 11일 ~ 2019년 11월 13일

5. 시험방법 : 의뢰자 제시기준

6. 시험결과 : '시험성적서 4. 시험결과' 참조

비고 : 1. 이 성적서는 의뢰자가 제출한 시료에 대한 시험결과이며, 한국시험인증원㈜의 사전 동의 없이 본성적서의 전부 혹은 일부를 복사할 수 없습니다.
2. 이 성적서의 시험결과 중 * 표시된 시험결과는 시험기관의 인정범위 밖의 것임을 밝힙니다.
3. 이 성적서의 진위확인은 기업지원플러스(http://www.g4b.go.kr) 웹페이지에서 G4B조회코드로 확인 가능합니다.

확 인	시험자 성 명 : [redacted] (서명)	기술책임자 성 명 : [redacted] (서명)

2019년 11월 21일

한국시험인증원㈜ 대표이사

TPS-0024-1(01)

G4B(www.g4b.go.kr)진위확인코드 : M+SpSIYEHkE=

1. 시험 개요

1.1 시험 기준

- 의뢰자 제시 기준

1.2 시험 장소

- 서울시 강서구 공항대로 219 808호 한국시험인증원 시험실

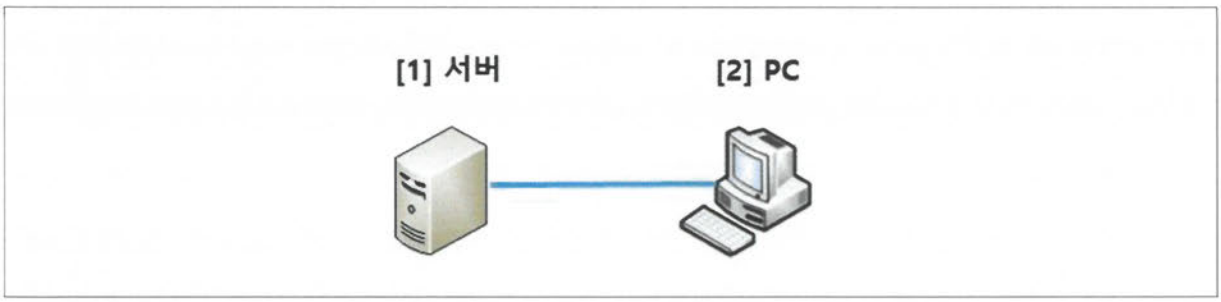
1.3 시험 대상 제품

- 제품명 : 확장 가능한 탈중앙화 보안성 ECCPoW 블록체인 v1.0
- 제조사 : 광주과학기술원

2. 시험 환경

시험환경은 한국시험인증원 시험실에 구축하였으며, 시험환경구성 및 세부사양은 아래와 같다.

<시험환경구성도>



<세부사양>

No	Role	OS	CPU	Memory	HDD/SSD	Pre-Requisite
1	서버	Ubuntu 18.04 LTS	AWS m5.xlarge vCPU 2	8GB	HDD 20GB	· Python
2	PC	Ubuntu 18.04 LTS	Intel(R) Core(R) CPU i5-7600U @ 2.60GHz	16GB	SSD 256GB	· Putty

G4B(www.g4b.go.kr)진위확인코드 : M+SpSIYEHkE=

3. 시험 목적 및 방법(의뢰자 제시)

시험내용(의뢰자 제시)

1. 시험항목명 : ETH-ECC 확장성

2. 시험목적

○ Ethash과 ECCPoW의 평균 TPS 측정

$$\text{Ethash과 ECCPoW의 TPS} = \frac{\text{블록의 보유 Transaction 수}}{\text{블록생성 소요시간(600)}}$$

* TPS : 초당 생성된 Transaction 수

3. 시험방법

1) Ethash과 ECCPoW을 위한 AWS Instance를 4개씩 생성
2) AWS Instance에 Ethash와 ECCPoW를 사용하는 네트워크를 2개씩 배포
3) 각 AWS Instance에서 'chainload -urls http://[노드IP] -id 58141 -tps 7 -v'를 실행하여 블록 생성 시작
3) 10분간 블록을 생성하고 로그에서 Transaction 생성 건수 확인

G4B(www.g4b.go.kr)진위확인코드 : M+SpSIYEHkE=

4. 시험 결과

시험항목	ETH-ECC 확장성
시험목적	Ethash과 ECCPoW의 평균 TPS 측정

시험결과

- Ethash 평균 TPS : 21.681건/초
- ECCPoW 평균 TPS : 22.325건/초

시험회차	구분	블록의 보유 Transaction 수	블록생성시간	TPS
1	Ethash	13,489건	600초	22.481건/초
	ECCPoW	13,299건	600초	22.165건/초
2	Ethash	13,312건	600초	22.186건/초
	ECCPoW	13,168건	600초	21.946건/초
3	Ethash	12,652건	600초	21.086건/초
	ECCPoW	13,764건	600초	22.940건/초
4	Ethash	11,084건	600초	18.473건/초
	ECCPoW	11,051건	600초	18.418건/초
5	Ethash	13,732건	600초	22.886건/초
	ECCPoW	15,776건	600초	26.293건/초
6	Ethash	13,783건	600초	22.971건/초
	ECCPoW	13,314건	600초	22.190건/초
평균	Ethash	13,008.6건	600초	21.681건/초
	ECCPoW	13,395.3건	600초	22.325건/초

G4B(www.g4b.go.kr)진위확인코드 : M+SpSIYEHkE=

시험결과

- Ethash 블록 및 Transaction 생성 정보
  - 1회차 시험

블록 Number	생성된 Transaction수	블록 Number	생성된 Transaction수
22600	155	22626	347
22601	347	22627	328
22602	345	22628	186
22603	347	22629	348
22604	322	22630	347
22605	210	22631	330
22606	334	22632	348
22607	29	22633	329
22608	54	22634	209
22609	28	22635	348
22610	0	22636	348
22611	345	22637	336
22612	340	22638	332
22613	345	22639	327
22614	345	22640	346
22615	70	22641	350
22616	345	22642	341
22617	67	22643	350
22618	342	22644	290
22619	250	22645	38
22620	48	22646	147
22621	347	22647	349
22622	136	22648	334
22623	346	22649	212
22624	79	22650	347
22625	346	-	-

G4B(www.g4b.go.kr)진위확인코드 : M+SpSIYEHkE=

시험결과

• 2회차 시험

블록 Number	생성된 Transaction수	블록 Number	생성된 Transaction수
22657	63	22683	7
22658	251	22684	118
22659	87	22685	1
22660	6	22686	1
22661	3	22687	332
22662	348	22688	330
22663	336	22689	348
22664	7	22690	349
22665	347	22691	339
22666	269	22692	349
22667	347	22693	345
22668	347	22694	348
22669	347	22695	323
22670	347	22696	329
22671	347	22697	204
22672	347	22698	334
22673	347	22699	171
22674	347	22700	18
22675	349	22701	102
22676	348	22702	166
22677	349	22703	330
22678	349	22704	334
22679	348	22705	329
22680	348	22706	329
22681	348	22707	28
22682	349	22708	167

G4B(www.g4b.go.kr)진위확인코드 : M+SpSIYEHkE=

시험결과

• 3회차 시험

블록 Number	생성된 Transaction수	블록 Number	생성된 Transaction수
22716	348	22740	346
22717	342	22741	346
22718	8	22742	228
22719	343	22743	346
22720	162	22744	303
22721	341	22745	37
22722	123	22746	37
22723	69	22747	190
22724	68	22748	102
22725	147	22749	348
22726	150	22750	348
22727	258	22751	348
22728	129	22752	331
22729	347	22753	348
22730	234	22754	348
22731	347	22755	347
22732	347	22756	347
22733	149	22757	347
22734	91	22758	331
22735	177	22759	332
22736	347	22760	335
22737	345	22761	346
22738	346	22762	349
22739	347	22763	347

G4B(www.g4b.go.kr)진위확인코드 : M+SpSIYEHkE=

시험결과

• 4회차 시험

블록 Number	생성된 Transaction수	블록 Number	생성된 Transaction수
22781	346	22802	346
22782	187	22803	346
22783	102	22804	123
22784	26	22805	36
22785	347	22806	348
22786	214	22807	346
22787	174	22808	346
22788	342	22809	288
22789	345	22810	153
22790	344	22811	69
22791	347	22812	195
22792	347	22813	344
22793	187	22814	347
22794	347	22815	346
22795	347	22816	348
22796	347	22817	329
22797	317	22818	347
22798	157	22819	347
22799	59	22820	339
22800	179	22821	334
22801	346	-	-

G4B(www.g4b.go.kr)진위확인코드 : M+SpSlYEHkE=

시험결과

• 5회차 시험

블록 Number	생성된 Transaction수	블록 Number	생성된 Transaction수
22839	350	22863	57
22840	351	22864	341
22841	349	22865	350
22842	244	22866	339
22843	270	22867	350
22844	311	22868	349
22845	310	22869	339
22846	350	22870	349
22847	350	22871	350
22848	244	22872	353
22849	74	22873	351
22850	81	22874	335
22851	233	22875	341
22852	196	22876	340
22853	115	22877	328
22854	175	22878	341
22855	350	22879	340
22856	350	22880	353
22857	351	22881	351
22858	351	22882	343
22859	347	22883	335
22860	142	22884	64
22861	347	22885	180
22862	38	22886	274

G4B(www.g4b.go.kr)진위확인코드 : M+SpSIYEHkE=

시험결과

• 6회차 시험

블록 Number	생성된 Transaction수	블록 Number	생성된 Transaction수
22892	352	22919	341
22893	351	22920	321
22894	351	22921	331
22895	351	22922	144
22896	349	22923	18
22897	347	22924	354
22898	349	22925	351
22899	353	22926	352
22900	352	22927	352
22901	351	22928	352
22902	352	22929	314
22903	352	22930	153
22904	351	22931	126
22905	352	22932	204
22906	263	22933	326
22907	98	22934	310
22908	164	22935	189
22909	32	22936	129
22910	40	22937	337
22911	283	22938	340
22912	273	22939	158
22913	135	22940	83
22914	110	22941	238
22915	282	22942	4
22916	90	22943	48
22917	184	22944	335
22918	353	22945	353

G4B(www.g4b.go.kr)진위확인코드 : M+SpSIYEHkE=

시험결과

- ECCPoW 블록 및 Transaction 생성 정보
  - 1회차 시험

블록 Number	생성된 Transaction수	블록 Number	생성된 Transaction수
18409	347	18432	175
18410	347	18433	87
18411	346	18434	312
18412	347	18435	336
18413	346	18436	349
18414	307	18437	347
18415	172	18438	348
18416	345	18439	316
18417	347	18440	77
18418	347	18441	292
18419	348	18442	330
18420	348	18443	234
18421	345	18444	349
18422	347	18445	332
18423	348	18446	238
18424	349	18447	292
18425	336	18448	270
18426	174	18449	335
18427	211	18450	340
18428	97	18451	267
18429	17	18452	167
18430	350	18453	289
18431	348	18454	348

G4B(www.g4b.go.kr)진위확인코드 : M+SpSlYEHkE=

시험결과

• 2회차 시험

블록 Number	생성된 Transaction수	블록 Number	생성된 Transaction수
18467	91	18493	347
18468	347	18494	348
18469	345	18495	254
18470	188	18496	262
18471	346	18497	49
18472	220	18498	0
18473	172	18499	99
18474	78	18500	18
18475	347	18501	68
18476	309	18502	26
18477	83	18503	57
18478	165	18504	346
18479	40	18505	347
18480	23	18506	330
18481	347	18507	347
18482	345	18508	346
18483	348	18509	330
18484	347	18510	333
18485	346	18511	346
18486	348	18512	167
18487	346	18513	337
18488	346	18514	346
18489	349	18515	181
18490	348	18516	292
18491	347	18517	342
18492	348	18518	331

G4B(www.g4b.go.kr)진위확인코드 : M+SpSIYEHkE=

시험결과

• 3회차 시험

블록 Number	생성된 Transaction수	블록 Number	생성된 Transaction수
18525	76	18549	344
18526	348	18550	67
18527	347	18551	306
18528	348	18552	337
18529	344	18553	328
18530	346	18554	207
18531	348	18555	350
18532	346	18556	349
18533	347	18557	349
18534	347	18558	350
18535	349	18559	317
18536	350	18560	349
18537	348	18561	312
18538	348	18562	335
18539	347	18563	349
18540	349	18564	317
18541	349	18565	351
18542	349	18566	345
18543	348	18567	244
18544	344	18568	314
18545	162	18569	203
18546	30	18570	60
18547	13	18571	28
18548	31	18572	339

G4B(www.g4b.go.kr)진위확인코드 : M+SpSIYEHkE=

시험결과

• 4회차 시험

블록 Number	생성된 Transaction수	블록 Number	생성된 Transaction수
18584	347	18603	348
18585	81	18604	256
18586	146	18605	161
18587	57	18606	347
18588	111	18607	349
18589	312	18608	346
18590	288	18609	347
18591	178	18610	348
18592	346	18611	349
18593	347	18612	349
18594	341	18613	322
18595	208	18614	349
18596	348	18615	316
18597	347	18616	349
18598	275	18617	351
18599	19	18618	351
18600	347	18619	349
18601	345	18620	324
18602	346	18621	351

G4B(www.g4b.go.kr)진위확인코드 : M+SpSIYEHkE=

시험결과

• 5회차 시험

블록 Number	생성된 Transaction수	블록 Number	생성된 Transaction수
18639	40	18666	2
18640	242	18667	13
18641	58	18668	0
18642	353	18669	714
18643	345	18670	415
18644	161	18671	415
18645	352	18672	691
18646	349	18673	676
18647	350	18674	689
18648	352	18675	617
18649	352	18676	459
18650	354	18677	246
18651	353	18678	125
18652	260	18679	425
18653	73	18680	350
18654	2	18681	351
18655	292	18682	349
18656	108	18683	222
18657	259	18684	223
18658	110	18685	13
18659	352	18686	325
18660	278	18687	352
18661	98	18688	333
18662	245	18689	353
18663	298	18690	323
18664	346	18691	324
18665	69	18692	320

G4B(www.g4b.go.kr)진위확인코드 : M+SpSIYEHkE=

시험결과

• 6회차 시험

블록 Number	생성된 Transaction수	블록 Number	생성된 Transaction수
18698	353	18723	352
18699	350	18724	352
18700	348	18725	350
18701	351	18726	341
18702	154	18727	244
18703	19	18728	202
18704	75	18729	86
18705	8	18730	355
18706	257	18731	353
18707	13	18732	354
18708	325	18733	355
18709	352	18734	351
18710	349	18735	343
18711	349	18736	170
18712	347	18737	209
18713	132	18738	328
18714	97	18739	223
18715	349	18740	0
18716	350	18741	219
18717	352	18742	226
18718	353	18743	333
18719	352	18744	203
18720	352	18745	318
18721	353	18746	354
18722	353	-	-

G4B(www.g4b.go.kr)진위확인코드 : M+SpSlYEHkE=

Time-Variant Proof-of-Work Using Error-Correction Codes

Sangjun Park, Haeung Choi, and *Heung-No Lee, *Senior Member, IEEE*

***Abstract*— The protocol for cryptocurrencies can be divided into three parts, namely consensus, wallet, and networking overlay. The aim of the consensus part is to bring trustless rational peer-to-peer nodes to an agreement to the current status of the blockchain. The status must be updated through valid transactions. A proof-of-work (PoW) based consensus mechanism has been proven to be secure and robust owing to its simple rule and has served as a firm foundation for cryptocurrencies such as Bitcoin and Ethereum. Specialized mining devices have emerged, as rational miners aim to maximize profit, and caused two problems: *i*) the re-centralization of a mining market and *ii*) the huge energy spending in mining. In this paper, we aim to propose a new PoW called Error-Correction Codes PoW (ECCPoW) where the error-correction codes and their decoder can be utilized for PoW. In ECCPoW, puzzles can be intentionally generated to vary from block to block, leading to a time-variant puzzle generation mechanism. This mechanism is useful in repressing the emergence of the specialized mining devices. It can serve as a solution to the two problems of recentralization and energy spending.**

***Index Terms*— Consensus, Cryptocurrency, Blockchain, Proof-of-Work, Error-Correction Codes, Hash Functions**

I. Introduction

In cryptocurrencies, the consensus part plays a role in leading an agreement among trustless nodes without any communications. This part is the most innovative because it can prevent the double spending attack [1] in a peer-to-peer network in the absence of trusted parties. In *Bitcoin* [2], as an example, more than ten thousand of nodes randomly scattered across the world aim to reach a consensus in each block time. The Internet is the only way to connect them; communication packets are delayed and sometimes dropped though the Internet that is designed to provide the best effort service. Cyberattacks frequently happen, making transactions over the Internet insecure. Nevertheless, Bitcoin has shown secure peer-to-peer transactions over the past 10 years. With the help of proof-of-work (*PoW*) which is fundamental to the consensus part, this can be possible.

In Bitcoin, each node does competitive work, called *mining*, to forge a block. The node which wins this competition has the right to mint a specified number of coins as this mining reward. If a node was re-forging all the blocks alone, it could spend the total amount of works done to all the mined blocks.

Without PoW, anybody with a computer can alter the content of the blockchain, implying unauthorized changes in any mined blocks can be possible. If PoW is attached to each mined block, attackers cannot make any unauthorized modifications without redoing all the works. No node can alone alter any mined block, meaning an immutability property.

In Bitcoin, miners make rational decisions to maximize their profits by following a two stage process in which *i*) the miners select a blockchain whose length is the longest and *ii*) they extend this longest one by adding a newly mined block. Suppose there are two blockchains where one is longer than the other one in terms of the length. Since the longer chain has the more accumulated works, altering it is more difficult. This longer chain shall be treated the more trustable and preferable by the miners. Thus, they select the longer chain. Making such a selection is rational for the sake of keeping the mining rewards. The mining reward is a delayed conditional payment, i.e., if a miner mines a block at a given time point t_1, the reward is delayed until the future moment t_2 of time. This time from t_1 to t_2 is measured in terms of the number of blocks, say 100 blocks. If this mined block was not a part of the longest chain at the future time point t_2, the reward vanishes. Thus, rational miners select the longest chain.

In Bitcoin, miners spend computational resources to forge a block by solving a puzzle carved in a bitcoin program as an on-chain policy. This puzzle is made using the secure hash algorithm 256 (SHA256) [3]. To solve the puzzle, the miners have to repeatedly execute SHA256 by varying an input to SHA256 until a good hash is given. This input is the header of the block, i.e., block header, including six fields such as version, previous hash, difficulty, timestamp and nonce, Merkle tree value. The version is fixed. Given a block at a certain height, i.e., the lth block, the previous hash and the difficulty are obtained from its previous block, i.e., $(l–1)^{\text{th}}$ block. They are constant. The rest varies until SHA256 returns a good hash. A good hash can be spotted since it shall possess a certain number of leading zero bits reflecting the difficulty. The block header of a mined block can serve as the proof that a given puzzle is solved without any falsehood.

Satoshi [2] intended for miners to execute SHA256 using a central processing unit (CPU). But, faster computing machines based on application-specific integrated circuit (ASIC) became available. As a result, the miners have chosen to exploit them to maximize their profits. To date, the miners equipped with ASIC

S. Park is with the Electronics and Telecommunications Research Institute (ETRI), South Korea (e-mail: sjpark86@etri.re.kr)

H. Choi and H.-N. Lee are with the School of Electrical Engineering and Computer Science (EECS), Gwangju Institute of Science and Technology (GIST), South Korea (e-mail: haeung@gist.ac.kr, heungno@gist.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean government (MSIP) (NRF-2018R1A2A1A19018665). The corresponding authors: Heung-No Lee (heungno@gist.ac.kr)

mining devices have dominated the mining business [4], leading to two problems:

M1. The mining markets have become re-centralized [5].
M2. The electrical energy spent to mine blocks is huge [6].

First, the miners have a large portion of the total hash power, implying that the plight of the blockchain is left to a handful of these influential miners. It can be possible to modify any mined blocks on their own rights, leading to shattered trust. Namely, they can break the immutability property. Second, new models of ASIC mining devices can surpass old models with respect to the hash power, which is measured as a hash rate. Each miner is forced to buy newer models to win the mining competition. As the new models are widely used, the total hash power inevitably grows. The difficulty level to solve the puzzle must increase to keep a certain predefined range of block generation time that is expectedly consumed to mine a single block. Today, this level has gotten huge, i.e., $O(10^{20})$ hash rate per second. As a result, using CPUs in mining has gotten no longer profitable. Besides, as the total hash power increases, the miners spend more and more electrical energy to mine blocks.

If we prevent the usage of ASIC mining, we can alleviate the problems M1 and M2. To this end, we use the error-correction codes and their decoder. In general, the aim of using the codes in modern communication systems is to combat errors occurring over noisy channels in which the errors introduced over a noisy channel can be corrected by running a decoding algorithm. The codes have a rich history where there are numerous classes of good codes available. They have been used to define both a good public-key crypto system and a good hash function.

The first result can be traced back to the late 1970. McEliece [8] used Goppa codes to make a McEliece cryptosystem where a message is encoded using a public key $\mathbf{A} := \mathbf{SGP}$ where $\mathbf{G}$ is the generator matrix of a binary Goppa code, $\mathbf{S}$ is a nonsingular random matrix and $\mathbf{P}$ is a permutation matrix. The hash, i.e., the encoded result, of a provided message is made as follows: *a*) a word is made by multiplying the message with the public key and *b*) adding a binary random word whose number of ones is at most t to this word[1] is to get the hash. Peters *et al.* [9] extended this system using non-binary Goppa codes to reduce the size of its public key. Even the size reduces, this system can achieve still the same security level as much as that of [8] could. Other codes such as low density generator matrix codes [10], low density parity check (LDPC) codes [11], [12], Reed-Solomon codes [13] and Reed-Muller codes [14] have been used to replace the Goppa codes. The aim of using them is to reduce the size of the public key.

Aside from the applications of the error-correction codes into the McEliece cryptosystem, the codes are used to construct new hash functions. Preneel [15] proposed a method to make new hash functions and proved that their hash functions can provide strong collision resistant properties. The codes in [15] are either the maximum distance separable codes or the Hamming codes. Selman *et al.* [16] used LDPC codes to make a hash function and proved this function as an average universal hash function defined in [17]. These results can motivate us to exploit error-correction codes in designing a new PoW framework.

The contributions of this paper are three folds. First, we propose a new PoW framework which we name as ECCPoW. As the name implies, we add the error-correction code part into PoW. To the best of our knowledge, this is the first work in which the error-correction codes are applied to blockchains. We then explain how we make puzzles, which we call ECC puzzles, and give routines to solve them.

Second, we conduct a probabilistic study to examine a random variable called First Success Hash Cycle (FSHC) representing the number of hash cycles required in solving a given ECC puzzle. Based on this study, we get the following results:

- FSHC follows a geometric distribution with a parameter in (21) that depends on the number of miners M and the code length n.
- The expected value of FSHC is a decreasing function with respect to the number of miners M.
- The expected value of FSHC is an increasing function with respect to the code length n.

Third, we define five properties for a good PoW and explain how ECCPoW satisfies these properties. We shall note that the most innovative property is the time-variant property, making ECCPoW suitable to resolve the problems of recentralization and energy spending.

We organize the rest of this paper as follows. Section II gives literature surveys regarding SHAs and PoWs. Section III elucidates LDPC codes and a decoder. Section IV addresses how ECCPoW works and gives its pseudo codes Section V presents theoretical results of ECCPoW. Section VI discusses properties of ECCPoW. Section VII presents the conclusions of this paper.

II. Literature On Blockchain Cryptographies

A. Secure hash standard and functions

The secure hash standard was formulated by NIST [3]. The purpose of this standard is to offer the specifications of SHAs that yield a hash of a given message. Even the message changes slightly, the hash of the changed message comes out completely different from that of the original message. Thus, a hash can be used to detect whether an original message was altered or not. SHAs with such a property can be used for the generation and the verification of digital signatures as well as for the message authentication.

A secure hash function takes an arbitrarily sized message and produces a fixed-size hash. Let a function h be

$$h : \mathcal{X} \to \mathcal{Y}$$

which is said to be a cryptographically secure hash function if it satisfies the three requirements defined in [17] below:

(***One-way function***) Given any hash y to which a corresponding message is not known, it is computationally infeasible to find a message x such that $h(x) = y$.

[1] For any Goppa code, there is a construction method to guarantee that the minimum distance d of that code is greater than a given positive integer. Thus, the value of t can be known in advance using Theorem 1 [32].

TABLE I. The routine of bitcoin

Inputs: $\mathcal{S}_0$ (block header except for nonce) and L	
Step 1:	for nonce = 0, 1, 2, … $2^{32} - 1$
Step 2:	$\mathbf{e} = \mathrm{SHA256}(\mathcal{S}_0 \cup \{\mathrm{nonce}\})$
Step 3:	If **e** begins with L zero bits, then go to Step 5.
Step 4:	end
Step 5:	Block generation & broadcast

(***Weak collision resistance***) Given an arbitrary message x, it is computationally infeasible to find any message x' which results in the same hash, i.e., $h(x) = h(x')$.
(***Strong collision resistance***) It is computationally infeasible to find any two different messages x and x' which make the same hashes $h(x) = h(x')$.

NIST [3] has proposed a family of SHAs including SHA1, SHA224, SHA256, SHA384, and SHA512. A message of any size less than 2^{64} bits can be given as an input for SHA1, SHA224, and SHA256, while that of less than 2^{128} bits for SHA384 and SHA512. The size of a hash ranges from 160 to 512 bits, depending on the algorithm.

B. PoW of Bitcoin

In Table I, we define a puzzle in Bitcoin and give routines to solve this puzzle. In Step 2, a miner puts a given block header with a selected nonce to SHA256 and obtain a hash of 256-bits. In Step 3, this puzzle is declared to be solved if a hash is smaller than a specified target value, or in a simpler argument it begins with L zero bits, where the target value or the value of L is given as the difficulty level of the PoW puzzle. The miner repeats the routines from Step 2 to Step 3 by varying the nonce. However, there exists a chance in which the miner can fail to hit a good hash even though the whole range of nonces, i.e., $0 \sim 2^{32} - 1$, is used. In such a case, the miner updates the block header, and repeats the routines from Step 2 and Step 3. There are two methods to update the block header. The first one is to update the timestamp field. The second one is to update the Merkle tree value by modifying the list of transactions being included in a block which the miner aims to construct.

C. PoW of Ethereum

Ethash [18] was created for the purpose of preventing the advent of ASIC mining devices in Ethereum. In Ethash, there is a memory structure called directed acyclic graph (DAG) where its data are randomly re-generated every 30,000 block.

Table II shows routines of Ethash. As we have shown in Step 2, the current block header with a nonce is taken by SHA3 to get a hash. This hash is taken by a predefined function to yield mix0 that is random. In Step 4, mix0 is used to determine which data from DAG are fetched. No one predicts which data shall be fetched from DAG because mix0 is random. The mixer takes both the fetched data and mix0 to get a random value in Step 5. In Step 6, mix0 is updated using the random value. The routines from Step 2 to Step 6 are repeated 63 times. Last, the decision is made using this final mix0, as we have shown in Step 8.

The ASIC resistant property in Ethash is originated from the fact that the operation time for the mixer is shorter than that of

TABLE II. The routine of Ethash

Inputs: $\mathcal{S}_0$ (block header except for nonce), L and DAG	
Step 1:	for nonce = 0, 1, 2, … $2^{32} - 1$
Step 2:	$\mathrm{mix0} = f\left(\mathrm{SHA3}(\mathcal{S}_0 \cup \{\mathrm{nonce}\})\right)$
Step 3:	for i = 1, 2, …, 63
Step 4:	$\mathrm{data1} = \mathrm{Fetch}(\mathrm{DAG}, \mathrm{mix0})$
Step 5:	$\mathrm{tmp} = \mathrm{Mixing}(\mathrm{mix0}, \mathrm{data1})$
Step 6:	$\mathrm{mix0} = f(\mathrm{tmp})$
Step 7:	end
Step 8:	If mix0 begins with L zero bits, then go to Step 10.
Step 9:	end
Step 10:	Block generation & broadcast

where f is a predefined function. Details on f is given in [18].

the fetch operation. To be specific, let A_i be the time duration (TD) to conduct the mixing operation in which the subscript i denotes a chip to run this mixing operation. It is clear that the mixing operation TD based on an ASIC chip is significantly less than that based on any CPU chip because the clock speed of an ASIC chip is much faster than that of any CPU chip, i.e.,

$$\mathrm{A}_{\mathrm{ASIC}} \ll \mathrm{A}_{\mathrm{CPU}}.$$

Next, let B_i be the TD to conduct the fetch operation. Unlike to the mixing operation TD, this TD depends on the communications bandwidth between the memory and the cache of the CPU in which the fetched data from DAG are passed through. In other words, the fetch operation time is connected mainly to the bandwidth but not to the clock speed. The purpose to use ASIC chips is to increase the processing speed; is not for obtaining a higher communications bandwidth. The fetch operation TD for CPU and ASIC are thus similar, i.e.,

$$\mathrm{B}_{\mathrm{ASIC}} \approx \mathrm{B}_{\mathrm{CPU}}.$$

We consider the inner routines of Ethash, i.e., Step 4 and Step 5. The mixing operation is conducted after the fetching operation is done. This operation TD at Step 4 can be significantly reduced using ASIC chips; but the fetching operation TD at Step 5 is not reduced even if ASCI chips are used. As a result, there is a bottleneck between Step 4 and Step 5. This bottleneck has the ASIC resistant property enabled.

Recently, a programmatic PoW (ProgPoW), which is planning to be used to replace Ethash, has been proposed to further improve the ASIC resistant property. This improvement is done by changing parameters related to DAG randomly from block to block. Such modifications can make the fetch operation time increased. But, the development of ProgPoW is not completed and ProgPoW is not proven to be secure at the time of writing this manuscript.

D. PoW of Dash

X11 was proposed in 2014 by Duffield [19]. In Table III, we give routines of X11 which consists of hash functions below:

Blake, Bmw, Groestl, Jh, Keccak, Skein, Luffa, Cubehash, Shavite, Simd and Echo.

TABLE III. The routine of X11

Inputs: $\mathcal{S}_0$ (block header except for nonce) and L	
Step 1:	for nonce = 0, 1, 2, … $2^{32} - 1$
Step 2:	$\mathbf{e} = \mathtt{Blake}(\mathcal{S}_0 \cup \{\text{nonce}\})$
Step 3:	$\mathbf{e} = \mathtt{Bmw}(\mathbf{e})$
	….
Step 12:	$\mathbf{e} = \text{Echo}(\mathbf{e})$
Step 13:	If **e** begins with L zero bits, then go to Step 15.
Step 14:	end
Step 15:	Block generation & broadcast

Blake first takes a given set of the current the block header with a selected nonce to get its hash. Next, Bmw takes this hash as its input to yield a hash. The same procedures are repeated until Echo, the last hash function, yields its hash. The decision is made using this last hash, as we have shown in Step 13.

However, the order of using the 11 hash functions is always fixed. This fixed order makes the development of ASIC mining devices possible and in fact an easy task. The development of ASIC mining solution can be done when these hash functions are implemented in a single device. Logic gates to sequentially connect the hash functions can be implemented. The first ASIC mining device targeting X11 was developed in 2016.

The idea behind X11 has been extended to other PoWs such as X13, X15, and X17. As the names suggest, they consists of 13, 15, and 17 hash functions, respectively. To date, there a set of ASIC mining devices for both X13 and X15 while there is no ASIC mining device yet for X17.

E. PoW of Raven

In 2018, a new extension of X11 was proposed in [20]. This is called X16r. It uses multiple hash functions given in Table IV to get the last hash like to other extensions of X11 that we have discussed in Section II D. But, unlike the others, the sequence of the hash functions in X16r can be made to vary from block to block. This variation seems to be a role for preventing the advent of ASIC mining devices for X16r.

We provide an example to address how X16r operates. The sequence X16r is determined upon the last 16 bytes of a previous hash. Let this previous hash be

0x0000...04def2c3eff6da11542ffcdabce.

The last 16 bytes are 6da11542ffcdabce. Then, the sequence is decided on the basis of Table IV below:

Luffa → Shabal → Echo → Bmw → Bmw → Skein → Keccak → Groestl → Sha512 → Sha512 → Fugue → Shabal → Echo → Hamsi → Fugue → Whirlpool.

A miner puts a given set of the block header with a selected nonce through Luffa to yield a hash. Shabal takes this hash as its input to yield its hash. This routine is repeated until the last hash is yielded. In the above example, the last hash is obtained through Whirlpool.

At the time of writing this manuscript, it seems, no one has officially succeeded in implementing ASIC mining devices for X16r, i.e., there is no announced commercial product. However,

TABLE IV. The map for X16r

Value	Hash	Value	Hash
0	Blake	8	Shavite
1	Bmw	9	Simd
2	Groestl	a	Echo
3	Jh	b	Hamsi
4	Keccak	c	Fugue
5	Skein	d	Shabal
6	Luffa	e	Whirlpool
7	Cubehash	f	Sha512

Black and Weight, the developers of X16r, in [20] stated that reordering the sequence cannot make the development of ASIC mining devices impossible. Recently, in [33] at Nov., 2019, Whitefire990 reported a simulation result which indicates the probability of k time-repetition, such that the same hash function is used at least 5 times consecutively when k is 5, exponentially decreases in k. Insisted further is that k greater than 5 can be ignored in designing of ASIC mining devices. As such, what claimed there is that the ASIC-resistant property of Raven can be broken by the said ASIC designing method. It shall be noted, however, that all these claims provided in [33] have not yet been carefully verified through a peer-review system.

F. Short summary from C to E

From the subsection II.C to II.E, we have reviewed the existing ASIC-resistant PoWs categorized as follows:

a. The usage of intentional memory access.
b. The usage of multiple hash functions.

Ethash and ProgPoW can belong to the first class while X11 and its variants such as X13, X15, X17 and X16r can belong to the second class. The basic idea of the first class is to use the bottleneck intentionally caused by randomly fetching data from a memory. The basic idea of the second one is to use the multiple hash functions, which can make the development costs of ASIC mining devices expensive.

At the time of writing this manuscript, ASIC mining devices for Ethash, X11, X13 and X15 are available. The development of ProgPoW is not yet available. X17 and X16r can be resistant to ASIC mining devices. But, as the ASIC-resistant property of the PoWs such as X11, X13 and X15 are broken, that of X17 can be cracked in the near future when the hardware development technology is improved. As mentioned in the subsection II.E, there is a claim that the anti-ASIC property of X16r could be broken; but to date no commercial ASIC mining device has been announced.

III. Literature Surveys on LDPC

An LDPC decoder plays an important role in ECCPoW. That is, the decoder is utilized to generate an unpredictable random output that can be later on used to give a proof whether a puzzle is solved or not. We prepare this section to give a quick summary regarding the LDPC codes and their decoders.

In 1963, LDPC codes were proposed by Gallager in his thesis [7]. But, the codes received no attention back then because computers were not sufficiently fast enough to check the per-

formance of a decoder. Mackay and Neal [22] reported in 1997 that the codes could achieve the Shannon limit [21] closely with a message passing decoder that uses a kind of belief propagation algorithms. Since then, numerous studies on the codes have been made. They are categorized as follows: *i*) constructing the codes to approach the Shannon limit [21] and *ii*) implementing the fast decoders based on either ASIC [23]–[26] or field programmable gate array (FPGA) [27]–[28] to support real-time decoding purpose.

A. LDPC codes

LDPC codes can be generalized to a non-binary alphabet for improving its error-correction capability. But, for the purpose of using these codes in ECCPoW to provide a new time-varying anti-ASIC PoW system, it hence is suffice to consider the binary alphabet version only.

An (n, k) LDPC code is a linear code constructed by supplementing each message $\mathbf{m}$ of size k with parity bits to get a codeword of size n. This code is often defined with respect to a parity check matrix $\mathbf{H}$ of size $m \times n$ such that each element is binary either zero or one and the number of ones is very small, where m is the number of parity bits, i.e., $m = n - k$.

For a given parity check matrix $\mathbf{H}$, its corresponding LDPC code can be either *regular* or *irregular*. If $\mathbf{H}$ contains a constant number w_c, called the column degree, of ones in each column and a constant number w_r, called the row degree, of 1s in each row, the code is called *regular*. For a given *regular* LDPC code, the parameters such as n, k, w_c, and w_r satisfy the following:

$$nw_c = (n-k)w_r = mw_r. \tag{1}$$

If $\mathbf{H}$ contains a different number of 1s in both each column and each row, the code is called *irregular*. In the perspective of the error-correction capability, irregular codes are better than regular codes. To serve our purpose of anti-ASIC PoW mechanism, we aim to consider the *regular* LDPC codes because

i. it is much easier to implement a decoder of regular LDPC codes and
ii. the aim of using this decoder is not to correct errors but to yield an unpredictable random output.

A bipartite graph is often used to represent an LDPC code, as we have shown in Fig. 1. The lower and upper nodes are called the variable nodes and the check nodes, respectively. Each edge shows the adjacency of the ith variable node and the jth check node and corresponds to a nonzero $(i, j)^{\text{th}}$ element in $\mathbf{H}$.

For a given LDPC code, its error-correction capability relies on the minimum (Hamming) distance d. This distance is given by solving an optimization problem in which we consider any pair of $2^k - 1$ different codewords below:

$$d = \min_{\mathbf{u} \in \{\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_{2^k}\} \setminus \mathbf{0}_n} \|\mathbf{u}\|_h \tag{2}$$

which is NP-complete, where $\|\mathbf{x}\|_h$ is the number of 1s in $\mathbf{x}$.

Thus far, studies on the computation of a good approximation to minimum distance for a given fixed $\mathbf{H}$ with reasonable size have been reported futile and as such it has remained as an open problem. Keha and Duman [29] proposed a branch and cut algorithm to obtain the minimum distance of LDPC codes. But, this algorithm requires a large amount of time and memory; it is thus only useful if n is small. Then, Hashemi and Banihashemi [30] proposed a method to find lower and upper bounds of the minimum distance of LDPC codes and obtained both of the bounds even when $n > 64{,}000$.

For regular LDPC codes with a particular pair of w_c and w_r, upper and lower bounds for a relative minimum distance which is the ratio of the minimum distance d to the code length n, are given in [31] and [7], respectively. We use them for our purpose in this paper in Section IV. Once the minimum distance is given, the number of correctable errors can be computed as follows:

Theorem 1 [32]: Let a linear code be defined as a given parity check matrix $\mathbf{H}$ which has the minimum distance d. Then, the number of correctable errors is

$$t = \lfloor (d-1)/2 \rfloor \tag{3}$$

where $\lfloor x \rfloor$ denotes the integer part of x.

We explain how to encode a message $\mathbf{m}$ for a given $\mathbf{H}$ of size $m \times n$. To this end, we build a generator matrix $\mathbf{G}$ of size $n \times k$ whose column space is orthogonal to the row space of $\mathbf{H}$ below:

Step 1: Conduct the Gaussian elimination to rewrite $\mathbf{H}$ as follows:

$$\mathbf{H} = \begin{bmatrix} \mathbf{A}^T & \mathbf{I}_{n-k} \end{bmatrix}$$

where $\mathbf{I}_{n-k}$ is the identity matrix of size $(n-k) \times (n-k)$.
Step 2: Form $\mathbf{G}$ of size $n \times k$ as follows:

$$\mathbf{G} = \begin{bmatrix} \mathbf{I}_k \\ \mathbf{A} \end{bmatrix}.$$

It is noted that $\mathbf{HG} = \mathbf{0}_{n-k,k}$ where $\mathbf{0}_{n-k,k}$ is the zero matrix of size $(n-k) \times k$. Note again $m = n - k$. The message $\mathbf{m}$ is encoded to produce a codeword $\mathbf{c}$ of size $n \times 1$ via $\mathbf{c} = \mathbf{Gm}$. Then, because of the definition of $\mathbf{G}$, it is always seen that the result of multiplying $\mathbf{H}$ with $\mathbf{c}$ is the zero vector of size m, i.e.,

$$\mathbf{Hc} = \mathbf{HGm} = \mathbf{0}_m.$$

A decoder takes both the parity check matrix $\mathbf{H}$ and the corrupted word $\mathbf{r}$, which is $\mathbf{r} = \mathbf{c} + \mathbf{e}$, where $\mathbf{e}$ is an error pattern. The decoder runs a message-passing algorithm [32] shown to be the standard decoding algorithm to remove $\mathbf{e}$.

The principle behind the message-passing algorithm is to iteratively propagate probabilistic information among the variable and check nodes. The iterations are terminated if either the number of iterations exceeds a given number or an output is a codeword. Detailed explanations on how this algorithm operates are provided in [32], i.e., Algorithm 5.1 on page 220. The algorithm takes parameters such as $\mathbf{H}$, $\mathbf{r}$, maxIter, and ε, where maxIter is the number of maximum iterations, and ε is the cross error probability that is used to determine the initial value of the algorithm.

The error-correction performance of the algorithm depends on both maxIter and the crossover error probability. If maxIter is small, the algorithm fails to obtain a converged solution. If it is large, the algorithm may take a considerably long computational time to obtain its solution. In the literature, maxIter is set from 10 to 20 in general. Next, the crossover error probability is set if the transition probability of a binary symmetric channel is either given or estimated. If this is improperly set, the decoding performance is degraded, leading to the poor error-correction capability. For the purpose in correcting errors, this parameter must be carefully considered.

The aim of using the decoder in ECCPoW, please make note of the fact that, is not to correct errors. Thus, there is no need to set maxIter and ε strictly. One condition that we shall aim to satisfy is that all the miners in ECCPoW system have to use the same values for these parameters. This condition can be easily satisfied by letting them to be published fixed constants in the proposed implemented program. As the miners verify a newly published block before accepting it, there is no benefit not to follow and use different ones for these published parameters. That is, any proof obtained from arbitrary parameters other than the published ones must be rejected.

We will give details on how to construct the other parameters, such as **H** and **r**, in Section IV.

B. FPGA and ASIC Implementation

LDPC decoders based on ASIC, a.k.a. ASIC-LDPC decoders are implemented to achieve low power consumption and fast processing. In the decoders, the check and variable nodes have to be physically connected using logical gates for a given parity check matrix. Fixed connections are to imply limited flexibility on the designs on the ASIC-LDPC decoders, making them only to support either a set of pre-defined parity check matrices or structured parity check matrices. We give our surveys related to existing ASIC-LDPC decoders as follows.

First, the ASIC-LDPC decoder in [23] supports quasi-cyclic parity check matrices decomposed into cyclic-shifted identity and zero matrices. These matrices have the same structure that is used in implementing the decoder. Second, the ASIC-LDPC decoder in [24] supports the parity check matrices included in the IEEE 802.16e system. These matrices are fixed; not change. Third, in [25], Hanzo *et al.*, reviewed the state-of-the-art of ASIC-LDPC decoders and stated that these decoders must take a bank of hardware to support many random parity check matrices. Namely, additional components such as memories, controllers and switchable interconnections are required, resulting in that these components occupy the most chip area in the decoders. They supported their statement by providing an example of [26] in which the ASIC-LDPC decoder supports about 100 parity check matrices, but its additional components occupy 75% of the total area of the decoder. This shows that there are no practical implementations on ASIC-LDPC decoders to support an infinite number of random parity check matrices.

There are FPGA-LDPC decoders that are the LDPC decoders implemented on FPGA chips. FPGA-LDPC decoders consume more power rather than ASIC-LPDC decoders do. But, it is much easier to reprogram the FPGA-LDPC decoders, which implies that they can achieve the more flexibility on the designs compared to ASIC-LDCP decoders. The FPGA-LDPC decoder in [27] supports parity check matrices up to n = 65,000. But, it is required to load a parity check matrix onto this decoder when it has to be changed, requiring additional time. In [28], Hanzo et al., stated that FPGA-LDPC decoders require additional routing and processing devices to support many parity check matrices. But, as they pointed, the use of these additional devices can lead to complex designs, increasing the cost of the decoders.

IV. Error-Correction Codes Proof of Work

In this section, we give details on ECCPoW. For simplicity, we organize this section into four subsections. In the first subsection, we list fields belonging to a block header of ECCPoW and provide their simple explanations. In the second subsection, we illustrate an overall structure of ECCPoW and its explanations. In the third subsection, we explain how we construct two inputs that appear as we use a decoder of error-correction codes. In the last subsection, we give the definition of an ECC puzzle generation function and present how to define an ECC puzzle using this ECC function. We end the last subsection by giving routines for solving this ECC puzzle.

A. Block header in ECCPoW

The block header of ECCPoW is defined to be a data structure that has eight fields such as timestamp, Merkle tree value, previous hash value, nonce, version, code length, row degree and column degree.

We use the fields such as version, timestamp, previous hash value, Merkle tree value and nonce to achieve in the purpose of guaranteeing the immutability property similar to Bitcoin.

We use the remains such as the code length, the row degree and the column degree to assign the size of hash vector and parity check matrix (PCM), which appears due to the usage of a decoder of a family of LDPC codes. As we will show in Section V, we change the difficulty level of a puzzle by varying the code length if the column degree and the row degree are fixed.

B. Overall structure of ECCPoW

Fig. 1 is prepared to present an overall structure of ECCPoW. This structure consists of three parts such as *i*) the hash vector generation (HVG) part, *ii*) the LDPC decoder part and *iii*) the decision part. We explain each part as follows.

In the HVG part, we randomly generate a hash vector of size n using a series of SHA256s taking the CBH with a given nonce generated by the nonce generator. The details how to generate this hash vector from the knowledge of the CBH will be given in the subsection IV.C.

In the LDPC decoder part, there is a decoder from a family of LDPC codes. This decoder takes the above hash vector and runs the message-passing algorithm [32] to yield a binary word **c**. It is noted that this decoder takes a parity check matrix (PCM) **H** determining the relation between an input and its corresponding output. The details how to construct this PCM will be given in the subsection IV.C.

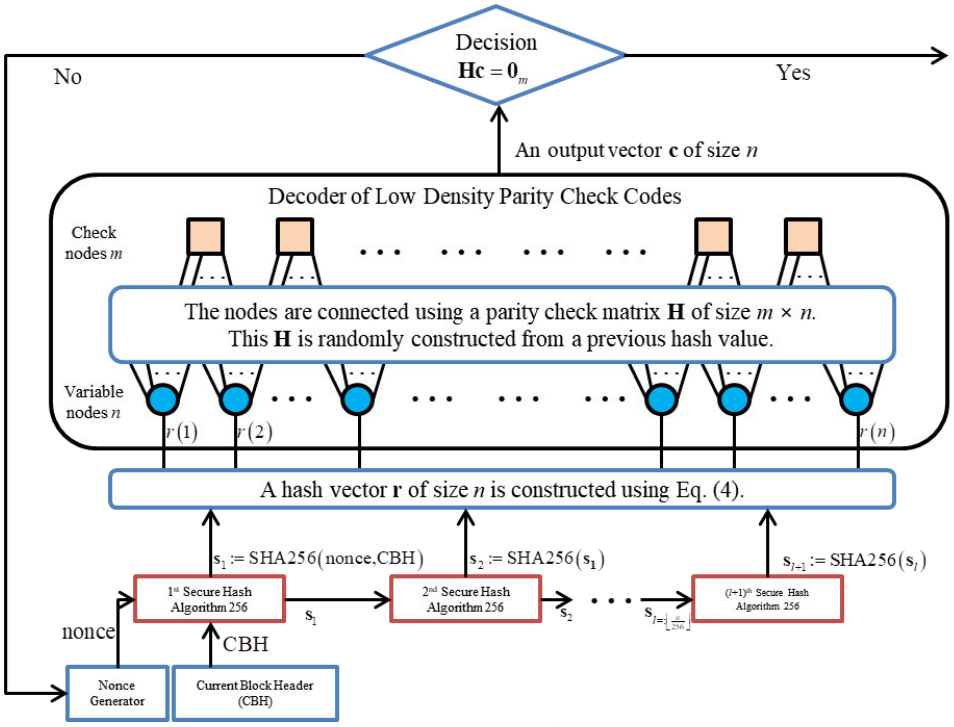


Fig. 1. An overall scheme of ECCPoW.

Last, in the decision part, the decision is made on the basis of the output provided by the decoder, as we have shown in Fig. 1.

C. Construction of hash vector and parity check matrix in ECCPoW

First, we provide the definition of hash vector **r** and give how we construct this hash vector using the current block header.

Definition 1 – *Hash Vector*: A hash vector **r**, which is a vector of concatenating outputs of SHA256s, of size n is defined to be:

$$\mathbf{r} := \begin{cases} \mathbf{s}_1[1:n] & \text{if } n \le 256 \\ \left[\mathbf{s}_1 \quad \cdots \quad \mathbf{s}_l \quad \mathbf{s}_{l+1}[1:j]\right] & \text{if } n > 256 \end{cases} \tag{4}$$

where $l = \lfloor n/256 \rfloor$, $j = n - 256 \times l$,

$$\mathbf{s}_1 := \mathrm{SHA256}(\mathrm{CBH}) \in \{0,1\}^{256} \tag{5}$$

and

$$\mathbf{s}_u := \mathrm{SHA256}(\mathbf{s}_1) \in \{0,1\}^{256} \tag{6}$$

where $u = 2, 3, \ldots, l + 1$ and CBH is the current block header.

The current block header represented to be CBH in (5) is the on-chain information that is stored in the Internet. Anyone thus can access this on-chain information, leading to that anyone can make the same hash vector that a miner made during his mining competition work.

Second, we give a construction method of PCM for a given previous hash. This construction method has to be designed to satisfy two conditions:

C1. Any verifier can reconstruct the PCM using on-chain information that the miner has used.

C2. Formation of a PCM can vary from block to block.

First, suppose that C1 is not met. One choice a miner can opt out is to include a constructed PCM in a block for making any verifier to check the validity of this block. This may result in, under the condition that the block size is fixed, reducing the number of transactions stored in the block. Second, suppose C2 is not met. We then remind the literature survey given in Section III. This survey is to indicate that developing ASIC-LDPC decoders for supporting a single PCM can be possible. Thus, if C2 is not satisfy, there is a possibility to develop ASIC mining devices. Thus, C1 and C2 must be satisfied simultaneously.

Now, we begin to explain the proposed construction method that can satisfy C1 and C2. First, we let you know that the PCM parameters such as the code length n, the row degree w_r and the column degree w_c are essentially required to construct a PCM **H**. More concretely, if they are provided, we can assign the value of m, the number of rows of **H** in (1). These parameters are included in the block header, as we have stated in the second paragraph in this section.

The proposed method is based on the method of Gallager [7]. This aims to construct **H** that can be decomposed into a set of sub-matrices of size $w_c \times n$ as follows:

$$\mathbf{H} = \begin{bmatrix} \mathbf{A} \\ \pi_1(\mathbf{A}) \\ \vdots \\ \pi_{w_c-1}(\mathbf{A}) \end{bmatrix} \in \{0,1\}^{\frac{nw_c}{w_r} \times n} \tag{7}$$

where $\pi_i(\mathbf{A})$ is the ith submatrix constructed by random per-

mutation of the columns of **A**, π_i is the ith permutation order, and

$$\mathbf{A} := \begin{bmatrix} \mathbf{1}_{w_r} & & \\ & \ddots & \\ & & \mathbf{1}_{w_r} \end{bmatrix} \in \{0,1\}^{w_c \times n}$$

whose the ith row has w_c 1s in a row from $(i-1)\times w_r$ to $i\times w_r$ and

$$\mathbf{1}_{w_r} := [1 \quad 1 \quad \cdots \quad 1] \in \mathbf{1}^{1\times w_r}.$$

Let w_c, w_r and n are fixed. Different PCM can be constructed by varying the permutation orders, leading to the serial creation of different PCMs. To come up with different random permutation from block to block, we use a previous hash. Namely, we let PBHV be assumed to be an array of 32 bytes to represent the given previous hash. We generate an initial seed value S below

$$S := \text{PBHV}[0] + \text{PBHV}[1] + \cdots + \text{PBHV}[31] \tag{8}$$

where PBHV[i] is the ith element of PBHV. The first permutation order is generated using a seed value S. The ith permutation order is then generated using $S - i + 1$. The pseudo code of this construction method of PCM is given in Table VI.

We explain how the pseudo code given in Table VI can satisfy the conditions C1 and C2 mentioned earlier. In Step 4, the ith permutation order is constructed using $S - 1 + i$. Any verifier can get the same value S without any communications because PBHVs are on-chain information and thus available within the chain. Thus, the verifiers can easily reconstruct what the miner has constructed in the past, which confirms that the proposed method satisfies C1. Second, PBHVs are hashes and thus possess the characteristics of random numbers; i.e., the initial seed value is a random number. All of the permutation orders are provided using the initial seed value. Thus, the orders vary from block to block, which confirms that the proposed method satisfies C2 as well.

D. Construction of ECC puzzle generation functions and ECC puzzles

Let the current block header (CBH) except nonce be given. We define an ECC puzzle generation function by concatenating the HVG part and the decoder part.

Definition 3 – *ECC puzzle generation function*: Let the current block header (CBH) except nonce be given. For this given CBH, an ECC puzzle generation function (ECCPGF) is defined to be a composite function as follows:

$$h_{\text{CBH}} : \{\text{nonce}\} \mapsto \mathbf{c} \in \{0,1\}^{n\times 1} \tag{9}$$

where **c** is the output of a decoder defined in Definition 4.

Definition 4 – *Decoder*: A decoder takes both the hash vector **r** in (4) and the PCM **H** in (7) as its inputs and runs the message passing algorithm given in [32] to yield a vector **c** of size n:

$$\mathcal{D}_{MP} : \{\mathbf{r}, \mathbf{H}\} \mapsto \mathbf{c} \in \{0,1\}^{n\times 1}. \tag{10}$$

TABLE V. The pseudo codes for ECCPoW

Inputs: CBH and PCM **H**	
Step 1:	A nonce is uniformly chosen from [0, $2^{32} - 1$]
Step 2:	Construct a HV **r** using (4) with a chosen nonce and the given CBH.
Step 3:	Obtain a vector **c** using (10) with the given PCM **H**.
Step 4:	If the constraint in (11) is satisfied, then go to Step 5.
Step 5:	Block generation & broadcast

TABLE VI. The pseudo codes to construct PCM

Inputs: n, w_c, w_r and BHV	
Output: **H**	
Step 1:	Construct S using (8).
Step 2:	Construct **A** by following the statements below (7) and **H** = **A**.
Step 3:	for i = 2 to $w_c - 1$
Step 4:	Construct π_i with the seed value $S - i + 1$.
Step 5:	$\mathbf{H} = \left[\mathbf{H} \quad \pi_i(\mathbf{A})^T\right]$.
Step 6:	end
Step 7:	$\mathbf{H} = \mathbf{H}^T$.

When the CBH with a selected nonce is given, we can make a hash vector **r** according to (4). Using the previous hash value included in the CBH, we can make a PCM **H** as well. The decoder then takes both of them to yield the binary word, as we have shown in (10). This binary word is the output of ECCPGF, as we have shown in (9)

A mapping rule in the LDPC decoder depends on the form of **H**. As we have discussed in the subsection IV.C, we can choose to vary **H** from block to block. As blocks are mined endlessly, infinitely many PCMs can be made. Thus, as we have mentioned in the subsection III.B, this can deter the development of an ASIC-LDPC decoder. Making ASIC chips to function as an ECCPGF becomes extremely difficult, if not impossible. Thus, we use this ECCPGF to define an ECC puzzle as follows.

Definition 5 – *ECC Puzzle*: An ECC puzzle constructed using a given ECCPGF defined in (9) is defined to be

$$\text{find} \quad \textit{nonce} \quad \text{subject to} \quad \mathbf{H}h_{CBH}(\textit{nonce}) = \mathbf{0}_m. \tag{11}$$

Namely, this puzzle is a problem where we aim to find a nonce satisfying the constraint given in (11).

We shall note that this puzzle can be time-variant from block to block and resistant to ASIC chips. We will provide details in Section VI.

Last, we provide codes to solve a puzzle in Table V. In Step 1, a nonce is selected from 0 to $2^{32} - 1$. We construct a hash vector **r** using (4), as we have shown in Step 2. In Step 3, we execute the decoder to give an output by taking both **r** and **H**. The decision is done using this output in Step 4. If the output is not a codeword, we repeat the routines from Step 1 to Step 4. Similar to Bitcoin, there is a case in which we cannot find the solution even we consider the whole nonces. In such a case, we modify the fields such as timestamp and Merkle tree value of the current block header. This modification can lead to the variation of the hash vector (4). We then repeat the whole routines with this modified block header.

V. Analysis on Error-Correction Codes Proof of Work

To solve an ECC puzzle, we repeat the routines from Step 1 to Step 4 many times until finding a nonce that satisfies the constraint in (11). A simple question such that what is the number of trials for finding this nonce naturally arises. To this end, we conduct a probabilistic study for giving answers to the following three questions:

Q1. What is the number of hash cycles needed to solve a given puzzle?
Q2. Which are the parameters which affect the number of hash cycles needed?
Q3. How does the number of miners affect the number of hash cycles needed?

We define the hash cycle, the success event, and the mining game, respectively.

Definition 6 – *Hash Cycle*: The single execution of the whole routines from Step 1 to Step 4 given in Table V is defined to be single hash cycle.

Definition 7 – *Success Event*: A success event occurs if a nonce such that the decoder defined in (10) can return a codeword is found, i.e., the constraint in (11) is satisfied.

Definition 8 – *Mining Game*: Let both a PCM $\mathbf{H}$ of size $m \times n$ and a CBH except for nonce be given. There are M miners, and each use a single computer with the same computing capacity. A mining game (MG)

$$\mathrm{MG}\left\{\mathbf{H}, \mathrm{CBH}, M, p\right\} \tag{12}$$

is defined that the miners compete with each other in a race to hit the success event first. We let p be the decoding success (DS) probability of the decoder for this given PCM $\mathbf{H}$, i.e.,

$$p \coloneqq \Pr\left\{\mathbf{r} : \mathbf{H}\mathbf{c} = \mathbf{0}_m\right\} \tag{13}$$

where $\mathbf{c}$ is the output of the decoder defined in (10).

For a given PCM $\mathbf{H}$ of size $m \times n$, there are 2^k codewords:

$$\left\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \cdots, \mathbf{c}_{2^k}\right\}.$$

Then, we define a sphere set for the given i^{th} codeword

$$\mathcal{A}\left(\mathbf{c}_i, l\right) \triangleq \left\{\mathbf{r} : \left\|\mathbf{r} - \mathbf{c}_i\right\|_h \le l \cap \mathbf{r} \in \left\{0,1\right\}^n\right\}$$

whose cardinality is

$$\left|\mathcal{A}\left(\mathbf{c}_i, l\right)\right| = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{l} = \sum_{l=0}^{s} \binom{n}{l}$$

where s is a positive integer and $\left\|\mathbf{x}\right\|_h$ is the number of 1s in $\mathbf{x}$.

We assume that the decoder defined in (10) is optimal, implying that it can correct up to t errors where t is obtained by (3). The decoder always yields the ith codeword when it takes an input belonging to the ith sphere, i.e.,

$$\mathcal{D}_{MP} : \left\{\mathbf{r}, \mathbf{H}\right\} \mapsto \mathbf{c}_i$$

where $\mathbf{r} \in \mathcal{A}\left(\mathbf{c}_i, t\right)$. Then, we have

$$\begin{aligned} p &= \sum_{i=1}^{2^k} \Pr\left\{\mathbf{c} = \mathbf{c}_i\right\} = \sum_{i=1}^{2^k} \sum_{l=0}^{t} \Pr\left\{\left\|\mathbf{r} - \mathbf{c}_i\right\|_h = l\right\} \\ &= \sum_{i=1}^{2^k} \Pr\left\{\mathbf{r} \in \mathcal{A}\left(\mathbf{c}_i, t\right)\right\} \end{aligned} \tag{14}$$

where $\mathbf{c}$ is an output of the decoder which takes a hash vector $\mathbf{r}$. Since the number of inputs that can be mapped into one of the codewords is

$$\sum_{i=1}^{2^k} \left|\mathcal{A}\left(\mathbf{c}_i, t\right)\right|,$$

the DS probability p can be expressed as follows:

$$p = 2^{-n} \sum_{i=1}^{2^k} \left|\mathcal{A}\left(\mathbf{c}_i, t\right)\right| = 2^{k-n} \left|\mathcal{A}\left(\mathbf{c}_i, t\right)\right| = 2^{k-n} \sum_{l=0}^{\left\lfloor \frac{d-1}{2} \right\rfloor} \binom{n}{l} \tag{15}$$

where the third equality comes from (3).

Intuitively, the number of trials increases as p decreases. It is natural to find which parameters make effects on p. To this end, we establish Proposition 1 to provide the behavior of p in terms of the code length n under the assumption that the row degree and the column degree are fixed.

Proposition 1 – Let $w_c \ge 3$, $w_r > w_c$ be fixed constants and their ratio be a fixed constant as well

$$w_c / w_r =: \alpha \in \left(0,1\right). \tag{16}$$

Let the size of a given PCM $\mathbf{H}$ be $m \times n$. For any $0 < \delta < 1/2$, we have

$$2^{-n\alpha} \le p \le 2^{-n\left(\alpha - H\left(\delta/2\right)\right)} \tag{17}$$

where $H(x)$ is the binary entropy function defined as follows:

$$H\left(x\right) = -x \log_2 x - \left(1-x\right) \log_2 \left(1-x\right).$$

Indeed, let the ratio further satisfy the following:

$$\alpha \in \left(H\left(0.25\right), 1\right). \tag{18}$$

Then, the DS probability p vanishes with increase in n.

Proof: From (1) with (16), we infer that

$$k - n = -n\alpha. \tag{19}$$

The results [7] state that the minimum distance of any regular LDPC code with constant w_r and $w_c \ge 3$ can linearly increase with increase in n. This statement indicates that the distance can be expressed as for any $0 < \delta < 1/2$,

$$d = \left\lfloor \delta n \right\rfloor.$$

Substituting (19) into (15) leads to an upper bound to p

$$p \le 2^{k-n} \sum_{l=0}^{\left\lfloor \frac{\delta n}{2} \right\rfloor} \binom{n}{l} \le 2^{k-n} 2^{nH\left(\frac{\left\lfloor \delta n \right\rfloor}{2n}\right)} \le 2^{-n\left(\alpha - H\left(\delta/2\right)\right)} \tag{20}$$

where the second inequality comes from the fact that any integers $n \geq k \geq 1$ with $k/n \leq 0.5$

$$\sum_{l=0}^{k} \binom{n}{l} \leq 2^{nH(k/n)}.$$

By assuming $t = 0$ in (14), the lower bound is obtained below:

$$p = 2^{k-n} \left| \mathcal{A}(\mathbf{c}_i, t) \right| \geq 2^{k-n} \left| \mathcal{A}(\mathbf{c}_i, 0) \right| = 2^{k-n} = 2^{-n\alpha}.$$

Last, let the ratio satisfy (18). Then, the term of the exponent in the upper bound in (17) is negative because for any $0 < \delta < 1/2$,

$$H(\delta/2) < H(0.25).$$

Thus, increasing n makes the upper bound on p reduced. This leads to that p vanishes as n goes to infinity. ■

We provide quick discussions on Proposition 1. First, for the fixed constant α, decreasing the code length n makes the lower bound on the DS probability p grow. Second, if we select any pairs of w_c and w_r which can satisfy (18), then p vanishes as n increases. This means that no one solves an ECC puzzle using these pairs for sufficiently large n. We thus have to avoid selecting such pairs for preventing this critical problem.

We define a random variable to represent the number of hash cycles required to end a MG and provide its statistical properties, respectively.

Definition 8 – *Random Variable*: For a given MG, X_M is defined as a random variable that represents the number of hash cycles to end this given MG, where the subscript M denotes the number of miners forging simultaneously and independently. We call this random variable First Success Hash Cycle (FSHC).

Theorem 2 – Let an MG{**H**, CBH, M, p} be given. Then the distribution that FSHC occurs at the lth hash cycle is

$$\Pr\{\mathrm{X}_M = l\} = p_{\mathrm{f,a}}^{l-1}\left(1 - p_{\mathrm{f,a}}\right)$$

which is a geometric distribution with a parameter

$$\left(1 - p_{\mathrm{f,a}}\right) \tag{21}$$

where $p_{\mathrm{f,a}} \coloneqq (1-p)^M$ is the probability that all of the miners fail to succeed in solving a given puzzle. Then, we have

$$\mathbb{E}[\mathrm{X}_M] = \left(1 - p_{\mathrm{f,a}}\right)^{-1} \tag{22}$$

and

$$\mathbb{V}[\mathrm{X}_M] = p_{\mathrm{f,a}}\left(1 - p_{\mathrm{f,a}}\right)^{-2}.$$

Proof: The proof is clear, as X_M follows the geometric distribution with (21). We thus omit it. ■

For the constant α defined in (16), Proposition 1 shows that the DS probability p grows with decrease in the code length n. When p grows, the parameter (21) reduces and converges to a real positive number. Thus, the expected value gets reduced and converges to a number, implying that an ECC puzzle becomes easier as n gets smaller.

TABLE VII. The lower bound given in (23)

$w_c = 4$ and $w_r = 5$	$M = 1$	$M = 5$	$M = 20$
Lower bounds, i.e., $\delta_1 = 0.3238$ in (23)			
$n = 80$, $k = 12$	1.58×10^{4}	0.31×10^{4}	0.08×10^{4}
$n = 120$, $k = 24$	6.03×10^{7}	1.20×10^{7}	0.30×10^{7}
$n = 160$, $k = 32$	2.46×10^{9}	0.49×10^{9}	0.12×10^{9}

Next, we consider a case in which n grows. In this case, the upper bound given in (17) is too loose to be considered unless α satisfies (18). Thus, we require another upper bound on p which is tighter than the previous upper bound. To this end, we invoke a table of [31]. For a certain pair of the column degree w_c and the row degree w_r, this table was obtained and given in the form of an upper bound δ_1 to the relative minimum distance δ_0, the ratio of the minimum distance d to the code length n. For $w_c = 4$ and $w_r = 5$, for example, the upper bound is 0.3238. For $w_c = 4$ and $w_r = 8$, it is 0.1765. The result was obtained from an asymptotic analysis; i.e., by letting n go to infinity.

We use the upper bound δ_1 to obtain a tighter bound on the DS probability p as follows:

$$p \leq g(n, k, \delta_1) \tag{22}$$

where δ_1 is given in the table of [31] for a certain pair of w_c and w_r and

$$g(n, k, \delta) \coloneqq 2^{k-n} \sum_{l=0}^{\left\lfloor \frac{\lfloor n\delta \rfloor - 1}{2} \right\rfloor} \binom{n}{l}.$$

The bound in (22) is obtained by simply replacing the minimum distance with the upper bound, respectively:

$$d \coloneqq n\delta_0 \leq \lfloor n\delta_1 \rfloor.$$

Once the upper bound on p has been obtained, we can use it to find a lower bound to the expected value of FSHC below:

$$\frac{1}{1 - \left(1 - g(n, k, \delta_1)\right)^M} \leq \mathbb{E}[\mathrm{X}_M] \tag{23}$$

We can examine the behavior of (22) with respect to the code length n and the number of miners M. In Table VII, we provide the lower bounds to the expected value by varying either n or M. They are obtained for a case in which the column degree w_c and the row degree w_r are 4 and 5, respectively. We can see that the lower bound increases with increase in n for the fixed M. That is, it increases from 1.58×10^4 to 2.46×10^4 when n is increased from 80 to 160. This result implies that increasing n makes an ECC puzzle more difficult to solve. For the other pairs of w_c and w_r given in [31], the same result is observed.

Now, for the fixed code length n, we consider the behavior of the expected value of FSHC by growing the number of miners M. Intuitively, as more miners are involved in solving a given puzzle, this puzzle has to end early. In addition, if an infinite number of miners work, any MG has to end at the 1st hash cycle. These intuitions can be confirmed by Corollary 1 given below.

Corollary 1: Let a MG{**H**, CBH, *M*, *p*} be given. The expected value of a FSHC given in (22) decreases with increase in the number of miners *M*. In particular, this value can converge to 1 as *M* goes to infinity.

Proof: It is immediately seen that

$$\frac{d\mathbb{E}[\mathrm{X}_M]}{dM} = -\frac{\log p_{\mathrm{f,a}}}{p_{\mathrm{f,a}}} \le 0$$

implying that the expected value given in (22) is a decreasing function of *M*. As *M* goes to infinity, the parameter defined in (21) goes to one. Thus, the expected value converges to one. ■

The decoding process, i.e., Step 3 in Table V, can occupy the most computational time in a single hash cycle. In this decoding process, matrix-vector products are required, implying that the computational cost to run a single hash cycle can be modeled as O(*mn*). Each miner uses the same single computer in a given MG. Thus, we can assume that each miner only runs τ operations per second. This assumption makes us define an expected value of a block generation time as follows.

Definition 10 – *Block Generation Time*: A MG{**H**, CBH, *M*, *p*} is given. Each miner is assumed to run τ operations per second. Then, the block generation time *T* can be defined as

$$T \coloneqq \tau^{-1}\mathbb{E}[\mathrm{X}_M]O(mn). \tag{24}$$

Both Proposition 1 and Theorem 2 can indicate that the expected value in (24) is an increasing function of the code length *n*. We thus immediately conclude that the block generation time *T* is an increase function of *n*.

VI. Discussions on ECCPoW

We prepare this section to provide plentiful discussions on ECCPoW. To this end, we begin to define general properties of PoWs and explain how ECCPoW can have them. We introduce a new property that ECCPoW can only have. This new property makes ECCPoW become a solution to the problems such as M1 and M2 which we have stated in Section I.

We now begin to define properties below.

P1. A puzzle has to be time-consuming, but it is easy to check whether a given solution is correct or not.
P2. Any previous solution cannot be used to find a current solution.
P3. A puzzle can be solved with overwhelming probability 1 if and only if miners follow the routines of PoW.
P4. The difficulty of a puzzle can change.
P5. A puzzle can be time-variant from block to block.

The existing PoWs have the properties from P1 to P4. Let us begin to consider how bitcoin satisfies them. First, each puzzle is expected to be solved per 10 minutes. In contrast, validating a given solution can be instantly done. Second, SHA256 takes the block header to get its hash. The block header in Bitcoin has the timestamp field. Due to this timestamp field, the contexts of the current block header can be different to those of any previously mined blocks. Thus, any solutions given in these mined blocks are useless in finding a current solution. Thus, P2 holds. Third, the number of possible hashes is 2^{256} while the number of solutions is about $2^{(256-L)}$ where *L* is a pre-defined value according to the difficulty of the provided block header. Thus, a possibility that a provided nonce is a solution is 2^{L}. In the 567,657th block of bitcoin, *L* is 72. This is to imply that for this block, the probability that a randomly given nonce can be a solution is 2^{-72}. Thus, all the miners have to follow the routines in Table I to solve a puzzle, meaning P3. Next, whenever 2016 blocks are mined, the difficulty periodically changes, implying P4. Last, existing PoW does not hold P5. The reasons for this issue will be given after explaining how ECCPoW satisfies P5.

Now, we begin to explain how ECCPoW can satisfy all the properties using the results established in the previous section.

Corollary 2: ECCPoW can satisfy the first property P1.

Proof: For a given nonce, we complete the verification whether this nonce is a solution or not by conducting the steps from Step 1 to Step 4 in Table V. This verification requires a single construction of a hash vector and a single execution of the decoder (10). In contrast, to solve an ECC puzzle, one has to do work by repeating them many times, as we have stated in Section V. ■

Corollary 3: ECCPoW can satisfy the second property P2.

Proof: As we have stated in Definition 5, we construct an ECC puzzle using an ECCPGF in Definition 3. This ECCPGF takes a hash vector as its input and produces a binary word as its output. The decision is made based on this binary word. To make this hash vector, we put a given set of the block header with a nonce through SHA256. Similar to Bitcoin block header, our block header includes the timestamp field. Thereby, one cannot effortlessly create a particular hash vector even if one has the full knowledge of all the previous solutions, i.e., all the collection of the nonces each of which was a solution to mine a block in the past. This makes ECCPoW satisfy P2. ■

Let us assume that there exists a knowledgeable miner who can solve ECC puzzles by referring to an input-output mapping table of ECCPoW without actually carrying out the decoding work. This miner can then mine blocks much faster than other honest miners can. To show the non-existence of such a miner in ECCPoW, we prepare Corollary 4. This corollary is to imply that all miners cannot but have to run the decoder to solve ECC puzzles.

Corollary 4: ECCPoW satisfies the third property P3.

Proof: Such a malicious miner can appear under an assumption that this miner has a mapping table that maps a given input to its corresponding output without actually carrying out the decoding work (10). That is, he puts each hash vector into the decoder to find what this decoder returns. He then uses this known information to construct the mapping table. However, the number of hash vectors is 2^{n}. For sufficiently large *n*, i.e. $2^{256} \sim 10^{77}$ for $n = 256$, we can easily see that this construction is impossible. Besides, this table depends on a PCM **H**, meaning that the mapping table is updated newly whenever **H** varies. As we have stated in the subsection IV.C, **H** can be set to vary from block to block. Thereby, whenever he aims to solve a new ECC puzzle, he has to construct a new mapping table. The above

assumption there exists the mapping table, we thus conclude, is invalid. Thereby, any miners in ECCPoW cannot but have to obey the routines in Table V to solve puzzles; this leads to the conclusion that ECCPoW satisfies P3. ■

Now, we remind the discussions regarding Theorem 2. In the discussions, we have proven that an ECC puzzle can become easier as the code length n gets smaller. We then have empirically shown that the increase in n makes this puzzle more difficult to solve. They are to indicate that ECCPoW holds P4.

Corollary 5: ECCPoW can satisfy the fourth property P4.
Proof: It is clear; we thus omit it. ■

A PCM **H** defines a mapping function of the decoder in (10). This decoder is used to define an ECCPGF in (9) which is used to construct an ECC puzzle. We conclude that this ECC puzzle can be a function of **H**. As we have mentioned in the subsection IV.C, we construct **H** using a previous hash, i.e., the ith PCM is made using a hash of the $(i–1)^{\text{th}}$ block. Such a construction can make **H** time-variant from block to block, leading to that the ECC puzzle can also be time-variant from block to block. This discussion leads to a conclusion that ECCPoW holds P5 below.

Corollary 6: ECCPoW can satisfy the fifth property P5.
Proof: It is clear; we thus omit it. ■

We now prepare to provide discussions how ECCPoW can be a solution to the problems such as *i*) M1 the re-centralization of the mining markets and *ii*) M2 the huge energy consumption for mining. As miners mine new blocks continuously, an infinite number of PCMs are constructed. Hence, we cannot count how many the number of PCMs is required in advance. Also, we cannot expect what PCMs will be generated. We remind the example given in [25], showing that there is no ASIC-LDPC decoder to support the infinite number of PCMs. By combining this example with the fact that there is an infinite number of PCMs are required, we thus conclude that the LDPC decoder in ECCPoW is operated by either graphical processing units or CPUs. This is the fundamental reason that ECCPoW becomes the solution remediating the aforementioned problems M1 and M2 which we have stated in Section I. This is the most remarkable contribution of ECCPoW.

We end this section by providing a reason that SHA based PoW does not hold P5. We consider SHA256 as a puzzle generation function (PGF) in Bitcoin because it is used to generate a puzzle. This PGF is fixed regardless of the height of a block that miners aim to mine. The puzzle made using a fixed PGF is not time-variant at all, leading to that such a puzzle cannot hold P5.

VII. Conclusions

PoW is fundamental to public blockchains, as it can be used to prohibit an unauthorized modification of mined blocks. For existing PoWs, ASIC mining devices have been introduced and used to mine blocks. The usage of such devices can cause the problems such as M1 and M2 that we have stated in Section I.

In this paper, as a solution to these problems, we proposed a new proof-of-work using error-correction codes which we call ECCPoW. To the best of our knowledge, this is the first study in which a decoder of LDPC codes is applied to the consensus part of blockchain. Specifically, we combined this decoder with SHA256 to construct a composite function named as ECCPGF (9). We used ECCPGF to define a corresponding ECC puzzle (11) and provided the routines to solve a given ECC puzzle.

We also studied the behavior of the expected value of the number of hash cycles for solving a given ECC puzzle. We showed that this value can be either increased or decreased as varying the code length, the size of a hash vector taken by the decoder, and the number of miners. Indeed, we discussed how ECCPoW can satisfy the five properties defined in Section VI, which shows the value of ECCPoW as a general PoW.

As we have reviewed in Section III, there is no ASIC decoder to support an infinite number of LDPC codes. By motivated this survey, we intended to vary the codes from block to block. As a result, we made ECCPGF time-variant, meaning that its mapping function can vary from block to block. This leads to the time-variant property P5 defined in Section VI. This is the most innovative part of ECCPoW in repressing the advent of ASICs, implying that the problems caused from the usage of ASICs can be solved using our ECCPoW.

We have implemented ECCPoW and forked two blockchains such as Bitcoin and Ethereum by replacing their consensus with the implemented ECCPoW. We name these forked versions as BTC-ECC and ETH-ECC, respectively. All of the source codes including ECCPoW, BTC-ECC and ETH-ECC can be available in a GitHub site [34]. We have also provided manuals that shows how to install them, how to compile them and how to run. We believe that anyone can easily operate their own BTC-ECC or ETH-ECC by following these manuals. We believe that this site can be the first repository that makes people in the error-correction codes community get involved in the blockchain community.

References

[1] U. W. Chohan, "The double spending problem and cryptocurrencies", SSRN Electronic Journal, Jan. 2019

[2] Nakamoto, S. *Bitcoin: A Peer-to-Peer Electronic Cash System*. Retrieved from: https://bitcoin.org/bitcoin.pdf

[3] Quynh H. Dang, "Secure Hash Standard FIPS PUB 180-4," U.S. Dept. of Commerce and NIST, Aug. 4th 2015. https://www.nist.gov/publications/secure-hash-standard, Accessed Sept. 25th, 2018.

[4] Jordan Tuwiner, "Bitcoin mining hardware ASICs," https://www.buybitcoinworldwide.com/mining/hardware/, June 30th, 2018. Accessed Sept. 26th, 2018.

[5] A. Gervais, G. Karame, V. Capkun and S. Capkun, "Is bitcoin a decentralized currency?", *IEEE Security and Privacy Magazine*, Vol. 12, No. 3, pp. 54 – 60, May, 2014

[6] Digiconomist. (2018). *Bitcoin Energy Consumption Index*. Retrieved April 3, 2018 from: https://digiconomist.net/bitcoin-energy-consumption

[7] R. G. Gallager, *Low Density Parity Check Codes*, Monograph, M.I.T. Press, 1963

[8] R. J. McEliece, "A public-key cryptosystem based on algebraic coding theory", DSN Progress Report, pp. 114-116, Jan.~Feb. 1978

[9] D. J. Bernstein, T. Lange and C. Peters, "Wild McEliece", in *Proc. Sel. Areas Cryptography*, vol. 6544, pp. 143 – 158, 2011

[10] M. Baldi, M. Bianchi, F. Chiaraluce, J. Rosenthal and D. Schipani, "Using LDGM codes and sparse syndromes to achieve digital signatures", in *Post-Quantum Cryptography 2013, ser. Lecture Notes in Comput. Sci.*, vol. 7931, Springer, pp. 1 – 15, 2013.

[11] Marco Baldi, "LDPC codes in the McEliece cryptosystem: attacks and countermeasures", NATO Science for Peace and Security – D: Information and Communication Security, vol. 23, pp. 160 – 174,

[12] M. Baldi, M. Bodrato and F. Chiaraluce, "A new analysis of the McEliece cryptosystem based on QC-LDPC codes", in 6th Int. Conf. on Sec. and Cryptography for Networks. Springer, pp. 246 – 262, 2008

[13] T. P. Berger, P-L. Cayerel, P. Gaborit and A. Otmani, "Reducing key length of the McEliece cryptosystem", in Progress in Cryptology, Africacrypt'2009, ser. LNCS, vol. 5580. Springer, pp. 77 – 97, 2009

[14] V. M. Sidelnikov, "A public-key cryptosystem based on binary Reed–Muller codes," Discrete Mathematics and Applications, vol. 4 no. 3, 1994.

[15] L. R. Knudsen and B. Preneel, "Construction of secure and fast hash functions using nonbinary error correcting codes", IEEE Transactions on Information Theory, vol. 48, no. 9, pp. 2254 – 2539, Sept., 2002.

[16] S. Ermon, C. P. Gomes, A. Sabharwal and B. Selman, "Low-Density Parity Constraints for Hasing-Based Discrete Integration", Proceedings of the 31st International Conference on Machine Learning, PMLR 32(1):271-279, 2014.

[17] Vadhan, S., *Pseudorandomness.* Foundations and Trends in Theoretical Computer Science, 2011.

[18] Vitalik Buterin. Dagger: *A memory-hard to compute, memory-easy to verify scrypt alternative*, Tech Report, hashcash.org website, 2013.

[19] https://github.com/dashpay/dash/wiki/Whitepaper

[20] T. Black and J. Weight, X16R ASIC resistant by design, 2018. https://ravencoin.org/assets/documents/X16R-Whitepaper.pdf

[21] C. E. Shannon, "A mathematical theory of communication", The Bell System Technical Journal, Vol. 27, 99. 379 – 423, 623 – 656, Jul., Oct., 1948

[22] D. J. C. Mackay and R. M. Neal, "Near Shannon limit performance of low density parity check codes", Electronics Letters, Vol. 33, No. 6, pp. 457 – 458, 1997

[23] M. Awais and C. Condo, "Flexible LPDC decoder architectures", VLSI Design, Vol. 2012, Doi: 10.1155/2012/730835

[24] C-H. Liu, S-W. Yen, C-L. Chen, H-C. Chang, C-Y. Lee and S-J. Jou, "An LDPC decoder chip based on self-routing network for IEEE 802.16e Applications" IEEE Journal of Solid-State Circuits, Vol. 43, No. 3, pp. 684 – 694, Mar. 2008

[25] S. Shao, P. Hailes, T-Y. Wang, J-Y. Wu, R. G. Maunder, B. M Al-Hashimi and L. Hanzo, "Survey of Turbo, LDPC and Polar Decoder ASIC Implementation", IEEE Communications Surveys & Tutorials 2019

[26] Y. L. Ueng, B. J. Yang, C. J. Yang, H. C. Lee, and J. D. Yang, "An Efficient Multi-Standard LDPC Decoder Design Using Hardware-Friendly Shuffled Decoding," IEEE Trans. Circuits Syst. I, vol. 60, no. 3, pp. 743–756, March 2013.

[27] C. Beuschel and H-J. Pfleiderer, "FPGA implantation of a flexible decoder for long LDPC codes", Proc. Int. Conf. Field Program. Logic Appl., Sep. 2008, pp. 185–190

[28] P. Hailes, R. G. Maunder and L. Hanzo, "A survey of FPGA-based LDPC decoders", IEEE Communications Surveys & Tutorials, Dec. 2015.

[29] A. B. Keha and T. M. Duman, "Minimum distance computation of LDPC codes using a branch and cut algorithm", IEEE Transactions on Communication, Vol. 58, No. 4, pp. 1072 – 1079, Apr., 2010

[30] Y. Hashemi and A. H. Banihashemi, "Tight Lower and Upper Bounds on the Minimum Distance of LDPC Codes", IEEE Communications Letters, Vol. 22, No. 1, pp. 33 – 36, Jan., 2018

[31] Y. Ben-Haim and S. Litsyn, "Upper bounds on the rate of LDPC codes as a function of minimum distance", IEEE Transactions on Information Theory, vol. 52, no. 5, pp. 2092 – 2100, May., 2006

[32] W. E. Ryan and S. Lin, *Channel Codes Classical and modern*, Cambridge, 2009

[33] Whitefire990, https://drive.google.com/file/d/14ZYwpI-t6rnQ3I_SEZHvvDwe2yQUnTJ5/view, Nov., 2019

[34] DeSecure Blockchain,https://github.com/cryptoecc/ETH-ECC, 2020

symmetry

MDPI

Article

ECCPoW: Error-Correction Code based Proof-of-Work for ASIC Resistance

Hyunjun Jung [1] and Heung-No Lee [2,*]

[1] Blockchain Internet Economy Research Center, Gwangju Institute of Science and Technology, Gwangju 61005, Korea; junghj85@gist.ac.kr

[2] School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, Korea

* Correspondence: heungno@gist.ac.kr; Tel.: +82-62-715-2237

Received: 20 May 2020; Accepted: 5 June 2020; Published: 9 June 2020



Abstract: Bitcoin is the first cryptocurrency to participate in a network and receive compensation for online remittance and mining without any intervention from a third party, such as financial institutions. Bitcoin mining is done through proof of work (PoW). Given its characteristics, the higher hash rate results in a higher probability of mining, leading to the emergence of a mining pool, called a mining organization. Unlike central processing units or graphics processing units, high-cost application-specific integrated circuit miners have emerged with performance efficiency. The problem is that the obtained hash rate exposes Bitcoin's mining monopoly and causes the risk of a double-payment attack. To solve this problem, we propose the error-correction code PoW (ECCPoW), combining the low-density parity-check decoder and hash function. The ECCPoW contributes to the phenomenon of symmetry in the proof of work (PoW) blockchain. This paper proposes the implementation of ECCPoW, replacing the PoW in Bitcoin. Finally, we compare the mining centralization, security, and scalability of ECCPoW and Bitcoin.

Keywords: error-correction codes proof-of-work (ECCPoW); proof-of-work (PoW); ECCPoW implementation; ASIC resistance

1. Introduction

We use digital signatures from third-party trust agencies to promote trust in Internet commerce. We warrant proof of data forgery using middlemen. Satoshi Nakamoto proposed an electronic money system without a middleman in a peer-to-peer (P2P) network through a Bitcoin white paper [1]. Bitcoin applies blockchain technology to electronic monetary systems to guarantee transactions without intermediaries (e.g., banks). Blockchain is a ledger management technology based on a distributed computing technology that cannot be arbitrarily modified by storing the transaction content in a chain-based distributed data storage environment in the form of a block [2,3].

The blockchain stores the same ledger on a global network and is designed to pay certain rewards to maintain the block. This is called mining, and mining creates blocks and obtains cryptocurrency by executing a hash function. Miners belong to mining pools because of the probability and convenience of being rewarded for mining cryptocurrency [4,5].

The hash rate is the number of hash values calculated per second as a measure of computational processing power for mining cryptocurrency. The hash rate of cryptocurrency is determined by the total number of the participating nodes. Most miners belong to a hash pool and occupy a high percentage of the hash rate. We should be concerned about the risk of a double-payment attack if the mining pool in the

blockchain accounts for a high percentage of the total hash rate of cryptocurrency [6]. A double-payment attack occurs when the mining pool seizes at least 51% of the total hash rate to determine the branch of the blockchain to the desired side. Recent studies have shown that a double-payment attack can be made to benefit from a low share of the hash rate [7].

Bitcoin miners receive blockchain information (version, previous block hash, Merkle root, bits, and others) and execute hash functions (e.g., the secure hash algorithm (SHA 256). Miners create a block if the output value of the hash function is less than the target level of difficulty. Currently, Bitcoin requires an increasingly high operation to create blocks. Miners purchase application-specific integrated circuits (ASICs) with high per-second computing power to mine Bitcoin and join the mining pool. Low-power miners using central or graphics processing units (CPUs or GPUs) have difficulty mining.

Bitcoin uses the number of zeros in the front digits of the output from the SHA 256 function to generate blocks. The mining difficulty increases with the number of zeros. Miners buy ASICs to compute SHA functions quickly. The mining machine Antminer S9 (13,000 MH/s) is approximately 8800 times faster than GTX1060 (1478 MH/s). The recently released S19 has a speed of 95 TH/s, and S19Pro exhibits performance of 110 TH/s [8]. Miners using CPUs or GPUs in Bitcoin, and miners using ASIC chips do not have equal chances for success because ASIC miners have an 8800-fold greater chance of success in mining.

Blockchain was proposed to allow nodes to participate as miners freely and to share mining rewards fairly. Blockchain is not free to participate in as a miner. The participating miners compete against equity. Several methods have been proposed to curb the development of ASIC miners, but in the end, these methods have not prevented ASIC development. As a new mining function to prevent the development of ASIC miners, we proposed the error-correction code proof of work (ECCPoW) concept, which combines the low-density parity-check (LDPC) decoder and hash function [9]. In addition, we analyzed the hash cycle of ECCPoW and demonstrated that it could be used for blockchain mining.

This paper contributes to the phenomenon of symmetry in the proof of work (PoW) blockchain. The PoW blockchain tends to increase the hash rate along with the total size of the blockchain. The size of the hash rate is related to the computing power required by block generation. Some people take issue with the large amount of wasted power used to create the PoW blockchain. However, the PoW blockchain guarantees high stability with the power required to create the block. The transition to other PoWs due to problems in the PoW blockchain causes dangerous problems. Thus, ECCPoW mitigates the power problems in the PoW blockchain and helps reduce the symmetry of the hash rate, which increases in proportion to the size of the blockchain in the PoW blockchain.

The contribution of this paper is two-fold. It introduces the proposed ECCPoW and proposes an implementation method. The second contribution is to introduce the process of experimenting in Bitcoin by replacing the SHA 256 function with the ECCPoW function. This paper proposes the creation of a cryptographic puzzle that changes every block and shows how to apply the crypto puzzle decoder to the solution. We present the implementation of the proposed method by replacing ECCPoW in Bitcoin. We also measure the block generation time of the ECCPoW. Finally, we compare ECCPoW and Bitcoin by implementing them in the same environment.

This study contributes to the previous literature on ASIC resistance in PoW blockchain. Previous studies have used forced memory access, leading to the inefficient behavior of ASICs. Another method is to make ASICs challenging to produce using various hash functions. The last method is to change the existing hash function to another hash function when ASICs are released. Therefore, ECCPoW combines the LDPC decoder with a hash function to create the effect of releasing different hash functions in each block. Implementing ASICs is difficult for the LDPC decoder due to cost issues. Furthermore, the PoW blockchain without ASICs reduces the mining centralization. In addition, ECCPoW can reduce power consumption by maintaining the advantages of the PoW.

This paper is structured as follows. Section 2 introduces the relevant research on the development of the prevention of ASIC miners. Section 3 presents ECCPoW and development methods. Section 4 reveals the results of the experiment by loading the ECCPoW into Bitcoin. Section 5 evaluates the mining centralization, security, and scalability of the ECCPoW. Finally, Section 6 presents the conclusion of the paper.

2. Related Work

2.1. Ethereum

Ethereum is a distributed computing platform developed to implement smart contract functions based on blockchain technology. In July 2015, Vitalik Buterin developed Ethereum in the C++ and Go languages. Ethereum uses the Ethash algorithm. Ethereum plans to change from a PoW to a proof of stake in the future [10]. Ether Solarium is against the use of the nonlinear directed acyclic graph (DAG) in ASICs. The initial size of the DAG was about 1 GB, and it was designed to increase in size linearly over time. In October 2019, the size of the DAG was 3.99 GB, which was maintained until December 20, 2020 [11,12].

Ethereum approved the application of programmatic proof of work (ProgPoW) [13] to respond to the centralization of mining in ASIC in 2019. First, ProgPoW regularly changes mining problems to a problem in which GPUs can adapt quickly. Second, ProgPoW makes the most of all the components of the graphics card for mining. Moreover, ProgPoW uses randomly generated problems based on block numbers and is designed for the efficient operation of GPUs. This reduces performance differences compared to ASICs.

2.2. The X-11 Series

The X-11 series is a cryptocurrency mining algorithm that uses as many hash functions as the number indicated behind the X [14]. The X-11 series used hash functions to add depth and complexity to curb ASICs. The X-11 series connects several hash functions and uses the output value of the hash as the input value of the next hash. Typically, the algorithm is used in Dash. The concept of the X-11 series uses multiple hashes to increase security and prevent ASIC mining. Currently, however, the X-11 series has been upgraded to increase the number of hash functions, such as X-13, X-14, X-15, X-16R, and X-17 because ASIC mining is still possible.

2.3. CryptoNote

CryptoNote was designed to be more inefficiently executed with a GPU than a CPU [15] to prevent ASIC mining. The performance of CryptoNote is susceptible to memory latency because memory creation and subsequent read operations occur repeatedly. This is similar to Etherium's Ethash function. CryptoNote creates blocks by determining the hash function to be used after memory-intensive tasks.

Despite such attempts, Bitmain released ASICs optimized for CryptoNote algorithms in March 2018. Monero uses CryptoNote to change its mining algorithm twice a year to prevent ASICs. Monero's algorithm change has prevented the use of ASICs. However, the frequent hard fork execution (radical changes) caused participants to break away from mining. In the end, there was a risk of the centralization of mining. To prevent frequent hard forks, RandomX proposed a key block concept that periodically changes mining methods [16].

3. The Proposed Method

This chapter provides an overview of and the implementation method of the proposed ECCPoW. We proposed the concept of ECCPoW to increase the resistance to ASIC, using a hash function that makes ASIC development difficult. Moreover, ASIC resistivity research reviews methods of inducing the loading

of memory; for example, Ethereum uses several hash algorithms, such as X-11. However, ASICs for Ethereum and the X-11 series have been developed. Thus, ECCPoW is a method for miners to release different hash functions for each block to curb the emergence of ASICs.

The ECCPoW can help ease the centralization of mining. The conversion to ASICs in mining is made by ordinary people to avoid being excluded from mining and by a small group of people with capital power to monopolize mining. Mining centralization of a small group is likely to lead to mining blocks for malicious purposes and forging and tampering with the mined blocks. The implementation of ASICs in the LDPC decoder lacks flexibility due to structural cost issues [17]. Moreover, ECCPoW proposes a method of combining the LDPC decoder with a hash function. We help reduce the risk in the blockchain by mitigating mining centralization.

3.1. ECCPoW Overview

The ECCPoW is a POW that uses the ECC decoder that is used in communication, which can be implemented using ASICs. As a simple example, cell phones use ASICs to implement an ECC decoder quickly and at low power. The parity-check Matrix H determines the design of the ECC decoder based on ASICs. In other words, ASIC equipment can produce a decoder using parity-check matrices. For mobile phones, ASICs have standardized parity-check matrices that allow an ECC decoder design. Building ASICs to match the decoder supporting countless parity-check matrices is difficult due to cost problems and decoder size problems.

Randomly, ECCPoW changes each block parity-check matrix (i.e., ECCPoW uses an infinite number of parity-check matrices). As a result, ECCPoW inhibits the development of ASICs for ECC decoders. The ECC decoding algorithm only runs on a CPU or GPU. For example, even if the SHA function used in the PoW is executed quickly, a bottleneck occurs in the execution of ECC decoding algorithms.

3.2. Create a Cryptographic Puzzle That Changes Every Block

The ECCPoW aims to create a cryptographic puzzle that changes every block. We changed the composite function used to create the cryptographic puzzle using the generation method by Gallagher [18] and the previous hash value. In other words, ECCPoW randomly generates an LDPC matrix used by the decoder of the composite function. The Gallagher method requires variables. Table 1 displays the definition of the variables used in the LDPC.

Table 1. Variables in the low-density parity-check Matrix H.

Variables	Definition
n	Number of columns in H
m	Number of rows in H
w_c	Number of 1s in each column
w_r	Number of 1s in each row

The LDPC matrix satisfies the equation $nw_c = (n-k)w_r$, where k is $n-m$ and 2^k is the total number of symbols that can be generated. When the variables are given, the LDPC Matrix H of size A is generated by the following method:

Step 1: Create a partial matrix of size $\frac{m}{w_c} \times n$

$$A_1 := \begin{bmatrix} \underbrace{1\,1\,\cdots\,1}_{w_r} & & & \\ & \underbrace{1\,1\,\cdots\,1}_{w_r} & & \\ & & \ddots & \\ & & & \underbrace{1\,1\,\cdots\,1}_{w_r} \end{bmatrix} \in \{0,1\}^{\frac{m}{w_c} \times n}$$

Step 2: Generate $w_c - 1$ submatrices by randomly permutating the following matrices:

$$A_i := \prod i(A_i) \in \{0,1\}^{\frac{m}{w_c} \times n},$$

where $\prod i$ is the ith sequence, and $i = 2, 3, \cdots, w_c$.

Step 3: Construct the final LDPC matrix using all submatrices above:

$$\mathrm{H} := \begin{bmatrix} A_1^T & A_2^T & \cdots & A_{w_c}^T \end{bmatrix} \in \{0,1\}^{m \times n}.$$

Moreover, ECCPoW changes the sequence of permutations through the previous hash values and uses the previous hash value as the seed value to determine the sequence of permutations. The sequence of permutations is random because the hash values are random. The code is implemented in Reference [19]. Table 2 compares Matrix H produced using different hash values. The lower part of the generated Matrix H in Table 2 is different.

Table 2. Form of the resulting Matrix H using different hash values.

Generated Matrix H $n = 24$ $m = 16$ $w_c = 3$ $w_r = 4$	[1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1] [0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 1 0 0 0] [0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1] [0 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0] [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 1 0 0] [1 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 1 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0] [0 1 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0] [0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 1 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0 0 1] [0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0] [1 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0]
Previous hash value	0x00000000000000000000000000000001
Generated Matrix H $n = 24$ $m = 16$ $w_c = 3$ $w_r = 4$	[1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1] [1 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0] [0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0] [0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 1 0 1 0 0 0 0 0] [0 0 0 1 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 1 0 0 0 0 0 0 0 1] [0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0] [0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 1 0 0 0 0 0] [0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1] [0 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0] [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 1 0 0] [1 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 1 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0]
Previous hash value	0x00000000000000000000000000000002

3.3. Crypto Puzzle Decoder That Changes Every Block

The LDPC decoder of ECCPoW was developed using a message-passing algorithm. The decoder receives an $m \times n$ LDPC matrix of hash values $r \in \{0,1\}^n$ of length n as input values. The decoder outputs a

value $c \in \{0,1\}^n$ of length n. The decoder can produce two types of answers depending on the input hash value r. The decoder outputs a sign $D_{MP} : \{r, H\} \mapsto c_i$ if the entered hash value r satisfies $\|r - c_i\|_h \leq t$ for any sign c_i, where t is the value determined by the LDPC matrix. If not satisfied, the decoder outputs a random vector $c \in \{0,1\}^n$. The code is implemented in Reference [19].

The two conditions for determining whether to solve a cryptographic puzzle are listed in Table 3. Condition 1 determines that the cryptographic puzzle has been solved if the decoder's output value c satisfies the conditions. In Condition 1, the output value is the code. Condition 1 is possible because the output value is less likely to code when the value is any input to the decoder. For example, there is a low probability of finding an answer to any input in SHA 256. In Condition 2, the Hamming weight of the output value is an element of the given set S. Set S is the range value of the output value. Condition 2 occurs because the Hamming weights of the possible codes may differ when Matrix H is given.

Table 3. Conditions for the determination of crypto puzzle resolution.

Condition 1	(Original method) If the result of the decoder is code and has a specific Hamming weight, the problem is solved.
Condition 2	(Existing proof of work) If the result of rehashing the result of the decoder is less than a specific value, the problem is solved.

The satisfactory probability of Condition 1 requires a minimum Hamming distance value of H. To calculate this value, all distinct codes in 2^k must be considered, which is possible when the number of codes is small, but it is impossible when the number is large. Litsyn [20] reported the upper and lower bounds of the minimum Hamming distance value of H at specific w_c and w_r values. Table 4 reveals the probability of finding the codes according to the variables in the LDPC matrix. In this way, the upper bound of the probability is small. This makes it less likely for the decoder to meet Condition 1 when any value is entered.

Table 4. Probability of finding a code according to the variables of the low-density parity-check matrix.

$w_c = 4, w_r = 5$	p_1 Upper Bounds	p_1 Lower Bounds
$n = 80, k = 12$	6.32×10^{-5}	2.12×10^{-8}
$n = 120, k = 24$	1.65×10^{-8}	1.49×10^{-13}
$n = 160, k = 32$	4.06×10^{-10}	1.34×10^{-17}

Condition 2 is used to increase the difficulty of cryptographic puzzles when variables n, m, w_c, and w_r are fixed. Table 5 displays the probability that Condition 2 is satisfied when given a set S and part of the distribution of Hamming weights of the codes that can be generated when $n = 256$, $m = 192$, and $w_c, w_r = 5$ are given.

Table 5. The probability that Condition 2 is satisfied.

Hamming Weight	Probability	Element of the Set S	Probability that Condition 2 is Satisfied
98	$\approx 5 \times 10^{-5}$	98	$\approx 5 \times 10^{-5}$
...	...	...	...
128	$\approx 9.7 \times 10^{-2}$	98, 100, ... , 126	$\approx 4 \times 10^{-1}$
128	$\approx 1 \times 10^{-1}$	98, 100, ... , 126, 128	$\approx 5 \times 10^{-1}$

The probability of satisfying both Conditions 1 and 2 is as follows:

$$p := Pr\{c|\mathrm{Hc} = 0\} \times Pr\{\|c\|_h \in S\}.$$

We produced the difficulty table, as shown in Table 6, by calculating the probability of meeting Conditions 1 and 2 at the same time when the variables n, w_c, and w_r and the set S are given. The probability value p in Table 6 is the difficulty level of the cryptographic puzzle. The closer the probability value is to zero, the higher the difficulty of the cryptographic puzzle.

Table 6. Difficulty table for ECCPoW.

Lv.	n	w_c	w_r	Set S	p
1	32	3	4	{10, 12, . . . , 20, 22}	$\approx 3.07 \times 10^{-5}$
2	32	3	4	{10, 12, . . . , 14, 16}	$\approx 2.02 \times 10^{-5}$
			...		
379	128	3	4	{34, 94}	$\approx 5.12 \times 10^{-23}$
380	128	3	4	{34}	$\approx 2.60 \times 10^{-23}$

Condition 2 determines the resolution of the cryptographic puzzle by comparing the output value of the decoder with the result value and comparing the nonce with the target. Figure 1 illustrates Condition 2 for determining whether ECCPoW cryptographic puzzles are resolved. If the composite function and the hash algorithm are recognized as one hash function, Figure 1 has the same structure as Bitcoin, so the difficulty control function of Bitcoin can be used.

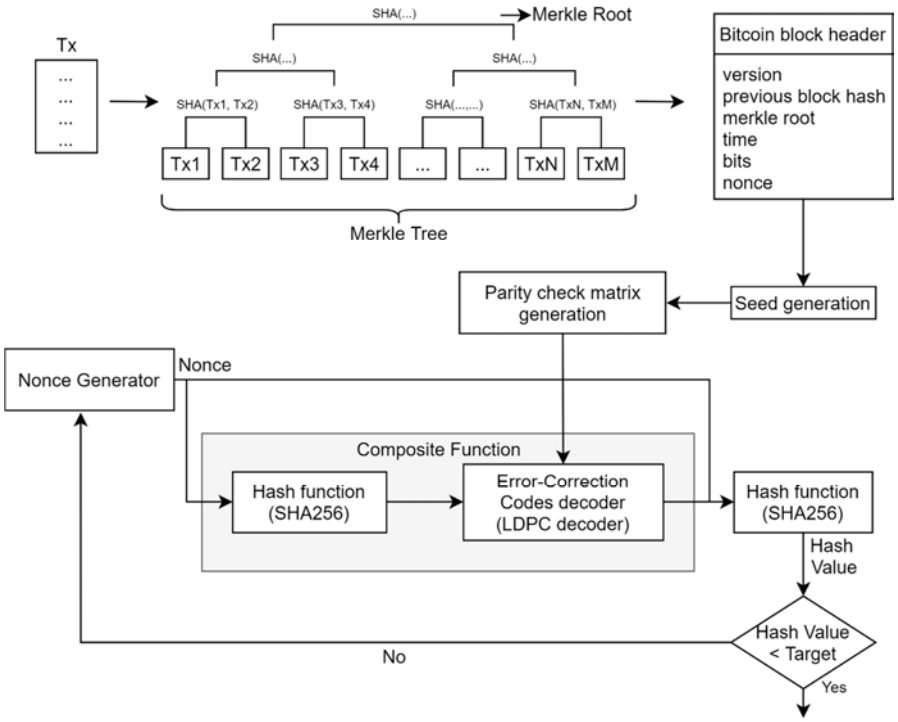


Figure 1. Condition 2 for determining the error-correction codes proof-of-work (ECCPoW) crypto puzzle resolution.

4. Experiment

This chapter presents experiments to verify ECCPoW. In the single-node experiment, the Bitcoin consensus algorithm was replaced by ECCPoW to verify the block generation function. In the multi-node experiment, we experimented with checking whether block generation, block synchronization, and transaction creation and transmission were performed correctly in a multi-node environment. In addition, the block generation time was tested in Bitcoin and Ethereum.

4.1. ECCPoW Operation Single/Multiple Node Experiment

We replaced Bitcoin's consensus algorithm with ECCPoW. The single-node experiment is a block generation experiment of the ECCPoW blockchain. Bitcoin uses the "generatetoaddress" command for block generation and receives the current address of the blockchain as a parameter. Then, a new block address is created using the "getnewaddress" command, and 10 blocks are created using "generatetoaddress." Figure 2 illustrates the change in 10 of the "blocks" after block generation.

Multiple node experiments use three random nodes (Nodes 1, 2, and 3) to experiment with block synchronization, to transmit transactions, and to verify them. This experiment demonstrates that each node mines different numbers of blocks, synchronizing to a long blockchain. Figure 3 indicates the results of mining and checking the blocks in Node 1.

Figure 4 reveals the results of mining and checking blocks on Node 2, and Figure 5 demonstrates the results of mining and checking blocks on Node 3.

Node 3 connects Node 1 (192.168.232.128) and Node 2 (192.168.232.129) using the "addnode" command. Then, Node 3 uses the "getadnednodeinfo" command to check the information on the connected nodes. Figure 6 illustrates the connection between Node 1 (192.168.232.128) and Node 2 (192.168.232.129) using the "addnode" command. Next, Figure 6 indicates the connected node information using the "getadnednodeinfo" command. Bitcoin's nodes synchronize to nodes that hold many blocks, and Bitcoin determines that nodes with many blocks are highly reliable. In the current experiment, Node 3 had the largest number of blocks. Nodes 1 and 2 synchronize to the blocks of Node 3, and these results are displayed in Figure 7.



Figure 2. Results of block generation.

```
getnewaddress
3FUFZ8ShBQpCGqqkZoFErG2aGo28QwaUDT
generatetoaddress 10 3FUFZ8ShBQpCGqqkZoFErG2aGo28QwaUDT
[
  "2b2d7485a399d93169ff8532de059db263e0ba42db73847489d4620a7b7510bf",
  "099640a372ec928e5af6aa1b2032014b24a04225724c5553c1a80bf53dd0d8c4",
  "1c44c7078afcdd49eb9ab8fd4ccbc64419cca96afda91c571a51de163b0ac9ce",
  "630b1b3127c74b8892ff7fcbae89f2733aeadf57f225354a05f60583200ac8a2",
  "74449f452fd35732dbfab5e3ada8f531d7df4ee1a7df23ef41ab91ca1b4c3ca3",
  "282768f094a38a6f7113e4ffc45f5cb93ae7fb01f9e164bd8482f59df49712ff",
  "4a16965fc7efab971b99cbd8cd807eb78b17852fe4be55a620cf00ca865bbafc",
  "5ea006ef89893ab591ba39d995c2b09bcba5ac21813a3c8ff9d7752765c820e7",
  "c43c6bb01c954f23b998f88ead512a7b7cbc1350e37225e6659f02f923ea8f17",
  "7a39ec86ae15a01d8d6930f2fcc5300d8830726fab16b4e733c59db1719f8e43"
]

getblockchaininfo
{
  "chain": "main",
  "blocks": 10,
  "headers": 10,
  "bestblockhash":
"7a39ec86ae15a01d8d6930f2fcc5300d8830726fab16b4e733c59db1719f8e43",
  "difficulty": 4.5230594683691 96e+73,
  "mediantime": 1569822743,
```

Figure 3. Block mining and verification in Node 1.

```
getnewaddress
34rqCn3Xf6PDtLoHLZ8wPp1wsRV6Gb9KFY
generatetoaddress 20 34rqCn3Xf6PDtLoHLZ8wPp1wsRV6Gb9KFY
[
  "404a6788ab0dcb5bd356b03b400c005fa8f49380ab978d4a7b3537dfc1f6ba23",
  "ca0404483d852a32abb8609073bbebab0d06c55c6fc5b9e141de211d570f5576",
  "129b0e34f22c2f745efcd708df8b653ac295ac7b08b9bf36de518d88e0cfaa9a",
  "807f58aa7d221d15c6acc9e3b0ea89bc96de2205c797547d72d10ea1d463a588",
  "d9265ae393201c60e988ccbeada6e82038e90997deaad64e5057d9c523d0abc3",
  "b518305ba918e8619c36c4d449e161dbe8ef9517bc2d2fd2cfaea92af4171008",
  "8d13e15542fda1e5950268f6d87b87a5ac737e4377231d06b79f683255d73869",
  "22c3b9e63d3b7f6fbc8959805fef052aea6fb136124fe5799097d61c6b29b4c4",
  "5658489c42ca5aecf600e324a8a07e4c1a38b5f0ec8c6367f4c6941f031fd496",
  "43cd9632b9879fa290c3f6c2e63818189d84e752791233ced6bc58e29977450e",
  "2c65177825619bbebc75caccc0afccb2e58fa69eef178cb472eff29c2eda6ca9",
  "67e39e0482caafc683b8e57c6961fd7a4a3f376117c4349acd5ba1a84574235a",

getblockchaininfo
{
  "chain": "main",
  "blocks": 20,
  "headers": 20,
  "bestblockhash":
"9e99fb700b43d5a59fa97a6d1550e536c3933c0899158d9300ea74ac50928279",
  "difficulty": 4.5230594683691 96e+73,
  "mediantime": 1569822820,
```

Figure 4. Block mining and verification in Node 2.

```
getnewaddress
3765hBQFFHMn8FProyKFuF4vHKayafka4d
generatetoaddress 30 3765hBQFFHMn8FProyKFuF4vHKayafka4d
[
  "28bbd5b26fd5b9530614279cf06a4e036056aa1b996d05d699f73d71b77311c9",
  "11902f08b5c19f9ce6551730311f076453ef620014f703643c9225d1279bf5ea",
  "0610136713ae049268c72dec6d32cd13346f12b9bcc20b1ae19d8d9492eeb955",
  "6848477bf878c2e38ff13d3aedcc047a8f68659da8454e1800ac963933c688c2",
  "1d20dd47c8400a5833fc84b4fe9d91d4f6dfc474fbaf939f45745121669d740c",
  "8f084bc0ea45ca54b81f0bc0d931591572209cf69210ccea9daf16edfc18ae56",
  "f8a52172a10c14283e9749c62aa9d4362d677f32d586b71e0ffced4ddf4f111f",
  "a0b47bde1bf2a11a35b0672e2a0be022eb94cd4dbb063bd35de44bc038cf0c43",
  "56221e5d9bc496a2cf2a61bc13efef967b87e74273e7dc6fbcd3da4f7486a532",
  "a93231e200e63364d56b5669144f2353ac22b018c1835220ada57ea471d68f58",
  "0f4913a21d6481cf7f476952dd6e5b7fc8c188c420569fd4365f5f12a766ab3c",
  "06d84fa3aa0b97b32194a5ff9533847250261d567e33c7b2b199cee0781f3cda",

getblockchaininfo
{
  "chain": "main",
  "blocks": 30,
  "headers": 30,
  "bestblockhash":
"0aaa7eb4a5bbd818e5a76383fd1f014df645ededd0ffb9568ee262b5080d7b5e",
  "difficulty": 4.5230594683691 96e+73,
  "mediantime": 1569822888,
```

Figure 5. Block mining and verification in Node 3.

```
addnode 192.168.232.128 add
null
addnode 192.168.232.129 add
null
```

```
getaddednodeinfo
[
  {
    "addednode": "192.168.232.128",
    "connected": true,
    "addresses": [
      {
        "address": "192.168.232.128:9777",
        "connected": "outbound"
      }
    ]
  },
  {
    "addednode": "192.168.232.129",
    "connected": true,
    "addresses": [
      {
        "address": "192.168.232.129:9777",
        "connected": "outbound"
      }
    ]
  }
]
```

Figure 6. Node connections in Node 3.

```
getblockchaininfo
{
  "chain": "main",
  "blocks": 30,
  "headers": 30,
  "bestblockhash":
"0aaa7eb4a5bbd818e5a76383fd1f014df645ededd0ffb9568ee262b5080d7b5e",
  "difficulty": 4.523059468369196e+73,
  "mediantime": 1569822888,
```

Figure 7. Blockchain synchronization of Nodes 1 and 2.

We checked the behavior of the blockchain by checking the synchronization. We sent the transactions between linked Nodes 2 and 3. Figure 8 indicates the balances held by Nodes 2 and 3. Currently, Node 2 has zero, and Node 3 has 1000.

```
getbalance
0.00000000

getbalance
1000.00000000
```

Figure 8. Balances for Node 2 (**top**) and Node 3 (**bottom**).

Figure 9 displays the address of Node 2 (in the Pay-To field) and the amount of coin to send (in the Amount field) for Node 3 to transmit 500 coins to Node 2. A transaction fee was set to the minimum cost of 0.00001. Figure 10 reveals the transaction transfer by identifying the amount of coin in Node 2 and displaying the recent transaction records. Figure 11 illustrates that the number of coins in Node 3 and the recent transaction record decreased.

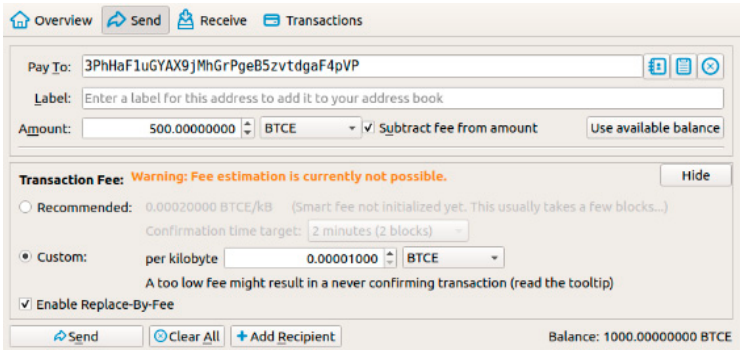


Figure 9. The input of the transaction transfer.

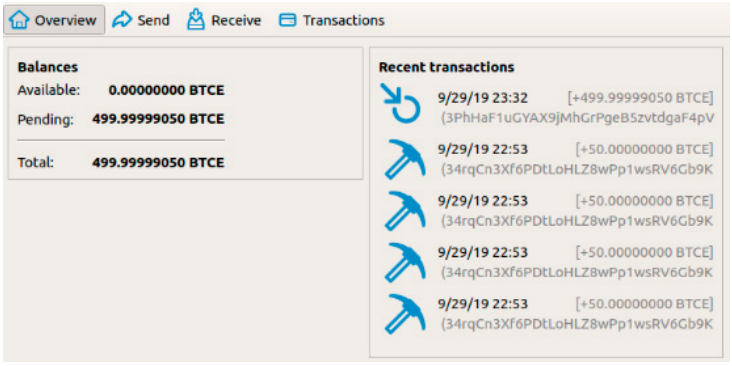


Figure 10. Balances and transaction logs for Node 2.

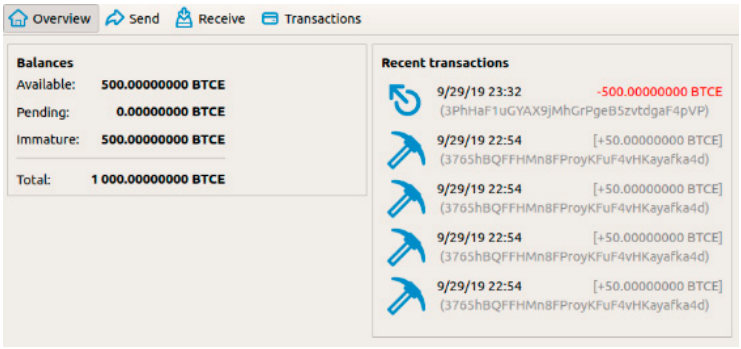


Figure 11. Balances and transaction logs for Node 3.

4.2. Block Generation Time

Figure 12 reveals the time-by-time block generation of Bitcoin with ECCPoW, using the difficulty table proposed in Section 3. Block generation time should be able to meet the target generation time for a stable blockchain. The blockchain adjusts the target block generation time by adjusting the difficulty level over time. The experimental environment used block generation with one node. As shown in Figure 12, the block generation time is unstable in the experiment. Sufficient nodes can confirm the stability of the block generation time. The difficulty table must be finely adjusted.

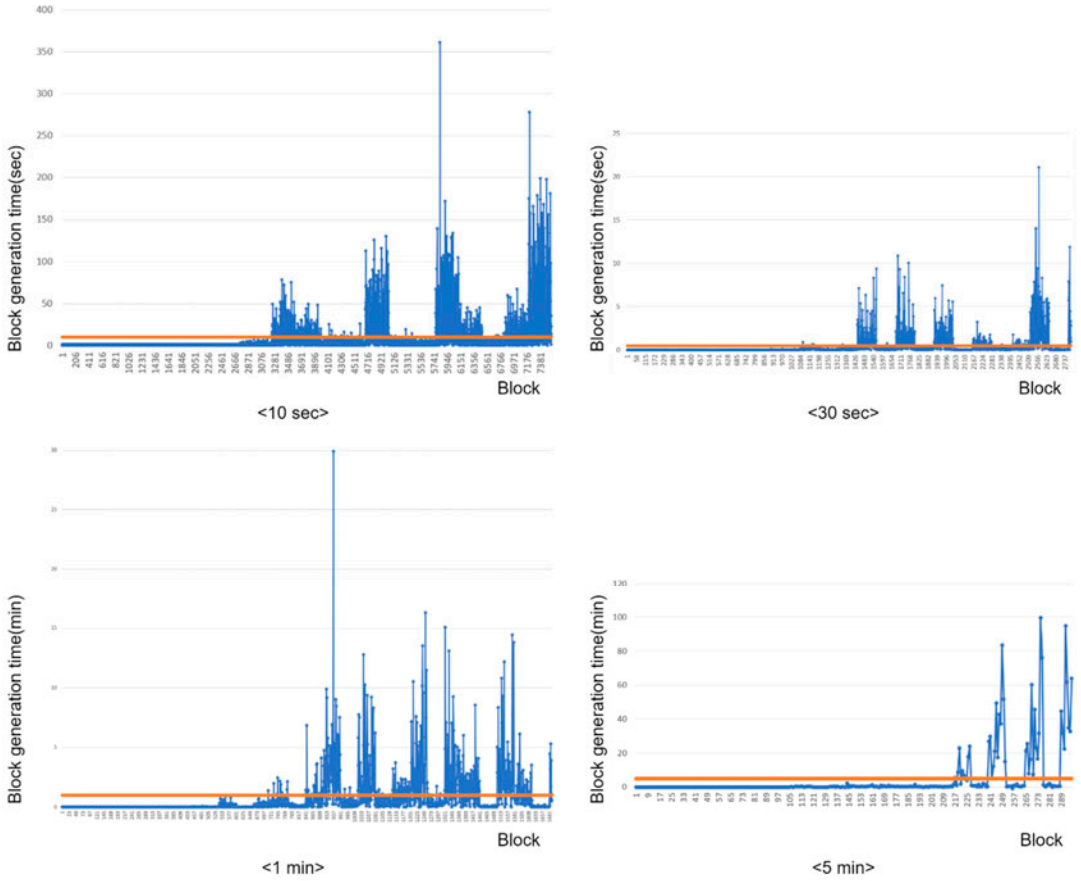


Figure 12. Block generation time of Bitcoin with ECCPoW.

5. Evaluation

We evaluated the mining centralization, security, and scalability of ECCPoW using Amazon Web Services. Table 7 presents the evaluation environment. A monitoring computer checked the network test results. The seed instance made the blockchain network connections between nodes. The mining instance was responsible for block mining. The implementation environment for each node was a Ubuntu Server 18.04 LTS, m5.large (vCPU processor 2 core with 8 GB of RAM, and a 20 GB solid-state drive).

Table 7. Implementation environment of the evaluation.

No	Role	CPU	Memory (GB)	HDD/SSD (GB)	Volume
1	Monitoring PC	Intel i7-8700 3.20 GHz	16	SSD 256	1
2	Seed Instance	AWS m5.xlarge vCPU 2	8	HDD 20	6
3	Mining Instance	AWS m5.xlarge vCPU 2	8	HDD 20	40

The blockchain trilemma problem is that trade-off relationships occur in the evaluation characteristics of the blockchain. The trilemma elements of a blockchain are decentralization, scalability, and security. Thus, we compared Bitcoin in terms of these elements (mining centralization instead of decentralization). Table 8 displays the evaluation standards and goals for evaluation.

Table 8. Description of evaluation features.

Evaluation Features	Unit of Measurement	Evaluation Standards	Evaluation Goal
Mining centralization	Distribution of mining success rate	40% (Estimate) (Bitcoin, Oct.–Dec. 2018)	40%
Security	Security of Bitcoin contrast	100% (Bitcoin)	100%
Scalability	Scalability of Bitcoin contrast	100% (Bitcoin)	100%

Bitcoin has a total hash rate of approximately 90 TH/s (March 2020). It is impossible to compare the current version of Bitcoin with ECCPoW. Moreover, ECCPoW (ver. 0.1.2) replaced the Bitcoin 0.17 version [21] of SHA 256 [22]. We compared two completely initialized blockchains (block count zero).

We compared the Bitcoin and ECCPoW initialized to set the evaluation environment (difficulty change cycle, target block generation time, 23 instances, among others). The blockchain typically sets a difficulty change cycle of 60 min and a target block creation time of 3 min for experimentation.

5.1. Mining the Centralization Evaluation

Mining centralization is when a network is no longer centralized and operates autonomously within a blockchain. We define mining centralization as when the nodes are fairly mined with the same performance. The indicators used in the mining centralization evaluation include the distribution of the mining success rates, which are defined by the formula below. According to the formula, the lower the dispersion of the probability of mining success, the higher the distribution of mining success. In other words, the distribution of mining success rates is higher for each participating node to have a uniform distribution of mining success rates, which means better dispersion. Bitcoin is estimated to have a 40% distribution of mining success rates (October to December 2018). Moreover, ECCPoW targets 40% dispersion:

$$A = \frac{C}{\sqrt{B + C^2}} \times 100 A = \text{The distribution of mining success rate } (\%),$$

$$B = \text{The dispersion of mining success rate,}$$

$$\text{C} = \text{Average of the mining success rate.}$$

We created three seed nodes for the ECCPoW blockchain and composed 20 mining nodes. Table 9 lists the number of mining successes of each mining node at the time of 100 blocks being mined. Table 10 indicates the distribution of mining success.

The dispersion of ECCPoW was measured by the distribution of the mining success rate and measured at 92.22%, which was 32% higher than the assessment target of 60% (Table 10). The results of the experiment revealed that the ECCPoW miner is more likely to be rewarded than a Bitcoin miner in an ideal environment where participating nodes exhibit the same performance. In other words, ECCPoW is stronger in mining centralization than Bitcoin.

Table 9. The number of mining success and number of nodes.

Number of Mining Nodes	Number of Mining Success	Number of Mining Nodes	Number of Mining Success
1	4	11	7
2	9	12	4
3	3	13	12
4	4	14	5
5	6	15	11
6	4	16	2
7	6	17	7
8	6	18	6
9	5	19	10
10	5	20	8

Table 10. Evaluation results of the mining centralization evaluation features.

Total Number of Mining Successes	Average Mining Successes	Square of the Average Number	Dispersion of the Average Number	Distribution of Mining Successes
124	6.2	38.44	6.76	92.21944

5.2. Security Evaluation

Security is associated with the difficulty of the blockchain being altered by a miner's attack. A typical attack on the blockchain is the double-payment issue. One of the causes of double-payment problems occurs in the branch of the blockchain. An orphan chain is a chain that has a branch other than the main chain, and an orphan block is a block belonging to an orphan chain. We assess the security as the ratio of orphaned blocks to the total number of blocks in the blockchain. The security assessment calculates the orphan block ratio of Bitcoin and ECCPoW using the expression below. We evaluated the security of ECCPoW based on the security of Bitcoin at 100%.

$$Orphan\ block\ ratio = \frac{number\ of\ orphan\ blocks}{Total\ Height\ of\ Blockchain} \times 100$$

We configured the blockchain test environment with three seed nodes and 20 mining nodes. In addition, we mined 100 blocks in the blockchain and checked the orphan blocks. Figure 13 depicts the orphan block identification for the security assessment. After mining 40 blocks, the evaluation identified blocks for stable synchronization of the blockchain. In the experiment, orphan blocks of a height of one frequently occurred in both environments. In our experiment, we classified the height of subchains of two or higher as orphan blocks. In Figure 13, orphan blocks correspond to Blocks 13, 14, 15, 16, and 18.

Figure 14 demonstrates the blockchain status based on the results of the security assessment in the Bitcoin environment.

Figure 15 reveals the blockchain status according to the results of the security assessment in the ECCPoW environment.

All blockchains were measured at 0% because of the assessment. The security of ECCPoW was the same on a Bitcoin and orphan block basis. Blockchain measurements inhibited the generation of orphan blocks.

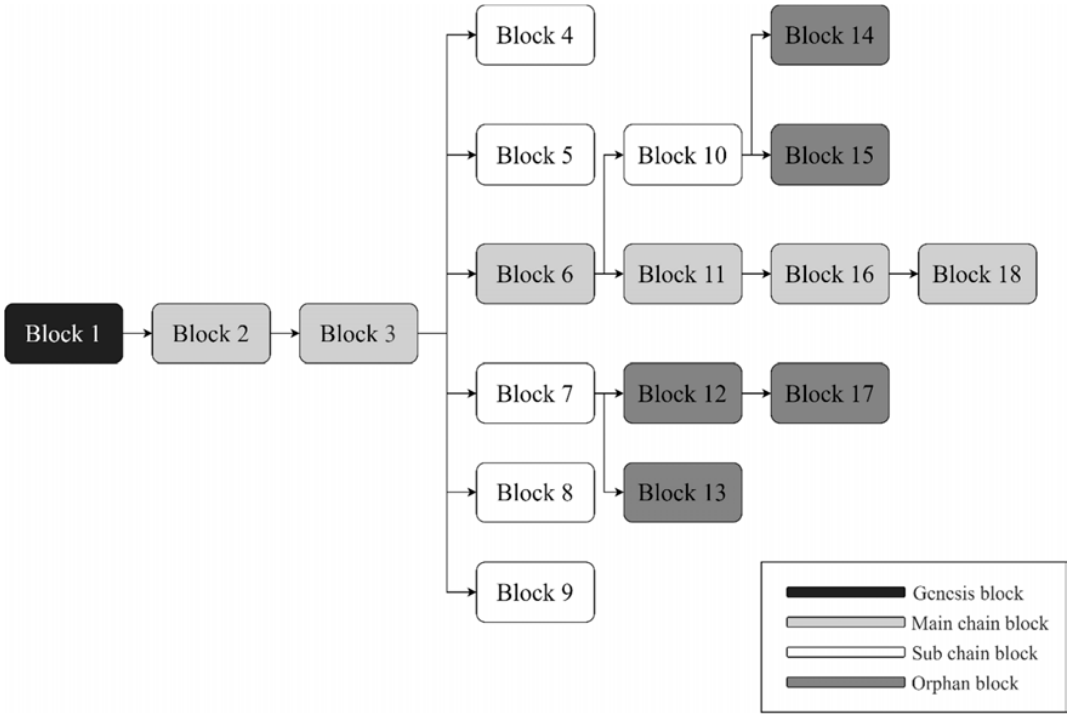


Figure 13. Diagram description of security evaluation.

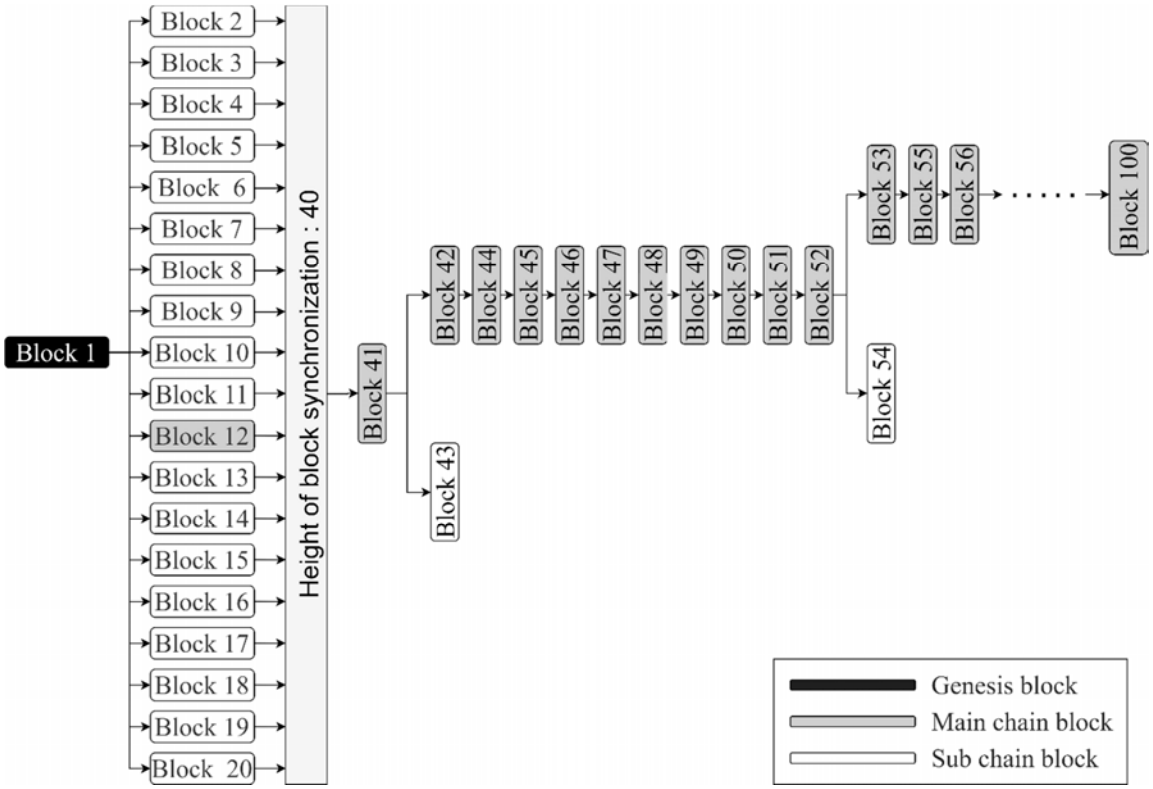


Figure 14. Evaluation diagram of Bitcoin security.

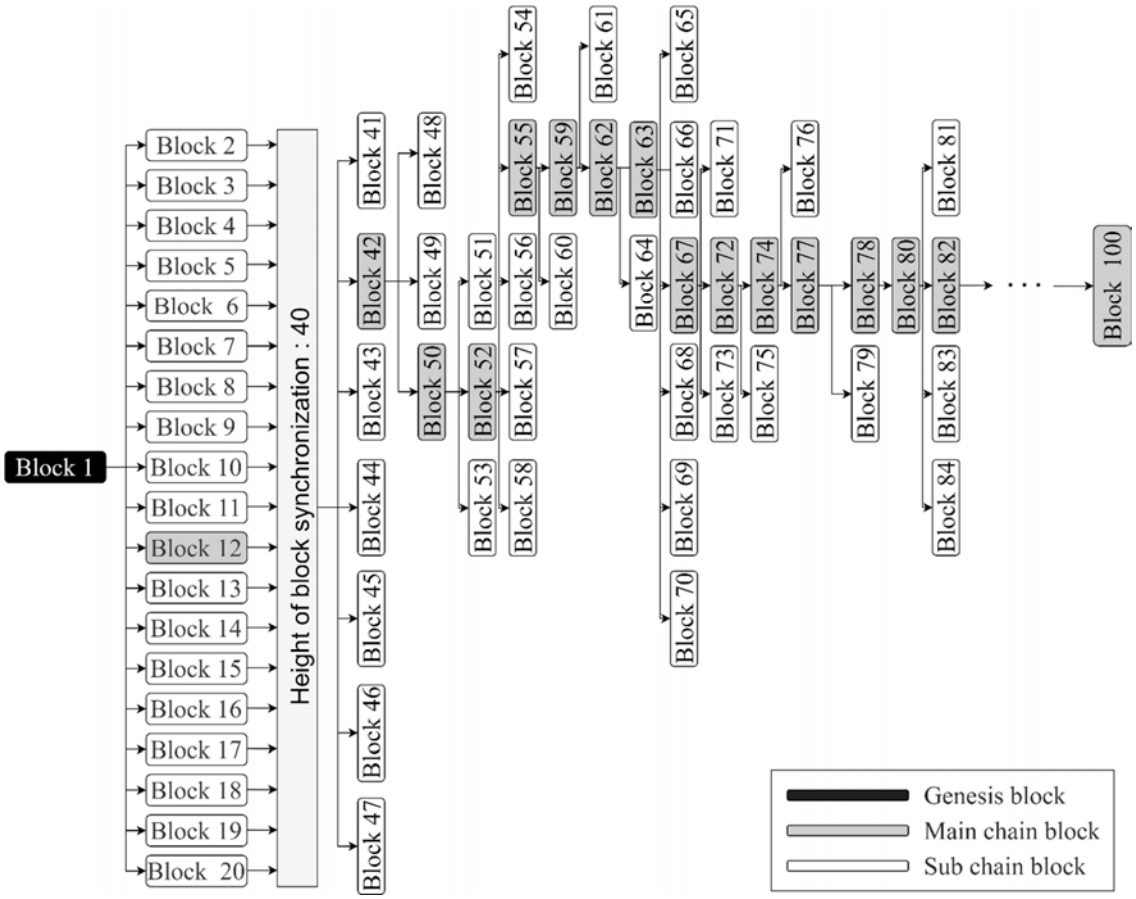


Figure 15. Evaluation diagram of ECCPoW security.

5.3. Scalability Evaluation

Scalability is the extent to which a system can flexibly respond to an increase in the number of users. Blockchain scalability is related to the transaction processing speed of blockchain. Transactions per second (TPS) is the transaction processing speed of the blockchain. Thus, we evaluated the scalability of the blockchain using TPS. The following equation defines how to obtain TPS using transactions in blocks.

$$\text{TPS} = \frac{Total\ number\ of\ transactions\ in\ block}{Generation\ time\ of\ total\ block}.$$

We configured the blockchain test environment with three seed nodes and 20 mining nodes. Transaction generation occurs continuously from height 91 to 100—the blockchain mines 100 blocks. We checked the number of transactions in the height blocks of the blockchain from 91 to 100. We calculated the TPS of the blockchain. Table 11 lists the generation time and number of transactions of blocks of height 91 to 100 in Bitcoin.

Table 12 displays the generation time and the number of transactions of blocks of height 91 to 100, in ECCPoW. The evaluation results revealed that Bitcoin's TPS is 0.95, and ECCPoW's average TPS is 0.94. Moreover, ECCPoW's scalability was measured at 98.95%, which was down 1.02% from Bitcoin's scalability. In addition, ECCPoW added a process to Bitcoin to resist ASICs. However, the scalability values of ECCPoW and Bitcoin were assessed at a similar level.

Table 11. The evaluation result of Bitcoin scalability.

Number	Height	Block Generation Time (Seconds)	Number of Transactions
1	91	719	672
2	92	19	22
3	93	94	88
4	94	144	141
5	95	275	263
6	96	40	40
7	97	313	296
8	98	6	11
9	99	574	534
10	100	146	137
Total		2330	2204
TPS		0.945922747 TPS	

Table 12. Evaluation results of ECCPoW scalability.

Number	Height	Block Generation Time (Seconds)	Number of Transactions
1	91	151	140
2	92	192	182
3	93	369	351
4	94	361	331
5	95	156	151
6	96	214	205
7	97	124	120
8	98	267	250
9	99	585	548
10	100	514	480
Total		2933	2758
TPS		0.940334129 TPS	

5.4. Comparison to Related Work

This section compares the features with the existing studies on ASIC resistance. Table 13 compares the proposed method and existing relevant research on ASIC resistance. The ASIC resistance study for the PoW blockchain has three approaches. The first method causes bottlenecks using memory access in a hash function—for example, Ethereum (Ethash) and Bytecoin. The second method prevents the generation of ASICs using a hash function overlay, for example, the X-11 series. Finally, the third method uses periodic hash function replacement for ASICs, for example, in Monero (2019.9, RandomX algorithm) and Ethereum (2020.7 applying scheduled, ProgPoW).

Table 13. Comparison results of the related work.

Comparison Features	Proposed Method	Memory Approach	Hash Function Overlay	Periodic Hash Function Replacement
Applied cryptocurrency	-	Ethereum, Bytecoin	X-11 series	Monero (2019.9, RandomX), Ethereum (2020.7 applying scheduled, ProgPoW)
Characteristic	Change hash function every block	Force memory access	Overlapping multiple hash functions	Hard fork manually or automatically with a hash function
Algorithm	ECCPoW	Ethash, ProgPoW, CryptoNight	Blake, Bmw, Groetl, etc.	RandomX, ProgPoW
ASIC appearance	-	Yes	Yes	Unknown
ASIC resistance induction method	Use of ASIC resistance by connecting the LDPC decoder and hash function	Induce cache miss when creating blocks (using a DAG, etc.)	Overlapping the difficulty of known hash functions	Hard fork every six months (formerly Monero). Convert the hash function for a period using the key-block concept (currently Monero)

The resistance induction of each ASIC in the existing studies is as follows. The memory approach induces a cache miss (using DAG, among others) when creating blocks. This method degrades ASIC performance as memory access increases. The hash function overlay method uses the overlap of the difficulty of known hash functions. This method relies on the security of the applied hash function. The periodic hash function replacement method periodically changes the hash function manually or automatically.

6. Conclusions

In this paper, we explained ECCPoW and applied the proposed method to Bitcoin. This paper addressed the problem of centralization of mining due to the emergence of ASICs. We proposed a PoW concept based on error-correction codes to solve this problem. The core of the ECCPoW is the connection of the hash function with the LDPC decoder. Blockchain applied with the ECCPoW determines the completion of the POW using the output value of the decoder. In addition, ASIC suppression is possible because ASICs use the LDPC decoder.

This paper contributes to the phenomenon of symmetry in the PoW blockchain. The total block size of the PoW blockchain symmetrically influences the hash rate. The PoW blockchain increases stability as the hash rate increases in size. However, a high hash rate causes the waste of computing power in block generation. Thus, we mitigate the causes of a high hash rate using ASIC resistance.

The ECCPoW offers a method of solving different puzzles in each block to avoid ASICs. This maximizes the benefits of how existing studies use a limited number of hash functions to solve each block of different hash functions. The proposed method offers more effective connections and the use of multiple hash functions. We presented the difficulty control, parity-check matrix generation method, hash vector generation, and code determination methods for implementing ECCPoW. Furthermore, ECCPoW was applied to Bitcoin to verify the proposed method. We assessed the mining centralization, security, and scalability of ECCPoW and Bitcoin. We found that ECCPoW maintained security and scalability, showing 32% higher mining centralization than Bitcoin. Using the proposed method, ECCPoW does not require high hash rates, and miners can compete more fairly.

This study contributes to the previous literature on ASIC resistance in the PoW blockchain. The ASIC resistance study of the PoW blockchain was conducted in response to hardware development by ASIC manufacturers. One of the studies on ASIC resistance induced forced access to memory in the hash function, thereby lowering the performance of ASICs. Other studies have suggested periodically altering the hash function to render ASICs useless. The ECCPoW provides a method of releasing different hash functions for each block, rather than changing the periodic hash function.

The ECCPoW revealed the limit of block generation time in the experiment. For the stable operation of the blockchain, the block generation time of the ECCPoW must be stable and controllable. The results of the block generation time test were unstable in the experiment using one node. Thus, ECCPoW requires further research regarding the difficulty of block generation for stable operation. We also need to increase the number of nodes in the ECCPoW to carry out mining.

Author Contributions: H.J. and H.-N.L. contributed equally to this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean government (MSIP) (NRF-2018R1A2A1A19018665).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nakamoto, S. *Bitcoin: A Peer-to-Peer Electronic Cash System*; 2008; Available online: https://bitcoin.org/bitcoin.pdf (accessed on 9 June 2020).
2. Nofer, M.; Gomber, P.; Hinz, O.; Schiereck, D. Blockchain. *Bus. Inf. Syst. Eng.* **2017**, *59*, 183–187. [CrossRef]
3. Saberi, S.; Kouhizadeh, M.; Sarkis, J.; Shen, L. Blockchain Technology and Its Relationships to Sustainable Supply Chain Management. *Int. J. Prod. Res.* **2019**, *57*, 2117–2135. [CrossRef]
4. Qin, R.; Yuan, Y.; Wang, F.-Y. Research on the Selection Strategies of Blockchain Mining Pools. *IEEE Trans. Comput. Soc. Syst.* **2018**, *5*, 748–757. [CrossRef]
5. Liu, X.; Wang, W.; Niyato, D.; Zhao, N.; Wang, P. Evolutionary Game for Mining Pool Selection in Blockchain Networks. *IEEE Wirel. Commun. Lett.* **2018**, *7*, 760–763. [CrossRef]
6. Karame, G.O.; Androulaki, E.; Capkun, S. Double-Spending Fast Payments in Bitcoin. In Proceedings of the 2012 ACM Conference on Computer and Communications Security, Raleigh, NC, USA, 16–18 October 2012; pp. 906–917. [CrossRef]
7. Jang, J.; Lee, H.-N. Profitable Double-Spending Attacks. *arXiv* **2019**. Available online: https://arxiv.org/abs/1903.01711 (accessed on 9 June 2020).
8. Antminer S19 Pro. Available online: https://support.bitmain.com/hc/en-us/articles/900000261726-S19-Pro-Specifications (accessed on 9 June 2020).
9. Park, S.; Kim, H.; Lee, H.N. Introduction to Error-Correction Codes Proof-of-WorK. *Mag. IEIE* **2019**, *5*, 26–32.
10. Buterin, V. A Next-Generation Smart Contract and Decentralized Application Platform. *White Pap.* Available online: https://ethereum.org/whitepaper/ (accessed on 9 June 2020).
11. Wood, G. Ethereum: A Secure Decentralised Generalised Sransaction Byzantium Version. *Yellow Pap.* **2014**. Available online: https://ethereum.github.io/yellowpaper/paper.pdf (accessed on 9 June 2020).
12. Zhou, Q.; Huang, H.; Zheng, Z.; Bian, J. Solutions to Scalability of Blockchain: A Survey. *IEEE Access* **2020**, *8*, 16440–16455. [CrossRef]
13. Greg, C.; Andrea, L.; Michael, C.; IfDefElse. ProgPoW, a Programmatic Proof-of-Work. Ethereum—EIPs, No. 1057. 2018. Available online: https://eips.ethereum.org/EIPS/eip-1057 (accessed on 9 June 2020).
14. Duffield, E.; Diaz, D. Dash: A PrivacyCentric Crypto-Currency. *White Pap.* **2014**. Available online: https://docs.dash.org/en/stable/introduction/about.html#whitepaper (accessed on 9 June 2020).
15. Saberhagen, N.V. Cryptonote v2.0. *White Pap.* **2013**. Available online: https://cryptonote.org/whitepaper.pdf (accessed on 9 June 2020).
16. RandomX. Available online: https://github.com/tevador/RandomX (accessed on 9 June 2020).
17. Shao, S.; Hailes, P.; Wang, T.-Y.; Wu, J.-Y.; Maunder, R.G.; Al-Hashimi, B.M.; Hanzo, L. Survey of Turbo, LDPC, and Polar Decoder ASIC Implementations. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 2309–2333. [CrossRef]
18. Gallager, R. Low-Density Parity-Check Codes. *IRE Trans. Inf. Theory* **1962**, *8*, 21–28. [CrossRef]
19. Bitcoin ECC LDPC. Available online: https://github.com/cryptoecc/bitcoin_ECC/blob/ecc-0.1/src/ldpc/LDPC.cpp (accessed on 9 June 2020).

20. Ben-Haim, Y.; Litsyn, S. Upper Bounds on the Rate of LDPC Codes as a Function of Minimum Distance. *IEEE Trans. Inf. Theory* **2006**, *52*, 2092–2100. [CrossRef]
21. Bitcoin Core 0.17 Version. Available online: https://github.com/bitcoin/bitcoin/tree/0.17 (accessed on 9 June 2020).
22. Bitcoin ECC 0.1.2 Version. Available online: https://github.com/cryptoecc/bitcoin_ECC/tree/ecc-0.1.2 (accessed on 9 June 2020).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 25, 2021, accepted September 7, 2021, date of publication September 16, 2021, date of current version October 11, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3113522

Error-Correction Code Proof-of-Work on Ethereum

HYOUNGSUNG KIM[1], JEHYUK JANG[1], SANGJUN PARK[2], AND HEUNG-NO LEE[1], (Senior Member, IEEE)

[1]School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea
[2]Electronics and Telecommunications Research Institute (ETRI), Gwangju 500-712, South Korea

Corresponding author: Heung-No Lee (heungno@gist.ac.kr)

This research was supported in part by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2021-0-01835) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation), in the part by the IITP Grant through Korean Government MSIT under Grant 2020-0-00958, and in part by the National Research Foundation of Korea (NRF) Grant through Korean Government MSIP under Grant NRF-2021R1A2B5B03002118.

ABSTRACT The error-correction code proof-of-work (ECCPoW) algorithm is based on a low-density parity-check (LDPC) code. ECCPoW can impede the advent of mining application-specific integrated circuits (ASICs) with its time-varying puzzle generation capability. Previous research studies on ECCPoW algorithm have presented its theory and implementation on Bitcoin. In this study, we have not only designed ECCPoW for Ethereum, called ETH-ECC, but have also implemented, simulated, and validated it. In the implementation, we have explained how ECCPoW algorithm has been integrated into Ethereum 1.0 as a new consensus algorithm. Furthermore, we have devised and implemented a new method for controlling the difficulty level in ETH-ECC. In the simulation, we have tested the performance of ETH-ECC using a large number of node tests and demonstrated that the ECCPoW Ethereum works well with automatic difficulty-level change capability in real-world experimental settings. In addition, we discuss how stable the block generation time (BGT) of ETH-ECC is. Specifically, one key issue we intend to investigate is the finiteness of the mean of ETH-ECC BGT. Owing to a time-varying cryptographic puzzle generation system in ECCPoW algorithm, BGT in the algorithm may lead to a long-tailed distribution. Thus, simulation tests have been performed to determine whether BGT distribution is not heavy-tailed and has a finite mean. If the distribution is heavy-tailed, stable transaction confirmation cannot be guaranteed. In the validation, we have presented statistical analysis results based on the two-sample Anderson–Darling test and discussed how the BGT distribution follows an exponential distribution which has a finite mean. Our implementation is available for download at https://github.com/cryptoecc/ETH-ECC.

INDEX TERMS Anderson–Darling test, ASIC-resistant, blockchain, error-correction codes, Ethereum, hypothesis test, LDPC, proof-of-work, simulation, statistical analysis.

I. INTRODUCTION

Blockchain is a peer-to-peer (P2P) network that consists of trustless nodes. In a reliable P2P network, no peers (nodes) would intentionally send wrong information to others. In contrast, in an unreliable P2P network (e.g., a group of trustless nodes), the possibility that some peers may send false information to others should be considered. For example, a node may spread wrong or fake information to others. To address these issues in an unreliable P2P network, Nakamoto proposed using blocks and chaining these blocks with a novel consensus algorithm [1].

In a blockchain, a peer sends a new block containing transactions to other peers. These peers validate the received block and link it to the previous block when there is no problem in the received block, i.e., when the authenticity of the block has been verified. A consensus algorithm is used to accomplish this verification task. If a peer has sent false information to others, such information is detected by the consensus algorithm as there is no collusion among the peers. A generated block contains information about previous blocks, i.e., all blocks are chained; thus, if someone wants to change one block in a chain, all previous blocks of the block to be changed must also be changed. Therefore, unless the network is centralized within a particular group, sending fake information about previous blocks to new peers is impossible.

The associate editor coordinating the review of this manuscript and approving it for publication was Giambattista Gruosso.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

Therefore, to prevent collusion, an unreliable network should avoid centralization.

Nakamoto proposed a proof-of-work (PoW) system for a consensus algorithm. In the PoW system, peers repeat a type of work to solve a cryptographic puzzle using a hash function (e.g., SHA256 [1] and Keccak [2]). When a peer successfully solves a cryptographic puzzle, the peer generates a block. In addition, the peer gets an incentive as a reward for the work done. In an ideal PoW system, new nodes can join to work and receive as much reward as they complete work. However, with an increase in the price of reward, attempts have been made to centralize the network to monopolize incentives.

Centralization is a phenomenon that occurs in PoW-based blockchain networks. In blockchains using PoW as a consensus algorithm, an oligarchy of miners with a disproportionate share of computation resources can monopolize block generation. Such centralization negatively impacts the credibility of a blockchain. For example, in a centralized network, a group of dominant nodes can selectively filter out some transactions belonging to others for their benefit. New nodes will find it difficult to earn trust and join the network in the fear of possible unfair treatment [3], [4].

The emergence of application-specific integrated circuits (ASICs) has accelerated the centralization of PoW. As more nodes use ASICs in generating blocks, the computation complexity in block generation increases. Thus, it has become difficult to generate blocks using general-purpose units, such as a central processing unit (CPU) and a graphics processing unit (GPU). As a result, a few groups equipped with powerful ASICs have surfaced and centralized the blockchain networks. To avoid centralization, researchers have proposed the use of ASIC-resistant PoW (e.g., Ethash of [2], X11 of [12], and Random X of [24]) and alternative consensus algorithms (e.g., proof-of-stake, delegated proof-of-stake, and Byzantium fault tolerance [25]). Networks using alternative algorithms have presented lesser decentralization effects than those have using ASIC-resistant PoW [25]. Specifically, in networks using alternative algorithms, only limited participants can generate blocks, but ASIC-resistant PoW has no limit on the number of participants. Thus, ASIC-resistant PoW presents a more decentralized network than do alternative algorithms.

For an ASIC-resistant PoW, an error-correction code proof-of-work (ECCPoW) algorithm was proposed [6], [7]. In ECCPoW algorithms, a hash value of a previous block generates a varying parity-check matrix (PCM) for error correction. This varying PCM works as a cryptographic puzzle in ECCPoW. These time-varying cryptographic puzzles make ECCPoW ASIC resistant. It is possible to use an ASIC for a specific fixed cryptographic puzzle. In ECCPoW, every newly created puzzle differs from all the previously created puzzles. As a result, if there was an ASIC for ECCPoW, such an ASIC must cover a wide range of cryptographic puzzle generation systems. Such a system, however, would incur huge chip space and cost [10], [11].

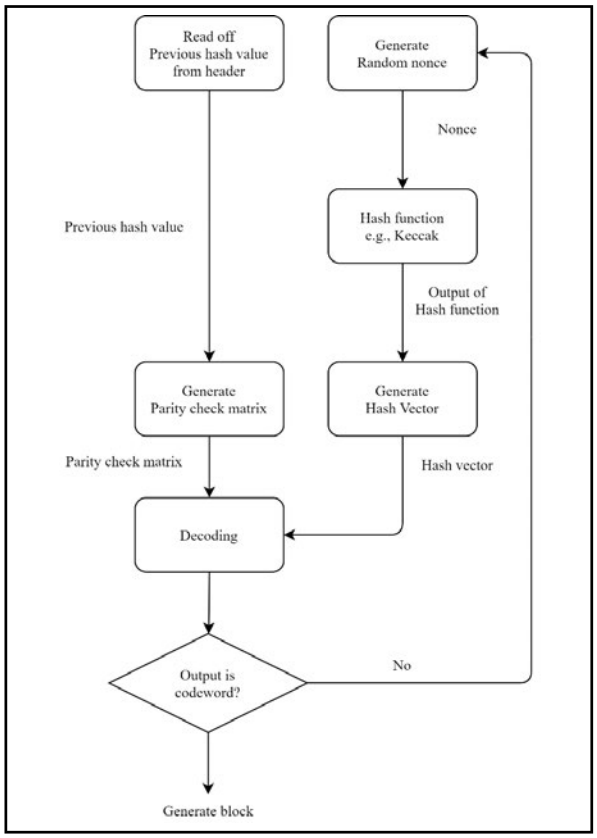


FIGURE 1. Flowchart of ECCPoW Ethereum. Every miner who generates blocks can construct a parity check matrix using a previous hash value. A generated nonce becomes an input of a hash function. A hash vector used for decoding can be generated using the output of a hash function. If decoding is successful, the block is generated; otherwise, a miner generates a new nonce to make a new hash vector for decoding.

In [7], the authors have reported that the time-varying puzzle system may generate large block generation time (BGT), i.e., outliers, for ECCPoW implemented on Bitcoin. If outliers occur frequently, it is of our interest to see whether or not the distribution of BGT is heavy-tailed with an infinite mean [15], [26]. As a result, the proposition made in [6] that BGT has a finite mean needs to be challenged. Previous works on ECCPoW [6], [7] did not include sufficient real-world experiments to conclude that BGT has a finite mean. If BGT does not have a finite mean, ECCPoW cannot be used as a consensus algorithm. In this paper, we aim to study the distribution of BGT of ECCPoW implemented on Ethereum (ETH-ECC). Our experimental results show that the BGT distribution is not heavy-tailed and has a finite mean.

The contributions of our work are as follows:

- We show how ECCPoW is implemented on Ethereum.
- We present a method for controlling the difficulty level in ETH-ECC and report the results of automatic difficulty level change with real-world experiments of ETH-ECC.
- We present a goodness-of-fit result using the Anderson–Darling (AD) test for distribution validation and discuss

the BGT distribution follows the exponential distribution which has a finite mean.

The remainder of this paper is organized as follows. Section II provides a background of the requirements of an ASIC-resistant PoW. Section III demonstrates the implementation of ETH-ECC. Section IV discusses the formulation of the problem. Section V provides the experimental result of the implementated ETH-ECC. Finally, Section VI summarizes our work and concludes the paper.

II. BACKGROUND

We introduce three approaches that can be used to avoid centralization problems in PoW. The first is an intentional bottleneck between an arithmetic logic unit (ALU) and memory, which is used by Ethash of Ethereum [2], [5]. It is also termed a memory-hard technique. The second is the *high complexity of ASIC design* used by Dash [12], Raven [13], and our method, ECCPoW. The third is *hybrid methods* of two methods; Random X of Monero uses *hybrid methods* [24].

A. INTENTIONAL BOTTLENECK

The most known PoW of the intentional bottleneck is Ethash of Ethereum [2], [5]. This method uses the difference between the throughput of ALU and the bandwidth of the memory. If there is a bottleneck between the ALU and memory, it is impossible to use the entire throughput of ALU. Specifically, if a miner needs to obtain data from memory to generate a block, the number of block generation attempts is determined by memory bandwidth. Ethash uses a directed acyclic graph (DAG), which is a set of randomly generated data for the bottleneck. The DAG is a huge dataset that cannot be stored in a cache memory; therefore, the DAG is stored in memory. To generate a block using Ethash, a miner must mix a part of the DAG that is stored in the memory. Owing to this procedure, the miner cannot avoid the bottleneck because of limited memory bandwidth. This method has been ASIC resistant for a long time; however, Bitmain released ASIC for Ethash in 2018.

B. HIGH COMPLEXITY OF ASIC DESIGN

Because of the high complexity of ASIC design, ASICs are less efficient. For example, if ASICs are less efficient than a general-purpose unit such as CPU or GPU, there is no reason to design ASIC. X11 of Dash [12] and X16R of Raven [13] use this method. Unlike PoW of Bitcoin, which uses only one hash function (SHA-256), X11 uses 11 hash functions consecutively: BLAKE, BMW, Grosetl, JH, Keccak, Skein, Luffa, Cubehash, SHAvite-3, SIMD, and ECHO. The BLAKE, which is the first hash function of X11, uses a block header with nonce as inputs; its output becomes the input of the next hash function. Similarly, the next hash function uses the output of the previous hash function. This procedure is repeated until a result is obtained for the last hash function. Miners determine whether they have found a valid nonce using the output of the final hash function.

Designing an ASIC for X11 was expensive; therefore, X11 was ASIC resistant. However, Bitmain released an ASIC for X11 in 2016. There are a few PoW algorithms that extend X11 (e.g., X13, X14, and X15); however, the ASICs for these have been released. X16R of Raven is an extended version of X11 of Dash. In X16R, unlike the previous extension of X11, the sequence of 16 hash functions is randomly changed. Therefore, it is costly to design an ASIC for X16R. However, T. Black, who designed X16R, mentioned that there is some evidence that ASICs for X16R exist [23]. Our ECCPoW employs a *time-varying puzzle generation system* to make ASCI design difficult. ECCPoW can make ASIC powerless as the puzzle generation system changes from block to block. We explain this further in Section III.

C. HYBRID METHODS

Random X of Monero combines the above two methods. Random X uses memory-hard techniques for the bottleneck with random code execution; Random X is optimized for CPU mining [24]. In [24], they mentioned that mining can be performed using a field-programmable gate array; however, it will be much less efficient than CPU mining. It implies that efficient mining hardware can be developed when the cost of developing chipsets is low in comparison to the mining reward. With the proposed ECCPoW, attempts in developing efficient mining hardware can be made when the reward-to-cost ratio increases. However, such attempts can be easily evaded since the parameters of ECCPoW can be easily changed, such as increasing the length of code and the code rate. The next section illustrates further the ASCI-resistance characteristic of ECCPoW.

III. ECCPoW IMPLEMENTED ON ETHEREUM

In this section, we briefly introduce ECCPoW and present how ECCPoW has been implemented on Ethereum using Fig. 1. Furthermore, we present how the difficulty level of ETH-ECC is automatically controlled.

A. OVERVIEW OF ECCPoW

In a blockchain employing the PoW consensus algorithm, a node solves cryptographic puzzles to publish a block. For a given puzzle, the node who solves the puzzle first obtains the authority to publish a block. For example, in the PoW of Bitcoin, the first node that finds a specific output of the secure hash algorithm (SHA) obtains the authority to publish a block. The PoW of Ethereum uses Keccak instead of SHA. The ECCPoW algorithm proposed in [6] is a PoW consensus algorithm that uses error-correction code, which comprises the low-density parity-check (LDPC) code [8], as a cryptographic puzzle. The ECCPoW algorithm consists of a pseudo-random puzzle generator (PRPG) and an ECC puzzle solver. Fig. 1 presents the flowchart of the ECCPoW algorithm. For every block, the PRPG generates a new pseudo-random LDPC matrix. A new LDPC matrix is distinct from the other previously generated matrices. Such a pseudo-random LDPC matrix takes the role of issuing an

independently announced cryptographic puzzle. The ECC puzzle solver uses the LDPC decoder to solve the given announced puzzle. Specifically, to publish a block, a node is required to run through an input header until the LDPC decoder hits a satisfying result; for instance, the output of the decoder is an LDPC codeword (with a certain Hamming weight). In the next subsection, we will discuss ECCPoW implementation on Ethereum with the flowchart presented in Fig. 1.

B. COMPARISON OF ETHASH AND ECCPoW

Ethereum uses `Ethash` for ASIC resistance, and `ETH-ECC` uses ECCPoW for ASIC resistance. In this subsection, we present how `Ethash` and `ETH-ECC` apply ASIC-resistance property to PoW with pseudo-codes.

`Ethash` uses a DAG for ASIC resistance. The DAG is a large size of data and is typically stored in a random access memory (RAM), not in cache memory. It implies that a miner must access the RAM to get the DAG data. Although the miner could be equipped with a high-throughput ALU, the bandwidth access from the RAM to the ALU is limited. That is, the bottleneck is the limited bandwidth of reading DAG information from the RAM; thus, any fast ALU, e.g., an ASIC implementation of `keccak512`, exceeding this bottleneck is of no use. This makes Ethash ASIC resistant.

When a miner reads DAG data from the RAM, the location where the data are read varies. The location of data reading is selected by the "mix"; the mix is a 128-byte hash value generated by the block header and a nonce. The mix is updated using the Fowler-Noll-Vo (FNV) hash function. The miner repeats this process 64 times. After updating the mix, the miner compresses the mix; for compression, the FNV hash is used again. The miner returns a hash value of the result of concatenating the compressed mix and the seed.

If this hash value is less than the desired target, the nonce is validated, and a new block is linked to the previous block. **Algorithm 1** denotes the pseudo-code of `Ethash`.

Algorithm 1 `Ethash`

Require: block header (BH), nonce, DAG

1: Initialize seed: seed = `keccak512`(BH, nonce)
2: Initialize 128 bytes mix:
   mix = `concatenate`(seed, seed)
3: **for** i = 0, 1, 2, . . . ,63:
4: Get data from DAG using mix:
   data = `DAG_lookup`(DAG, mix, i)
5: update mix: mix = `FNV_hash`(mix, data)
6: **end for**
7: **for** i = 0, 4, 8, . . . , `length`(mix):
8: Compress mix: cmix = `compress_mix`
   (mix, i)
9: **end for**
10: return `keccak256(concatenate`(seed, cmix))

`Ethash` uses *the intentional bottleneck* for ASIC resistance, but `ETH-ECC` aims to use *a time-varying puzzle generation system* for ASIC resistance. In `ETH-ECC`, two factors make the design of ASICs very difficult. One is flexible code lengths and randomly generated PCMs. The `ECC_puzzle_solver` generates a hash vector of length-n (subsection C) using a nonce; this n determines the code length. The development of an ASIC for a PCM with length n cannot be realized, as the `ETH-ECC` network changes n and the PCM from one block to another block. The `PRPG` creates a PCM $\mathbf{H}$. A PCM uses a BH as a seed; thus, it is randomly generated. All miners that work to extend the same previous block use the same PCM to solve the ECCPoW puzzle. Thus, it is highly expensive, if not impossible, to implement an ASIC that can handle a time-varying PCM [10], [11]. After generating a hash vector and a PCM, a miner works out how to generate an output word. If this output word satisfies a specific condition, the miner is successful at completing ECCPoW; e.g., the output word can be a codeword, and then, a new block is linked to the previous block. **Algorithm 2** denotes the pseudo-code of `ETH-ECC`. In our implementation, we have replaced `Ethash` and all its relevant peripheral systems with ECCPoW; thus, it has the same requirement as `Ethash` except for the DAG. We present more details about `ETH-ECC` in the following subsections.

Algorithm 2 `ETH-ECC`

Require: block header (BH), nonce

1: Generate hash vector:
   hash_vector = `ECC_puzzle_solver`(nonce)
2: Generate parity check matrix: PCM = `PRPG` (header)
3: output_word = `decoder`(PCM, hash_vector)
4: return output_word

C. ECCPoW ON ETHEREUM

In this subsection, we present how the error-correction process is applied to ETH-ECC using Fig. 1.

$$C := \{\mathbf{c}|\mathbf{Hc} = \mathbf{0} \cap \mathbf{c} \in \{0, 1\}^{n\times 1}\} \tag{1}$$

when a PCM $\mathbf{H}$ is given, a code $\mathbf{c}$, satisfying (1), is referred to as an LDPC code. The goal of the ECCPoW algorithm is to find an LDPC code $\mathbf{c}$ using the PCM $\mathbf{H}$, which is derived by PRPG, and a hash vector $\mathbf{r}$, which is obtained using the ECC puzzle solver. For the PRPG, we employ the previous hash value; the previous hash value, known as the parent hash in the Ethereum block header, randomly generates a PCM. Specifically, we use Gallagher's method to create random PCM [9]; we use the previous hash value as a seed of randomness. Thus, PCMs are changed for every block; because every node has the same seed, they use the same PCM until a block is generated [6].

1) ECC PUZZLE SOLVER ON ECCPoW ETHEREUM

Here, we introduce the ECC puzzle solver process in ETH-ECC. Our definitions are based on [6]. The equations below follow the right-hand side of Fig. 1.

Definition 1 (ECC Puzzle Solver): Hash vector $\mathbf{r}$ in which the size of n can be obtained as follows:

$$s_1 := Keccak(nonce) \in \{0, 1\}^{256} \tag{2}$$

where *Keccak* denotes the hash function used in Ethash of Ethereum [5]. We generate a *nonce* in the same way that Ethereum does. Furthermore, for a longer length of a hash vector, we use $s_u := Keccak(s_1) \in \{0, 1\}^{256}$ with $u = 2, 3, \ldots, l + 1$. We slice or concatenate the result of *Keccak* to generate a flexible length hash vector $\mathbf{r}$:

$$\mathbf{r} := \begin{cases} s_1[1:n] & \text{if } n \leq 256 \\ [s_1 \cdots s_l s_{l+1}[1:j]] & \text{if } n > 256 \end{cases} \tag{3}$$

where $l = \lfloor n/256 \rfloor$ and $j = n - 256 \times l$. For example, when n is less than 256, $\mathbf{r}$ obtains the same length as n, whereas when n is not less than 256, $\mathbf{r}$ concatenates the results of *Keccak*. This flexible length hash vector is used for ASIC resistance.

2) PoW OF THE LDPC DECODER

The goal of the LDPC decoder is to find a hash vector $\mathbf{c}$ that satisfies $\mathbf{Hc} = 0$. The definition below explains the decoding presented in Fig. 1.

Definition 2 (Decoder): Given a PCM $\mathbf{H}$, which is the size of $m \times n$, and hash vector $\mathbf{r}$, which is the size of n, are given, the LDPC decoder uses $\mathbf{H}$ and $\mathbf{r}$ as inputs and obtains output $\mathbf{c}$ using the message-passing algorithm [6], [14]. When $\mathbf{c}$ satisfies (1), $\mathbf{c}$ becomes an LDPC code, and a miner completes LDPC decoding.

$$D_{np} : \{\mathbf{r}, \mathbf{H}\} \mapsto \mathbf{c} \in \{0, 1\}^{n \times 1} \tag{4}$$

A PCM $\mathbf{H}$ is randomly generated; however, all miners use the same previous hash value, which is derived from the previous block. Therefore, predicting the next PCM to mine a block in advance is impossible. In the PoW of Ethereum, miners change a nonce when they obtain a wrong output. We follow the same procedure as Ethereum to obtain a hash value from *Keccak* with a *nonce*, but ETH-ECC uses one more step (3) to generate a hash vector for decoding. When the code derived by (4) does not satisfy (1), the miner generates a new *nonce* and repeats all steps.

Our method is based on the *high complexity of ASIC design* in Section II for an ASIC-resistant PoW. However, unlike the mentioned method in Section II, ECCPoW generates varying cryptographic puzzles of *high complexity*. Specifically, ECCPoW uses two factors to achieve *high complexity*: flexible length LDPC code $\mathbf{c}$ and randomly generated PCM $\mathbf{H}$. ASICs can be released for the n length of code. However, extending the length of code (e.g., $n + 1$) makes ASICs powerless. Furthermore, in [10], [11], it has been proven that implementing an ASIC that can handle variable PCMs is expensive and occupies a lot of space. If developing an ASIC costs more than buying a CPU or GPU, there is no incentive to develop an ASIC. In other words, the ECCPoW algorithm is ASIC resistant as implementing an ASIC that can handle various lengths of changing codes and randomly generated PCMs is inefficient.

D. DIFFICULTY-LEVEL CONTROL OF ETH-ECC

In this subsection, we demonstrate the implementation of ETH-ECC's difficulty-level control. Bitcoin [1] and Ethereum [2] have different difficulty-level control methods. Furthermore, we present one way to add fine difficulty control.

In Bitcoin, the Bitcoin network changes the difficulty level every 2016 block; the desired BGT is 10 min for a block. If miners generate a block every 10 min, generating 2016 blocks takes precisely 2 weeks. Thus, if generating 2016 blocks takes more than 2 weeks, the difficulty level decreases; otherwise, the difficulty level increases. Unlike Bitcoin, the Ethereum network changes the difficulty level every block. Ethereum network allows for a block to be generated between 9 and 18 s. If a block is generated within 9 s, then the difficulty level increases. If it exceeds 18 s, then the difficulty level decreases. Because of this difference between Bitcoin and Ethereum, ECCPoW-based Bitcoin (BIT-ECC) and ETH-ECC also have different difficulty-level control methods. Thus, ETH-ECC cannot use BIT-ECC's method. Because of the need for a new method, we demonstrate the implementation of ETH-ECC's difficulty level control with a difference from Ethereum's method.

Ethereum uses the number of attempts to generate a block per second, termed hash rate, and a probability of block generation. Similarly, ETH-ECC uses the hash rate but considers a probability of decoding success. In [5], the difficulty of Ethereum is defined by the probability of block generation. The difficulty is as follows:

$$n \leq \frac{2^{256}}{Diff} \tag{5}$$

It indicates that

$$Diff \leq \frac{2^{256}}{n} \tag{6}$$

where n denotes the result of PoW and *Diff* denotes the difficulty of Ethereum. Thus, (6) means that when the difficulty level increases, the number of n that satisfies (6) decreases. Furthermore, we can consider that the reciprocal of difficulty is a probability of block generation. Ethereum uses this probability and hash rate to control BGT. For example, without replacement, when the probability of block generation is 1/150 and hash rate is 10 hash per second, brute force takes 15 s. If the hash rate increases, such as 20 hash per second, Ethereum's method adjusts the probability of BGT to 1/300. Thus, brute force takes 15 s even though the hash rate increases.

For ECCPoW, if we can calculate a probability of decoding success, it is possible to control the difficulty level similar to the process in Ethereum. Thus, it is important to know the probability of a successful LDPC decoding according to the LDPC parameter. We use the pseudo-probability of a successful LDCP decoding according to the parameters to test the difficulty level change using the BGT [7]. That is, ETH-ECC uses the probability of decoding success and

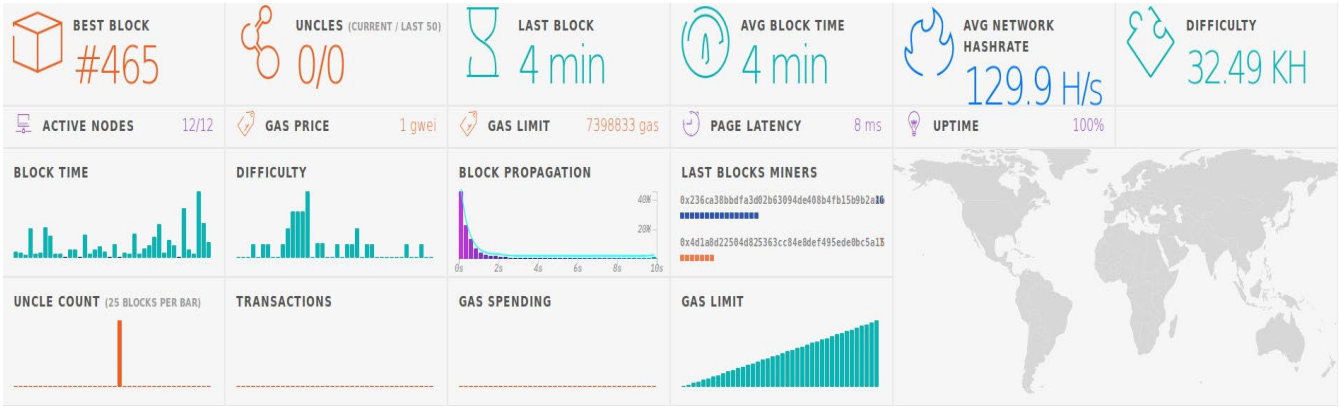


FIGURE 2. This figure shows the simulation results of ECCPoW Ethereum on Amazon Web Services AWS). Twelv nodes are used in the simulation. The tw nodes are bootnodes that help connect th nodes, and the other 10 nodes are sealnodes that participate i block generation. We use the m5.xlarge of AWS EC2 for the simulation. In the charts, BLOCK TIME shows the block generation times for the last 40 blocks, and DIFFICULTY shows the difficulty levels of the last 40 blocks. BLOCK PROPAGATION shows th percentage of the block propagation time corresponding to time.

hash rate to control the difficulty level. For example, without replacement, when the probability of decoding success is 1/150 and the hash rate is 10 hash per second, it takes 15 s, as in the above example of Ethereum's method. However, unlike Ethereum, when the hash rate increase, ETH-ECC tunes parameters of LDPC to adjust the probability of decoding success. By tuning parameters, ECCPoW achieves both difficulty-level control and ASIC resistance. These parameters can be found at https://github.com/cryptoecc/ETH-ECC/blob/master/consensus/eccpow/LDPCDifficulty_utils.go#L65. In Fig. 2, the difficulty of ETH-ECC is 32.49 KH, indicating that the probability of block generation is 1 of 32,490 hash.

One Way to Add Fine Difficulty Control. ECCPoW controls difficulty using integer and discrete variable n. Thus, it may look inappropriate to manage difficulty precisely. However, as the number of blocks increase, block generation time (BGT) converges to the ideal BGT time, which is suitable for a network. For example, when there exist two difficulties: n and $n+1$, we can define average BGT of each difficulty as t_n and t_{n+1}. Thus, we can define the average BGT:

$$averageBGT = \frac{\alpha t_n + \beta t_{n+1}}{k} \tag{7}$$

where α denotes the number of generated blocks with difficulty t_n, β denotes the number of generated blocks with difficulty t_{n+1}, and k denotes the total number of generated blocks (TNGB). Thus, α can be replaced as $\alpha = k - \beta$. As a result, equation (7) is:

$$averageBGT = \frac{(k-\beta)t_n + \beta t_{n+1}}{k} \tag{8}$$

When TNGB k is kept constant, the average BGT is determined by the number of generated blocks β in Eq. (8). Thus, the ideal average BGT, which is suitable for the number of nodes in a network, depends on β. In other words, when TNGB $k$is low, the average BGT cannot meet the ideal average BGT because there are not enough blocks of each difficulty. However, as TNGB k increases, the number of blocks corresponding to the difficulty, such as β, getting closer to the proportion that fits the probability of block generation. As a result, average BGT converges to the ideal average BGT; this convergence confirms our proposition that the network can control difficulty precisely.

IV. PROBLEM FORMULATION

In PoW, there is a case that nodes generate blocks at the same time. Bitcoin allows only one block to be generated at a time; Ethereum allows three blocks to generate at the same time. However, in Ethereum, only one block can be canonical; the other blocks cannot. Blocks that cannot be canonical are called uncle blocks. In Ethereum, nodes rollback transactions of uncle blocks [5]. Therefore, the transaction's participants must wait for block confirmation to prevent a rollback. That is, in the blockchain using PoW, the BGT must have a finite mean for the block confirmation time. For example, if the BGT has an infinite mean, the waiting time for the confirmation of transactions cannot be determined. Therefore, to apply the ECCPoW algorithm in a real network, the BGT must have a finite mean.

In [6], the authors presented the definition of the block generation of the ECCPoW algorithm using a hash rate with a geometric distribution. That is, they assumed that nodes generate a block with specific block generation attempts. However, if the BGT has an infinite mean, there is no guarantee that nodes generate a block with specific attempts. In [7], the authors presented a practical experiment using the ECCPoW algorithm. However, they only mentioned that the BGT of ECCPoW is "unstable." That is, they mentioned that the BGT of ECCPoW has outliers; however, they did not present a discussion on the BGT. Thus, in this study, we present a discussion on the BGT. Specifically, our experimental result presents evidence that the exponential distribution describes the distribution of the BGT of ECCPoW.

V. EXPERIMENT ON ETH-ECC

In this section, we conduct experiments using ETH-ECC. First, we simulate the difficulty level change using multinode networks. Second, we conduct a goodness-of-fit experiment using the AD test [16]–[18] to discuss the distribution of the BGT with a fixed difficulty level.

A. SIMULATION OF THE DIFFICULTY CHANGE

We simulate the difficulty-level change employing Amazon Web Services (AWS) using 12 nodes. Two nodes are *bootnodes* that help connect the nodes, and the other 10 nodes are *sealnodes* that participate in block generation. In the charts presented in Fig. 2, *BLOCK TIME* presents the BGT of the last 40 blocks, and *DIFFICULTY* shows the difficulty level of the last 40 generated blocks. *BLOCK TIME* and *DIFFICULTY* show that because of the large standard deviation, a block is gradually generated despite the low difficulty level, as mentioned in [7]; in the next subsection, we discuss the BGT. In the charts presented in Fig. 2, *LAST BLOCK* shows the BGT of the previous block, and *AVG BLOCK TIME* shows the average of the BGT. In addition, *AVG NETWORK HASHRATE* shows the average hash rate of all miners. *BLOCK PROPAGATION* shows the block propagation time from a miner who generated a block to other miners. We used two different regions: Seoul and US East for *sealnodes*. Specifically, 3 of the 10 *sealnodes* are in the US East region, whereas the rest are in the Seoul region. *BLOCK PROPAGATION* also shows the percentage of blocks that are propagated at corresponding times. *BLOCK PROPAGATION* indicates that the propagation of approximately all blocks between Seoul and US East regions takes less than 2 s. The block propagation is the same method as that of Ethereum.

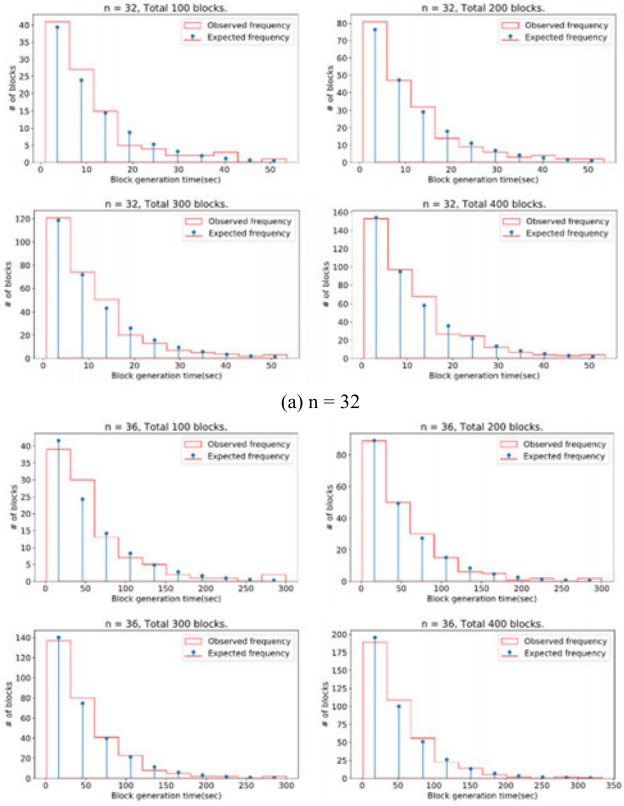
B. STABILITY OF THE BLOCK GENERATION TIME

Fig. 2 demonstrates the importance of determining whether varying puzzles may result in outliers. That is, in *BLOCK TIME* and *DIFFICULTY* of Fig. 2, slow block generations are observed despite the low difficulty level. In other words, the observation of BGT shows outliers. If the outliers are uncontrollable, the BGT distribution has an infinite mean similar to the heavy-tailed distribution. An infinite mean cannot guarantee transaction confirmation. Thus, to achieve a stable BGT that can guarantee transaction confirmation, the BGT must have a finite mean.

We obtain the BGT of ECCPoW Ethereum with a fixed difficulty level to observe the type of distribution with a finite mean the BGT follows. Specifically, if BGT follows an exponential distribution, it has a finite mean. However, if the BGT follows a heavy-tailed distribution, it has an infinite mean [15]. Thus, through the goodness-of-fit experiment, we aimed to discuss what type of distribution the BGT follows. For the goodness-of-fit experiment, we set a null hypothesis H_0 and an alternative hypothesis H_A:

H_0: BGT has the exponential distribution

H_A: BGT does not have the exponential distribution



(a) n = 32

(b) n = 36

FIGURE 3. We did experiments for 100, 200, 300, and 400 blocks to observe the distribution over the number of blocks. As the number of blocks increases, the standard error decrease. That is, when the number of blocks increases, sample distribution reflects an actual distribution of sample distribution. In these figures, experiment results show the tendency; the distributions of observed frequency, known as sample distribution, follow the distribution of expected frequency.

For the goodness-of-fit experiment, we use the AD test [16]–[18]. Other available tests can be used in the goodness-of-fit experiment, such as the chi-square test [19], Kolmogorov–Smirnov test [20], and AD test [16]. The chi-square test has a restrictive assumption that all expected frequencies should be greater or equal to 5 [21]. However, there is no guarantee that our samples will achieve this assumption. If we collect more samples, the chi-square test can be used. However, the p-values used to validate the hypotheses are affected by the number of samples. When the number of samples increased in the chi-square test, the p-values tend to decrease. Therefore, the assumption of the chi-square test is inappropriate for verifying our distributions. The Kolmogorov–Smirnov test is unaffected by sample sizes; however, it is more sensitive to the center of the distribution rather than the tail [22]. We must consider verifying the tail of the distribution to cover all possibilities. Therefore, we have chosen to use the AD test [16], which gives more weight to the tail than does the Kolmogorov–Smirnov test.

C. AD TESTS

In this subsection, we discuss the AD test and verify its usage using test examples. The AD test is used to verify if a sample

TABLE 1. Example of the Anderson–Darling test results.

The number of samples	Standardized A^2_{MN}	p-value
10	-0.59	$p \geq 0.25$
20	0.44	$p = 0.21$
30	0.69	$p = 0.17$

(a) F ~ Exp(1), G ~ Normal(1, 1)

10	1.20	$p = 0.11$
20	3.57	$p = 0.02$
30	4.67	$p = 0.01$

(b) F ~ Exp(1), G ~ Exp(2)

10	1.11	$p = 0.12$
20	-0.41	$p \geq 0.25$
30	-0.08	$p \geq 0.25$

(c) F ~ Exp(1), G ~ Exp(1)

follows a specific distribution. We discuss one-sample and two-sample AD tests. In our work, we use the two-sample AD test; however, to clearly present our contribution, we briefly introduce the one-sample AD test first.

1) ONE-SAMPLE AD TEST

The one-sample AD test is suitable to verify a hypothesis that a sample set comes from a population. The one-sample AD test is as follows. When the cumulative distribution function (CDF) of the population distribution is $F(x)$, and the CDF of the empirical distribution is $F_M(x)$, the one-sample AD test [18] is used as follows:

$$A^2_M = M \int_{-\infty}^{\infty} (F_M(x) - F(x))^2 \, w(x) dF(x) \tag{9}$$

and

$$w(x) = [F(x)(1 - F(x))]^{-1} \tag{10}$$

where M denotes the number of samples and A^2_M denotes the results of the one-sample AD test. Intuitively, in (9), if $F_M(x) - F(x)$ is 0 for all x, A^2_M is 0. This means that when A^2_M is small, the empirical distribution $F_M(x)$ is close to the population distribution $F(x)$. As we have noted, we focus on the tail of the distribution; it can be accomplished using (10). The one-sample AD test result A^2_M can be used to verify if a given sample comes from a population with a specific distribution.

2) TWO-SAMPLE AD TEST

In our work, we want to verify that two-sample sets come from the same unknown population. The two-sample AD test is appropriate for such verification. The two-sample AD test [17], [18] is as follows. There are two-sample empirical distributions $F_M(x)$ and $G_N(x)$. The $F_M(x)$ is an empirical distribution derived from the set $\mathcal{F}$ with a cardinality of the sample set $M = |\mathcal{F}|$. The $G_N(x)$ is also an empirical distribution derived from the set $\mathcal{G}$ with a cardinality of the sample set $N = |\mathcal{G}|$. $F_M(x)$ and $G_N(x)$ are the respective sample sets

TABLE 2. The observed frequency is calculated using the histogram in Fig. 4, and the expected frequency is calculated using the CDF of the exponential distribution derived from the mean in Fig. 4.

Interval(%)	Observed frequency	Expected frequency
[0, 10)	107	118.70
[10, 20)	82	71.73
[20, 30)	56	43.35
[30, 40)	20	26.19
[40, 50)	14	15.83
[50, 60)	7	9.56
[60, 70)	5	5.78
[70, 80)	4	3.49
[80, 90)	2	2.11
[90, 100]	3	1.27

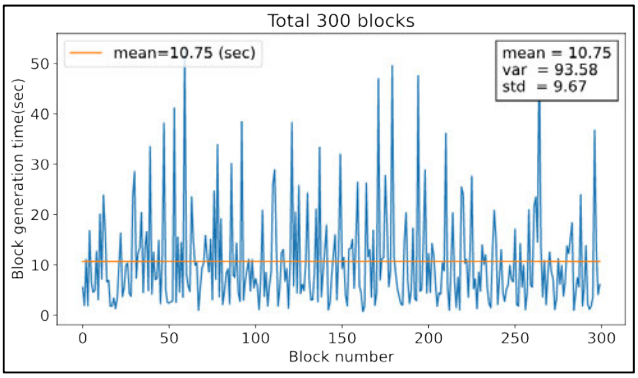


FIGURE 4. This figure presents block generation time of 300 blocks when n is 32. The mean block generation time of 300 blocks is 10.75 s, and it is presented as a horizontal line. Such a result is converted to a histogram. The observed frequency of Table 2 denotes the histogram of Fig. 4. The legend at the top right shows the mean, variance, and standard deviation of the BGT.

independently obtained from two different testing locations. The two-sample AD test can be used to determine whether both sample distributions come from the same distribution. In [17], [18], the two-sample

version is defined as follows:

$$A^2_{MN} = \frac{MN}{K} \int_{-\infty}^{\infty} \frac{(F_M(x) - G_N(x))^2}{H_K(x)(1 - H_K(x))} dH_K(x) \tag{11}$$

where $H_K(x) = (MF_M(x) + NG_N(x))/K$ with $K = M + N$. A^2_{MN} is standardized to remove the dependencies derived by the number of samples. This standardized form is used to calculate the p-value [17], [18]. The p-value evidences the hypothesis test.

The two-sample AD test is suitable to verify a hypothesis that two-sample sets come from the same population. As a null hypothesis H_0 for the two-sample AD test, we set $F_M(x)$ to have the same population as $G_N(x)$. In addition, we set $G_N(x)$ as an exponential distribution. Thus, if $F_M(x)$ and $G_N(x)$ comes from same the population, that is, H_0 is true, we may consider that $F_M(x)$ is the exponential distribution. If the p-value of the AD test is sufficiently large, it proves that H_0 is true.

The p-value is the false positive probability under the assumption that the null hypothesis is true. A low p-value indicates that a test result provides evidence against the null hypothesis; a large p-value does not. That is, a large p-value denotes the probability of a true negative is low. The p-value is determined from the observation of the sample data. Thus, before observing the data, we first set the threshold significance level (*TSL*), $TSL \in [0, 1]$. The *TSL* can be used to determine the critical value. Given a *TSL* and the number of samples that are used in the AD test, the *TSL* table in [18] is used to read off a value corresponding to the TSL and the number of samples. This read-off value is called the critical value. If the standardized A^2_{MN} is smaller than the critical value, this result indicates that the p-value is larger than the predefined *TSL*. In the *TSL* table of [18], the maximum *TSL* is 0.25. Thus, when standardized A^2_{MN} is less than the critical value corresponding to the 0.25 *TSL*, the p-value is capped at 0.25.

3) VERIFICATION OF THE AD TEST

In this subsection, we verify the two-sample AD test method. Verification is performed under the assumption that the input distributions are *a priori* known. This will clearly illustrate how we will use the AD test and interpret its test results.

In Table 1, we present three examples to give an insight into the p-value of the AD test; in this example, we use true distributions for $F_M(x)$ and $G_N(x)$. In Table 1, Exp(θ) indicates the exponential distribution with mean θ and Normal(μ, σ) indicates the normal distribution with mean μ and standard deviation σ. That is, $\mathcal{F} \sim \text{Exp}(\theta)$ denotes the sample set $\mathcal{F}$ of $F_M(x)$; samples are derived from the exponential distribution with mean θ. In Table 1 (a), we use the exponential distribution for $F_M(x)$ and the normal distribution for $G_N(x)$; these distributions have the same mean. This example shows that as the number of samples increases, the p-value tends to decrease if samples are drawn from different distributions. In Table 1 (b), we set both $F_M(x)$ and $G_N(x)$ as the exponential distribution, but each with different mean values. This example shows that as the number of samples increases even though samples are drawn from the same exponential distribution, the p-value tends to decrease if the means of distributions are different. In Table 1 (c), we set both $F_M(x)$ and $G_N(x)$ to be exactly the same exponential distribution. That is, the two-sample sets $\mathcal{F} \sim F_M(x)$ and $\mathcal{G} \sim G_N(x)$ come from the same population. This example shows that, as the number of samples increases, the p-value tends to increase when two-sample sets are drawn from the same population. From these examples in Table 1, we note that the closer the two distributions $F_M(x)$ and $G_N(x)$ are to each other, the larger p-value is obtained.

We determine whether the AD test result of our experiments indicates that $F_M(x)$ is sufficiently close to $G_N(x)$. That is, given there are two-sample sets, one of $F_M(x)$ and the other of the exponential $G_N(x)$, we want to determine whether we can make a quality statement about how close the two-sample sets are to each other according to the AD test.

TABLE 3. Anderson-Darling test result. The test result presents a large p-value. It means that if we reject the null hypothesis, the probability of a true negative is low.

n	# of blocks	Observed mean(sec)	std	Standardized A^2_{MN}	p-value
32	100	10.86	9.84	-1.12	$p \geq 0.25$
32	200	11.24	10.16	-1.20	$p \geq 0.25$
32	300	10.74	9.67	-1.18	$p \geq 0.25$
32	400	11.08	9.84	-1.09	$p \geq 0.25$
32	500	10.91	9.62	-1.11	$p \geq 0.25$
32	600	10.87	9.48	-0.80	$p \geq 0.25$
32	700	10.84	9.41	-0.36	$p \geq 0.25$
32	800	10.76	9.40	-0.36	$p \geq 0.25$
36	100	56.00	55.20	-1.11	$p \geq 0.25$
36	200	51.04	49.71	-1.19	$p \geq 0.25$
36	300	47.84	45.49	-1.12	$p \geq 0.25$
36	400	49.97	47.80	-1.19	$p \geq 0.25$
36	500	49.24	46.95	-1.11	$p \geq 0.25$
36	600	48.23	46.96	-1.18	$p \geq 0.25$
36	700	48.36	47.68	-1.18	$p \geq 0.25$
36	800	48.03	46.89	-1.18	$p \geq 0.25$

The AD test result presents a significant p-value, i.e., $p \geq 0.25$; it is a necessary condition but not a sufficient one for the case that the two distributions are the same. In other words, if a decision is made to reject the null hypothesis, that is, the distribution $F_M(x)$ is not close to the exponential distribution $G_N(x)$, such a decision will result in an error with a probability greater than 0.25.

D. APPLICATION OF AD TEST TO BGT DISTRIBUTION

In this subsection, we use the AD test to determine the distribution of the BGT of ETH-ECC. For this experiment, 90 threads were used to generate a block. We experimented using a fixed code length to observe the BGT without changing the difficulty level. In the test, two kinds of code length n are used: 32 and 36. These are the two lowest types of code length n in our pseudo-difficulty table used in the simulation. We divided the BGT into 10 intervals between the minimum BGT and maximum BGT for a histogram. For example, when the minimum BGT is 10 and the maximum BGT is 20, there are 10 intervals, i.e., [10,11], [11,12], . . . , [19,20]. Using these intervals, we count the observed frequency of the BGT data. We set $F_M(x)$ using the observed frequency and set $G_N(x)$ using the mean of the BGT data. The mean in Fig. 4 is used for the expected frequency of $G_N(x)$ in Table 2. That is, the mean in Fig. 4 is used as $1/\lambda$ for the CDF of the exponential distribution $G_N(x)$:

$$G_N(x) = 1 - e^{-\lambda x} \tag{12}$$

The expected frequency of Table 1 is calculated using the integral of $G_N(x)$ corresponding to the interval time. Because $G_N(x)$ is the exponential distribution, if $F_M(x)$ is close to $G_N(x)$ we may consider $F_M(x)$ is an exponential distribution.

E. DISCUSSION ON AD TEST RESULTS

Fig. 4 shows the example result of the BGT over different blocks. Each block denotes the trial to obtain the BGT. We converted the test results, such as those in Fig. 4, to a distribution over time to analyze the BGT. These converted distributions are presented in Fig. 3. Fig. 3 presents the plots of the distribution of the observed and expected frequencies. These frequencies are calculated using the method described in Section V-D.

When we obtain a distribution using a sample set, there is a standard error; the standard error is high when the number of samples in the set is small. The standard error is expressed as

$$\frac{\sigma}{\sqrt{N}}$$

where σ denotes the standard deviation of a population and N denotes the cardinality of the sample set. The standard error decreases as the number of samples increases. Thus, the sample distribution becomes closer to the actual distribution of the observed samples. If the sample distribution, which reflects the actual distribution, differs from the expected distribution, we can observe that the sample distribution differs from the expected distribution. To observe the tendency of distribution over some blocks, we experimented with 100, 200, 300, and 400 blocks. Fig. 3 shows that the distribution of the observed frequency tends to follow the distribution of the expected frequency. In addition, Table 3 shows that the observed mean and standard deviation tend to converge as the number of blocks increases.

Furthermore, for the quantitative analysis, we use the AD test. Table 3 presents the AD test results to discuss hypotheses H_0 and H_A. These results show a similar result in Table 1 (c). In Table 1 (c), we drew samples from the same true distribution; the results present the largest possible p-value. All p-values in Table 3 are larger than or equal to 0.25, regardless of the number of blocks. In other words, if the null hypothesis is rejected, this decision will cause an error with a probability greater than 0.25. That is, the decision that the BGT distribution $F_M(x)$ does not follow the exponential distribution could be made with a high decision error.

VI. DISCUSSION

The purpose of ECCPoW is not to replace the current PoW of Ethereum. We propose our algorithm to present as one of the options for the Ethereum network. Ethereum can be utilized, for example, not only in a large-scale network but also in local-scale networks. To support a local-scale network, Ethereum provides PoW and PoA(Proof-of-Authority) as consensus algorithms. These algorithms have limitations for the local-scale network. For instance, PoW based local network has a risk of a double-spending attack by ASIC miners; PoA based network has a limitation of a participant because the time complexity of a PoA increases exponentially when the number of participants increases. Our algorithm, ECCPoW, can be utilized in such cases for the benefit of offering a novel PoW that allows numerous participants with deterrence to ASIC-borne attacks. In addition, our novel ECCPoW may open up for an expected use and thus untraveled future to Ethereum.

Extensive Simulation Set up at AWS. We have recruited twelve instances on Amazon Web Service (AWS) EC2; each instance of EC2 instances works as a node in a blockchain network. The cost of using AWS EC2 increases rapidly because PoW utilizes all the resources of instances. We were able to confirm that this scale of the experiment was good enough to achieve our main goal, which is aimed at verifying the stability of the block generation time of ECCPoW Ethereum. AWS simulation was done to obtain the trace data of block generation times. The twelve nodes employed in our simulation were divided into two different kinds of nodes. One kind is *bootnodes* which help the nodes connected. Nodes that want to join a network are connected to *bootnodes* first. After connection, *bootnodes* relay nodes to other nodes. In Ehtereum, *bootnodes* addresses are hardcoded on source codes, but it is possible to set *bootnodes* addresses manually for private networks. We have chosen two *bootnodes*. The other kind of nodes are *sealnodes* that participate in block generation as a miner in the PoW network. We have chosen the number of *sealnodes* to be 10. We use the m5.xlarge of AWS EC2, which has conventional node specification: four virtual CPUs and 16 GB memory for the real-world simulation. All nodes are deployed by Docker according to the guidance of Ethereum. Thus, all of our simulation results, which are shown in Fig. 2, are reproducible.

VII. CONCLUSION

In this work, we present the implementation, simulation, and validation of ETH-ECC. In the implementation, we showed how Ethereum can be updated with ECCPoW as its new consensus algorithm. In the simulation, we conducted a multinode experiment using AWS EC2. The results showed that ETH-ECC with its adaptive difficulty-level controllability is successfully implemented. In the validation, we showed statistical results in which the necessary condition for a finite mean BGT is satisfied such that the distribution of the ECCPoW block generation time is exponential.

REFERENCES

[1] S. Nakamoto. *Bitcoin: A Peer-to-Peer Electronic Cash System*. Accessed: Nov. 3, 2021. [Online]. Available: https://bitcoin.org/bitcoin.pdf

[2] V. Buterin. *A Next-Generation Smart Contract and Decentralized Application Platform*. Accessed: Nov. 3, 2021. [Online]. Available: https://ethereum.org/en/whitepaper/

[3] M. Rosenfeld, "Analysis of hashrate-based double spending," 2014, *arXiv:1402.2009*. [Online]. Available: http://arxiv.org/abs/1402.2009

[4] J. Jang and H.-N. Lee, "Profitable double-spending attacks," 2019, *arXiv:1903.01711*. [Online]. Available: http://arxiv.org/abs/1903.01711

[5] G. Wood, "Ethereum: A secure decentralised generalised transaction ledger," *Ethereum Project Yellow Paper*, vol. 151, pp. 1–32, Apr. 2014.

[6] S. Park, H. Choi, and H. Lee, "Time-variant proof-of-work using error correction codes," Jun. 2020, *arXiv:2006.12306*. [Online]. Available: https://arxiv.org/abs/2006.12306

[7] H. Jung and H. Lee, "Error-correction code based proof-of-work for ASIC resistance," *Symmetry*, vol. 12, no. 6, p. 988, Jun. 2020.

[8] W. Ryan and S. Lin, "Low-density parity-check codes," in *Channel Codes: Classical and Modern*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[9] R. G. Gallager, *Low Density Parity Check Cdes*. Cambridge, MA, USA: MIT Press, 1963.
[10] S. Shao, P. Hailes, T.-Y. Wang, J.-Y. Wu, R. G. Maunder, B. M. Al-Hashimi, and L. Hanzo, "Survey of turbo, LDPC, and polar decoder ASIC implementations," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2309–2333, 3rd Quart., 2019.
[11] Y.-L. Ueng, B.-J. Yang, C.-J. Yang, H.-C. Lee, and J.-D. Yang, "An efficient multi-standard LDPC decoder design using hardware-friendly shuffled decoding," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 60, no. 3, pp. 743–756, Mar. 2013.
[12] E. Duffield and D. Diaz. *Dash: A Payments-Focused Cryptocurrency*. Accessed: Nov. 3, 2021. [Online]. Available: https://github.com/dashpay/dash/wiki/Whitepaper
[13] T. Black and J. Weight. *X16R ASIC Resistant by Design*. Accessed: Nov. 3, 2021. [Online]. Available: https://ravencoin.org/assets/documents/X16R-Whitepaper.pdf
[14] W. E. Ryan and S. Lin, *Channel Codes—Classical and Modern*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[15] S. Foss, D. Korshunov, and S. Zachary, *An Introduction to Heavy-Tailed and Subexponential Distributions*. New York, NY, USA: Springer, 2011.
[16] T. W. Anderson and D. A. Darling, "Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes," *Ann. Math. Statist.*, vol. 23, no. 2, pp. 193–212, Jun. 1952.
[17] A. N. Pettitt, "A two-sample anderson-darling rank statistic," *Biometrika*, vol. 63, no. 1, pp. 161–168, 1976.
[18] F. Scholz and M. Stephens, "K-sample anderson-darling tests," *J. Amer. Stat. Assoc.*, vol. 82, no. 399, pp. 918–924, Sep. 1987.
[19] W. G. Cochran, "The X^2 test of goodness of fit," *Ann. Math. Statist.*, vol. 23, no. 3, pp. 315–345, Sep. 1952.
[20] J. L. Hodges, "The significance probability of the Smirnov two-sample test," *Arkiv Matematik*, vol. 3, no. 5, pp. 469–486, Jan. 1958.
[21] W. G. Cochran, "Some methods for strengthening the common tests," *Biometrics*, vol. 10, pp. 417–451, Dec. 1954.
[22] J. J. Filliben. *1.3.5.16. Kolmogorov–Smirnov Goodness-of-Fit Test*. NIST/SEMATECH e-Handbook of Statistical Methods. Accessed: Nov. 3, 2021. [Online]. Available: http://www.itl.nist.gov/div898/handbook/
[23] T. Black. *Ravencoin–ASIC Thoughts—Round Two*. Accessed: Nov. 3, 2021. [Online]. Available: https://medium.com/@tronblack/ravencoin-asic-thoughts-round-two-f4f743942656
[24] Tevador. *Random X*. Accessed: Nov. 3, 2021. [Online]. Available: https://github.com/tevador/RandomX
[25] M. Belotti, N. Bozic, G. Pujolle, and S. Secci, "A vademecum on blockchain technologies: When, which, and how," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3796–3838, 4th Quart., 2019.
[26] S. Asmussen, "Steady-state properties of of GI/G/1," in *Applied Probability and Queues*. New York, NY, USA: Springer, 2003, pp. 266–301, doi: 10.1007/0-387-21525-5_10.



HYOUNGSUNG KIM received the B.S. degree in electronics and computer engineering from Chonnam National University, Gwangju, South Korea, in 2019, and the M.S. degree from the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju. Since 2021, he has working with Korea Electronics Technology Institute. His research interests include blockchain, distributed systems, and statistical analysis.



JEHYUK JANG received the B.S. degree in electronic engineering from the Kumoh National Institute of Technology, Gumi, South Korea, in 2014, and the M.S. degree in information and communication engineering from GIST, South Korea, where he is currently pursuing the Ph.D. degree with the School of Electrical Engineering and Computer Science. His research interests include sub-Nyquist sampling and compressed sensing.



SANGJUN PARK received the B.S. degree in computer engineering from Chungnam National University, Daejeon, South Korea, in 2009, and the Ph.D. degree in electrical engineering and computer science from Gwangju Institute of Science and Technology, Gwangju, South Korea, in 2019. Since 2020, he has been working with the Electronics and Telecommunications Research Institute. His research interests include information theory, numerical optimization, compressed sensing, blockchain, deep-neural networks, and finite state machine.



HEUNG-NO LEE (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees from the University of California, Los Angeles, CA, USA, in 1993, 1994, and 1999, respectively, all in electrical engineering. From 1999 to 2002, he was a Research Staff Member with HRL Laboratories, LLC, Malibu, CA, USA. From 2002 to 2008, he was an Assistant Professor with the University of Pittsburgh, PA, USA. In 2009, he was with the School of Electrical Engineering and Computer Science, GIST, South Korea, where he is currently affiliated. His research interests include information theory, signal processing theory, blockchain communications/networking theory, and their application to wireless communications and networking, compressive sensing, future internet, and brain–computer interface. He was a recipient of several prestigious national awards, including the Top 100 National Research and Development Award, in 2012; the Top 50 Achievements of Fundamental Researches Award, in 2013; and the Science/Engineer of the Month, in January 2014.

Gwangju Institute of
Science and Technology

School of Electrical Engineering and Computer Science



DeFi 강의 및 실습

<한국금융ICT융합학회 추계워크샵: 빅뱅하는 디지털금융—디파이 NFT 메타버스>
일시: 11.27토-28일

광주과학기술원 (GIST)
INFONET 연구실
이흥노 교수
최해웅 박사과정
박하영 석사과정

강의 목차

1. 스마트 컨트랙트, 토큰 정의 (1)
2. DEX 개념 (10)
   - 기존 중앙거래소 와 DEX 차이점
   - Automatic Market Maker (AMM)
   - 스왑, 가격영향, 슬리피지 예제
   - 수수료
   - 유동성 제공 및 제거
   - 비영구적 손실 설명 및 예제
3. DEX 서비스 예시 (2)
4. Flash loan 개념 (2)
5. Flash loan attack (1)
6. Lending 프로토콜 소개 및 예시 (4)
7. Stable coin 소개 및 예시 (DAI) (1)
8. SWAP 실습

스마트 컨트랙트, 토큰 정의

이더리움: 비트코인과 비교하면, 네트워크 상에 구현할 수 있는 계약의 종류가 무궁무진함.

- 스마트 컨트랙트의 작성과 배포, 사용이 편리함. 대부분의 DeFi 서비스가 이더리움 프로토콜 기반으로 동작함.

스마트 컨트랙트: 계약을 프로그램으로 작성하여 블록체인에 등록함으로써, 정확히 프로그래밍 한 그대로 실행되는 것을 보장함.

ERC20 토큰: 이더리움 상에서 스마트 컨트랙트를 사용하여, 발행, 교환, 소각 과정이 ERC20 표준에 명시된 대로 작동하는 가상 자산. 암호화폐 프로토콜 속에 있는 또다른 암호화폐.

DEX

기존 중앙거래소 와 DEX 차이점

- 블록체인은 중앙 서버나 기관이 존재하지 않기 때문에 사용자들이 네트워크를 유지할 필요가 있다.
- 네트워크를 유지하는데 기여하는 사용자들은 보상으로 암호화폐를 받을 수 있다.
- 암호화폐 간 교환을 위해 거래소가 필요
- 중앙화된 거래소를 사용할 경우 거래소의 신뢰성 문제, 단일 장애점 공격의 위험 같은 문제가 존재함.
- 같은 블록체인 프로토콜을 사용하는 토큰간의 교환의 경우, 해당 블록체인 상의 스마트 컨트랙트를 이용하여 탈중앙화된 거래소 (DEX) 를 구현하여 이와 같은 문제를 해결할 수 있음.

DEX

Automatic Market Maker (AMM) 개념 설명

- 전통적인 시장에서는 오더북(order book) 방식 사용.
- 오더북 방식은 판매자 및 구매자가 직접 거래 희망 가격을 제시하고, 적절한 거래 희망가를 제시한 판매자와 구매자 간에 거래가 체결되는 방식이다.
- 하지만 희망 판매가와 희망 구매가가 맞지 않을 경우 거래가 체결이 무한정 지연될 수 있고, 거래가 체결되기 전 까지 제시된 거래 희망가와 물량 정보를 계속 저장 및 추적해야 하므로, 용량이 한정된 블록체인 환경에 적절치 못함.
- 대신에, 현재 거래소에 예치된 유동성 양에 따라 가격이 자동 결정되는 AMM 방식이 많이 사용됨.
- 유동성 제공자는 유동성 (토큰 A 및 토큰 B) 를 거래소에 위탁함. 사용자가 거래소에 거래를 요청하면, AMM 알고리즘에 의해 결정된 가격으로 즉시 거래가 체결된다.
- 장점: 가격이 알고리즘에 의해 자동 결정되며, 거래가 즉시 체결됨.
- 단점: 거래량에 따라 거래 가격이 달라지는 price impact 현상, 거래 요청 시점과 거래 체결 시점의 가격이 달라지는 slippage 현상이 일어날 수 있음

DEX

Automatic Market Maker (AMM) 개념 설명

- 가장 유명한 AMM 알고리즘은 Prediction market에서 주로 사용된 Logarithmic market scoring rules
- 그러나 이는 리소스가 많이 필요해 블록체인 상에서 적절하지 않음.
- 블록체인상에서는 주로 Constant Function이 사용됨
- 유니스왑과 같은 주요 DEX들에서는 Constant Product function 을 사용한 AMM (CPMM) 을 사용함

CPMM 예제

- 토큰 A 와 토큰 B가 예치되어 있는 유동성 풀(pool) 을 가정해보자.
- x 가 토큰 A의 예치량이고, y 가 토큰 B의 예치량이라고 하면, 거래 전후로 두 토큰의 예치량의 곱 k 가 유지되도록 가격을 결정한다. 즉,

$$x \cdot y = k$$

DEX

DEX 스왑 예제: 토큰 A 가 10개, 토큰 B 가 20 개 예치된 풀에 토큰 A를 1개를 판매하면 토큰 B를 몇 개 받을 수 있을까?

sol)

$$x = 10,\ y = 20$$

- 받을 수 있는 토큰 B의 양을 y_{out} 이라고 하면, 스왑 거래가 체결된 이후 풀의 예치량은 다음과 같음

$$x' = (10+1),\ y' = (20 - y_{out})$$

- 스왑 전후로 예치량의 곱이 일정해야 하므로,

$$10 \times 20 = 11 \times (20 - y_{out})$$

$$y_{out} = 20 - \frac{200}{11}$$

$$= 1.82$$

- 즉, 1.82 개의 토큰 B를 받을 수 있다.

DEX

Price Impact 예제

- 앞선 예제에서, 예치된 토큰 A 및 토큰 B의 비율은 1:2 이다. 직관적으로 생각했을 때, 토큰 A를 1개 지불하면 토큰 B를 2개 얻을 수 있어야 하는데, 실제 결과는 1.82 개만 얻을 수 있었다. 이는 예치량 대비 거래량이 너무 많아서 일어나는 price impact 현상이다.
- 같은 예제에서, 토큰 A 의 초기 예치량이 1000개, 토큰 B의 초기 예치량이 2000개라고 가정해보자.

스왑 전: $x = 1000,\ y = 2000$

스왑 후: $x' = 1001,\ y' = 2000 - y_{out}$

$y_{out} = 1.998$ 개의 B 토큰을 받을 수 있다.

- 즉, 풀의 예치량이 많을 수록 거래 가격은 예치량 비율과 비슷해지는 것을 확인할 수 있다.

DEX

슬리피지 예제

- 앞선 스왑 예제에서, $x = 10, y = 20$ 인 토큰 A/B 풀에서, 토큰 A를 1개 지불하여 토큰 B를 1.82 개 얻을 수 있음을 계산할 수 있다.
- 이 계산에 기반하여, 사용자가 토큰 B를 1.82개 얻을 것을 예상하고 스왑 거래를 등록할 것이다.
- 하지만, 네트워크 딜레이, 혹은 먼저 제출된 스왑거래로 인해 사용자가 거래 할 때의 실제 교환비가 달라질 수 있다.
- 실제 교환 결과로 사용자가 예상보다 많은 토큰을 얻는다면, 아무런 문제가 없으므로 그대로 실행해도 된다.
- 하지만 실제 교환 결과값이 예상 결과값 보다 적다면, 손실이 일어나게 된다.
- 사용자는 이러한 손실 (슬리피지) 의 허용 한도를 설정하여, 일정 이상의 슬리피지가 일어날 경우 스왑 거래 자체를 취소하도록 할 수 있다.
- 단 슬리피지 허용 오차를 너무 작게 설정할 경우, 스왑 거래가 지연될 수 있다.

DEX

수수료

- 위와 같은 DEX 가 원활히 운용되려면, DEX에 예치된 유동성이 많아야 한다.
- 유동성을 확보하기 위해서 유동성 제공자에게 보상을 주어야 한다.
- 이에 더해, DEX 운영 비용도 충당해야 한다.

해결법: 거래마다 수수료를 징수하여 유동성 제공자에게 보상(유동성 제공 수수료) 및 DEX 운영진에게 지불 (프로토콜 수수료) 한다.

- Uniswap 의 경우 사용자에게 입력 토큰의 0.3%를 수수료로 징수하여 모두 유동성 제공자에게 보상으로 지불한다. (현재 프로토콜 수수료는 없음.)
- Pancakeswap 의 경우 사용자에게 입력 토큰의 0.25%를 수수료로 징수한다. 그 중 0.17% 는 유동성 제공자에게 보상하고, 나머지 0.03%는 프로토콜 수수료, 0.05% 는 CAKE (거버넌스 토큰) 의 소각에 사용.

DEX

유동성 제공 및 제거

- 유동성을 제공하면, 현재 풀의 예치량 대비 지분 비율량에 해당하는 LP 토큰을 얻게 된다. 자신의 지분 비율은 (자신의 LP토큰량)/(LP토큰의 총 발행량).
- 제공중이던 유동성을 제거하면, 현재 풀의 총 토큰중 자신의 지분 비율 만큼 출금한다. LP토큰은 소각한다. 유동성 중 일부만 제거하는 것도 가능.
- 유동성을 제공하고 있는 도중, 스왑 거래가 일어날 경우 유동성 제공 수수료가 풀에 적립되어 풀의 전체 가치가 상승한다.
- 제공중이던 유동성을 제거할 경우 풀 가치의 상승비율 만큼의 차익을 실현할 수 있다.

DEX

비영구적 손실

- 스왑 수수료가 풀에 계속 적립되므로, 토큰의 가치가 일정하게 유지된다면 유동성 제공으로 차익실현이 가능함
- 하지만 토큰의 가치가 급격히 변할 경우, 유동성 제공으로 인해 손실을 입을수 있음
- 이 손실은 토큰 가치가 이전 수준으로 다시 복구될 경우 회복될 수 있음
- 이를 비영구적 손실이라 지칭함

DEX

비영구적 손실 예제

- 토큰 A가 10개, 토큰 B가 20개 예치된 풀을 가정하자. ($x = 10, y = 20, xy = 200$)
- 그 중 토큰 A 1개, 토큰 B 2개는 자신이 제공한 유동성이다. 즉, 풀의 지분 10% 를 가지고 있다.
- 토큰 A의 가격이 $10, 토큰 B의 가격이 $5 라 가정하자. 그렇다면 풀에 예치된 내 자산의 총 가치는 ($10 * 1) + ($5 * 2) = $20 이다.

이 때, 토큰 A의 가격이 $1로 폭락한다고 생각해보자.

- 만약 내가 유동성 제공을 하지 않고, 토큰 A 1개와 토큰 B 2개를 그대로 가지고 있었다고 가정하면, 내 자산의 총 가치는 ($1 * 1) + ($5 * 2) = $11 가 되었을 것이다.
- 폭락 직후, 풀의 유동성을 즉시 제거한다면 내 자산의 총 가치는 여전히 $11일것이다.

하지만 토큰 A가 폭락한다면, 많은 사람들이 스왑 거래를 통해 토큰 A를 팔고 토큰 B를 사갈 것이다.

- 풀의 예치량 비율이 가격 비율 (1:5) 에 가까워 질 때 까지 교환이 진행될 것이므로, 평형 상태에 도달한 풀의 최종 예치량은 다음과 같을 것이다.

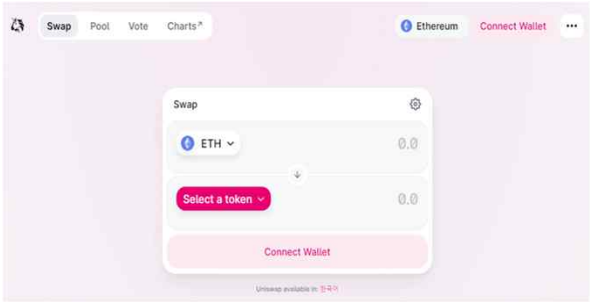
$x = 31.56, y = 6.32$ (xy=200, 수수료가 없다고 가정)

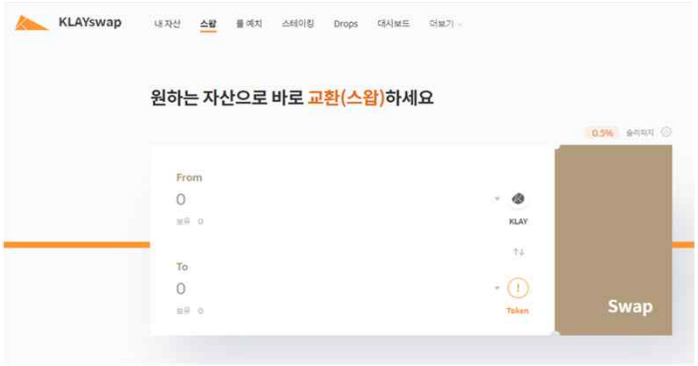
- 이 풀에 대한 내 지분은 10% 이므로, 나의 자산은 토큰 A 3.156 개 및 토큰 B 0.632개이다. 총 자산 가치는 ($1 * 3.156) + ($5 * 0.632) = $6.32 가 되었다.
- 즉, 유동성 제공을 하지 않았을 때의 자산가치에 비해 $4.68 의 손실이 발생하였다.

DEX 서비스 예시

유니스왑: https://app.uniswap.org/#/swap

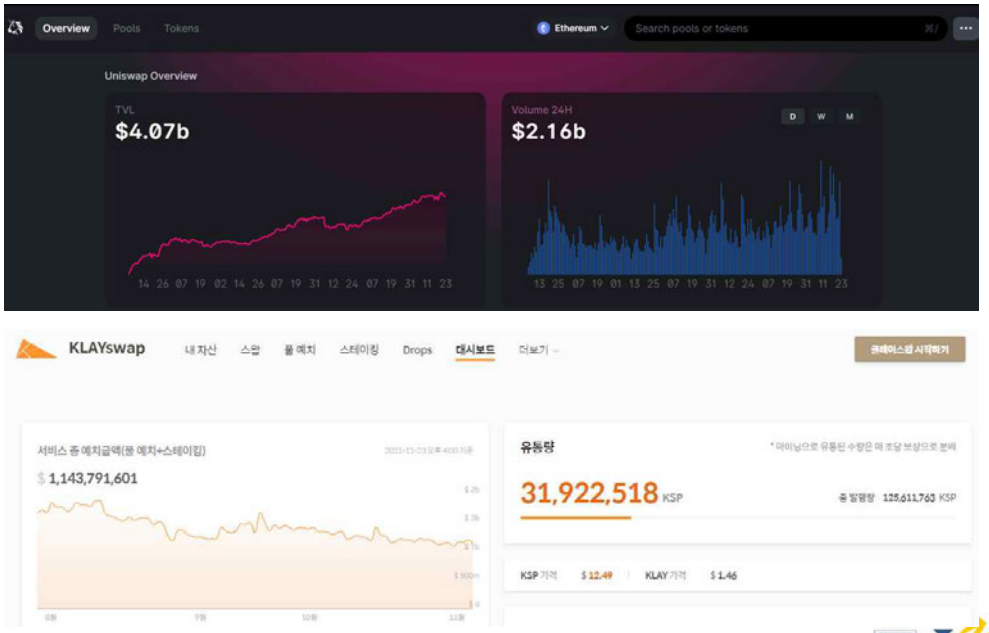
클레이스왑: https://klayswap.com/exchange/swap





DEX 서비스 예시

총 예치량 (Total Value Locked; TVL) : 40억 USD (Uniswap), 11억 USD (Klayswap)



Flash Loan

사용자가 유동성 풀에 토큰 A를 입금하고, 토큰 B를 출금하는 시나리오에서:

- 사용자가 풀에 토큰 A를 입금한 후, 그에 상응하는 가치의 토큰 B를 출금해가는 과정이 하나의 트랜잭션에서 처리됨.
- 하지만, 사실 하나의 트랜잭션에서 처리되기만 한다면, 사용자가 풀에서 토큰 B를 출금한 이후, 그에 상응하는 가치의 토큰 A를 풀에 입금 해도 문제 없음.
- 여기에 더해, 사용자가 토큰 B를 출금하여, 다른 탈중앙 서비스를 통해 토큰 A를 얻은 이후, 그 토큰 A를 풀에 상환할 수 도 있음.
- 즉, 사용자가 토큰 A를 전혀 가지고 있지 않은 상황에서도, 토큰 B를 빌려 차익거래로 토큰 A를 얻어 상환하는 대출이 가능함.

단, 대출과 상환을 하나의 거래로 묶어서, 상환에 오류가 발생하면 대출또한 무효가 되도록 해야함.

하나의 거래 안에 대출과 상환이 동시에 이루어지므로 플래시론 이라는 이름이 붙음.

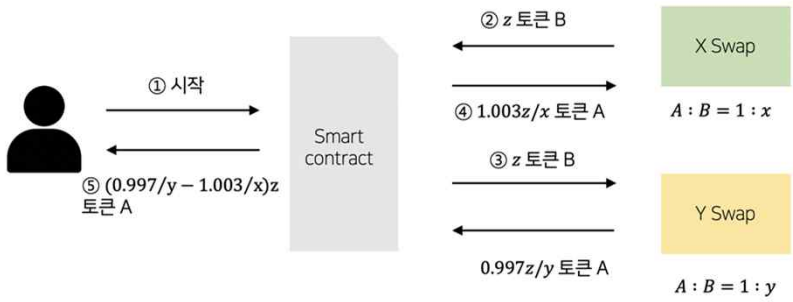
플래시론을 사용하면 아무런 자산이 없는 상태에서도 차익거래가 가능함. 차익거래가 쉬워지므로 여러 DEX 간의 가격안정화에 도움이 될 수 있음.

Flash Loan

플래시론 예시

X Swap과 Y Swap을 통해 토큰 A/토큰 B 풀에 대해 플래시론으로 차익거래를 시도하는 시나리오에서:

- A와 B의 교환 비율은 X Swap에서 1:x, Y Swap 에서 1:y 로 가정
- 가격 영향(Price Impact)을 고려하지 않고 플래시론과 스왑 수수료는 0.3%로 가정
- 먼저 X Swap에서 토큰 B를 z만큼 빌려, 이를 Y Swap에서 토큰 A로 교환한다. 이 때 얻을 수 있는 토큰 A의 개수는 $0.997z/y$ 이다.
- 이를 X Swap에 상환한다. 이때 상환은 토큰 A를 이용해 이뤄지며, 수수료를 포함하여 $1.003z/x$개 만큼의 토큰 A를 상환해야 한다.
- 사용자는 $0.997z/y$ - $1.003z/x$ 개의 토큰 A 만큼의 차익을 얻는다.



Flash Loan Attack

플래시론 기능 자체는 순기능이 있으나, 플래시론을 악용한 공격이 가능함.

- 해커가 플래시론을 사용하여 대량의 암호화폐를 확보함.
- 이를 사용해 암호화폐 시세를 순간적으로 조작하여 이득을 얻음.

최근 플래시론 공격 예시: Pancake BUNNY 공격 (2021.05.19, 바이낸스 체인)

Pancake BUNNY: 암호화폐 예치량에 따라 보상으로 BUNNY 토큰을 지급하는 DeFi 서비스.

- 해커는 Pancake BUNNY가 보상을 산정할 때 Pancakeswap 의 현재 가격을 참조하는 취약점을 공격함.
- Pancakeswap 및 다른 플래시론 프로토콜에서 대량의 코인을 빌려 Pancakeswap 의 BNB-USDT 가격을 조작함.
- Pancake BUNNY 서비스에 토큰을 예치하고, 조작된 가격에 기반하여 막대한 양의 BUNNY 토큰을 발행받음.
- 얻은 BUNNY 토큰을 즉시 처분하여 플래시론을 상환하고 차익을 탈취함.
- 그 결과 BUNNY 토큰의 가격이 폭락함.

DeFi Lending 프로토콜 개념

중간자 없이 담보 대출 서비스 제공

기존 암호화폐 대출 서비스의 문제점:

1) 중앙화된 서비스: 신뢰하기 어려움
2) P2P 대출 프로토콜: 대출과정이 번거롭고 유동성이 극히 제한됨

해결책

1) 이용자들이 자산을 예치하고 이자를 보상받을수 있는 머니 마켓 구축함으로써 유동성 확보
2) 머니 마켓의 예치량에 따라 프로토콜에 의해 자동 계산된 이자율로 즉시 대출
3) 모든 과정이 스마트 컨트랙트 및 거래내역상에 투명하게 공개

DeFi Lending 프로토콜 예제 (Compound)

이자율 자동 계산

- incentive 모델: 대출량이 많을수록 이자율을 높여서 대출억제 및 예치 유도
- 모든 사용자에게 동일한 이자율 적용
- 대출, 예치량이 변화할 때 마다 재계산

$$U_a = Borrows_a \,/\, (Cash_a + Borrows_a)$$

$$Borrowing\ Interest\ Rate_a = 2.5\% + U_a * 20\%$$

DeFi Lending 프로토콜 예제 (Compound)

대출 담보 설정 및 청산

- 대출자는 대출받을 암호화폐 이상의 가치를 가지는 다른 암호화폐를 담보로 설정하여야 함. 최대 담보율은 프로토콜의 거버넌스에 의해 결정됨.
- 가치 책정은 상위 10개 거래소를 기준으로 한 가격 오라클에 의해 수행됨
- 암호화폐의 가격 변동으로 인해 담보가치가 일정 이상 줄어들면, 담보물의 일부를 청산하여 파산을 방지할 수 있음.
- 이 때 담보물을 구입하려는 청산자는 할인된 가격으로 구입가능함.

DeFi Lending 프로토콜 예제 (Compound)

대출 담보 설정 및 청산 예제

- ETH 가격이 $5000, DAI 가격이 $1 라고 가정.
- ETH의 최대 담보율이 50%라고 가정. 즉, 1 ETH를 담보로 최대 2500 DAI를 대출가능.

1) 사용자가 1 ETH를 담보로 2500 DAI를 대출함
2) 그런데 ETH 가격이 $4500 으로 하락하면, 대출자의 담보가치는 $2250 인 데 반해 대출가치는 $2500 인 부실대출로 변함.
3) 이 때, 청산자가 담보물인 ETH 중 일부를 머니 마켓으로부터 구입 가능함
4) 청산자가 0.2 ETH를 약 10% 할인된 가격 810 DAI로 구매함
5) 결과적으로 대출자는 0.8 ETH 를 담보로 1690 DAI를 대출한 것이 되며, 환산 담보가치 (0.8 ETH * $4500 * 0.5 = $1800) 대비 대출가치 $1690 으로 건강한 대출로 전환됨.

Stable coin (DAI) 개념 소개 및 예시

다른 화폐 또는 실물 자산 가격과 연동하여 코인의 가격이 거의 변동하지 않고 안정된 암호화폐

법정화폐 또는 다른 암호화폐를 담보로 잡거나 알고리즘에 의해 공급량 조정을 통해 가격이 안정됨.

법정화폐 (USD)를 담보로 발행함으로써 가치 유지: USDT, USDC

이더리움을 담보로 발행 및 가격 오라클에 기반하여 가치 유지: DAI

수요와 공급을 조절해 가격안정성 유지: Terra

Gwangju Institute of
Science and Technology

School of Electrical Engineering and Computer Science

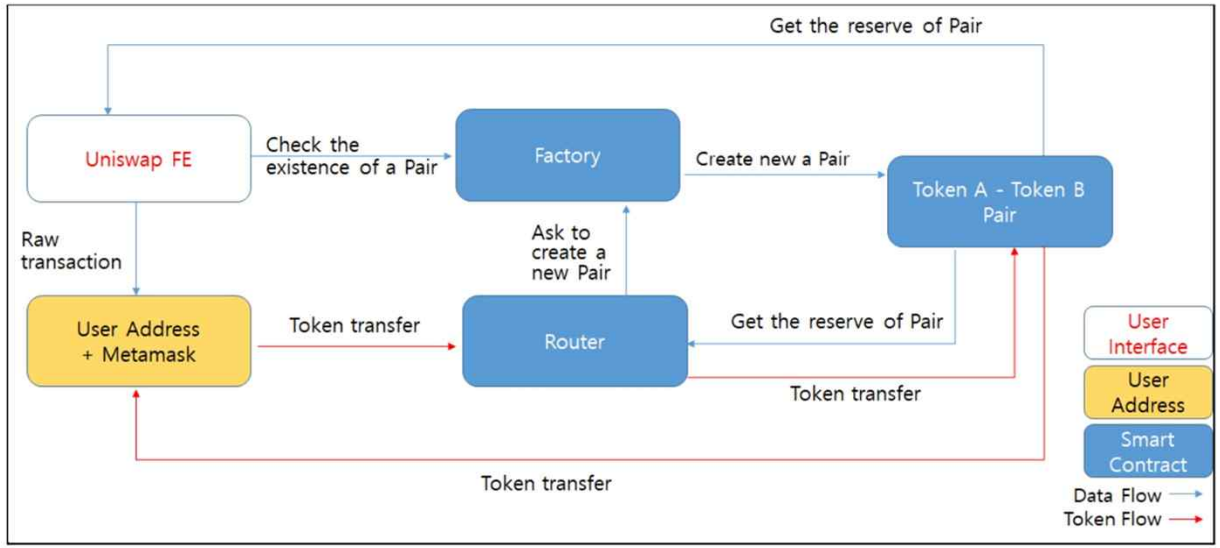


SWAP 실습 (GIST swap)

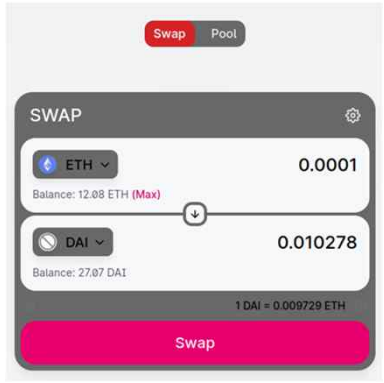
실습 목차

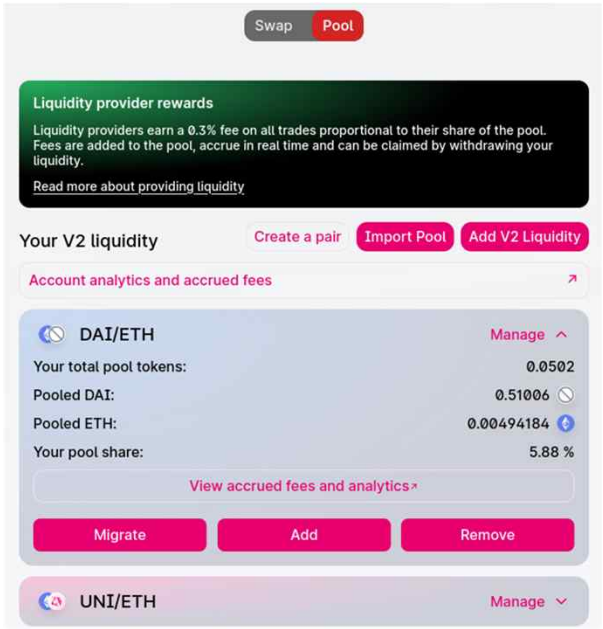
- 실습 전 준비사항: 메타마스크 설치
- 실습 전 준비사항: 테스트 이더 분배
- GIST swap 기능 및 구조 상세
- GIST swap 인터페이스 살펴보기
- 주요 기능Swap/Add Liquidity/Remove Liquidity 실행해보기
- 이더스캔 상에서 확인해보기

GIST swap 기능 및 구조 상세



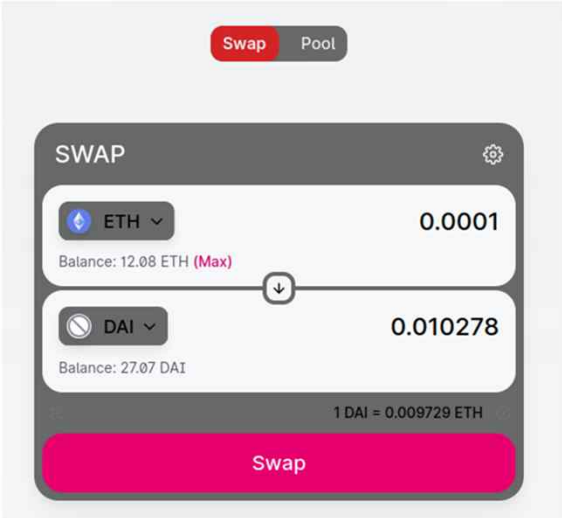
GIST swap 인터페이스 살펴보기





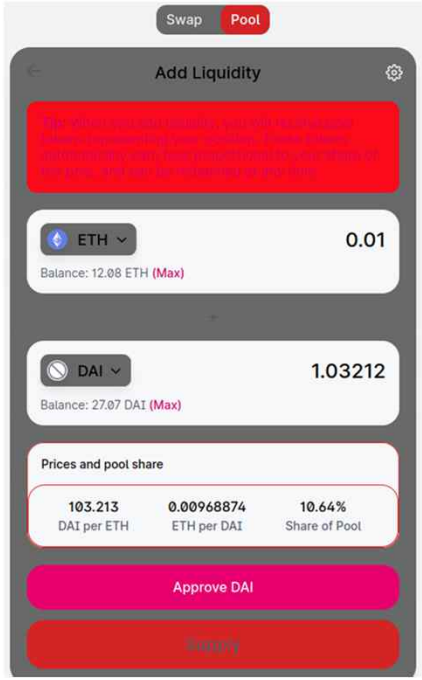
Swap/Add Liquidity/Remove Liquidity 실행해보기

Swap



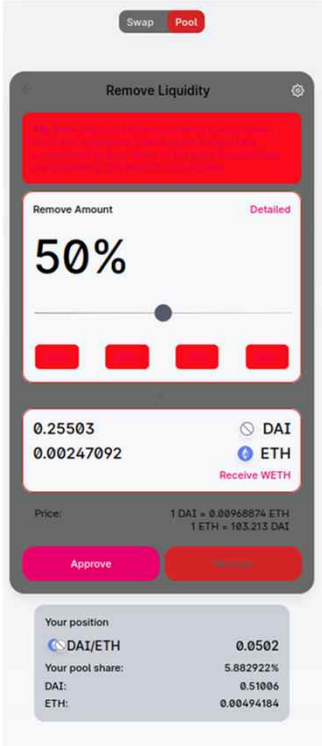
Swap/Add Liquidity/Remove Liquidity 실행해보기

Add Liquidity



Swap/Add Liquidity/Remove Liquidity 실행해보기

Remove Liquidity



이더스캔 상에서 확인해보기

Router:

https://ropsten.etherscan.io/address/0x073b662f25d22a1cda1e7a7100bffd6aef37d394

Test token address (INFONET):

Code: https://github.com/Haeung9/myERC20.git

Deployed: 0xD23E2320cA33178Fd6651A99e61cBBC1612909FF

https://ropsten.etherscan.io/address/0xD23E2320cA33178Fd6651A99e61cBBC1612909FF

Consensus Mechanisms in Major Cryptocurrencies

Lecture in B-Capitalist Course



이흥노, GIST, South Korea
Home page: http://infonet.gist.ac.kr
Facebook/Publication ID: Heung-No Lee
E-mail: heungno@gist.ac.kr

Agenda of this talk

- 강사소개
- Bitcoin Consensus
- Other Consensuses

이흥노 교수



GIST **블록체인 인터넷 경제연구센터 / 센서지능화센터 센터장**

LiberVance 리버밴스㈜ 설립자

GIST 이흥노

국가스마트도시위원회 민간위원, (사)대한전자공학회 통신소사이어티 회장
우정사업본부 블록체인이니셔티브 민간위원

연구실적

- **국제 논문 (SCI) : 200+편(71+편)**
- 국제 특허 : 출원 9건, 등록 7건
- 국내 특허 : 출원 7건, 등록 22건
- **기술이전 : 2건 (2억 1천만원)**
- 저술 : 8건
- 기술가치평가 : 1건(중앙기술평가원)

대표언론보도

- 프레시안, "GIST, 글로벌 블록체인 개발회사 '컨센시스' 와 MOU" (2019.07.21)
- 스마트시티랩, "블록체인 스마트시티의 유기적 연결의 중추 역할 할 것" (2018.08.24)
- 전자신문, "대한전자공학회 통신소사이어티(회장 이흥노), 19일 블록체인으로 여는 미래' 워크숍 개최" (2018/04/16)
- 블록체인허브, "블록체인·AI·빅데이터 융합산업 육성 정책토론회", (2018/12/06)

연구분야

- 블록체인경제
- 지능형 네트워크
- 센서지능화

관련수상이력

- **2019년 대한전자공학회 해동학술상 수상**
- 2016년 GIST 연구상
- **2016년 GIST 대표기술상**
- **2014년 1월 이달의 과학기술자상, 미창부**
- 2013년 기초연구 우수성과 50선, 미창부
- **2012년 국가연구개발 우수성과 100, 미창부**

© 2019 GIST

보유기술현황

보유기술

- GIST는 블록체인의 재중앙화 문제 해결을 위한 **부호-암호 기반의 작업증명** 방식 연구 및 특허 확보 및 출원한 **부호-암호 작업증명 특허**
  (GIST IP)를 기반으로 블록체인 시스템 개발

특허

특허	출원인	출원번호/등록일	핵심기술
부호-암호 화폐 시스템	광주과학기술원	10-2019-0151246 / 2019.11.12	블록체인의 재중앙화 문제 해결을 위한 핵심 기술
블록체인거버넌스	광주과학기술원	10-2019-0084800 / 2019-07-12	블록체인 거버넌스와 규제
블록체인의 거래검증시스템, 및 블록체인이 거래검증방법	광주과학기술원	10-2019-0120655 / 2019-09-30	블록체인거래검증

© 2019 GIST

Consensus

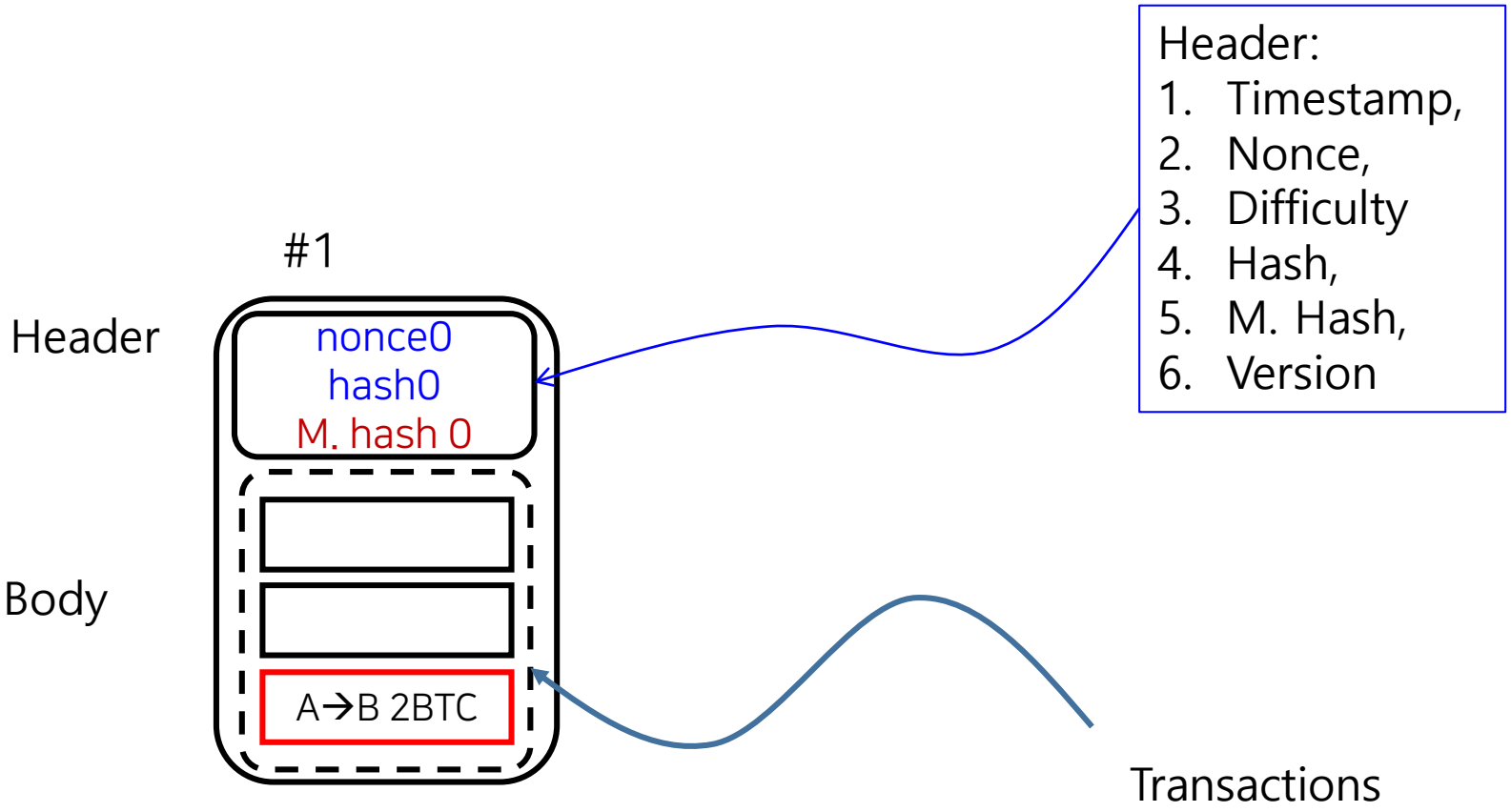
- **Goal: 하나의 동일한 원장 유지 관리**

- Nodes: 글로벌 영역에 분산되어 있는 노드 N개
  1. State: 각 노드의 현재 상태
     A. Work: (각자) 새로운 TXs 담아 블록 생성
     B. Announce: 발표
     C. Inspection: 검사
     D. Approval: 승인
  2. State Update: 새로운 블록을 기존 블록체인에 순서대로 추가

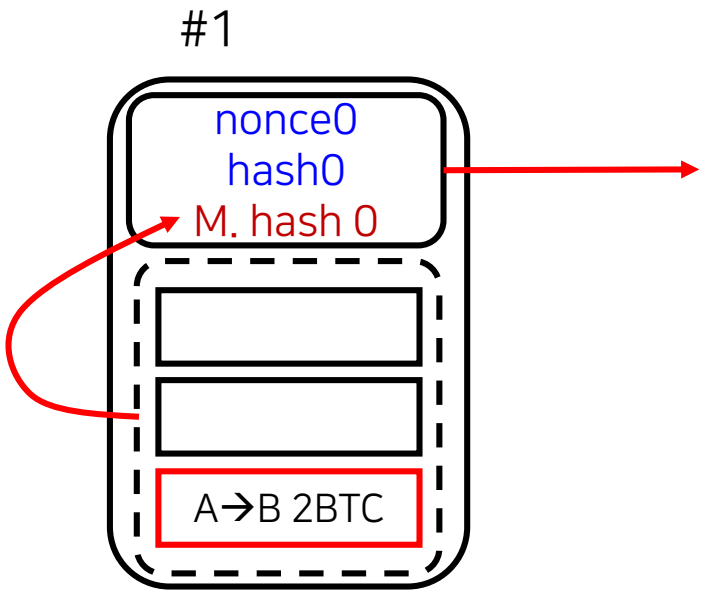
Consensus Algorithm설계의 어려움

- 노드들은 인터넷 상에서 협력하여 합의에 도달해야 함
- OS SW
- Cryptocurrency의 경우 다양한 공격 방법이 존재
  - Sybil Attack
  - DDoS Attack
  - Byzantine faults

Block: Header 와 Body



Secure Hash Function: 블록요약을 만드는 함수



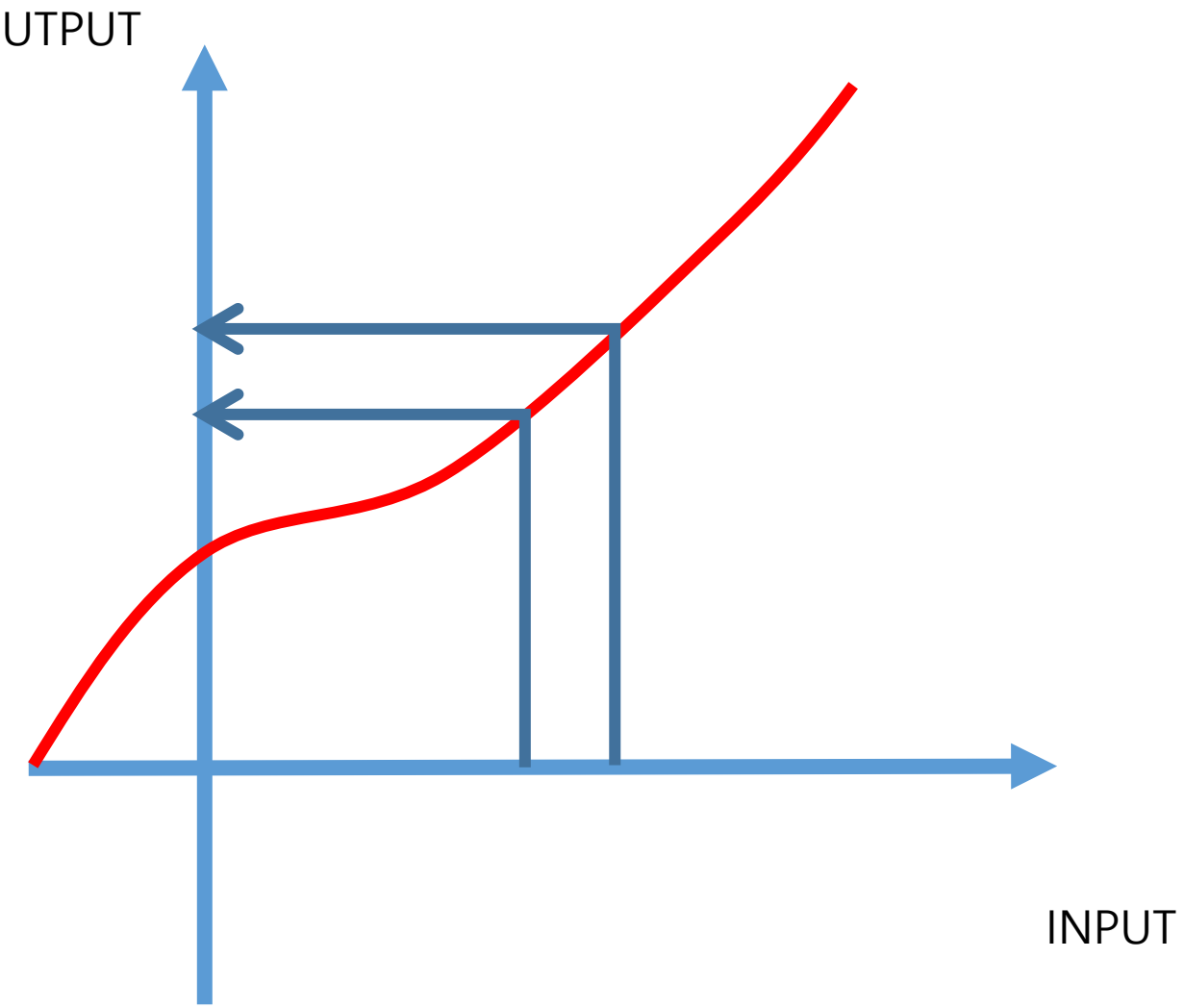
특징: Oneway 앞으로만 가는 함수.
뒤로 가려면 대입법 밖에 없다.

함수는 INPUT 을 받고 OUTPUT을 만든다
INPUT: 파일
OUTPUT: 256 bit 숫자

쓰이는 곳
1. 은행 서버는 Passwd의 hash값만 저장해 놓는다.
2. Spam mail 방지
3. PoW

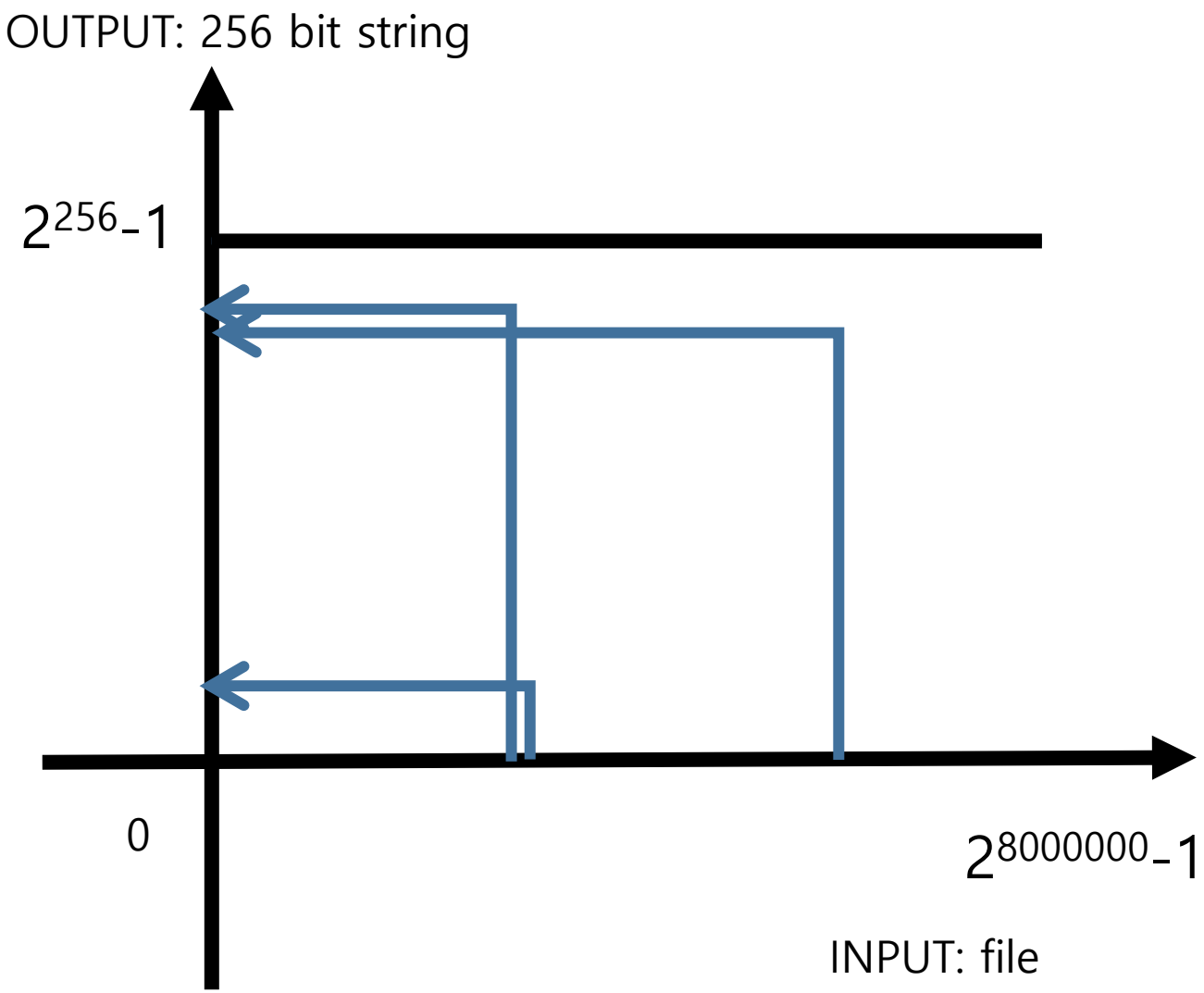
일반 함수

- Output을 본 후,
  Input 예측 가능함



SHA 함수

- 동일한 Input은 언제나 동일한 Output 출력.
- Input이 조금만 차이가 나도, 마치 무작위로 선택된 것 같은 Output을 줌.
- 결과: Output 값을 얻은 후 Input값을 전혀 예측 할 수 없게 됨.
- Many to 1 함수, but collision free!



Modulo Operation

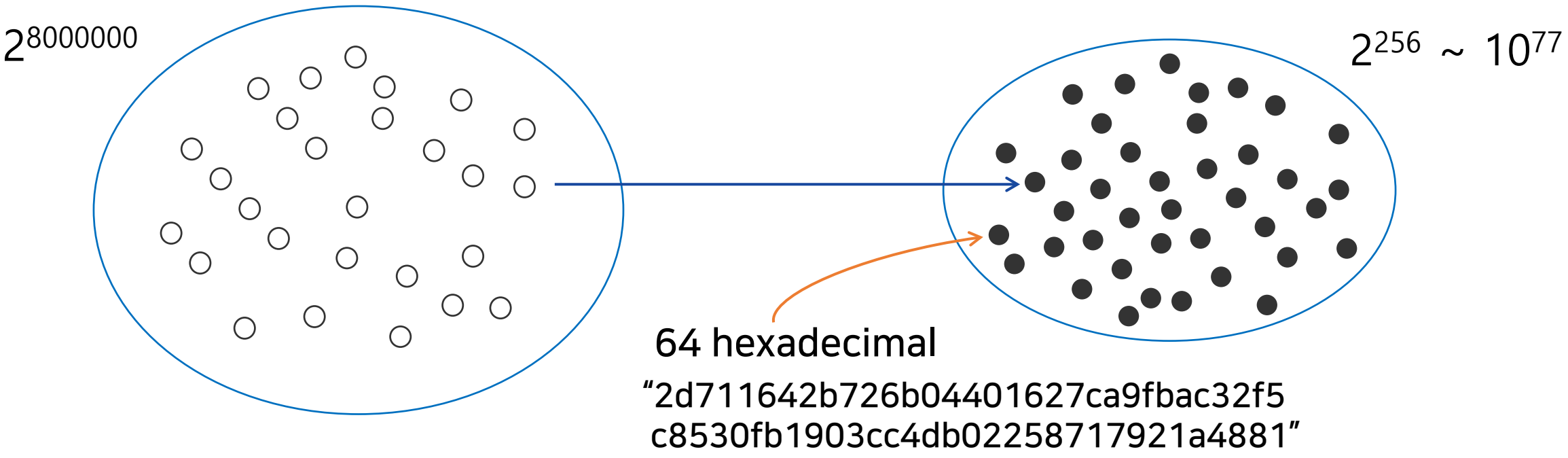
- F(x) = y
- F(x) = 7x + 6 mod 5

- Let x = 5. Then, y = 7*5 + 6 = 41 mod 5 = 1.
- Let x = 4. Then, y = 7*4 + 6 = 34 mod 5 = 4.

Hash 함수

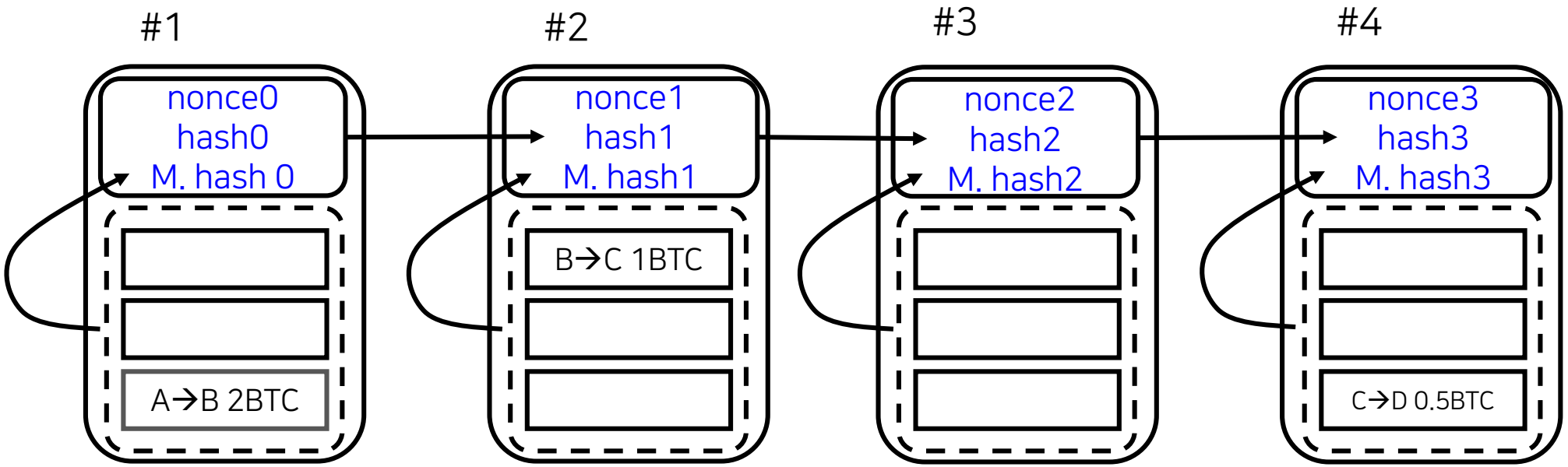
- SHA256, $F(x) = y$

$X = \{x | x \text{ is a message up to 1 Mbyte in size}\}$ $Y = \{y | y \text{ is a 256bit string}\}$

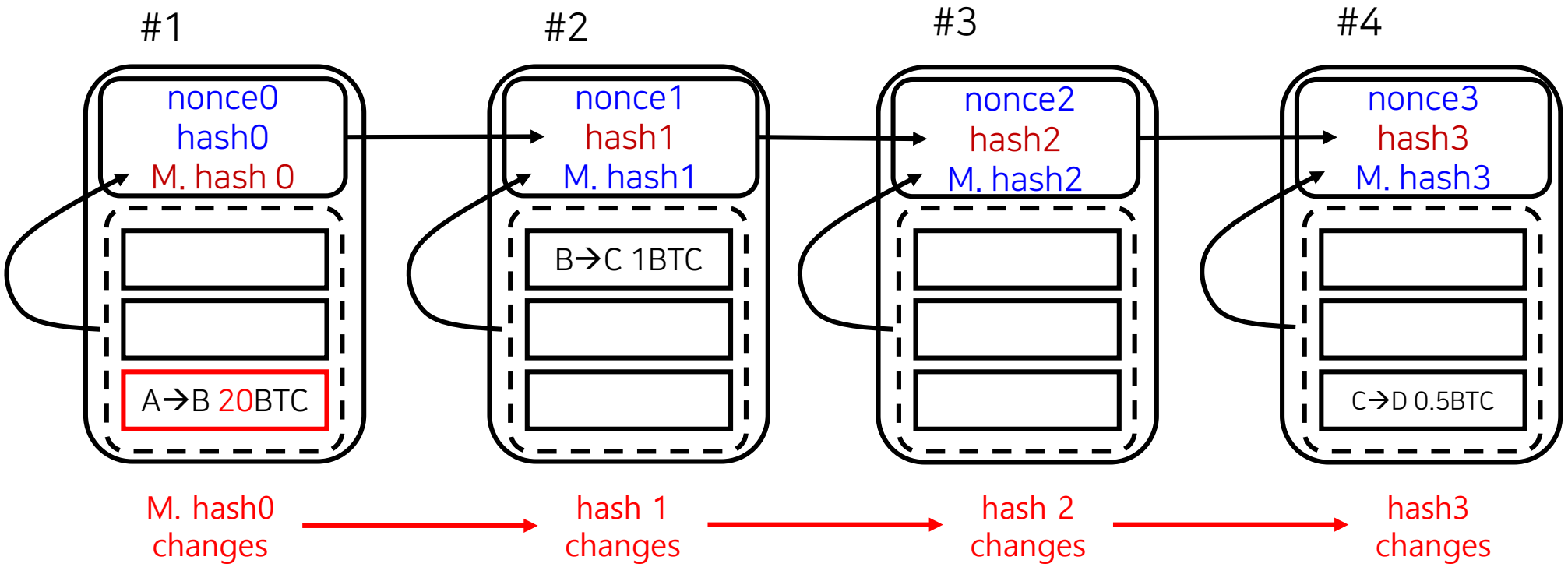


Blockchain이란?

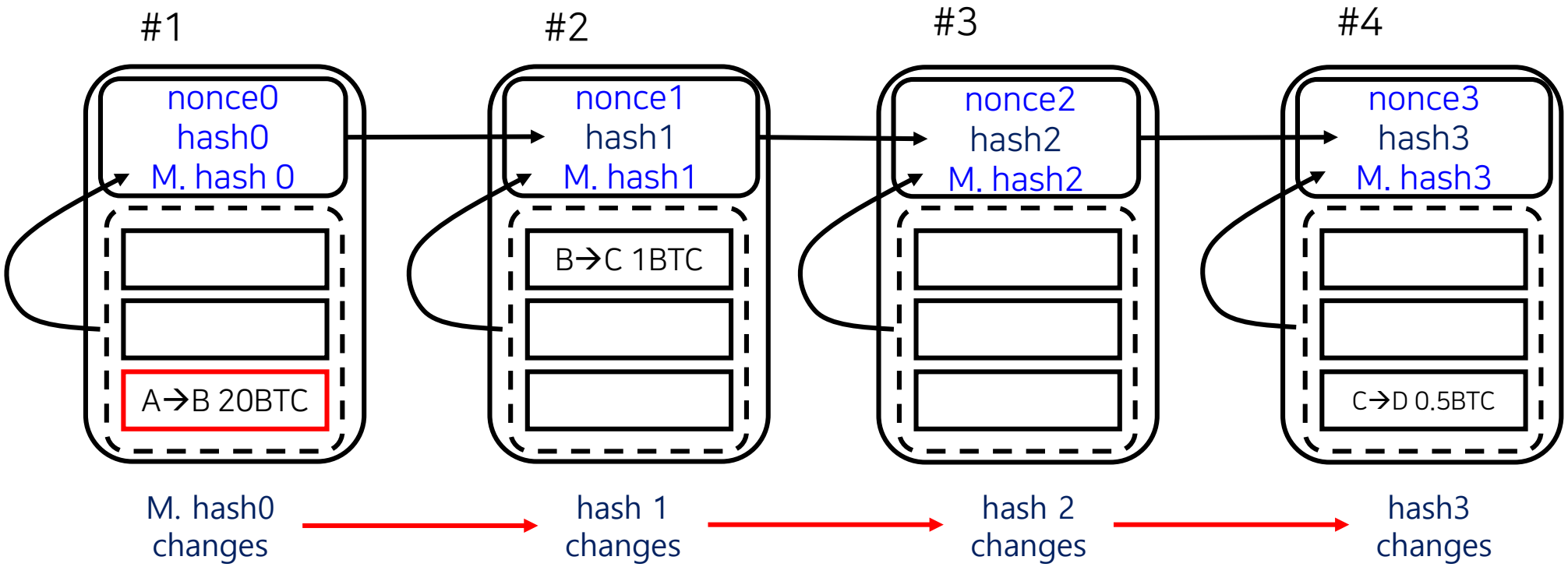
- Block body의 hash값을 Block header에 넣고
- Oneway function으로 block 과 block을 연결하는 것



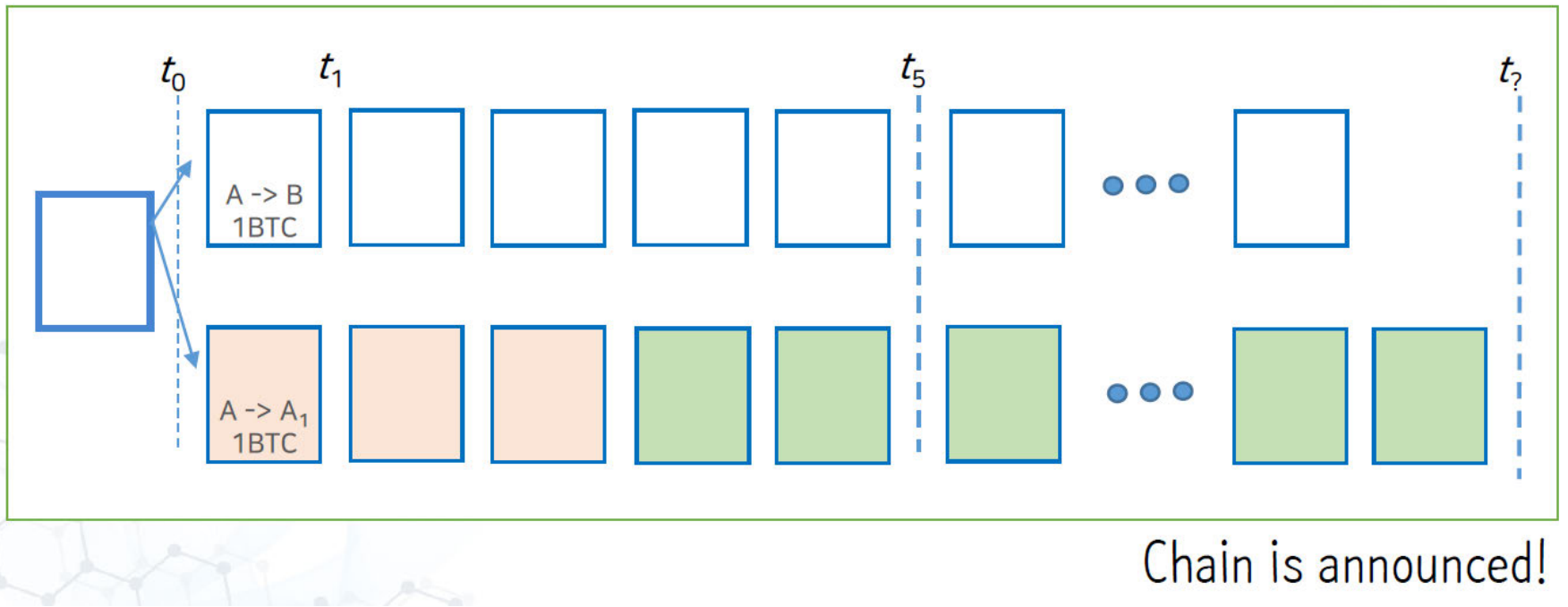
Block을 Chain으로 연결하여 얻는 효과는?



Block을 Chain으로 연결하면 컨텐츠를 못 바꿀까?



Double Spending은 어떻게 성립하는가?
방지하려면 무엇이 필요한가?

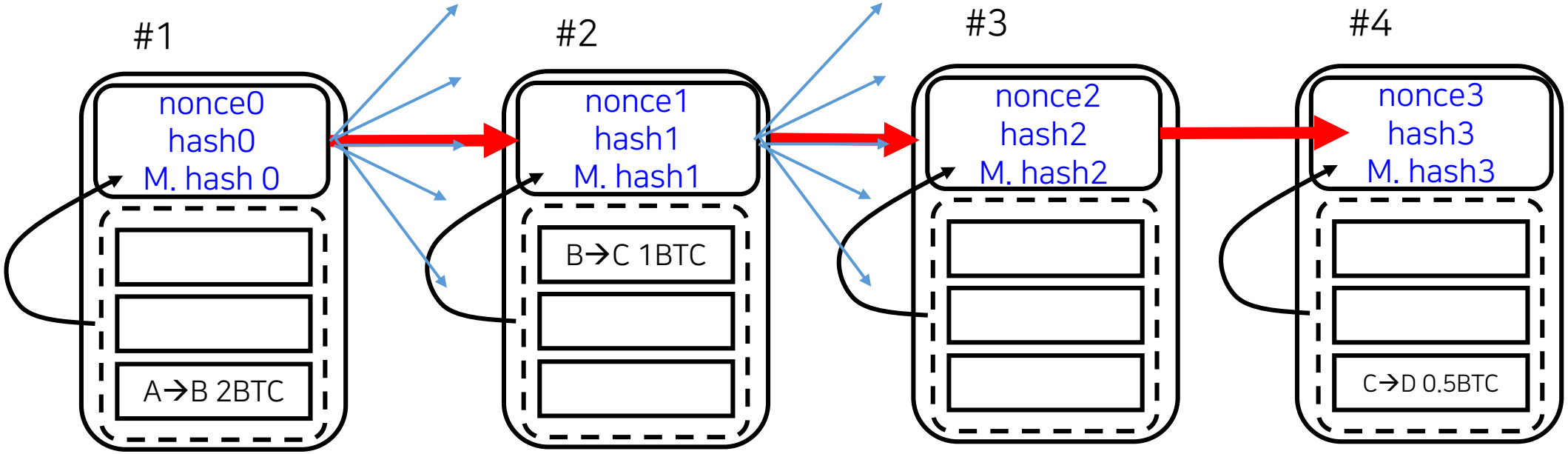


정상 체인

공격자 체인

자세한 내용은 이 링크를 눌러 확인하세요 https://arxiv.org/abs/1903.01711

컨텐츠를 바꿀 수 없게 하려면 무엇이 필요한가?



- Block Chain + PoW
  - Revolutionary new idea!
    - Al-Im-To-Po Theory!
    - The more come to get involved, the safer the network becomes!
    - Reward
    - Fresh new race for each block.
    - Race means competition.

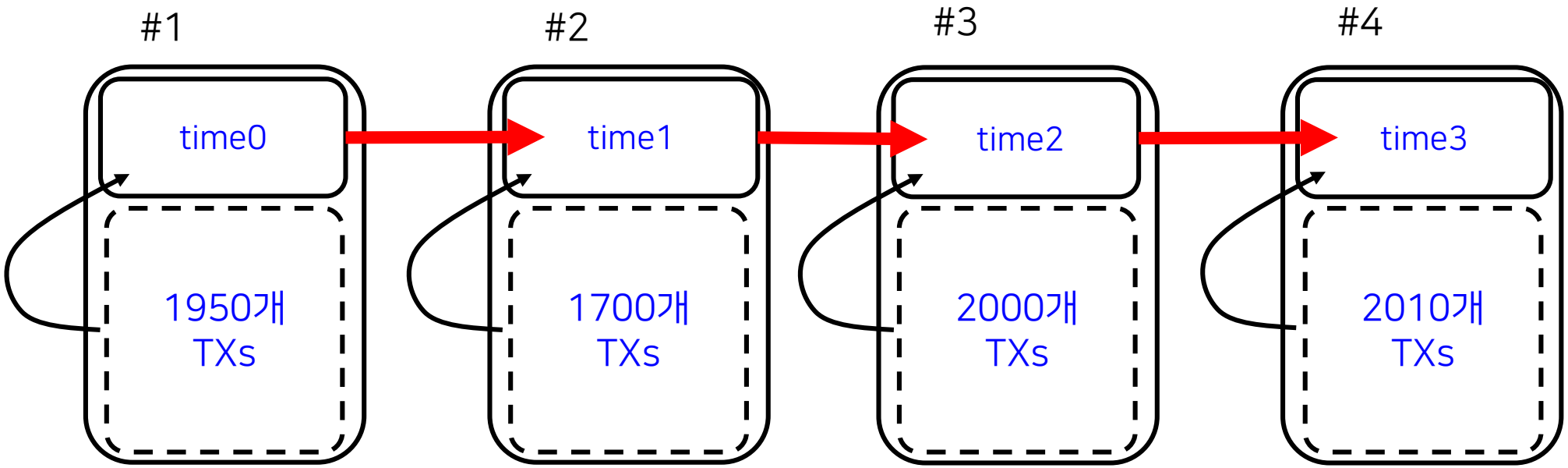
Block #613573

BlockHash 00000000000000000000288ae77f5b8b3c99784f55827431a23f524adf69d0e3a

Summary

Number Of Transactions	3170	**Difficulty**	14776367535688.639
Height	613573 (Mainchain)	**Bits**	17130c78
Block Reward	12.5 BTC	**Size (bytes)**	869316
Timestamp	Jan 20, 2020 1:15:17 AM	**Version**	1073733632
Mined by		**Nonce**	3774951927
Merkle Root	523c52ad89c6ac67bcc01a6...		
Previous Block	613572		

Block Interval, TXs/Block, TPS



TPS = TXs/Block / BI
= 2000 TXs/600초
= 3.3 TXs/sec

TPS제고 필요성 제기

- 인기가 많은 블록체인에만 해당
- TPS를 높이려면?
  1. 블록사이즈를 키우거나
  2. 블록인터벌을 줄이거나
  3. Both

블록인터벌 **BI** 을 줄이려면?

- 노드 수를 줄이는 프로젝트들이 성행
  - 적은 수의 노드에서 잘 돌아가는 합의알고리즘 사용

  - 그러나 노드 수가 적으면
    - 탈 중앙화는????
    - DS 공격 취약성은???
    - Sybil Attack 취약성은???

블록인터벌 **BI** 을 줄이려면? (2)

- PoW도 Block Interval을 줄일 수 있음
  - Litecoin: 2.5분
  - ETH: Block Interval 14초

  - 수 많은 fork가 매 번 발생하는 문제 발생
    - DS 공격 취약성은???
    - 그러나, Profitable DS Attack분석으로 해결 가능 (See DeSecure블록체인)



PoW Puzzles

- Making PoW puzzles
  - Bitcoin uses SHA256.
  - SHA is *oneway* and *collision free*.
  - Difficulty and Nonce are in the BH.

2 PoW Puzzles

- Finding Good Block Summary
  - Function F takes input x and gives output y :

$$y = F(x)$$

  - x is block header (BH), i.e., F(BH) = hash.
  - Then, it can be written as

F(BH: nonce) < Target PoW Ineq.

  - For a block, find a nonce that satisfies the above inequality (Work)
  - Record the nonce in the block header. (Proof)

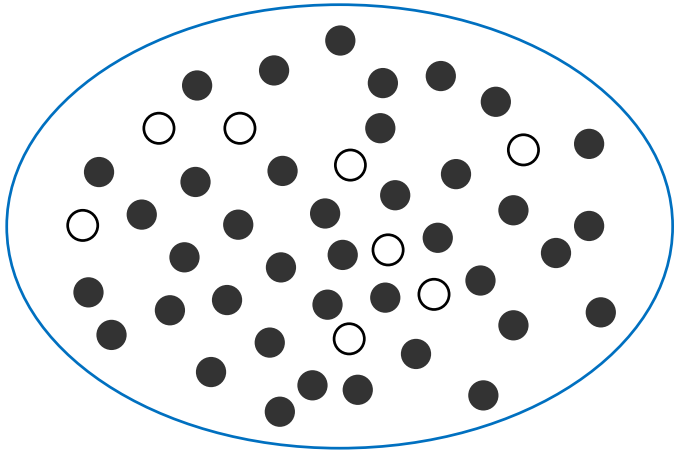
2 PoW Puzzles

- Toy puzzle
  - White and black balls.
  - There are 2^6 balls.
  - Balls are numbers, i.e., hashes.
  - Let Target be 2^3=8.
  - Pick a nonce and run SHA-256.

Total no. of balls $2^6 = 64$
Target = 2^3 0 0 1 0 0 0
White balls = {Balls < Target}
$2^3 - 1 = 7$ 0 0 0 1 1 1
6 0 0 0 1 1 0
5 0 0 0 1 0 1
…

What is the probability that a while ball is picked?

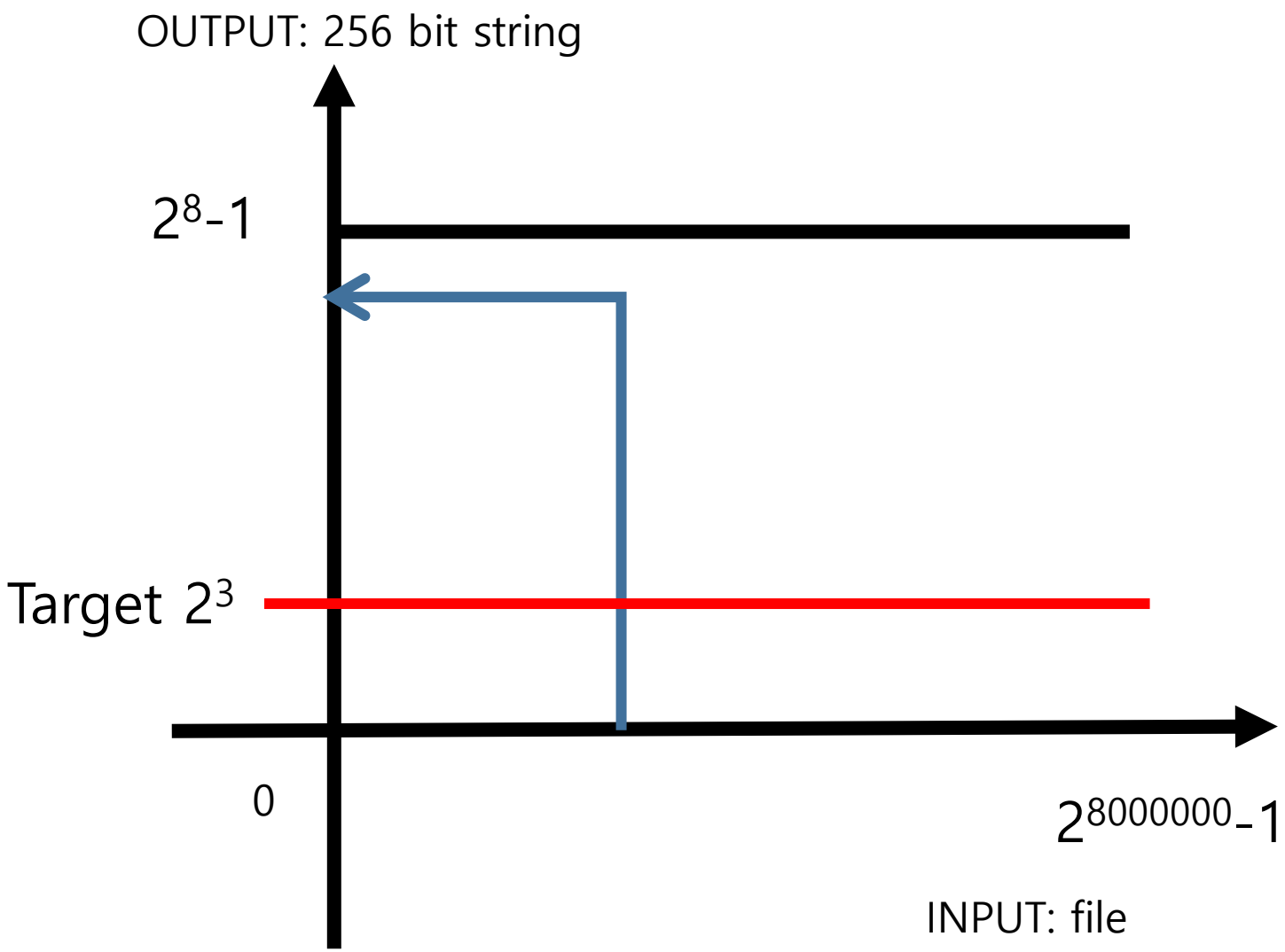
$$p = 2^3/2^6 = 1/8$$



Toy PoW 문제

- SHA-8 사용
- PoW 성공
  - Hash가 8 보다 작은 수가 되는 Input을 찾으면 성공
- 한번 뽑고 PoW 성공 확률?

$$p = 2^3/2^6 = 1/8$$



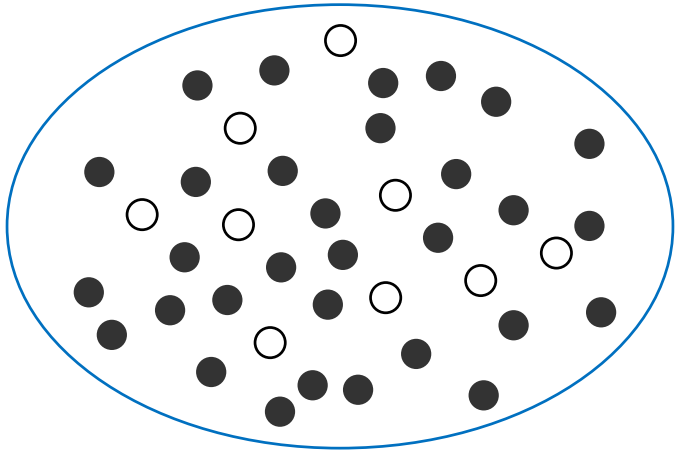
2 PoW Puzzles

- Bitcoin puzzle
  - Hashes are strings of 256 bits.
  - There are 2^{256} hashes in Y.
  - Let Target be $2^{256-16}=2^{240}$.

What is the probability that the hash satisfies the PoW?

$$p = 2^{240}/2^{256}$$
$$= 2^{-16}$$
$$= 1/64000$$

$$Y = \{y | y \text{ is a 256bit string}\}$$



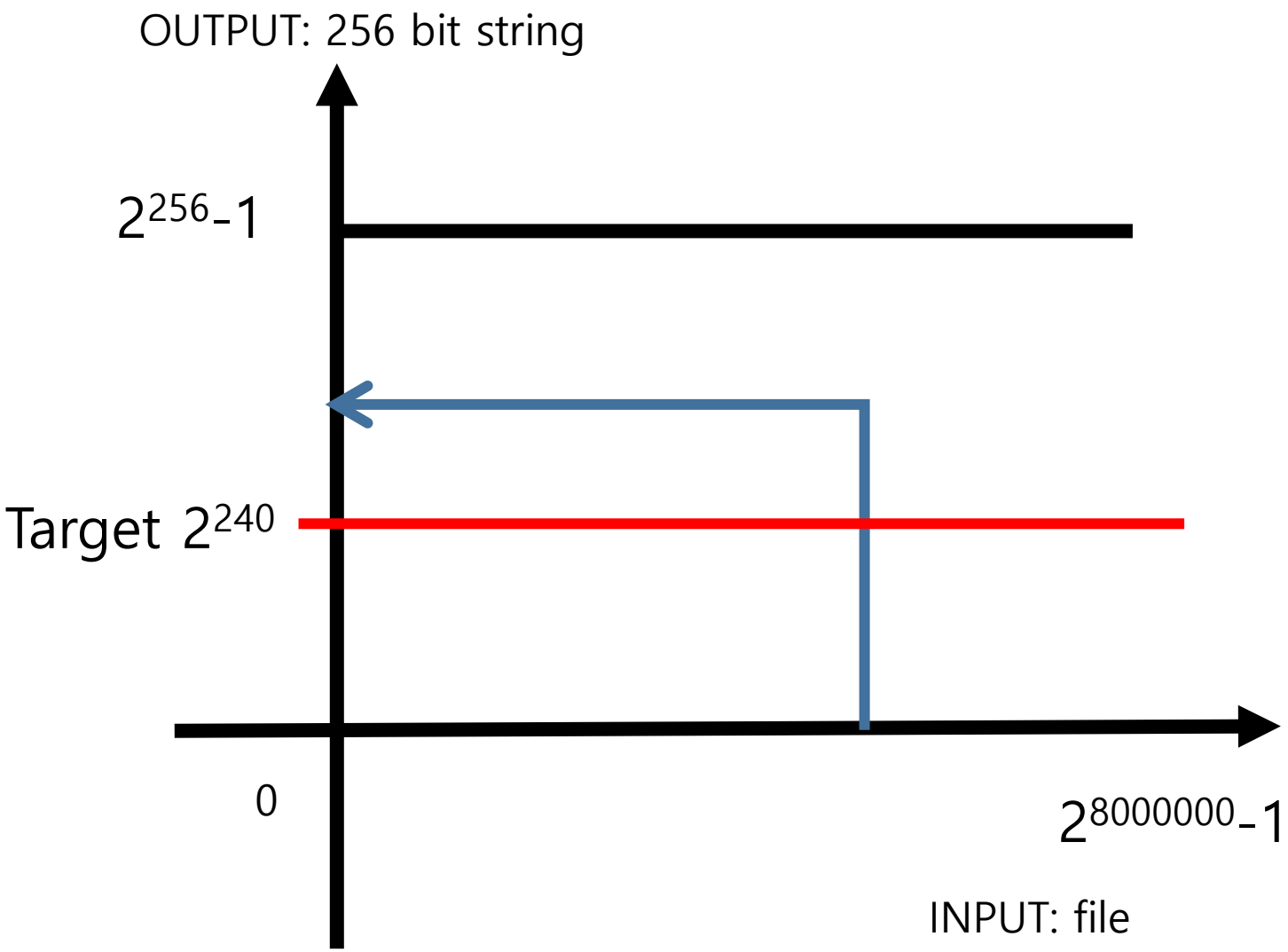
White balls are
64 hexadecimals with 4 leading zeros

"00001642b726b04401627ca9fbac32f5
c8530fb1903cc4db02258717921a4881"

PoW

- SHA-256 사용
- PoW 성공
  - Hash 값이 2^{240} 보다 작은 Input 을 찾으면 성공
- 한번 뽑고 PoW 성공 확률?

$$p = 2^{240}/2^{256}$$
$$= 2^{-16}$$
$$= 1/64000$$



4 Probability of Mining Success

- Given the difficulty p, we aim to find Probability of Mining Success.

- 한 번 뽑을 때 PoW성공 확률이 p 일 때,
  1. k 번째 뽑았을 때 비로서 성공할 확률은?
  2. 평균적으로 몇 번 뽑아야 성공할까?

4 Probability of Mining Success

- (PMF) What is the probability that a CPU solves PoW exactly at the k-th hash?

$$
\begin{aligned}
P_{pmf}(p,\ k) &:= P_p\{K \le k\} - P_p\{K \le k-1\} \\
&= P_p\{K = k\} \\
&= p + (1-p)\,p + (1-p)^2\ p + \cdots + (1-p)^{k-1}\ p \\
&\qquad -\left(p + (1-p)\,p + (1-p)^2\ p + \cdots + (1-p)^{k-2}\ p\right) \\
&= (1-p)^{k-1}\ p \qquad \text{for any } k = 1, 2, 3, \cdots
\end{aligned}
$$

4 Probability of Mining Success

- Average no. of hashes for a PoW success.
  - What is the average number of hashes for a PoW success for a given puzzle difficulty p?

$$\mathbb{E}\{K\} = \sum_{k=1}^{\infty} P_{pmf}(p,\ k)\ k$$

$$= \sum_{k=1}^{\infty} (1-p)^{k}\ p\ k$$

$$= \frac{1}{p}$$

$$= 2^{16} \quad [\text{hashes/block}]$$



Probability of Mining Success

- Definition: Hash Rate of CPU.
  - The hash rate of a CPU is defined as number hashes in a unit time.
  - For example, the hash rate of a CPU which can do 10^6 hash cycles per second is 10^6 hashes/sec.

4 Probability of Mining Success

- ASIC Mining Hardware

Bitcoin Mining Hardware Comparison

Pic	Miner	Hash Power	Price	Buy
	Antminer S9	14.0 TH/s	$3,000	
	Antminer R4	8.6 TH/s	$1,000	

출처: https://www.buybitcoinworldwide.com/mining/hardware/

Bitcoin 난이도 조절

- 평균 **BI**을 10분으로 유지하기 위해 Target값을 조절함

- 문제
- **마이너노드**의 수가 증가하면
  - Network Hash Rate이 낮아/높아 진다.
  - Block Interval이 줄어/길어 진다.
  - 평균 10분 BI을 유지하기 위해서는
  - Target을 크게/작게 해야 한다.

1 Bitcoin Difficulty

- Bitcoin Difficulty (D)

  - The aim is to keep the average block generation time be 10 min.
    - Ex) The time span to mine 2016 blocks is set to take 2 weeks.
  - Difficulty is adjusted for every 2016 block.
  - Measure the time span, T [min], during which the past 2016 blocks were mined.
  - Let T_D be 2 weeks [min], i.e., T_D = 2016x10 = 20160 [min].
  - If T is different from T_D, adjust the Difficulty D :

$$D = D_{prev} \times \frac{T_D}{T}$$

In Bitcoin, initial D is set to 1 with 8 leading hexa zeros.



- Given a Target, one can determine the network hash rate.
- Suppose you bring your own mining chip.
- You can determine your chance of winning a puzzle.
- It is the ratio of your hash rate to the total hash rate:

$$= \frac{\text{Your Hash Rate}}{\text{Your Hash Rate} + \text{Network Hash Rate}}$$

Bitcoin Difficulty

Ex

- Target 이 2^{204}.
- 채굴에 1 Tera hash/sec mining chip을 구매하여 참여.
- 이때 내가 이길 확률은?
  - The network hash power is $2^{256-\log 2target}/600 = 2^{52}/600 = 7.51e12$ [hash/sec].
  - The hash rate percentage is:

$$= \frac{\text{Your Hash Rate}}{\text{Your Hash Rate} + \text{Network Hash Rate}}$$

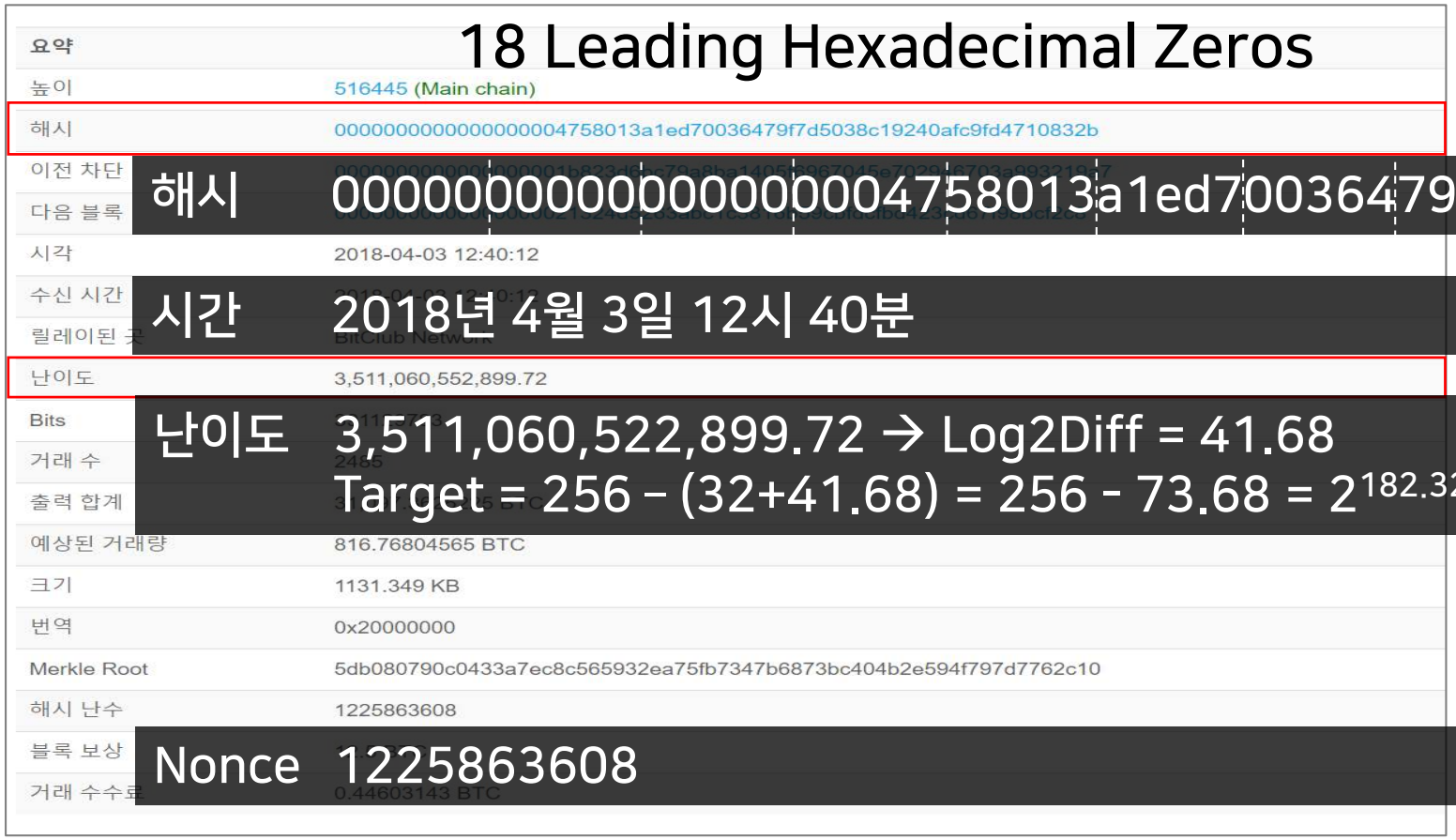
$$= \frac{1.00e12}{1.00e12 + 7.51e12}$$

$$= 11.8\%$$

Bitcoin 실전 문제

Bitcoin Difficulty

- 블록 높이 516445 비트코인 블록체인 내 깊이 값 513445에서의 블록들



18 Leading Hexadecimal Zeros

해시 000000000000000000047580l3a1ed70036479f7d5038c19240afc9fd4710832b

시간 2018년 4월 3일 12시 40분

난이도 3,511,060,522,899.72 → Log2Diff = 41.68
Target = 256 – (32+41.68) = 256 - 73.68 = $2^{182.32}$

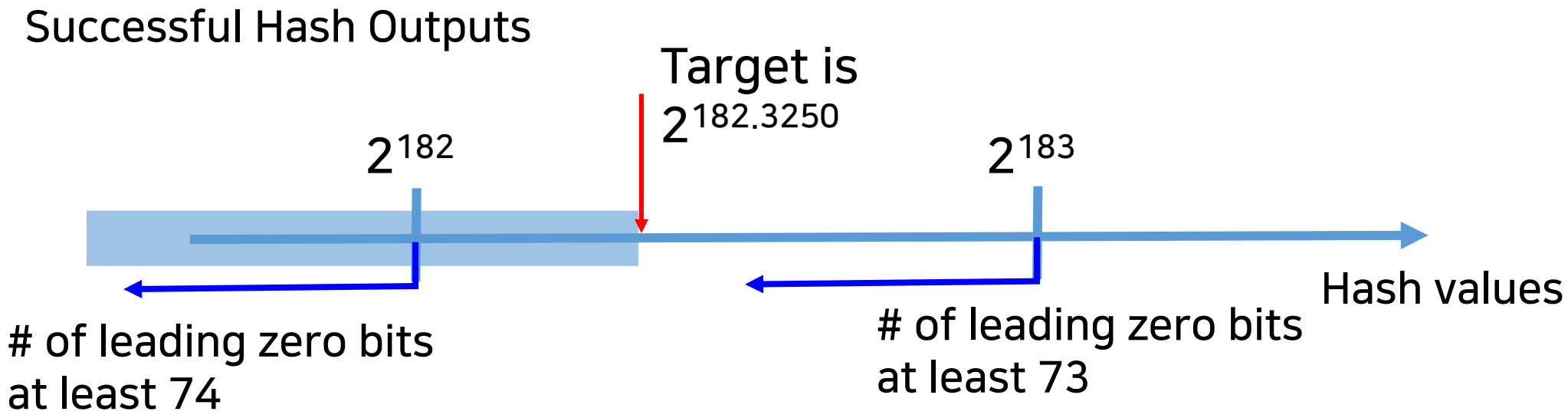
Nonce 1225863608

Bitcoin Difficulty

- Example of Difficulty and Target
  - Block #516445
  - BlockHash 0000 0000 0000 0000 0047 5801 … … … 832b
    - 18 hex zeros * 4 bits/hex + 1 bit = 72 + 1 = 73 bit zeros
  - Difficulty D is 3,511,060,552,899.7197 = 3.5e12
  - Target is $Target_0$ * (1/Difficulty)
    - Log2(D) = 41.68
    - Target = $2^{224.000}$ $2^{-41.675}$
      = $2^{182.325}$

Bitcoin Difficulty

Recall PoW Success
is
SHA Hash Output < Target



Bitcoin Difficulty

- Network Hash Rate : Block#516445
- With D=3.5e12, the probability p is about $2^{-(32+41.675)} = 2^{-73.6750}$.
- Then, it would take $1/p = 2^{73.6750} \sim$ 1.5080 e22 hashes to mine a single block.
- Dividing it by 10 min = 600 sec, the network hash rate is obtained, 25.13 Exa hash/sec.

Exa = 10^{18}

10% 해쉬 파워를 위해 투자해야 할 돈은?

Ex

- Antminer S9s (14 Thps) 이 몇 개 필요할까, 만약 내가 hash power를 0.01 % 갖기를 원한다면?
- 네트워크 hash rate 은 25 Exa hash/sec 이다.

  - You need to bring at least 179 AS9 chips.

$$\text{Your Hash Rate} \geq \frac{0.01\% \;\; \text{Network Hash Rate}}{100\% - 0.01\%}$$

$$= \frac{1}{9999} 25e18 = 25e14$$

$$= 178.6(14e12)$$

- 179000개
- 단가 3000 USD
- 537 MUSD

문제

네트워크 hash rate은 25 Exa hash/sec 이다.

- 현재 네트워크에서 작동하는 마이닝 노드의 최소 개수는 몇 개인가?

답

네트워크 hash rate은 25 Exa hash/sec 이다.

- 현재 네트워크에서 작동하는 마이닝 노드의 최소 개수는 몇 개인가?

$$\text{최소갯수} = \frac{\text{Net HR}}{\text{가장빠른칩 HR}}$$

$$= \frac{25 \times 10^{18}}{14 \times 10^{12}}$$

$$\sim 1.8 \times 10^{6}$$

문제

네트워크 hash rate은 25 Exa hash/sec 이다.

- 한개의 가장 빠른 마이닝칩으로 하나의 블록 작업증명에 성공하려면 걸리는 시간은?

문제

네트워크 hash rate은 25 Exa hash/sec 이다.

- 한개의 가장 빠른 마이닝칩으로 하나의 블록 작업증명에 성공하려면 걸리는 시간은?

$$\text{시간} = \frac{\text{Net HR} \times 600\text{초}}{\text{가장빠른칩 HR}}$$

$$= \frac{25 \times 10^{18}}{14 \times 10^{12}} \times 600\text{초}$$

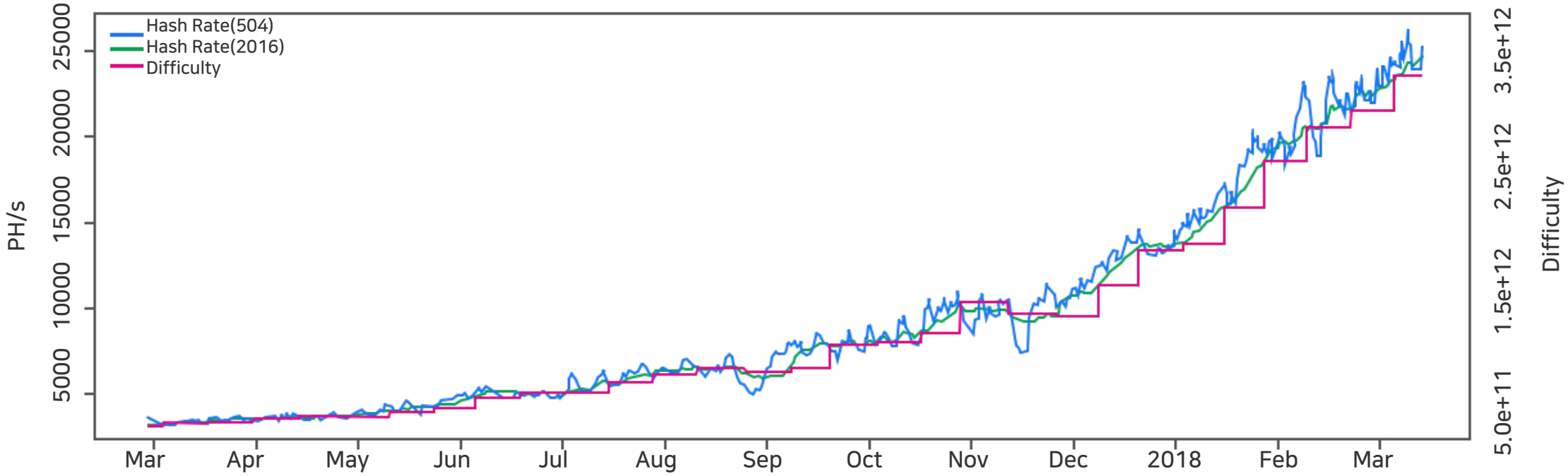
$$\sim 1.07 \times 10^{9}\text{초}$$

$$= 31.7 \text{ 년}$$

일년 = 3153만6천 초

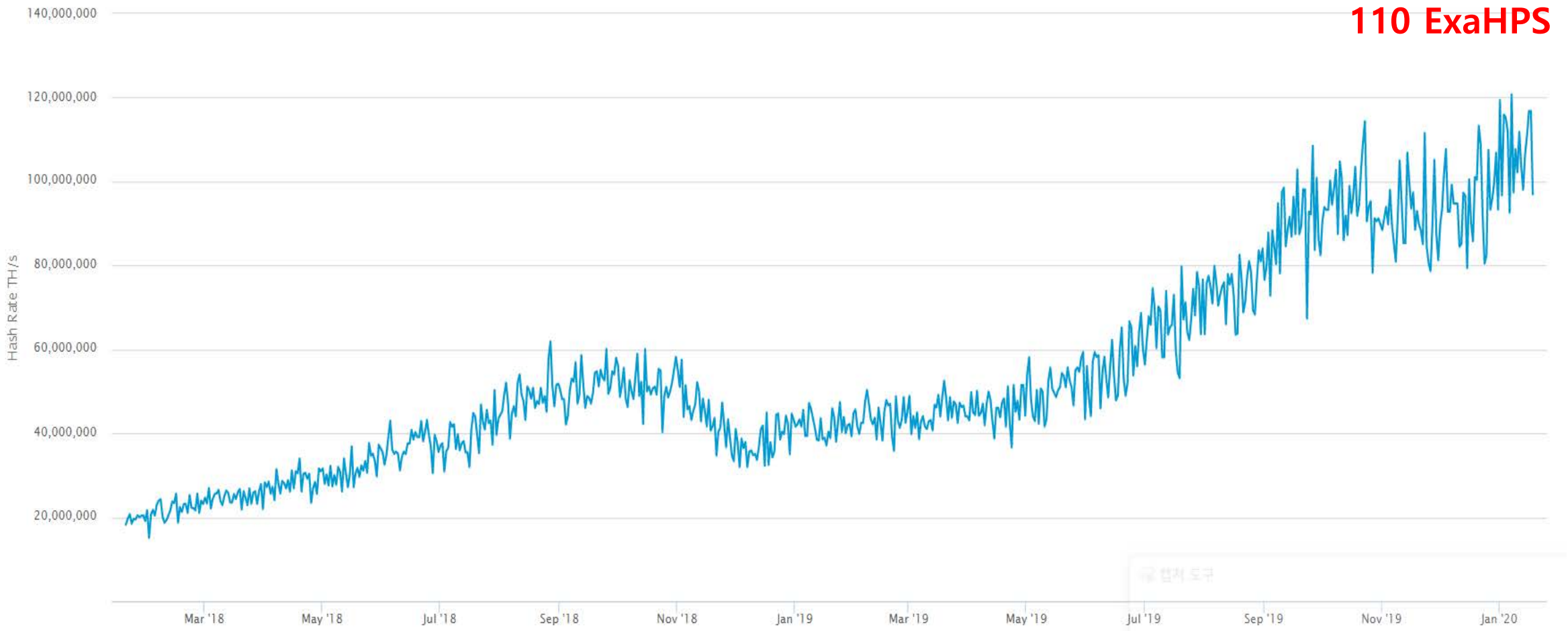
2 History of Bitcoin difficulty

- Bitcoin Hash Rate vs Difficulty (Mar/17 ~ Apr 18)



출처: https://bitcoinwisdom.com/bitcoin/difficulty

Network Hash Rate = 1/p



Goodness of Bitcoin PoW

- PoW provides flexibility in TPS solutions!
  - Small difficulty → fast TPS
  - Large difficulty → slow TPS

- PoW is sufficient for data immutability!

- PoW and blockchain is a technological breakthrough.

현재 PoW의 문제점

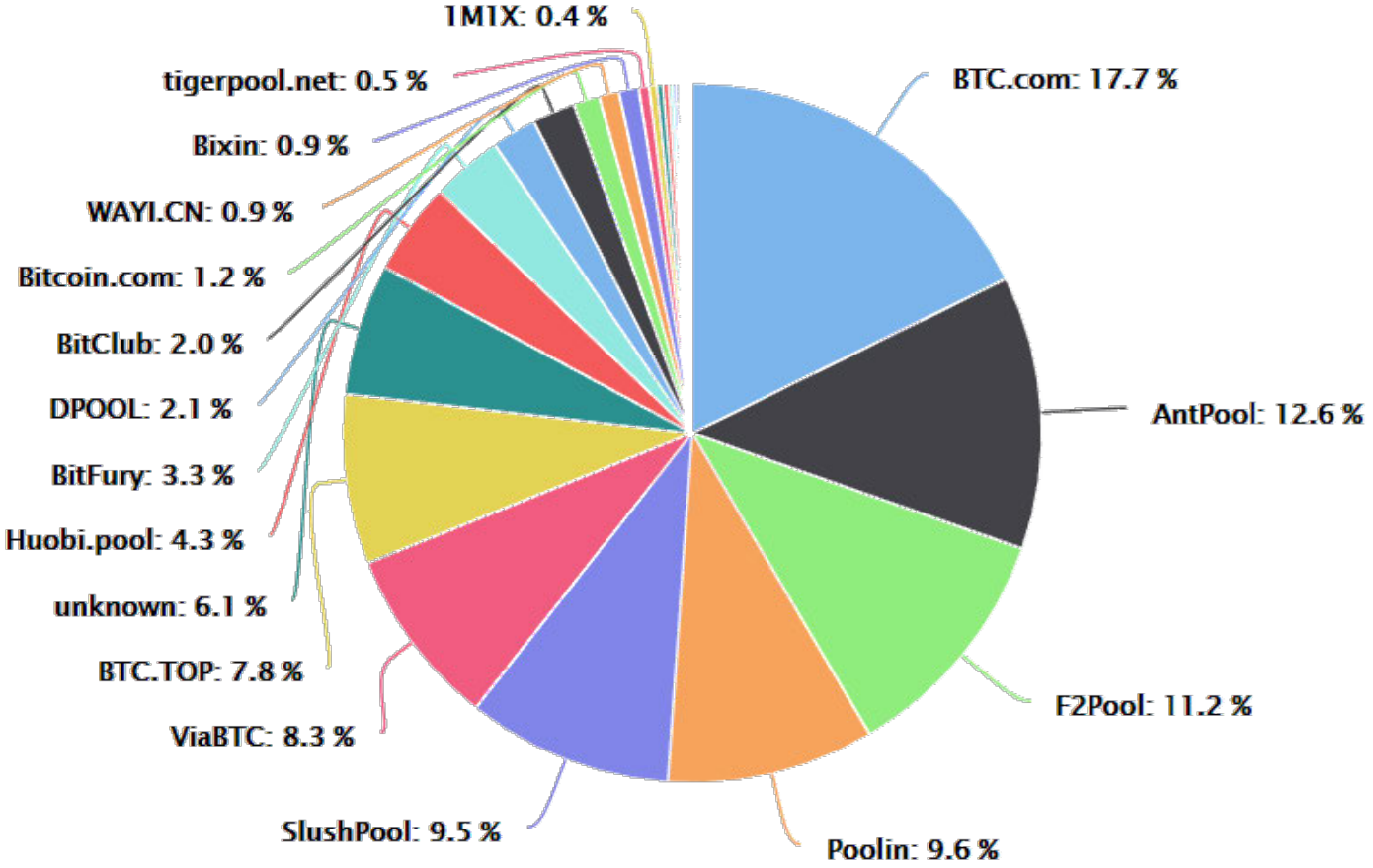


Proof of Work-the Monopoly Problem

- Miners are rational profit seekers.
- They strive to make as much profit as possible.
- Mining operations are highly parallelizable, GPU mining quickly replaced CPU mining:
- FPGAs did GPUs.
- ASICs did FPGAs.

3 Proof of Work-the Monopoly Problem

- Proof of Work, any alternative?
  - PoW, monopolized today.
  - Handful of mining sites dominating.
  - The trust has been degraded.
  - No more one CPU one vote.
  - Rational profit seeking miners use ASICs now.
  - Huge energy spending



https://btc.com/ (기간: 2018. 09 ~ 2019. 09)



Proof of X Schemes

- Proof of Stake (PoS)
- Delegated PoS (DPoS)
- Proof of Activity (PoA)
- Byzantine Fault Tolerant (BFT)



Proof of X Schemes

- Proof of Stake (PoS)
  - Give higher PoW chance to a node with a higher stake (more coins).
  - Good: No high energy consumption
  - Bad: Rich gets richer problem
  - What if the node stays off line?
  - Centralized.
- Delegated PoS
  - Lend coin and share profit.
  - Small number of nodes, off-chain politics

Reference: See page 2114, IEEE COMMUNICATIONS SURVEYS & TUTORIALS, VOL. 18, NO. 3, THIRD QUARTER 2016



Proof of X Schemes

- Proof of Stake based on Coin Age
  - *Coin age* is no. coins times the holding period.
    - Implemented in Peercoin (peercoin.net).
  - The difficulty of PoW is individually determined, inversely proportional to one's *coin age*.
  - If one finds a solution, one's *coin age* is reset.
  - Slowly increasing the chances of solving the puzzle next time.

Reference: See page 2114, IEEE COMMUNICATIONS SURVEYS & TUTORIALS, VOL. 18, NO. 3, THIRD QUARTER 2016



Proof of X Schemes

- Proof of Stake
  - In contrast to PoW, where the longest block chain survives, *coin age* PoS declares the block chain with the highest total sum of *destroyed coin age* as the main chain.
  - An attacker must hold a huge amount of coins.



Proof of X Schemes

- Proof of Stake
  - Good: Energy consumption is minimized.

[229] N. Houy, “It will cost you nothing to ‘kill’ a proof-of-stake cryptocurrency,” Econ. Bull., vol. 34, no. 2, pp. 1038–1044, 2014.



Proof of X Schemes

- Proof of Stake
  - Bad
    - Coin age accumulates even when the node is not connected to the network.
    - Come online for reward go offline afterwards.
    - The lacking of sufficient number of online nodes, may facilitate attacks.

5 Summary of Altcoins

- Table IV – Summary of Altcoins and Extensions

	Approach	Distinct Feature (incl. References)	Sec.
Precursor	B-Money	Mining reward proportional to proof of work difficulty; requires a broadcast channel [7]	II-B, V-D, V-E
	Bit Gold	Chained proof of work [10]; Byzantine-resilient quorum [13]	III-B, V-D, V-E
	Karma	Distributed currency maintained by a bank set [8]	V-E
	RPOW	Centralized (reusable) proof of work exchange/ bank [9]	V-E
Altcoins	Bitshares (BTS)	Delegated proof of stake [231]	V-F
	Bytecoin (BCN)	Implements CryptoNote [190], which aims for unlinkable and untraceable transactions	V-C, V-E
	Counterparty (XCP)	Colored coin; used proof of burn	V-H, V-H
	Cryptonite (XCN)	Implements the mini block chain scheme [127]	IV-D
	Dash (DASH)	Formerly known as Darkcoin; implements native CoinJoin-like transactions [178]	V-C
	Dogecoin (DOGE)	Block payload holds TXIDs only; fast block generation	IV-D, V-E
	Litecoin (LTC)	Uses scrypt [214] to foster distributed power among miners	V-E
	Mastercoin (MSC)	Colored coin; exodus address	V-H
	Nextcoin (NXT)	Entirely proof of stake based	V-F
	Peercoin (PPC)	Identified coin age as alternative measure; proof of stake [227]	V-F
	Primecoin (XPM)	Proof of work with intrinsic value i. e. prime chains [218]	V-E
	Reddcoin (RDD)	Proof of stake velocity [234]	V-E
	RSCoin	Centrally controlled money supply with distributed verification [126]	IV-D
	Ripple (XRP)	Implements a novel Byzantine agreement protocol [200]	V-D
	Zerocash	Full-fledged altcoin, carrying on the ideas of Zerocoin [189]	V-C

5 Summary of Altcoins

- Table IV – Summary of Altcoins and Extensions

	Approach	Distinct Feature (incl. References)	Sec.
Altchains	Bitmessage	Secure messaging service [145]	IV-G
	Ethereum (Ether)	Turing complete smart contract processing [44], [45]	II-E
	Namecoin (NMC)	Key-value storage; realizes decentralized domain name coordination [143]	IV-G
	Permacoin	Decentralized file storage; proposes proof of retrievability [100]	V-E
Protocols / Extensions	CoinJoin	Uses multi-signature transactions to enhance privacy [160]	V-C
	CoinShuffle	Decentralized protocol to coordinate CoinJoin transactions [180]	V-C
	CoinSwap	Enables P2P-based trustless mixing [41]	V-C
	CommitCoin	Secure timestamping protocol [40]	V-H
	Mini block chain	Identifies individual block chain components [127]	IV-D
	Mixcoin	Mixing with accountability [174]	V-C
	Zerocoin	Unlinkable and untraceable transactions by employing zero knowledge proofs [187]	V-C

CoinMarketCap

Accessed on 1/21/2020

	#	Name	Market Cap
PoW	1	Bitcoin	$157,523,327,976
PoW+PoS	2	Ethereum	$18,237,664,973
BFT	3	XRP	$10,173,030,507
PoW	4	Bitcoin Cash	$6,324,111,387
PoW	5	Bitcoin SV	$5,514,168,411
PoS	6	Tether	$4,642,239,400
PoW	7	Litecoin	$3,703,645,030
DPoS	8	EOS	$3,489,493,081
BFT+PoS	9	Binance Coin	$2,703,188,581
BFT	10	Stellar	$1,257,162,885
PoS	11	Cardano	$1,133,667,715
PoW	12	Monero	$1,132,339,953
DPoS	13	TRON	$1,110,529,913
PoS	14	Tezos	$1,049,197,574
PoW	15	Ethereum Classic	$1,033,632,329
PoW+PoS	16	Dash	$1,025,878,916
	17	Chainlink	$942,976,348
PoW	18	UNUS SED LEO	$888,524,224
BFTPoS	19	Cosmos	$867,402,795
dBFT	20	Neo	$793,931,677

Proof-of-XXX, Alternatives to PoW

	Pros	Cons	Coins within top 50 rank
PoW (Proof-of-Work)	▪ Strongest security ▪ Difficult to produce ▪ Easy to verify	▪ Extreme computing power ▪ 51% attacks ▪ Transaction speed / Transaction throughput	Bitcoin, Ethereum
PoS (Proof-of-Stake)	▪ Energy & hardware efficiency ▪ Expensive 51% attacks	▪ Recentralization ▪ The rich-get-richer ▪ "Nothing at stake" problem	ETH2.0, Cardano, Tezos
DPoS (Delegated PoS)	▪ Scalability and speed ▪ Energy & hardware efficiency ▪ Encouraging good behavior by real-time voting	▪ Recentralization ▪ DDoS attacks	EOS, TRON, NEO
BFT	▪ Resistant to n/3 attacks ▪ Validators are randomly selected.	▪ $O(n^2)$ complexity	XRP, NEO, BinanceCoin, Cosmos

 © 2019 GIST

Libra

- 합의알고리즘: **Byzantine Fault Tolerance** O(n^2) + 지분증명
- 노드수 100개
- 1000 ~1500 TPS
- 중앙화 솔루션
- Stable Coin ~ Reserve
- Facebook network has 2B users.
- The era of digital currency may arrive sooner!



Marcus noted that blockchain technology is inevitable and if the U.S. doesn't lead in building and regulating it, the tech will come from places "out of reach of our national security apparatus," raising the spectre of China

LibraBFT

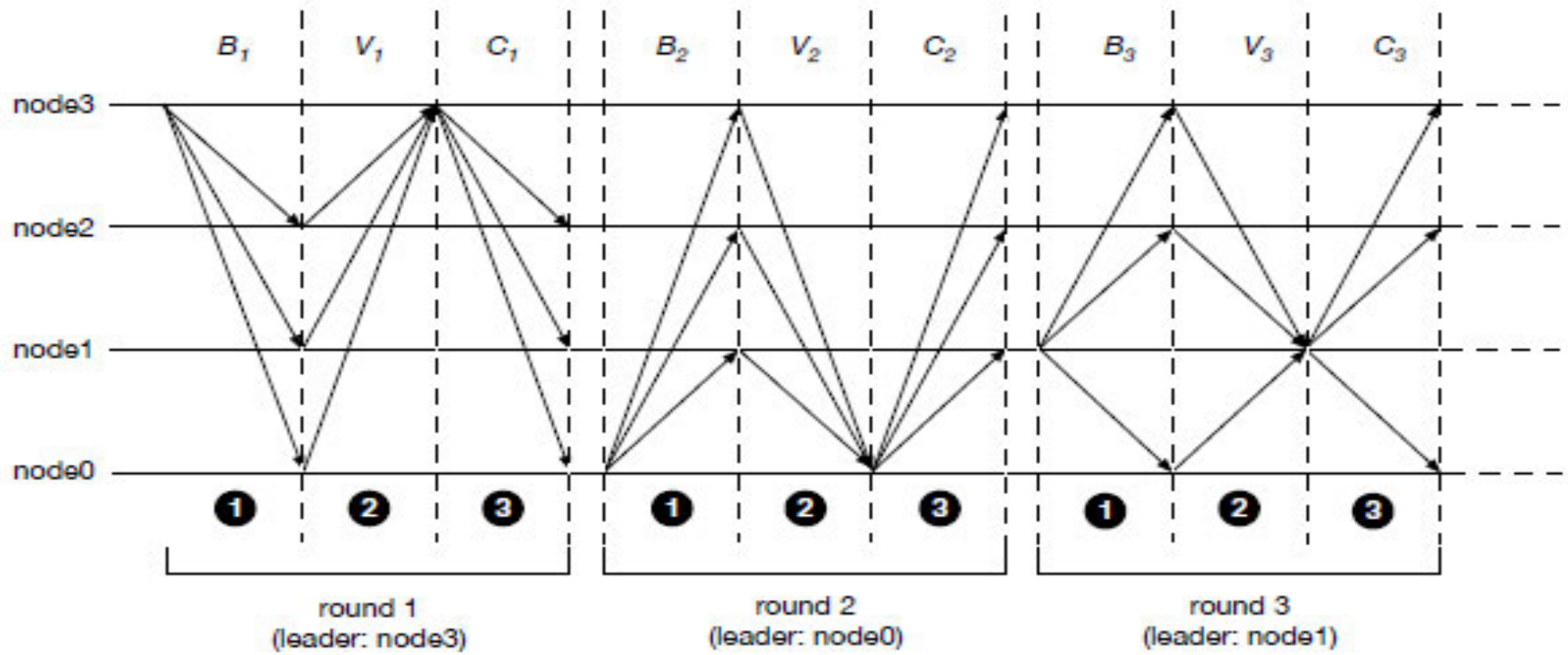


Figure 3: Overview of the LibraBFT protocol (simplified, excluding round synchronization)

PoW는 근본적인 보안성 해결방안

- 탈 중앙성과 보안성 모두 확보 가능한 해결방안!

- 주요 합의알고리즘은 탈 중앙성을 희생하여 보안성과 확장성을 얻으려 목표, 혹은 off-chain 정책에 의존함.
  - Staking, Delegation, Activity-checking, Leader Selecting, Random Selecting, etc.

- PoW의 현재 문제점을 새로운 기술로 해결하면 여전히 가장 좋은 해결방안이 됨!

블록생성과 검증이 분리되면 어떤 일이?

- 랜덤 선출된 자가, 지분을 많이 가진 자가, 활동이 많은 자가, 장군으로 선출된 자가, …

블록을 생성하세요.

- 나머지는 검증을 하겠습니다.
- 선출 → 생성 → 검증 등 여러 단계로 분리 됨.
- 각 단계에서 일이 제대로 되었을까?
  - Validators를 뽑아서 합시다. 보상은 나눕시다.
- 생성자 혹은 검증자가 일을 대충해서 좋지 않은 TX가 들어갈 경우?
  - 각 단계에서 혹시 실패한 경우가 발생할 경우는…
  - 평판도 측정, Staking 압수, …

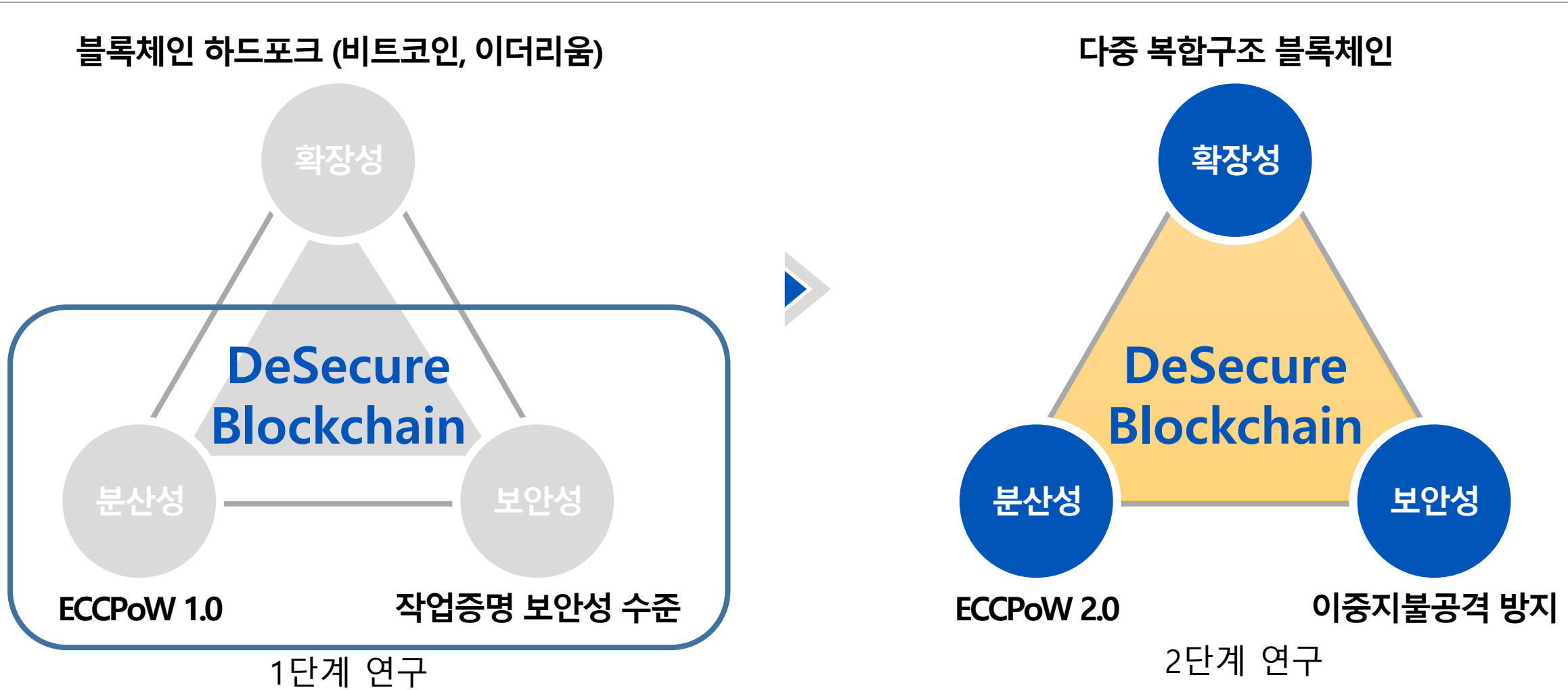
PoW는 생성과 검증이 동시에 이루어짐

- 수 많은 노드가 생성과 검증을 경쟁적으로 (동시에) 함
- 끊김이 없이 블록이 생성됨
- 앞으로 전진만 있음
- 잘못된 TX이 포함될 경우 헛고생, 스스로 검증하고 생성함

블록체인 트릴레마 (V. Buterin)

구분	내용	사례
확장성 (Scalability)	사용자 수 증대에 유연하게 대응할 수 있는 능력 단위 시간당 거래건수가 늘어 날 때 무리 없이 거래를 증대시킬 수 있는 능력을 의미	이더리움 Plasma, Sharding
탈중앙성 (Decentralization)	중앙 집중화 구조를 벗어나 노드들이 소규모 네트워크로 **자율적으로 모여 운영주체가** 되는 것을 말함 기존의 서버-클라이언트 관계를 벗어나 개별 노드들이 자발적이고 자율적으로 피투피(P2P) 방식으로 연결함 블록체인 기술이 사회적으로 널리 확산됨에 따라 **기존의 중앙집중식 조직, 기업, 단체, 기구** 등은 **탈 중앙 구조로** 변경되고 있음	초기 Bitcoin Network
보안성 (security)	블록체인 **데이터의 무결성 확보**하거나, 오픈 프로토콜이 원래 의도한 바 데로 시스템이 작동하도록 하고, 외부로 부터의 공격으로 부터 **시스템을 보호**하는 능력 을 의미	

Research at GIST Blockchain Economy Center



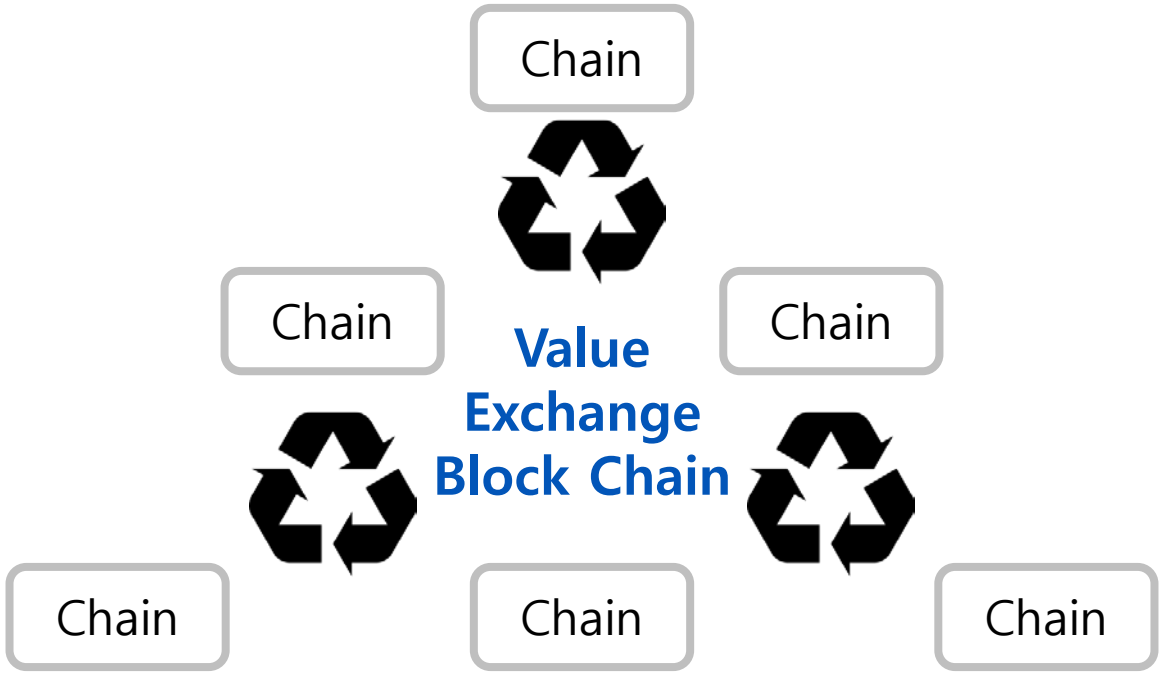
© 2019 GIST

확장성 강화 연구 with DeSecure Blockchains

- Evolutionary Space에서 DeSecure 블록체인을 용도에 따라 특화
- 용도별 특화 블록체인들을 연결하여 다중 복합구조 DeSecure 블록체인 형성

Low TPS
Global Nets
High Delay

Large Tx (Assets)



High TPS
Local Nets
Low Delay

Small Tx
(Utility Coins)

<용도에 따라 세분화된 다중 복합구조 블록체인>

© 2019 GIST

Comparison to Existing Scalability Solutions

DeSecure Blockchain aims to resolve scalability issue with promotion of competition.

Type	DeSecure	Bitcoin		Ethereum	
Name	Multi-level, multiple chains	Seg-Wit	Lightening Network	Plasma	Sharding
How	Many ECCPoW based chains can talk to each other via value-exchange service	Realize by modifying a block data structure	Allow off-chain transactions and record the end result of these transactions into the main blockchain	Allow transactions in child chains, TX records end up at the main chain are limited.	Divide BC DB with multiple shards
Pro	Many different services and levels of chains can co-work.	Easy to realize	Faster transactions Small TX fees	Faster transactions Small TX fees	Faster transaction
Con	No single chain solution Requires an ecosystem	Small improvement	The content of off-chain transactions lost	Some TX content lost Only full node can run this	Increased SW complexity

블록인터벌을 10분에서 1분으로 줄이려면 어떻게 될까?

블록인터벌을 10분에서 1분으로 줄이려면 어떻게 될까?

수 많은 fork가 매 번 발생하는 문제 발생, 비 동기화
DS 공격 취약성을 따져봐야
그러나, Profitable DS Attack분석으로 해결 가능 (See DeSecure블록체인)
소규모 네트워크, 용도별, 지역별로 쓰는 용도를 다르게 할 수 있다.

네트워크 별, 용도별, 지역별 로 서로 다른 블록체인 만들고 연동 시킨다!

이 것을 가능하게 하는 필요 조건은 무엇인가?

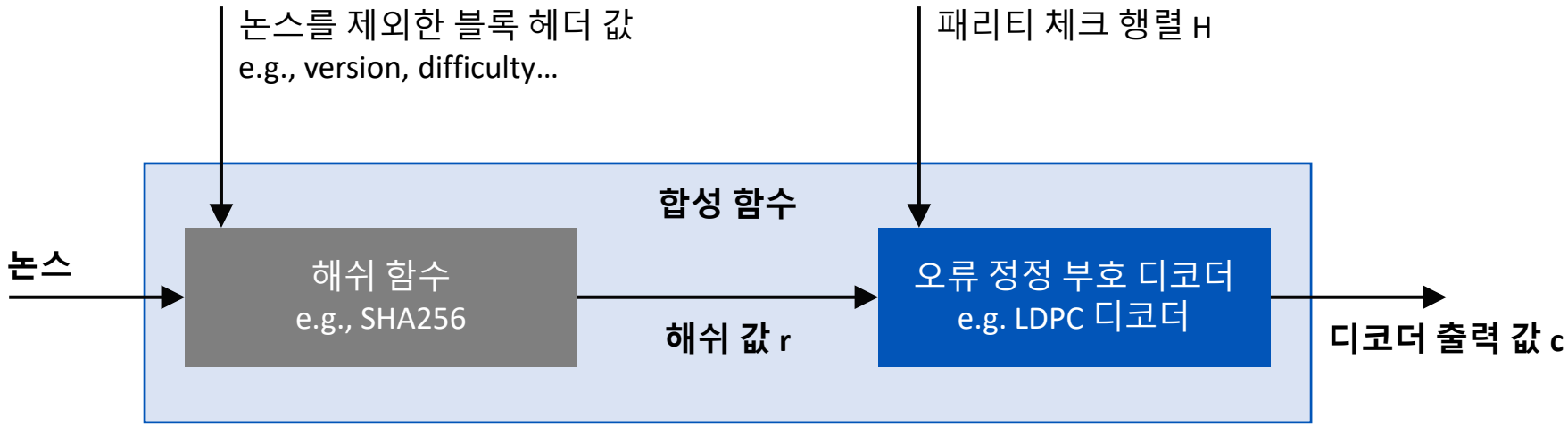
Key New Idea is a TV Puzzle System!

- 새로운 PoW 알고리즘 개발을 위해 고려해야 할 사항
  - Puzzles 문제를 다양화 하고, 문제 출제 방식을 시시각각 바꿀 수 있게 한다.

  - A puzzle should be difficult to solve but very easy to check.
  - The puzzle should be resistant to attacks.
  - Solution to the puzzle for a block should not be reusable.
  - Puzzle difficulty should be adjustable.
  - Anyone with a CPU who wishes to participate should be able to join.
  - Consensus must eventually be reached; there must be a common rule to resolve forks and to determine the main blockchain.

부호 – 암호 작업증명

- 부호-암호 작업증명 (ECCPoW) 정의 :
  오류 정정 부호와 작업증명을 동시에 사용하여 분산성/보안성을 확보하기 위해 개발된 작업증명 기술임.
- ECCPoW은 매 블록마다 변하는 합성 함수를 사용함.

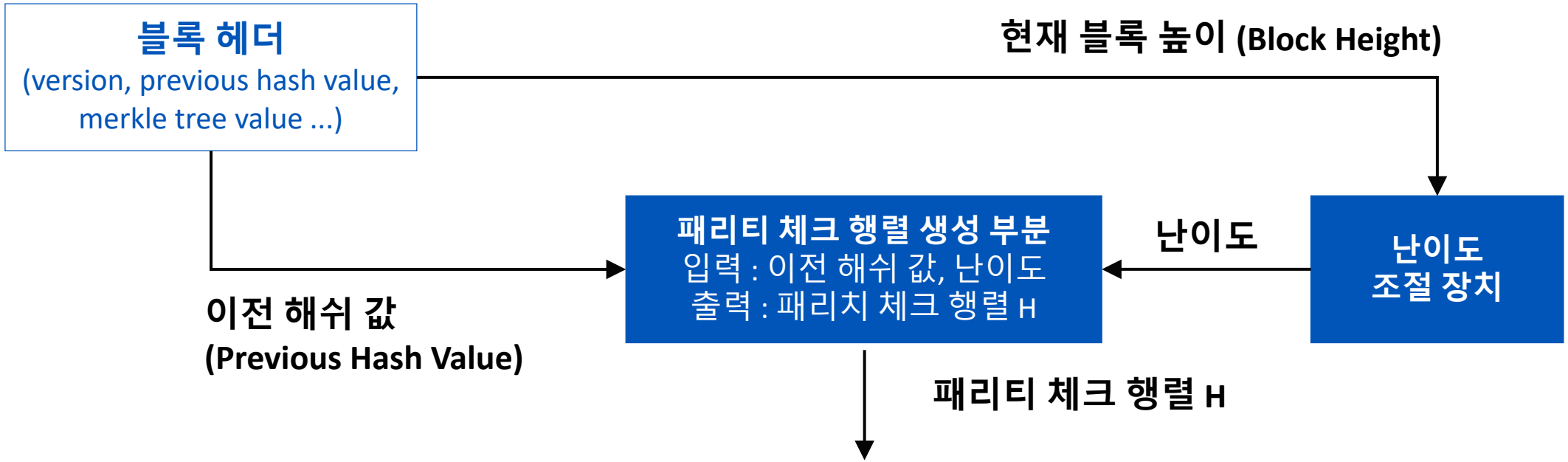


- (**시변성**) 패리티 체크 행렬을 매 블록 마다 변경함으로써, 함성 함수를 매 블록 마다 변경 할 수 있음.
- (**무한성**) 무한히 많은 패리티 체크 행렬을 사용하면, 무한히 많은 합성 함수를 생성 할 수 있음.

Ⓒ 2019 GIST

매 블록마다 바뀌는 암호 퍼즐 생성 (1)

- 난이도는 "블록 길이" 및 "부호의 비율"을 결정
- 이전 해쉬 값을 이용해 Gallager code 생성 방법의 순열 순서를 결정



Ⓒ 2019 GIST

Time-Variant Proof-of-Work Using Error-Correction Codes

Sangjun Park, Haeung Choi, and Heung-No Lee, *Senior Member, IEEE*

***Abstract*— The protocol for cryptocurrencies can be divided into three parts, namely consensus, wallet, and networking overlay. The aim of this consensus part is to deal with coming to an agreement among nodes to the current status of their blockchain. This status must be updated only through valid transactions and it is achieved among trustless rational peer nodes. A proof-of-work (PoW) based consensus has been proven to be secure and robust owing to its simple rule and has served as a firm foundation for cryptocurrencies such as bitcoin and ethereum. However, the usage of specialized mining devices for the existing PoWs causes two problems: *i*) the re-centralization issue of a mining market and *ii*) the usage of a considerable amount of energy in mining. In this paper, we propose a new PoW called Error-Correction Codes PoW (ECCPoW) where the error-correction codes and their decoder can be utilized by concatenating the decoder with a hash function. In ECCPoW, puzzles can be intentionally generated to vary from block to block, leading to a time-variant property. This property is useful in repressing the emergence of the specialized mining devices, which can be a solution to the problems that the existing PoWs face currently with**

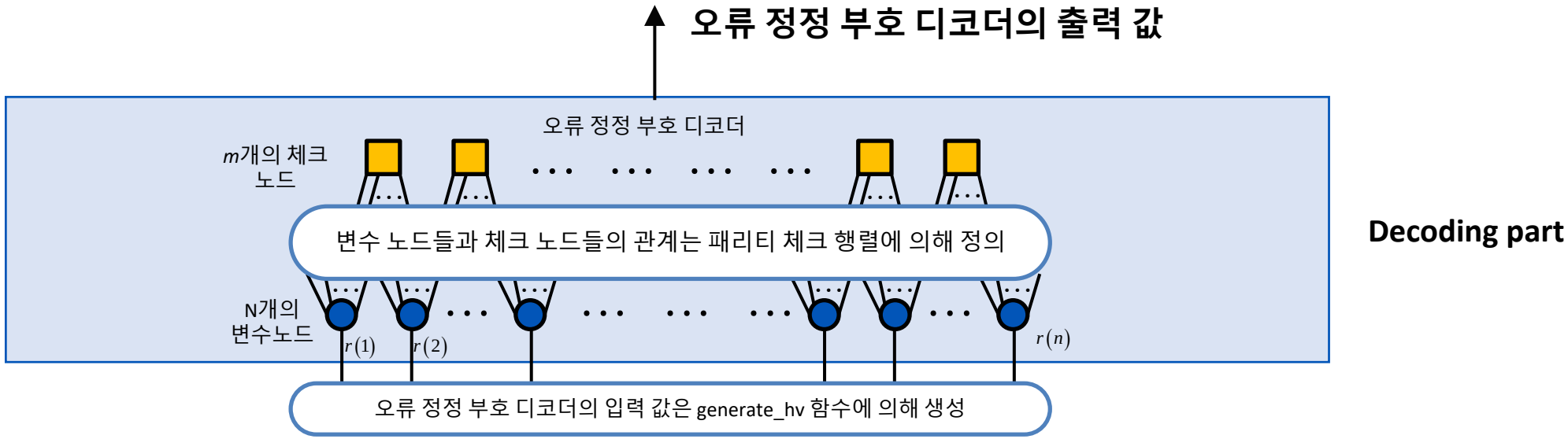
***Index Terms*— Consensus, Cryptocurrency, Blockchain, Proof-of-Work, Error-Correction Codes, Hash Functions**

blocks can be possible In contrast, if PoW is attached to each mined block, attackers cannot make any unauthorized modifications without redoing all the works. No one thus can alter any mined blocks, meaning an immutability property.

In bitcoin, miners make rational decisions to maximize their profits. These decisions are conducted by following two stages. First, the miners select a blockchain in which its height is the longest. Second, they extend this longest one by adding a newly mined block which is a result of solving a puzzle generated in PoW. To evaluate the quality of the decisions, we consider an example. At a certain time point, suppose there are two chains in which one chain is longer than the other chain. The longer chain shall be treated preferable and more trustable because it has the more PoW accumulated, meaning that it is more difficult to alter. Thus, miners shall select and extend the longer chain. Making such a selection is good for keeping the mining rewards. The mining reward is actually a delayed conditional payment. Given the miner who mined a certain block at a certain time point t_1, the decision to execute the reward is delayed only after a certain amount of time has passed since the time point t_1. This holding time is measured in terms of number of

매 블록마다 바뀌는 암호 퍼즐 디코더 개발 (2)

- 오류 정정 부호 디코더 입력 값 생성



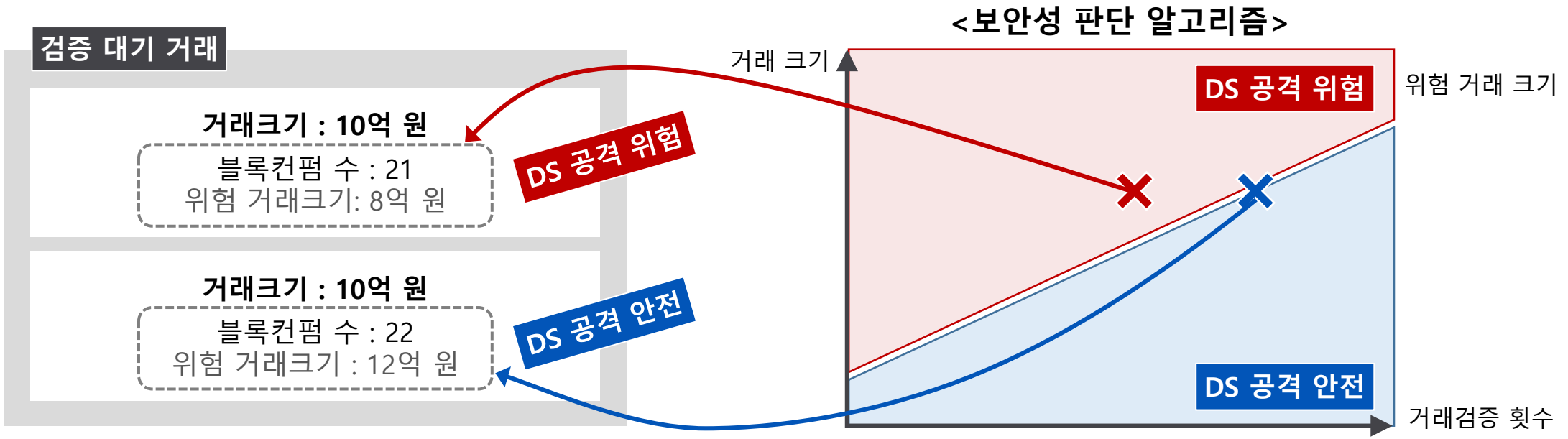
- 패리티 체크 행렬의 변경은 변수/체크 노드들의 관계의 변경을 야기함.
- 즉, 동일한 입력 값을 취득하더라도, 패리티 체크 행렬이 변경되면 상이한 값을 산출함.
- 디코딩은 메시지 전달 알고리즘에 의해 이뤄짐.

Ⓒ 2019 GIST

보안성 강화 연구

TX Size별로 보안성을 판단해 각자 대처

- 거래 별로 DS공격의 위험성을 평가하여 사용자가 보안위험 자각 및 대처 가능
- 거래검증 횟수와 거래크기가 DS 공격자에게 이윤을 발생시키는지를 거래 별로 평가 가능



© 2019 GIST

Scalable DeSecure Blockchain 개발/보급

- 부호-암호 화폐 시스템 (GIST특허 2018)
  - Error-Correction Code 기반 PoW
  - 재-중앙화 문제 해결 (SHA함수에 전적으로 의존 ASIC칩)
  - 누구나 화폐 발행에 참여 가능

- 세월의 검증을 이겨낸 Bitcoin과 Ethereum의 합의알고리즘에 적용, 보급, 확산 계획

- 탈 중앙적이며 안정적인 Bitcoin과 Ethereum 보급
  - Scalability 문제 해결

ECCPoW resolves the issue of Recentralization!
[See Ref 1]

- Uses time-variant Error-Correction Code(ECC) PoW
  - ASIC/FPGA/Pool resistant
- Achieves a decent decentralization
  - People with a cpu or a gpu can participate!
- Provides energy efficient operation
  - Small hash power needed: O(10^10), instead of O(10^20)
- Offers ECCPoW puzzles completely different from SHA-PoW
  - There are many mining chips available for SHA-PoWs.
  - They cannot be utilized to launch a DS attack to ECCPoW.

Aim to launch BTC-ECC first.

LiberVance

- BTC-ECC hardfork has been 90% completed.
- Raise funding with BTC-ECC.
- Pressing to launch BTC-ECC in 2nd Qtr 2020.

- Fund raised from BTC-ECC will be used to help a series of follow-up projects.

Today money is not honest!

[Hayek has predicted in Denationalization of Money]

- P.24. Free trade in money. As soon as the public became familiar with the new possibilities, any deviations from the straight path of providing an honest money would at once lead to the rapi d displacement of the offending currency by others.
- P. 115. There is no better case for preventing the decrease of the quantity of money circulating in a region or sector of a larger community than there is for governmental measures to preven t a decrease of the money incomes of particular individuals or groups—even though such mea sures might temporarily relieve the hardships of the groups living there. It iseven essential for honest government that nobody should have the power of relieving groups from the necessity of having to adapt themselves to unforeseen changes, because, if government can do so, it wil l be forced by political necessity to do so all the time.
- P. 116. Such areas in which one currency predominates would however not have sharp or fixed boundaries but would largely overlap, and their dividing lines would fluctuate. But once the pri nciple were generally accepted in the economically leading countries, it would probably spread rapidly to wherever people could choose their institutions. No doubt there would remain encla ves under dictators who did not wish to let go their power over money—even after the absenc e of exchange control had become the mark of a civilized and honest country.

The Hayek's Proposal

- The concrete proposal for the near future, and the occasion for the examination of a much more far-reaching scheme, is that
- the countries of the Common Market, preferably with the neutral countries of Europe (and possibly later the countries of North America) mutually bind themselves by formal treaty not to place any obstacles in the way of the free dealing throughout their territories in one another's currencies (including gold coins) or of a similar free exercise of the banking business by any institution legally established in any of their territories.

Which sort of currency would the public select? [Hayek pg 66]

- Four uses of money
  - (i) Cash purchases
  - (ii) Saving
  - (iii) Standard of deferred payments
  - (iv) A reliable unit of account

Hayek argues that money is a commodity that would be better off supplied through competition.

- Monopoly of government vs. competition by private issuers.
- The advantages of competitive currencies are not only removing the power of government to inflate the money supply but also that they would go a long way to prevent the destabilizing fluctuations that government monopoly of money has precipitated over the last century.
- In addition, it makes it difficult for government to inflate its own expenditures.
- The central theme is crystal clear: government has failed, must fail, and will continue to fail to supply good money.

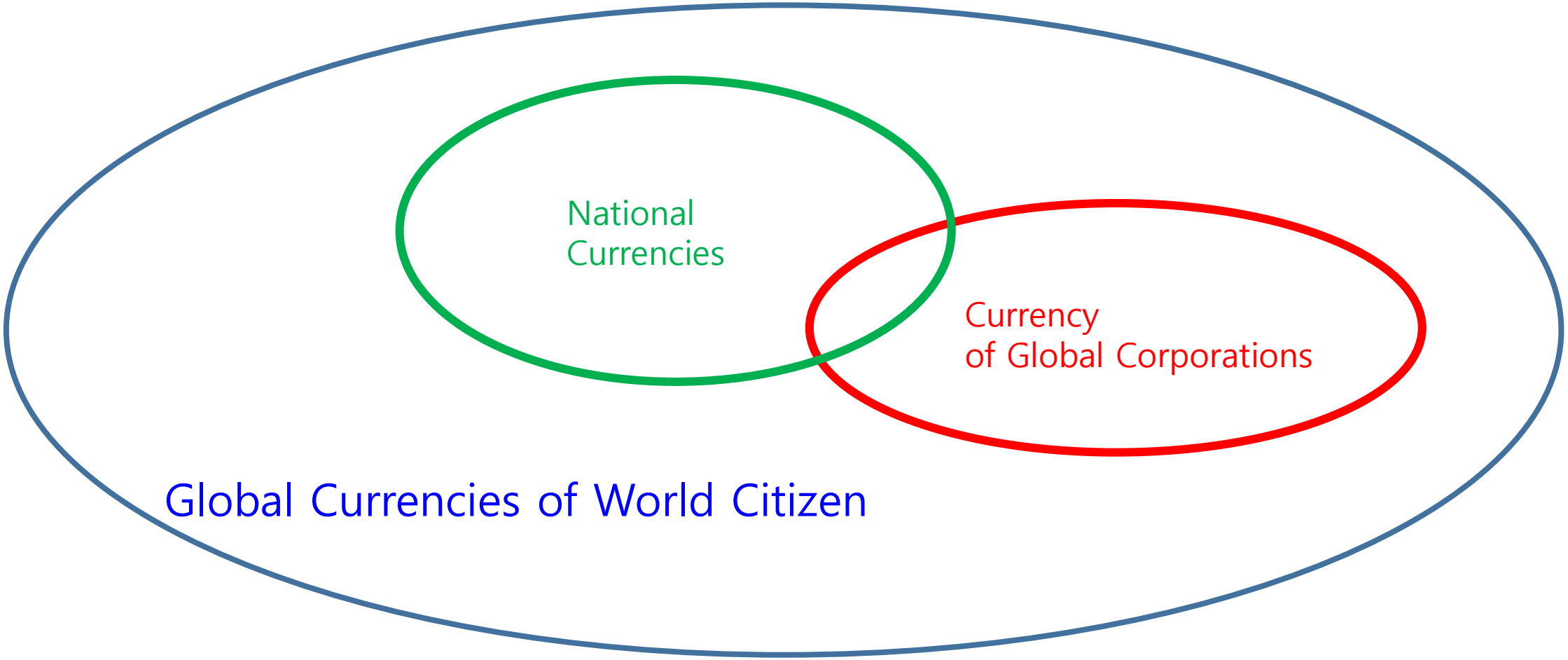
BIT-ECC is money of people!

- We are people from the globe.
- We have our computers shared.
- This computer network maintains the book of transactions.
- This computer network mints coins.
- Currency is created by us the people.
- We the people get the seigniorage.
- Supply of this global currency is fixed.

BIT-ECC provides honest money!

- It is a money freely traded worldwide.
- Supply is fixed.
- Participation is open.
- BIT-ECC is more decentralized than BIT.

10 years from now, competing currencies



Concluding Remarks

- Bitcoin is the currency of world citizen.
- PoW is fundamental to decentralization and security.
- Centralized altcoins are not as valuable as bitcoin.
- DeSecure blockchains and multi-level blockchains provide new technological advance.
- It is inevitable to use cryptocurrencies.

References

1. **Heung-No Lee, "DeSecure Blockchains," ETH-CON, Seoul, Reb. of Korea., May 2019.**
2. **Heung-No Lee, "Blockchain Consensus and Governance," July Meet-Up, Institute of Blockchan and Law, July 11th, 2019. YouTube Video Available at** https://www.youtube.com/watch?v=7ujkFgsKPdY.

Selected References of GIST Blockchain Economy Center

- [Lee1] JH Jang and Heung-No Lee, "Profitable Double Spending Attacks," March 5th, 2019 submitted to IEEE Trans. Information Forensics and Securities, downloadable from https://arxiv.org/abs/1903.01711.
- [Lee2] 장재혁, 이흥노, "50%미만 이중 지불 공격", OSIA S&TR Journal, Vol. 32, No. 1, Mar. 2019. (pdf)
- [Lee3] 정현준, 이흥노, "암호화폐 투자와 규제 현황", 한국정보과학회, 정보과학회지, 제 36권, 제 12호, pp. 49-56, Dec, 2018. (pdf)
- [Lee4] 박상준, 김형성, 이흥노, "Introduction to Error-Correction Codes Proof of Work," 블록체인경제 특집호, 대한전자공학회지, June 2019.
- [Lee5] Sangjun Park, HS Kim, Heung-No Lee, "Time-Variant Proof-of-Work Using Error-Correction Codes," to be submitted to IEEE Trans. Information Forensics and Securities.
- [Lee6] Mohamed Yaseen.J, Giljun Jung and Heung-No Lee."Decentralized Framework for Medical Images Based on Blockchain and Inter Planetary File System", The 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society(EMBC 2019), Berlin, Germany, Jul. 23-27, 2019.
- [Lee7] Please visit INFONET home page https://infonet.gist.ac.kr/?page_id=14 for more references.

IITP 정보통신기획평가원

2019년 블록체인 융합기술 개발사업

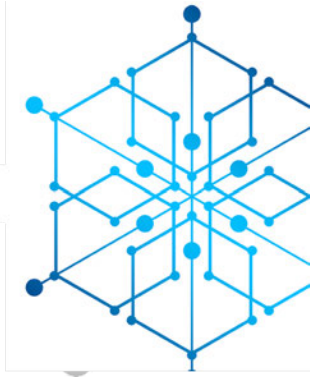
확장가능한 탈중앙화 보안성(DeSecure)

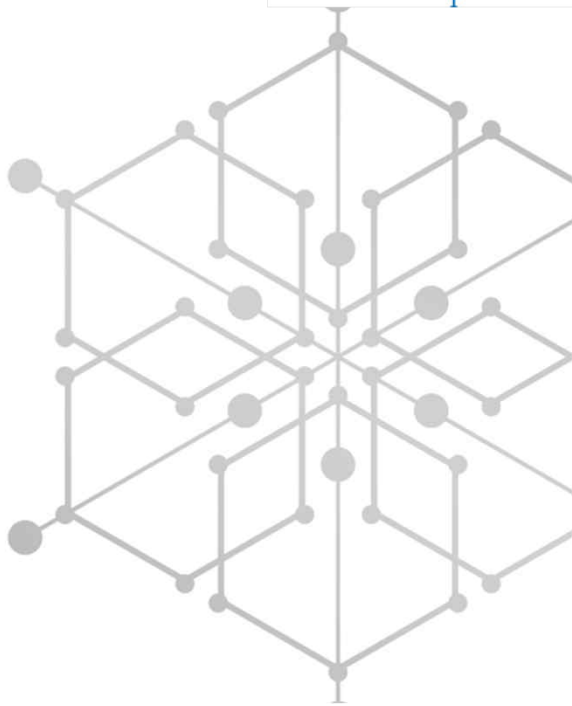
ECCPoW 블록체인

2019.12.05

블록체인 인터넷 경제연구센터 / 센서지능화센터
발표자 : 이흥노 교수

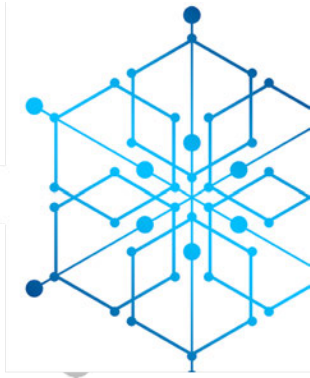
광주과학기술원
Gwangju Institute of Science and Technology

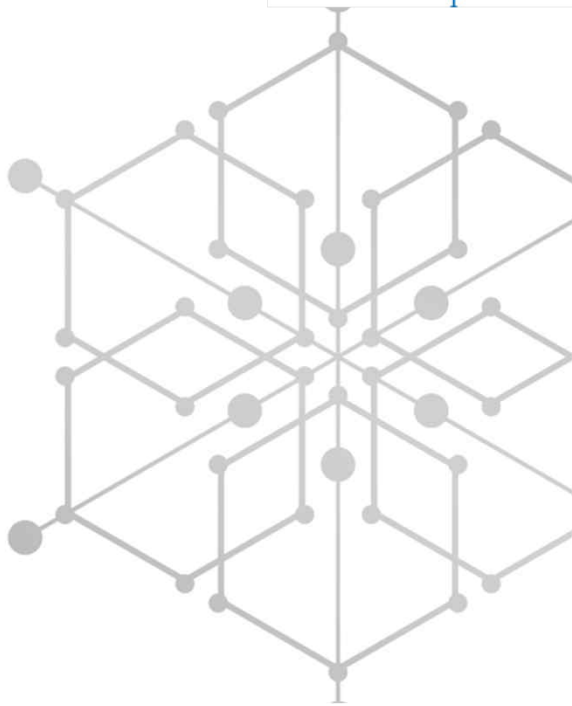




목 차

1. 수행기관 현황
2. 기술개발의 필요성
3. 기술개발의 목표
4. 성과내역 (1차년도 보고)
5. 1단계 기술개발 내용
    5.1 ECCPoW 버전 1.0 개발
    5.2 ECCPoW 개발 및 기존 암호화폐(비트코인) 시스템에 적용
    5.3 ECCPoW 개발 및 기존 암호화폐(이더리움) 시스템에 적용
    5.4 BTCECC/ETHECC 평가 결과
6. 2단계 연구 계획
7. 기술개발의 추진전략·방법 및 추진체계





01

수행기관 현황

수행기관 현황

총괄책임자 연구 역량



연구책임자 | 광주과학기술원 **블록체인 인터넷 경제연구센터 / 센서지능화센터 센터장**

GIST 이흥노 교수

국가스마트도시위원회 민간위원, (사)대한전자공학회 통신소사이어티 회장
우정사업본부 블록체인이니셔티브 민간위원

연구실적

- **국제 논문 (SCI) : 200+편(71+편)**
- 국제 특허 : 출원 9건, 등록 7건
- 국내 특허 : 출원 7건, 등록 22건
- **기술이전 : 2건 (2억 1천만원)**
- 저술 : 8건
- 기술가치평가 : 1건(중앙기술평가원)

대표언론보도

- 프레시안, “GIST, 글로벌 블록체인 개발회사 ‘컨센시스’ 와 MOU” (2019.07.21)
- 스마트시티랩, “블록체인 스마트시티의 유기적 연결의 중추 역할 할 것” (2018.08.24)
- 전자신문, “대한전자공학회 통신소사이어티(회장 이흥노), 19일 블록체인으로 여는 미래’ 워크숍 개최” (2018/04/16)
- 블록체인허브, “블록체·인AI·빅데이터 융합산업 육성 정책토론회”, (2018/12/06)

연구분야

- 블록체인 융합기술
- 블록체인경제
- 지능형 네트워크
- 센서지능화

관련수상이력

- **2019년 대한전자공학회 해동학술상 수상**
- 2016년 GIST 연구상
- **2016년 GIST 대표기술상**
- **2014년 1월 이달의 과학기술자상, 미창부**
- 2013년 기초연구 우수성과 50선, 미창부
- **2012년 국가연구개발 우수성과 100, 미창부**

Ⓒ 2019 GIST

보유기술현황

보유기술

- GIST는 블록체인의 재중앙화 문제 해결을 위한 **부호-암호 기반의 작업증명** 방식 연구 및 특허 확보 및 출원한 **부호-암호 작업증명 특허** (GIST IP)를 기반으로 블록체인 시스템 개발
- ㈜온더는 이더리움관련 개발 프로젝트를 수행, **블록체인 관련 특허 확보**

특허

특허	출원인	출원번호/등록일	핵심기술
부호-암호 화폐 시스템	광주과학기술원	10-2019-0151246 / 2019.11.12	블록체인의 재중앙화 문제 해결을 위한 핵심 기술
블록체인거버넌스	광주과학기술원	10-2019-0084800 / 2019-07-12	블록체인 거버넌스와 규제
블록체인의 거래검증시스템, 및 블록체인이 거래검증방법	광주과학기술원	10-2019-0120655 / 2019-09-30	블록체인거래검증

Ⓒ 2019 GIST

02

기술개발의 필요성

블록체인 트랜잭션에 따른 최적 합의 알고리즘 문제 해결

제안 요청서의 문제 정의

01 블록체인 성능 및 신뢰성을 향상시키기 위하여 트랜잭션 종류와 목적에 따라 최적 합의 알고리즘 문제 해결

02 거래처리 속도, 확장성, 보안 등 현 블록체인의 트릴레마 문제를 해결하고 분산 신뢰 인프라 서비스 제공에 필요한 트랜잭션에 따른 합의 알고리즘 필요

블록체인 트릴레마?

01 블록체인에서 분산성/보안성/확장성의 3가지 특성을 확보하는 문제

02 Decentralized Secure (DeSecure) 블록체인으로 문제 해결!

Ⓒ 2019 GIST

분산성과 보안성 문제 해결의 중요성

분산성을 제고하여 Satoshi Nakamoto가 의도 한 바와 같이
one-cpu-one-vote에 접근 가능하도록 하는 새로운 작업증명 방식의 개발이 필요함

ASIC/FPGA개발, 그리고 채굴 Pool 형성의 억제가 가능하면, 규모의 경제의 효율을 떨어뜨릴 수 있고,
절대 다수의 소규모 채굴 노드의 참여 동기를 부여함

작업증명(Proof-of-Work, PoW)의 효과

01 보안성이 검증된 합의 알고리즘은 작업증명(Proof-of-Work, PoW)

02 가장 규모가 큰 블록체인인 비트코인과 이더리움이 PoW를 사용
지난 수년간 보안 관련 사고 사례가 보고된 적이 없음

03 PoW는 시빌 공격(Sybil Attack)과 이중 지불 공격(Double-Spending Attack)등에 매우 강인하여 화폐 및 금융응용에 적합

© 2019 GIST

분산성과 보안성 문제 해결의 중요성

작업증명(Proof-of-Work, PoW)의 해결과제

01 **ASIC (Application Specific integrated Circuit)의 등장으로 인한 분산성 훼손**

02 ASIC장비와 CPU 혹은 GPU와의 연산속도는 수십만 배 수준으로 차이 발생

03 ASIC을 다수 보유한 전문 채굴업체와 그렇지 않은 소규모 채굴 노드 간의 해시 파워 격차가 크게 발생

04 **규모의 경제 효과에 의한 분산성 훼손**

05 자금력이 풍부한 ASIC장비 투자자들은 소규모 채굴자보다 압도적인 이익을 얻음(기대수익 초선형 양상)

06 반면 투자 여력이 크지 않은 절대다수의 소규모 채굴자들은 참여 동기를 상실함

07 **분산성 약화는 보안성 약화를 유발함**

Ⓒ 2019 GIST

확장성 문제 해결의 도전성

분산성과 보안성을 훼손하지 않고, TPS를 상향하는 확장성 솔루션이 필요함

01 초당 거래 속도(TPS) 상향의 한계

02 TPS는 블록생성주기를 짧게 설정하거나 블록의 크기를 늘려서 달성 가능

03 비트코인, 이더리움 등의 블록체인 (세계적 네트워크 보유)은 전송지연 문제 보유

04 전송지연 비 고려시 수많은 분기 발생으로 이중 지불 공격에 취약함

05 다중 복합구조 블록체인의 필요

06 다수의 블록체인을 용도/지역/서비스 등 목적에 따라 최적화 후 연결하여 확장성 개선 가능

07 분산성과 보안성이 확장된 작업증명은 파라미터 변경으로 용도별 적용가능

08 다중 블록체인들간의 정보를 교환하는 프로토콜을 개발 할 필요가 있음

Ⓒ 2019 GIST

작업증명 이외의 다른 합의알고리즘 사용 가능성은?

분산성/보안성/확장성 3요소를 모두 충족하는 합의알고리즘은 없음

01 Proof-of-Stake, Proof-of-Activity, Proof-of-Publication, Delegated PoS, BFT, PBFT 등 다양한 합의알고리즘들이 제안됨

02 이런 합의알고리즘들의 공통된 특징은 블록생성 및 검증에 관여하는 노드의 숫자를 제한하고, 필요한 절차를 간소화해 TPS를 상향함

03 그러나 선별된 소수의 노드가 블록 생성권한을 갖기 때문에 분산성과 보안성이 훼손될 수 있음

04 합의 알고리즘을 분류하면 크게 세 가지로 나눌 수 있음

05 **블록생성과 검증 단계가 통합된 모델 (비트코인 PoW 모델)**

06 블록생성과 블록검증 단계가 분리된 모델(하이퍼레저, PoA 모델, DPoS 모델)

07 블록생성과 블록검증 단계를 분리해서 시작하지만, 결과적으로 통합하는 모델(이더리움2.0 PoS)

ⓒ 2019 GIST

작업증명 이외의 다른 합의알고리즘 사용 가능성은?

작업증명에는 블록생성과 블록검증이 통합된 모델이 주는 장점이 있음

01 **블록생성과 검증 단계가 통합된 PoW의 이점**

02 참여한 모든 노드는 블록전파 과정에서 비잔티움 장애, Sybil 공격, 이중지불 문제 등에 노출

03 Longest chain wins 룰과 같은 간단한 정책으로 위의 문제를 모두 해결함

04 나아가 전세계에 분포된 모든 노드에 '일관성'있는 상태 갱신을 보장함

05 즉, 매 블록 생성과 검증을 새로운 상태에서 다시 시작할 수 있는 글로벌 참조 값도 제공함

06 인센티브 메커니즘과 결합하여 지속적이고 자발적 참여를 이끌어 냄

07 분산성을 기반으로 네트워크 지연, DDoS공격, 단일 취약점 장애, 비잔티움 문제에 강인하게 대처

ⓒ 2019 GIST

Lex Cryptographia

작업증명은 Lex Cryptographia 실현을 촉진함

01 **Lex Cryptographia란 On-chain 정책의 사회적 확대에 의해 국제사회에서 P2P를 통해 계약을 맺고, 이행과 결산을 신뢰받는 제 3자 개입 없이 합의 할 수 있게 되는 이상적 미래 사회를 의미함 (P. De Filippi 교수)**

02 전통적 사회적 규범에 의존하지 않고 블록체인 프로토콜에 의해 합의를 의미

03 재 시점에서 Lex Cryptographia의 완전한 실현은 현실적이지 않음

04 On-chain 정책의 확대에도 불구하고 사람들 사이에 새로운 권력의 출현은 불가피함 (e.g., 영향력이 큰 소수의 개발자, 압도적 해시 파워의 채굴자)

05 따라서 Off-chain 기존의 정치/경제/법제도 등에 의존하게 됨

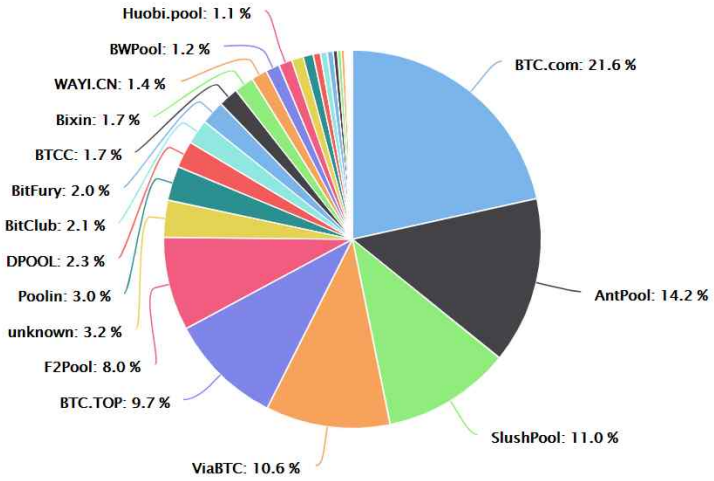
06 만약 작업증명의 분산성이 보장되면 Off-chain 정책의 의존도를 크게 완화할 수 있음

 Ⓒ 2019 GIST

PoW 문제점을 해결, 제안기술의 도전성, 혁신성

PoW 재 중앙화, 전력소비 문제를 해결하는 새로운 작업증명 기술 개발 필요

- ASIC 채굴기의 출현, 소수의 채굴 점유, 재 중앙화 문제 발생
- 소수 채굴업자 블록체인 보상 독점
- 채굴업자 담합, 블록 ·위변조 위험 대두



중앙화 된 블록체인 네트워크



탈 중앙화 된 블록체인 네트워크

Ⓒ 2019 GIST

Proof-of-XXX, Alternatives to PoW

	Pros	Cons	Coins within top 50 rank
PoW (Proof-of-Work)	▪ Strong security ▪ Difficult to produce ▪ Easy to verify	▪ Extreme computing power ▪ 51% attacks ▪ Transaction speed / Transaction throughput	Bitcoin, Ethereum
PoS (Proof-of-Stake)	▪ Energy & hardware efficiency ▪ Much more expensive 51% attacks	▪ Recentralization ▪ The rich-get-richer ▪ “Nothing at stake” problem	Qtum, Stratis
DPoS (Delegated PoS)	▪ Scalability and speed ▪ Energy & hardware efficiency ▪ Encouraging good behavior by real-time voting	▪ Recentralization ▪ DDoS attacks	EOS, NEO
PoA (Proof-of-Activity)	▪ Much more expensive 51% attacks ▪ Decentralization ▪ Validators are randomly selected.	▪ Extreme computing power ▪ Recentralization ▪ The rich-get-richer	decred

 © 2019 GIST

03

기술개발의 목표

기술 개발 개요

GIST 블록체인

- 블록체인의 재중앙화 문제 해결을 위한 부호-암호 기반의 작업증명 방식 연구
- 출원한 암호-부호 작업증명 특허(GIST IP)를 기반으로 블록체인 시스템 개발
- 개발한 블록체인 시스템을 기존 암호화폐(비트코인/이더리움)에 적용 및 확산

1차년도

ECCPoW 코어 1.0 연구 및 개발

확장성	기존 암호화폐(비트코인, 이더리움) 적용
보안성	매 블록 변경 암호퍼즐, 디코더 개발
분산성	블록 생성 주기 제어

2차년도

ECCPoW 코어 2.0 연구 및 개발

확장성	다중 복합구조 블록체인, 플라즈마 EVM 적용
보안성	거래검증 최적화 알고리즘, 데이터 가용성 개선
분산성	블록 생성 제어 방법 고도화

Ⓒ 2019 GIST

트릴레마 문제 해결 전략

1단계	ASIC 채굴에 저항성이 있는 ECCPoW를 개발 **(분산성 향상)**
1단계	비트코인과 이더리움에 ECCPoW를 탑재 및 하드포크
2단계	이중 지불 공격 등 블록체인에서 문제가 되는 공격들에 대한 방지 알고리즘을 개발 **(보안성 향상)**
2단계	다중 복합구조의 DeSecure 블록체인들을 설계하여 **거래의 종류/목적에 따라 독립적으로 최적화**하고, 이들을 계층적으로 연결함 **(확장성 향상)**
2단계	이더리움의 레이어2 확장성 솔루션인 플라즈마를 DeSecure 블록체인에 적용 **(유연성 향상)**

- 오류 정정 부호 (Error Correction Codes)를 활용한 새로운 Anti-ASIC PoW 알고리즘인 ECCPoW를 개발하고자 함
- DeSecure 블록체인은 비트코인과 이더리움에 ECCPoW를 탑재하여 하드포크한 새로운 블록체인을 의미함

Ⓒ 2019 GIST

04

성과내역 (1차 년도 보고)

연구 개발 성과 _ 특허

(목표) 2건의 국내 출원, 2건의 국외 출원
(성과) 4건의 국내 출원, 2건의 국외 출원 – 1건은 현재 진행 중
(달성도) 국내 : 200% 달성 확정, 국외 : 100% 달성 예정

번호	지식재산권 등 명칭	국명	출원		
			출원인	출원일	출원번호
1	부호-암호 화폐 시스템[증빙자료 p.154]	KOR	광주과학기술원	2019-11-12	10-2019-0151246
2	블록체인의 거래검증시스템 및 블록체인이 거래검증방법[증빙자료 p.76]	KOR	광주과학기술원	2019-09-30	10-2019-0120655
3	블록체인거버넌스[증빙자료 p.113]	KOR	광주과학기술원	2019-07-12	10-2019-0084800
4	블록체인의 거래검증시스템, 및 블록체인이 거래검증방법[증빙자료 p.76]	PCT	광주과학기술원	진행 중	

 Ⓒ 2019 GIST

연구 개발 성과 _ 국내 및 국외 논문 게재

(목표) 0건의 국내, 0건의 국외
(성과) 2건의 국내 학회, 2건의 석사 학위 논문, 4건의 국외 논문 준비 중

번호	논문명	학술지명	국명	발행기관	SCI 여부	게재일
1	A Blockchain For The Collision Avoidance And The Recovery of Crashed UAVs[증빙자료 p.199]	대한전자공학회 2019년도 하계종합학술대회	국내	대한전자공학회	비 SCI	2019-06-07
2	A Permissioned ISPs Blockchain For The End to End Quality of Service Enhancement[증빙자료 p.192]				비 SCI	2019-06-07
3	Time-Variant Proof-of-work using Error-Correction codes	Under revision	국외			
4	Profitable Double-Spending Attacks	Under revision	국외			
5	A Blockchain-Based Distributed Patient-Centric Image Management System	Under revision	국외			
6	Fine difficulty control for ECCPoWs	Under revision	국외			

Ⓒ 2019 GIST

연구 개발 성과 _ 국내 및 국제 학술대회 발표

8건의 국내 학술대회 발표
1건의 학술 대회 조직

번호	논문명	발표자	발표일시	장소
1	블록체인 법학회 컨퍼런스	이흥노	2019.10.22	서울
2	블록체인 법학회 컨퍼런스	이흥노	2019.08.28	서울
3	대한전자공학회 2019년도 하계종합학술대회	Yaseen Jabarulla	2019.06.27	제주도
4	대한전자공학회 2019년도 하계종합학술대회	Michele Scarlato	2019.06.27	제주도
5	블록체인으로 여는 미래사회 Workshop	이흥노	2019.06.17	서울
6	ETHcon Korea 2019	이흥노	2019.05.27	서울
7	ETHcon Korea 2019	장재혁	2019.05.27	서울
8	JCCI 2019	이흥노	2019.05.01	강릉
9	한국 ICT 금융학회 세미나	이흥노	2019.04.19	국회 도서관

블록체인으로 여는 미래 사회
Workshop



Ⓒ 2019 GIST

성과내역 (1단계 보고)

연구 개발 성과 _ 기술 이전

(목표) 1건의 기술 이전, 1000만원의 기술료
(성과) 1건의 기술 이전, 2000만원의 기술료
(달성도) 기술료 200%달성

번호	기술 이전 유형	기술 실시 계약명	기술 실시 대상 기관	기술 실시 발생일	기술료 (해당 연도 발생액)	누적 징수 현황
1	전용실시	DeSecure 블록체인 기술	LiberVance (리버밴스)	2019.12 (체결 예정)	20,000,000 (선급기술료)	20,000,000 (선급기술료)

LiberVance는 본 과제의 연구책임자가 설립한 회사
12월 중으로 체결 예정 – (기술 이전 의향서 완료) [증빙자료 p.204]
대상은 본 과제의 실적으로 보고된 광주과학기술원의 특허
중기부 예비창업자패키지 과제선정 (2019.10.1 ~ 2020.7.31, 5500만원)


기 술 이 전 의 향 서

기술명	DeSecure 블록체인 기술
연구책임자	이홍노
출원인	광주과학기술원(GIST)
관련특허	1 블록체인의 거래검증시스템, 및 블록체인의 거래검증방법 / 대한민국 / 광주과학기술원 / 2019-09-30 / 10-2019-0120655 2 블록체인거버넌스 / 대한민국 / 광주과학기술원 / 2019-07-12 / 10-2019-0084800 3 Scalable DeSecure EDPoR Blockchains Algorithm 블록체인의 거래검증시스템, 및 블록체인의 거래검증방법 / 대한민국 / 광주과학기술원 / 2019-08-13 / 10-2019-0099153 4 / PCT / 광주과학기술원 / 진행 중
기술이전형태	소유권 이전 ( ) 전용실시권 (⇒) 통상실시권 ( ) 기타 ( )
대상기업	㈜리버밴스
주요협의내용	선급금: 2천만원 경상기술료: 협의 중 기술이전 계약시기: 2019.12 이내

상기 기술(특허)에 대하여 광주과학기술원(GIST)은 향후 ㈜리버밴스와 협의를 통하여 기술이전 및 기술사업화에 협력할 의사가 있음을 알려드립니다.

2019년 11월 21일
광주과학기술원 과학기술응용연구단장 박 우 진 (직인)

(주)리버밴스 대표이사 귀하

Ⓒ 2019 GIST

연구 개발 성과 _ 연구 인력 양성 및 고용 창출

(목표) 0건의 연구 인력 양성, 2건의 고용 창출
(성과) 2건의 연구 인력 양성, 5건의 고용 창출
(성과도) 연구 인력 양성은 초과 달성, 고용 창출은 250% 달성

번호	소속	이름	학위	제목
1	광주과학기술원	정길준 [증빙자료 p.523]	석사	MyData Blockchain
2	광주과학기술원	박상준 [증빙자료 p.522]	박사	Compressed Sensing for Support Set Reconstruction : Information-Theoretic Results and Design of Practical Algorithm

번호	채용기관	이름	채용 직위	채용	직무내용
1	광주과학기술원	이수민	위촉 연구원	2019.8.1. ~	블록체인 개발/경제/하드포크
2	광주과학기술원	이명은	위촉 연구원	2019.11.20. ~	블록체인 연구 및 기획
3	광주과학기술원	조성민	인턴	2019.3.1 ~ 2019.7.31	블록체인 개발
4	광주과학기술원	김석주	인턴	2019.3.1 ~ 2019.7.31	블록체인 개발
5	광주과학기술원	Michele Scarlato	인턴	2019.4.1 ~	블록체인 관련조사 및 연구


MyData Blockchain

Advisor: Heung-No Lee

by

Giljun Jung

School of Electrical Engineering and Computer Science
Gwangju Institute of Science and Technology

A thesis submitted to the faculty of the Gwangju Institute of Science and Technology in partial fulfillment of the requirements for the degree of Master of Science in the School of Electrical Engineering and Computer Science

Gwangju, Republic of Korea
2019. 06. 10.
Approved by

Professor Heung-No Lee
Committee Chair


Compressed Sensing for Support Set Reconstruction: Information-Theoretic Results and Design of Practical Algorithm

Advisor: Heung-No Lee

by

Sangjun Park

School of Electrical Engineering and Computer Science
Gwangju Institute of Science and Technology

A thesis submitted to the faculty of the Gwangju Institute of Science and Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Electrical Engineering and Computer Science

Gwangju, Republic of Korea
May 31, 2019
Approved by

Professor Heung-No Lee
Committee Chair


기간제근로자 고용계약서


연구직 증사 확인서

Ⓒ 2019 GIST

연구 개발 성과 _ 기술요약정보 & 보고서 원문

(목표) 1건의 기술요약정보, 1건의 보고서 원문
(성과) 1건의 기술요약정보, 3건의 보고서 원문
(성과도) 기술요약정보 100% 달성, 보고서 300% 달성

연도	기술명	요약 내용	기술 완성도
2019	DeSecure 기술	LDPC 코드명세서, DeSecure Blockchain 매뉴얼 (Code Level, Bitcoin-ECC Build 및 Docker 사용 법) [증빙자료 p.192] **총 58페이지**	V. 1.0

연도	보고서 구분	발간일
2019	DeSecure 블록체인의 미래 성장 전략 도출 [증빙자료 p.265] **총218페이지**	2019. 11
2019	ECCPoW 적용 영역 연구	2019. 12월 중
2019	ECCPoW Blockchain 확장가능 비즈니스 모델 개발 [증빙자료 p.483] **총38페이지**	2019. 11

 Ⓒ 2019 GIST

연구 개발 성과 _ 프로그램 산출물

부호-암호 작업증명 (ECCPoW)
부호-암호 작업증명이 탑재된 비트코인/이더리움

이름	목적	Github 주소
ECCPoW	암호 퍼즐 생성 (디코딩) 암호 퍼즐 해결 유무 판단 암호 퍼즐 난이도 조절	github.com/cryptoecc/bitcoin_ECC 최초생성일 : 2019.5.23 commit : 66회
BTCECC	ECCPoW가 탑재된 비트코인	
ETHECC	ECCPoW가 탑재된 이더리움	github.com/cryptoecc/go-ethereum_ECC 최초생성일 : 2019.5.20 commit : 82회

Ⓒ 2019 GIST

05

기술개발 내용 (1단계 보고)

5.1 ECCPoW 버전 1.0 개발
5.2 ECCPoW 개발 및 기존 암호화폐(비트코인) 시스템에 적용
5.3 ECCPoW 개발 및 기존 암호화폐(이더리움) 시스템에 적용
5.4 BTCECC/ETHECC 평가 결과

당해 (연·도단계) 연구 개발 목표 및 결과

ECCPoW 코어 버전 1.0 연구 및 개발

세부 연구 목표	연구개발 수행 내용	연구 결과	담당
ECCPoW 버전1.0 개발	▪ 매 블록마다 바뀌는 암호 퍼즐 생성 연구 ▪ 매 블록마다 바뀌는 암호 퍼즐 디코더 개발	▪ 매 블록마다 바뀌는 암호 퍼즐 생성/디코더 개발 완료	주관 기관
ECCPoW 개발 및 기존의 암호화폐(비트코인) 시스템에 적용	▪ 블록 생성 주기 제어 방법 연구 ▪ 시스템 테스트 및 고도화	▪ 블록 생성 주기 제어 방법 연구 및 구현 완료 ▪ 시스템 테스트 및 고도화 완료	
ECCPoW 개발 및 기존의 암호화폐(이더리움) 시스템에 적용)	▪ go-ethereum 패키지 연구 및 분석 (합의 계층 내부구조 변경, 공통 계층 최적화, 응용 계층 개발, 데이터 계층 개발) ▪ 블록 생성 주기 제어 방법 연구 ▪ 시스템 테스트 및 고도화	▪ go-ethereum에 ECCPoW 탑재 완료 ▪ 블록 생성 주기 제어 방법 연구/구현 완료 시스템 테스트 및 고도화 완료	참여 기관

 Ⓒ 2019 GIST

부호 – 암호 작업증명 버전 1.0 주요 함수

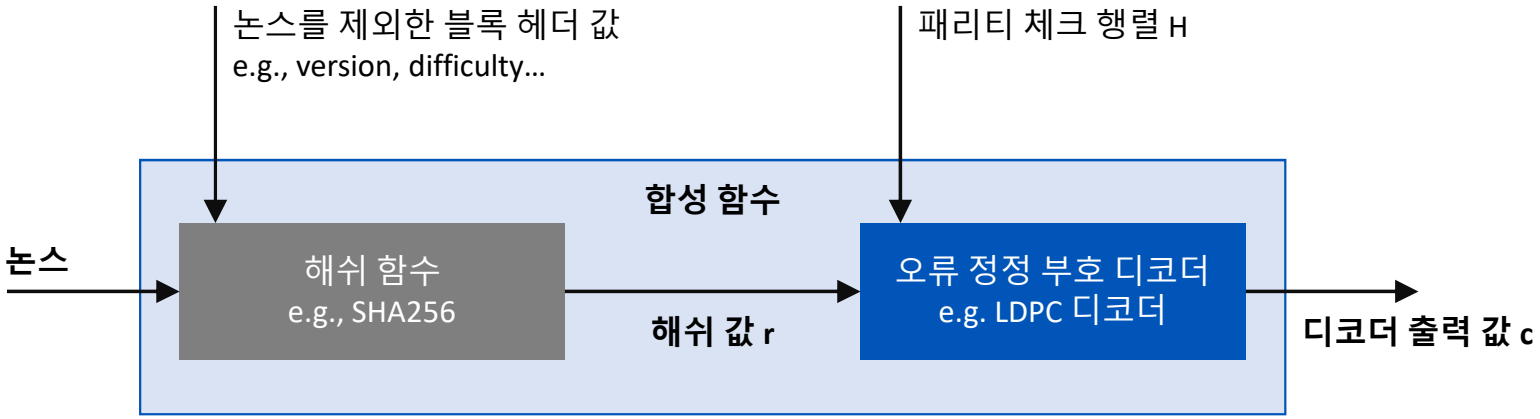
함수 이름	인력 값	설명
set_difficulty	int level	암호 퍼즐의 난이도 설정
generate_seeds	uint64_t hash	이전 해시 값으로부터 시드를 결정
generate_H	none	시드를 이용하여 패리티 체크 행렬을 생성. 행렬의 크기는 level에 의해 결정
generate_hv	Unsigned char hash[]	해시 값을 0/1로 구성된 문자들로 전환하여 hash_vector에 저장
decoding	none	메시지 전달 알고리즘을 사용하여 디코딩하고 그 결과를 output_word에 저장
decision	none	output_word가 암호 퍼즐의 해결 유무를 만족하는지 판단
binary_to_hex	none	2진 문자열인 디코더의 출력 값을 16진수 문자열로 치환

- 부호-암호 작업증명의 주요 함수
- https://github.com/cryptoecc/bitcoin_ECC/blob/ecc-0.1/src/ldpc/ldpc.h
- C++언어로 작성

ⓒ 2019 GIST

부호 – 암호 작업증명

- 부호-암호 작업증명 (ECCPoW) 정의 :
  오류 정정 부호와 작업증명을 동시에 사용하여 분산성/보안성을 확보하기 위해 개발된 작업증명 기술임.
- ECCPoW은 매 블록마다 변하는 합성 함수를 사용함.



- (**시변성**) 패리티 체크 행렬을 매 블록 마다 변경함으로써, 함성 함수를 매 블록 마다 변경 할 수 있음.
- (**무한성**) 무한히 많은 패리티 체크 행렬을 사용하면, 무한히 많은 합성 함수를 생성 할 수 있음.

© 2019 GIST

부호 – 암호 작업증명 _ 5가지 특성

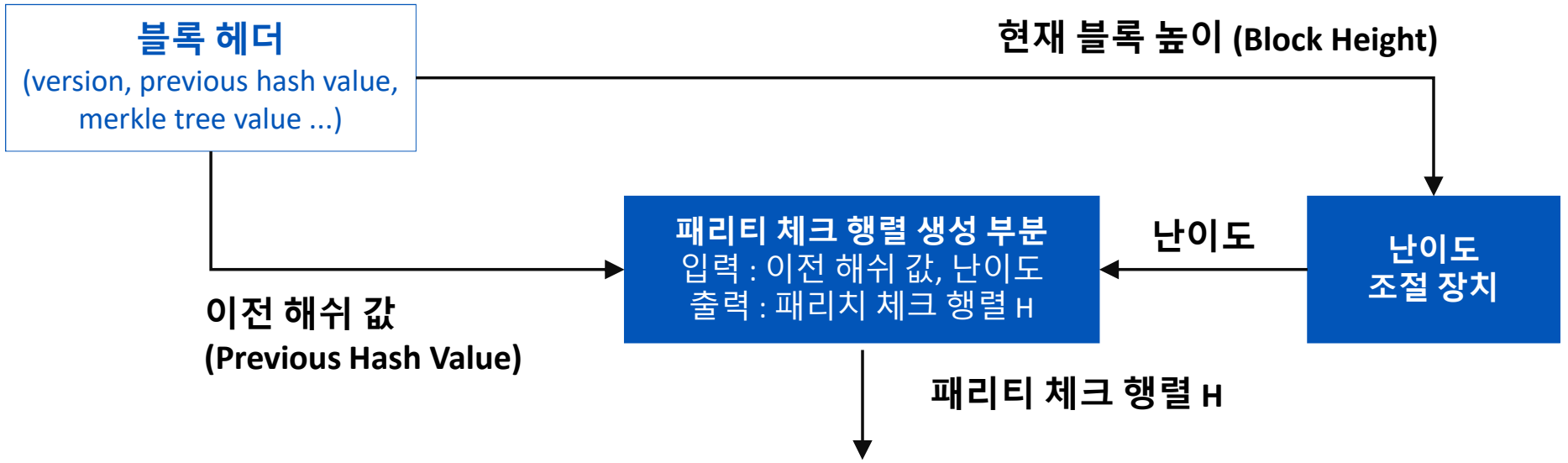
요건 1	암호 퍼즐은 풀기 어려우나, 반대로 검증은 쉬워야 함.
요건 2	풀린 암호 퍼즐의 정답은 재사용되지 않아야 함.
요건 3	암호 퍼즐의 난이도는 조절 가능해야 함.
요건 4	암호 퍼즐은 외부 공격으로부터 견고한 저항성을 가져야 함.
요건 5	**(시변성/무한성) 작업증명에 쓰이는 함수는 매 블록마다 변화하여 바꿀 수 있어야 함.** 부연 설명 : 매 블록마다 패리티 체크 행렬이 변경됨. 따라서 디코더의 입력 값과 출력 값의 관계가 바뀌므로 합성 함수의 입력 값과 출력 값이 관계 또한 바뀜. 작업증명에 사용되는 함수의 변경을 의미함.

- 요건 1 – 4는 작업증명이라면 가져야 할 기본 요건.
- 요건 5는 고유 특징으로써 시변성과 무한성을 뜻함.
- 패리티 체크 행렬은 "블록 길이", "부호 비율" 등을 변경함으로써 서로 다른 행렬들을 무한히 많이 생성 가능함.

ⓒ 2019 GIST

매 블록마다 바뀌는 암호 퍼즐 생성 (1)

- 난이도는 "블록 길이" 및 "부호의 비율"을 결정
- 이전 해쉬 값을 이용해 Gallager 생성 방법의 순열 순서를 결정

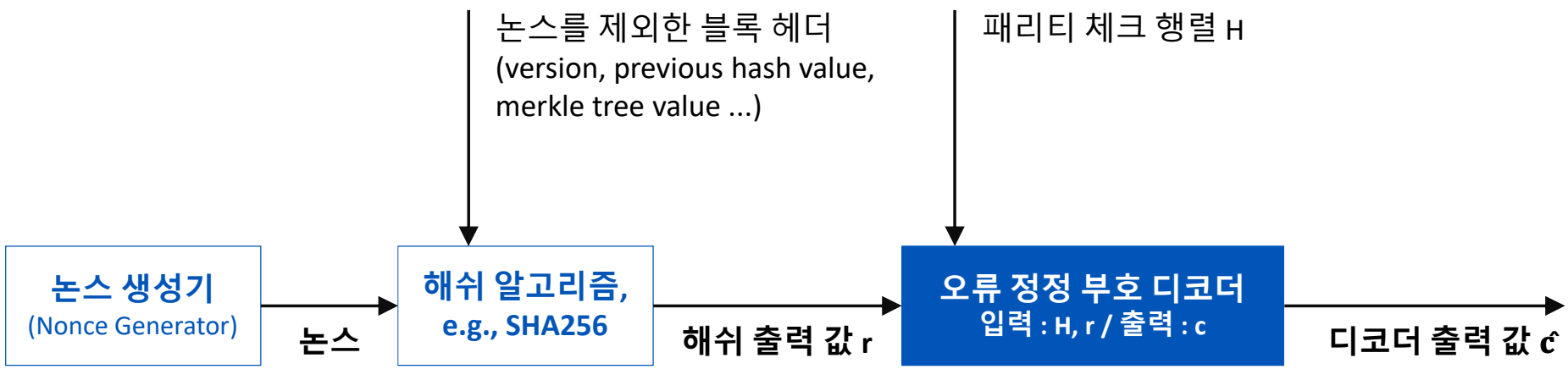


 Ⓒ 2019 GIST

광주과학기술원 Gwangju Institute of Science and Technology

매 블록마다 바뀌는 암호 퍼즐 디코더 개발 (1)

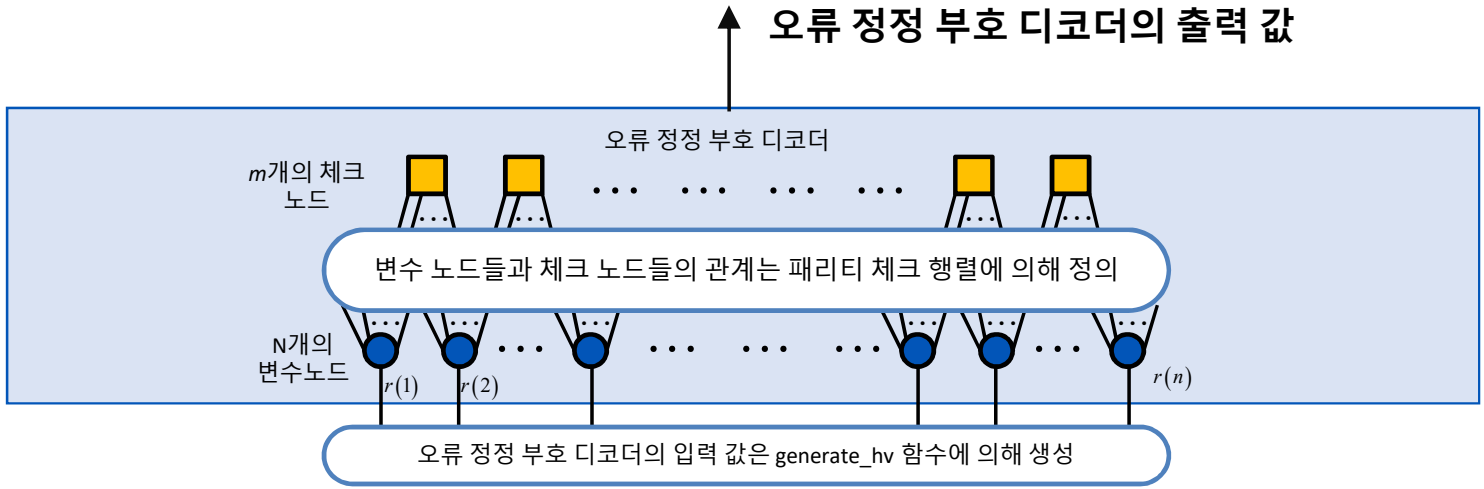
- 오류 정정 부호 디코더 입력 값 생성



© 2019 GIST

매 블록마다 바뀌는 암호 퍼즐 디코더 개발 (2)

- 오류 정정 부호 디코더 입력 값 생성



Decoding part

- 패리티 체크 행렬의 변경은 변수/체크 노드들의 관계의 변경을 야기함.
- 즉, 동일한 입력 값을 취득하더라도, 패리티 체크 행렬이 변경되면 상이한 값을 산출함.
- 디코딩은 메시지 전달 알고리즘에 의해 이뤄짐.

Ⓒ 2019 GIST

암호 퍼즐 해결 유무 판단 기준

- 2가지 서로 다른 암호 퍼즐 해결 유무 판단 기준을 제시함

기준 1	**독자적인 방식 :** 디코더의 결과 값이 a) 부호이고, b) 특정 해밍 가중치를 가지면 해결 한 것으로 판단.
기준 2	**기존 작업증명 방식 :** 디코더의 결과 값을 다시 해싱 하여 나온 값이 특정 값 보다 작으면 해결 한 것으로 판단.

- **(기준 1)** 디코더의 출력 값 c가 다음 2가지 조건들을 만족시키면, 암호 퍼즐을 해결 하는 것으로 판단
  (조건 1) 출력 값이 부호일 것.
  (조건 2) 출력 값의 해밍 가중치가 주어진 집합 S의 원소 일 것.

조건 1이 만족할 확률 $p_1 := \Pr\{\mathbf{c} | \mathbf{Hc} = \mathbf{0}\}$

조건 2가 만족할 확률 $p_2 := \Pr\{\|\mathbf{c}\|_h \in S\}$

 Ⓒ 2019 GIST

기준 1 _ 부호의 특징을 이용한 것

- 조건 1과 조건 2를 만족시킬 확률을 구하고, 이를 통해 암호 퍼즐의 난이도를 총 380개로 세분화 함.

Level	n	w_c	w_r	집합 S	$p := \Pr\{\mathbf{c}\|\mathbf{Hc}=\mathbf{0}\} \times \Pr\{\|\mathbf{c}\|_h \in S\}$
1	32	3	4	{10,12,…,20,22]	$\approx 3.07 \times 10^{-5}$
2	32	3	4	{10,12,…,14,16}	$\approx 2.02 \times 10^{-5}$
				…	
379	128	3	4	{34,94}	$\approx 5.12 \times 10^{-23}$
380	128	3	4	{34}	$\approx 2.60 \times 10^{-23}$

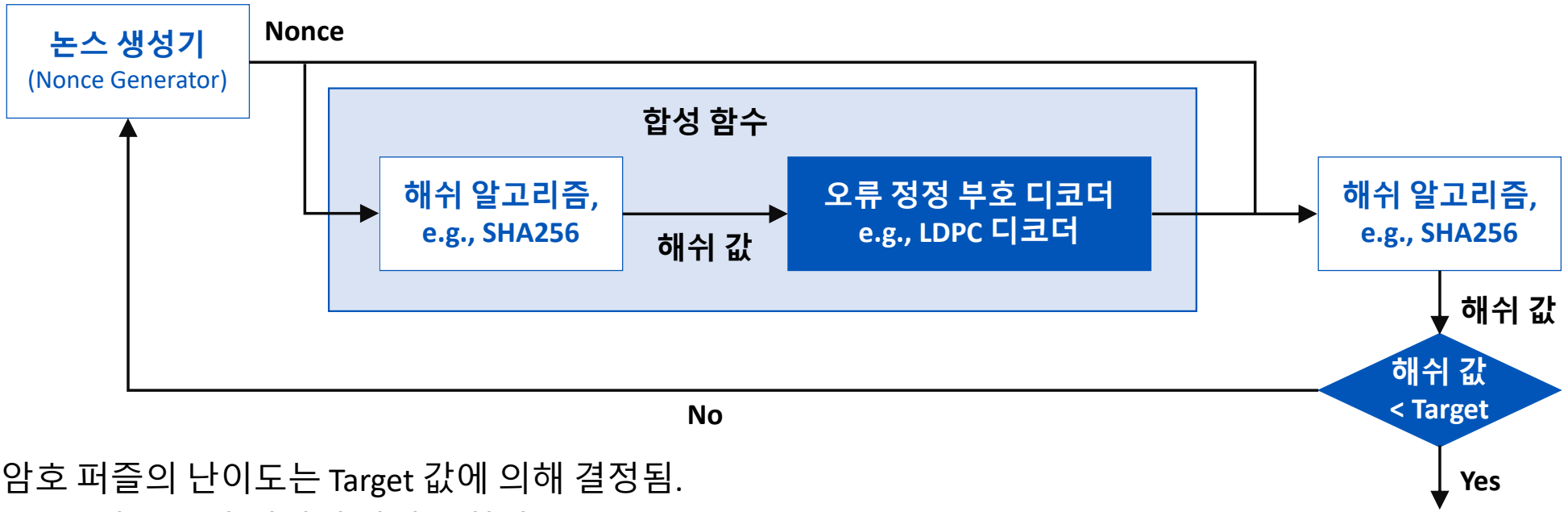
난이도 하락

난이도 상승

© 2019 GIST

기준 2 _ 기존 작업 증명 것

- 합성 함수의 출력 값을 다시 한번 해싱하여 나온 결과 값으로 판단함.



- 암호 퍼즐의 난이도는 Target 값에 의해 결정됨.
- Target 값을 증가 시키면 난이도 하락
- Target 값을 감소 시키면 난이도 상승

ⓒ 2019 GIST

부호-암호 작업증명 버전 1.0 주요 함수

	함수 이름	입력 값	설명
암호 퍼즐 난이도 조절	set_difficulty	int level	암호 퍼즐의 난이도 설정
암호 퍼즐 생성	generate_seeds	uint64_t hash	이전 해시 값으로부터 시드를 결정
	generate_H	none	시드를 이용하여 패리티 체크 행렬을 생성 행렬의 크기는 level에 의해 결정
암호 퍼즐 디코더	generate_hv	unsigned char hash[]	해시 값을 0/1로 구성된 문자들로 전환하여 hash_vector에 저장
	decoding	none	메시지 전달 알고리즘을 사용하여 디코딩하고 그 결과를 output_word에 저장
암호 퍼즐 해결 유무 판단	decision	none	output_word가 암호 퍼즐의 해결 유무를 만족하는지 판단
	binary_to_hex	none	2진 문자열인 디코더의 출력 값을 16진수 문자열로 치환

- https://github.com/cryptoecc/bitcoin_ECC/blob/ecc-0.1/src/ldpc/ldpc.h
- C++언어로 작성

Ⓒ 2019 GIST

ECCPoW 하드포크의 의미는 합의엔진 교체

블록체인 프로그램 코어 (블록체인 구성 요소 3가지)

1. Web server interface networking of peers

- Node registration, get-address, give-address
- Full node or light node
- Communication among the wallets and the miners

2. Wallet for TX generations

- Make private and public keys, address, store UTXOs, make TX, put signature, announce it to the neighbor, check to see if the TX is supported by the blockchain.

합의엔진

3. Consensus Mechanism

- **Data :** Genesis block + regular blocks, one block every 10 min, block-size 1Mbyte
- **Protocol :** consensus, block header, difficulty level adjustment, …
- **Mining :** Get the longest chain, validate it and all transactions within it, get transactions from mempool and form a block, run SHA repeatedly until you hit a good hash, put the proof into the block header, and attach the proofed block to the longest chain, and make announcement ASAP.

- 비트코인/이더리움의 작업증명을 ECCPoW로 교체

Ⓒ 2019 GIST

ECCPoW가 적용된 비트코인 (BTCECC)

- 비트코인의 블록 헤더를 그대로 사용.
- BTCECC (ECCPoW가 적용된 비트코인)의 제네시스 블록 생성.
- 비트코인의 작업증명과 관련된 부분을 ECCPoW의 것으로 교체.
- 비트코인의 난이도 조절과 관련된 부분을 변경.

<주요 변경 및 추가 함수>

타입	함수 이름	위치	함수의 역할
변경	CMainParams	chainparms.cpp	제네시스 블록 생성 함수
추가	CheckProofOfWork	pow.cpp	암호 퍼즐 해결 유무 판단 함수
	GetLevelfromnBits	pow.cpp	비트코인의 난이도를 ECCPoW 암호 퍼즐 난이도로 맵핑

Ⓒ 2019 GIST

기본 비트코인의 난이도 조절

- 2016개의 블록 채굴 후에 난이도 조절이 발생하게 설계됨.
- 관련 함수는 CalculateNextWorkRequired

```
unsigned int CalculateNextWorkRequired(const CBlockIndex* pindexLast, int64_t nFirst
{
    int64_t nActualTimespan = pindexLast->GetBlockTime() - nFirstBlockTime;
    if (nActualTimespan < params.nPowTargetTimespan / 4)
        nActualTimespan = params.nPowTargetTimespan / 4;
    if (nActualTimespan > params.nPowTargetTimespan * 4)
        nActualTimespan = params.nPowTargetTimespan * 4;

    const arith_uint256 bnPowLimit = UintToArith256(params.powLimit);
    arith_uint256 bnNew;
    bnNew.SetCompact(pindexLast->nBits);
    bnNew *= nActualTimespan;
    bnNew /= params.nPowTargetTimespan;

    if (bnNew > bnPowLimit)
        bnNew = bnPowLimit;

    return bnNew.GetCompact();
```

01 급격한 난이도 변경을 막기 위함

02 난이도 변경을 위해 Target 값을 증감

03 만일 변경된 Target이 최소 값 보다 작으면, 최소 값으로 설정

- CalculateNextWorkRequired는 매 블록을 생성/검증 할 때 마다 호출됨.

Ⓒ 2019 GIST

BTCECC 난이도 조절

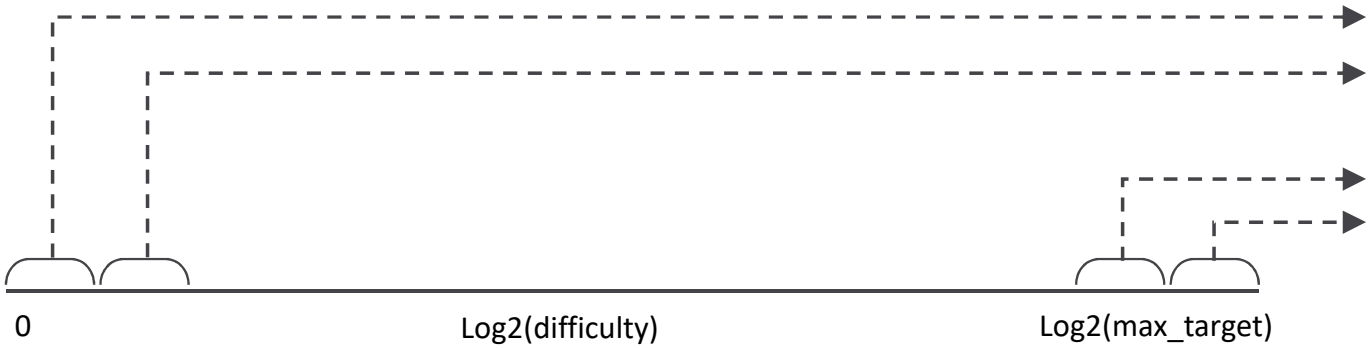
- ECCPoW 암호 퍼즐의 난이도를 380개로 세분화
- nBits를 이용해 difficulty를 계산하고, 이 difficulty 값을 이용해 기 제작된 난이도 테이블로 맵핑

```
int GetLevelfromnBits(uint32_t nBits)
{
    int level = 1;
    double dDiff = 0.0, log_dDiff = 0.0;
    dDiff = GetDifficulty(nBits);
    log_dDiff = log2(dDiff);
    level = floor(log_dDiff / 0.3) + 1;

    if (level >= 380)
        level = 380;
    if (level <= 1)
        level = 1;

    return level;
}
```

01 ECCPoW 암호 퍼즐 난이도로 맵핑

02 현재는 380레벨만 지원

Level	n	w_c	w_r	집합 S
1	32	3	4	{10,12,…,20,22]
2	32	3	4	{10,12,…,14,16}
				…
379	128	3	4	{34,94}
380	128	3	4	{34}



 © 2019 GIST

BTCECC 작업증명

- pow.cpp의 CheckProofOfWork 함수를 변경

```
bool CheckProofOfWork(CBlockHeader block, const Consensus::Params& params)
{
    bool fNegative; bool fOverflow; bool result = false;

    arith_uint256 bnTarget;
    bnTarget.SetCompact(block.nBits, &fNegative, &fOverflow);

    // Check range
    if (fNegative || bnTarget == 0 || fOverflow || bnTarget > UintToArith256(params.powLimit))
        return result;

    int level = GetLevelfromnBits(block.nBits);
    LDPC *ldpc = new LDPC;
    ldpc->set_difficulty(level);
    ldpc->initialization();
    ldpc->generate_seeds(UintToArith256(block.hashPrevBlock).GetLow64());
    ldpc->generate_H();
    ldpc->generate_Q();
    ldpc->generate_hv((unsigned char*)block.GetHash().ToString().c_str());
    ldpc->decoding();
    if (ldpc->decision())
        result = true;

    delete ldpc;
    return result;
```

01 nBits의 유효성 검사

02 암호 퍼즐 난이도 결정

03 암호 퍼즐 생성 및 검증

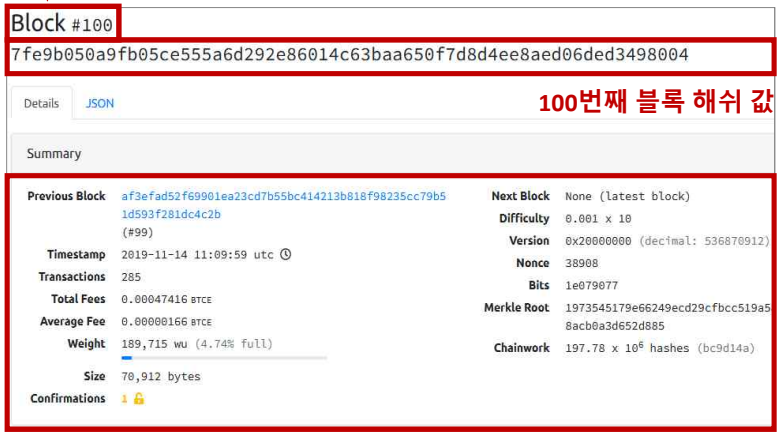
 Ⓒ 2019 GIST

BTCECC 블록 탐색기

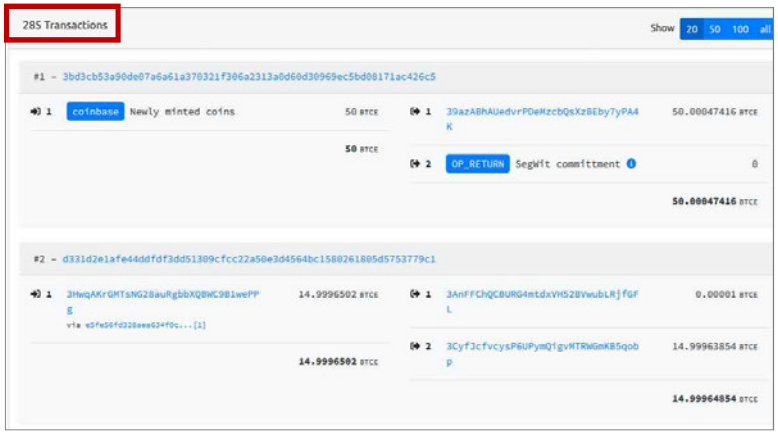


100번째의 블록

100번째 블록에 담긴 트랜잭션 수



100번째 블록 해쉬 값



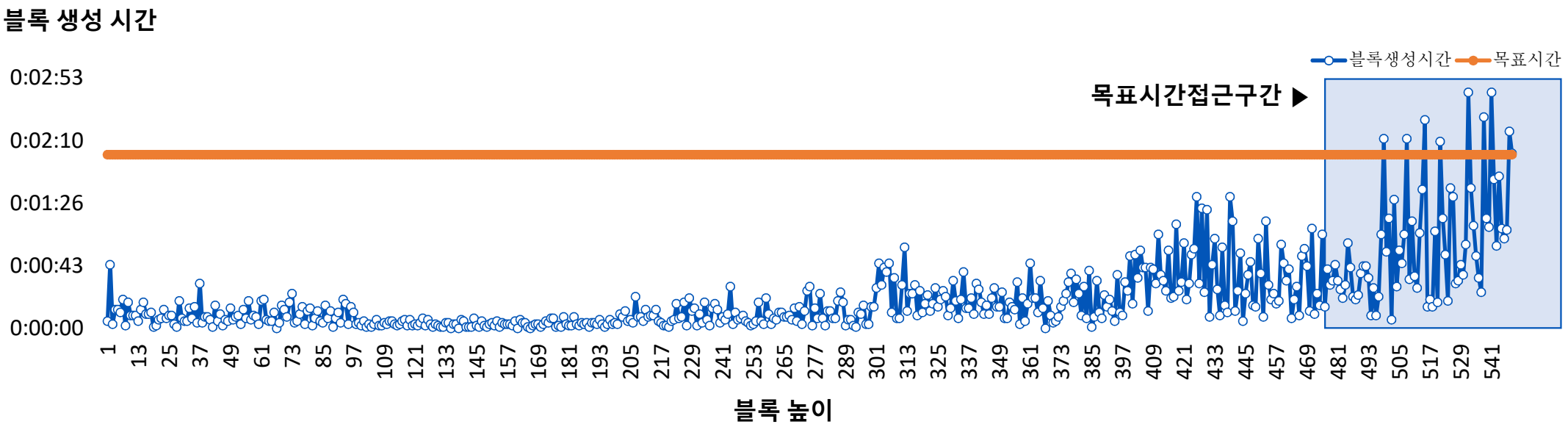
블록 상세 정보

* 블록체인 탐색기(http://15.164.244.139/blocks)에서 확인 가능

© 2019 GIST

BTCECC 난이도 조절 기능 검증

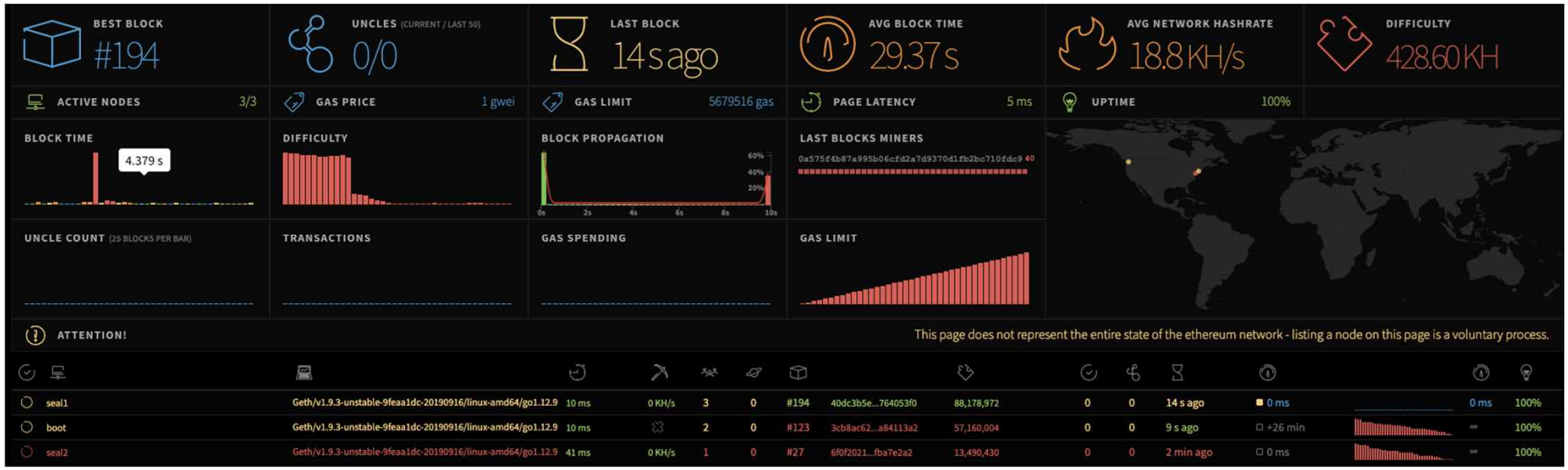
<블록생성시간-블록높이 그래프(블록 생성 목표시간 2분)>



- 난이도 변경 높이 이전의 총 블록 생성시간을 계산하여 생성시간이 빠르면 난이도를 증가시키고 낮으면 난이도를 감소시킴
- 난이도 증감을 통해 목표 시간 접근 구간 확인 가능
- 위의 그래프는 블록체인 탐색기(http://13.209.74.13/blocks)를 통해 확인 가능

Ⓒ 2019 GIST

puppeth를 이용하여 배포된 프라이빗 네트워크



- puppeth를 통해 배포된 ETHECC의 대시보드
- 평균 블록 생성시간 현재 블록 넘버, 마지막 블록 생성 후 경과한 시간 등을 확인 가능
- 대시보드와 통신하고 있는 여러 노드들의 상태를 한번에 관찰 가능
- 노드숫자 4대, 네트워크 생성초기라 29.37초이며 시간에 지남에따라 10~15초 형성

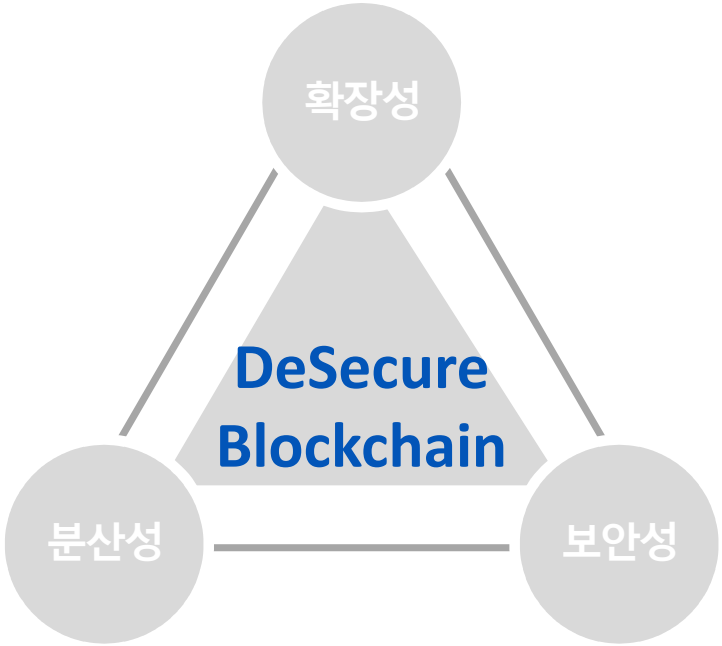
 Ⓒ 2019 GIST

06

2단계 연구계획

2단계 연구 계획

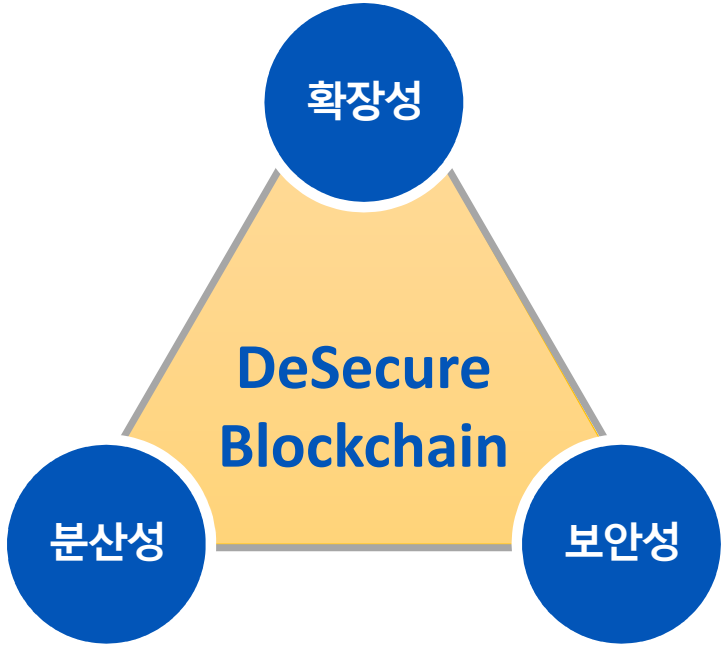
블록체인 하드포크 (비트코인, 이더리움)



ECCPoW 1.0 **작업증명 보안성 수준**

1단계 연구

다중 복합구조 블록체인, 플라즈마



ECCPoW 2.0 **이중지불공격 방지**

2단계 연구

 Ⓒ 2019 GIST

2단계 연구 계획 요약

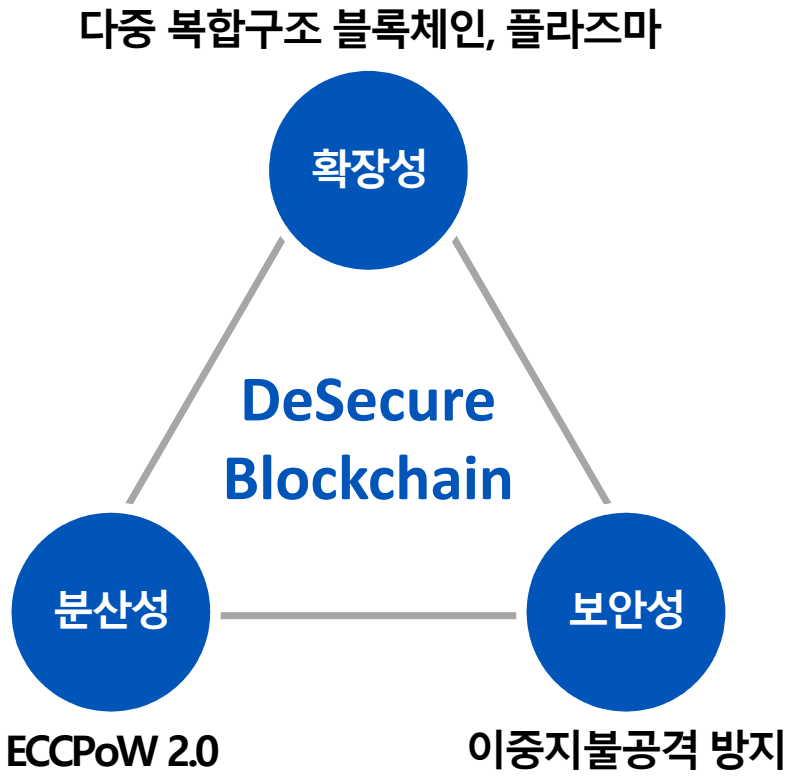
확장성 강화
다중 복합구조 블록체인 연구

보안성 강화
거래검증횟수 최적화 알고리즘 개발

유연성 확대
플라즈마 모델 연구, 개발 및
DeSecure 블록체인에 적용

보안성 강화
플라즈마의 데이터 가용성
문제 개선에 관한 연구 및 개발

ECCPoW 코어 2.0 개발
블록생성 제어 방법 고도화



 Ⓒ 2019 GIST

2단계 연구 계획 목표

세부 연구 목표		연구개발 수행 내용	담당
분산성	**ECCPoW 코어 2.0 개발**	▪ 블록생성주기 및 난이도제어 고도화 ▪ 난이도별 블록생성 확률 분석 ▪ 난이도 조절 규칙 연구	주관 기관
확장성	**다중 복합구조 블록체인 연구**	▪ 다중 복합구조 블록체인 설계 ▪ 트랜잭션의 용도별 블록체인 파라미터 최적화 ▪ 다중 블록체인간 정보교환 프로토콜 연구	
보안성	**거래 검증 횟수 최적화 알고리즘 개발**	▪ 이윤발생 (Profitable) 이중 지불 공격을 방지하는 동시에 빠른 거래완료 속도를 달성하기 위한 조건 연구 ▪ 트랜잭션의 용도별 검증 횟수 최적화 알고리즘 개발	
유연성	**플라즈마 모델 개발**	▪ 이더리움의 레이어2 솔루션인 플라즈마 모델을 개발하여 DeSecure 블록체인에 적용 ▪ 부적절 거래 방지를 위한 플라즈마 데이터 가용성 개선 알고리즘 개발	참여 기관

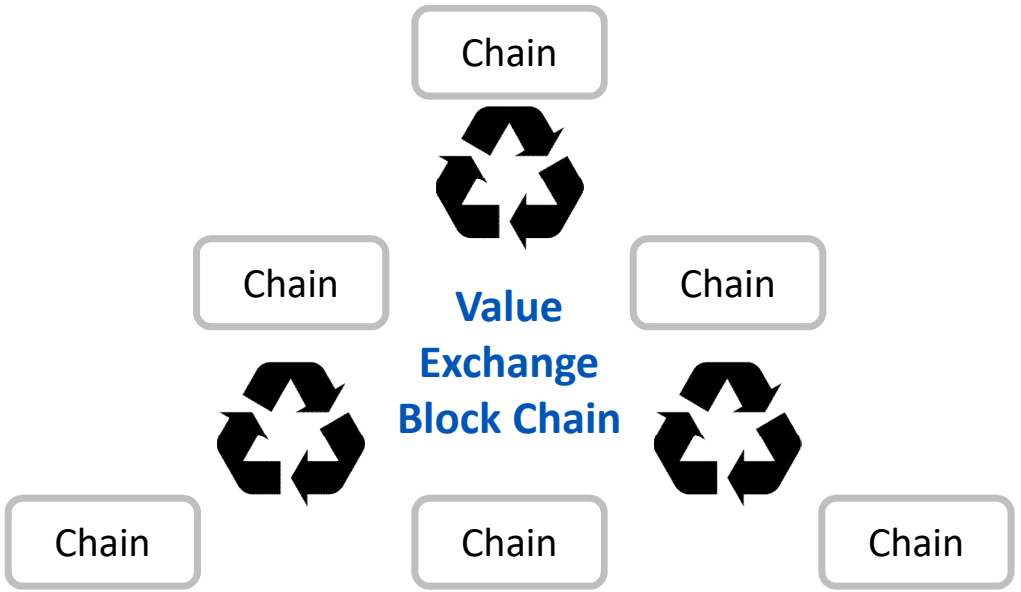
 ⓒ 2019 GIST

확장성 강화 연구

- Evolutionary Space에서 DeSecure 블록체인을 용도에 따라 특화
- 용도별 특화 블록체인들을 연결하여 다중 복합구조 DeSecure 블록체인 형성

Low TPS
Global Nets
High Delay

Large Tx (Assets)



High TPS
Local Nets
Low Delay

Small Tx
(Utility Coins)

<용도에 따라 세분화된 다중 복합구조 블록체인>

ⓒ 2019 GIST

확장성 강화 연구

- 사용 용도에 따라 블록전파지연시간과 참여 노드 수가 다르므로 용도에 따른 파라미터 최적화를 연구함
- **Enterprise DeSecure Blockchain**은,
  Private Consortium으로 구성되어 기업간의 정보를 공유하는데 사용될 수 있음
  - Enterprise와 지역규모 DeSecure Blockchain들은 빠른 거래 처리속도 (TPS)에 집중하여 채굴난이도를 낮은 수준으로 설정함
  - 동시에 이중 지불 공격 방지 및 이상 행위 퇴출을 위해 멤버십 제도와 같은 Off-Chain Policy를 적용함
- **지역규모 DeSecure Blockchain**은 시/구/군 규모의 블록체인으로 작은 규모의 거래가 주를 이룸
- **국가규모와 국제규모 DeSecure Blockchain**은,
  각각 지역 간, 그리고 국가 간의 거래를 기록하며, 거래의 규모가 큼
  - 국가규모와 국제규모 DeSecure 블록체인들은 거래의 규모가 큰 만큼 거래 처리속도 보다 보안성 제고를 위해 채굴난이도를 높은 수준으로 설정함

Ⓒ 2019 GIST

확장성 강화 연구

피라미터/용도	Enterprise / Consortium	지역 규모	국가규모	국제규모
네트워크 전파시간	<1초	<10ms	<100ms	<1초
블록생성 시간	<0.1초	1분	5분	10분
채굴난이도	최소	하	중	상
Off-chain 정책	O	O	X	X
참여 노드 수	허가된 소수 (10개 내외)	<100개	<1000개	<10000개
보안성	낮음	낮음	중간	높음
TPS	>42000	<70	<14	<7

※ TPS 계산 기준 : 1MB 블록 사이즈, 0.24KB 트랜잭션 사이즈

Ⓒ 2019 GIST

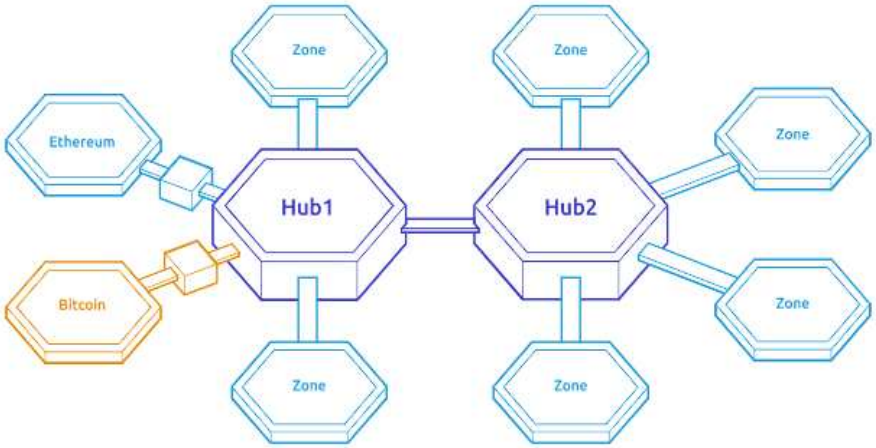
확장성 강화 연구

- **DeSecure 다중 복합구조 블록체인은 용도에 따른 최적화만으로도, PoS/DPoS/BFT 등보다 유연하게 작동할 수 있음**
  - PoS/DPoS/BFT 등의 합의방식은 노드의 Generic ID가 완벽히 신뢰할 수 있거나, 네트워크 규모가 아주 작을 때 정상적으로 작동함
  - 노드의 Generic ID가 완벽하지 않다면 소수의 노드가 많은 ID를 보유하여 네트워크에 큰 영향력을 행사하는 Sybil Attack이 가능함
  - Generic ID기술 중 하나인 PUF의 경우, 하나의 전자장비로 여러 개의 ID를 만들어 낼 수 있음
  - 또 다른 기술인 SSID의 경우 중앙 시스템 (정부 혹은 플랫폼 제공 기업)에 의해 통제되기 때문에, 안전하지 않음
- 반면 ECCPoW를 사용하는 **DeSecure 블록체인**은 Generic ID에 의존하지 않기 때문에 **상기 보안 문제로부터 자유로움**
- 이뿐만 아니라, **DeSecure 블록체인**을 다중 복합구조로 설계함으로써 PoS/DPoS/BFT가 가지는 장점인 **빠른 거래처리 속도까지 확보할 수 있음**

 Ⓒ 2019 GIST

확장성 강화 연구

- **다중 복합구조 DeSecure 블록체인 간의 통신 프로토콜 연구**
  - 용도별 DeSecure 블록체인간의 가치 및 정보교환을 위한 통신 프로토콜을 연구함
  - 다중 복합구조 DeSecure 블록체인들은 모두 같은 합의 알고리즘을 탑재하고 있으므로 블록체인 간 통신이 용이함
  - Sidechain의 개념으로 다중 구조 블록체인을 제안한 Polkadot과 Cosmos와 같이 통신의 Hub Station 역할을 해주는 블록체인을 도입하는 방법도 적용 가능함



<Polkadot의 블록체인 간의 연결 구상도>

ⓒ 2019 GIST

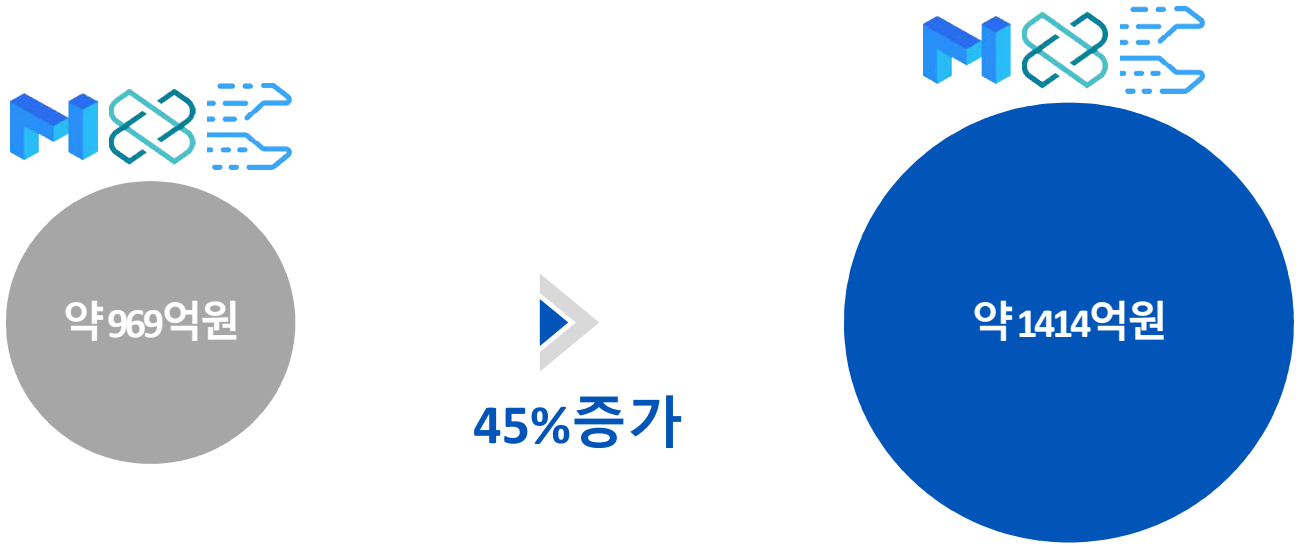
ECCPoW와 다른 확장성 방법과 비교

- DeSecure Blockchain aims to resolve the re-centralization problem without sacrificing the secureness!

체인 이름	DeSecure	Bitcoin		Ethereum	
방법	다층멀티체인 ECCPoW	세그윗	라이트닝 네트워크	플라즈마	샤딩
구현	ECCPoW 기반 독립체인들을 여러 계층으로 묶음	블록 데이터 구조를 변경하여 구현	오프체인 거래 진행, 최종 결과값을 메인 블록체인 기록	하부 체인 생성, 거래 진행 후 최소한의 기록만 메인 블록체인 기록	블록체인의 DB에 해당하는 스테이트를 여러 샤드로 분할, 트랜잭션 별 분리 처리
장점	서로 다른 블록체인 연결해 다양한 기능과 역할 구현	쉽게 구현이 가능함	결제 속도 제고, 즉각적인 완결성, 수수료 절감	수수료 절감	트랜잭션 처리 속도 증가
단점	No single chain solution/생태계필요	트랜잭션 처리속도 증가 효과 미비	오프체인 거래기록 없음	Full노드 만 플라즈마 사용 가능	S/W 복잡도 상승

Ⓒ 2019 GIST

주요 Layer2 프로젝트 규모 증가



- 주요 layer2 프로젝트의 시가총액의 합은 2019년초 약 969억원 대비 2019년 12월 현재 약 1414억원으로 약 45% 성장
- 다른 프로젝트의 시가 총액이 감소, layer2 솔루션의 성장으로 기대와 전망을 확인할 수 있음

Ⓒ 2019 GIST

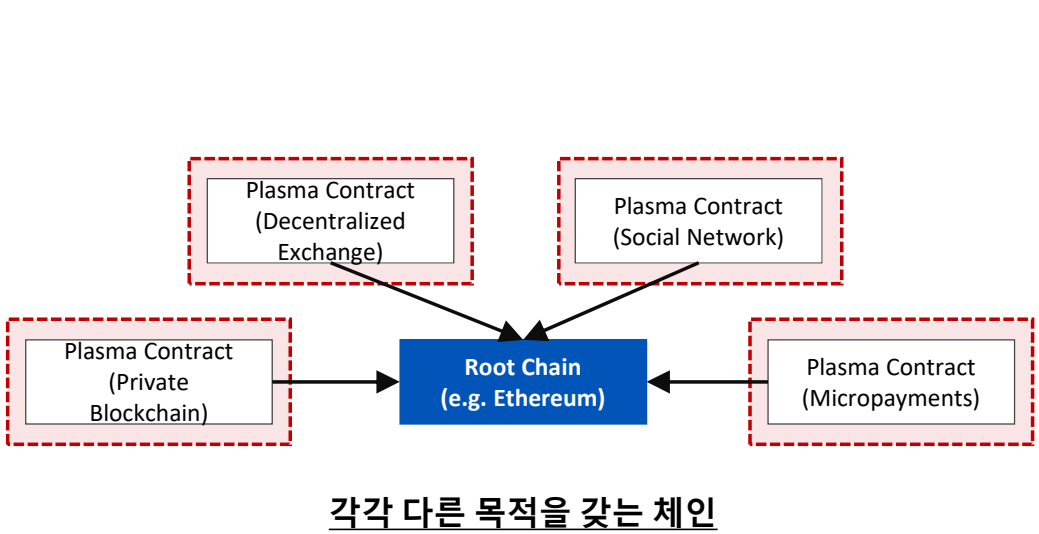
유연성 강화를 위한 플라즈마 도입

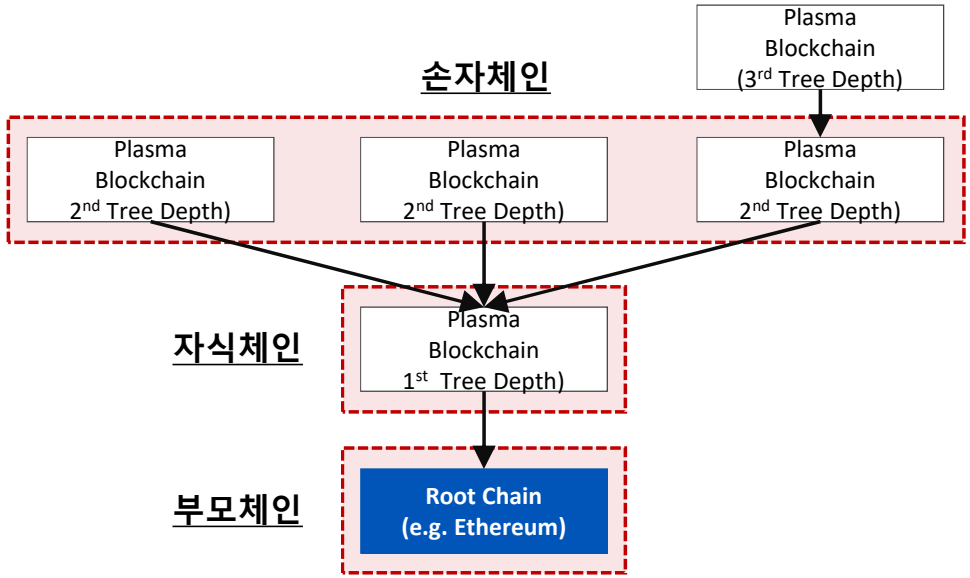
- **플라즈마**
  - 스마트 계약(이하 플라즈마 컨트랙)의 인센티브 및 강제 실행(incentiveized and enforced execution)을 위한 프레임워크로 초당 상당한 양의 상태 업데이트(잠재적으로 수십억 개)로 확장 가능하여, 블록체인(별도의 플라즈마 체인)이 전세계적으로 상당한 양의 분산형 금융 애플리케이션 구동을 가능하게함

  - 스마트 계약은 네트워크 거래 수수료를 통해 자율적으로(autonomously) 운영을 지속하도록 유도되며, 이는 궁극적으로 트랜잭션 상태 전환을 강제(enforce)하기 위해 메인 체인(예: Ethereum)에 의존함

 ⓒ 2019 GIST

유연성 확보를 위한 플라즈마 도입





- 플라즈마를 도입 했을때 DeSecure체인을 부모체인으로 하는 다수의 체인을 생성할 수 있음
- 각 체인은 **목적에 맞는 스펙**을 가질 수 있으며 필요에 따라 언제든지 체인을 열고 닫을 수 있음
- 또한, 자식체인이 손자/손녀 체인들의 부모 체인이 되는 **계층적 구조로 확장**할 수 있기 때문에 이론상으로는 무한대의 확장성 또한 확보 가능

Ⓒ 2019 GIST

플라즈마 체인의 보안성 강화 연구

플라즈마 체인의 보안성 문제 정의

- 자식체인의 상태 강제를 위해서 진실에 공판에 올라갈 증거가 필요함
- 플라즈마 체인의 운영자(Operator)가 블록의 내용을 내어주지 않을 가능성이 존재
- 증거(Proof)가 없으면 앞에서 언급된 어떠한 검증장치들도 작동하지 못함
- 이러한 문제를 데이터 가용성(Data Availability, 이하 DA) 문제라고 함

더 큰 문제는 이 문제가 상당히 주관적(Subjective) 특징을 지님

- 데이터를 받았다는 것을 증명하는 것은 쉬우나, 데이터의 내용을 모르는 상태에서 데이터를 받지 않았다는 것을 증명하는 것은 **상당히 어려움**
- 받았음에도, 받지 않았다고 주장할 수도 있기 때문에 데이터 인질이 있는지 없는지 여부에 대한 판단은 매우 **주관적**이 될 수밖에 없음

Ⓒ 2019 GIST

플라즈마 체인의 보안성 강화 연구

플라즈마 체인의 보안성에 대해 보유한 연구 성과

[Solevm] – 스마트 컨트랙트 EVM

- 연산 챌린지를 위해서는 EVM의 런타임에 대한 검증게임이 이뤄져야 함
- 그래서, EVM자체를 스마트 컨트랙트로 구현하는 것이 매우 중요함
- 이는 마치 파이썬으로 만든 파이썬, 자바로 만든 자바와 같음
  → 다만 커스텀한 체인의 실행모델이라면, 그 환경에 맞춰진 변형된 EVM 컨트랙트 코딩이 필요함

[Continuous Rebase] – 플라즈마 체인 적용 핵심기술

- Continuous Rebase 모델은 플라즈마 체인의 정상적인 작동과정에 Rebase를 포함시킴.따라서 지속적이고 주기적인 Rebase를 통해 사용자들의 Escape Request를 반영
  (Rebase : 기존 블록을 다른 블록 기준으로 마이닝)
- 이를 통해 사용자들은 DA문제가 있을 경우 Escape Request를 제출하여 안전하게 탈출 가능. 또한 제출된 블록이 올바르지 않을 경우 연산 챌린지를 통해 해당 블록들이 Finalize 되는 것을 막을 수 있음

ⓒ 2019 GIST

보안성 강화 연구

보안성 문제 정의

- 거래검증 횟수가 많을수록 거래처리속도는 느려지나, 이중 지불(DS) 공격에 강인해짐
- 지금까지는 거래검증 횟수의 설정 효과가 수치화 되어 있지 않음

※ 거래검증 횟수 : Block confirmation number



보안성 분석 보유 연구성과

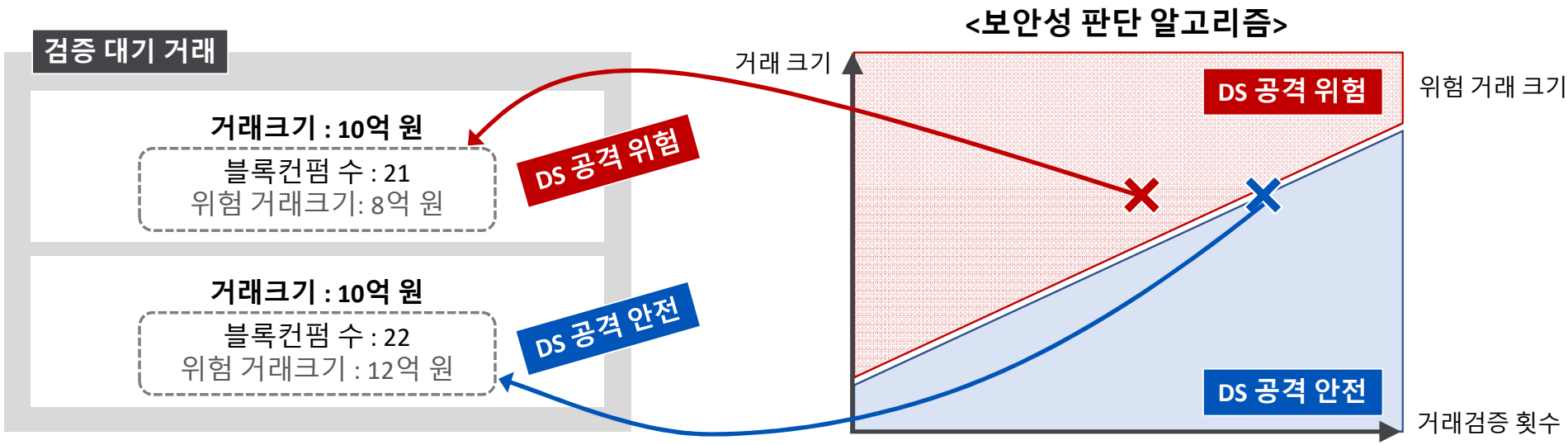
- Satoshi Nakamoto의 백서를 포함한 기존 연구는 대규모 (50% 이상)컴퓨팅파워 확보에 의한 DS 공격의 위험성에 초점을 맞추었음
- 소규모 (50%미만) 컴퓨팅 파워 확보로도 위협적인 DS 공격이 가능함을 수학적으로 밝힘
- **DS공격의 수익(Profit)과 공격목표 거래규모, 거래검증 횟수와의 상관관계를 제시함**
- 연구 성과를 논문으로 작성하여 국제 학술지에 제출

Ⓒ 2019 GIST

보안성 강화 연구

보유 연구성과 활용 방안

- 거래 별로 DS공격의 위험성을 평가하여 사용자가 보안위험 자각 및 대처 가능
- 거래검증 횟수와 거래크기가 DS 공격자에게 이윤을 발생시키는지를 거래 별로 평가 가능



ⓒ 2019 GIST

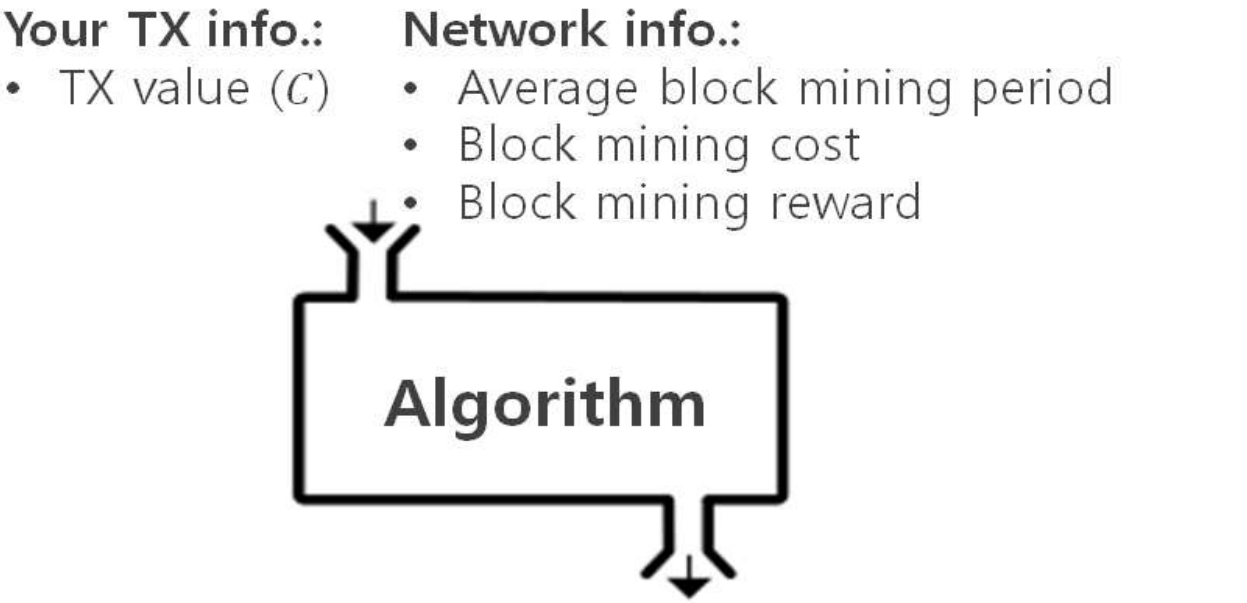
보안성 강화 연구

거래 검증횟수 최적화 알고리즘 개발

- **거래 용도에 따라 최적의 검증 횟수를 결정해주는 알고리즘을 개발함**
- 최적의 검증횟수란 DS 공격을 방지하는 동시에 거래 완료에 걸리는 시간을 최소화하는 검증 횟수임
  - 거래금액에 따라 공격자의 이윤 발생을 방지하는 검증 횟수가 존재함
- **최대 거래 속도(확장성)가 거래의 용도에 따라 유동적으로 결정되도록 함**
  - 소액 거래일 때는 공격자가 DS 공격을 통한 이윤을 기대하기 어려우므로 보안성보다 확장성을 위한 작은 검증횟수가 적절함
  - 고액 거래일 때는 이중 지불 공격의 목표가 될 가능성이 있으므로 큰 검증횟수를 설정해 보안성을 유지함
- 거래 검증 횟수 최적화 알고리즘을 지갑에 탑재하여 거래를 생성할 때마다 자동 실행되도록 함

Ⓒ 2019 GIST

보안성 강화 연구



<검증횟수 최적화 알고리즘>

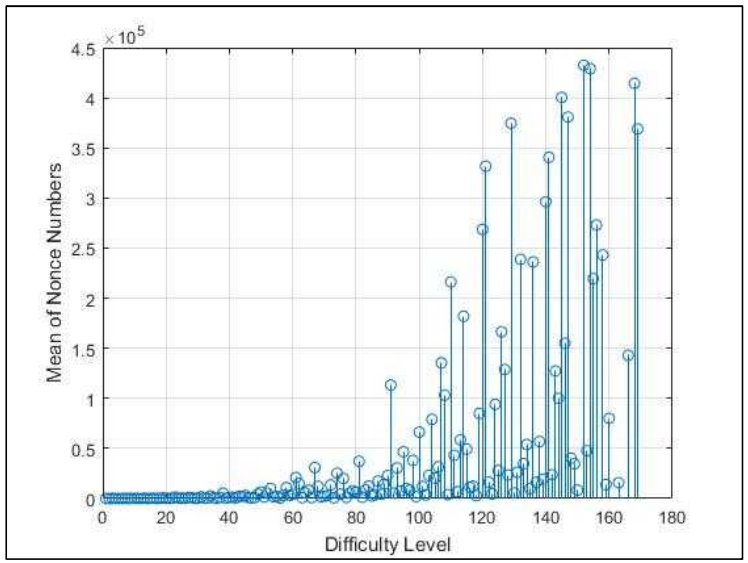
 Ⓒ 2019 GIST

ECCPoW 코어 2.0 개발

1단계에서 확보한 연구 성과

Level	Length	wc (the number of ones	wr (the number of ones	from	to	type	Prob. Calculated	Average Nonce Measured
1	32	3	4	10	22	2	3.08E-05	20.46
2	32	3	4	10	16	2	2.02E-05	30.51
3	32	3	4	16	16	1	9.68E-06	66.24
4	32	3	4	14	14	1	6.78E-06	97.22
5	36	3	4	12	24	2	4.83E-06	29.40
6	36	3	4	12	18	2	3.13E-06	44.19
7	32	3	4	12	12	1	2.86E-06	222.73
8	44	3	4	14	30	2	1.64E-06	62.10
9	36	3	4	18	18	1	1.42E-06	103.60
10	36	3	4	16	16	1	1.05E-06	141.06

<채굴 난이도별 설정방법>



<채굴 난이도별 블록생성 시도 횟수 측정 값>

 Ⓒ 2019 GIST

광주과학기술원 Gwangju Institute of Science and Technology

ECCPoW 코어 2.0 개발 – 연구문제의 도전성 및 중요성

블록생성 확률 계산의 도전성

- 1단계 연구에서 블록생성 확률의 상한과 하한 값을 찾음
- 기존 작업증명 알고리즘은 그 원리가 단순하여 블록생성 확률 계산이 쉬움
- Anti-ASIC 특성을 확보하기 위해 ECC라는 고도화된 알고리즘을 활용하므로 확률 계산이 도전적임

블록생성 주기 제어 고도화의 중요성

- 네트워크 총 계산 량의 변동에 따른 블록생성 주기의 안정화 시간 (settling time)을 더욱 단축
- 급격한 계산 량 변동을 유도하여 부당 이득을 취하려는 공격 시도를 억제 할 수 있음

Ⓒ 2019 GIST

ECCPoW 코어 2.0 개발 – 연구문제의 해결 전략

블록생성 확률 계산

- ECC에서 사용되는 LDPC 행렬의 특성과 확률 분포에 관한 연구 문헌을 조사하여 정확한 확률을 계산 가능
- 많은 양의 실험을 통해 난이도별 블록생성 통계를 측정하여 활용 가능

계산 량의 변동에 따른 난이도 설정 규칙 정의

- 비트코인과 이더리움은 각각 자체 정의한 난이도 설정 규칙을 적용함
- 두 규칙은 오랜 기간 사용되어 왔으며 그 안정성이 실증되었음
- 블록생성 확률을 정확하게 계산 할 수 있다면, 기존의 두 규칙을 활용 가능함

$$D_{next} = D_{prev} \times \frac{\text{2 weeks (20160 minutes)}}{\text{Time spent for mining 2016 blocks}}$$

<비트코인의 난이도 설정 규칙>

$$D(H) \equiv \begin{cases} D_0 & \text{if} \quad H_i = 0 \\ \max\left(D_0, P(H)_{H_d} + x \times \varsigma_2 + \epsilon\right) & \text{otherwise} \end{cases}$$

<이더리움의 난이도 설정 규칙>

 Ⓒ 2019 GIST

공인 인증 평가 목적

세계 최고 수준의 블록체인인 비트코인, 이더리움과 비교 평가함
동일한 평가 환경(난이도 변경 주기, 목표 블록 생성시간,
23개의 Instance 등)에서 비교

[평가의 객관성]
평가 데이터와 블록체인 소스코드를 Github에 공개하여 평가 결과는
누구나 검증 가능함

Bitcoin 평가 데이터 : https://github.com/paaabx3/BTCBGT3-2
Bitcoin 평가 데이터 : https://github.com/paaabx3/BTCeccBGT3
ECCPoW 블록체인 탐색기 : http://13.125.223.181/blocks
ECCPoW 블록체인 탐색기 : http://15.164.244.139/blocks

Ⓒ 2019 GIST

1차년도 공인 인증 평가 요소

01

분산성

각 채굴 노드들이 전체 채굴량에 대해 일정하게 채굴하는 정도를 의미함.
분산성 수치가 100%에 가까울 수록 각 채굴자들의 채굴 확률이 균등하게 분포 된 것을 뜻함.

02

보안성

이중 지불 공격의 원인인 고아 블록의 출현을 방지하고 이를 보안을 유지할 수 있는 것을 의미함
즉, 고아 블록 비율이 낮으면 보안성이 좋다고 할 수 있음

03

확장성

블록체인 환경에서의 처리할 수 있는 트랜잭션 (TPS)을 증가시킬 수 있음을 의미함.
동일한 환경에서 비트코인/이더리움과 ECCPoW의 TPS를 비교는 확장성을 비교할 수 있음

Ⓒ 2019 GIST

공인 인증 평가 결과

- 국제 공인인증 기관인 "한국시험인증원"으로부터 기술력을 인정 받음

평가 항목 (주요성능 Spec)	단위 (%)	전체 항목에서 차지하는 비중 (%)	세계최고 수준 보유기업 (비트코인, 이더리움)	연구개발 전 국내수준	목표		실적	평가방법
			성능수준	성능수준	최종목표	당해년도	당해년도	
① 분산성	채굴성공률 분포도	35%	40% (추정 값) (기준 : 비트코인, 2018년 10~12월)	40%	80%	60%	**92.22%**	공인인증기관
② 보안성	비트코인 대비 보안성	35%	100% (기준: 비트코인)	100%	200%	100%	**100%**	공인인증기관
③ 확장성	비트코인 대비 확장성	15%	100% (기준: 비트코인)	100%	200%	100%	**98.95%**	공인인증기관
④ 확장성	이더리움 대비 확장성	15%	100% (기준: 이더리움)	100%	200%	100%	**102.97%**	공인인증기관

① 분산성 : 당해 년도 목표 대비 **+32.22%**
② 보안성 : 당해 년도 목표 대비 **0%**
③ 확장성(BTCECC) : 당해 년도 목표 대비 **-1.05%**
④ 확장성(ETHECC) : 당해 년도 목표 대비 **+2.97%**

© 2019 GIST

2차년도 공인 인증 평가 요소

01

분산성

각 채굴 노드들이 전체 채굴량에 대해 일정하게 채굴하는 정도를 의미함.

분산성 수치가 100%에 가까울 수록 각 채굴자들의 채굴 확률이 균등하게 분포 된 것을 뜻함.

02

보안성

동일한 환경의 비트코인과 DeSecure 블록체인에 이중지불 공격 모의 실험을 수행하여 피해 금액을 측정함.

피해 금액이 적을수록 보안성이 우수함.

두 블록체인의 피해 금액을 비교하여 보안성을 상대 평가함.

03

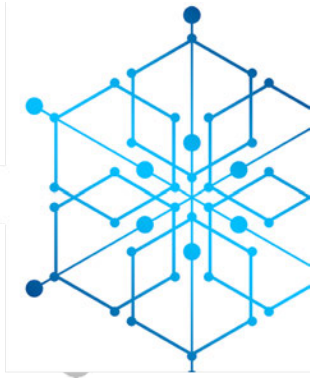
확장성

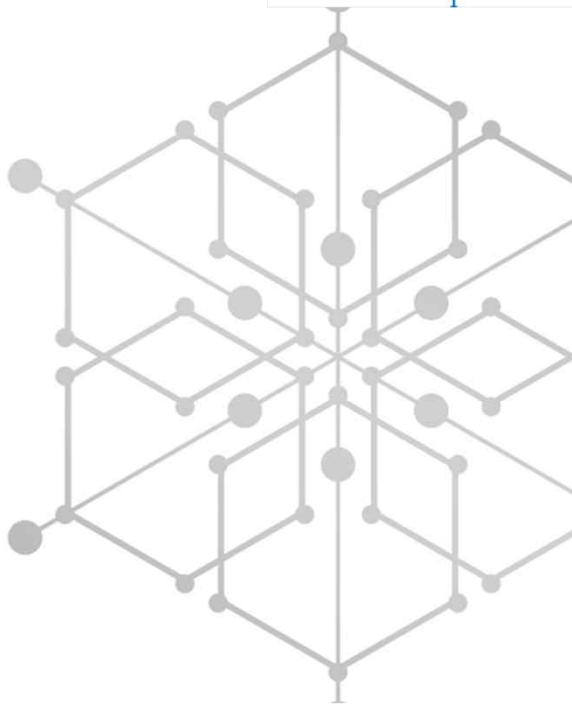
블록체인 환경에서의 처리할 수 있는 트랜잭션(TPS)을 증가시킬 수 있음을 의미함.

동일한 환경에서 비트코인/이더리움과 ECCPoW의 TPS를 비교는 확장성을 비교할 수 있음

Ⓒ 2019 GIST

감사합니다





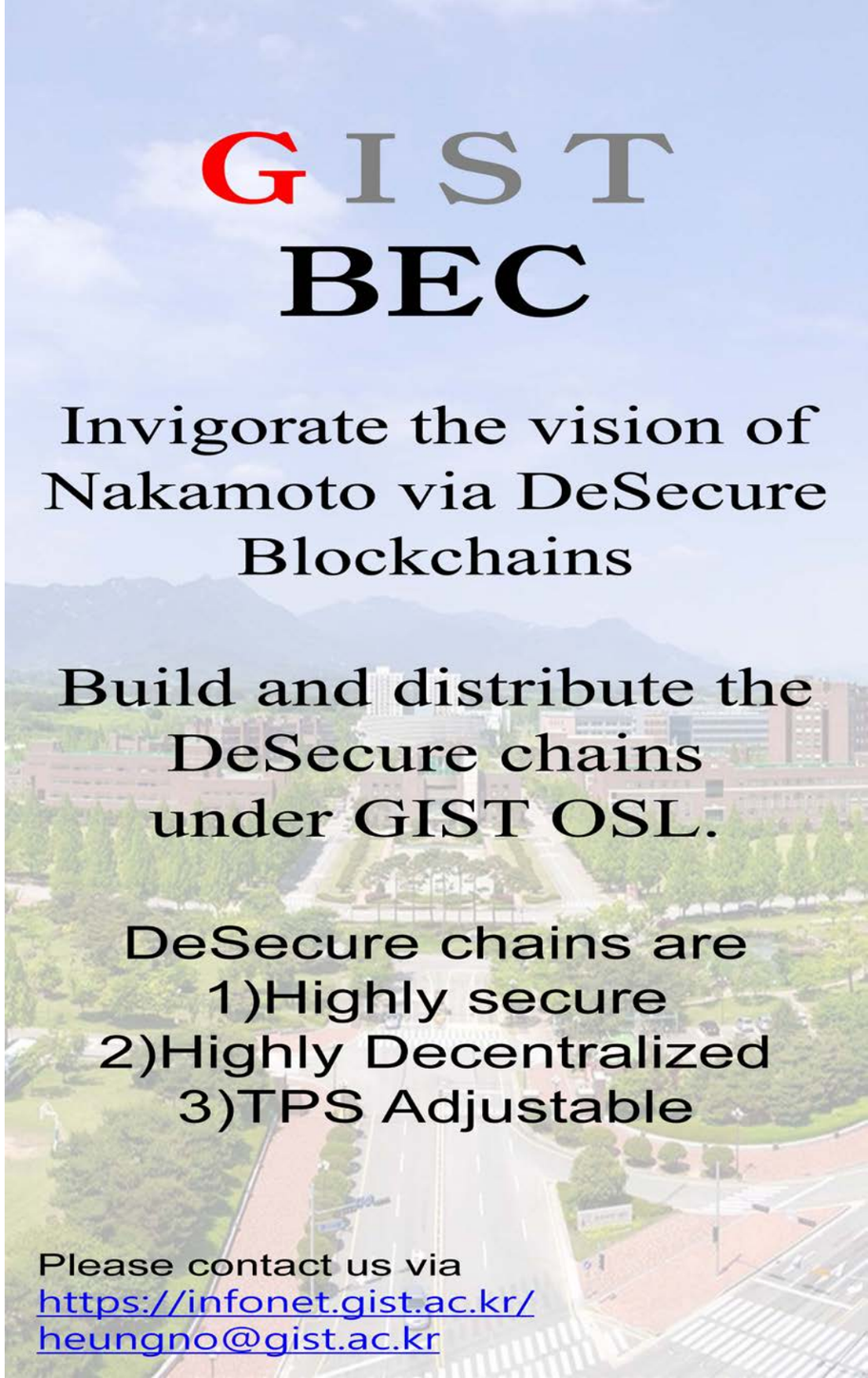
Scalable DeSecure ECCPoW Blockchains



Heung-No Lee, GIST, South Korea
Home page: http://infonet.gist.ac.kr
Facebook/ Publication ID: Heung-No Lee
E-mail: heungno@gist.ac.kr

Abstract

GIST Blockchain-Economy Center (BEC, Director Heung-No Lee) aims to introduce the Decentralized Secure(DeSecure) blockchains it has been developing since 2018. They aim to resolve the re-centralization problem of today's mining market. One of the key ideas is to have the proof-of-work (PoW) puzzle time-varying from block-to-block, using the error-correction-codes (ECC). Two new blockchains based on Bitcoin and Ethereum are to be developed using new consensus algorithm based on this new ECC-PoW. Time-varying puzzles make it very difficult to develop an ASIC mining chips. As the result, with the size of network growing, the difficulty level needs not be growing as well. As such, energy spent for mining can be controlled. The proposed ECC-PoW mechanism is to be explained in details. In addition, our plan to hardfork Bitcoin and Ethereum, by replacing the SHA based PoW with the proposed ECC-PoW, and by developing two new DeSecure blockchains, i.e. BTC-ECC and ETH-ECC, is discussed. The two DeSecure blockchains will be openly shared under an open source license at Github. We address how DeSecure blockchains can be used to resolving the issue of scalability. Our schedule to release the cores (C++ and Go) and technical meet-ups will be addressed.



Short Bio of Dr. Heung-No Lee



Heung-No Lee graduated from University of California, Los Angeles (UCLA), U.S.A. with Ph.D., M.S., and B.S. degrees all in Electrical Engineering, 1999, 1994 and 1993 respectively. He has written more than 270 journal and conference publications. In the past, he worked at HRL Laboratory, Malibu, California, U.S.A., as Research Staff Member and as Assistant Professor at the University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A. He is currently a full tenured professor at Gwangju Institute of Science and Technology (GIST), Republic of Korea.

His research lies in the areas of Information Theory, Signal Processing Theory and their application to Communications and Networking systems, Biomedical systems, and Signal Processing systems.

Awards he has received recently include Top 50 R&D Achievements of Fundamental Research in 2013 (National Research Foundation), Top 100 National R&D Research Award in 2012 (the Ministry of Science, ICT and Future Planning) and This Month Scientist/Engineer Award (National Research Foundation) in January 2014.

He was the Director of Electrical Engineering and Computer Science within GIST College in 2014. Administrative positions he has held at GIST include the Dean of Research and the Director of GIST Research Institute.

Talk today

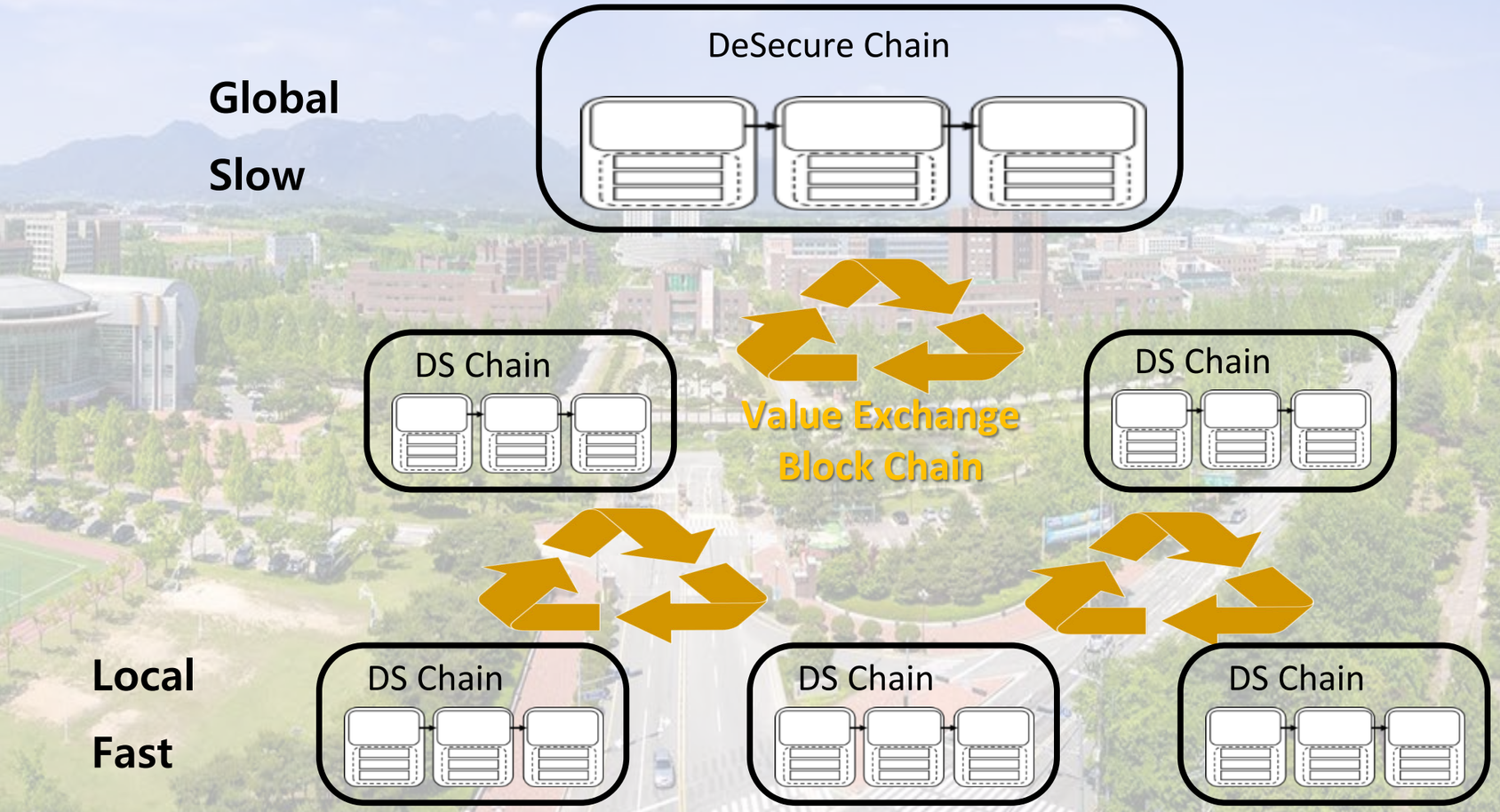
- DeSecure Blockchains
- Error Correction Codes based PoW
- Safe Transactions enabled by novel DS analysis
- Release Plan

Gwangju Institute of Science and Technology
Blockchain Economy Center

DeSecure chains are
1) **Highly secure**
2) **Highly Decentralized**
3) **TPS Adjustable**

The aim is to build and distribute the DeSecure chains under GIST OSL.

Please contact us via
https://infonet.gist.ac.kr/
heungno@gist.ac.kr

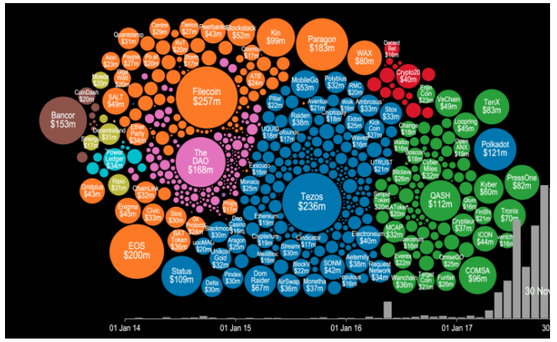


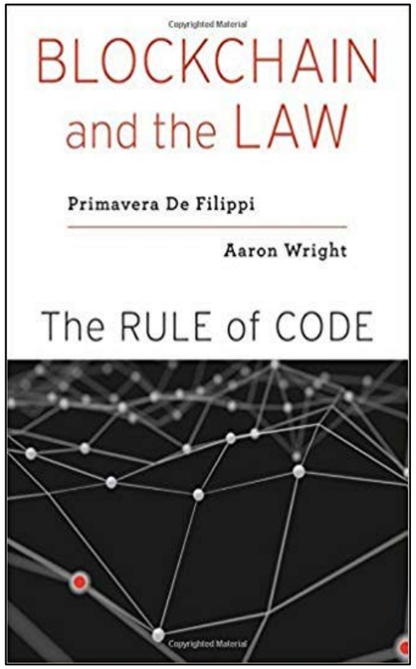
Bitcoin’s Ideals

- Since birth in 2009, Bitcoin has never stopped breathing and alive currency system.
- It is a global digital currency which works beyond national boundaries.
- It was the time when trust on the banks and governments were severely degraded.
- Ideals around bitcoin are
  - Decentralization
  - Reforming Wall street
  - Unbundling big corporations
  - Reduction of inequality

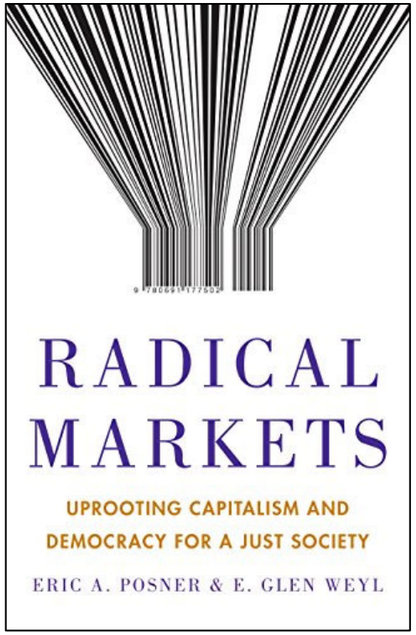
Ethereum's Ideals







- Ethereum network allows not only coin TXs, but also doc files and computer codes.
- A *decentralized* app (Dapp) runs a front end code; a backend code runs in *the Eth Net*.
  - ✓ cf) For an ordinary app, the backend code is running on *a centralized server*.
- Smart contracts
  - ✓ A computer code can be executed and advanced to the next stage each time a contractual term matures.
- Decentralized autonomous organization has its bylaw written in smart contracts.
  - ✓ The organization spends tokens and makes governance decisions w.r.t. smart contracts.
- *Lex Cryptographia!*
- *Uprooting capitalism and democracy for a just society!*
- *Sharing Economy!*



Novel DeSecure Blockchains

- BTC and ETH are great BUT they are
  - Re-Centralized
  - Scalability Issue
  - Said to be too slow and small

- We aim to approach these two issues with DeSecure blockchains
  - Anti-ASIC ECC PoW
  - Ecosystem of DeSecure blockchains

- DeSecure blockchains uses novel Error-Correction Code PoW.

- We aim to provide two DeSecure blockchains, ETH-ECC and BTC-ECC.

They Have Sought Alternatives to SHA-PoW, BUT

	Pros	Cons	Coins within top 50 rank
PoW (Proof-of-Work)	• Strong security - Difficult to produce - Easy to verify	• Extreme computing power • 51% attacks • Transaction speed / Transaction throughput	Bitcoin, Ethereum
PoS (Proof-of-Stake)	• Energy & hardware efficiency • Much more expensive 51% attacks	• Recentralization • The rich-get-richer • “Noting at stake” problem	Qtum, Stratis
DPoS (Delegated PoS)	• Scalability and speed • Energy & hardware efficiency • Encouraging good behavior by real-time voting	• Recentralization • DDoS attacks	EOS, NEO smart economy
PoA (Proof-of-Activity)	• Much more expensive 51% attacks • Decentralization - Validators are randomly selected.	• Recentralization • Extreme computing power • The rich-get-richer	decred

Comparison to Existing Scalability Solutions

DeSecure Blockchain aims to resolve the re-centralization problem without sacrificing the securedness and decentralization!

Type	DeSecure	Bitcoin		Ethereum	
Name	Multi-level, multiple chains	Seg-Wit	Lightening Network	Plasma	Sharding
How	Many ECCPoW based chains can talk to each other via value-exchange service	Realize by modifying a block data structure	Allow off-chain transactions and record the end result of these transactions into the main blockchain	Allow transactions in child chains, TX records end up at the main chain are limited.	Divide BC DB with multiple shards
Pro	Many different services and levels of chains can co-work.	Easy to realize	Faster transactions Small TX fees	Faster transactions Small TX fees	Faster transaction
Con	No single chain solution Requires an ecosystem	Small improvement	The content of off-chain transactions lost	Some TX content lost Only full node can run this	Increased SW complexity

We aim to Replacing SHA-PoW with ECC-PoW!

Blockchain Core Program

Three key parts

1. Web server interface networking of peers

- Node registration, get-address, give-address
- Full node or light node
- Communication among the wallets and the miners

2. Wallet for TX generations

- Make private and public keys, address, store UTXOs, make TX, put signature, announce it to the neighbor, check to see if the TX is supported by the blockchain.

3. Consensus Mechanism

Consensus Engine

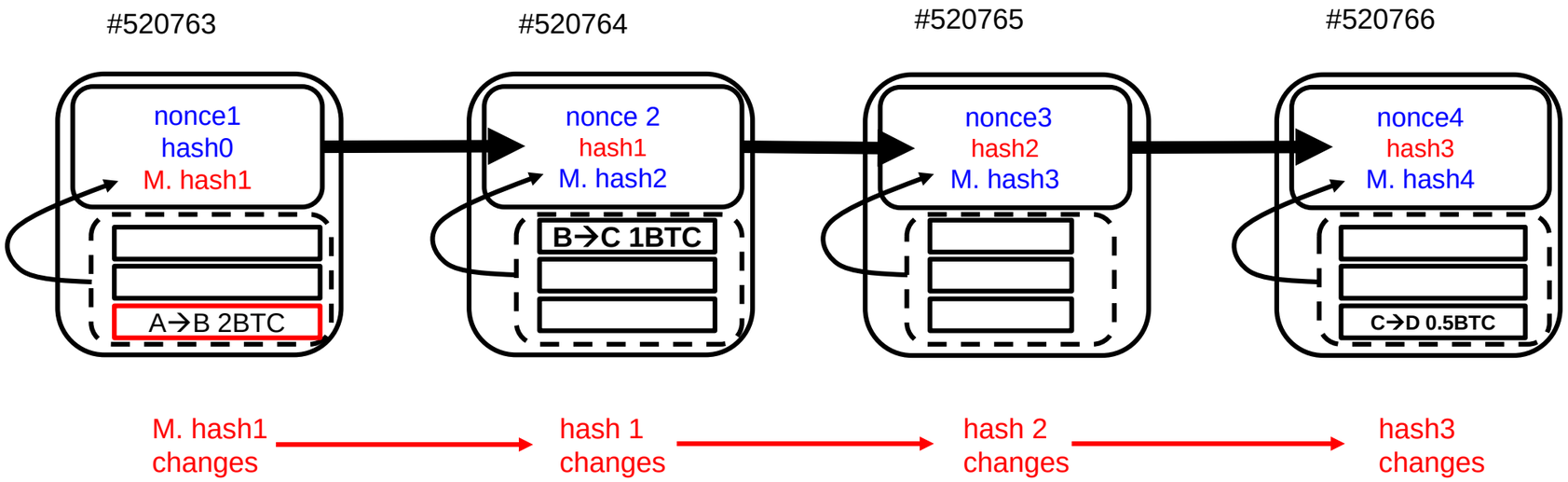
- Data: Genesis block + regular blocks, one block every 10 min, block-size 1Mbyte
- Protocol: consensus, block header, difficulty level adjustment, …
- Mining: Get the longest chain, validate it and all transactions within it, get transactions from mempool and form a block, run SHA repeatedly until you hit a good hash, put the proof into the block header, and attach the proofed block to the longest chain, and make announcement ASAP.

Program Suite

- **C++, Python, Go, Java, Flask, http**
- **Download and run, then you have a blockchain server.**

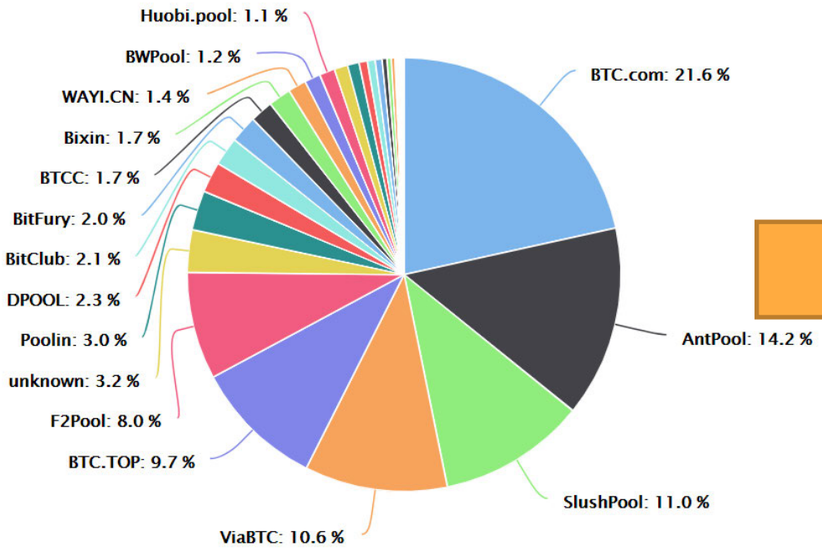
Pow is fundamental to OPEN blockchains!

- What happens when any alteration is made?
- Proof-of-Work (PoW)
- Immutability and openness allow transactions.
  - A → B 2 BTC
  - B → C 1BTC
  - C → D .5 BTC

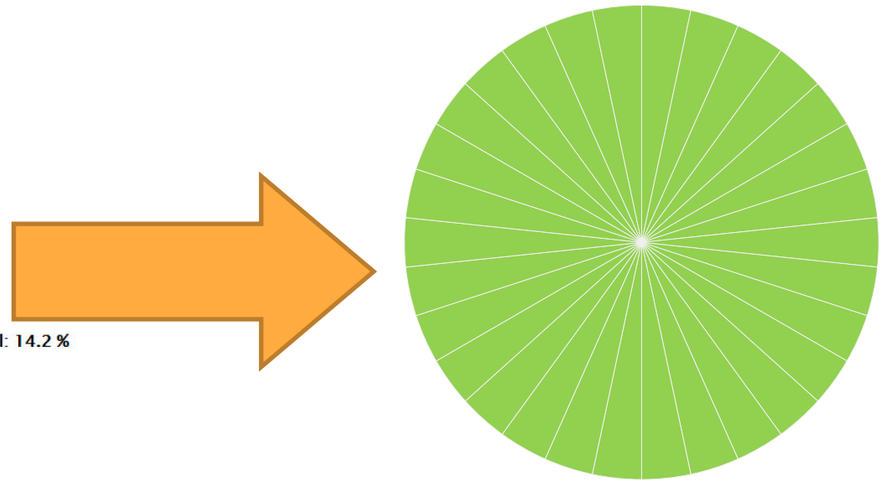


ECC-PoW aims to resolve Recentralization Issue.

- ASIC → Mining Moguls → Discourage Average Miners
- Prone to Collusion, Censorship



Recentralized



Decentralized again

1. ASIC resistant
2. Vulnerability to DS attacks reduced

There are items to consider for a new PoW!

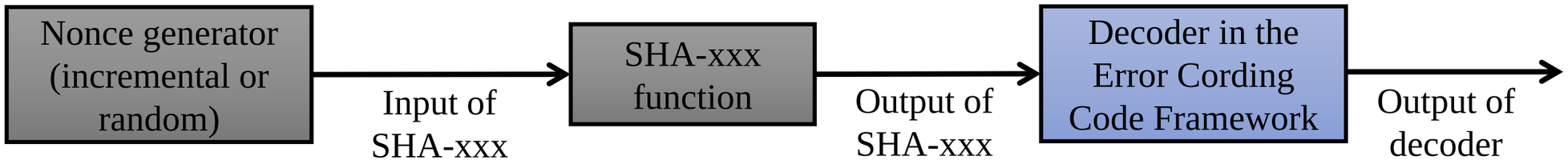
- A new puzzle generation system is capable of varying puzzles from block to block with the following properties:

  P1: Easy to verify but difficult to prove
  P2: Robust to detect block modification attacks
  P3: Controllable in changing the difficulty level
  P4: Open to anyone with a CPU
  P5: Unfixed and changeable from block to block

- The re-centralized problem can be resolved thanks to P5.

Novel Error Correction Codes PoW (ECCPoW)

- There are many one-way functions in Inverse Problems such as Error Correction Codes, Sparse-Signal Recovery, Space-Time Coding, Sphere-Decoding, Digital Communications Receiver algorithms.
- In these problems, encoding is easy but decoding is time-consuming!
- We combine a Error Correcting Code framework with SHA-xxx.

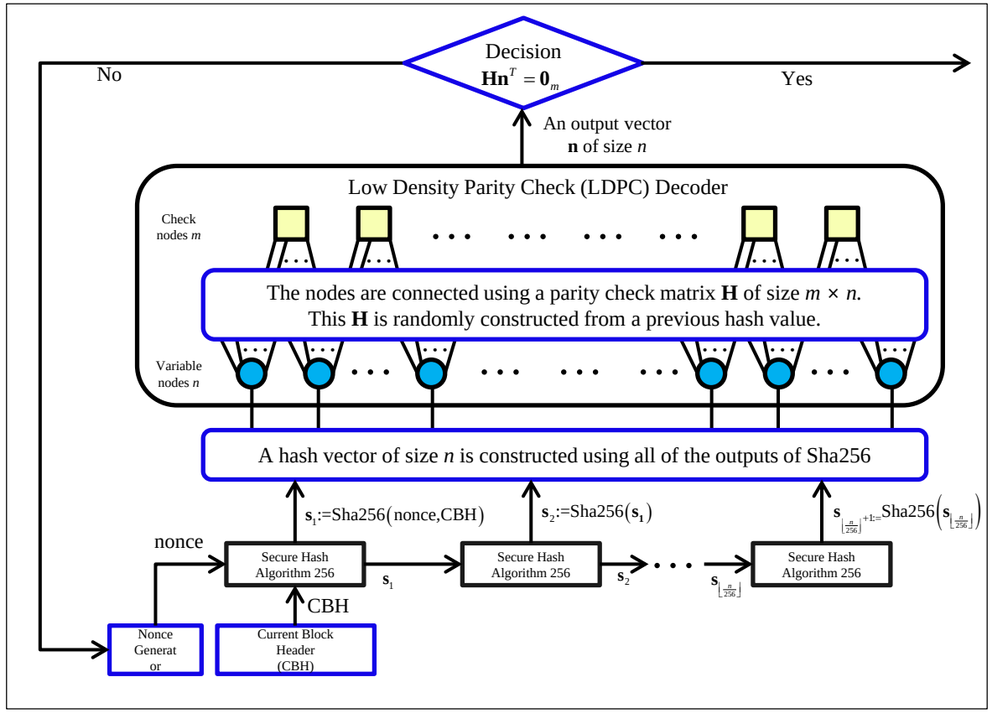


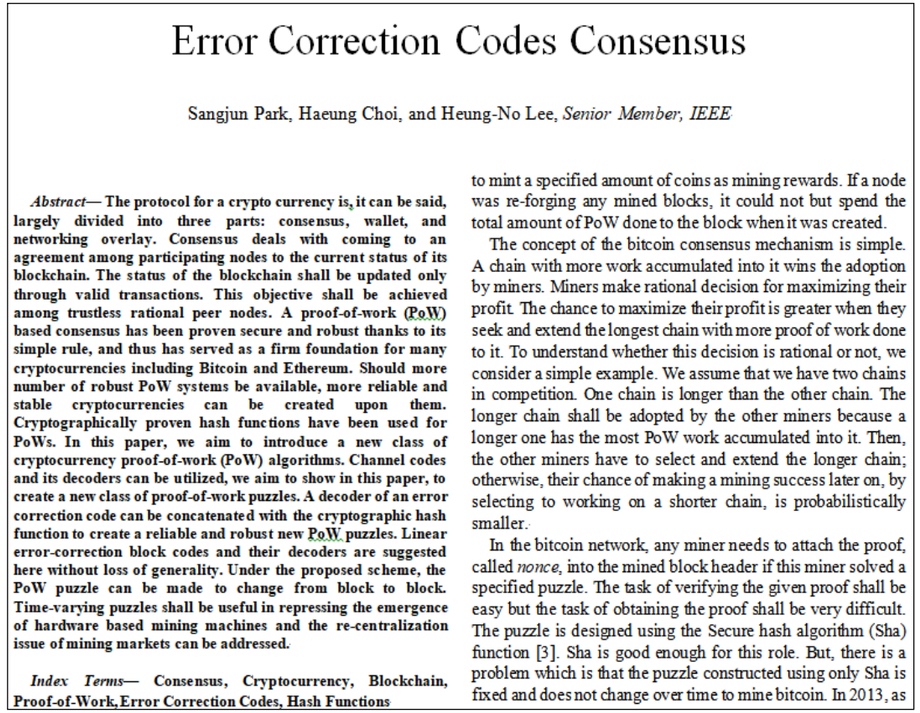
- The decision of mining success is made with the output of the above decoder.

Novel ECCPoW Consensus mechanism, how!

▣ ECCPoW Engine

- **Compound code of SHA and LDPC decoder.**
- **Variable size of Parity Check Matrix (PCM) → Amt of resource (mem, comp) varies.**
- **PCM is varied by the hash of the previous block.**





Error Correction Codes Consensus

Sangjun Park, Haeung Choi, and Heung-No Lee, *Senior Member, IEEE*

***Abstract*— The protocol for a crypto currency is, it can be said, largely divided into three parts: consensus, wallet, and networking overlay. Consensus deals with coming to an agreement among participating nodes to the current status of its blockchain. The status of the blockchain shall be updated only through valid transactions. This objective shall be achieved among trustless rational peer nodes. A proof-of-work (PoW) based consensus has been proven secure and robust thanks to its simple rule, and thus has served as a firm foundation for many cryptocurrencies including Bitcoin and Ethereum. Should more number of robust PoW systems be available, more reliable and stable cryptocurrencies can be created upon them. Cryptographically proven hash functions have been used for PoWs. In this paper, we aim to introduce a new class of cryptocurrency proof-of-work (PoW) algorithms. Channel codes and its decoders can be utilized, we aim to show in this paper, to create a new class of proof-of-work puzzles. A decoder of an error correction code can be concatenated with the cryptographic hash function to create a reliable and robust new PoW puzzles. Linear error-correction block codes and their decoders are suggested here without loss of generality. Under the proposed scheme, the PoW puzzle can be made to change from block to block. Time-varying puzzles shall be useful in repressing the emergence of hardware based mining machines and the re-centralization issue of mining markets can be addressed.**

***Index Terms*— Consensus, Cryptocurrency, Blockchain, Proof-of-Work, Error Correction Codes, Hash Functions**

to mint a specified amount of coins as mining rewards. If a node was re-forging any mined blocks, it could not but spend the total amount of PoW done to the block when it was created.

The concept of the bitcoin consensus mechanism is simple. A chain with more work accumulated into it wins the adoption by miners. Miners make rational decision for maximizing their profit. The chance to maximize their profit is greater when they seek and extend the longest chain with more proof of work done to it. To understand whether this decision is rational or not, we consider a simple example. We assume that we have two chains in competition. One chain is longer than the other chain. The longer chain shall be adopted by the other miners because a longer one has the most PoW work accumulated into it. Then, the other miners have to select and extend the longer chain; otherwise, their chance of making a mining success later on, by selecting to working on a shorter chain, is probabilistically smaller.

In the bitcoin network, any miner needs to attach the proof, called *nonce*, into the mined block header if this miner solved a specified puzzle. The task of verifying the given proof shall be easy but the task of obtaining the proof shall be very difficult. The puzzle is designed using the Secure hash algorithm (Sha) function [3]. Sha is good enough for this role. But, there is a problem which is that the puzzle constructed using only Sha is fixed and does not change over time to mine bitcoin. In 2013, as

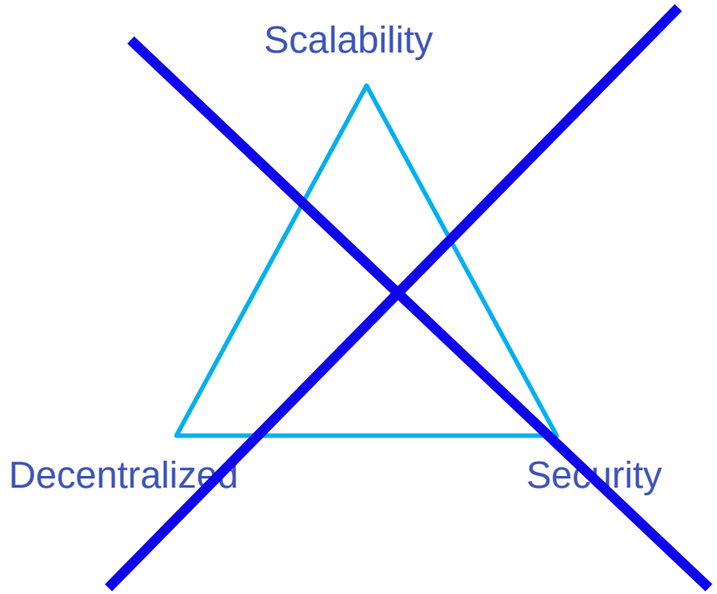
※ 국제 학술지 IEEE trans. Information Forensics and Security에 제출예정

Blockchain Trilemma?

“
blockchain systems can only at most have two of the following three properties

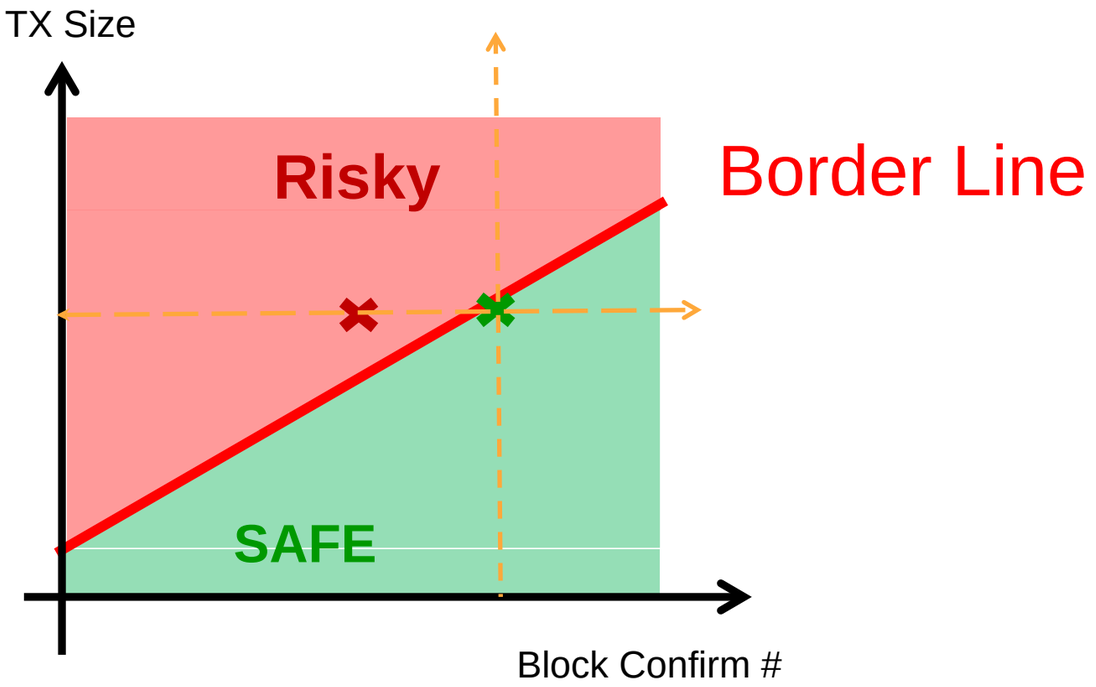
- Vitalik Buterin, Sharding FAQ
https://github.com/ethereum/wiki/wiki/Sharding-FAQ

”



- Wrong approach!
- Not in a single blockchain, can it be achieved!
- ***We shall promote many decentralized secure (DeSecure) blockchains and approach the scalability problem!***

Profitable DS Risk Analysis



<Risk Analysis Algorithm>

Profitable Double-Spending Attacks

Jehyuk Jang and Heung-No Lee, *Senior Member, IEEE*

***Abstract*—Our aim in this paper is to investigate the profitability of double-spending (DS) attacks that manipulate a priori mined transaction in a blockchain. Up to date, it was understood that the requirement for successful DS attacks is to occupy a higher proportion of computing power than a target network's proportion; i.e., to occupy more than 51% proportion of computing power. On the contrary, we show that DS attacks using less than 50% proportion of computing power can also be vulnerable. Namely, DS attacks using any proportion of computing power can occur as long as the chance to making a good profit is there; i.e., revenue of an attack is greater than the cost of launching it. We have novel probability theory based derivations for calculating time finite attack probability. This can be used to size up the resource needed to calculate the revenue and the cost. The results enable us to derive sufficient and necessary conditions on the value of a target transaction which make DS attacks for any proportion of computing power profitable. They can also be used to assess the risk of one's transaction by checking whether or not the transaction value satisfies the conditions for profitable DS attacks. Two examples are provided in which we evaluate the attack resources and the conditions for profitable DS attacks given 35% proportion of computing power against *Syscoin* and *BitcoinCash* networks, and quantitatively shown how vulnerable they are.**

***Index Terms*— Blockchain, Bitcoin, Double-Spending Attack, Profit, Gambler's Ruin Theorem, Poisson Counting Process.**

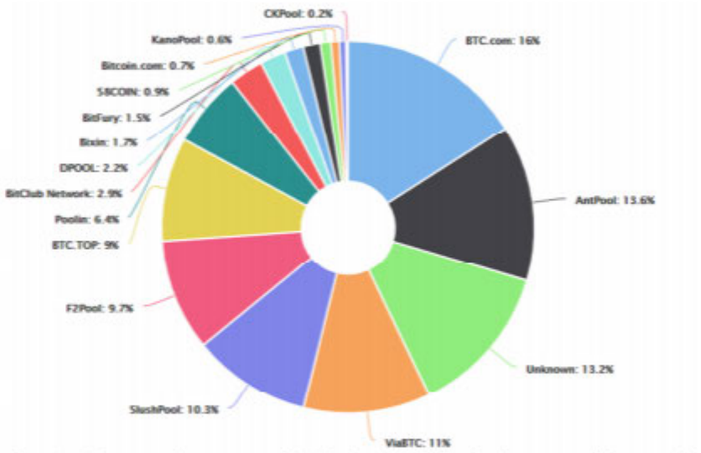


Fig. 1. Computation power distribution among the largest mining pools provided by *blockchain.com* (date accessed: 22 Oct. 2018).

peers in a network to share a common chain. If a full node succeeds in generating a new block, he/she has the latest version of the chain. All of the nodes in the network continuously communicate with each other to share the latest chain. If a node suffers from a conflict between two or more different chains, the consensus rule provides a rule that a single chain is selected. Satoshi Nakamoto suggested the *longest chain consensus* for *Bitcoin* protocol which conserves the longest chain among the conflictions [1]. There are also

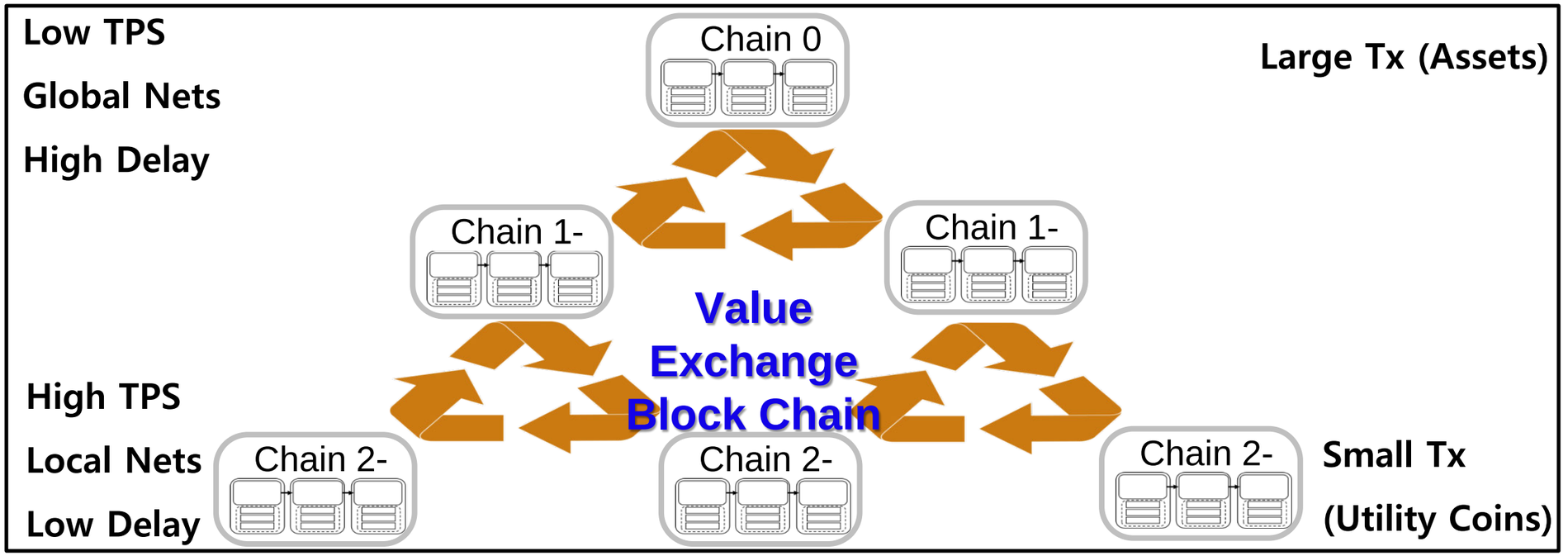
※ Jahyuk will present tomorrow afternoon!

※ 국제 학술지 IEEE trans. Information Forensics and Security에 제출됨

https://arxiv.org/ftp/arxiv/papers/1903/1903.01711.pdf

Provision of DeSecure chains, use ecosystem to solve Scalability issue!

- ▣ Global chains → national chains → local chains
- ▣ One chain is designed to hold only up to 20 DApps



<Multi-level DeSecure chains>

Block code

- A block code C(N, Rate, **G**, **F**, ENC, DEC, GF(q)) is well defined as a collection of codewords. When, $q = 2$, it is a binary system.
- N is the dimension of the code (e.g. $N = 512$)
- Rate = $(N - M)/N$ is the rate of the code, where M < N.
- For example, with $N = 1024$ and $M = 256$, Rate = 3/4.

- **G** is the Generator matrix with dimension $N \times (N - M)$.
- **F** is the Check matrix with dimension, $M \times N$.
- **G** and **F** are orthogonal to each other, i.e., $\mathbf{FG} = \mathbf{0}$.

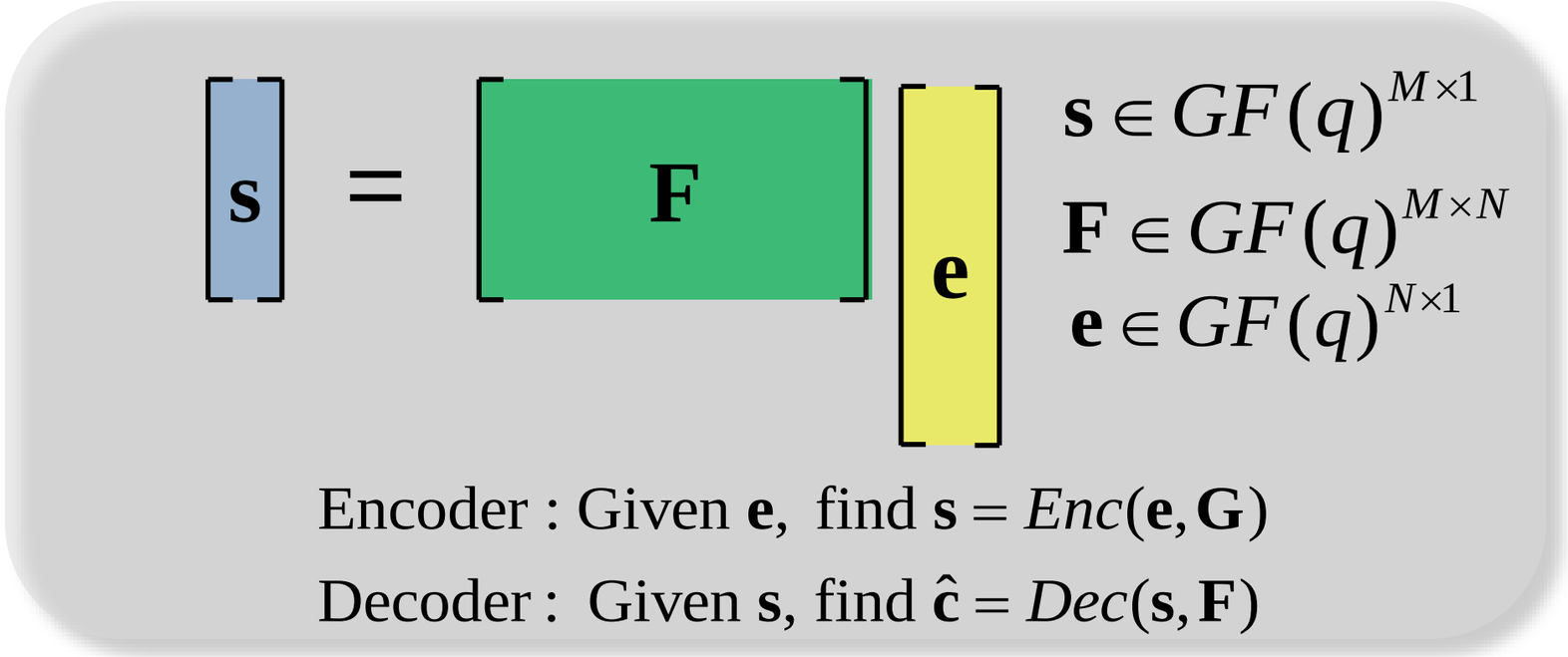
- A message vector m is an $(N - M) \times 1$ vector.
- A codeword **c**, an N x 1 vector, is an element of the code and can be generated by multiplying a message vector m to the Generator matrix **G**, i.e., $\mathbf{c} = \mathbf{Gm}$.

- Galois Field of size q, GF(q), is used for addition and multiplication operations and storage of numbers in the system.

Block code, encoder and decoder

- ENC implies the encoder function, i.e., ENC takes the message vector **m** as the input and produces a codeword vector corresponding to it, e.g. **c** = ENC(**G**, **m**).

- DEC implies the decoding function; DEC takes an arbitrary vector e and returns a closest codeword $\mathbf{c}\hat{}$, i.e., $\mathbf{c}\hat{} = \mathrm{DEC}(\mathbf{F}, \mathbf{e})$.



$$M < N$$

Decoder

- DEC is to find a codeword $\mathbf{c}\hat{}$ most close to the input word $\mathbf{e}$.
- For the concept of distance, the Hamming distance can be used.
  For example, $DH(\mathbf{e}, \mathbf{c}\hat{}) = ||\mathbf{e} - \mathbf{c}\hat{}||_0$ is the number of non-zero values in the $(\mathbf{e} - \mathbf{c}\hat{})$ vector.

- There are many ways to find $\mathbf{c}\hat{}$ satisfying $\mathbf{F}\mathbf{c}\hat{} = \mathbf{0}$.

- We propose to use the message passing graph decoder for its excellency in accuracy and superiority in decoding speed.
  This is to prevent a cheating attack in which a smart miner comes up with a new decoder algorithm of his own developed and outpaces the regular miners using the designated decoder. If this is allowed, a hidden advantage goes to the smart miner.

Geometrical Explanation

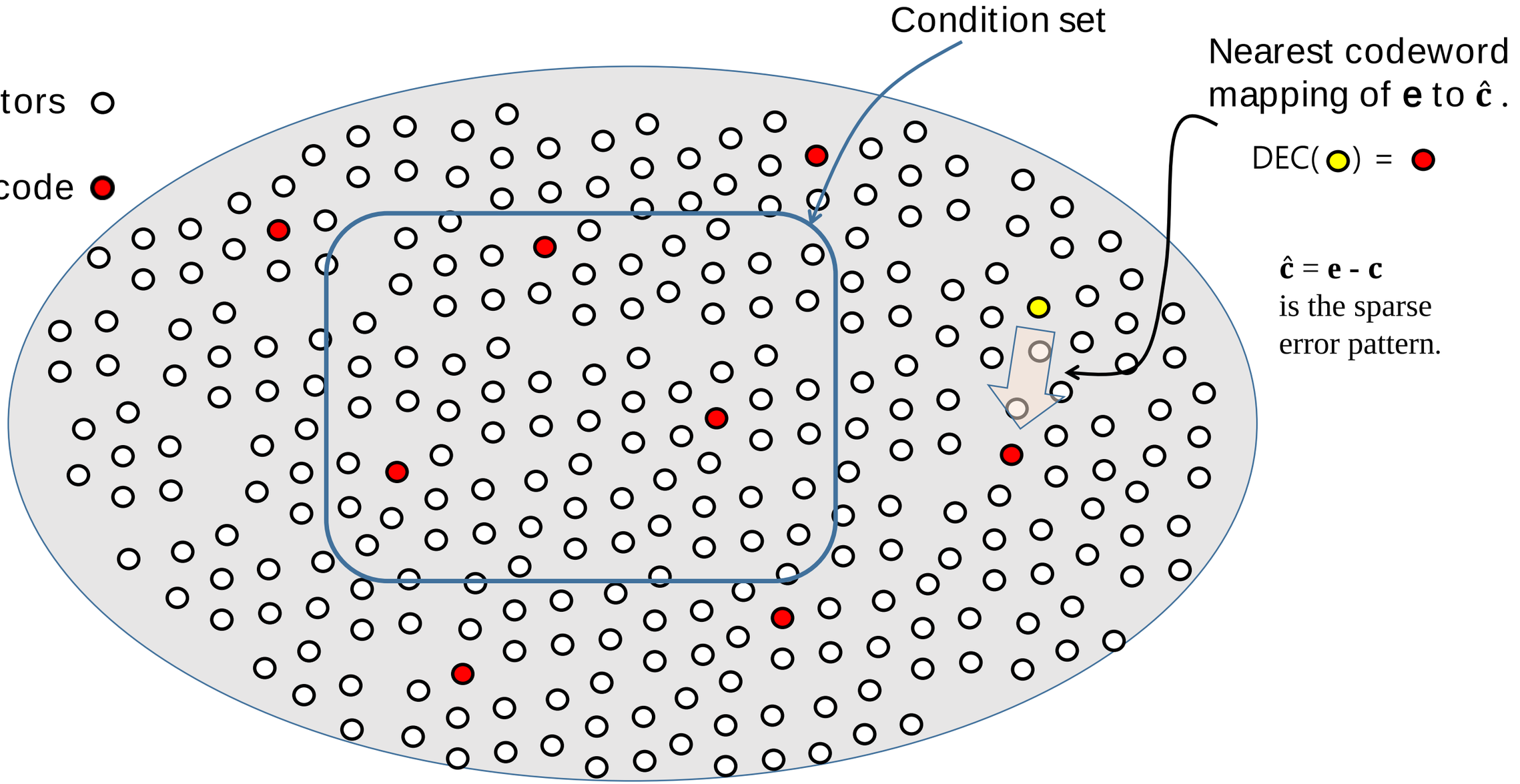
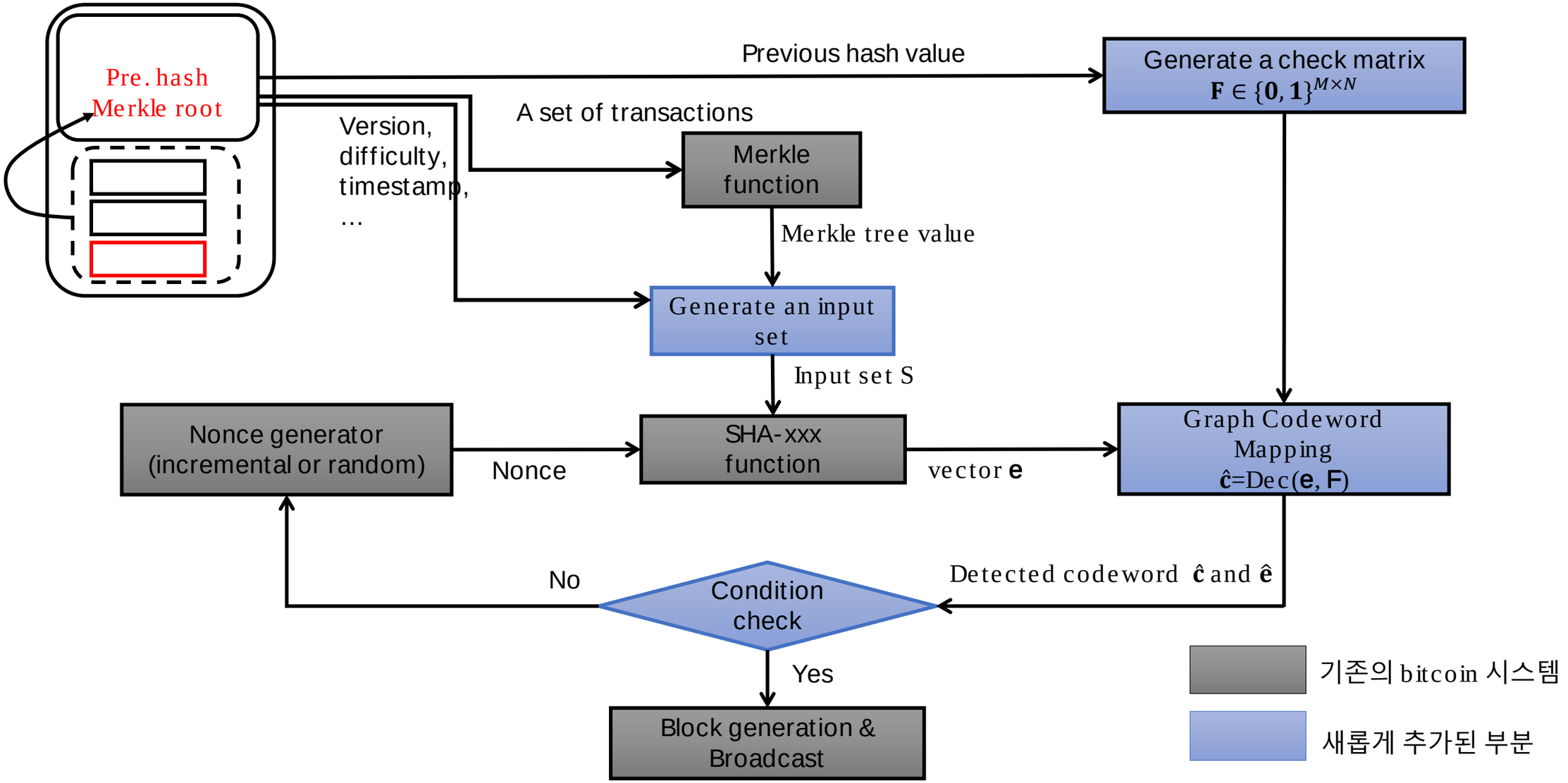
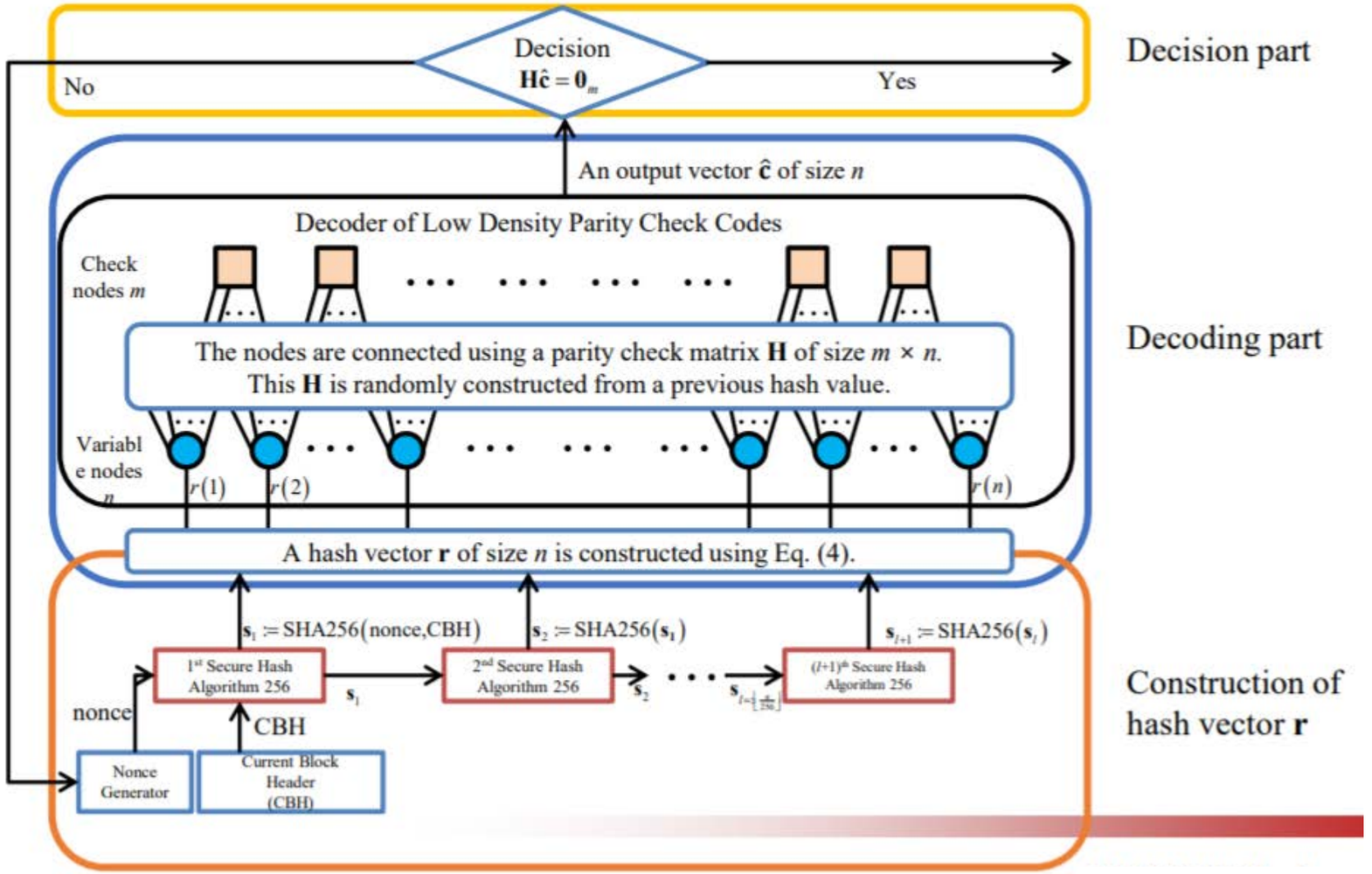


Diagram of ECCPoW

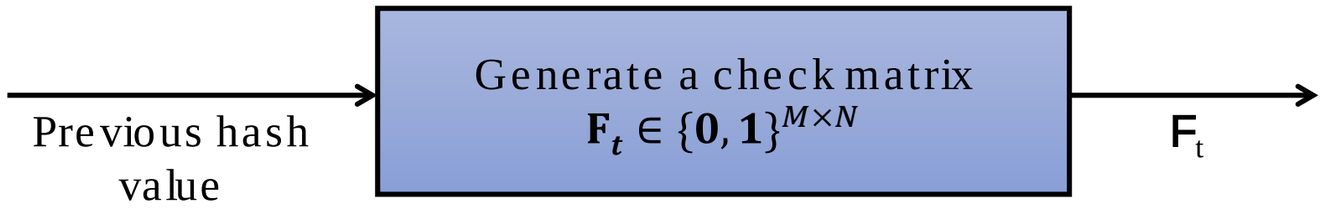


Big Picture of ECCPoW



Generate a Check Matrix

- Parameter set $S_t = \{h_t\text{-}1,\ code\ parameters\}$;
- GenCheckMatrix(S_t) = $\mathbf{F}_t$
- Generate a check matrix $\mathbf{F}_t$ w.r.t. previous hash h_t-1.
- Takes the previous hash h_t-1 as the input to this routine.
- That is, $\mathbf{F}_t$ changes from block to block.



Pseudo Code of the Decoder

- Input:
- ✓ Hard decision of a priori LLR: $L_a^t = \mathbf{e}[t]$

- Iteration: repeat until converse
- Update variable-to-check node messages for $t = 1, 2, .., N$ and $\forall l \in Q1(t)$:

$$L^{t \to l} = \left[\sum_{l' \in Q1(t) \setminus l} (L_a^t \oplus L^{l' \to t}) \Big/ (j-1) \right]$$

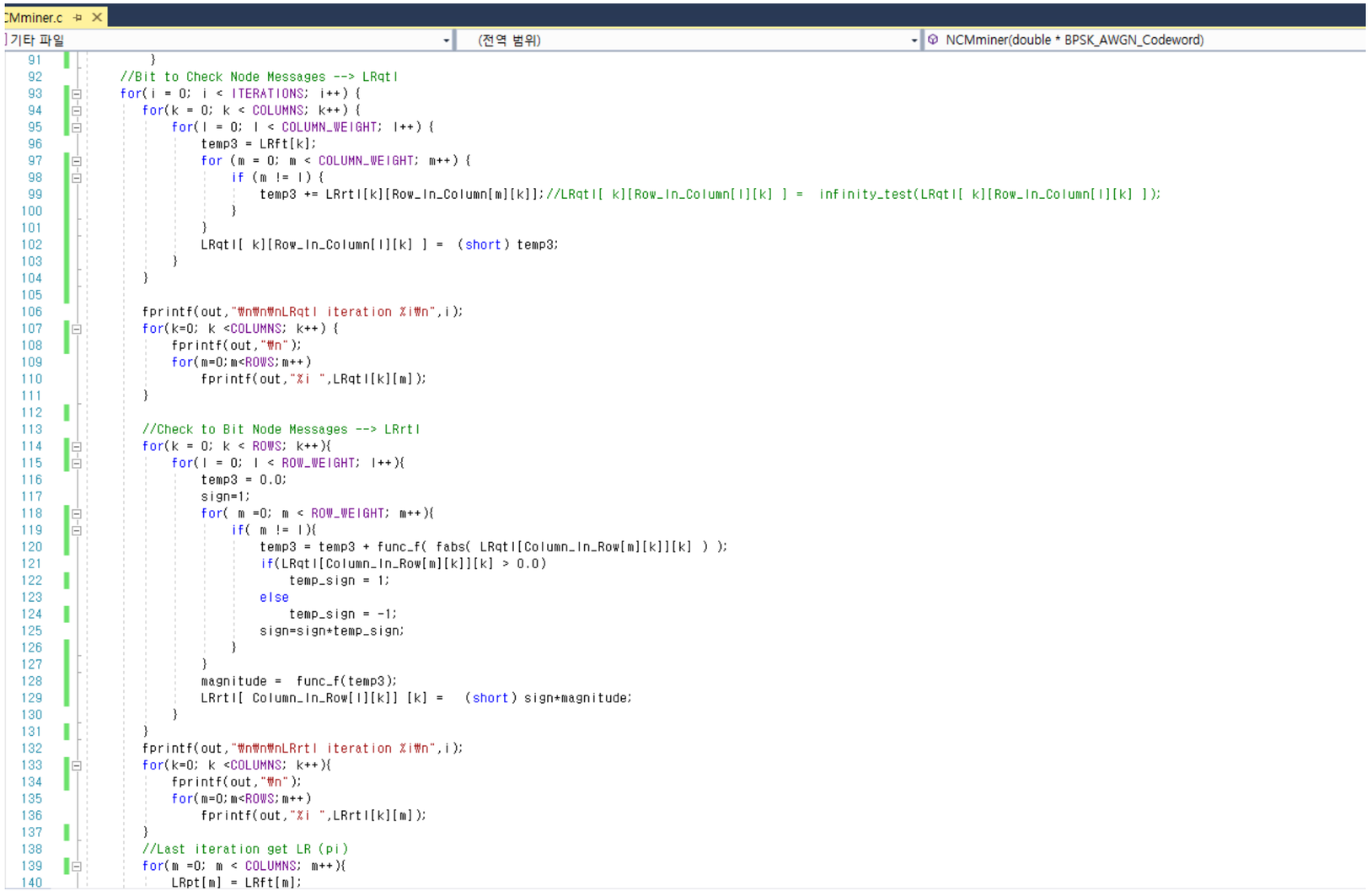
- Update check-to-variable node messages for $l = 1, 2, .., M$ and $\forall t \in Q2(l)$:

$$L^{l \to t} = \oplus \sum_{t' \in Q2(l) \setminus t} L^{t' \to l}$$

- Output
- ✓ Hard decision of a posteriori LLR:

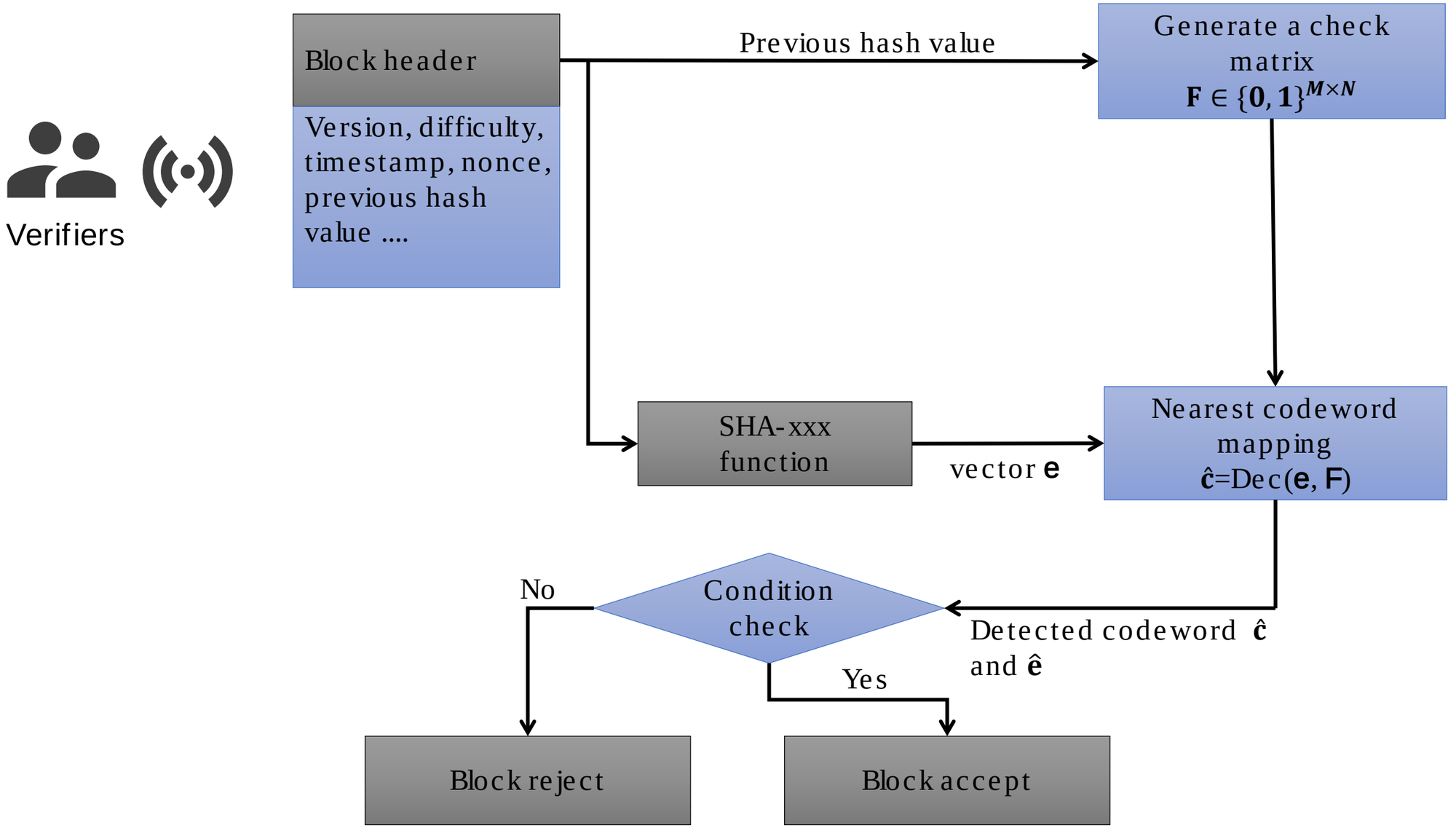
$$L^{t \to l} = L_a^t \oplus \left[\sum_{l' \in Q1(t)} L^{l' \to t} \Big/ j \right] \rightarrow \hat{\mathbf{c}}[t]$$

Implemented code in C




```
91          }
92      //Bit to Check Node Messages --> LRqtl
93      for(i = 0; i < ITERATIONS; i++) {
94          for(k = 0; k < COLUMNS; k++) {
95              for(l = 0; l < COLUMN_WEIGHT; l++) {
96                  temp3 = LRft[k];
97                  for (m = 0; m < COLUMN_WEIGHT; m++) {
98                      if (m != l) {
99                          temp3 += LRrtl[k][Row_In_Column[m][k]];//LRqtl[ k][Row_In_Column[l][k] ] =  infinity_test(LRqtl[ k][Row_In_Column[l][k] ]);
100                     }
101                 }
102                 LRqtl[ k][Row_In_Column[l][k] ] =  (short) temp3;
103             }
104         }
105
106         fprintf(out,"\n\n\nLRqtl iteration %i\n",i);
107         for(k=0; k <COLUMNS; k++) {
108             fprintf(out,"\n");
109             for(m=0;m<ROWS;m++)
110                 fprintf(out,"%i ",LRqtl[k][m]);
111         }
112
113         //Check to Bit Node Messages --> LRrtl
114         for(k = 0; k < ROWS; k++){
115             for(l = 0; l < ROW_WEIGHT; l++){
116                 temp3 = 0.0;
117                 sign=1;
118                 for( m =0; m < ROW_WEIGHT; m++){
119                     if( m != l){
120                         temp3 = temp3 + func_f( fabs( LRqtl[Column_In_Row[m][k]][k] ) );
121                         if(LRqtl[Column_In_Row[m][k]][k] > 0.0)
122                             temp_sign = 1;
123                         else
124                             temp_sign = -1;
125                         sign=sign*temp_sign;
126                     }
127                 }
128                 magnitude =  func_f(temp3);
129                 LRrtl[ Column_In_Row[l][k]] [k] =   (short) sign*magnitude;
130             }
131         }
132         fprintf(out,"\n\n\nLRrtl iteration %i\n",i);
133         for(k=0; k <COLUMNS; k++){
134             fprintf(out,"\n");
135             for(m=0;m<ROWS;m++)
136                 fprintf(out,"%i ",LRrtl[k][m]);
137         }
138         //Last iteration get LR (pi)
139         for(m =0; m < COLUMNS; m++){
140             LRpt[m] = LRft[m];
```

Diagram of New Verifiers



New Functions in ECCPoW

- New functions
  1. int **H = GenCheckMatrix(int n, int w_c, int w_r, int seed);
  2. bool DEC(int **H, int *e, int n, int w_c, int w_r, int *c);
  3. void Dec_Difficulty(int &n, int &wc, int &wr, int level);

- These functions are the key parts of the proposed solution.
  1. They are implemented in C++.
  2. They are used to implement a mining routine

- An example of mining

1. generate block header with zero nonce.
2. Dec_Difficulty(&n,&wc,&wr,difficulty)
3. Seed = f(phv)
4. H = GenCheckMatrix(n, wc, wr, seed)
5. nonce = nonce + 1
6. e = SHA256(version, time, difficulty, nonce, mtv)
7. flag = DEC(H,e,n,wc,wr)
8. If flag == 0; go to step 4
9. Update chv and nonce.
10. Generate block and broadcast.

ECCPoW Hardfork

- New ECCPoW

  A new structure of the block header has been introduced and,

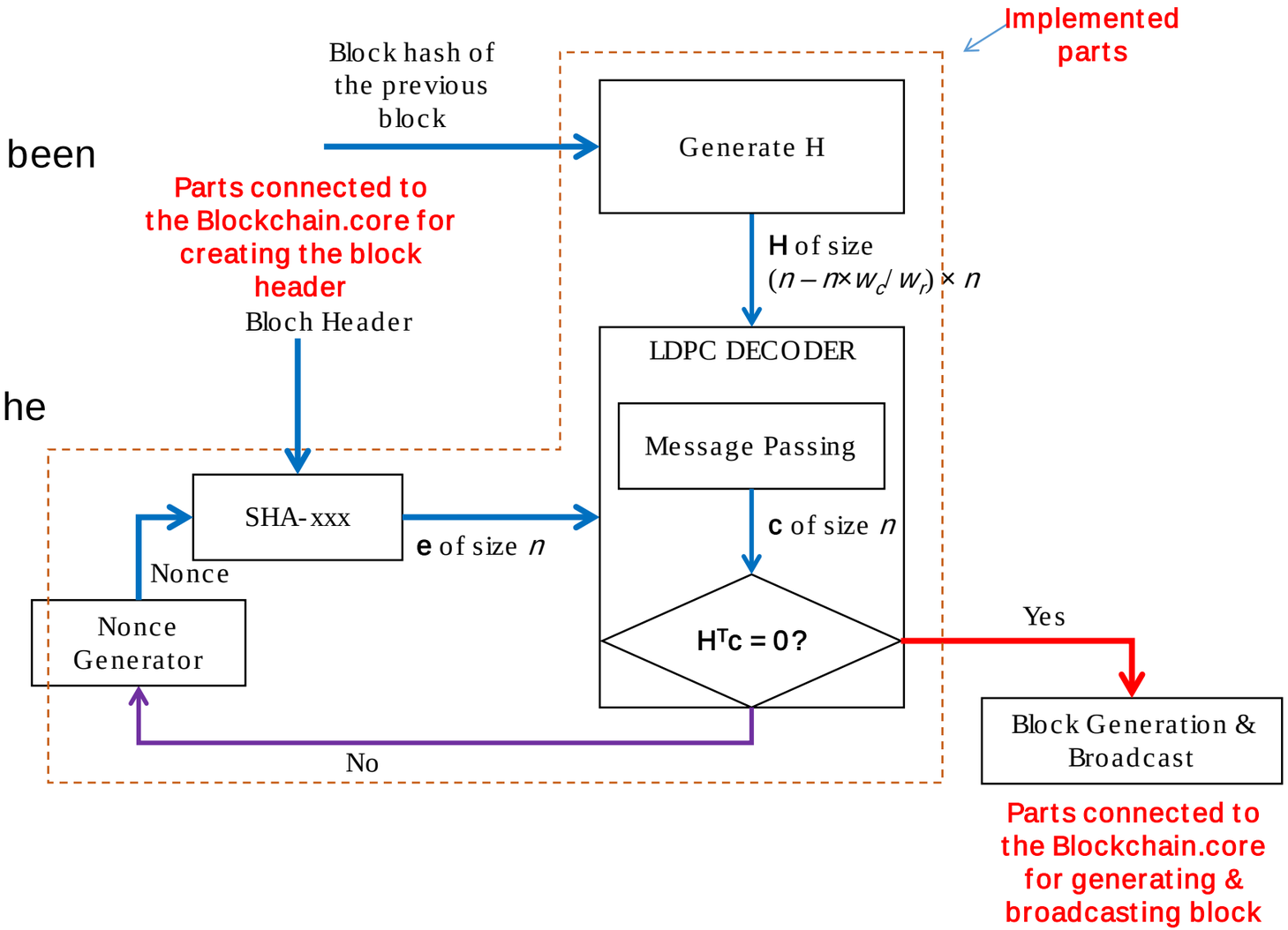
  three new functions are also have been introduced.

- We aim to link these functions to existing the blockchain. For example,

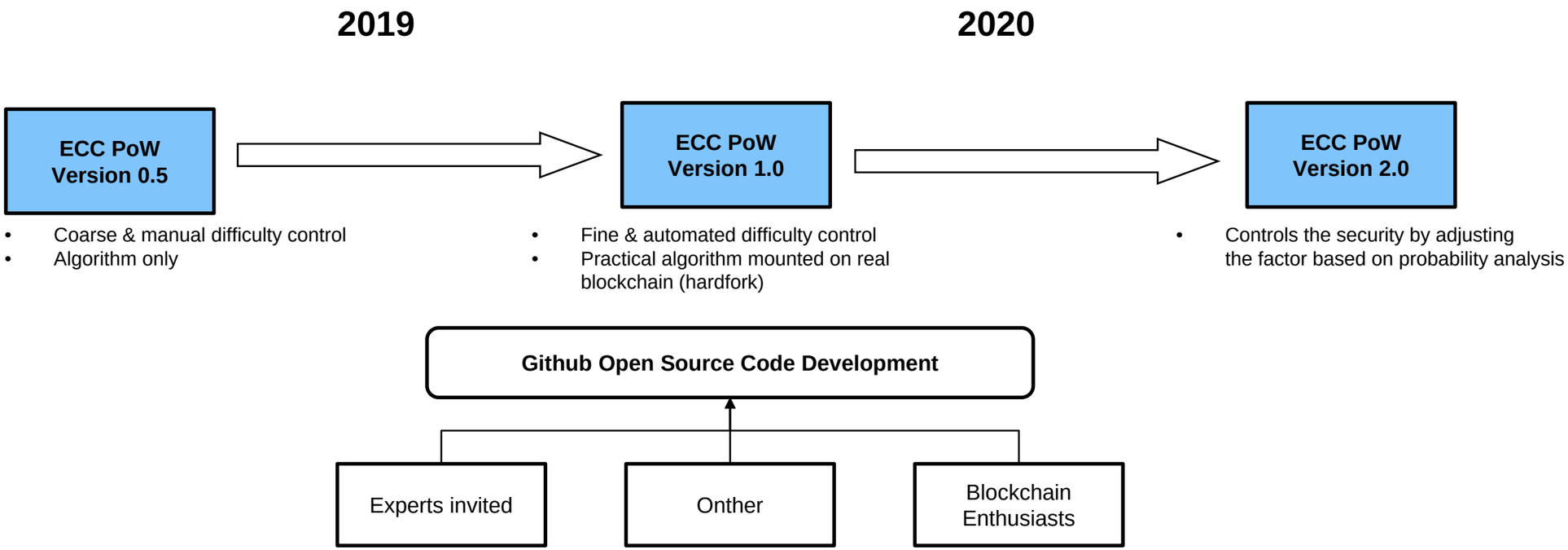
  mining function,

  chain validation function,

  consensus function and so on.



DeSecure Blockchain Release Plan



Impact of ECCPoW 1:
It is easier to start a new blockchain network.

- A large blockchain network is stable and not easy to disrupt.

- Today there are mining equipment renting sites.
- A new borne blockchain network needs to grow, but newbies are much more vulnerable to 51% attacks.

- DeSecure blockchain networks with ECCPoW do not suffer from such problems since there are no mining equipment available for ECCPoW.

Impact of ECCPoW 2:
One can make multiple blockchain networks

- It is easy to make a new blockchain with ECCPoW.

- Suppose hardforking a Bitcoin, and an Ethereum, with ECCPoW.
- Let us call them BTC-ECC and ETH-ECC protocols.

- Make the first blockchain network by running ETH-ECC over a network (Pusan coin)
- Make the second blockchain network by running BTC-ECC over other network (Gwangju coin)
- Make the third blockchain network by running ETH-ECC over another network (Seoul coin)
- Make the fourth blockchain network by running BTH-ECC over yet another network (Korea coin)

- Each cryptocurrency is independent with its own genesis block and random starting seed, and can be adjusted sufficiently strong for its regional requirement in the sense of scalability, security and decentralization.
- These blockchains are inter-connected at the local, regional, and national, transnational level.

Impact of ECCPoW 3: Resolving the Scalability Trilemma

- Trilemma by V. Buterin is well known: Only up to **two out of the three** virtues such as Scalibility, Decentralization and Security **can be achieved simultaneously**.

- With ECC, each blockchain is already very strong in decentralization.
- Each EEC blockchain is flexible enough to fit into various settings of transaction speeds and security levels.
  - Campus ECC blockchain networks can be set to work very fast allowing up to 100s of thousands of TXs per second since the delay of the underlying communications network is very small.
  - Regional ECC blockchain networks can be set to work fast, i.e. allowing up to 10s of thousands of TXs per sec.
  - National ECC blockchain networks can be set sufficiently fast for covering inter-regional transactions.
  - Transnational ECC blockchain networks shall be set to work slow due to large delays.

- All these DeSecure chains started up with its own seed and decentralized levels are mutually independent and each one can be set to work at the required level of security and speed to serve its purpose.
- These DeSecure chains can be inter-connected via *distributed value-exchange* networks.
- The connected ECC blockchains can be named the ECC Blockchain International.
- *ECC Blockchain International* as a whole can serve to resolve the Scalability Trilemma.

Impact of ECCPoW 4:
It is safe to use a time-proven blockchain protocol.

- Bitcoin protocol has withstood the tough test of time.
- Thus, the networking part and the wallet part are robust enough.

- PoW is problem. Yes.
- But it is not the problem of PoW.
- It is the **fixedness** of the PoW puzzle.

- ECCPoW puzzles can be made to vary over time.

- The problematic consensus part with a fixed PoW can be replaced with the new ECC PoW consensus.

Impact of ECCPoW 5:
The complexity of ECCPoW puzzles can be set to grow very large; thus the cost for hardware acceleration is boundless.

- ECCPoW is a computer algorithm!
- Thus it is not impossible to find a hardware acceleration solution for it.

- ECCPoW puzzle can be represented as a randomly connected bipartite graph.
- In order to parallelize the algorithm, more memory and computation resource need to be allocated.
- The size of ECCPoW puzzle can grow very large.
- As the size of the puzzle grows, the more needed is the memory and computation resource.
- With ECCPoW puzzles, therefore, one can easily deter the emergence of hardware acceleration solution.

- Deterrence to hardware acceleration offers a blockchain network with small power consumption requirement.

Development Schedule

- Open research platform
  - Source codes github uploaded
  - Open development

- 2019 plan
  - ECCPoW 0.5 Version
  - Ethereum and Bitcoin Hardforks with ECCPoW 0.5v
  - Develop them into Ethereum ECCPoW 1.0v and Bitcoin ECCPow 1.0v

- 2020 plan
  - Network growth at least by 10,000 nodes worldwide
  - Co-working with Bitcoin and Ethereum communities



Concluding Remarks

- PoW is fundamental for blockchains' immutability.
  - You put PoW to a block, you get the benefit of data immutability.
  - Recentralization issue is problem due to fixeness of PoW puzzles, not due to PoW itself.

- Trilemma by V. Buterin is well known. We seek to get two Security and Decentralization.
  - Flexible puzzles enabled by ECCPoW can resolve the recentralization problem;
  - PoW has shown to be the most secure.

- Scalability is left to the ecosystem of DeSecure blockchains.
  - Multiple layers of ECCPoW blockchains can operate simultaneously resolving the issues of scalability and thus breaking the trilemma.

- ECCPoW blockchains can play a crucial role in ushering in the ideals of blockchains and advance our society to the next level!

Selected References of GIST Blockchain Economy Center

- [Lee1] JH Jang and Heung-No Lee, "Profitable Double Spending Attacks," March 5th, 2019 submitted to IEEE Trans. Information Forensics and Securities, downloadable from https://arxiv.org/abs/1903.01711.
- [Lee2] 장재혁, 이흥노, "50%미만 이중 지불 공격", OSIA S&TR Journal, Vol. 32, No. 1, Mar. 2019. (pdf) (GIST 연구원 GRI 사업, 과학기술응용연구단 실용화사업)
- [Lee3] 정현준, 이흥노, "암호화폐 투자와 규제 현황", 한국정보과학회, 정보과학회지, 제 36권, 제 12호, pp. 49-56, Dec, 2018. (pdf)
- [Lee4] 박상준, 김형성, 이흥노, "Introduction to Error-Correction Codes Proof of Work," 블록체인경제 특집호, 대한전자공학회지, June 2019.
- [Lee5] Sangjun Park, HS Kim, Heung-No Lee, "Time-Variant Proof-of-Work Using Error-Correction Codes," to be submitted to IEEE Trans. Information Forensics and Securities.
- [Lee6] Mohamed Yaseen.J, Giljun Jung and Heung-No Lee."Decentralized Framework for Medical Images Based on Blockchain and Inter Planetary File System", The 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society(EMBC 2019), Berlin, Germany, Jul. 23-27, 2019.
- [Lee7] Please visit INFONET home page https://infonet.gist.ac.kr/?page_id=14 for more references.

- Thank you!



Heung-No Lee, GIST, South Korea
Home page: http://infonet.gist.ac.kr
Facebook/ Publication ID: Heung-No Lee
E-mail: heungno@gist.ac.kr

- Q&A

- We are actively looking for blockchain students to join us.
- Send me an e-mail!

ITRC 블록체인 지능 융합센터
2021-10 워크샵

영지식 센싱 탈중앙 고성능 분산 원장 기술 개발

GIST 이흥노 교수

2021.10.29

CONTENTS

1. 연구 목표
2. 1차년도 연구 내용
3. 2차년도 연구 계획

1. 연구 목표

제1 세부과제 연구 목표

세부과제 1: 영지식 센싱 탈중앙 고성능 분산원장 기술 개발

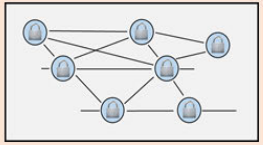
3단계

2단계

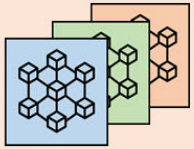
1단계

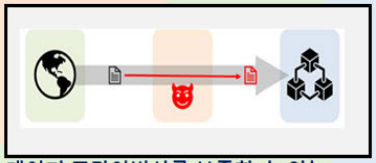
다중 복합 구조 계층형 블록체인 구조 확장 기술

[3단계] 다중 복합구조 계층형 블록체인 플랫폼 기술 운용

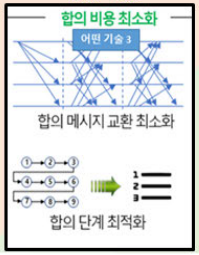


[2단계] 다중 복합구조 계층형 블록체인 개발 및 테스트

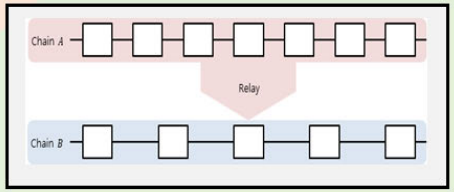




[1단계] 분산합의 기술 분석 및 개발



[2단계] 데이터 프라이버시를 보존할 수 있는 연합학습 방법론 개발

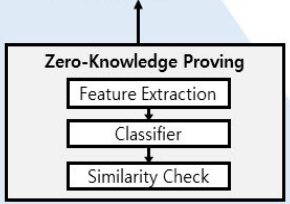


[1단계] 초경량・초확장 블록체인 개발

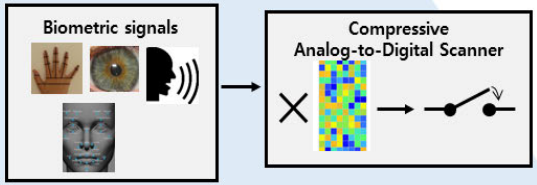
고성능 합의알고리즘 등 블록체인 기반기술

[3단계] 영지식 아날로그 생체 인증 시스템 연구



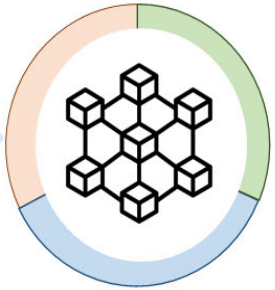


[2단계] 영지식 증명 및 검증 프로토콜 연구



[1단계] 신원 인증 키 아날로그-디지털 스캐너 연구

DID를 위한 초분광 카메라, 화자인식 기반 ZKP 생체인증기법

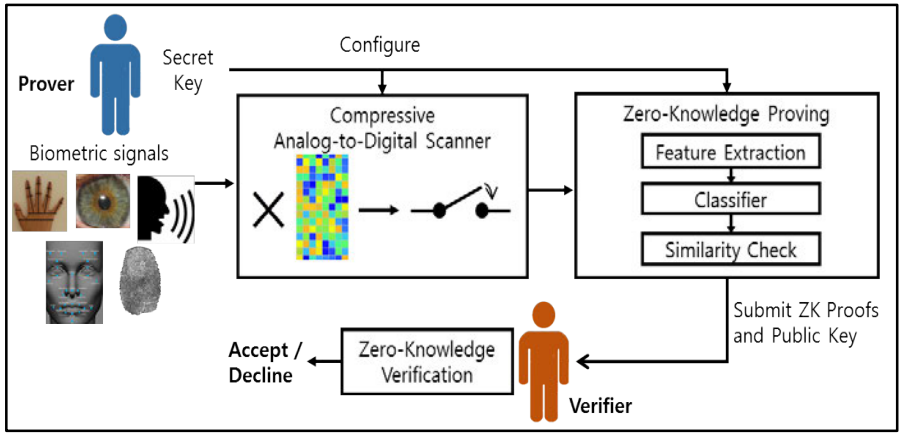


블록체인 센싱 탈중앙 고성능 분산원장 기술

제1 세부과제 연구 목표

최종 목표: 영지식 센싱 탈중앙 고성능 분산원장 기술 개발

- DID를 위한 초분광 카메라, 화자인식 기반 ZKP 생체인증 기법 개발 (이흥노 교수)
- 고성능 합의 알고리즘 등 블록체인 기반기술 연구
- 다중 복합구조 계층형 블록체인 구조 확장 기술 연구
- 블록체인 및 클라우드 상에서 동작하는 연합학습 플랫폼 개발



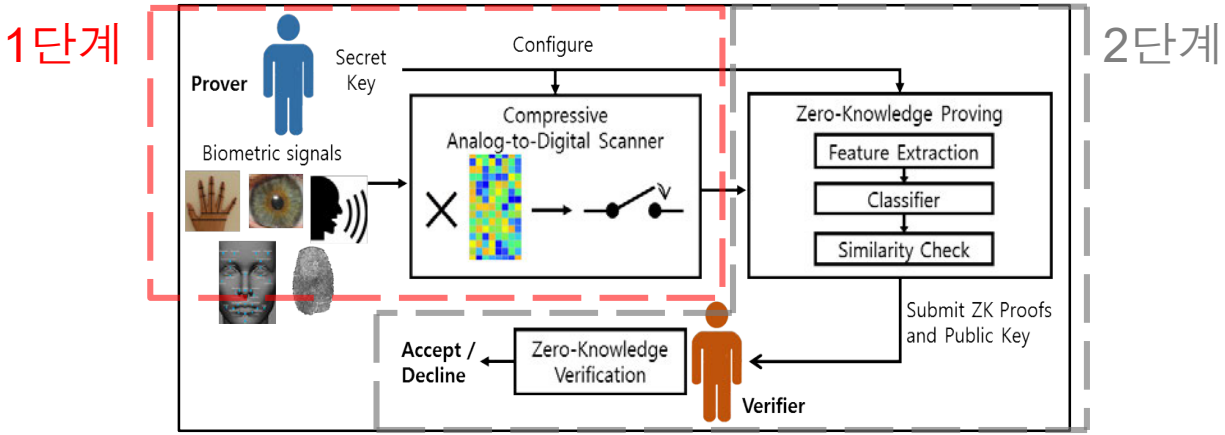
<ZKP 아날로그 생체인증 시스템 (GIST 발명공개)>

연구실 1단계 (1년차+2년차) 목표

1단계 목표: 초분광 이미지 비가역적 아날로그-디지털 스캐너 연구

- ✓아날로그 신원 인증 키 연구
- ✓비가역적 아날로그-디지털 스캐너 설계
- ✓초분광압축이미징 장치 구현 및 데이터 확보

2단계 목표: 영지식 증명 및 검증 프로토콜 연구



<ZKP 아날로그 생체인증 시스템 (GIST 발명공개)>

2. 1차년도 연구 내용

초분광이지미 신원 인증 키 연구

목표: 초분광이미지 사용 신원인증 (Biometrics) 동향 분석

목적: 소형 압축 분광기의 요구사항에 적합한 이미지 소스 결정

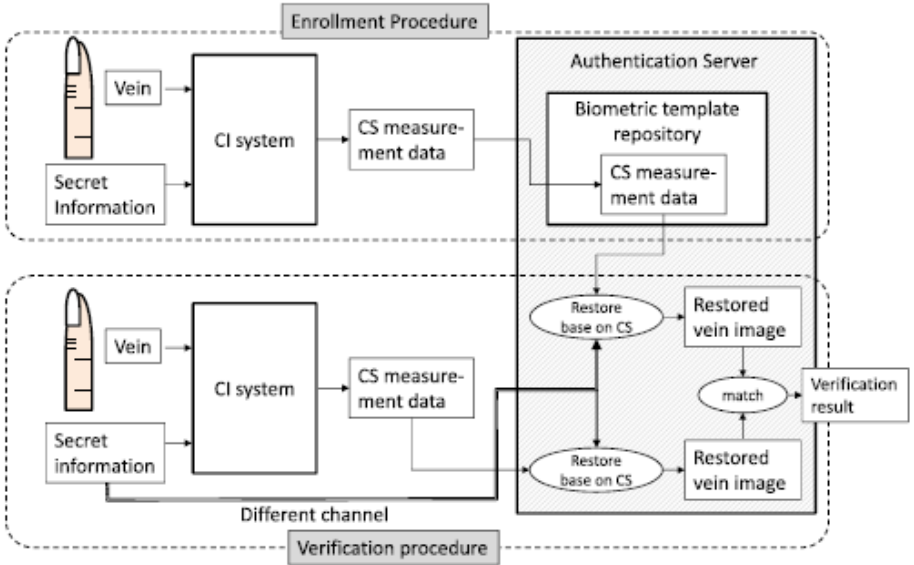
- 피부의 표피하부 다층 레이어간의 색소 분포 (멜라닌, beta-carotene, bilirubin, 헤모글로빈) 모델과 분류 활용을 소개함 [Moshen20]
- 피부 하위구조가 개인마다 충분히 다르며, 시간에 따라 안정적임을 확인함 [Pan03]
- 다수의 눈 주위 이미지를 활용한 신원분류를 제안하였고, 여러 알고리즘을 ensemble 하여 342명을 분류하는 문제에서 최대 99.7%의 성공률을 얻음 [Muhammad15]
- 얼굴 이미지를 활용한 신원분류를 제안하였고, 초분광 대역 중 일부 대역만을 희소하게 사용하여 CNN의 분류성능을 개선하였으며, 25명을 분류하는 문제에서 최대 97%의 성공률을 얻음 [Zhihua20]

비가역적 아날로그-디지털 스캐너 설계

목표: 비가역적 아날로그-디지털 스캐너 설계 및 활용 생체인증 동향 분석

랜덤압축센싱 스캐너의 생체인증 활용 사례

- 생체 Biometric 이미지의 유출 (spoofing attack) 방지를 위해, DMD compressive sensing 손가락 혈관 이미지를 사용하는 암호키 기반 생체인증 시스템 제안 [Hiroyuki13]



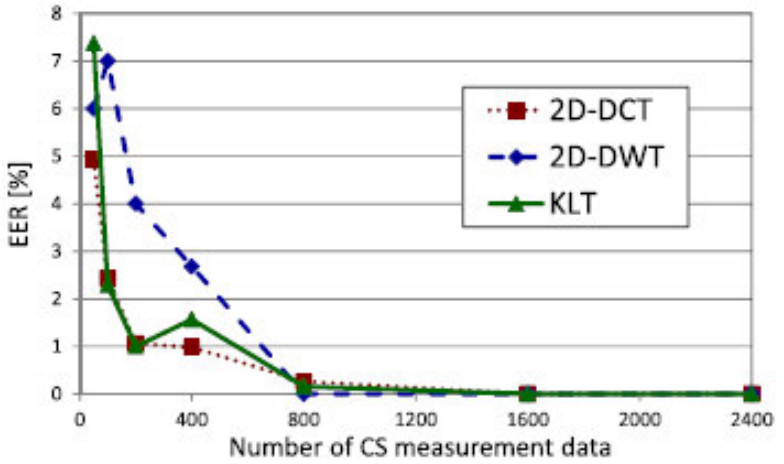


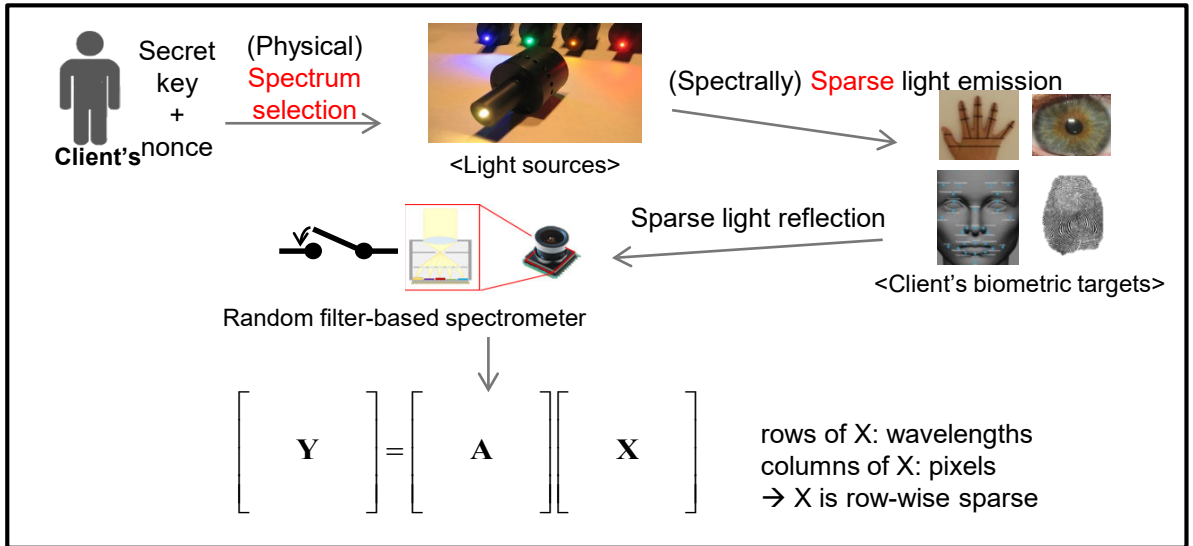
Fig. 12. EER depending on the number of CS measurement data in TPT.

비가역적 아날로그-디지털 스캐너 설계

목표: 비가역적 아날로그-디지털 스캐너 설계 및 활용 생체인증 동향 분석

제안하는 비가역적 초분광 스캐닝 시스템의 특징

- 광원의 파장 대역을 무작위로 조합함 → 분광기의 압축 특성이 무작위로 선택됨
- 광원 조합을 알아야만(client만이) 원 신호를 복원 가능
- 후에 client가 자가인증 후 영지식 증거와 public key를 server에 제출



3. 2차년도 연구 계획

비가역적 아날로그-디지털 스캐너 설계

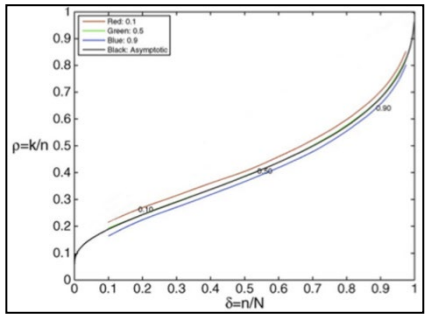
목표: 비가역적 아날로그-디지털 스캐너 설계 및 활용 생체인증 동향 분석

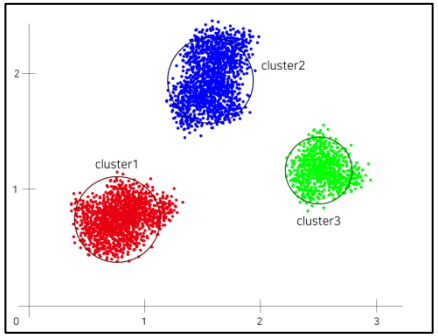
제안하는 비가역적 초분광 스캐너의 요구사항

- ✓제한된 수와 제한된 조합의 광원을 사용
- ✓광원 조합을 모를 때, 압축된 신호로부터 원 신호를 복원 할 수 없어야 함
- ✓원 신호의 희소 정도가 개인을 특징지을 수 있을 정도로 커야 함

요구사항을 만족하는 시스템 파라미터 탐색 필요

- 광원의 희소 정도
- 압축 신호(snapshot)의 개수
- 압축 신호의 길이



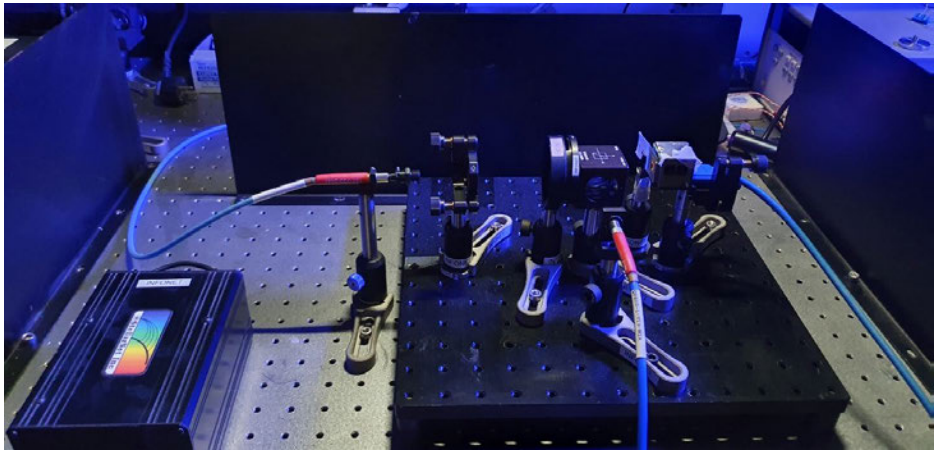


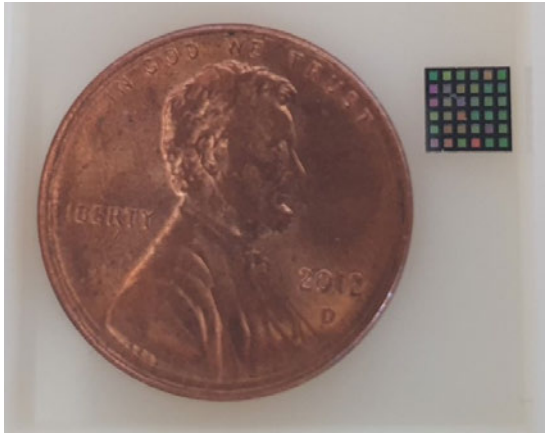
<압축 신호가 복원될 확률은 낮으면서 희소 신호의 inter-class incoherence 는 충분히 큰 희소 정도를 찾아야 함>

초분광압축이미징 장치 구현 및 데이터 확보

목표: 기 보유한 랜덤필터기반 compressive 초분광 기술을 실제 구현

- 랜덤필터기반 소형 초분광 기술을 보유하고 있음 [Lee20], [Lee13A], [Lee13B]
- 이를 바탕으로 실물 장치를 구현
- 실물 장치를 사용하여 데이터 확보 기대





Q&A

REFERENCES

[Pan03] Pan, Z., et al.: Face recognition in hyperspectral images. IEEE Trans. Pattern Anal. Mach. Intell. 25(12), 1552–1560 (2003)

[Moshen20] Ardabilian M., Zine AM., Li S. (2020) Multi-, Hyperspectral Biometrics Modalities. In: Nait-ali A. (eds) Hidden Biometrics. Series in BioEngineering. Springer, Singapore. https://doi.org/10.1007/978-981-13-0956-4_8

[Muhammad15] Muhammad Uzair, Arif Mahmood, Ajmal Mian, Chris McDonald, “Periocular region-based person identification in the visible, infrared and hyperspectral imagery,” Neurocomputing, Volume 149, Part B, 2015

[Zhihua20] Zhihua Xie, Yi Li, Jieyi Niu, Ling Shi, Zhipeng Wang, and Guoyu Lu, "Hyperspectral face recognition based on sparse spectral attention deep neural networks," Opt. Express 28, 36286-36303 (2020)

[Hiroyuki13] Hiroyuki Suzuki, Masamichi Suzuki, Takuya Urabe, Takashi Obi, Masahiro Yamaguchi, and Nagaaki Ohyama, "Secure biometric image sensor and authentication scheme based on compressed sensing," Appl. Opt. 52, 8161-8168 (2013)

[Lee20] Cheolsun Kim. Woong-Bi Lee, Soo Kyung Lee, Yong Tak Lee, and **Heung-No Lee**, “Fabrication of 2D thin-film filter-array for compressive sensing spectroscopy”, *Optics and Lasers in Engineering*, Vol. 115, pp. 53-58, Apr. 2019

[Lee13A] J. Oliver, WoongBi Lee, and **Heung-No Lee**, **“**Filters with random transmittance for improving resolution in filter-array-based spectrometers**,”** Optics Express, Vol. 21, No. 4, pp. 3969–3989, Feb. 2013

[Lee13B] Woong-Bi Lee, Cheolsun Kim, Gun Wu Ju, Yong Tak Lee, and **Heung-No Lee**, “Design of thin-film filters for resolution improvements in filter-array based spectrometers using DSP”, SPIE Defense + Commercial Sensing 2016, Baltimore, USA, Apr. 17-21, 2016

Seeing Through Computation and Artificial Intelligence

Heung-No Lee, Ph.D.

Professor, GIST, KOREA

2016.09.22

Agenda

- Seeing Through Computation 기술과 AI
- StC 기술
- 센서지능화 기술

강의 요지 및 방향

- Seeing through Computation (StC)?

- Artificial Intelligence



특 · 집 · 편 · 집 · 기

압축센싱: Seeing through Computation

이홍노 편집위원
(GIST 정보통신공학과)

지난 2011년 1월에는 승실대학교 신요안 교수님께서 편집자로 수고해 주신 압축센싱에 관한 대한전자공학회 특집호가 출판되었습니다. 거기에서는 압축센싱 기술의 기초이론과 통신 및 초음파 시스템 응용에 대한 연구 동향을 살펴보는 기회를 가졌습니다. 지난 특집호 이후 삼년 여가 지난 지금 어떤 새로운 연구가 진행되었는지 되돌아 보며, 압축센싱이라는 주제를 다시 한 번 다루어 보고자 합니다. 이렇게 압축센싱이라는 연구분야가 계속해서 여러 연구자들의 관심을 끄는 이유는 압축센싱 분야에서 오랫동안 사용되어 온 전통적 방식의 신호 획득 방법 및 그에 의한 시스템 디자인 방식을 무너뜨리는 전혀 새로운, 와해적인(interruptive) 시스템과, 혁신적(breakthrough)인 결과들이 속속 발표되고 있기 때문입니다.

그 예로서, 렌즈가 필요 없어서 어디든지 부착하고 사용할 수 있는 초소형 Lens Free 카메라, 위치 측위 정확도의 전통적 한계를 무너뜨리는 레이더, 전기 코일 동속에 앉아서 기다리는 시간을 10배 이상 획기적으로 줄인 고해상도 fMRI, Shannon-Nyquist 표본화와 같은 전자공학 상식을 뛰어넘는 Analog-to-Digital Converter, 기존의 방식으로는 상상할 수 없을 만큼의 넓은 주파수 대역을 모니터링 할 수 있는 인지무선 라디오의 스펙트럼 센싱 방법, 광 대역 동신 시스템에서 간섭을 제거하는 방법, 정확도를 크게 향상시킨 무선 채널 추정 방법 등을 들어볼 수 있습니다.

이번호에서는 이런 여러 가지 연구결과들을 Seeing through Computation이라는 부제를 붙이고 바라보고자 합니다. 압축센싱 이론을 응용측면에서 간략히 말하면, 신호측정은 간결히, 신호복원은 신중하고 정확하게 라고 할 수 있습니다. 후 신호 처리 Computation을 통해 복원을 잘 할 수 있으므로, 신호 측정을 하는 센서부분은 상대적으로 훨씬 간결하게 만들 수 있다는 의미입니다. 즉, 센서 숫자를 줄이거나 측정시간을 줄이거나 하는데 쓰일 수 있습니다. 혹은 센서 수나 신호 측정량을 고정시킬 경우, 센서 수 대비 정보 해상도 또는 정확도를 향상시킬 수 있다는 것입니다. 현미경, 영상장치, ADC, 스펙트럼 센싱, 무선 채널 추정, 간섭신호 추적 등은 사물을 보는 것, 즉 Seeing 을 의미합니다. 후 처리 연산을 통해 신호를 잘 복원할 준비가 되어 있으면, 즉 Computing 자원을 확보하고 있으면, 센서를 통해 사물을 보는 시스템을 간결하고 쉽게 만들 수 있다는 것입니다.

이번 특집호에서는 저의 연구실과 Seeing through Computation이라는 주제에 잘 맞는 연구를 진행해 오신 다섯 분의 연구자를 선별하여 초청하고 특집호를 구성 하였습니다. 먼저 최진호 교수님께서는 "Compressed Sensing Radar 연구 동향"을 통해 압축센싱의 레이더 시스템 적용 및 레이더의 분해능 증대 방안 연구 동향을 살펴보았습니다. 전병우 교수님께서는 "압축센싱과 영상처리 응용"의 제목으로, 단일 화소 카메라, Lens Free 카메라 등 압축센싱의 영상처리 응용 최신 동향을 살펴봐 주셨습니다. 황도식 교수님께서는 "다중변수 압축 MRI 기술 연구 동향"의 제목으로 다중변수 압축 MRI의 기본 원리와 관련 분야 연구 동향을 살펴 주셨습니다. 신요안 교수님께서는 "무선통신에서의 압축센싱 응용" 이란 제목의 논문에서, 압축센싱이 무선 통신에서 Analog-to-Digital Conversion, 무선 채널 추정, 인지 라이오 스펙트럼 센싱, 무선랜 다중 표적 무선 측위 등에서 정확도를 높일 수 있는 데 쓰일 수 있음을 보여주었습니다. 예종철 교수님께서는 "신호처리를 이용한 초고해상도 형광 현미경"이라는 제목의 논문에서 압축센싱이 형광 현미경의 해상도 증진에 쓰일수 있음을 보여주었습니다. 마지막으로 저의 연구실에서 광 대역 통신 시스템의 협 대역 간섭 제거를 위한 압축센싱 기술 연구 동향을 살펴 보았습니다.

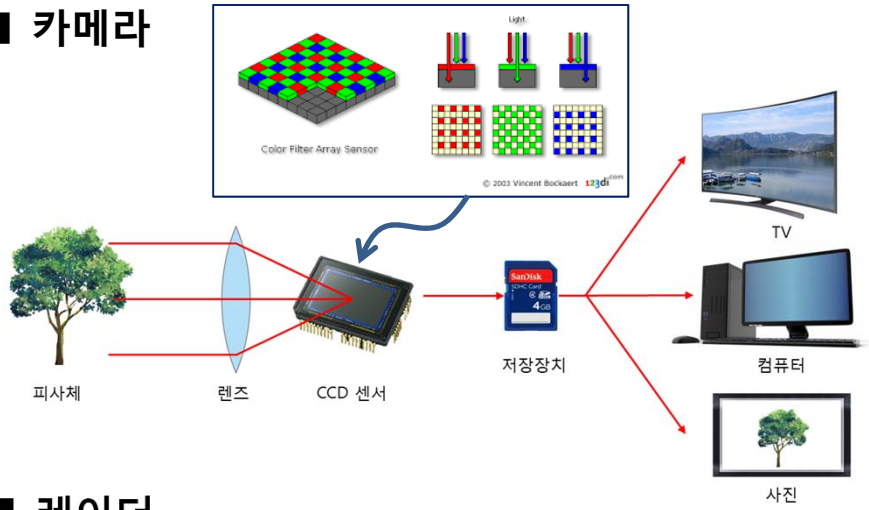
이 자리를 빌어 연구에 전념하느라 바쁘신 중에도 저의 원고 초대에 흔쾌히 응해 주시고, 촉박한 일정에도 불구하고, 기한 내에 원고를 마감해주신 다섯 분의 교수님들과, 심의와 조언을 아끼지 않아 주신 강의성 학회지 편집위원장님과 편집위원님들께 감사의 말씀을 드립니다.

Seeing through Computation (StC)

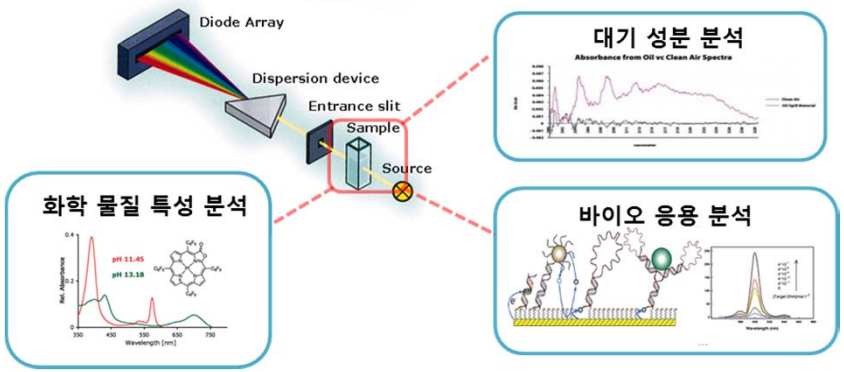
■ **분광기, 현미경, 카메라, 레이다 등 영상 센서의 성능을 크게 제고하는 Seeing thru Computation (StC) 기술을 개발하고, 시작품 제작을 통해 이론의 실제성을 입증 함.**

- (전기전자컴퓨터공학부 이흥노 교수)
- StC 기술은 센서를 지능화 하는 신호 처리 수학을 중심으로 한 Software 시스템 기술 임.
- StC를 통해 Hardware의 성능이 극대화되기 때문에, Lensfree현미경, 곤충눈 카메라, 혹은 Grating없는 분광기 등 HW 구조가 매우 간략하면서도, 높은 성능을 유지 할 수 있는 매우 혁신적인 제품의 출현이 가능케 됨.

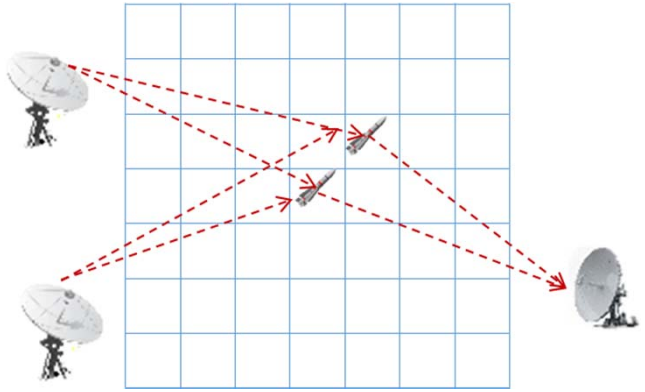
■ 카메라



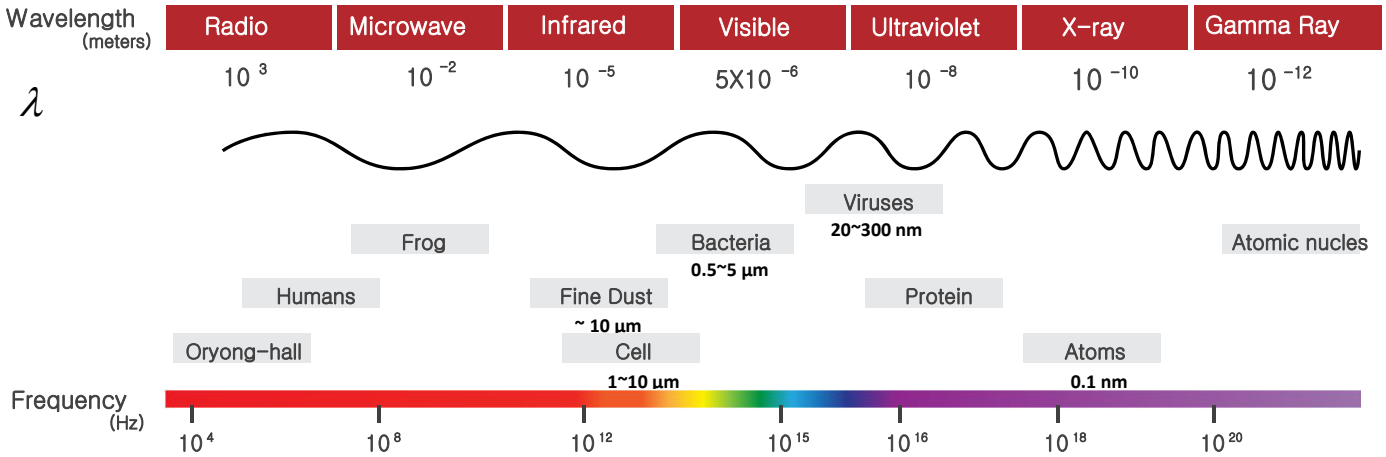
■ 분광기



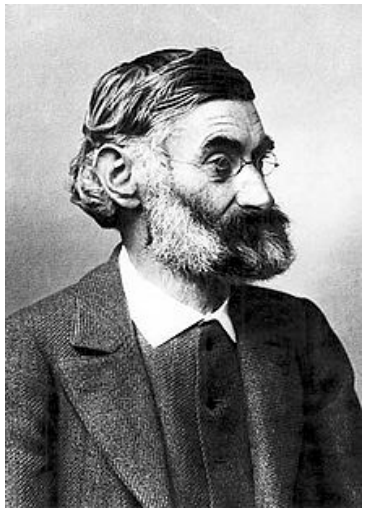
■ 레이더



빛의 성질



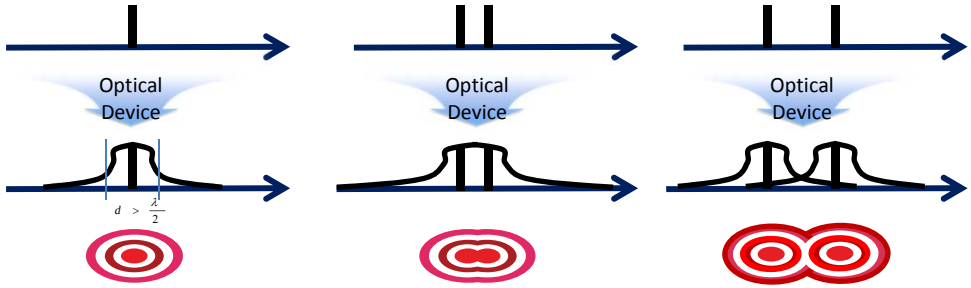
Ernst Abbe (1840 ~ 1905)





Diffraction Limit

$$d > \frac{\lambda}{2}$$



수학을 이용한 해상도 제고 방법

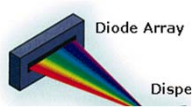
Interference free

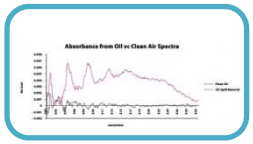
$$\begin{pmatrix} 0 \\ 0 \\ 1.0 \\ 1.0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1.0 & 0 & 0 & & & \\ 0 & 1.0 & 0 & 0 & & \\ & 0 & 1.0 & 0 & 0 & \\ & & 0 & 1.0 & 0 & 0 \\ & & & 0 & 1.0 & 0 \\ & & & & 0 & 1.0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1.0 \\ 1.0 \\ 0 \\ 0 \end{pmatrix}$$



$$\mathbf{y} = \mathbf{I}\,\mathbf{x}$$







Interfered optics

$$\begin{pmatrix} 0.1 \\ 0.9 \\ 1.8 \\ 1.8 \\ 0.9 \\ 0.1 \end{pmatrix} = \begin{pmatrix} 1.0 & 0.8 & 0.1 & & & \\ 0.8 & 1.0 & 0.8 & 0.1 & & \\ 0.1 & 0.8 & 1.0 & 0.8 & 0.1 & \\ & 0.1 & 0.8 & 1.0 & 0.8 & 0.1 \\ & & 0.1 & 0.8 & 0.1 & 0.8 \\ & & & 0.1 & 0.8 & 0.1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1.0 \\ 1.0 \\ 0 \\ 0 \end{pmatrix}$$

$$\mathbf{y}_1 = \mathbf{A} \times \mathbf{x}$$

$$\min \|\mathbf{x}\|_2 \quad \text{s. t.} \quad \mathbf{y}_1 = \mathbf{A}\mathbf{x}$$

$$\Rightarrow \quad \tilde{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\,\mathbf{y}_1$$

수학적 해상도 제고 방법: Super Resolution

DSP 사용

$$\begin{pmatrix}0.1\\0.9\\1.8\\1.8\\0.9\\0.1\end{pmatrix}=\begin{bmatrix}1.0&0.9&0.8&0.3&0.1&&&&&&&\\0.8&0.9&1.0&0.9&0.8&0.3&0.1&&&&&\\0.1&0.3&0.8&0.9&1.0&0.9&0.8&0.3&0.1&&&\\&&0.1&0.3&0.8&0.9&1.0&0.9&0.8&0.3&0.1&\\&&&&0.1&0.3&0.8&0.9&1.0&0.9&0.8&0.3\\&&&&&&0.1&0.3&0.8&0.9&1.0&0.9\end{bmatrix}\begin{bmatrix}0\\0\\0\\0\\1.0\\0\\1.0\\0\\0\\0\\0\\0\end{bmatrix}$$

$$\mathbf{y}_2 = \mathbf{A}_2\mathbf{x}$$

$$\tilde{\mathbf{x}} = \arg\min \|\mathbf{x}\|_1 \quad \text{s. t.} \quad \mathbf{y}_2 = \mathbf{A}_2\mathbf{x}$$

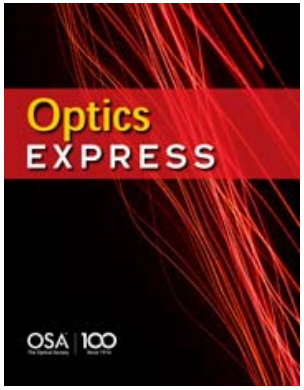
랜덤 필터 및 DSP 사용

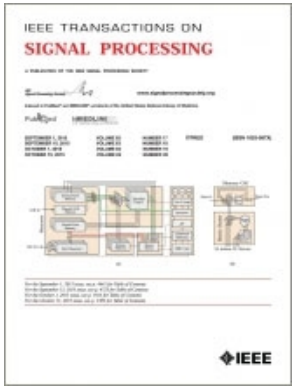
$$\begin{pmatrix}0.3\\0.9\\0.9\\1.2\\0.7\\1.8\end{pmatrix}=\begin{bmatrix}0.6&0.4&0.2&0.3&0.1&0.5&0.2&0.9&0.2&0.6&0.3&0.1\\0.9&0.7&0.9&0.4&0.3&0.1&0.6&0.1&0.5&0.6&0.7&0.6\\0.8&0.9&0.7&0.5&0.8&0.3&0.1&0.3&1.0&0.2&0.2&0.7\\0.3&0.8&0.3&0.1&0.9&0.8&0.3&0.1&0.6&0.6&1.0&0.3\\0.1&0.3&0.8&0.9&0.2&0.9&0.5&0.3&0.1&0.1&0.3&0.9\\0.6&0.1&0.3&0.2&0.9&0.4&0.9&0.8&0.3&0.1&0.6&0.4\end{bmatrix}\begin{bmatrix}0\\0\\0\\0\\1.0\\0\\1.0\\0\\0\\0\\0\\0\end{bmatrix}$$

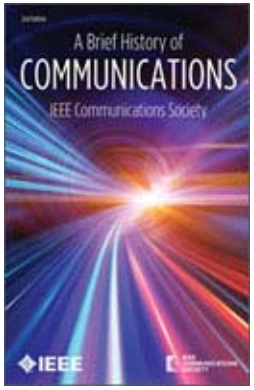
$$\mathbf{y}_3 = \mathbf{A}_3\mathbf{x}$$

$$\tilde{\mathbf{x}} = \arg\min \|\mathbf{x}\|_1 \quad \text{s. t.} \quad \mathbf{y}_3 = \mathbf{A}_3\mathbf{x}$$

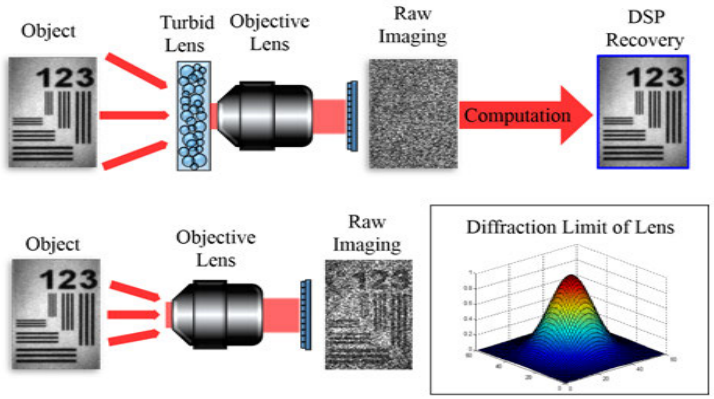
최근 4년 간 성과



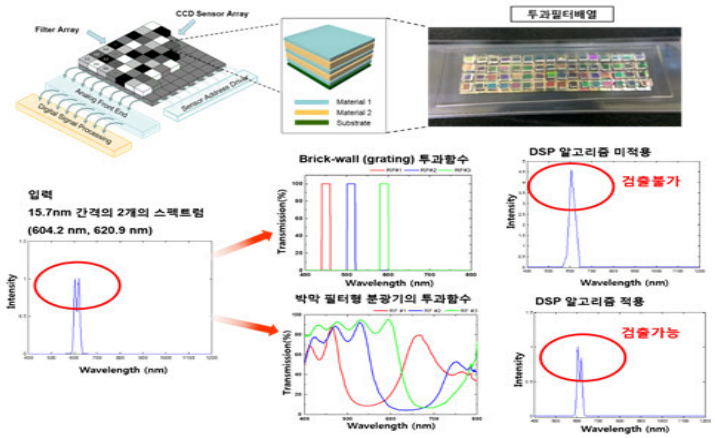




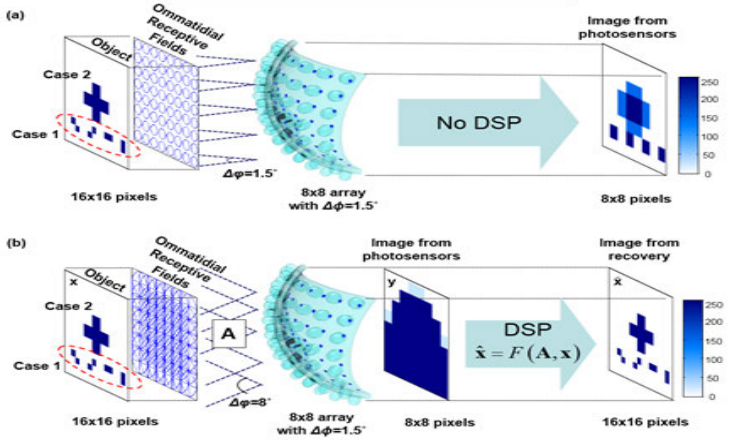
피부 속을 보는 현미경



Grating free 초소형 분광기



탈 부착형 곤충눈



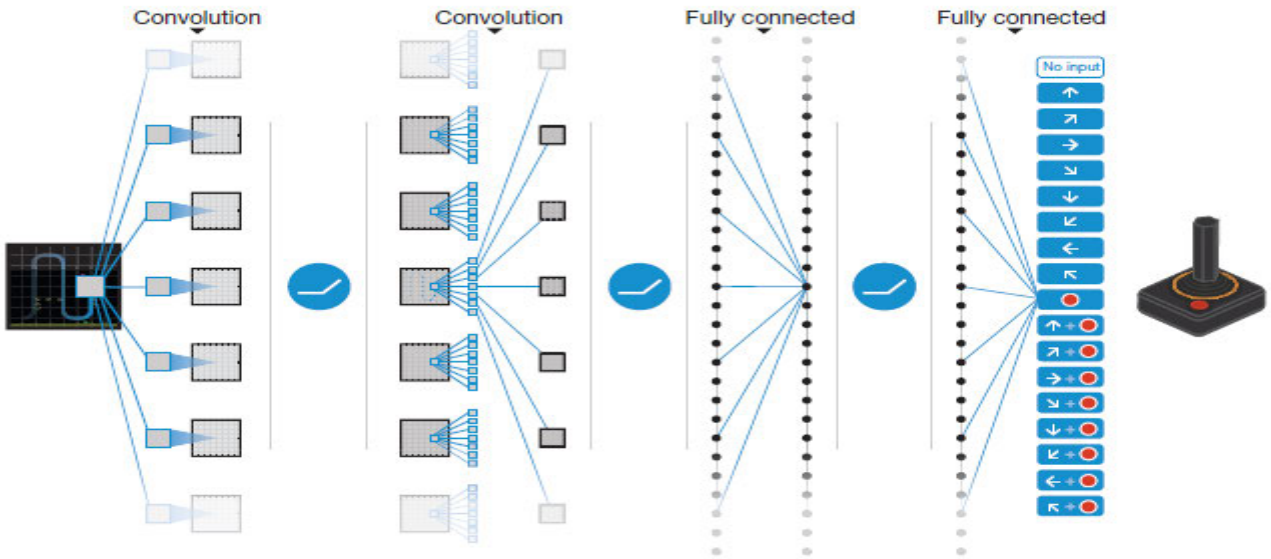
Artificial Intelligence

- 지능이란?

  - 주어진 상황을 인지하고 목표를 효과적으로 달성하는 결정, 임무 수행 (한정된 자원하에)

- Machine learning
  - Supervised Learning
  - Unsupervised Learning
  - Reinforcement Learning (RL)
  - Deep Neural Network

AlphaGo Deep Neural Network

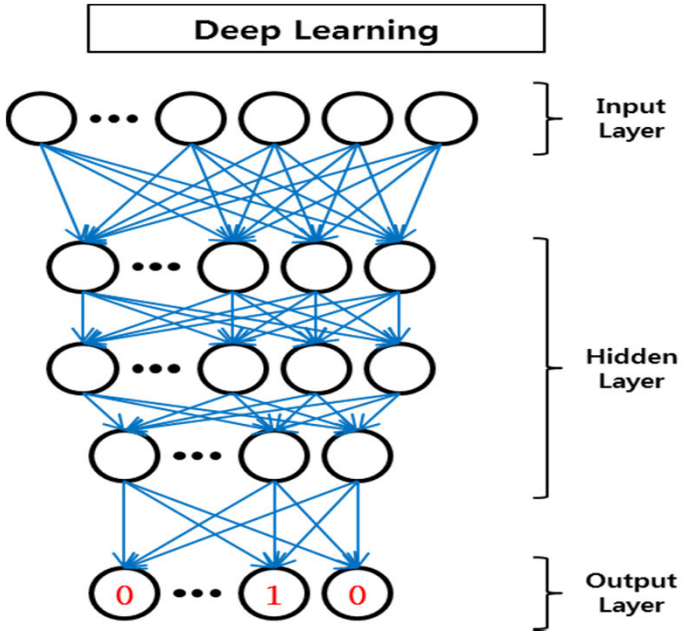


상황인지: Input: Screen
Output: Joystick Action

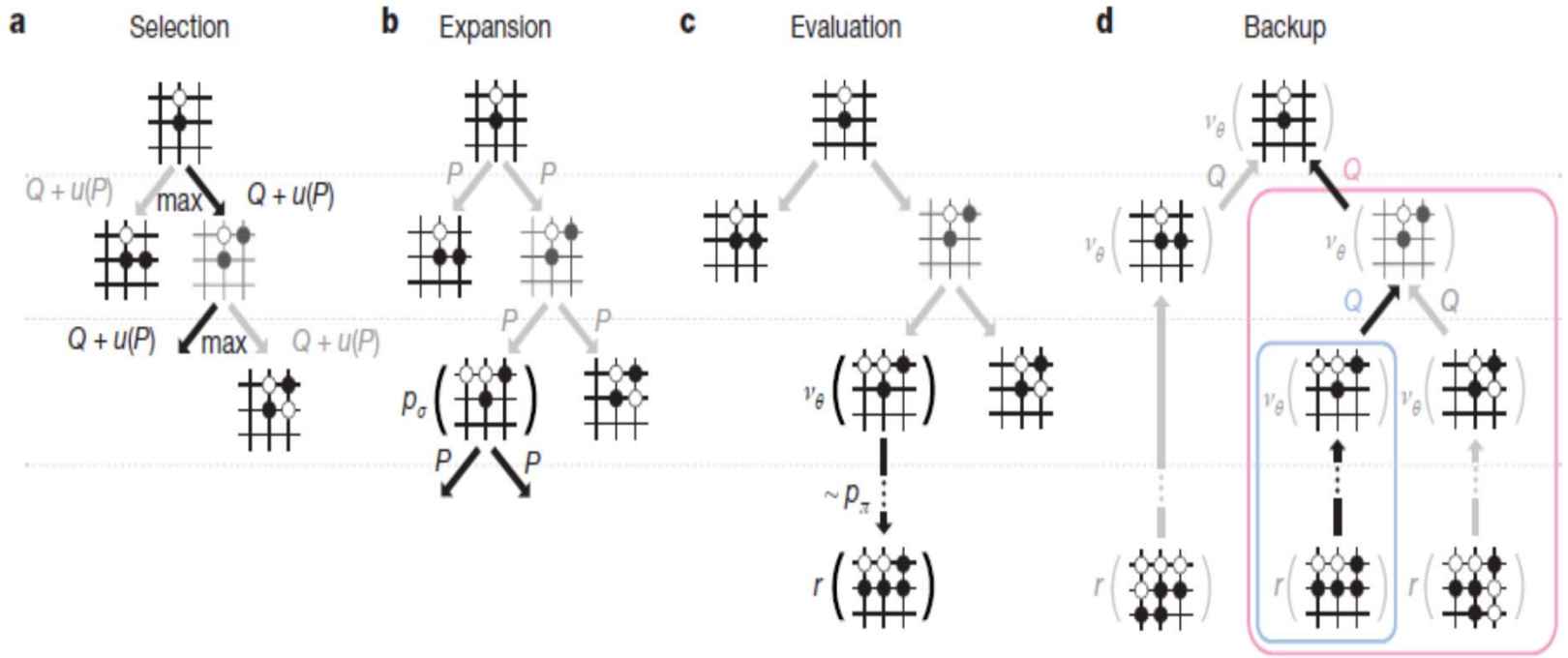
Deep Neural Network

- **Deep Neural Network (DNN)**
  - 여러 개의 Perceptron들이 모여 다수의 입/출력 node를 지니는 layer를 구성함.
  - DNN은 Input Layer와 Output Layer와 함께 하나 이상의 Hidden Layer을 지니는 인공 신경망(ANN)을 의미함.
  - Training signal 에 대한 Backpropagation algorithm을 통해 Weight optimization을 수행함.

- **Beginning of DNN**
  - DNN은 1980년에 소개된 Neocognitron 로부터 시작됨 [Fukushima 80].
  - 초기에는 시간 복잡도, Local Minimum 수렴, Overfitting 문제 등 여러 기술 장벽으로 인해 기술의 상용화가 힘들었음.
    - E.g. 10개의 숫자를 인식하기 위해 Training하는 과정이 3일 걸림 [Jackel 89].



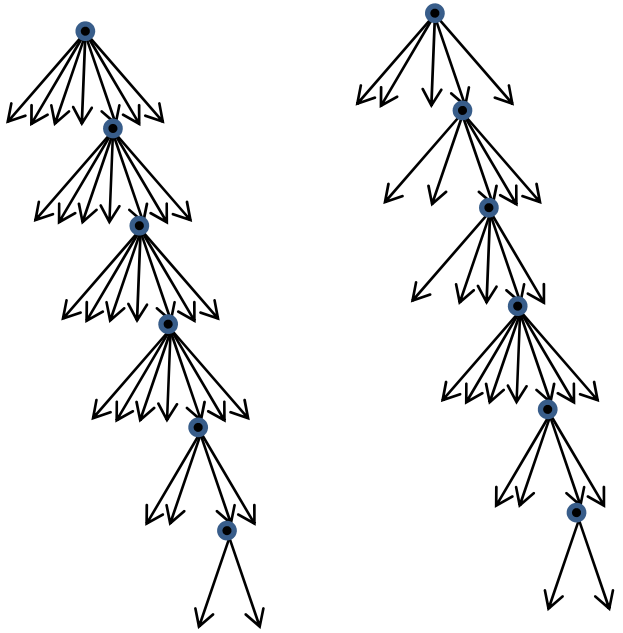
AlphaGo Tree Search



AlphaGo

- Tree Search
- Supervised (human players), Reinforced learning (Self-play 학습)
- Trim the possibilities with Deep Neural Network (총 action의 가지 수를 줄임)
- Breadth Search
- Depth Search
- Aim to make the best decision among sequences of actions, rather than single actions

RL 및 기보 학습의 결과



정보이론/신호처리이론

- MIMO Channels
- Sphere-decoding
- Viterbi decoding
- Tree-search
- Tree pruning!
- Aim to make decision on the sequence of transmitted symbols, rather than on single symbol transmission

IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, VOL. 23, NO. 5, MAY 2005 1013

Robust Iterative Tree-Pruning Detection and LDPCC Decoding

Heung-No Lee, *Member, IEEE,* and Xinde Hu, *Student Member, IEEE*

***Abstract*—A novel suboptimal low-complexity equalization and turbo-iterative decoding scheme is proposed in this paper. The scheme is developed for multiple transmit- and multiple receive-antenna systems operating over severe *frequency-selective* fading *intersymbol interference* channels. The signal-processing complexity may be of a concern for such systems. The complexity of a full-search equalization grows in a power-law manner $O(M^{N_t L})$, where M denotes for M-ary channel symbols, N_t the number of transmit-antennas, and L the number of delay channel taps. A low-complexity solution can be obtained by pruning an equalizer tree. The two main operations include a *sphere list detection* and a *threshold-based tree-search*. In the operation of extracting extrinsic messages from the pruned tree, a set of explored paths with different survival lengths poses a fairness problem: a longer-lived path naturally builds a larger discrepancy-metric than a shorter lived path does. A novel survival-length *compensation-rule* is devised so that all explored paths with different survival lengths are utilized fairly in generating the output message. Simulation results are obtained for multi-input and multi-output systems equipped with four transmit and four receive antennas. They indicate the performance of the receiver is very robust.**

***Index Terms*—Joint equalization and decoding, low-density parity-check (LDPC) codes, maximum *a posteriori* (MAP), multi-input and multi-output (MIMO) systems, reduced complexity receiver, turbo-iteration.**

single-input–single-output techniques [4]. In addition, this improvement is intended for dynamic urban nonline-of-sight multipath channel conditions, where the networked forces are frequently deployed for carrying out tactical missions. Among many design challenges for MIMO systems operating in such environment, the most notable issue relevant to this paper is how to design computationally efficient MIMO signal processing algorithms while maintaining a robust receiver performance. The appropriate size of an antenna-array is considered to be four transmit and four receive antennas for a vehicle mounted system.

The receiver studied in this paper uses a joint turbo-iterative equalization and decoding scheme with a low-density parity-check code (LDPCC) as the coding scheme. The turbo-iterative receiver paired with the powerful outer block code at the transmitter for robust performance has a high potential for achieving the optimal capacity-diversity tradeoff performance. Several turbo-like code-based space–time coded modulation systems have been reported in the literature and their performances have been found to be robust [5]–[8]. However, for higher spectral-efficiency (obtained by increasing the constellation size) and for robust performance over intersymbol

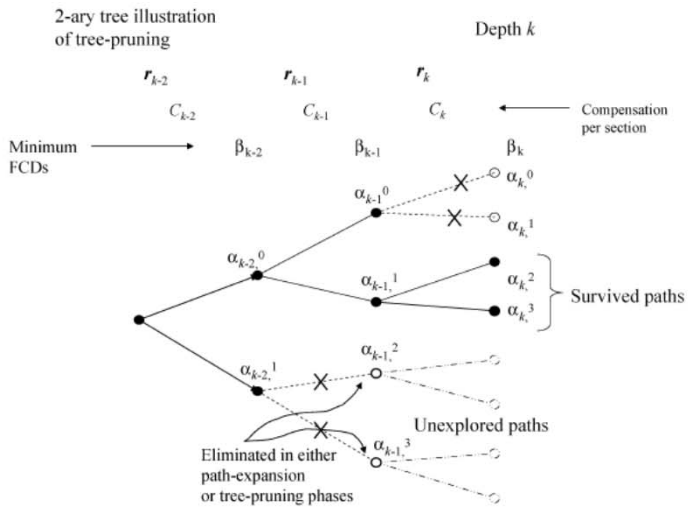
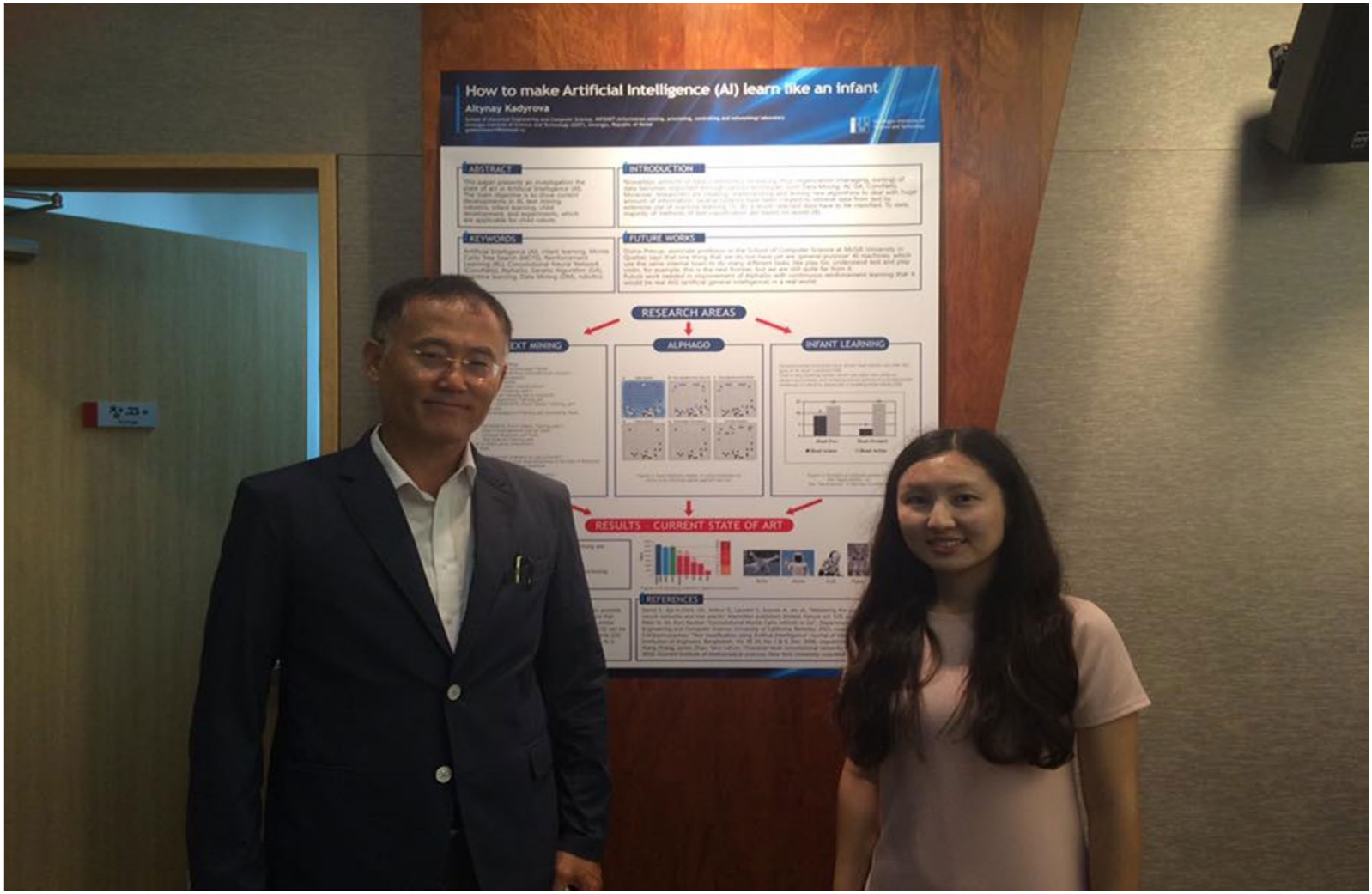


Fig. 2. Illustration of tree-pruning with $(Q = 2)$ binary tree. For MIMO settings, the tree is $Q = M^{N_t}$-ary.

Seeing Through Computation기술은

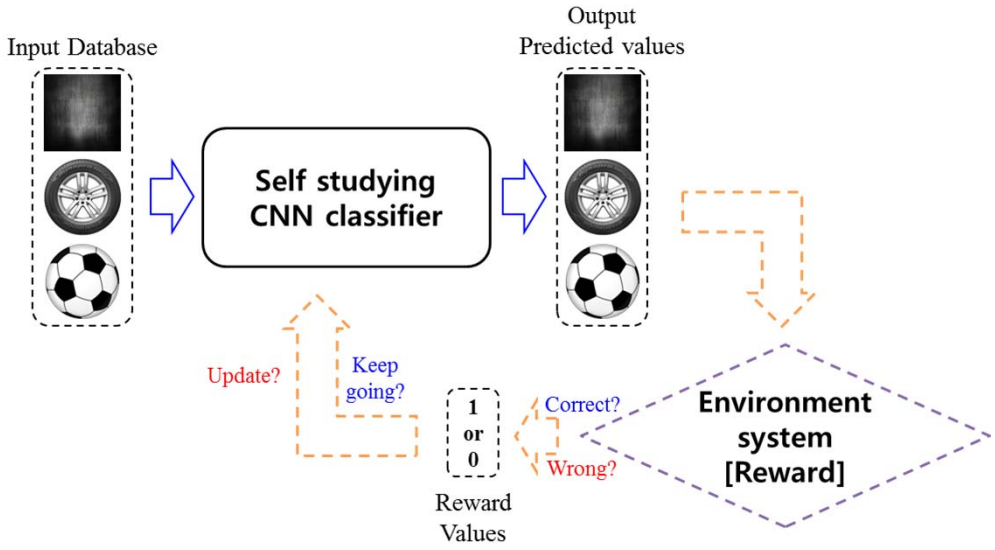
- 눈, 코, 귀, 촉각, 뇌파 등을 측정하는 센서를 DSP기술을 통해 그 성능을 향상 시키는 것에 초점을 맞추어 왔읍니다.
- 최근 인공지능의 발전은, 시스템을 한 번에 최적화 하려 하기 보다는, 스스로 학습할 수 있게 만들어 스스로 상황에 맞추어 개선할 수 있게 만드는 것이 더 좋다는 것을 시사 합니다.
- 이에, 센서가 스스로 학습을 통하여 성능을 개선 시키는 시스템으로 StC기술을 개발 하고자 목표 합니다.

How to make sensors learn like an infant?



Self Learning CNN

- Reward Function을 통해 스스로 학습하는 CNN
- Self studying classifier :
  'Scene → Action'에 해당하는 CNN 함수
- Environment system :
  'Action → Reward'에 해당하는 환경 모델 함수
- Action에 따른 Reward 값을 모니터링하고, 결과에 따라 CNN 재학습 수행
  - Step 1 : 변경된 Input population 입력
  - Step 2 : Reward 값 하락
  - Step 3 : Input population 변화 감지
  - Step 4 : CNN 재학습 결정
- Input population이 변경 되는 application에 적용 가능
- System Memory에 In/Output 저장 및 CNN 재학습



Self-Studying CNN machine

Self Learning CNN

Input Population

<Phase 1>

$P(\text{⚽}) = \frac{1}{2}$

$P(\text{▪}) = \frac{1}{2}$

<Phase 2>

$P(\text{⚽}) = \frac{1}{2}$

$P(\text{▪}) = \frac{1}{4}$

$P(\text{⚙}) = \frac{1}{4}$

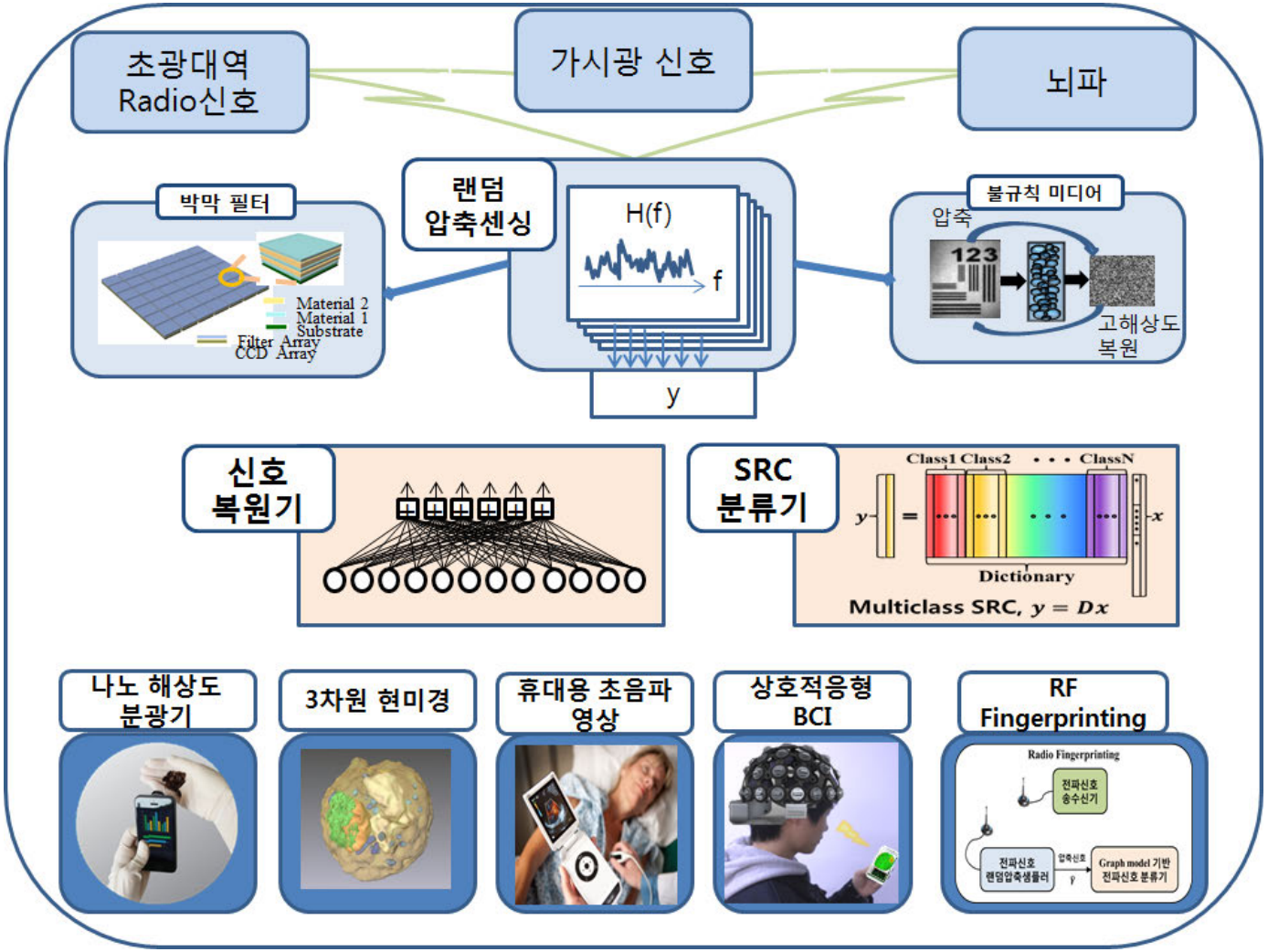
Input	Output	Reward	Optimized Weight
⚽	⚽	1	W_1
▪	▪	1	
⋮			
⚙	⚽	-1	$W_1 \rightarrow W_2$
⚽	⚽	1	
▪	▪	1	
⋮			

Expected Reward values (Calculation)

$$\text{Re}ward = E(R(output|input))$$
$$= P(\text{⚽}) \times R(\text{⚽}|\text{⚽}) + P(\text{▪}) \times R(\text{▪}|\text{▪})$$
$$= \frac{1}{2} \times 1 + \frac{1}{2} \times 1 = 1$$

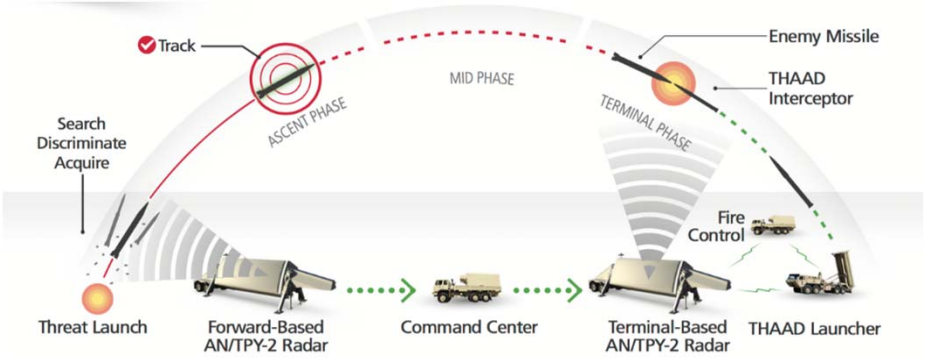
$$\text{Re}ward = E(R(output|input)$$
$$= P(\text{⚽}) \times R(\text{⚽}|\text{⚽}) + P(\text{▪}) \times R(\text{▪}|\text{▪})$$
$$+P(\text{⚙}) \times R(\text{⚙}|\text{⚽})$$
$$= \frac{1}{2} \times 1 + \frac{1}{4} \times 1 + \frac{1}{4} \times (-1) = \frac{1}{2}$$

Focus Areas in this talk

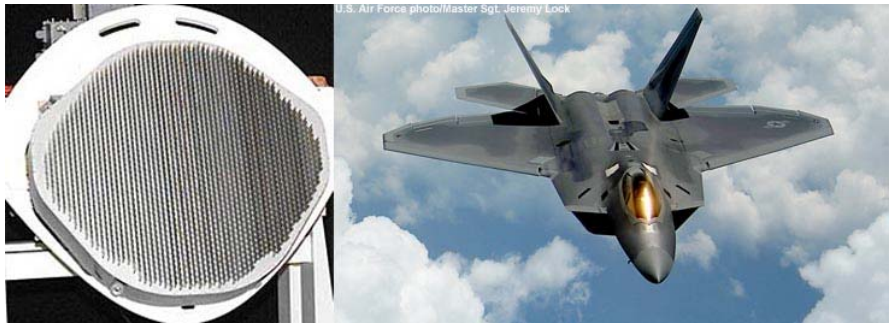


미래 응용 분야

■ Thaad Radar



■ Raptor 22



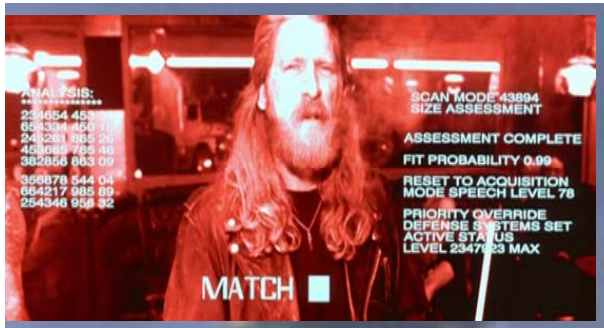
■ 분광기

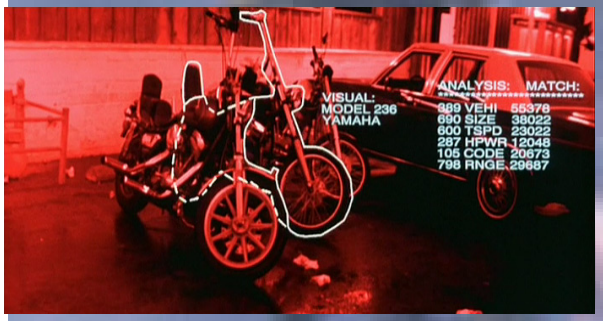




■ 360 VR




SCAN MODE 43894
SIZE ASSESSMENT
ASSESSMENT COMPLETE
FIT PROBABILITY 0.99
RESET TO ACQUISITION
MODE SPEECH LEVEL 78
PRIORITY OVERRIDE
DEFENSE SYSTEMS SET
ACTIVE STATUS
MATCH


VISUAL:
MODEL 236
YAMAHA
ANALYSIS: MATCH:
689 VEHI 55378
690 SIZE 39022
600 TSPD 23022
287 HPWR 12048
105 CODE 20673
798 RNGE 29687


VISUAL ANALYSIS
HOMETOWN: HARIPUR
PAKISTAN
SINGING: TRUE
GENDER: MALE
BIRTHDAY: 07/16
LOCATION: NEW YORK,
NEW YORK
CLAS
HARIS AMIN









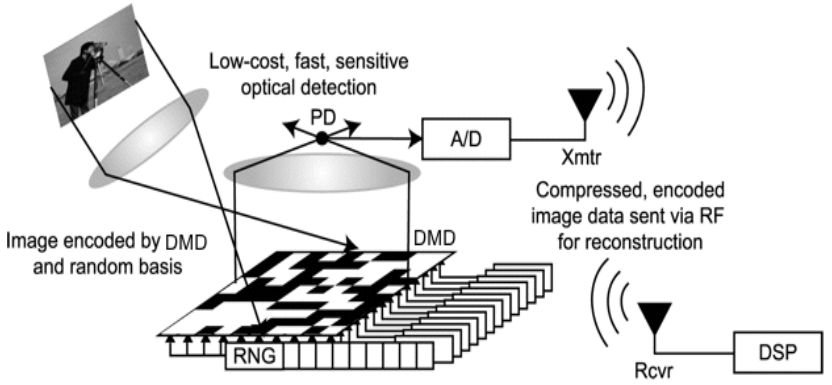





미세먼지 크기 비교
2.5㎛ 이하
10㎛
50~70㎛
초미세먼지
(PM 2.5)
미세먼지 · 황사
(PM10)
머리카락

Background

- Compressed sensing (CS)
  - New signal acquisition techniques [Donoho06], cited >4000 times.
  - MIT 2007 Tech Review, “Top 10 Emerging Technologies”
- CS is to find sparse solution from an under-determined linear system.
  - Real, complex field
- Many application areas: Cameras, Medical Scanners, ADCs, Radars, …



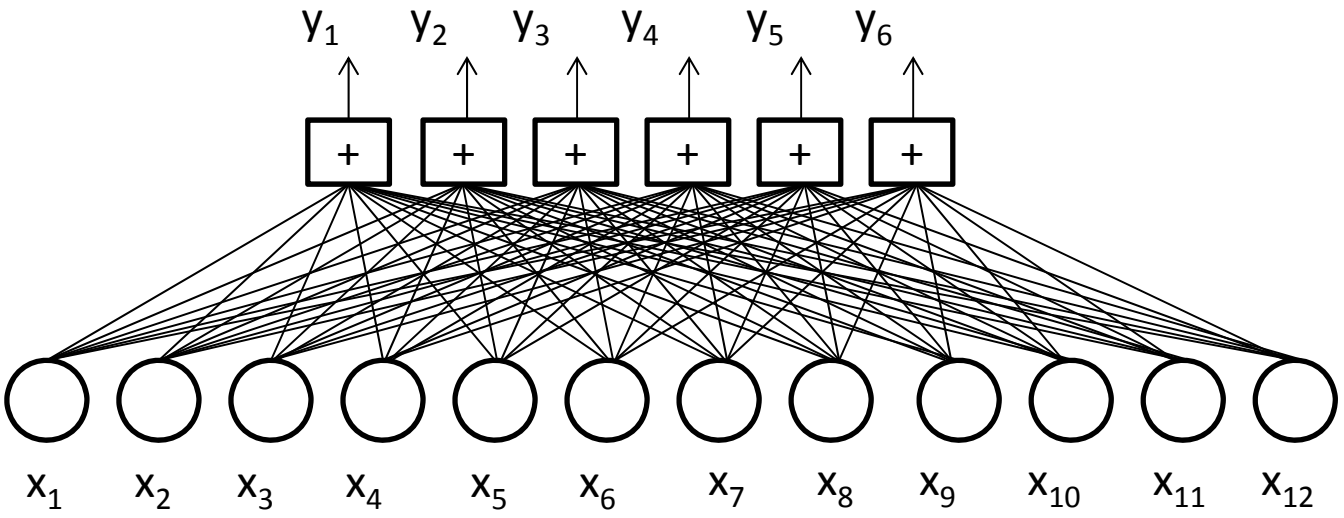
$$M < N$$

$$\begin{bmatrix} y \end{bmatrix} = \begin{bmatrix} F \end{bmatrix} \begin{bmatrix} x \end{bmatrix}$$

$$y \in \mathbb{R}^M$$
$$F \in \mathbb{R}^{M \times N}$$
$$x \in \mathbb{R}^N$$

$$\hat{x} = \arg\min_x \|x\|_{\ell_1} \quad \text{subject to } Fx = y$$

Have you seen this before?



- We can draw a graph for y = Fx, shown above.
- Yes!!!
- Channel codes—Syndrome decoding!

- Two applications:
  - Compressed measurement **y** and recovery of uncompressed data **x**
  - Super-resolution of hidden **x** from limited measurement **y**

전세계적 연구 동향

- (압축센싱이란) 2006년 정보이론과 신호처리 분야에 소개된 압축센싱이론은 응용 측면에서 **한 마디로 요약하면 '영상 및 음성 등의 자연신호를 압축적으로 센싱(Encoding) 할 수 있으며**, 그렇게 했을 때 **적은 량의 측정샘플 만으로도 신호를 복원(Decoding)할 수 있다'** 정도로 표현 할 수 있음.
  - 2008 Rice대학 Single-Pixel Camera, 단 한 개의 포토센서로 영상을 압축하고 복원할 수 있음 시연하고 이론의 실제성 입증.

- 압축적 샘플링(Encoding)과 신호 복원(Decoding) 문제는 하나의 간단한 연립방정식 $y = Ax$ 으로 표현 가능함.
  - $A \sim M \times N$ 측정행렬이라 칭함.
  - $M<N$ 일 때 압축이 일어남.

- **신호복원**은 측정행렬 A를 알고 있는 상황에서, y 로부터 x를 찾아내는, 즉 **역문제를 푸는 문제임**.

- 이 역문제가 잘 풀리는 경우는?

혁신 기술 출현

- (혁신기술 출현) 압축센싱은 또한 "**적은 수의 센서로도**" 혹은 "**짧은 시간 동안의 측정으로도**" 신호를 **높은 해상도로 복원**하는데에도 응용 가능함이 밝혀졌음.

- 레이더, fMRI, 현미경등 영상장치들은 **신호가 센서에 도달하기까지 어떻게 변화하는지를 나타내는 광물리학적 전달 함수가** 잘 알려져 있는데, 이 점을 사용하여 **Decoding에 필요한 측정행렬 *A*를 만들어내고**, Diffraction Limit등을 돌파하는 **초고해상도 신호복원이 가능함**이 알려짐.

- 이런 응용연구 결과의 발표는, 2010년 전후로부터 현재까지 지속적으로 이루어지고 있으며, 초 고해상도 레이더, 초 광대역 신호 측정, 촬영시간이 크게 단축된 fMRI 머신, 초 고해상도 현미경 등의 **혁신 기술의 출현을 가능케** 하고 있음.

Sparse Representation

- 압축센싱분야에서 개발된 Decoder 알고리즘은 또한 **신호를 희소하게 표현 하는 방법** (Sparse Representation) **으로도 응용되어 왔음**.
- 즉, 신호의 고유정보를 잃지 않고 신호의 **Dimensionality를 크게 Reduction 하는 방법**으로 쓰이고 있음.
- 나아가, 신호를 알려진 몇 가지의 **클래스로 구분하는데 뛰어난 성능**을 보임이 보고됨.
- CCTV 카메라에 찍힌 사람의 이상 행동 양식 구분, 사람의 얼굴 검색 등 Big Data 처리 및 분석 등에도 Sparse Representation Classification이라는 이름으로 응용되고 있음.
- 한 예로, 사용자의 뇌파 신호를 구분하여 컴퓨터 명령을 내릴 수 있는 Brain-Computer Interface 용 분류알고리즘을 들 수 있음.

Encoding 연구의 혁신성

- 압축센싱의 **Encoding 문제는 아날로그 신호 $x(t)$를 아날로그 도메인 상에서 압축적으로 샘플링하여 디지털 샘플 벡터 y를 얻는 전기/전자/재료 공학적 장치를 만들어 내야하는 문제임**.

- 다시 말해, 압축센싱이론 논문에서 흔히 다루는, 이미 디지털 샘플링이 완료된 신호 벡터 x를 측정행렬 A와 곱하여 y를 얻는 것은 Encoding 문제가 아님.

- 아날로그 도메인에서 연속적인 신호 $x(t)$를 랜덤한 반응 패턴을 가진 함수와 곱하고, 그 결과를 더해서 샘플을 취하는 장치를 만드는 연구 및 실제 구현이 혁신적 성과를 내기 위해 꼭 필요함.

Decoding 연구 지속의 중요성

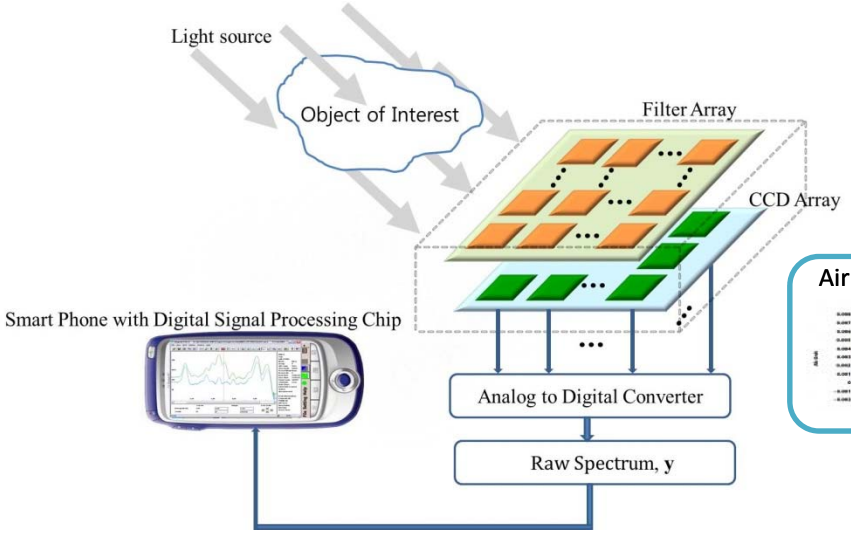
- 많은 Decoder가 현재 개발 되었음
  - Gaussian, FFT, Bernoulli 센싱등 문제가 수학적으로 잘 알려진 형태 일 때

- 특정 문제에 특화된 Decoder 개발 필요 함
  - Photonics, Spectrometers, Electronic Eyes, …
  - Non negative signal sensing
  - No general sensing matrices

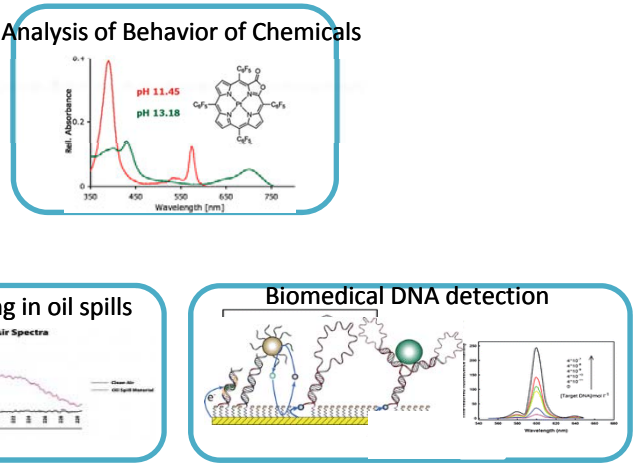
Compressive Spectrometers for Super Resolution

- Spectrometer: Used to find the spectrum of an optical signal
- It takes in the light, breaks it into its spectral components, and displays them in a portable device such as smart phones.

State-of-the art portable spectrometers



Applications



- The ability of the spectrometer in revealing fine information is determined by its "Resolution"
- Problem: Resolution is limited by the number of filters

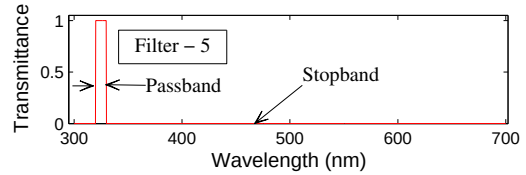
Compressive spectrometers Contd.

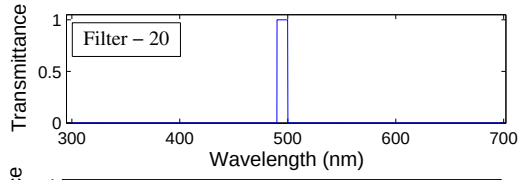
- How to improve the resolution for a fixed set of filters in a spectrometer?
- Solution: Compressive Spectrometers!!!
- Innovations
  - Spectrum acquisition using random filters design (using thin-films) [Lee13s]
    - Analog domain acquisition (Our design is first of a kind)
    - A set of *M* filers with good auto and cross covariance
  - Spectrum reconstruction using a new L_1 algorithm [Lee12s]
- We model the spectrometer output as underdetermined linear system $\mathbf{y}=D\mathbf{x}$
- The matrix *D* is transmittance matrix is determined from the filter manufact uring process
  - MEMS, non-ideal filters (conventional)
  - Thin-film, random filters (our method)

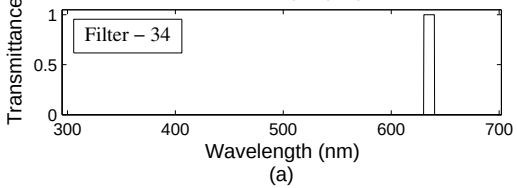
Various Transmittance Functions

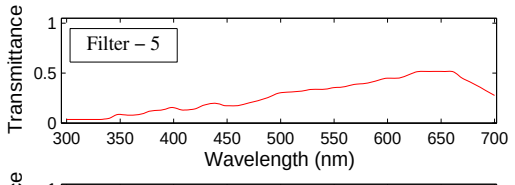
Conventional Design

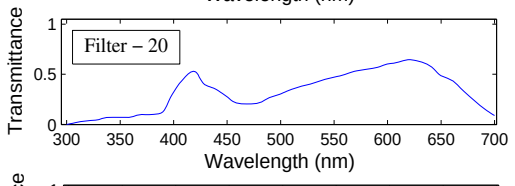
1. Stringent filter design
2. Local sampling

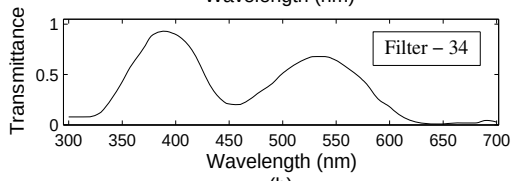






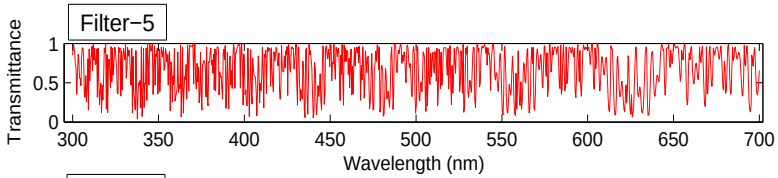


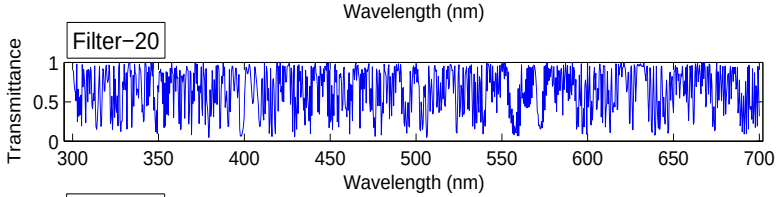


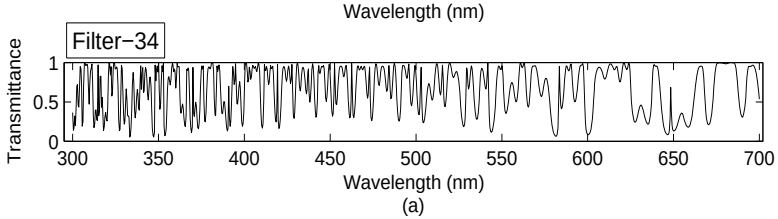


Our approach

1. Ease of filter design
2. Holistic sampling





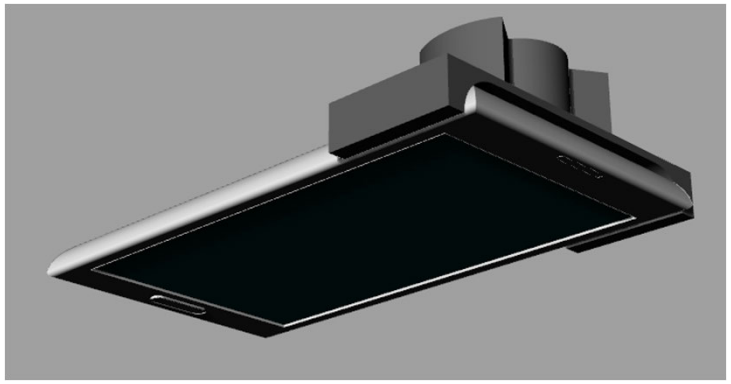


Summary

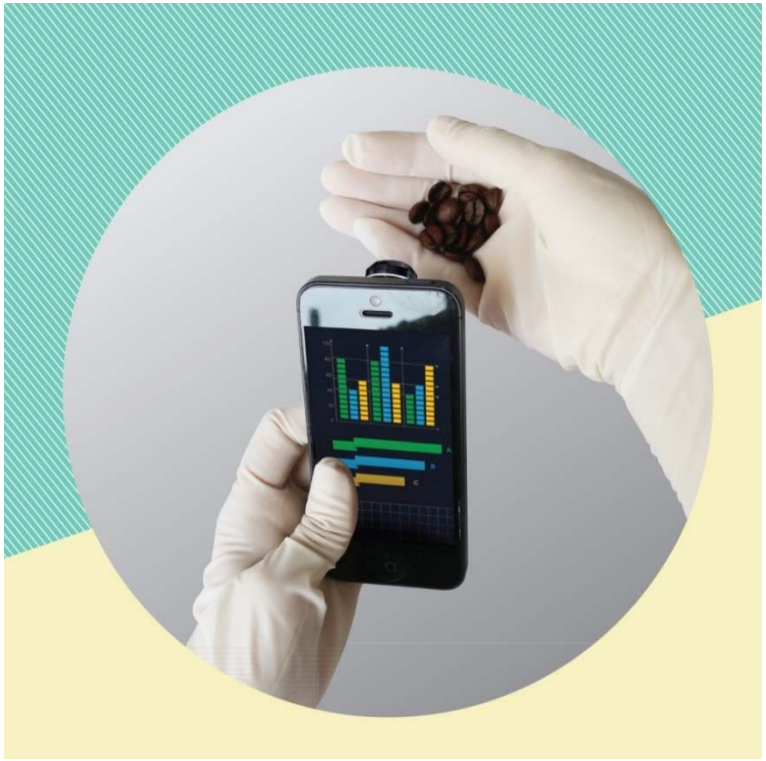
- 신호처리로 hardware의 단점을 보완하는 시스템 개발 Needs가 증가 함.
- 좋은 신호 처리 알고리즘을 사용하여 센서 및 센서 시스템을 간단하면서도 정확하게 작동할 수 있게 할 수 있음.
- Lensfree camera, single pixel camera등의 application 등 이 있음.

Future works

- Implementing random filters with thin-film technology varying thickness and reflective indices
- Ultimate Goal: Smartphone attachable high resolution spectrometers and microscopes



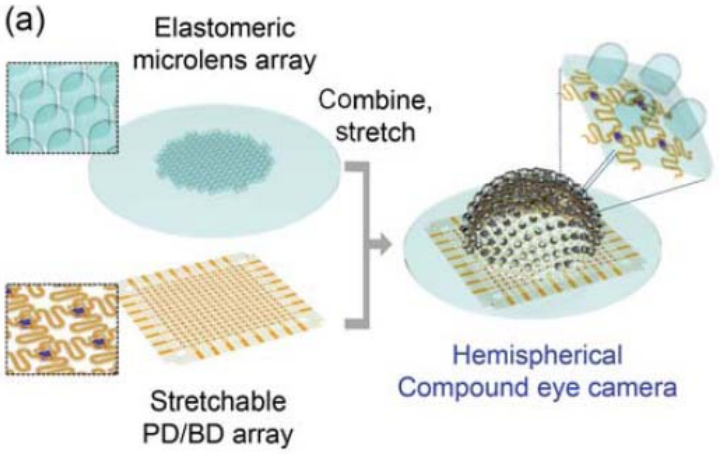


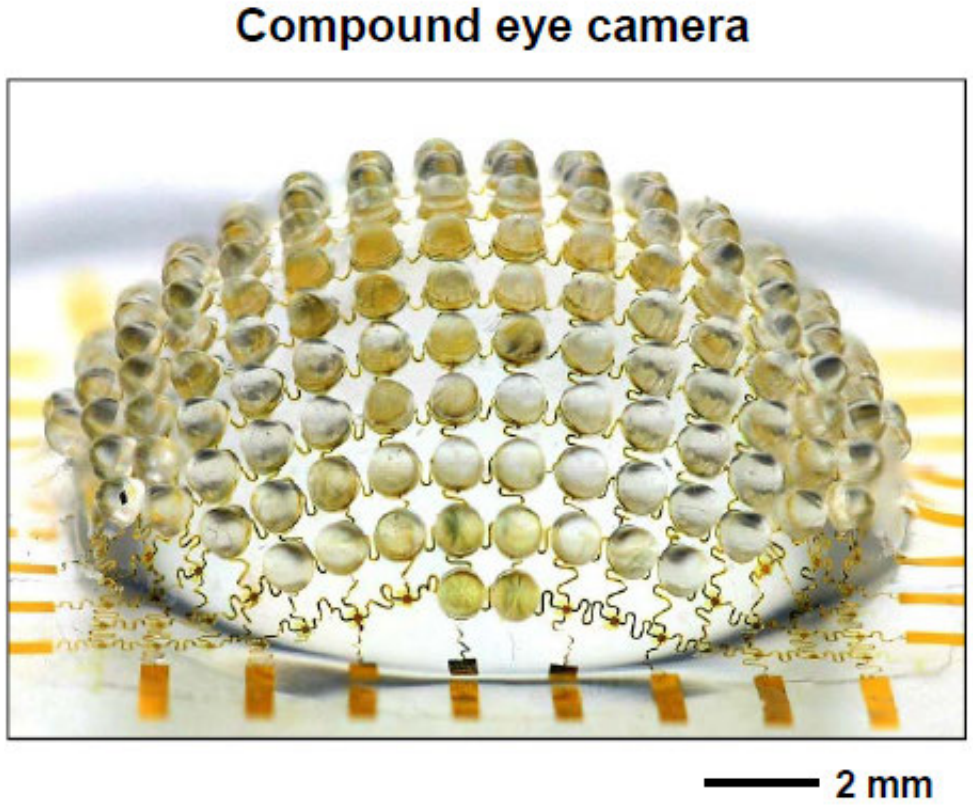


COMPutational Compound EYE (COMP-EYE) imaging system

Hemispherical Apposition Compound Eyes

- Implemented by stretchable microlens array and photodiodes
- Limitation: 180 pixels (16x16 photo diodes)



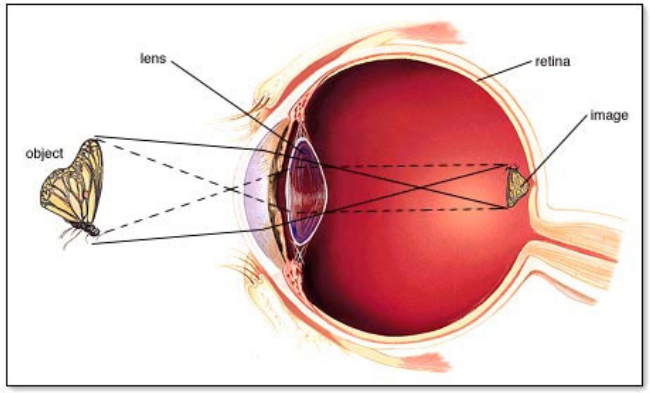


Computational Compound EYE imaging system

- Eyes in nature
  - Camera-type eye vs. Compound eye

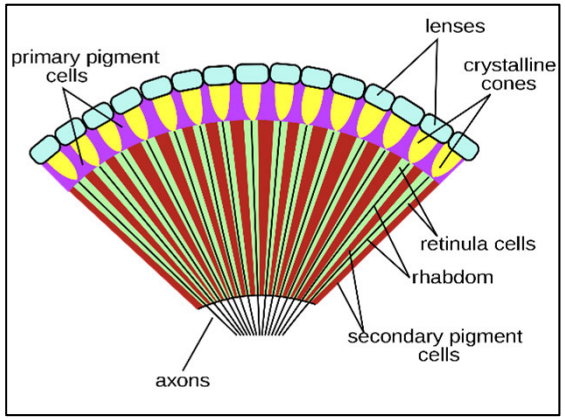


Camera-type Eye



- Single lens system
- High resolution
- Pattern recognition

Compound Eye



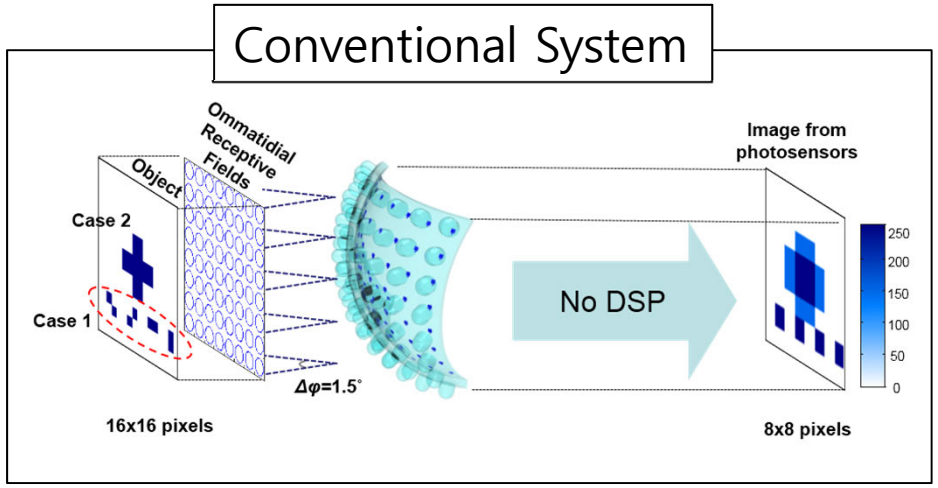


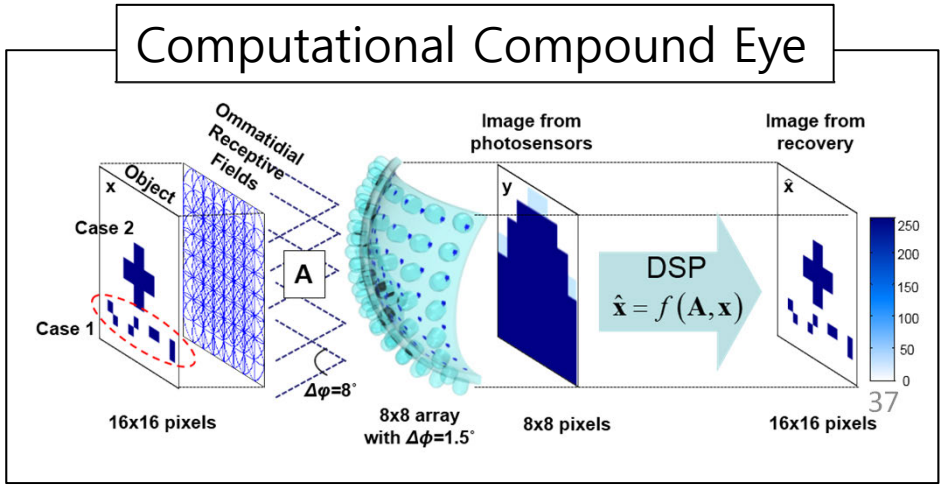
- Multi-lens system
- Wide field of view (FOV), Infinite depth of field (DOF)
- Motion detection

- Due to diffraction limit and low density of photoreceptors, the resolution of compound eyes is limited.

Computational Compound EYE imaging system

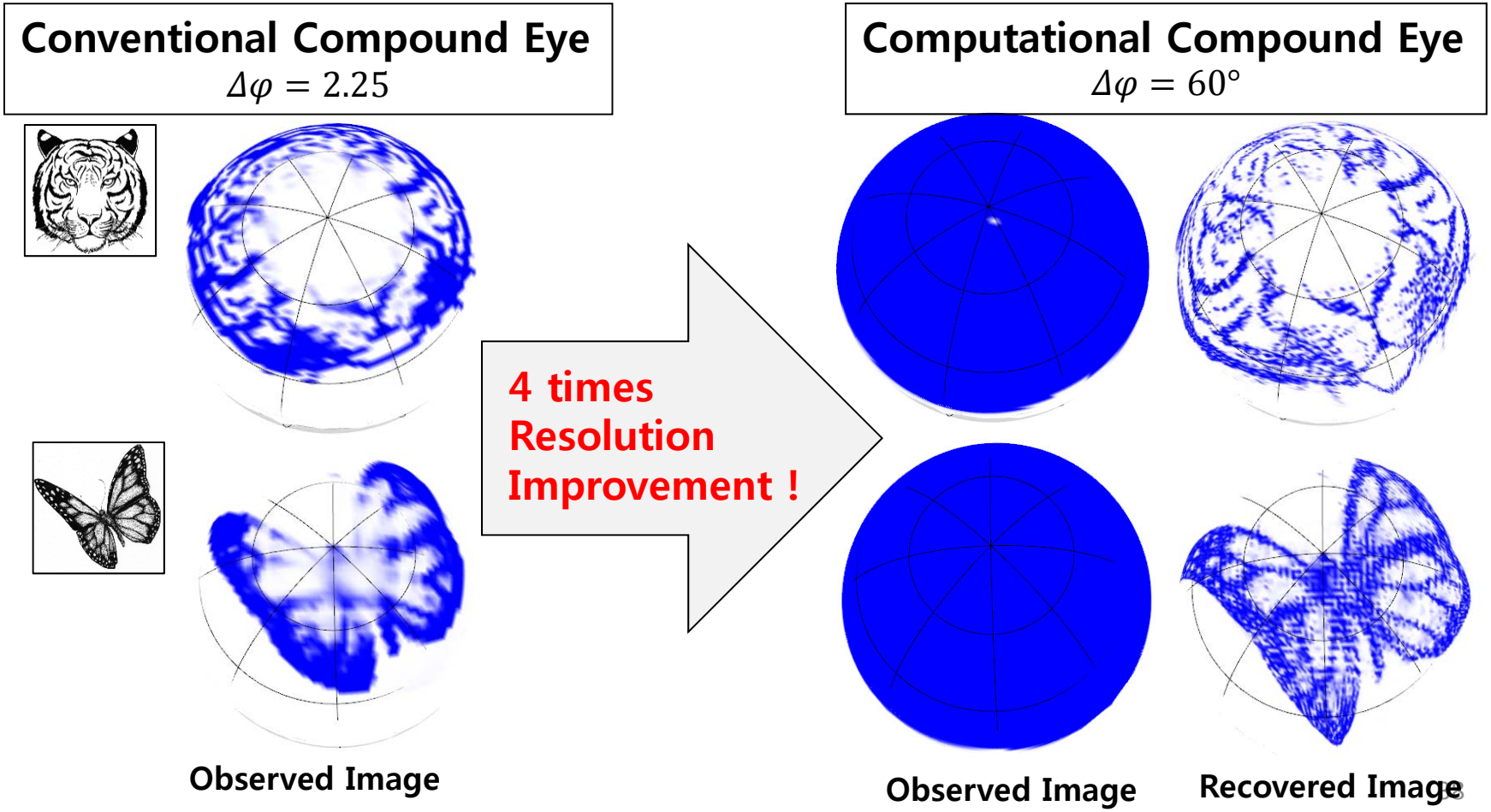
- COMP-EYE
  - We aim to improve the resolution of the compound eye imaging system by designing larger acceptance angles of ommatidia and using a digital signal processing (DSP) technique
  - Larger acceptance angles enable each ommatidium to observe multiple pieces of information all at once.
  - Each piece of information is observed multiple times by multiple ommatida each with different perspectives.
  - By exploiting this, the DSP technique recovers the object image with high resolution.





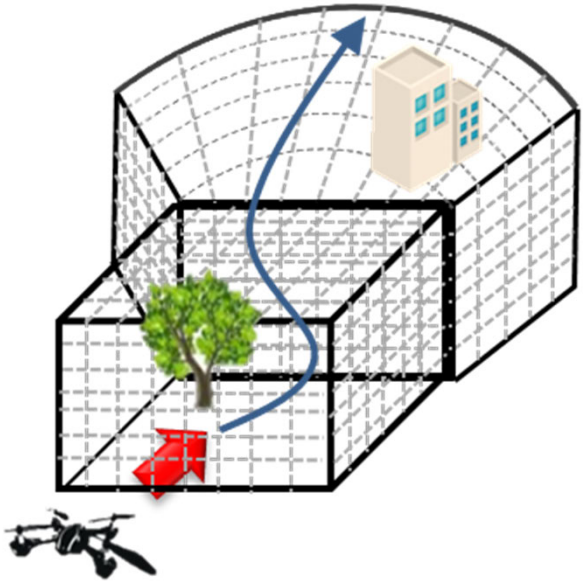
Computational Compound EYE imaging system

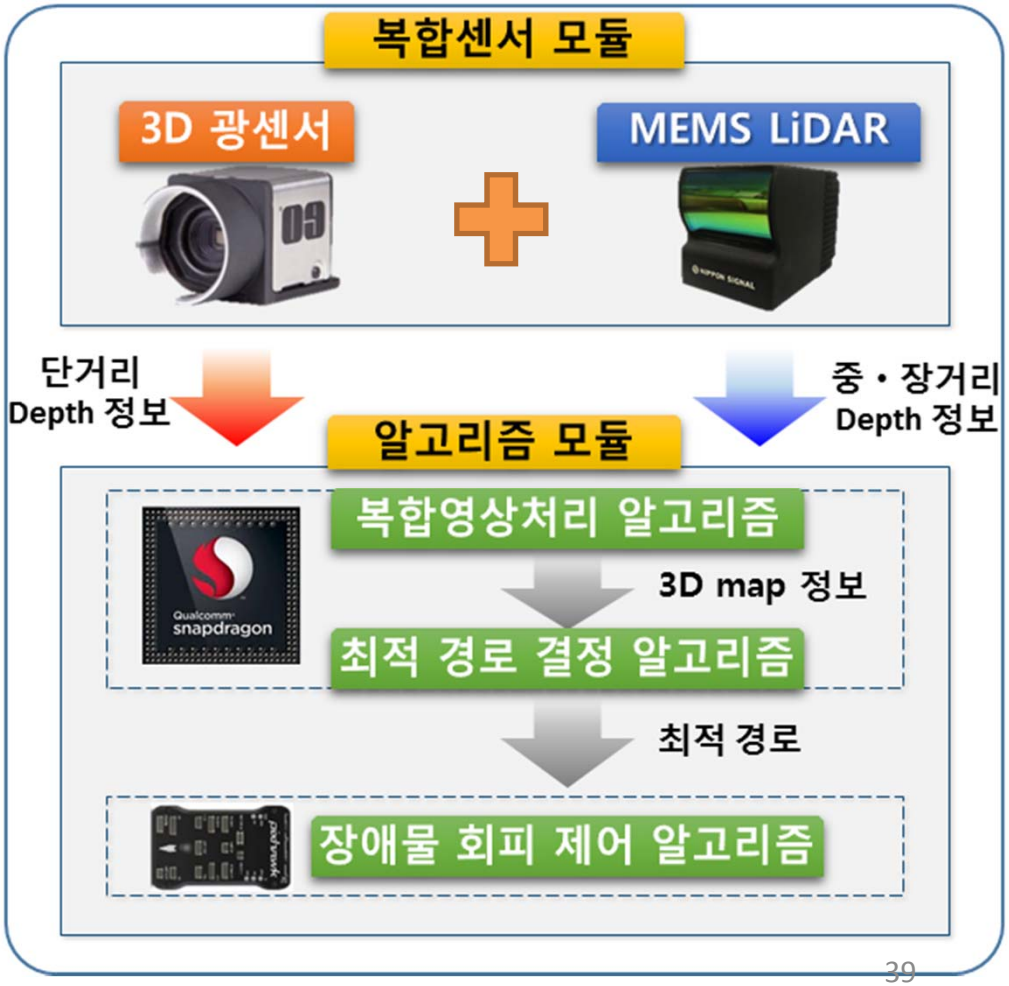
- Simulation results
  - M=80 x 80 ommatidia, N= 160 x 160 pixels



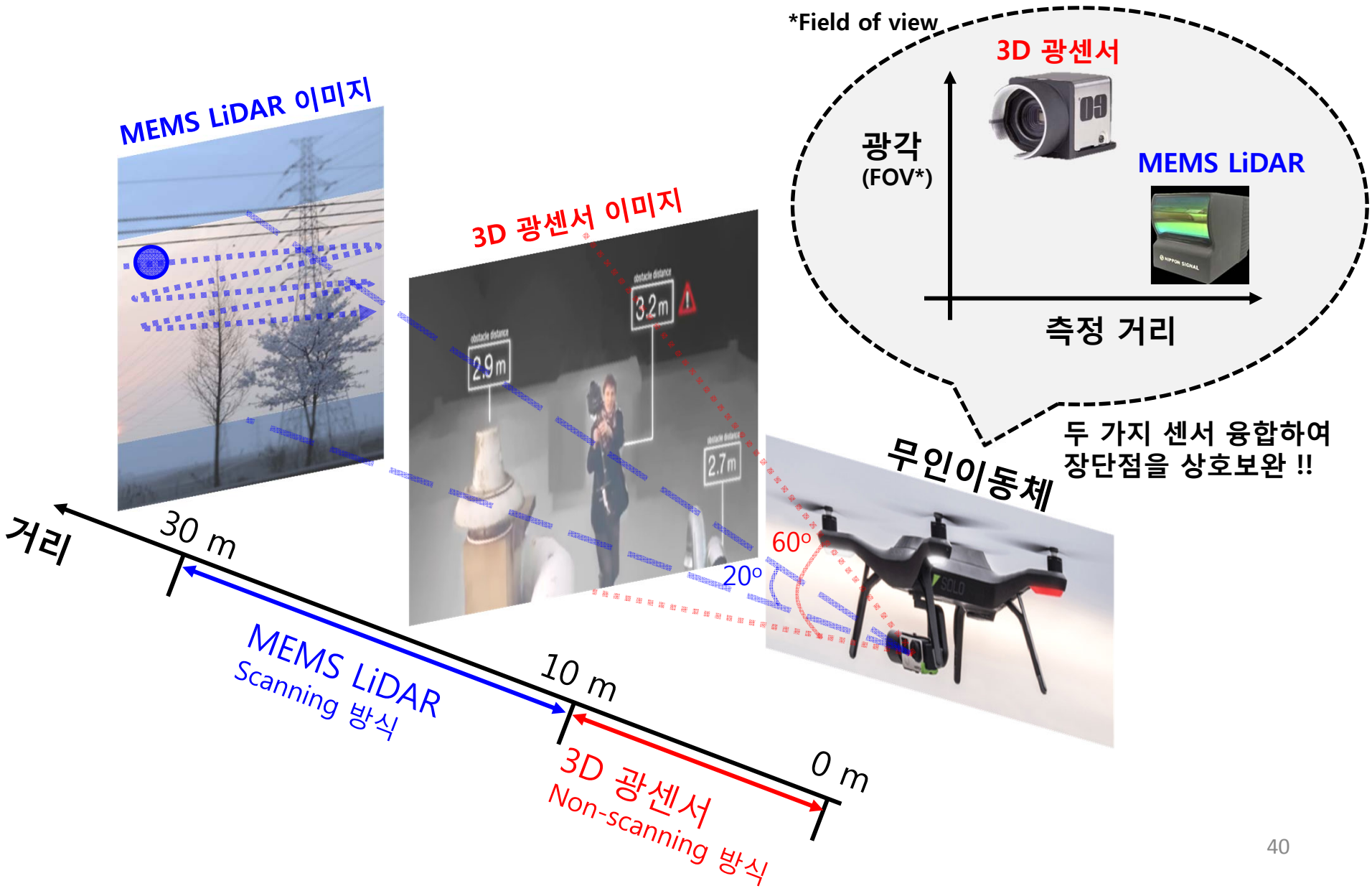
3D 영상센서와 초소형 라이다(LiDAR)를 결합한 차세대 3D인식 복합센서모듈 기술개발

- **최종 연구 목표**
  - ✓ **차세대 3D 인식 복합센서 및 장애물 충돌 회피 알고리즘 개발**
    - 3D 광영상 센서와 초소형 LiDAR를 결합한 복합센서 개발
      - 3 Lows: 무게, 부피, 소비전력
    - 복합영상처리 알고리즘 및 최적 경로 결정 알고리즘 개발
      - 탑재 가능 영상/신호 처리 알고리즘
    - 장애물 충돌회피 알고리즘 개발
    - 소형 드론 테스트베드 개발



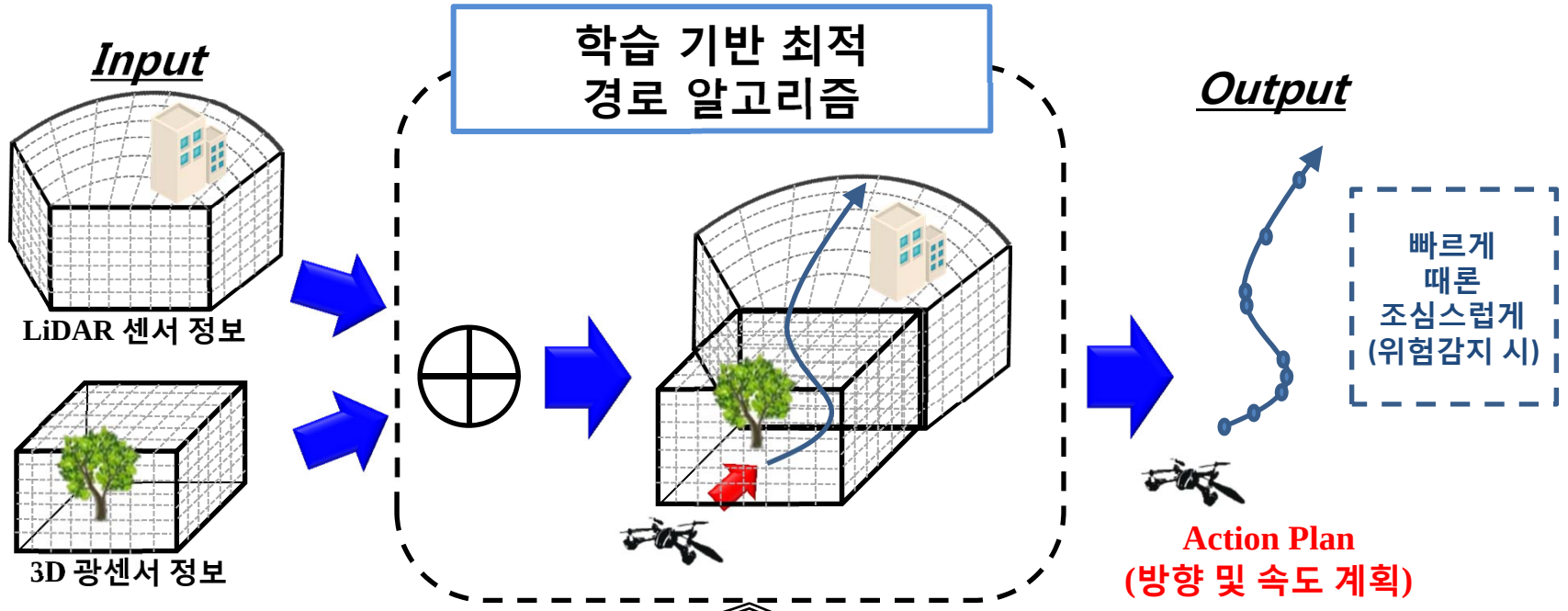


복합영상센서 개요


*Field of view
3D 광센서
광각
(FOV*)
MEMS LiDAR
측정 거리
두 가지 센서 융합하여
장단점을 상호보완 !!
MEMS LiDAR 이미지
3D 광센서 이미지
무인이동체
obstacle distance
2.9m
3.2m
2.7m
60°
20°
거리
30 m
10 m
0 m
MEMS LiDAR
Scanning 방식
3D 광센서
Non-scanning 방식

영상 통합 및 경로 설계 알고리즘 개발 1/2

- 목표: 해상도/제어불확실성/객체추적 속도 등이 변화하는 조건하에 충돌확률을 최소화하며 목적지에 가장 빠르게 도착하는 스스로 학습하여 성능을 개선하는 최소시간 경로 추적 알고리즘 개발 및 탑재 Test



y_L = LiDAR 센서 정보, y_O = 3D 광센서 정보, x = possible action

$$\textbf{Action} = \arg\max_{x} f(x|y_O, y_L, past\ 정보)$$

- 단기 Action 계획은 3D 광센서, 장기적 시선은 LiDAR정보를 활용 최적 경로 계산
- 외부 환경 요소 및 제어 불확실성을 고려
- 매 프레임 당 최적 경로 선택
- HD해상도(1Mpixels) * $\left({}^{\text{거리범위}}/_{\text{거리해상도}} \approx 11bits\right)$*프레임율(10)=110Mbps의 정보처리

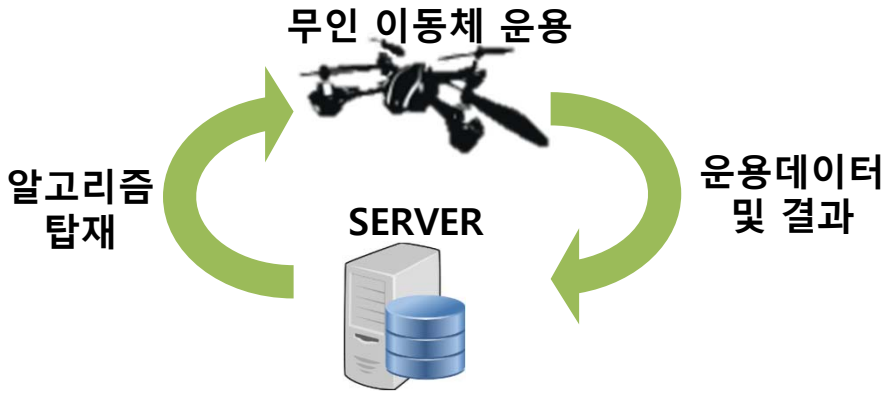
영상 통합 및 경로 설계 알고리즘 개발 2/2

✓ Tree Search 주행 경로 탐색 알고리즘

- ***Tree search 문제*** 모델링
  - 위치(p_t), 동작(a_t), 충돌 확률(v_{p_t})
- **Breadth 감소 방법** 개발 (**3D 광센서**)
  - 가능한 동작 제한 → Breadth 복잡도 감소
- **Depth 감소 방법** 개발 (**MEMS LiDAR**)
  - 충돌 확률 Regression → Depth 복잡도 감소

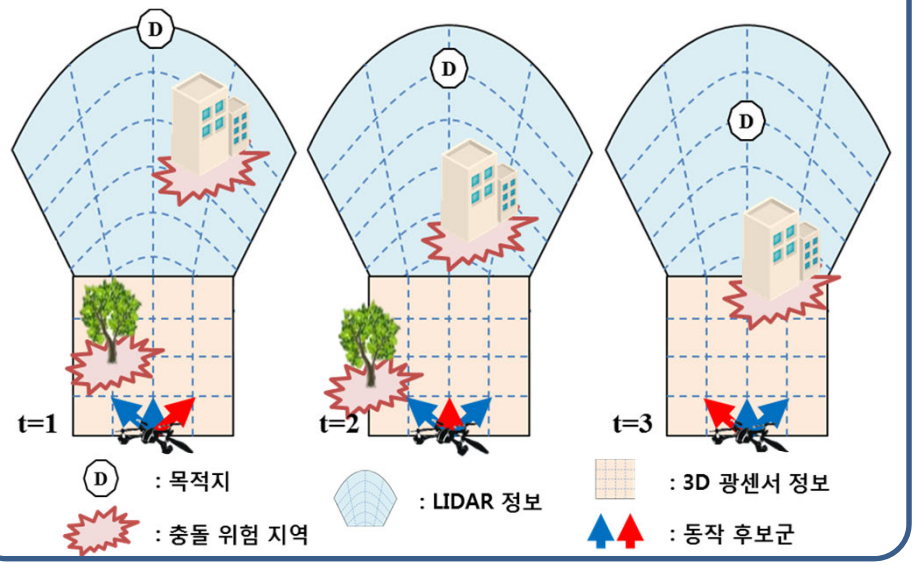
➡ 제한된 컴퓨팅 자원 및 외부 방해 요소 하에서 ***경로안정성*** 유지

✓ Reinforcement Learning (RL)

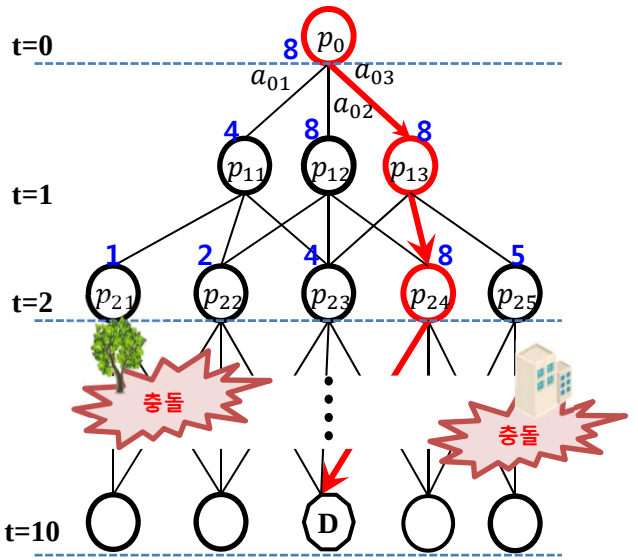


주행 경로 탐색 알고리즘 upgrade

✓ 주행 경로 탐색 알고리즘 예제



✓ Tree search 예제



대학ICT연구센터육성지원사업

연구개발계획서

2021. 4. 28.

신청분야	블록체인 · 융합 - 블록체인
주관기관명	광주과학기술원
센 터 명	블록체인 지능 융합 센터
과 제 명	영지식 센싱, 암호인증, 블록체인 기반 클라우드 서비스 융합 기술 개발

연구개발계획서

[√] 신청용
[ ] 협약용

보안등급
일반[√], 보안[ ]

중앙행정기관명	과학기술정보통신부	사업명	사업명	정보통신방송혁신인재양성	
전문기관명	정보통신기획평가원		내역사업명	대학ICT연구센터육성지원	
공고번호	제2021-296호	총괄연구개발과제번호		(기재 불필요)	
		연구개발과제번호		(기재 불필요)	
선정방식	정책지정[ ] 공모: 지정공모[ ] 품목공모[√] 분야공모[ ] 자유공모[ ]				

기술분류	국가과학기술표준분류	EE0401 네트워크 구조 설계/운영지원	60%	EE0303 서비스/응용 보안	20%	EE0304 산업보안/융합보안	20%
	ICT기술분류	블록체인 기반 기술	60%	블록체인 확장 기술	30%	블록체인 응용 지원	10%

총괄연구개발과제명 (해당 시 작성)	국문	(기재 불필요)
	영문	(기재 불필요)
연구개발과제명	국문	영지식 센싱, 암호인증, 블록체인 기반 클라우드 서비스 융합 기술 개발
	영문	Development of zkp sensing, cryptographic authentication, and blockchain based cloud service convergence technology

주관연구개발기관	기관명	광주과학기술원		사업자등록번호	410-82-07550
	주소	(우)61005 광주광역시 북구 첨단과기로123		법인등록번호	204271-0000871
연구책임자	성명	이흥노		직위	정교수
	연락처	직장전화	062-715-2237	휴대전화	[redacted]
		전자우편	[redacted]	국가연구자번호	[redacted]

연구개발기간	전체			2021. 07. 01 - 2028. 12. 31(7년 6개월)
	단계	1단계	1년차	2021. 07. 01 - 2021. 12. 31(0년 6개월)
			2년차	2022. 01. 01 - 2022. 12. 31(1년 0개월)
		2단계	1년차	2023. 01. 01 - 2023. 12. 31(1년 0개월)
			2년차	2024. 01. 01 - 2024. 12. 31(1년 0개월)
			3년차	2025. 01. 01 - 2025. 12. 31(1년 0개월)
			4년차	2026. 01. 01 - 2026. 12. 31(1년 0개월)
		3단계	1년차	2027. 01. 01 - 2027. 12. 31(1년 0개월)
			2년차	2028. 01. 01 - 2028. 12. 31(1년 0개월)

연구개발비 (단위: 천원)	정부지원 연구개발비	기관부담 연구개발비		그 외 기관 등의 지원금				합계			미포함 예산
				지방자치단체		기타( )					
	현금	현금	현물	현금	현물	현금	현물	현금	현물	합계	
총계	[redacted]										
1단계 1년차											
1단계 2년차											
2단계 1년차											
2단계 2년차											
2단계 3년차											
2단계 4년차											
3단계 1년차											
3단계 2년차											

공동연구개발기관 등 (해당 시 작성)	기관명	책임자	직위	휴대전화	전자우편	비고	
						역할	기관유형
공동연구개발기관							
위탁연구개발기관							
연구개발기관 외 기관	광주광역시청	이정선	주무관	[redacted]	[redacted]	지자체	공립연

연구개발과제 실무담당자	성명	최해웅	직위	박사과정 학생
	연락처 직장전화	062-715-2282	휴대전화	[redacted]
	전자우편	[redacted]	국가연구자번호	[redacted]

관련 법령 및 규정과 모든 의무사항을 준수하면서 이 연구개발과제를 성실하게 수행하기 위하여 연구개발계획서를 제출합니다. 아울러 이 연구개발계획서에 기재된 내용이 사실임을 확인하며, 만약 사실이 아닌 경우 연구개발과제 선정 취소, 협약 해약 등의 불이익도 감수하겠습니다.

2021년 4월 28일

연구책임자: 이흥노 (인)

주관연구개발기관의 장: 총장직무대행 김인수 (직인)

과학기술정보통신부장관 귀하

< 요 약 문 >

※ 요약문은 5쪽 이내로 작성

<table>
<tr><td colspan="2">사업명</td><td colspan="3">정보통신방송혁신인재양성</td><td>총괄연구개발과제번호</td><td colspan="3">(기재 불필요)</td></tr>
<tr><td colspan="2">내역사업명</td><td colspan="3">대학ICT연구센터육성지원</td><td>연구개발과제번호</td><td colspan="3">(기재 불필요)</td></tr>
<tr><td rowspan="2">기술분류</td><td>국가과학기술 표준분류</td><td>EE0401 네트워크 구조 설계/운영지원</td><td>60%</td><td>EE0303 서비스/응용 보안</td><td>20%</td><td>EE0304 산업보안/융합 보안</td><td>20%</td></tr>
<tr><td>ICT기술분류</td><td>블록체인 기반 기술</td><td>60%</td><td>블록체인 확장 기술</td><td>30%</td><td>블록체인 응용 지원</td><td>10%</td></tr>
<tr><td colspan="2">총괄연구개발과제명</td><td colspan="6">(기재 불필요)</td></tr>
<tr><td colspan="2">연구개발과제명</td><td colspan="6">영지식 센싱, 암호인증, 블록체인 기반 클라우드 서비스 융합 기술 개발</td></tr>
<tr><td colspan="2">전체 연구개발기간</td><td colspan="6">2021. 07. 01 - 2028. 12. 31(7년 6개월)</td></tr>
<tr><td colspan="2">총 연구개발비</td><td colspan="6">[redacted]</td></tr>
<tr><td colspan="2">연구개발단계</td><td colspan="3">기초[√] 응용[ ] 개발[ ]
기타(위 3가지에 해당되지 않는 경우)[ ]</td><td>기술성숙도
(해당 시 작성)</td><td colspan="2">착수시점 기준(2)
종료시점 목표(7)</td></tr>
<tr><td colspan="2">연구개발과제 유형
(해당 시 작성)</td><td colspan="6">인력양성-연구지원</td></tr>
<tr><td colspan="2">연구개발과제 특성
(해당 시 작성)</td><td colspan="6">(기재 불필요)</td></tr>
<tr><td rowspan="4">연구개발목표 및 내용</td><td>최종 목표</td><td colspan="6">지능적 자원관리와 효율성과 탄력성이 제고된 영지식 센싱, 암호인증, 클라우드 서비스 융합 기술 개발 및 석·박사급 고급인력 양성</td></tr>
<tr><td>전체 내용</td><td colspan="6">영지식 센싱, 암호인증, 블록체인 기반 클라우드 서비스 융합 기술 개발을 위해 다음과 같은 세부과제를 연구함
제1 세부과제: 영지식 센싱 탈중앙 고성능 분산원장 기술 개발 (혁신도약형)
제2 세부과제: Edge 컴퓨팅 기반 블록체인 서비스 인프라 기술
제3 세부과제: DID 블록체인 기반 IoT 인증 및 보안 기술 개발</td></tr>
<tr><td rowspan="2">1단계</td><td>목표</td><td colspan="6">(제1 세부과제)
o 신원 인증 키 (초분광이미지, 화자) 비가역적 아날로그-디지털 스캐너 연구
o 강건하고 보안성 있는 분산합의 기술 분석 및 개발
o 초경량・초확장 블록체인 개발
(제2 세부과제)
o 블록체인 대응 Cloud-native Connected DataLake 연구
o 신뢰 정의 네트워크(TDN) 구조 연구
(제3 세부과제)
o DID 블록체인 기반의 인증 및 안전한 Data 전송 방식 연구
o 물리계층 고유성을 활용한 DID 기술 개발</td></tr>
<tr><td>내용</td><td colspan="6">(제1 세부과제)
o 아날로그 신원 인증 키 연구
o 비가역적 아날로그-디지털 스캐너 설계
o 초분광압축이미징 장치 구현 및 데이터 확보
o 대용량 화자 데이터 확보, 분류, 정제 및 특징추출
o 최신 딥러닝 기반 화자인증 베이스라인 구축 및 기초 모델 개발
o 분산 합의 기술 분석 및 특징 파악
o 이종 분산합의 기술에 의한 환경 영향 분석
o 강건한 블록체인 인프라를 위한 분산합의 기술 분석 및 개발
o 트랜잭션 크기 별 안전 거래 분석 알고리즘 개발
o 초경량 블록체인 기술 개발
o 초확장 블록체인 기술 설계
o 신뢰실행환경 기반 데이터 검증 경량화 기술 개발</td></tr>
</table>

			(제2 세부과제) o 블록체인 대응 Cloud-native Connected DataLake 구성설계 및 PoC검증 o 신뢰 정의 네트워크 구조 정의 및 데이터 제어 방식 구성 o 지속 가능한 데이터 스트림 경로의 신뢰도 수준 등급화 및 이를 기반으로 트래픽 분류 및 처리 방법 모델링 **(제3 세부과제)** o DID 기반 인증/인가 서비스 요구사항 분석 o IoT에 적합한 DID 블록체인 인증 기법 연구 o 안전한 IoT Data 전송 방법 연구 o 블록체인 기반 보안 시스템을 위한 사전연구 o IoT의 물리적 특성을 활용한 DID 생성기술 개발 o DID 블록체인 연계 장치 인증 프로토콜 연구
	2단계	목표	**(제1 세부과제)** o 영지식 증명 및 검증 프로토콜 연구 (초분광 이미지, 화자인증) o 다중 복합구조 계층형 블록체인 플랫폼 기술 개발 및 테스트 o 빅데이터 및 분산형 인프라를 이용한 분산 애플리케이션 개발 o 데이터 프라이버시를 보존할 수 있는 연합학습 방법론 개발 **(제2 세부과제)** o 블록체인 대응 멀티사이트 Connected DataLake 연구 o 블록체인 플랫폼 지원을 위한 TDN 구성 기술 개발 **(제3 세부과제)** o DID 블록체인 기반의 인가 및 접근제어 방법 연구 o 복제불가능한 DID 기반 보안기술 개발
		내용	**(제1 세부과제)** o 영지식 증명 및 검증 프로토콜 설계, 성능분석 및 취약점 분석 o 초분광 이미지에서 비정상 데이터를 인지/추출/분류 알고리즘 개발 o 초분광 이미지를 복원하는 고급 딥러닝 기법 연구 o 시간 영역 Raw 데이터를 이용한 딥러닝 기반 고신뢰 화자인증 기술개발 o 신뢰도 향상을 위한 화자인증에 특화된 추가 모듈 개발 o 원거리 상황, 잡음 환경 및 짧은 음성 입력에 강인한 모델 개발 o 다중 복합구조 블록체인 연구 o 블록체인·샤드·클라우드 간의 통신 프로토콜 및 플랫폼 연동 기술 연구 o 블록체인 기반 빅데이터 생성 및 관리 기술 개발 o 분산형 애플리케이션을 위한 비즈니스 로직 및 스마트 계약 기술 개발 o 탈중앙화 연합학습 플랫폼 설계 o 비잔틴 내구성 연합학습 방법론 설계 o 블록체인 기반 연합학습 플랫폼 개발 o IoT 기기를 위한 인공지능 모델 경량화 **(제2 세부과제)** o 블록체인 대응 멀티사이트 Connected DataLake 구현 o 다중경로 초고속 데이터 전달 기능 개발을 통한 Connected DataLake 개선 o 신뢰도 기반의 네트워크 구현 및 이에 최적화된 오버레이 네트워크 지원 o 블록체인 노드 간의 선택적 및 지능적 멀티캐스팅을 TDN 측면에서 지원 o TDN의 Time Sensitive Networking(TSN) 기능 지원 o 장애 허용성을 가지는 분산형 TDN 컨트롤러 기반 트래픽 분산 방안 연구 **(제3 세부과제)** o DID기반의 자기주권형 액세스 토큰 발행 연구

			o IoT 환경의 DID기반 접근 제어 방법 연구 o 영지식 증명 기반의 자격증명 방법 연구 o 대규모 IoT 환경에서 신뢰성을 가지는 동적 DID 생성 기술 개발 o 블록체인 연계 DID 기반 저전력/고신뢰 보안인증 프로토콜 연구 o 소규모 private 블록체인을 활용한 보안통신 기술 검증
	3단계	목표	**(제1 세부과제)** o 영지식 아날로그 생체인증 시스템 연구 o 다중 복합구조 계층형 블록체인 플랫폼 기술 운용 **(제2 세부과제)** o 블록체인 대응 Connected DataLake 고도화 o 블록체인 서비스와 연동한 TDN PoC 레벨 구현 **(제3 세부과제)** o DID 인증/인가 기반 블록체인 운용 시스템 통합/평가 o DID 기반 블록 생성 및 검증 기술 연구
		내용	**(제1 세부과제)** o 영지식 아날로그 생체인증 시스템 개발, 테스트베드 구축 및 평가 o 개발한 모델과 베이스라인 화자인증 성능 비교를 통한 검증 o 개발한 모델 최적화 및 고도화, 경량화 연구 진행 o 사용 용도별 블록전파지연시간, 참여노드 수에 따른 분산원장 파라미터 최적화 o 샤드, 2nd 레이어, 크로스체인(cross-chain) 기술 융합 운용 o 이종 IoT 네트워크 설계 및 크로스체인 기술 개발, 통합 기술 개발 **(제2 세부과제)** o 블록체인 대응 Connected DataLake 고도화를 위한 분산된 데이터셋 동기 기능 개발 o TDN 요소 기술 통합 및 테스트베드 개발 o TDN 시스템 통합 및 PoC 레벨 구현 **(제3 세부과제)** o DID 인증/인가 기반 블록체인 연구/개발 o DID 인증/인가 기반 블록체인 응용 시스템 성능 측정/ o 복제불가능한 DID 기반 DID 문서 생성 및 제어 기법 개발 o IoT를 활용한 테스트베드 구성 및 성능 평가
연구개발성과 활용계획 및 기대 효과			**(기대효과)** 블록체인 기반기술 개발을 통하여 다음과 같은 효과를 기대할 수 있음. (1) 기술적 측면 o DID 기반기술, 초경량·초확장 블록체인 기술, 대용량 데이터 저장/공유/관리 기술, 신뢰정의 네트워크 기술 확보 o 확보한 기술들을 통해 디지털 트윈 구현의 장애물이 되는 디지털 환경과 현실 환경의 차이를 줄일 수 있음. 이를 통해 디지털 트윈기술의 활용성 제고 (2) 경제적, 산업적 측면 o 블록체인 데이터 저장 원천 기술 확보로 누구나 쉽게 블록체인 데이터 기반 융합 서비스 제공 역량 향상, 산업간 융합 IT 서비스 개발을 활성화하는 등 데이터 경제 활성화를 통해 4차산업혁명 견인 o 다양하고 혁신적인 융합서비스를 창출할 수 있는 신뢰 정의 네트워크 기반 환경을 조성함으로써, 차세대 글로벌 네트워크 전략에서 국가 경쟁력 확보 o (IoT 보안 통신 시장 선점) IoT 장치를 위한 보안통신 원천기술 확보를 통해 로열티 및 기술 수출로 새로운 시장확보

	o 높은 보안성을 요구하는 산업 (원자력, 의료, 전력 등)의 시설을 IoT 장치로 대체할 수 있게 되어 관련 사업의 효율성을 극대화 o 정보기기의 신뢰성 문제를 해결함으로써 AR, VR, 메타버스 등 디지털 트윈 기술이 필요한 새로운 애플리케이션 시장을 활성화 (3) 사회적 측면 o 본 과제의 DID 기반 인증과 영지식 증명 기술로 사생활 민감정보, 의료정보, 생체인증 등에 대한 인증/증명 시스템에서 개인정보 보호능력 제고 o 위변조가 불가능하고 원본의 무결성을 보증하여 투명하고 신뢰받는 사회를 구현할 수 있으며, 특히 정부/공공서비스에 위와 같은 기술을 활용하여 강력한 신뢰사회를 구축 o 신뢰 정의 네트워크를 통해 사회시스템의 신경망 임무를 수행함으로써 시공간을 초월한 자유로운 참여와 공유 확산 효과 o 암호화폐에만 집중된 블록체인 기술의 산업·경제 구조의 불균형 문제 개선 o 데이터, 블록체인 기술 전문인재를 양성 및 신규 일자리 창출 가능 o 데이터를 보호하는데 중요한 기술 개발을 통해 마이데이터 산업 활성화에 역할을 하며 데이터 자기주권 확보 특히 ICT 융복합을 위한 전문지식을 갖춘 박사급 글로벌 인재 양성함으로써 다음과 같은 파급효과를 기대함. o 학문 융복합이 이뤄질 것으로 예견되는 미래사회를 위한 성장 동력확보 o 배출된 전문 인력을 중심으로 관련 기업에게 맞춤 인력을 공급하고 기술개발의 공동 참여기회를 확대 o 블록체인 융복합 학문을 산업체에 공급함으로써, '기술개발-상품화-수익창출-전문인력 양성'의 순환 고리를 연결하는 교육·연구개발 통합클러스터 구축 o 철저한 실무 교육으로 ICT 인력의 질적 수준을 향상시켜 4차산업혁명 시대 국가 경쟁력 제고 o 안전하고 신뢰성 있는 분산형 블록체인 기반 공공 서비스 제공에 필요한 인재 양성을 통해, 목적별 공공분야 발전 및 국민 편익 증대 o 산업현장 진출 인력에 대한 체계적 사후관리 및 수련 중인 학생들과 지속적인 연계를 통해 관련분야 고급기술 인력간의 인적 관계망 구축 **(활용계획)** 연구과제의 성과의 활용계획은 다음과 같음. (1) 연구개발성과의 특허관리, 기술이전, 사업화 시행 o 광주과학기술원 과학기술응용연구소(GTI)와 연계한 기술이전, 상용화, 시제품개발 시행; 연구성과를 GTI에 제공하고 GTI는 특허관리, 사업 확장, 애로기술지도 등을 지원하고 수익금은 연구 사업에 재투자하여 지속가능한 R&D형 인력양성 사업 기반 마련 (2) 고급 인력 배출을 통한 고용촉진 장려 o 참여 학생들의 지속적 관리를 통한 취업 촉진 및 기업과 연계한 고용 장려 시스템 구축하고 지적 인프라를 활용하여 협력업체의 기술 개발에 도움을 줄 수 있는 연결고리 역할 수행

국문핵심어 (5개 이내)	블록체인	탈중앙화 신원증명	클라우드	인증	분산원장
영문핵심어 (5개 이내)	Blockchain	Decentralized Identity	Cloud	Authentication	Distributed Ledger

< 본문 1 >

목 차

1. 연구개발과제의 필요성 ····· 1

1) 서론 ····· 1

2) 필요성 ····· 2

3) 동향 ····· 5

2. 연구개발과제의 목표 및 내용 ····· 6

1) 연구개발과제의 최종 목표 ····· 6

2) 연구개발과제의 단계별 목표 ····· 12

3) 연구개발과제의 내용 ····· 16

4) 연구개발과제의 수행일정 및 주요 결과물 ····· 23

5) 세부 프로젝트 연구계획 ····· 28

6) 혁신도약형과제 내용 ····· 28

3. 연구개발과제의 추진전략·방법 및 추진체계 ····· 45

1) 연구개발과제의 추진전략·방법 ····· 45

2) 연구개발과제의 추진체계 ····· 48

3) 연구책임자 및 참여인력의 역량 ····· 53

4) 수행기관의 역량 ····· 55

4. 인력양성 계획 ····· 59

1) 전주기적 인력양성 계획 ····· 59

2) 창의·혁신 인재양성 방안 ····· 63

5. 연구개발과제의 활용 방안 및 기대효과 ····· 66

1) 연구개발성과의 활용방안 ····· 66

2) 연구개발성과의 기대효과 ····· 66

3) 인재양성 파급효과 ····· 68

6. 연구개발성과의 사업화 전략 및 계획(해당시) ····· 69

1) 블록체인 지능 융합 센터 사업의 SWOT 분석 ····· 69

2) 블록체인 지능 융합 센터 사업의 전략목표 및 시사점 ····· 69

3) 블록체인 지능 융합 센터의 기술이전(사업화) 전략 ·····70
4) 연구개발성과 활용계획 ·····70
5) 연구개발성과 상용화 계획 ·····71

7. 연구개발 안전 및 보안조치 이행계획 ·····72
1) 안전조치 이행계획 ·····72
2) 보안조치 이행계획 ·····74

< 본문 2 >

8. 연구개발기관 현황 ·····77
1) 연구책임자 등 현황 ·····77
2) 연구개발기관 일반 현황 ·····93

9. 연구개발비 사용에 관한 계획 ·····94
1) 연구개발비 지원·부담계획 ·····94
2) 연구개발비 사용계획 ·····94
3) 연구시설·장비 구축·운영계획 ·····97

10. 평가기준 및 평가방법 ·····99
1) 연차별 자율성과목표 ·····99
2) 그 외 주요 성과목표 ·····100

1 연구개발과제의 필요성

1) 서론

○ **(디지털 전환의 시대)** 코로나 19는 인공지능(AI), 블록체인(Blockchain), 클라우드(Cloud), 데이터(Data)와 같은 기술이 이끄는 새로운 뉴노멀을 형성. 코로나 19의 확산으로 모든 산업과 생활양식의 변화는 비대면 디지털 전환을 초래했으며 특히 전 세계적으로 디지털 경제활동을 가속화함. 도시의 여러 데이터를 활용해 디지털 인프라와 4차 산업혁명 기술을 도시 설계, 개발, 운영 및 유지관리의 기반으로 둠으로써 물리적 환경과 디지털 환경을 통합한 살아있는 생활환경을 구축하는 것이 중요함.

○ **(프로토콜 경제)** 언택트에 기반하여 급성장하기 시작한 플랫폼 경제는 데이터 독점과 폐쇄성 문제 등 여러 문제를 발생시키기 시작함. 이를 극복할 수 있는 차세대 경제모델로 프로토콜 경제가 각광을 받기 시작. 프로토콜 경제는 플랫폼 경제에 대한 문제의식으로 생겨난 새로운 경제개념으로 시스템이 데이터를 교환하기 위해 사용하는 규약을 의미. 2021년 정부의 경제정책 방향에도 프로토콜 경제 발전전략 수립이 과제로 담김.

○ **(블록체인의 역할: 신뢰확보)** 4차 산업혁명의 핵심기술인 인공지능과 블록체인의 관계는 혁신과 신뢰, 성장과 분배라고 볼수 있음. 인공지능 발달에 있어 데이터 신뢰성 문제가 발생하면 학습 알고리즘이 훌륭해도 의미가 퇴색함. 블록체인은 모든 거래와 정책, 의사 표현을 이해당사자와 참여자 모두에게 공개하면서도 변조나 해킹이 불가능한 기술이기에 초신뢰 경제혁명을 불러올 뿐만 아니라 신뢰성을 획기적으로 신장시킬 역량을 보유함. 즉, 블록체인 기술은 이해관계자 간 민주적 절차를 통해 이용자가 함께 플랫폼 규칙을 설정할 수 있어 개방성과 투명성을 확보할 수 있으며 참여자에게 공정한 분배시스템을 보장함으로써 네트워크 확장과 경제적 성장을 함께 얻을 수 있음.

○ **(ABCD 융합)** 특히 인공지능(AI), 블록체인(Blockchain), 클라우드(Cloud), 데이터(Data) 기술이 결합된 ABCD 네트워크를 구성, 지역내에 발생하는 데이터를 축적하고 해당 데이터를 통해 AI 소프트웨어, 모델 등은 인공지능 자동학습을 통해 업그레이드할 수 있음. 학습을 통해 향상된 사항은 교통신호와 연동된 안전한 도로주행(자동차), 정확한 장소에 배달하는 배달로봇, 효율적인 전력 생산 및 거래, 행정정보 제공 등 시민의 안전과 편의를 극대화할 수 있음.

○ **(블록체인 지능 융합 연구개발 및 우수인력 확보의 필요성)** 무한한 가능성을 현실화할 수 있는 블록체인 기반 생체 및 디바이스 인증 방식은 정보 탈취의 취약성을 드러내고 있으며, 생체정보 유출을 방지할 보안방법 또한 매우 부족한 상황임. 컴퓨팅 파워를 보유한 네트워크는 클라우드 컴퓨팅 기반의 데이터 스토리지 공간 확보를 통해 실증 대상인 데이터 축적과 인공지능 학습활용이 가능하지만, 과연 어떻게 인공지능, 클라우드, 데이터 기술과 합의 프로토콜을 융합하여 투명성, 신뢰성, 경량성, 확장성, 가용성 등이 지속적으로 제고되고 수요에따라 탄력적인 운용이 가능한 네트워크 설계하고 구현할 것인지에 대한 학문적 연구 성과가 전무함. 어떻게 탈중앙성을 확보하고, 모두의 참여를 독려하며, 프로토콜 설정 합의에 따른 포용과 투명성을 보장할 것인지 구체적인 방안이 없음. SCI급 논문 성과도 전세계적으로도 매우 미흡한 수준이고, 석박사급 인력양

성 노력도 부족한 현실임.

○ 위와 같은 문제점을 아래와 같은 세가지 방향에서 분석하고 연구개발의 필요성을 도출하고자 함.

1) 영지식 증명을 활용한 새로운 생체인증 방식 및 탈중앙화된 인증방식 개발의 필요성

2) 클라우드-네이티브 한 데이터 레이크, 신뢰 가능한 네트워크 관리 기술 개발의 필요성

3) 블록체인과 물리계층 기술을 결합한 지능형 IoT기기 보안관리 시스템 개발의 필요성

2) 필요성

○ **(디지털 뉴딜 정책과의 정합성)** 정부에서는 2025년까지 총 58.2조원을 투자해 경제 및 사회 전반의 디지털 대전환을 가속화하는 국가 혁신 프로젝트 '디지털 뉴딜 정책'을 펼치고 있으며, 디지털 전환의 핵심기술 중 하나로 블록체인을 꼽을 수 있음.

○ **(디지털 전환 시대의 네트워크 필요성: 신뢰성)** 블록체인에 기록된 데이터는 위·변조가 거의 불가능하므로 디지털 및 프로토콜경제로의 전환에 필수적인 신뢰확보 기술임. 하지만 해당 블록에 기록되는 센서 데이터 자체가 왜곡될 경우 문제가 발생할 수 있음. 따라서 기존 네트워크 형태와는 구분되는 신뢰-정의 네트워크 개념 도입이 필요함.

○ **(디지털 전환 시대의 네트워크 필요성: 보안성)** 전력사용정보, 의료정보 등 개인정보를 수집하는 센서 뿐만 아니라 중요한 시설(스마트 그리드 등)을 제어하는 IoT 기기의 보급이 증가함에 따라서 보안 문제가 대두되고 있음. 현재 IoT의 데이터 전송방식은 멀티 홉 네트워크를 통해 데이터 전송을 수행하기 때문에 데이터의 기밀성에 대한 취약점이 있음. 기존의 중앙화된 신원증명 시스템은 많은 개인정보 주권 문제 및 보안 문제를 안고 있어 이를 해결하기 위한 방안으로 블록체인을 이용한 탈중앙화된 신원증명(Decentralized identity, DID) 기술의 중요성이 떠오르고 있음. DID는 정보에 대한 통제권이 중앙에 있지 않으며, 개개인이 자신의 정보에 완전한 통제권을 갖게 함.

○ **(디지털 전환 시대의 네트워크 필요성: ABCD 융합)** 전방위적인 ICT 중심의 디지털전환을 시도하는 국제적인 노력들이 활발함. 즉 다양한 산업 도메인들에서 산출된 개방형 데이터를 활용하면서, Data-Networking-AI (D-N-A) 개념의 융복합 서비스 창출을 중심으로 지속적인 가치-지향의 생태계를 체계화하는 노력이 확산중임. 따라서 목표하는 가치-지향 서비스 생태계를 선제적으로 만들어내기 위해서는 데이터, 인공지능등을 유기적으로 연계해가는 Data—AI 통합 프레임워크와 같은 구조적인 접근이 필요함.

[영지식 센싱 탈중앙 고성능 분산원장 기술 개발]

○ **(블록체인 합의 메커니즘 연구 필요성)** 블록체인에서 합의 메커니즘은 분산 노드 간의 신뢰를 가져오고 트랜잭션을 처리할 수 있게 해주며, 작업 증명(PoW) 기반의 합의 메커니즘은 암호화 측면에서 안전한 것으로 입증됨. 그러나 ASIC 채굴기가 도입되고 집중되면서 블록체인 네트워크는 더욱 중앙 집중화되었고, 그 결과로 블록체인 네트워크를 유지하는 데 필요한 에너지가 폭증했음. 이러한 문제에 더해 느린 거래 시간은 지분 증명과 지식 증명과 같은 대안적 합의 메커니즘에 대한 관심이 증가하는 계기가 되었

음. 따라서 보다 분산적이고, 안전하고, 강력하고, 친환경적인 합의 메커니즘이 필요함.

○ **(블록체인의 기반 연합학습 필요성)** 탈중앙화 환경에서도 신뢰를 형성할 수 있는 데이터 분산원장기술인 블록체인(Blockchain)을 기반으로, 중앙 집권형 컴퓨팅에서 탈피하고 IoT 기기 스스로가 대용량 데이터를 처리할 수 있으며 효율적으로 스토리지 자원을 활용하는 데이터 관리기술이 필요함. 또한 본 플랫폼에서 축적된 데이터를 활용해, 탈중앙성을 해치지 않으면서도 인공지능 모델을 학습해 IoT 기기를 통한 사물제어 성능을 제고할 수 있는 연합학습(Federated Learning) 방안이 필요함.

○ **(화자인증 기술의 필요성)** DID에서 고도의 보안성을 제공하기 위해 사용자 인증에 개인의 생체정보를 활용하는 기술이 개발되고 있음. 기존에 보편적으로 사용되는 생체정보 데이터는 지문과 얼굴 사진 등을 포함함. 한편, 화자인증(Speaker verification, SV) 기술은 화자의 음성을 분석하여 데이터에 대한 접근권한 부여 및 차단을 결정하는 기술임. SV 기술은 멀티모달(Multi-modal) 인증 방식에서 하나의 생체정보와 관련한 인증 modalitiy로 추가되어 인증 신뢰도 향상에 기여할 것으로 기대함.

○ **(생체인증을 위한 초분광 이미징 장치의 필요성)** 사용자 인증을 위한 생체정보로 중 홍채, 지문, 장문등을 사용할 경우 초분광 이미징 장치가 사용됨. 하지만, 기존의 초분광 이미징 장치는 크기가 크고, 값이 비싸며, 측정 시간이 긴 단점이 있음. 따라서 필터 어레이 기반 소형 초분광 압축 이미징 장치의 개발이 필요함.

○ **(생체인증의 보안 문제)** 생체인증은 신체정보를 디지털 데이터로 변환하여 사용하기 때문에 데이터가 외부로 유출 될 경우 보안 문제가 발생함. 2017년 송희경 자유한국당 의원은 국정감사에서 직접 실리콘으로 만든 가짜 지문으로 손쉽게 생체인증을 뚫는 시연을 하였음. 이 때, 가짜 지문 실리콘 모형은 주민등록증에 있는 지문 이미지를 디지털 파일로 변환하여 생산하였음.

○ **(생체정보의 보안방법 개발 필요)** 생체인증시 또는 생체정보 등록과정에서 개인정보 노출의 가능성이 있으며, 생체정보가 유출될 경우 생체정보가 제공자의 동의 없이 다른 서비스나 응용에서 사용되거나, 생체정보로부터 그 사람의 부가적인 정보가 노출될 수 있으며, 등록된 생체정보가 제 3자에 의해 오용될 위험이 있음. 게다가 생체정보는 변경 불가능하므로 한번 해킹당할 경우 피해가 지속될 수 있음. 따라서 생체정보를 안전하게 보호할 수 있는 보안방법의 개발이 필요함.

[Edge 컴퓨팅 기반 블록체인 서비스 인프라 기술]

○ **(Cloud-native 컴퓨팅에 기반한 상호호환적인 에지 컴퓨팅 인프라와 이에 기반한 블록체인 대응 데이터 수집·정제 및 공유 기능의 필요성)** 오픈소스 클라우드-네이티브(Cloud-native) 컴퓨팅을 활용하여 특정 벤더의 의존성을 줄이는 하이브리드/멀티 클라우드 대응 에지 인프라의 구축, 그리고 D-N-A 패러다임에 충실하게 대응하는 유연한 데이터 공유가 가능한 에지 인프라의 구축의 필요성이 커지고 있음. 이때 D-N-A 패러다임에서 요구하는 데이터의 확장성있는 수집·공유를 위해서는 'Connected DataLake' 개념에 따른 에지 및 코어 클라우드 인프라간의 유기적인 협력에 따른 블록체인 대응 데이터 공유 지원도 매우 중요함.

○ **(블록체인을 통한 신뢰할 수 있는 데이터 스트림 제공 필요성)** 사물인터넷으로부터 수집되는 각종 센서 데이터를 위변조 혹은 손실 없이 안전하게 블록체인 플랫폼으로 전달하기 위한 온전히 신뢰할 수 있는 데이터 전송 경로가 요구됨. 기존 네트워크의 보안은 종단 중심으로 발전해오고 라우팅 경로에 대해 온전한 신뢰를 담보할 수 없음. 여기에 기타 네트워크 컨트롤러의 개입까지 더해지면 그 불확실성은 더욱 커짐. 따라서 이러한 문제를 타파할 수 있는 네트워크 패러다임 개발이 요구됨.

○ **(중앙화된 네트워크 컨트롤러의 장애에 대한 해결책 필요)** 현재 네트워크 제어 기능은 중앙화된 컨트롤러를 기반으로 소프트웨어 정의 네트워크 형태로 발전하고 있음. 이는 네트워크의 신뢰성 확보에 적합한 형태임. 하지만 이에 대한 트레이드오프로 중앙화된 컨트롤러에 권한이 집중되고 이것은 곧 단일 장애점을 유발함. 따라서 이러한 문제를 절충할 수 있는 새로운 네트워크 패러다임 개발이 요구됨.

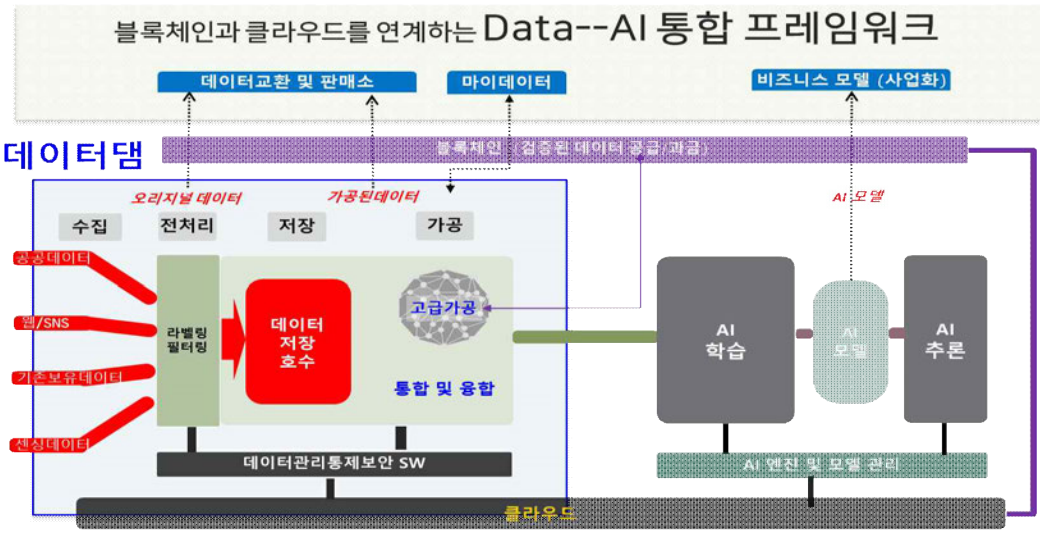


그림 2. D-N-A 개념에 따른 융복합 서비스 창출을 지원하는 통합 프레임워크

[DID 블록체인 기반 IoT 인증 및 보안 기술 개발]

○ **(IoT 보안의 필요성)** IoT 기술은 디지털 전환을 위한 필수적인 기술임. IoT 네트워크에 대한 안전과 신뢰를 위해서는 인증, 허가 기술이 필수적이어야 함. 인증, 허가를 위해서는 높은 수준의 암호화 방식을 사용해야 하지만, IoT 기기는 제한된 자원(배터리, 계산 자원 등)을 가지기 때문에 많은 계산비용을 필요로 하는 보안 방식은 적용하기 어려움.

○ **(DID 블록체인 기반 IoT 보안)** 기존 중앙화된 인증 방식은 single-point-failure attack에 취약하다는 단점이 있지만, 블록체인 시스템은 탈중앙화된 방식으로써 시스템의 보안성을 크게 향상시킬 수 있음. 블록체인 기반 보안통신 구현을 위하여, 각 IoT에 고유한 DID를 생성하고 블록체인에 등록하여 IoT 장치 인증을 수행할 수 있음. 기존 IoT 장치 인증은 암호 알고리즘으로 생성된 키를 통해 인증을 수행하게 되지만, 이러한 키 관리 시스템은 높은 계산비용을 필요로하기 때문에 통신지연 등의 문제가 발생할 수 있음. 따라서 대규모 IoT 환경에 적합한 저복잡도를 가지는 DID 생성방식이 필요함.

3) 동향

○ **(국내 블록체인 연구 동향)** 블록체인 보안연구센터 (고려대학교) 에서는 블록체인 안전성, 프라이버시 분석, 퍼블릭・프라이빗 블록체인에 대한 합의 알고리즘 분석 및 설계, 스마트 계약 분석 및 설계, 블록체인 지향 프로그래밍 언어, 블록체인 애플리케이션, 프로토콜 관련 주제를 중점적으로 다루고 있음. 블록체인 & 빅데이터 연구소 (ETRI) 에서는 빅데이터 생성, 데이터 탈중앙화 관리, 인공지능을 활용한 가치추출을 위한 블록체인 애플리케이션・플랫폼 연구가 진행되고 있음.

○ **(생체인증 사용 동향)** 전 세계의 정부 기관은 생체인증을 국가 차원 인프라로 활용하는 추세임. 유엔 산하 국제전기통신연합(ITU), 중앙아프리카 공화국, 적도 기니, 나이지리아, 탄자니아에서 추진중임. 영국은 생체인증을 도입한 전자 여행 허가 시스템을 계획중임. 네덜란드는 Self-sovereign identity 프로젝트를 통해 블록체인과 영지식 증명을 적용한 생체인증방식을 추진하고 있음. 이에 반해, 한국은 금융과 유통 시장에서 일부 생체인증을 도입하고 있지만 걸음마 단계에 머물고 있음. 따라서 한국도 생체 분야 기술력 확보를 위한 노력이 필요함.

○ **(생체정보 보호 기술 동향)** 한국스마트인증에서는 홍채 정보가 유출되지 않고 안전하게 사용 될 수 있는 동형암호기술을 개발함. 블록체인 적용을 통해 출입국 관리 등 대규모 인증 시스템에 적용 할 계획임. 디지털화된 원본 생체정보를 암호학적 방법으로 숨기는 방법들로 Fuzzy commitment, Fuzzy Vault, Fuzzy Extraction, 폐기형 인증, 동형암호 기반 인증 등이 연구됨.

○ **(소형 초분광 이미징 장치 연구 동향)** 학계에서 저가의 소형 분광 영상 장치 개발에 몰두하고 있음. 본 연구팀은 해상도와 크기 문제를 해결하기 위해 압축센싱 이론을 분광기에 적용하여 작고 해상도가 높은 랜덤압축 분광기를 제안함. 최근에는 이차원 필터 배열을 실리콘웨이퍼 증착 방법을 이용하여 제작하고 CMOS 이미지센서와 결합하여 분광장치를 만들고 분광기의 성능과 기존의 분광기의 성능을 비교하는 실험을 수행함. 또한, DL 기반 알고리즘을 개발하여 분광 신호에 대한 사전정보 없이 신호를 복원할 수 있는 지능형 랜덤압축 분광 시스템을 개발함.

○ **(분산된 에지 컴퓨팅 인프라, 이를 중심으로 한 데이터-인공지능의 결합)** 초연결 클라우드 중심의 ICT 인프라를 기반으로 사용자-주도형 AI융합 서비스를 신속하고 경제적으로 실증하도록, 오픈소스 소프트웨어-정의 인프라를 종단-에지-코어 영역에 걸쳐 DevOps(개발/운영병행체제) 자동화가 가능한 형태로 구축하여 제공하는 것이 바람직함. 지난 20여년을 되돌아보면 장기적인 비전에 따른 ICT 인프라 구축과 운영의 변화는 클라우드가 선도한 것으로 분석되며, 클라우드의 확산에 따라 사용자 및 소프트웨어 중심형 인프라가 중앙(코어)에서 종단으로 퍼져나가면서 확산중임. 이러한 ICT 인프라의 코어-종단 확산 과정에서 종단/코어의 접점에 해당하는 에지(edge) 영역의 전략적인 중요성이 증대되었으며, 에지 컴퓨팅·클라우드·지능 등의 키워드들이 모바일, 멀티-액세스, 컴퓨팅, 클라우드 기술과 연결되어 MEC[1] 개념으로 자리 잡고 있음.

1) MEC(Mobile Edge Computing): 사용자 단말과 가장 가까운 통신망 내부 주요 거점에 클라우드 컴퓨팅 형태의 대용량 분산 인프라를 구성하여, 초저지연·초대용량의 고품질 서비스를 제공하는 차세대 네트워킹 핵심 기술.

2 연구개발과제의 목표 및 내용

1) 연구개발과제의 최종 목표

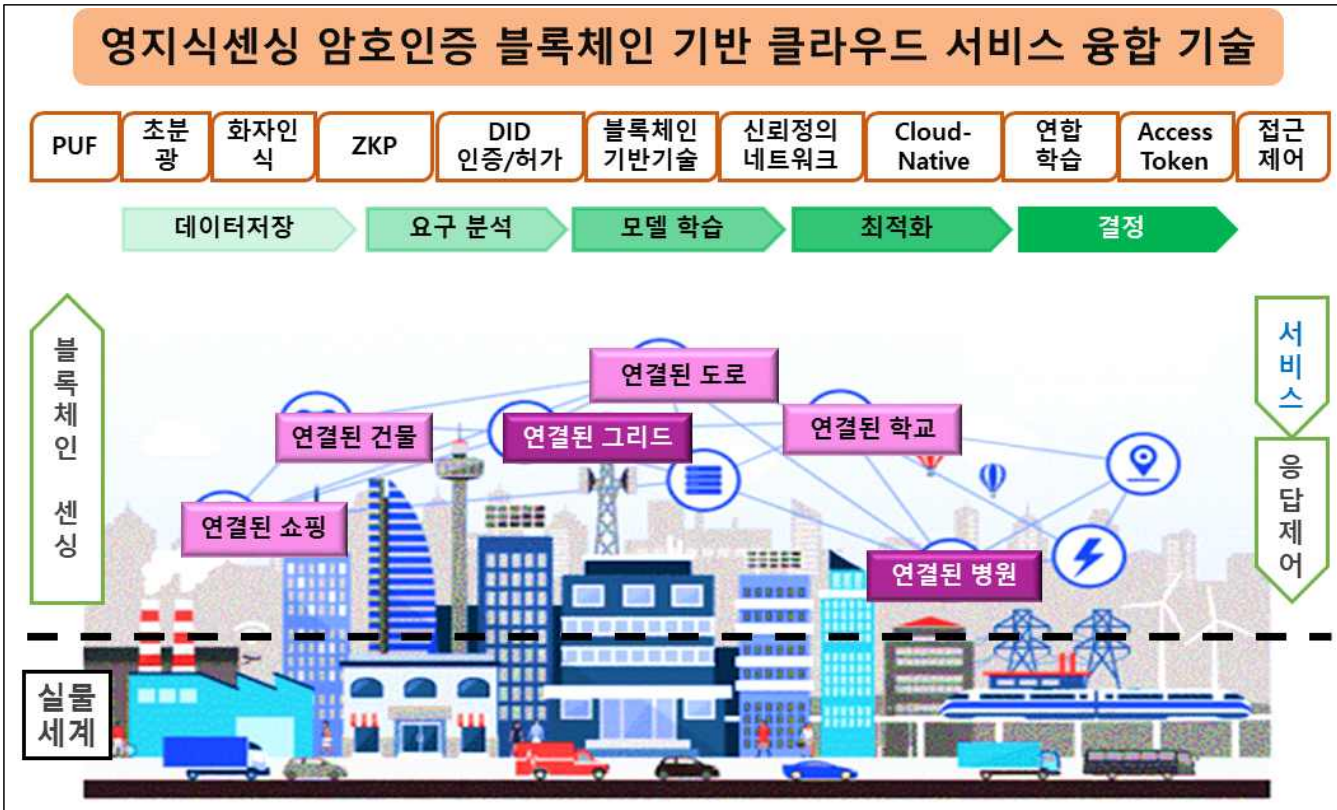


그림 3. 영지식 센싱 암호인증 블록체인 기반 클라우드 서비스 융합 기술 개념도

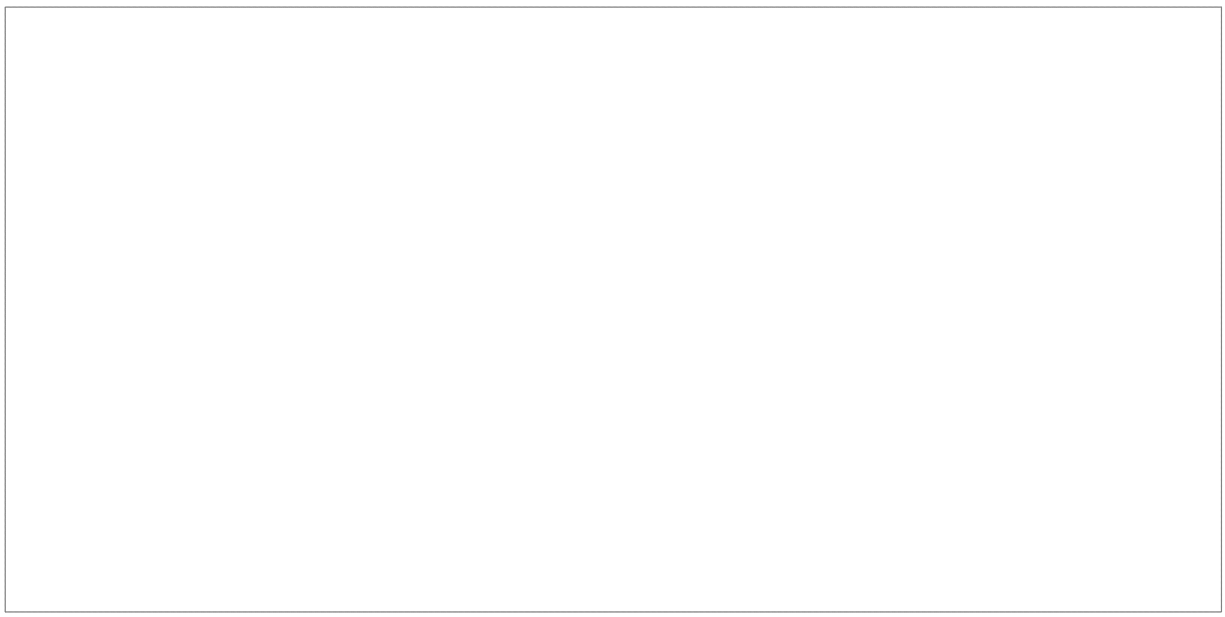


그림 4. 영지식 센싱 암호인증 블록체인 기반 클라우드 서비스 융합 기술 개발 체계도

○ 블록체인 지능 융합 센터의 연구개발과제 및 인력양성 개요는 다음과 같음.

○ 물리적 세상에서 디지털세상으로의 관문은 IoT센서들임. 전파, 음성, 생체신호등 물리적 신호는 신호측정시간, 방위적 위치, 물리계층 채널의 상태등 많은 정보를 내포하고 있음. 이런 물리적 특성을 간직한 정보들은 신호의 고유성을 보존할 수 있고 디지털 인증

이나 주요자원 액세스의 허가시 활용할 수 있음. 그러므로, IoT센서들은 신호 고유성을 잘 보존하면서, 신호를 데이터로 바꾸어야 함.

○ 디지털 세상으로 들어온 센싱정보들은 위변조가 용이한 데이터로 전환되는 문제가 발생하므로, 이것을 방지하기 위해 네트워크 기반 인증 및 블록체인을 활용하고자 함. 블록체인은 동일한 프로토콜을 클라우드 컴퓨터 네트워크에서 실행하여 데이터를 안전하게 저장하고, 순전무결하게 보관하고 관리함. 문제는 블록체인은 데이터를 여러 서버에 분산관리하고 주기적인 상호 대조를 통해 데이터를 위변조를 방지하는 기술이기 때문에, 확장성 문제와 속도제어 문제가 발생함.

○ 이때 활용할 수 있는 기술은 2nd 레이어 블록체인(플라즈마, 샤드, 사이드체인등)과 영지식 증명 (재귀증명, ZK rollup) 등을 쓸 수 있음. 영지식 증명은 현재상태와 미래상태의 변화를 짧은 증명(요약)만을 남김으로 상태변화의 정당성을 확보하므로, 확장성 문제를 개선할 것으로 기대할 수 있고, 또한 과정증명에도 쓰일 수 있음.

○ 과정증명이란, 신호측정 시 물리적 고유성이 블록체인에 증거가 올라가고, 네트워크내에서 이동시에도, 모든 과정이 프로토콜에 정의된 그대로 수행이 되었다는 증거를 작은 용량으로, 남길 수 있음. 그러므로 고도의 신뢰성 확보가 필요한 매우 중요한 정보(DID 혹은 보안이 요구되는 데이터)의 보관 및 관리에 중요한 역할을 할 수 있음.

○ 본 과제는

1) 화자인식, 초분광기를 활용한 생체인식, 물리계층 지식을 활용한 복제 불가능 함수 인식 등 다수의 방식 (multi-modal)으로 인증의 신뢰성과 보안성을 제고하고,

2) 지리적으로 분산된 여러 데이터 폰드, 레이크, 클라우드 등에 보관되어 있는 데이터가 어느 한 지점에서 다른 지점으로 이동할 때, 데이터의 진위성이 깨지지 않도록 에지 컴퓨팅 기발 블록체인 서비스 인프라 기술로 보호하며,

3) 클라우드에 축적된 신뢰성 있는 디지털 데이터로 연합학습, 강화학습, 이전학습 등을 적용하여 모델성능을 제고하고, 서비스 요구사항 분석 및 탄력적 클라우드 리소스 배치 및 관리, 끊김없는 고신뢰 서비스 제공을 위한 주요 지능형 성능 지표를 주기적으로 개선하며,

4) DID를 기반으로 디지털 트윈을 생성하고 안전하게 인증하며 주요자원에 접근을 허가해주는 기술의 개발과 한정된 자원을 가진 IoT 기기를 위한 간결하고 안전한 인증 프로토콜의 개발을 통해,

→ 지능적 자원관리와 효율성과 탄력성이 제고된 블록체인 지능 융합 플랫폼 관리 기술을 개발하는 것임.

○ 본 과제의 최종 목표는 그러므로 다음과 같음:

고신뢰 센싱과 데이터 축적, 지능 모델 학습, DID 인증, 블록체인 기반 클라우드 서비스 플랫폼 및 관리 기술의 개발과 석박사급 고급 인력 양성

○ 위 센서의 연구목표를 이루기 위해서 총 세 개의 세부과제를 구성함

- 제1 세부과제: 영지식 센싱 탈중앙 고성능 분산원장 기술 개발 (혁신도약형)
- 제2 세부과제: Edge 컴퓨팅 기반 블록체인 서비스 인프라 기술
- 제3 세부과제: DID 블록체인 기반 IoT 인증 및 보안 기술 개발

구분	내용
최종목표	**세부과제 1** **최종목표: 영지식 센싱 탈중앙 고성능 분산원장 기술 개발 (혁신도약형)** o DID를 위한 초분광 카메라, 화자인식 기반 ZKP 생체인증기법 개발 o 고성능 합의알고리즘 등 블록체인 기반기술 연구 o 다중 복합 구조 계층형 블록체인 구조 확장 기술 연구 o 블록체인 및 클라우드 상에서 동작하는 연합학습 플랫폼 개발 **세부과제 2** **최종목표: Edge 컴퓨팅 기반 블록체인 서비스 인프라 기술** o Connected DataLake에 기반한 지능형 Edge Infra 기술 개발 o 멀티사이트 에지/코어 클라우드간 유기적 협력에 따른 블록체인 대응 초고속/대용량 데이터 공유 실현 기술 개발 o 블록체인 플랫폼을 위한 신뢰 정의 네트워크 (Trust-Defined Network) 개발 및 분산형 신뢰 정의 네트워크 관리 기술 개발 o 블록체인 서비스 지원을 위한 지능형 TDN 요소기술 통합 구현 및 검증 **세부과제 3** **최종목표: DID 블록체인 기반 IoT 인증 및 보안 기술 개발** o DID 블록체인 기반의 인증 및 안전한 Data 전송 방식 연구 o DID 블록체인 기반의 인가 및 접근제어 방법 연구 o 제한된 하드웨어 자원을 갖는 IoT를 위한 고신뢰/경량 DID 생성 기술 개발 o 물리계층을 활용한 블록체인 기반의 탈중앙화 IoT 기기 인증기술 및 프로토콜 개발
세부목표	**세부과제 1** **o 주요 기능(또는 규격)** - 블록체인 스토리지 경량화 기능 - 블록체인과 상호작용을 위한 데이터 검증 경량화 기능 - 비잔틴 참여자가 있는 상황에서의 효율적인 연합학습 기능 - 입력된 음성의 화자인증 모델 - 공간 정보 추출 및 잡음 제거를 통한 화자 신호 강조를 수행하는 화자인증 전처리 모델 **o 주요 성능치**

구분	내용

핵심기술/제품 성능지표		단위	달성목표	국내 최고수준	세계최고수준 (보유국, 기업/기관명)
ZKP 기술	1. 증명 프로토콜의 증거의 길이	kB	5	-	5 (미국/UC Berkeley, Fractal)
	2. 증명 프로토콜의 증거의 생성 시간	초	5	-	5 (미국/UC Berkeley, Fractal)
	3. 증명 프로토콜의 증거의 검증 시간	초	0.004	-	0.004 (미국/UC Berkeley, Fractal)
딥러닝 기반 화자인증 기술	1. 딥러닝 기반 음성 화자인증 모델 정확도	에러율	2%	-	1.4% (러시아, STC-innovations)
	2. 화자인증전처리 모델의 음성향상도	PESQ	3.20	-	3.12 (미국, Chin-Hui Lee 연구그룹)
블록체인 기반 기술	AWS 인스턴스 2000개 노드로 구성	TPS	1000	N/A	900 (미국, Algoland)
	글로벌 전체 네트워크 처리 성능	TPS	10,000	N/A	7,380 (미국, RapidChain)
	합의지연	초	15	N/A	10 ~ 22 (미국, Algoland; 미국, RapidChain)
	블록체인 스토리지 데이터 감소 비율	%	10~20%	-	-

주요 성능 지표 두 개

1. TPS 성능 수치는 합의체 노드 수, 블록 크기 등의 패러미터에 따라 달라질 수 있는데, Algorand는 2,000 노드 합의체 네트워크를 대상으로 최적의 조건에서 약 900 TPS에 근접하는 성능을 달성함.

* 참고 : Algorand: Scaling Byzantine Agreements for Cryptocurrencies," Proceedings of the SOSP, Oct. 2017.

2. 병렬처리 방식의 합의 프로토콜의 TPS 성능 수치는 전체 노드 수, 블록 크기 등의 패러미터에 따라 달라질 수 있는데, RapidChain은 4,000 노드 네트워크에서 샤드 크기가 250일 때 7,380 TPS 성능을 달성함.

* 참고 : RapidChain: Scaling Blockchain via Full Sharding," Proceedings of the CCS, Oct. 2018.

o 핵심 기술

- 고성능 분산 합의알고리즘 기술 (ECCPoW, PoS, DPoS, PoA, BPF계열)
- 블록체인 비활성 상태정보 데이터 제거를 통한 스토리지 경량화 기술
- 신뢰실행환경과 블록 헤더 검증을 통한 블록체인 데이터 검증 경량화 기술
- 신뢰실행환경을 통한 오프체인에서의 트랜잭션 처리량 향상 기술
- 비잔틴 저항 인공지능 모델 연합학습 방법론 기술
- 딥러닝 기반 음성 화자인증 기술
- 화자인증 전처리를 위한 다채널 잡음 제거기술

o 적용범위(또는 서비스)

- 효율적 스토리지 및 검증을 통한 블록체인 기반 디지털 트윈 플랫폼에 활용 가능
- IoT 기반 네트워크에서의 탈중앙화 인공지능 모델 학습에 활용 가능
- 탈중앙화된 신원 증명 시스템 신뢰도 향상에 활용 가능
- 다양한 음성 기반 애플리케이션의 성능향상을 위한 전처리로써 활용 가능

구분	내용

세부과제 2

o 주요 기능(또는 규격)

- 블록체인 대응 Connected DataLake의 Cloud-native 및 멀티사이트 지원
- 블록체인 대응 Connected DataLake 구성을 위한 데이터 네트워킹 오프로딩 기능
- 블록체인 신뢰도 기반 Cloud-native Connected DataLake 데이터 공유 기능
- 네트워크 구성 요소에 대한 신뢰도 측정 모델
- 신뢰도 기반에 따른 동적 네트워크 제어 기능
- 결정적 추적 가능 데이터 스트림 전송 보장 기능
- 신뢰 기반 네트워크 내의 부하 추적 및 분산 최적화 기능

o 주요 성능치

핵심기술/제품 성능지표		단위	달성목표	국내 최고수준	세계최고수준 (보유국, 기업/기관명)
Connected DataLake에 기반한 지능형 Edge Infra	1. Cloud-native/멀티사이트 에지 지원 여부	-	지원	부분 지원	지원 (미국, SAS)
	2. 사이트간 데이터 네트워킹 오프로딩 속도	Gbps	100	-	100Gbps (미국, IBM Aspera)
	3. 신뢰도 기반 데이터 공유(동기화) 용량	TB/Day	10	-	100TB/Day (미국, IBM Aspera)
신뢰 기반 네트워크 구성 기술	1. 네트워크 전송 후 원본과의 데이터 오차 비율	%	1 (이하)	-	-
	2. 신뢰 정의 네트워크에서의 데이터 전송 비율	[0,1]	0.95	-	-
	3. 신뢰 정의 네트워크 적용 사례	건	1건 (이상)	-	-

<신뢰 기반 네트워크 구성 기술 주요 지표 항목 설명>

1. 신뢰 정의 네트워크를 통해 IoT 센싱 데이터와 디지털 트윈 간에 전달된 데이터 간 오차 비율.
2. 신뢰 정의 네트워크 내에서 IoT 디바이스, 블록체인 노드, 디지털 트윈 시스템 사이에서 패킷이 성공적으로 전달될 비율 (delivery ratio)
3. 신뢰 정의 네트워크 적용을 통해 개선한 네트워크 서비스 사례

o 핵심 기술

- Cloud-native 스토리지 소프트웨어를 적용한 Connected DataLake 통합관리 기술
- 멀티사이트 에지 클라우드 연계형 블록체인 대응 Connected DataLake 운영 기술
- Connected DataLake 간의 다중경로 초고속/대용량 데이터 전달 기술
- 신뢰 정의 네트워크에 기반한 오버레이 네트워크 구성 기술
- 데이터 무결성을 위한 시민감 데이터 전송 기술
- 가십형 프로토콜 및 분산형 네트워크 컨트롤러 기술
- 지능형 네트워크를 위한 동적 멀티캐스팅 기술

o 적용범위(또는 서비스)

- 에지 인프라 기반 신뢰성 있는 데이터 공유 (저장 및 교환)의 기반 구축에 활용
- 신뢰 데이터에 기반한 D-N-A 개념의 미래형 AI융합 서비스의 안정성 제고
- 블록체인 운영을 지원하기 위한 고신뢰 데이터 스트림 인프라 구축
- 블록체인 플랫폼을 위한 오버레이 네트워크 구축
- 인공지능 학습용 고신뢰 데이터 파이프라인 구축

<table>
<tr><th>구분</th><th>내용</th></tr>
<tr><td></td><td>
- 기타 빅데이터 플랫폼의 시큐어 데이터 전송 경로 구축
세부과제 3
o 주요 기능(또는 규격)
- 영지식 증명의 자격증명 기능
- DID 블록체인 기반의 인증/인가 정책 생성 기능
- Access Token 관리 및 사용 기능
- IoT 고유의 특성을 활용한 보안키 생성 기능
- 제한된 하드웨어 자원을 갖는 IoT에 적합한 경량 DID 생성 기능
- 환경에 따라 변화하는 동적 DID 생성 기능
- 블록체인의 DID 보안 요구사항을 반영한 DID 인증 프로토콜
- 동적 DID에 적합한 DID 제어 프로토콜
o 주요 성능치
<table>
<tr><th colspan="2">핵심기술/제품
성능지표</th><th>단위</th><th>달성목표</th><th>국내최고수준</th><th>세계최고수준
(보유국,
기업/기관명)</th></tr>
<tr><td rowspan="4">DID기반
인증/인가
기술</td><td>1. DID 블록체인 네트워크 반영 시간</td><td>TPS</td><td>2,000</td><td>-</td><td>2,700
미국/Hyperledger</td></tr>
<tr><td>2. DID 블록체인 기반의 인증 검증 시간</td><td>ms</td><td>1,000</td><td>-</td><td>-</td></tr>
<tr><td>3. DID 블록체인 기반의 인가 검증 시간</td><td>ms</td><td>1,000</td><td>-</td><td>-</td></tr>
<tr><td>4. 동적 DID의 인증 오류율</td><td>%</td><td>≤0.15</td><td>-</td><td>LTE의 장치 인식 오류율:
약 0.2 %</td></tr>
<tr><td>DID 응용
시스템
테스팅</td><td>시나리오에 따른 DID기반의 디지털 트윈 응용 검증</td><td>use-case</td><td>1건 이상</td><td>-</td><td>-</td></tr>
<tr><td rowspan="2">복제불가능한
DID 기반
인증 기술</td><td>1. 분산 네트워크 내 DID 유일성 검증</td><td>%</td><td>~50%</td><td>ICTK holdings:
49.99%</td><td>한국/ICTK holdings:
49.99%</td></tr>
<tr><td>2. 분산 네트워크 내 DID 무작위성 평가</td><td>randomness
(normalized hamming distance)</td><td>≥0.49</td><td>ICTK holdings:
0.4972</td><td>한국/ICTK holdings:
0.4972</td></tr>
</table>
o 핵심 기술
- DID 블록체인 기반의 인증/인가 단일화 기술
- 안정성이 보장된 OAuth2.0기반의 Access Token 관리 기술
- DID 인증/인가 기반 블록체인 응용 시스템 접근제어 기술
- 분산 네트워크 환경에서 유일성을 가지는 DID 생성 기술
- 분산 네트워크 통신 환경에 따른 동적 DID 생성 및 업데이트 기술 (세계 최초)
o 적용범위(또는 서비스)
- DID 블록체인 기반의 IoT 네트워크 환경
- 안전한 IoT Data를 위한 인증/인가 기술이 필요한 서비스
- 경량화 기술이 요구되는 드론 등에 DID 인증 서비스 적용 가능
- 스마트 그리드, 스마트 팩토리 등 고신뢰 무선 보안 통신이 요구되는 산업에서 사용할 수 있는 무선 네트워크 보안 인증 서비스 제공
</td></tr>
</table>

2) 연구개발과제의 단계별 목표

① 1단계: 2021. 07. ~ 2022. 12.

구분	내용
단계 목표	**세부과제 1** o 신원 인증 키 (초분광이미지, 화자) 비가역적 아날로그-디지털 스캐너 연구 o 강건하고 보안성 있는 분산합의 기술 분석 및 개발 o 초경량ㆍ초확장 블록체인 개발 **세부과제 2** o 블록체인 대응 Cloud-native Connected DataLake 연구 o 블록체인 데이터 스트림 전송을 위한 신뢰 정의 네트워크 구조 연구 **세부과제 3** o DID 블록체인 기반의 인증 및 안전한 Data 전송 방식 연구 o 물리계층 고유성을 활용한 DID 기술 개발
세부 내용	**세부과제 1** o 신원 인증 키 (초분광이미지, 화자) 비가역적 아날로그-디지털 스캐너 연구 - 아날로그 신원 인증 키 연구 - 비가역적 아날로그-디지털 스캐너 설계 - 초분광압축이미징 장치 구현 및 데이터 확보 - 대용량 화자 데이터 확보, 분류, 정제 및 특징추출 - 최신 딥러닝 기반 화자인증 베이스라인 구축 및 기초 모델 개발 o 강건하고 보안성 있는 분산합의 기술 분석 및 개발 - 분산 합의 기술 분석 및 특징 파악 - 이종 분산합의 기술에 의한 환경 영향 분석 - 강건한 블록체인 인프라를 위한 분산합의 기술 분석 및 개발 - 트랜잭션 크기 별 안전 거래 분석 알고리즘 개발 o 초경량ㆍ초확장 블록체인 개발 - 초경량 블록체인 기술 개발 - 초확장 블록체인 기술 설계 - 신뢰실행환경 기반 데이터 검증 경량화 기술 개발 **세부과제 2** o 블록체인 대응 Cloud-native Connected DataLake 연구 - 블록체인 대응 Cloud-native Connected DataLake 구성 설계 - 블록체인 대응 Cloud-native Connected DataLake PoC 검증 o 블록체인 데이터 스트림 전송을 위한 신뢰 정의 네트워크(TDN) 구조 연구 - 신뢰 정의 네트워크 구조 정의 및 데이터 제어 방식 구성 - 지속 가능한 데이터 스트림 경로의 신뢰도 수준 등급화 및 이를 기반으로 트래픽 분류 및 처리 방법 모델링 **세부과제 3** o DID 블록체인 기반의 인증 및 안전한 Data 전송 방식 연구 - DID 기반 인증/인가 서비스 요구사항 분석 - IoT에 적합한 DID 블록체인 인증 기법 연구

구분	내용
	- 안전한 IoT Data 전송 방법 연구 o 물리계층 고유성을 활용한 DID 기술 개발 - 블록체인 기반 보안 시스템을 위한 사전연구 - IoT의 물리적 특성을 활용한 DID 생성기술 개발 - DID 블록체인 연계 장치 인증 프로토콜 연구

② 2단계: 2023. 01. ~ 2026. 12.

구분	내용
단계 목표	**세부과제 1** o 영지식 증명 및 검증 프로토콜 연구 (초분광 이미지, 화자인증) o 다중 복합구조 계층형 블록체인 플랫폼 기술 개발 및 테스트 o 빅데이터 및 분산형 인프라를 이용한 분산 애플리케이션 개발 o 데이터 프라이버시를 보존할 수 있는 연합학습 방법론 개발 **세부과제 2** o 블록체인 대응 체계화된 데이터 저장·교환을 지원하는 멀티사이트 Connected DataLake 구현 및 개선 o 다중경로 초고속 데이터 전달 기능 개발을 통한 블록체인 대응 Connected DataLake 개선 o 블록체인 플랫폼 지원을 위한 TDN 구성기술 개발 **세부과제 3** o DID 블록체인 기반의 인가 및 접근제어 방법 연구 o 복제불가능한 DID 기반 보안기술 개발
세부 내용	**세부과제 1** o 영지식 증명 및 검증 프로토콜 연구 - 영지식 증명 및 검증 프로토콜 설계, 성능분석 및 취약점 분석 - 초분광 이미지에서 비정상 데이터를 인지/추출/분류 알고리즘 개발 - 초분광 이미지를 복원하는 고급 딥러닝 기법 연구 - 시간 영역 Raw 데이터를 이용한 딥러닝 기반 고신뢰 화자인증 기술 개발 - 신뢰도 향상을 위한 화자인증에 특화된 추가 모듈 개발 - 원거리 상황, 잡음 환경 및 짧은 음성 입력에 강인한 모델 개발 o 다중 복합구조 계층형 블록체인 플랫폼 기술 개발 및 테스트 - 다중 복합구조 블록체인 연구 - 블록체인·샤드·클라우드 간의 통신 프로토콜 및 플랫폼 연동 기술 연구 o 빅데이터 및 분산형 인프라를 이용한 분산 애플리케이션 개발 - 블록체인 기반 빅데이터 생성 및 관리 기술 개발 - 분산형 애플리케이션을 위한 비즈니스 로직 및 스마트 계약 기술 개발 o 데이터 프라이버시를 보존할 수 있는 연합학습 방법론 개발 - 탈중앙화 연합학습 플랫폼 설계 - 비잔틴 내구성 연합학습 방법론 설계 - 블록체인 기반 연합학습 플랫폼 개발 - IoT 기기를 위한 인공지능 모델 경량화

구분	내용
	세부과제 2 o 블록체인 대응 체계화된 데이터 저장·교환을 지원하는 멀티사이트 Connected DataLake 연구 - 멀티사이트 Connected DataLake를 위한 블록체인 대응 체계화된 데이터 저장·교환 기능 구현 - 블록체인 대응 체계화된 데이터 저장·교환을 지원하는 멀티사이트 Connected DataLake 구현 - 블록체인 대응 Connected DataLake를 위한 초고속 데이터 전달 기능 개발 - 다중경로 초고속 데이터 전달 기능 개발을 통한 블록체인 대응 Connected DataLake 개선 o TDN 구성을 통해 블록체인 데이터 스트림을 위한 네트워크 구성 기술 개발 - 신뢰도 기반의 네트워크 구현 및 이에 최적화된 오버레이 네트워크 지원 - 블록체인 노드 간의 선택적 및 지능적 멀티캐스팅을 TDN 측면에서 지원 - 시민감 데이터 전송을 보장하는 동기화 네트워크를 확장한 Time Sensitive Networking(TSN) 기능 지원 - 장애 허용성을 가지는 분산형 TDN 컨트롤러 기반 효율적인 트래픽 분산 방안 연구 **세부과제 3** o DID 블록체인 기반의 인가 및 접근제어 방법 연구 - DID기반의 자기주권형 액세스 토큰 발행 연구 - IoT 환경의 DID기반 접근 제어 방법 연구 - 영지식 증명 기반의 자격증명 방법 연구 o 복제불가능한 DID 기반 보안기술 개발 - 대규모 IoT 환경에서 신뢰성을 가지는 동적 DID 생성 기술 개발 - 블록체인 연계 동적 DID 기반 저전력/고신뢰 보안인증 프로토콜 연구 - 소규모 private 블록체인을 활용한 보안통신 기술 검증

③ 3단계: 2027. 01. ~ 2028. 12.

구분	내용
단계 목표	**세부과제 1** o 영지식 아날로그 생체인증 시스템 연구 o 다중 복합구조 계층형 블록체인 플랫폼 기술 운용 **세부과제 2** o 분산된 데이터 셋 동기 및 공유 기능을 통합한 블록체인 대응 Connected DataLake 고도화 o 블록체인 서비스와 연동한 TDN PoC 레벨 구현 **세부과제 3** o DID 인증/인가 기반 블록체인 운용 시스템 통합/평가 o DID 기반 블록 생성 및 검증 기술 연구
세부 내용	**세부과제 1** o 영지식 아날로그 생체인증 시스템 연구 - 영지식 아날로그 생체인증 시스템 개발, 테스트베드 구축 및 평가 - 개발한 모델과 베이스라인의 화자인증 성능 비교를 통한 검증 - 개발한 모델 최적화 및 고도화, 경량화 연구 진행

구분	내용
	o 다중 복합구조 계층형 블록체인 플랫폼 기술 운용 - 사용 용도별 블록전파지연시간, 참여노드 수에 따른 분산원장 파라미터 최적화 - 샤드, 2nd 레이어(샤드, zk rollup), 크로스체인(cross-chain) 기술 융합 운용 - 이종 IoT 네트워크 설계 및 크로스체인 기술 개발, 통합 기술 개발 **세부과제 2** o 분산된 데이터 셋 동기 및 공유 기능을 통합한 블록체인 대응 Connected DataLake 고도화 - 블록체인 대응 Connected DataLake 고도화를 위한 분산된 데이터셋 동기 기능 개발 o 블록체인 서비스와 연동한 TDN PoC 레벨 구현 - TDN 요소 기술 통합 및 테스트베드 개발 - TDN 시스템 통합 및 PoC 레벨 구현 **세부과제 3** o DID 인증/인가 기반 블록체인 응용 시스템 통합/평가 - DID 인증/인가 기반 블록체인 응용 연구/개발 - DID 인증/인가 기반 블록체인 응용 시스템 성능 측정/평가 o DID 기반 블록 생성 및 검증 기술 연구 - 복제불가능한 DID 기반 DID 문서 생성 및 제어 기법 개발 - IoT를 활용한 테스트베드 구성 및 성능 평가

3) 연구개발과제의 내용

① 세부과제 1 : 영지식 센싱 탈중앙 고성능 분산원장 기술 개발 (혁신도약형)

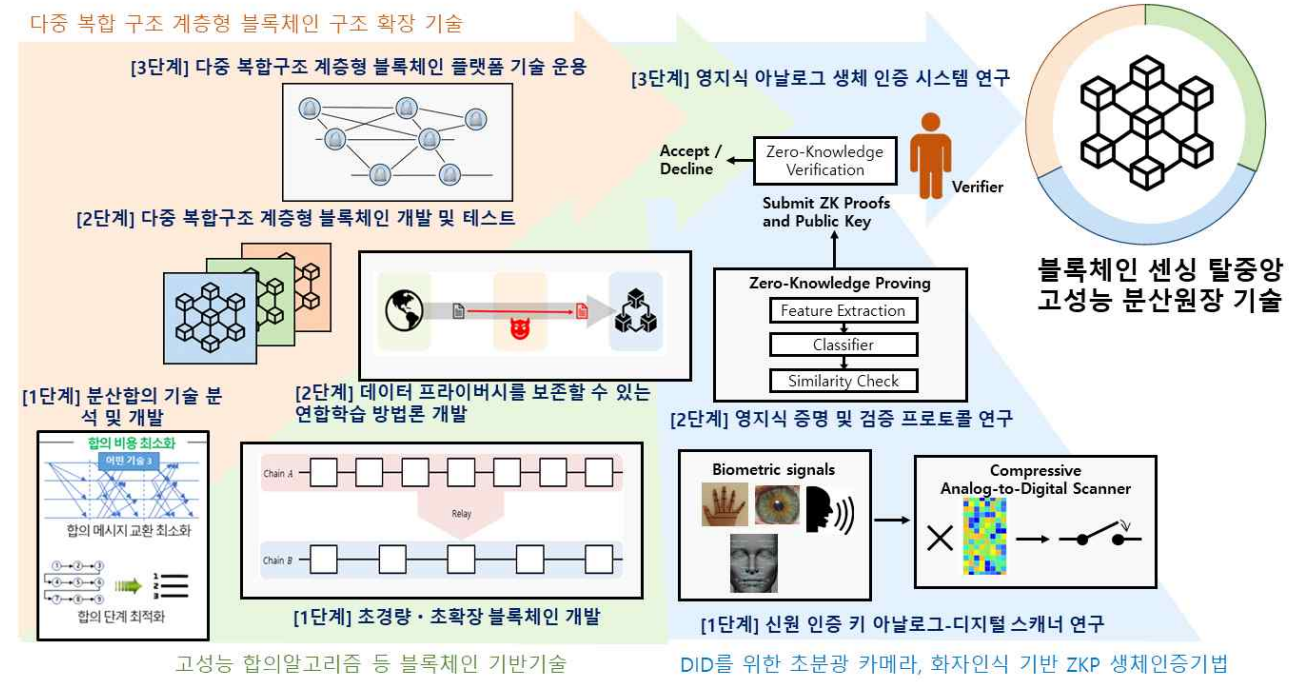


그림 5. 영지식 센싱 탈중앙 고성능 분산원장 기술 개발

가. 1단계 개발 내용 및 범위

a) 광주과학기술원 이홍노 교수 연구팀 (세부과제1 책임)

- 신원 인증 키 (초분광이미지, 화자) 비가역적 아날로그-디지털 스캐너 연구
  - ▸아날로그 신원 인증 키 연구
  - ▸비가역적 아날로그-디지털 스캐너 설계
  - ▸초분광압축이미징 장치 구현 및 데이터 확보
  - ▸최신딥러닝 기반 화자인증 베이스라인 구축 및 기초 모델 개발
- 강건하고 보안성 있는 분산합의 기술 분석 및 개발
  - ▸분산 합의 기술 분석 및 특징 파악
  - ▸이종 분산합의 기술에 의한 환경 영향 분석
  - ▸강건한 블록체인 인프라를 위한 분산합의 기술 분석 및 개발
  - ▸트랜잭션 크기 별 안전 거래 분석 알고리즘 개발

b) 서울대학교 문수묵 교수 연구팀

- 초경량 · 초확장 블록체인 개발
  - ▸초확장 블록체인 기술 설계
  - ▸경량화된 스토리지와 검증 방법론을 가지는 라이트 클라이언트 연구
  - ▸신뢰실행환경에 기반한 블록 검증 연산 부담 경감 연구

c) 광주과학기술원 신종원 교수 연구팀

- 화자인증 시스템 학습을 위한 데이터 확보 및 딥러닝 기반 화자인증 베이스라인 구축
  - ▸VoxCeleb1, VoxCeleb2 등 데이터셋과 유튜브 영상 추출을 통한 데이터셋 확보 및 구조화
  - ▸화자의 음성 정보를 임의로 선별 및 정제 과정 수행
  - ▸시간영역 데이터를 이용하여 입력 음성에 추가적인 전처리 기법 연구
  - ▸Tex-dependent/independent 방식에 따른 화자인증 기술 조사
  - ▸딥러닝 기술을 이용하여 탈중앙화된 신원증명 화자인증 모델 구현 및 베이스라인 구축
  - ▸딥러닝 기반 화자인증 기초 모델 개발 및 성능 확보

나. 2단계 개발 내용 및 범위

a) 광주과학기술원 이홍노 교수 연구팀

- 영지식 증명 및 검증 프로토콜 연구
  - ▸영지식 증명 및 검증 프로토콜 설계, 성능분석 및 취약점 분석
  - ▸초분광 이미지에서 비정상 데이터를 인지/추출/분류 알고리즘 개발
  - ▸초분광 이미지를 복원하는 고급 딥러닝 기법 연구
  - ▸원거리 상황, 잡음 환경에 강인한 모델 개발
- 다중 복합구조 계층형 블록체인 플랫폼 기술 개발
  - ▸다중 복합구조 블록체인 연구
  - ▸블록체인·클라우드 간의 통신 프로토콜 및 플랫폼 연동 기술 연구
- 빅데이터 및 분산형 인프라를 이용한 분산 애플리케이션 개발
  - ▸블록체인 기반 빅데이터 생성 및 관리 기술 개발
  - ▸분산형 애플리케이션 및 빅데이터를 위한 비즈니스 로직 및 스마트 계약 기술 개발

b) 서울대학교 문수묵 교수 연구팀

- 데이터 프라이버시를 보존하며 인공지능 모델을 학습하는 연합학습 기술
  - ▸데이터 프라이버시를 보존하며 인공지능 모델을 고도화하는 플랫폼 개발
  - ▸탈중앙화 비잔틴 저항성 연합학습 시스템 구현

c) 광주과학기술원 신종원 교수 연구팀

- 고신뢰 화자인증 모델 개발
  - ▸시간영역 Raw 데이터를 이용한 딥러닝 기반 화자인증 기술 개발 및 성능 제고
  - ▸다채널 처리 기법을 활용하여 배경 잡음에 강건한 화자인증기술 개발
  - ▸짧은 음성 입력신호를 통한 화자인증 성능 확보 연구

다. 3단계 개발 내용 및 범위

a) 광주과학기술원 이홍노 교수 연구팀 (세부과제1 책임)

- 영지식 아날로그 생체인증 시스템 연구
  - ▸영지식 아날로그 생체인증 시스템 개발, 테스트베드 구축 및 평가
  - ▸개발한 모델과 베이스라인의 인증 성능 비교를 통한 검증
  - ▸개발한 모델 최적화 및 고도화, 경량화 연구
- 다중 복합구조 계층형 블록체인 플랫폼 기술 운용
  - ▸사용 용도별 블록 전파 지연시간, 참여노드 수에 따른 분산원장 파라미터 최적화

b) 서울대학교 문수묵 교수 연구팀

- 이종 네트워크 간 상호작용을 위한 크로스체인 기술 개발
  - ▸서로 다른 디지털 도메인을 갖는 블록체인 통합 연구
  - ▸2-way-relay 기술을 사용한 크로스체인 기술 개발
  - ▸자산 및 데이터의 호환이 가능한 통합 IoT 플랫폼 기술 개발

c) 광주과학기술원 신종원 교수 연구팀

- 개발된 고신뢰 화자인증 시스템 검증
  - ▸베이스라인과 개발된 화자인증 모델을 검증
  - ▸실측 데이터를 적용하여 각 모델 최적화
  - ▸실용적 활용을 고려한 모델 파라미터 최적화 연구

② 세부과제 2 : Edge 컴퓨팅 기반 블록체인 서비스 인프라 기술

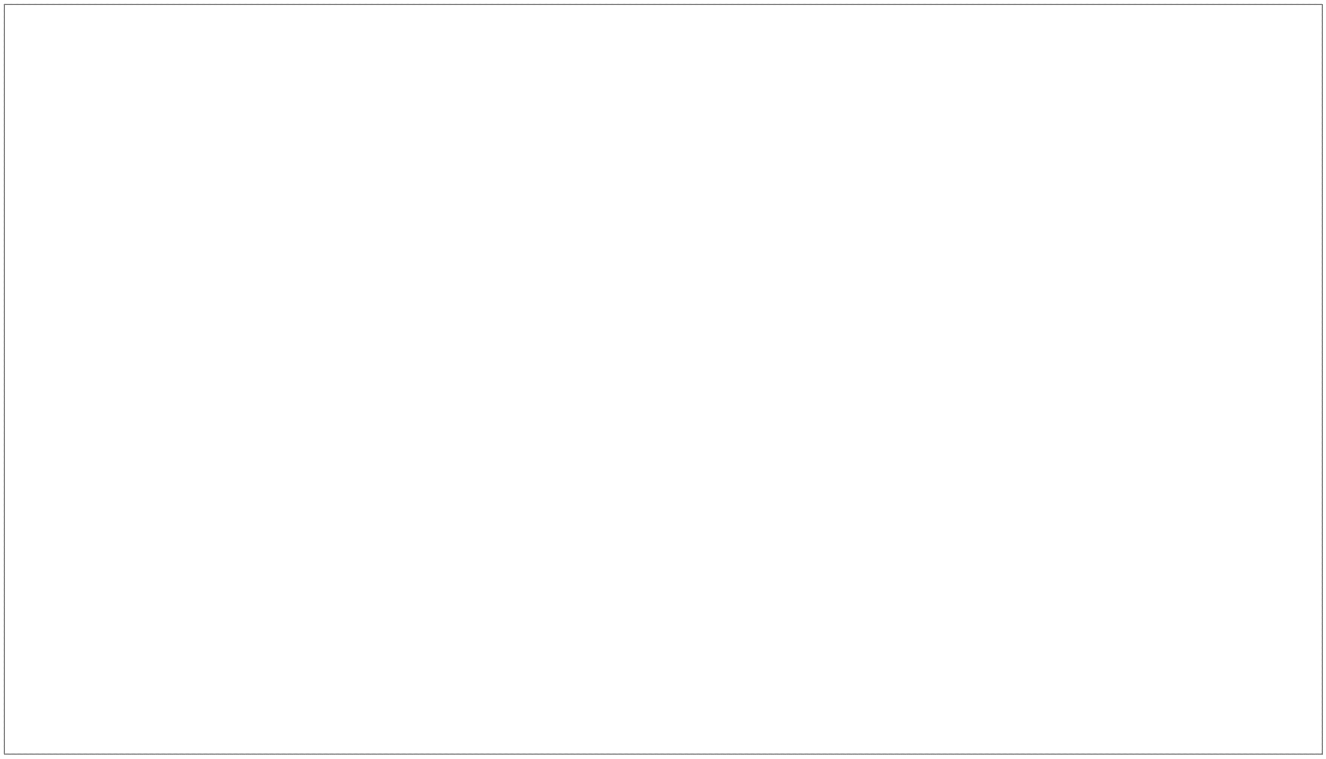


그림 6. Edge 컴퓨팅 기반 블록체인 서비스 인프라 기술

가. 1단계 내용 및 범위

a) 광주과학기술원 김종원 교수 연구팀 (세부과제2 책임)

- 블록체인 대응 Cloud-native Connected DataLake 구성 설계
- 블록체인 대응 Cloud-native Connected DataLake PoC 검증

b) 고려대학교 김황남 교수 연구팀

- 블록체인에 기반한 차세대 디지털 트윈에 대한 유사 사례와 장단점 분석 및 조사
- 블록체인 데이터 스트림 지원을 위한 신뢰 정의 네트워크 요소 기술 선정
  - ▸신뢰 정의 네트워크 구축에 필요한 관련 기술 조사
  - ▸네트워크 구성요소 선별을 위한 신뢰도 평가 기술 개발
  - ▸신뢰 정의 네트워크 구현을 위한 기존 네트워크 구성요소 운영 개선점 도출
- 데이터 스트림의 지속 가능성을 보장하기 위한 신뢰 기반 네트워크 제어 연구

나. 2단계 내용 및 범위

a) 광주과학기술원 김종원 교수 연구팀

- 멀티사이트 Connected DataLake를 위한 블록체인 대응 체계화된 데이터 저장·교환 기능 구현
- 블록체인 대응 체계화된 데이터 저장·교환을 지원하는 멀티사이트 Connected DataLake 구현
- 블록체인 대응 Connected DataLake를 위한 초고속 데이터 전달 기능 개발
- 다중경로 초고속 데이터 전달 기능 개발을 통한 블록체인 대응 Connected DataLake 개선

b) 고려대학교 김황남 교수 연구팀

- 네트워크 구성요소 평가를 위한 전송된 제어 데이터에 대한 신뢰도 평가 기술 개발
- TDN 구성을 통해 지속 가능한 디지털 트윈 수행에 필요한 스트림용 네트워크 구성 기술 개발
- 블록체인 플랫폼을 위한 오버레이 네트워크 지원 기술 개발 및 최적화
- 외부 위협 발생 가능성 최소화와 블록체인 데이터 무결한 전송 지원 기술 개발
- 선택적 및 지능적 멀티캐스팅을 신뢰 정의 네트워크 측면에서 구현
- 시민감 메시지 전송을 보장하는 Time Sensitive Networking (TSN) 기능 개발
- 적응형/가십형 프로토콜 및 분산형 TDN 컨트롤러 기반 효율적인 트래픽 분산 방안 연구
- TDN 내의 부하 추적 및 기계학습 기반 지능형 부한 분산 최적화 기술 연구

다. 3단계 내용 및 범위

a) 광주과학기술원 김종원 교수 연구팀

- 블록체인 대응 Connected DataLake 고도화를 위한 분산된 데이터셋 동기 기능 개발
- 분산된 데이터셋 동기 및 공유 기능을 통합한 블록체인 대응 Connected DataLake 고도화

b) 고려대학교 김황남 교수 연구팀

- 블록체인 서비스와 연동한 TDN PoC 레벨 구현
- TDN 구조 기반 블록체인 검증 데이터 스트림 지속/안정적으로 블록체인 노드에 전송하는 기능 검증
- 확립된 TDN 구조 기반 PoC 형태의 코드 레벨 모듈화

③ 세부과제 3 : DID 블록체인 기반 IoT 인증 및 보안 기술 개발

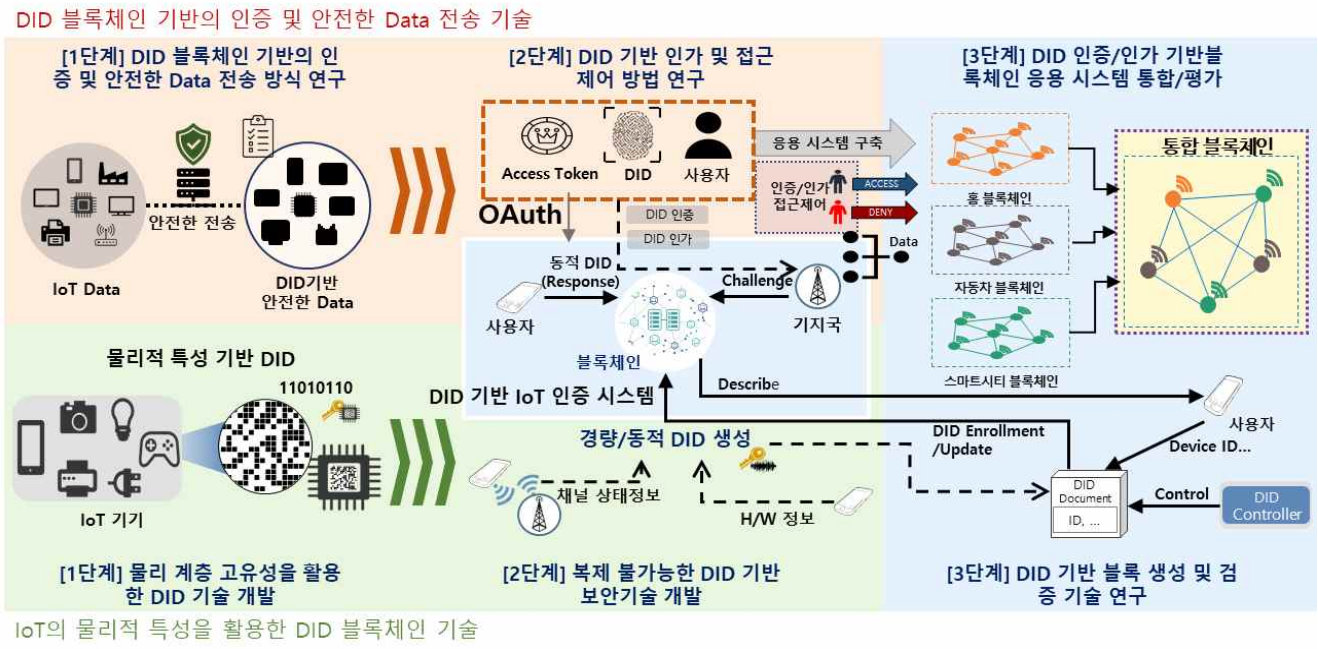


그림 7. DID 블록체인 기반 IoT 인증 및 보안 기술

가. 1단계 개발 내용 및 범위

a) 아주대학교 김기형 교수 연구팀 (세부과제3 책임)

- IoT 생태계에 최적화된 DID 블록체인 기반의 인증/인가 기술 연구
- DID 블록체인 기반의 안전한 서비스 요구사항 연구
- 안전한 IoT Data 전송 기법 연구
  - ▸Data 위변조 방지기술을 연구
  - ▸DID 블록체인 보안성 강화 연구

b) 광주과학기술원 황의석 교수 연구팀

- 보안통신을 위한 기존 암호화 방식의 보안취약점 분석
- IoT의 물리적 특성을 활용한 DID 생성기술 개발
  - ▸다양한 하드웨어의 고유한 물리적 정보(IC chip 등) 수집 및 분석
  - ▸RF-PUF 및 device-PUF를 활용한 경량 DID 생성 알고리즘 개발 및 보안 성능검증
- DID 블록체인 연계 IoT 인증 프로토콜 연구
  - ▸블록체인의 DID 보안 요구사항을 반영한 DID 인증 프로토콜 설계

나. 2단계 개발 내용 및 범위

a) 아주대학교 김기형 교수 연구팀

- DID 블록체인 기반의 자기주권형 액세스 토큰 발급구조 연구
  - ▸OAuth2.0에서 인증/인가에 사용되는 액세스 토큰의 개선된 발급 · 관리구조를 연구

▸자기주권형 액세스 토큰의 저장 기법을 연구

- DID 블록체인 기반의 IoT 단말 접근제어 기법 연구
  ▸IoT 환경에 적합한 DID Document 구조 연구
  ▸네트워크 외부에서도 식별 및 활용이 가능한 안전한 인증/인가 및 접근제어기법의 연구

b) 광주과학기술원 황의석 교수 연구팀

- 대규모 IoT 환경에서 신뢰성을 가지는 DID 생성 기술 개발
  ▸환경에 따라 변화하는 채널 상태 정보 등을 활용한 동적 DID 생성 기술 개발
  ▸구축된 secure 전송 채널을 이용하여 DID 인증 업데이트를 통한 보안성 강화 연구
- 블록체인 연계 동적 DID 기반 저전력/고신뢰 보안인증 프로토콜 연구
  ▸다수의 IoT의 DID 등록/관리/인증을 위한 프로토콜 개발
  ▸한정된 배터리 자원을 가지는 IoT를 위한 동적 DID 업데이트 프로토콜 개발
- 소규모 private 블록체인 연계 보안통신 기술 검증

다. 3단계 개발 내용 및 범위

a) 아주대학교 김기형 교수 연구팀

- DID 블록체인에 저장되는 정보를 보호하는 기술 연구
  ▸개인정보 · 민감정보를 식별하여 영지식 증명기술을 통해 보호
  ▸인증/인가 시도에 과정 정보 노출 최소화 기술 연구
- DID 인증/인가 기반 블록체인 응용 시스템 구축
  ▸블록체인 응용 시스템 연구/개발 및 성능 측정과 평가

b) 광주과학기술원 황의석 교수 연구팀

- 복제불가능한 DID 기반 DID 문서 생성 및 제어 기법 개발
- IoT를 활용한 테스트베드 구성 및 성능 평가

4) 연구개발과제의 수행일정 및 주요 결과물

① 수행일정

가. 세부과제 1 단계별 수행 일정

1단계														
일련번호	개발내용	추진 일정												책임자 (소속 기관)
		2021. 01. 01 - 2021. 12. 31						2022. 01. 01 - 2022. 12. 31						
1-1	아날로그 신원 인증 키 연구													이홍노 (GIST)
1-2	비가역적 아날로그-디지털 스캐너 설계													이홍노 (GIST)
1-3	초분광압축이미징 장치 모델링 및 구현 및 데이터 확보													이홍노 (GIST)
1-4	대용량 화자 데이터 확보, 분류, 정제 및 특징추출													신종원 (GIST)
1-5	최신 딥러닝 기반 화자인증 베이스라인 구축 및 기초 모델 개발													신종원 (GIST)
1-6	분산 합의 기술 분석 및 특징파악													이홍노 (GIST)
1-7	이종 합의 메커니즘에 의한 환경 영향 분석													이홍노 (GIST)
1-8	강건한 블록체인 인프라를 위한 분산합의 기술 분석 및 개발													이홍노 (GIST)
1-9	트랜잭션 크기 별 안전 거래 분석 알고리즘 개발													이홍노 (GIST)
1-10	초경량 블록체인 기술 개발													문수묵 (서울대학교)
1-11	초확장 블록체인 기술 설계													문수묵 (서울대학교)
1-12	신뢰실행환경 기반 데이터 검증 경량화 기술 개발													문수묵 (서울대학교)

2단계																								
일련번호	개발내용	추진 일정																						책임자 (소속 기관)
		2023. 01. 01 - 2023. 12. 31						2024. 01. 01 - 2024. 12. 31						2025. 01. 01 - 2025. 12. 31						2026. 01. 01 - 2026. 12. 31				
2-1	영지식 증명 및 검증 프로토콜 설계, 성능분석 및 취약점 분석																							이흥노 (GIST)
2-2	초분광 이미지에서 비정상 데이터를 인지/추출/분류 알고리즘 개발																							이흥노 (GIST)
2-3	초분광 이미지를 복원하는 고급 딥러닝 기법 연구																							이흥노 (GIST)
2-4	시간 영역 Raw 데이터을 이용한 딥러닝 기반 고신뢰 화자인증 기술 개발																							신종원 (GIST)

2-5	신뢰도 향상을 위한 화자인증에 특화된 추가 모듈 개발																								신종원 (GIST)
2-6	원거리 상황, 잡음 환경 및 짧은 음성 입력에 강인한 모델 개발																								신종원 (GIST)
2-7	다중 복합구조 블록체인 연구																								이흥노 (GIST)
2-8	블록체인·샤드·클라우드 간의 통신 프로토콜 및 플랫폼 연동 기술 연구																								이흥노 (GIST)
2-9	블록체인 기반 빅데이터 생성 및 관리 기술 개발																								이흥노 (GIST)
2-10	분산형 애플리케이션 및 빅데이터를 위한 비즈니스 로직 및 스마트 계약 기술 개발																								이흥노 (GIST)
2-11	탈중앙화 연합학습 플랫폼 설계																								문수묵 (서울대학교)
2-12	비잔틴 내구성 연합학습 방법론 설계																								문수묵 (서울대학교)
2-13	블록체인 기반 연합학습 플랫폼 개발																								문수묵 (서울대학교)
2-14	IoT 기기를 위한 인공지능 모델 경량화																								문수묵 (서울대학교)

3단계														
일련 번호	개발내용	추진 일정												책임자 (소속 기관)
		2027. 01. 01 - 2027. 12. 31						2028. 01. 01 - 2028. 12. 31						
3-1	영지식 아날로그 생체인증 시스템 개발, 테스트베드 구축 및 평가													이흥노 (GIST)
3-2	개발한 모델과 베이스라인의 화자인증 성능 비교를 통해 검증													신종원 (GIST)
3-3	개발한 모델 최적화 및 고도화 실시													신종원 (GIST)
3-4	딥러닝 기반 화자인증 모델의 경량화 연구 진행													신종원 (GIST)
3-5	사용 용도별 블록전파지연시간, 참여노드 수에 따른 분산원장 파라미터 최적화													이흥노 (GIST)
3-6	샤드, 2nd 레이어, 크로스체인(cross-chain) 기술 융합 운용													이흥노 (GIST)
3-7	이종 IoT 네트워크 설계													문수묵 (서울대학교)
3-8	크로스체인 기술 개발													문수묵 (서울대학교)
3-9	통합 IoT 플랫폼 기술 개발													문수묵 (서울대학교)

나. 세부과제 2 단계별 수행 일정

1단계				
일련번호	개발내용	추진 일정		책임자 (소속 기관)
		2021. 07. 01 - 2021. 12. 31	2022. 01. 01 - 2022. 12. 31	
1-1	블록체인 대응 Cloud-native Connected DataLake 구성 설계			김종원 (GIST)
1-2	블록체인 대응 Cloud-native Connected DataLake PoC 검증			김종원 (GIST)
1-3	신뢰 정의 네트워크 (Trust-Defined Network) 구조 정의 및 데이터 제어 방식 구성			김황남 (고려대학교)
1-4	지속 가능한 데이터 스트림 경로의 신뢰도 수준 등급화 및 이를 기반으로 트래픽 분류 및 처리 방법 모델링			김황남 (고려대학교)

2단계						
일련번호	개발내용	추진 일정				책임자 (소속 기관)
		2023. 01. 01 - 2023. 12. 31	2024. 01. 01 - 2024. 12. 31	2025. 01. 01 - 2025. 12. 31	2026. 01. 01 - 2026. 12. 31	
2-1	멀티사이트 Connected DataLake를 위한 블록체인 대응 체계화된 데이터 저장·교환 기능 구현					김종원 (GIST)
2-2	블록체인 대응 체계화된 데이터 저장·교환을 지원하는 멀티사이트 Connected DataLake 구현					김종원 (GIST)
2-3	블록체인 대응 Connected DataLake를 위한 초고속 데이터 전달 기능 개발					김종원 (GIST)
2-4	다중경로 초고속 데이터 전달 기능 개발을 통한 블록체인 대응 Connected DataLake 개선					김종원 (GIST)
2-5	신뢰도 기반의 네트워크 구현 및 이에 최적화된 오버레이 네트워크 지원					김황남 (고려대학교)
2-6	블록체인 노드 간의 선택적 및 지능적 멀티캐스팅을 TDN 측면에서 지원					김황남 (고려대학교)
2-7	시민감 데이터 전송을 보장하는 동기화 네트워크를 확장한 Time Sensitive Networking (TSN) 기능 지원					김황남 (고려대학교)
2-8	장애 허용성을 가지는 분산형 TDN 컨트롤러 기반 효율적인 트래픽 분산 방안 연구					김황남 (고려대학교)

3단계				
일련번호	개발내용	추진 일정		책임자 (소속 기관)
		2027. 01. 01 - 2027. 12. 31	2028. 01. 01 - 2028. 12. 31	
3-1	블록체인 대응 Connected DataLake 고도화를 위한 분산된 데이터 셋 동기 기능 개발			김종원 (GIST)
3-2	분산된 데이터 셋 동기 및 공유 기능을 통합한 블록체인 대응 Connected DataLake 고도화			김종원 (GIST)
3-3	TDN 요소 기술 통합 및 테스트베드 개발			김황남 (고려대학교)
3-4	TDN 시스템 통합 및 PoC 레벨 구현			김황남 (고려대학교)

다. 세부과제 3 단계별 수행 일정

1단계				
일련번호	개발내용	추진 일정		책임자 (소속 기관)
		2021. 07. 01 - 2021. 12. 31	2022. 01. 01 - 2022. 12. 31	
1-1	DID 블록체인 기반 인증/인가 서비스 요구사항 분석			김기형 (아주대학교)
1-2	IoT에 적합한 DID 블록체인 인증 기법 연구			김기형 (아주대학교)
1-3	안전한 IoT Data 전송 방법 연구			김기형 (아주대학교)
1-4	IoT 보안통신의 보안취약점 분석			황의석 (GIST)
1-5	IoT의 물리적 특성을 활용한 DID 생성기술 연구			황의석 (GIST)
1-6	DID 블록체인 연계 IoT 인증 프로토콜 연구			황의석 (GIST)

2단계						
일련번호	개발내용	추진 일정				책임자 (소속 기관)
		2023. 01. 01 - 2023. 12. 31	2024. 01. 01 - 2024. 12. 31	2025. 01. 01 - 2025. 12. 31	2026. 01. 01 - 2026. 12. 31	
2-1	DID 블록체인 기반의 자기주권형 액세스 토큰 발행 연구					김기형 (아주대학교)
2-2	IoT 환경의 DID 블록체인 기반 접근제어 방법 연구					김기형 (아주대학교)
2-3	영지식 증명 기반의 자격증명 방법 연구					김기형 (아주대학교)
2-4	대규모 IoT 환경에서 신뢰성을 갖는 DID 생성 기술 개발					황의석 (GIST)
2-5	동적 DID 기반 보안인증 프로토콜 연구					황의석 (GIST)
2-6	소규모 private 블록체인 연계 보안통신 기술 검증					황의석 (GIST)

3단계				
일련번호	개발내용	추진 일정		책임자 (소속 기관)
		2027. 01. 01 - 2027. 12. 31	2028. 01. 01 - 2028. 12. 31	
3-1	DID 인증/인가 기반 블록체인 응용 시스템 연구/개발			김기형 (아주대학교)
3-2	DID 인증/인가 기반 블록체인 응용 시스템 성능 측정/평가			김기형 (아주대학교)
3-3	복제불가능한 DID 기반 DID 문서 생성 및 제어기법 개발			황의석 (GIST)
3-4	IoT를 활용한 테스트베드 구성 및 성능 평가			황의석 (GIST)

② 주요 결과물

가. 세부과제 1 주요 결과물

1단계 주요결과물	2단계 주요결과물	3단계 주요결과물
▪ 비가역적 아날로그-디지털 스캐너 명세서 및 시뮬레이션 코드 ▪ 필터기반 초분광 압축 이미지 데이터셋 및 영상 장치 ▪ 미래 블록체인 인프라를 위한 강건한 합의 메커니즘 분석 도구 ▪ 화자인증을 위한 정제된 데이터 및 베이스라인 및 화자인증 기초 모델 ▪ 블록체인 스토리지 데이터 경량화를 위한 비활성 상태 정보 삭제 및 복구 코드 ▪ 신뢰실행환경과 블록 헤더 검증을 통한 경량화된 블록체인 데이터 검증 알고리즘 ▪ 화자인증을 위한 정제된 데이터 및 베이스라인 및 화자인증 기초 모델	▪ 영지식 증명 프로토콜 명세서 및 시뮬레이션 코드 ▪ 초분광 압축 이미징 장치를 활용한 분광 신호 복원 및 비정상 데이터 인지 알고리즘 ▪ 빅데이터 기반 스마트 계약 및 분산 애플리케이션 ▪ 비잔틴 저항성을 가지는 연합학습 방법론 ▪ IoT 기기를 위한 인공지능 모델 경량화 알고리즘 ▪ 고신뢰 화자인증 모델 (에러율 <= 3%)	▪ 영지식 아날로그 생체인증 시스템 및 테스트베드 ▪ 초분광 이미징 장치를 위한 디지털 지능 트윈 피드백 알고리즘 ▪ 분산 블록체인 기반 빅데이터 처리가 가능한 디지털 트윈 시스템 ▪ 이종 블록체인간 상호작용을 위한 2-way-relayer 코드 ▪ 블록체인 기반 초확장 IoT 네트워크를 위한 디지털 트윈 플랫폼 ▪ 고도화된 고신뢰 화자인증 모델 (에러율 <= 2%)

나. 세부과제 2 주요 결과물

1단계 주요결과물	2단계 주요결과물	3단계 주요결과물
▪ 블록체인 대응 Cloud-native Connected DataLake 프레임워크 설계 기술문서 ▪ Blockchain 대응 Cloud-native Connected DataLake 지원 멀티사이트 클러스터 시험 구축 ▪ 확장가능한 신뢰 정의 네트워크 명세화 및 구체적 제어 방식 제안 ▪ 네트워크 구성요소의 신뢰도 기준 확립 및 평가 모델 개발	▪ 멀티사이트 Connected DataLake를 위한 데이터 태깅 및 메타데이터 병합/분할 기능 소프트웨어 ▪ 3개 사이트에 분산된 DataPond 클러스터들을 연동하여 구축한 Connected DataLake 초기형 인프라 ▪ 블록체인 대응 Connected DataLake 구성을 위한 데이터 네트워킹 오프로딩 기능 소프트웨어 ▪ 다중경로 패브릭 네트워킹 기반의 병렬 전달 기능을 포함한 Connected DataLake 개선형 인프라 ▪ 기존 성능지표 기반의 네트워크 구성 기법을 개선한 신뢰도 기반의 네트워크 구성 기법 구현 ▪ 지능적 멀티캐스트 방식 개발을 통해 블록체인 노드들에 신속한 합의 환경 제공 ▪ 확정적 데이터 전송을 보장하는 시민감 네트워크 체계 확립 ▪ 분산형 TDN 컨트롤러 개념 확립 및 구성, 이를 지원하는 소규모 사이드체인 도입	▪ 블록체인 대응 Connected DataLake의 고도화를 위한 분산된 데이터 셋 동기 기능 소프트웨어 ▪ 기능 통합을 포함한 블록체인 대응 멀티사이트 Cloud-native Connected DataLake의 고도화 인프라 ▪ TDN 요소 기술 통합 및 테스트베드 개발 ▪ TDN 시스템 통합 및 PoC 레벨 구현, 타 세부 과제와의 시스템과의 연계

다. 세부과제 3 주요 결과물

1단계 주요결과물	2단계 주요결과물	3단계 주요결과물
▪ DID 블록체인 기반 인증/인가 요구사항 분석서 ▪ 단계별 DID 적용 방안 분석서 ▪ IoT의 물리적인 고유한 특성 정보 데이터셋 ▪ 물리적인 특성을 활용한 DID 생성 알고리즘 기술문서 ▪ 물리계층 DID 인증 프로토콜 명세서 및 시뮬레이션 소프트웨어 (S/W)	▪ DID 인증을 위한 DID Document 구조 설계서 ▪ 자기주권형 액세스 토큰 발행 및 관리 시스템 기술 문서 ▪ 접근제어 시나리오 ▪ DID기반 인가 및 접근제어 결과 평가서 ▪ 동적 DID 생성 알고리즘 및 소프트웨어 (S/W) ▪ 동적 DID 관리를 위한 인증 프로토콜 명세서	▪ DID 인증/인가 기반 블록체인 응용 시스템 프로토타입 소프트웨어 ▪ 응용 시스템 자체 시험 평가 보고서 ▪ 복제불가능한 DID 기반 IoT 통신보안 테스트베드 ▪ 테스트베드를 활용한 DID 접근제어 테스트 결과 분석 기술 문서

5) 세부 프로젝트 연구계획

① 제1프로젝트 연구계획 (혁신도약형)

가. 혁신도약형 연구과제 개요

프로젝트명	국문	영지식 센싱 탈중앙 고성능 분산원장 기술 개발 (혁신도약형)		
세부 프로젝트 책임자	성명	이홍노	직위	교수
	소속기관	광주과학기술원	부서명(학과명)	전자전기컴퓨터공학부
	연락처	[redacted]	E-Mail	[redacted]
혁신도약형 과제 여부		O	기술분야 (소분류 기준)	블록체인-블록체인 기반기술

나. 연구내용

○ 제1 세부 연구팀은 **영지식 센싱 탈중앙 고성능 분산원장 기술을 연구함.** 분산합의, 분산원장 기술 및 고성능 트랜잭션 처리 기술과 같은 블록체인 기반 기술을 연구함과 동시에 2nd layer 및 클라우드등을 사용하는 블록체인 구조 확장 기술을 연구함. 또한 블록체인 응용 지원 연구로써 DID를 위한 초분광 카메라, 화자인식 기반 ZKP 생체인증기법을 개발하고, 블록체인 및 클라우드 상에서 동작하는 연합학습 플랫폼을 개발함.

○ **제1세부 과제를 혁신 도약형 과제로 수행함.**

○ **(혁신성)** 제 1세부 과제의 혁신성은 다음과 같음. 즉, 첫 째, 세계 최고 수준의 블록체인 합의알고리즘 기술을 개발하고, 글로벌 기업(Algoland, RapidChain등)의 테스트 성과치를 비교하고(pg. 9 참조), 2단계 완료 전에 100% 달성하고자 목표로 함. 둘째, 영지식 증명을 초분광이미징 센싱과 화자인식 기술과 융합 활용하여, DID 및 데이터 인증에 활용하고, 생체정보의 오남용 문제와 오라클 문제를 해소하는 원천 기술획득을 목표로 제시함.

○ **(도전성)** 상기한 목표는 확장성, 보안성, 탈중앙성 트릴레마 문제 및 오라클 문제 등 블록체인 업계의 핵심부를 다루는 고도의 숙련과 투자를 요구하는 원천 기술을 개발하는 것임. 국내 기업의 기술개발 수준은 세계적인 수준과 매우 큰 편차가 있는 것으로 평가됨. 본 센터는 연구 및 교육을 통해 세계 최고 수준의 합의 기술과 영지식 센싱 기술로 훈련된 석박사급 고급 인력을 양성하고, 원천기술을 확보하게 되면, 대한민국의 블록체인 분야 전반에 경쟁력 제고 및 신산업 기반 확보의 발판을 마련할 수 있게 됨. 세계 각국에서 블록체인 합의알고리즘 및 영지식 관련 연구개발 경쟁이 가속화되고 있는 이때, 상기 핵심기술 미확보 시 기술 종속성 및 기술사용 로열티 문제가 예상되어 기술 개발 초기 단계에 정부 지원이 절실함.

○ **(목표치 설정, 평가 및 환류 체계 수립)** 혁신적 성과목표를 제시하고 미달성 시에는 실패원인을 분석하고 그 결과를 기록하여 다음 연구에 피드백 할 수 있는 평가 관리체계 마련함. 최종 과제 목표를 아래와 설정하고, 단계별로 목표치를 설정하고 진행함. 가령 TPS는 1단계는 70% 달성, 2단계까지는 100% 달성하는 목표치를 제시함. 3단계 종료 까지는 확보한 기술로 교원 및 학생이 창업을 하는 것을 목표로 함.

○ 영지식 센싱 탈중앙 고성능 분산원장 기술 개발을 위해 다음과 같은 세부연구를 진행함.

[1-1] 고성능 합의알고리즘등 블록체인 기반기술 연구

[1-2] 다중 복합구조 계층형 블록체인 구조 확장 기술 연구

[1-3] DID를 위한 초분광 카메라, 화자인식 기반 ZKP 생체인증기법 개발

[1-4] 블록체인 및 클라우드 상에서 동작하는 연합학습 플랫폼 개발

○ **[1-1] 고성능 합의알고리즘등 블록체인 기반기술 연구**

- **(ECCPoW란?)** 이흥노교수 연구팀이 2018년 개발하여 논문들과 Github SW 패키지를 공개한(https://github.com/cryptoecc, 포크 2천여건 발생) 새로운 분산합의 알고리즘임. ECCPoW는 오류정정부호를 사용하여 PoW 수행을 위한 퍼즐을 시간에 따라 변하도록 함. 그 결과 ASIC 개발을 불가능하게 하여 탈중앙성을 강화시킴. ECCPoW 에서는 각 블록 생성 노드들이 오류정정부호 디코더가 특정 부호를 출력하게 만드는 넌스값을 찾기 위해 노력함. ETH-ECC는 이더리움과 ECCPoW 프로토콜이 결합한 미래형 탈중앙 블록체인임. ETH-ECC 블록체인 프로토콜은 스마트 계약 지원이 가능한 가상머신을 갖고 있음. 따라서 본 과제에서 필요한 지능 융합 DApp 및 응용 서비스 제공이 가능함. 즉 생체정보 인증, 디지털 신분증, 디지털 트윈 기술, 사물인터넷, 금융 애플리케이션에 적용 가능함.

- **(합의 알고리즘의 보안성 강화 연구)** 블록체인은 분산화 특성으로 인해 다양한 공격에 노출될 수 있음. 대표적으로 Selfish Mining 공격, Sybil 공격, 이중지불 공격 등의 공격이 가능함. BFT와 그 변종 알고리즘들, 그리고 PoA와 같이 Identity (ID)를 기반으로 한 합의 방식의 경우 Sybil 공격에 취약한 반면, PoW는 Sybil 공격에 강인함. 본 연구팀은 ECCPoW PoS, DPoS, BFT계열 합의 알고리즘의 보안성 결정 요인 분석함.

- **(ECCPoW 보안성 연구)** Selfish Mining 공격은 PoW를 사용하는 블록체인을 겨냥한 공격이지만, 실용성이 검증되지 않음. 이중지불공격은 모든 합의알고리즘에서 문제가 될 수 있음. 결론적으로 ECCPoW는 PoW의 한 종류로써 Sybil 공격과 Selfish Mining 공격에 대한 안전성을 내재하고 있기 때문에, 이중지불공격에 대한 대책마련이 보안성 제고에 지대한 영향을 끼칠 것으로 기대함.

- **(부호-암호 작업증명의 보안성 강화)** 최적의 검증횟수란 이중지불공격을 방지하는 동시에 거래완료에 소요되는 시간을 최소화 하는 검증횟수임. 검증횟수를 충분히 크게 설정하면 이중지불 공격자의 이윤 발생이 방지됨. 반면 검증횟수가 클수록 거래완료에 소요되는 시간이 길어짐. 따라서 거래 금액에 따라 공격자의 이윤 발생을 방지하는 최소한의 검증횟수를 찾아야 함. 즉, 최대 거래 속도(확장성)가 거래의 용도에 따라 유동적으로 결정되는 것임. 공격 대상 거래가 소액일 때는 이중지불 공격을 통한 이윤을 기대하기 어려움. 따라서 보안성보다 확장성을 위한 작은 검증횟수가 적절함. 반면, 고액 거래는 이중지불 공격의 목표가 될 가능성이 있음. 따라서 거래량이 이윤발생 위험 거래량보다 작아지도록 하는 검증횟수를 결정함으로써 보안성을 유지함. 본 연구에서는 거래금액에 따라 최적의 검증횟수를 결정해주는 알고리즘을 개발함. 검증횟수 최적화 알고리즘을 지갑에 탑재하여 거래를 생성할 때 마다 자동 실행되도록 하는 DApp을 개발함으로써 이중지불 공격에 대한 보안성을 제고함.

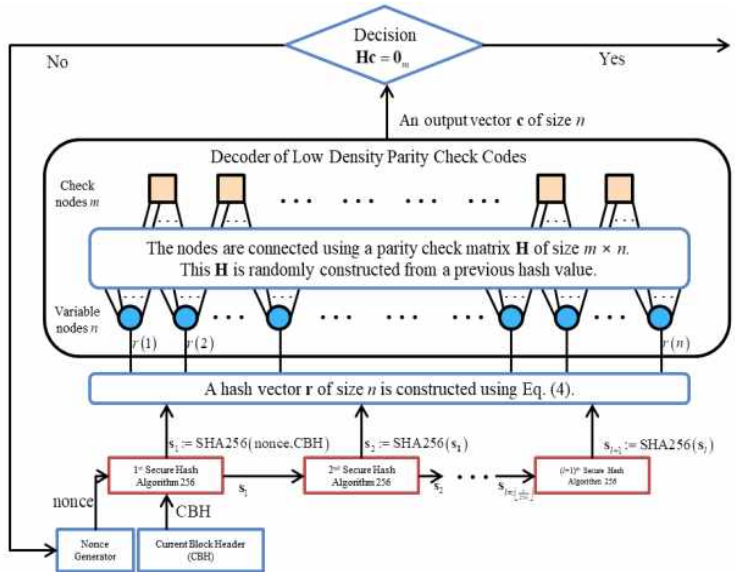


그림 8. ECCPoW 기술 개략도

○ **다중 복합구조 계층형 블록체인 구조 확장 기술 연구**

- **(다중 복합구조 계층형 블록체인 연구)** 합의 알고리즘이 전 세계 공개형 블록체인에서부터 지역 블록체인, 혹은 Private 블록체인으로 Hybrid의 형태로 진화하고 있음. Evolutionary Space에서 다중 복합구조 계층형 블록체인들을 설계하고, 각각을 용도별로 특화시킬 필요가 있음.

- **(블록체인간의 통신 프로토콜 연구)** 다중 복합구조 계층형 블록체인간의 가치 및 정보 교환을 위한 통신 프로토콜을 연구함. 다중 복합구조 계층형 블록체인들은 모두 같은 합의 알고리즘을 탑재하고 있으므로 블록체인 간 통신이 용이함. 사이드체인의 개념으로 다중 구조 블록체인을 제안한 Polkadot과 Cosmos와 같이 통신의 Hub Station 역할을 해주는 블록체인을 도입하는 방법도 적용 가능함. 다중 복합구조 계층형 블록체인을 국제체인, 국내체인, 지역체인 등의 다중 구조로 설계하여 용도에 따라 각 블록체인의 주요 파라미터를 독립적으로 최적화함. 예를 들어, 국제체인에서는 큰 규모의 거래가 이루어지므로 TPS 보다는 보안성 강화에 초점을 맞추어 파라미터를 설정함. 반면 지역체인은 소규모의 거래에 맞추어 최소한의 보안성을 유지하며 TPS를 극대화함.

- **(블록 생성 주기 제어 방법 연구)** 블록 생성 주기는 하나의 블록이 생성되고, 그 다음 블록이 생성이 될 때까지 걸리는 시간으로 정의 할 수 있음. 블록 생성 주기는 미리 결정되고 유지되도록 설계되어야 함. 예를 들어 비트코인은 평균 블록 생성 주기가 10분이 되도록 설계됨. 2016개의 블록 생성이 완료될 때마다 블록 생성 난이도가 증감되도록 구현함. ECCPoW의 노드들은 넌스(nonce) 값을 바꾸면서 디코더가 부호를 출력하게 만드는 특정 넌스값을 찾기 위해 노력함. 따라서 디코더가 부호를 출력할 확률을 통제함으로써 블록 생성 난이도를 제어함.

- **(블록 생성 테스트)** 개발된 시스템을 단일 노드 / 복수 노드 시나리오 2가지 시나리오에서 테스트하여 문제점을 찾고 보완하고자 함. 각 시나리오에서 블록작성 검증, 체인 형성 검증, 난이도 조절 검증, 블록 동기화 검증, 합의 검증 및 거래 검증을 수행함.

- **(용도에 따른 블록 파라미터 독립 최적화)** Enterprise 블록체인은 Private Consortium으로 구성되어 기업간의 정보를 공유하는데 사용될 수 있음. 지역규모 블록체인은 시/구/군 규모의 블록체인으로 작은 규모의 거래가 주를 이룸. 국가규모와 국제규모 블록체인은 각각 지역간, 그리고 국가간의 거래를 기록하며, 거래의 규모가 큼. 사용 용도에 따라 블록전파지연시간, 참여노드 수가 다르므로 용도에 따른 파라미터 최적화를 연구함. Enterprise와 지역규모 블록체인들은 빠른 거래 처리속도 (TPS)에 집중하여 블록 생성 난이도를 낮은 수준으로 설정함. 동시에 이중지불공격 방지 및 이상행위 퇴출을 위해 멤버십 제도와 같은 정책을 적용함. 국가규모와 국제규모 블록체인들은 거래의 규모가 큰 만큼 거래 처리속도 보다 보안성 제고를 위해 블록 생성 난이도를 높은 수준으로 설정함.

- **(비잔틴 저항성 확보를 통한 오프체인 데이터 신뢰성 확보 기술)** 블록체인은 온체인(on-chain) 데이터에 대해서는 무결성을 보장함. 그러나 밖으로부터 들어오는 오프체인(off-chain) 데이터에 대해서는, 그림 8에서와 같이 데이터를 입력하는 주체를 신뢰할 수밖에 없는 오라클(oracle) 문제가 발생함. 이러한 오라클 문제는 근본적인 해결법이 없으므로, 악의적/불안정한 네트워크로 인해 위변조된 데이터를 감내하고 학습할 수 있는 비잔틴 저항성(Byzantine tolerance)을 갖춘 인공지능 학습 방법이 필요함. 특히 본 과제에서는 데이터 프라이버시를 보장하기 위해 연합학습(Federated Learning, FL)의 방법론을 차용하는데, 연합학습에서 비잔틴 저항성을 달성하기 위한 방법을 연구할 필요성이 있음. 대표적인 방법으로는 최대 50%의 비잔틴까지 감내가 가능한 '중앙값 사용' 의 방법이 있음. 또한, 연합학습은 모델이 가벼워 IoT 기기에서도 취급이 가능하고, 네트워크 비용을 효율적으로 사용할 수 있으며, 데이터 프라이버시를 지키는 환경 하에서 운용될 수 있음.

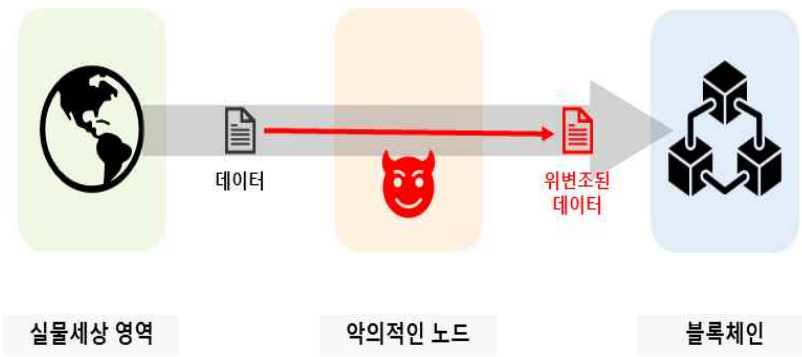


그림 9. 오라클 문제의 도식

- **(크로스체인 기반 초경량 블록체인)** IoT 기기로 형성된 네트워크는 지역, 성능 등의 기준을 따라 여러 계층으로 구성하면 효율적임. 이 때 서로 다른 네트워크(블록체인)가 통신하기 위해서는 별도의 장치가 필요한데, 이를 이종 블록체인 간 크로스체인(cross-chain) 기술이라 함. 대표적인 크로스체인 기술 중에는 릴레이(relay)가 있는데, 릴레이에서는 한 블록체인의 블록 헤더를 상대 블록체인에서 추적하고 검증하며, 단순지불검증(Simplified Payment Verification, SPV)을 통해 트랜잭션의 존재성을 입증함. 한 블록체인의 트랜잭션을 상대 블록체인에서 검증할 수 있게 되므로, 자산의 이동이나 이벤트의 발발, 데이

터의 전송 등이 가능함. 양방향 릴레이 기술인 2-way relay를 통해 서로 다른 IoT 네트워크로 이루어진 분할된 환경이라도 통합해서 관리할 수가 있음. 이로부터 사물제어 제고를 관장하는 초확장 플랫폼의 구축이 가능함.

○ **비잔틴 저항성 확보를 통한 오프체인(off-chain) 데이터 신뢰성 확보 기술**

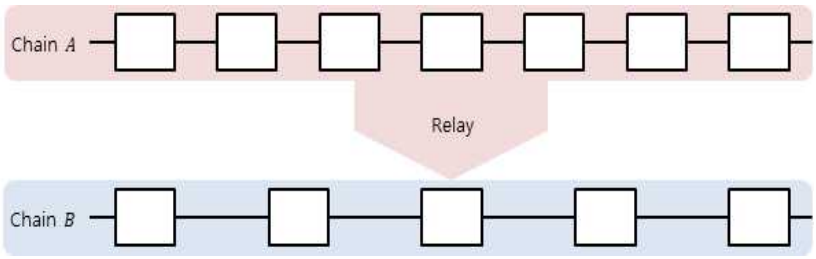


그림 10. 이종 블록체인 간 릴레이(Relay) 도식

- **(신뢰실행환경기반 초경량 블록체인)** IoT 기기에서도 사용 가능하게 가볍고 효율적이고 신뢰할 수 있는 데이터베이스를 설계 및 개발함. 블록체인은 신뢰할 수 있는 분산 데이터베이스라는 점에서 차용할 수 있으나, 스토리지 관련 확장성 한계가 존재하므로 비트코인, 이더리움과 같은 현존하는 블록체인을 그대로 IoT 기기에 적용할 수 없음. 가령 이더리움의 경우 7TB에 달하는 대용량 스토리지를 요구하며, 현재도 매우 빠른 추세로 증가하고 있음. 이는 비단 스토리지 제한뿐 만 아니라 동기화 과정을 늦추는 병목을 야기함. 본 과제에서는 블록체인 데이터 중 가장 많은 비중을 차지하는 상태(state) 데이터를, 비활성화된 계정에 대해서는 일시적으로 삭제하는 기법을 통해 스토리지 확장성을 확보함. 이러한 방법을 이더리움에 적용하면 상태를 저장하는 상태 트리가 약 220MB 정도로 수렴할 수 있음. 본 과제에서는 발생하는 데이터의 총량이 이더리움 대비 훨씬 적을 것이므로, 그 크기는 수 MB 정도로 수렴할 것으로 기대함. 이러한 기법은 동기화 과정에서 소요되는 시간도 감축시키는데, 각 노드는 가장 최근 주기의 상태 트리만으로 부트스트랩이 가능하므로, 오래된 트랜잭션과 덜 중요한 블록 헤더를 생략하는 최적화 방법과 결합되면 기존 대비 훨씬 적은 데이터만으로도 부트스트랩이 가능해짐. 또한, 신뢰실행환경(Trusted Execution Environment, TEE)을 사용하면 가벼운 블록 헤더 검증만으로도 블록체인의 전체 데이터를 검증할 필요 없이 신뢰 하에서 데이터를 활용할 수 있음. TEE에서는 기본적으로 모든 데이터가 암호화 되어 외부에선 함부로 보거나 조작할 수 없고, 작성된 프로그램 대로만 동작하고, 임의로 내부 데이터를 조작할 수 없음. 이러한 특성을 이용하면 IoT 기기들은 상대적으로 가벼운 헤더 체인만 저장하고 TEE 내에 있는 헤더 체인이 올바르다는 것을 확인하는 것으로, TEE가 보유하고 있는 블록체인 전체 데이터가 유효하고 올바른 것으로 신뢰할 수 있음.

○ DID를 위한 초분광 카메라, 화자인식 기반 ZKP 생체인증기법 개발

- **(초분광 생체정보 기반 DID 기술)** 본 과제에서는 초분광 이미지(HSI) 를 사람의 생체정보 인식 분야에 적용하여, 고유한 생체정보를 초분광 영역에서 파장 대역별로 분석하고, 이를 특징삼아 사용자를 분별할 수 있는 DID 로 활용함.

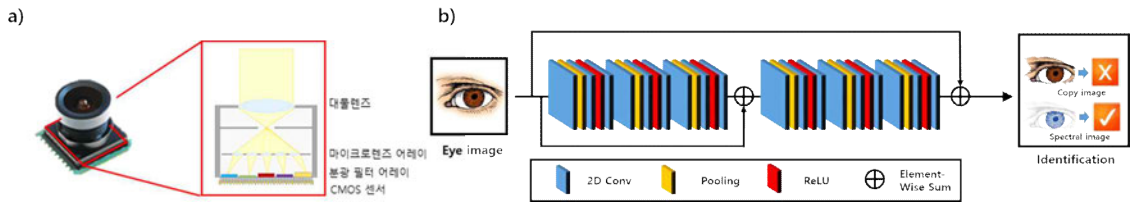


그림 11. 초분광 생체정보 기반 DID 기술;
a) 초분광 압축 이미징 장치, b) 딥러닝 기반 분광 해상도 개선 기법

- **(HSI 기술의 문제점)** HSI 기술은 10년 이상 사용되어 온 기술이지만 지금까지는 고차원 H/W 구성으로 인해 상대적으로 높은 비용을 요구해 왔으며, 때문에 NASA 등 국방분야에의 정부 연구기관에 한해 적용되옴. 이러한 이유로 카메라, 라이다와 같은 다른 이미징 기술과 비교하여 상대적으로 아직 알려지지 않은 기술이 됨. 본 연구팀은 이러한 초분광 이미징 기술의 저변 확대를 위해 필터 어레이 기반 소형 초분광 압축 이미징 장치를 연구 개발하고, H/W 적으로 열화되는 영상 해상도를 딥러닝 기반 이미지 처리를 통해 개선하고자 하며, 초분광 특성을 이용한 영상 보정 기술을 통해 초분광 생체정보를 획득할 수 있는 기술을 개발하고자 함.

- **(HSI 연구 주제)** 본 연구팀은 초분광 생체정보 기반 Decentralized IDs 기술 개발을 위하여 필터 어레이 기반 소형 초분광 압축이미징 장치 연구, 딥러닝 기반 초분광 이미지 해상도 개선 알고리즘 연구 및 딥러닝 기반 분광 이미지 Dehazing 알고리즘 개발을 수행함. 먼저 2차원 랜덤 압축 센싱 필터, 광학렌즈 및 CMOS 이미지 센서를 활용하여 초분광 압축 이미징 장치를 개발하고, 데이터를 확보할 계획임. 랜덤 압축 센싱 필터를 CMOS 이미지 센서 위에 바로 부착하는 접근 방식을 통해, 소형 이미징 장치를 구성하며, 한번의 노출로 다양한 정적/동적 장면에 대한 측정을 수행하고자 함. 다음으로 딥러닝 기술을 이용한 초분광 이미지 해상도 개선 알고리즘을 연구하고자 함.

- **(화자인식기반 DID 기술)** 딥러닝의 발전으로 인해, 심층 신경망 (Deep neural networks) 을 이용하여 화자 정보를 추출하는 시스템들이 제안되었으며, 전통적인 화자인증 방식인 I-vector 보다 좋은 성능을 보임. 추가로, 화자 정보간의 차이를 두드러지게 강조하기 위해 다양한 손실함수를 적용한 연구·개발이 진행됨. 지금까지의 화자인증을 위한 최신 딥러닝 모델 기법들은 통상적으로 MFCC, spectrogram 등 주파수 영역으로 변환한 특징 추출 기법을 사용하여 진행된 바가 대다수이며, 시간 영역에 특화된 화자인증 모델의 기술의 연구·개발은 충분히 이뤄지지 않은 상태임. 게다가 이러한 전통적인 특징 추출 기법이 화자인증 모델에 특화되어 설계된 특징들이 아니기 때문에 최적의 특징이 아닐 여지가 충분함. 본 연구과제에서는 이러한 기존의 특징 추출 기법들을 대용량 데이터를 이용하면서 대체할 수 있는 심층 신경망기반의 시간영역 end-to-end 화자인증 모델 기술의 개발을 진행함. 최근에 시간영역 음원분리 기법에서

는 기존의 주파수-시간 변환 영역 특징 추출을 data-driven transformation을 이용하여 대체하는 기술이 개발되어 좋은 성능을 보인바 있음.

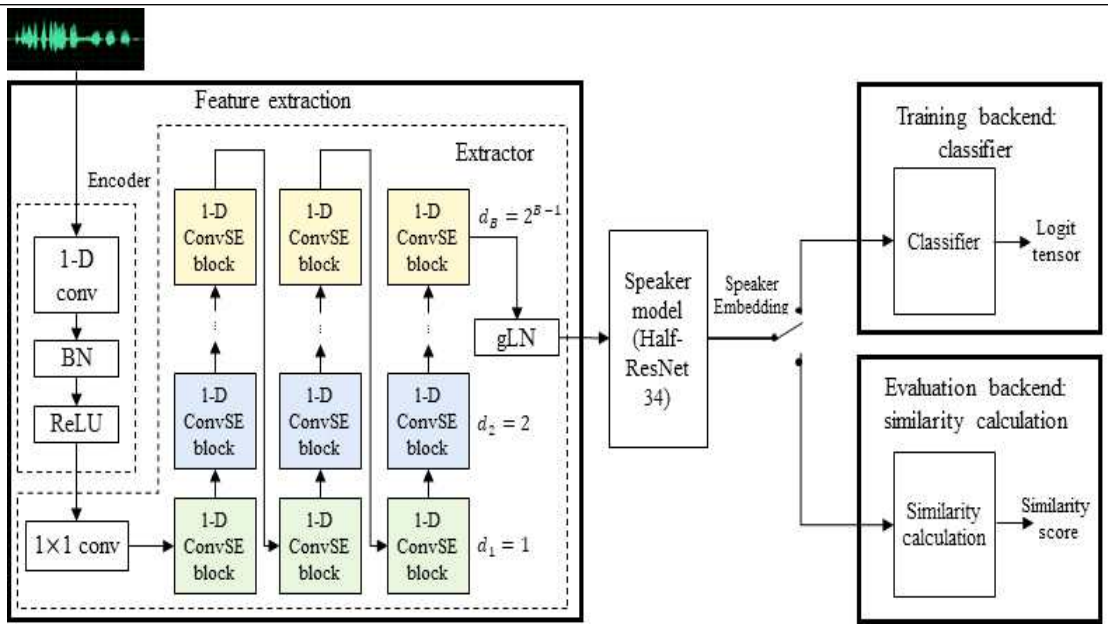


그림 12. 딥러닝 기반 화자인증 기술 개략도

- **(ZKP 기반 생체정보 보호)** 영지식 아날로그 생체인증 시스템의 개발을 통해 개인 생체정보의 유출을 원천 차단하는 생체인증 시스템의 개발하고자 함. 생체정보가 유출될 경우 개인정보 유출문제 뿐만 아니라 인증 시스템의 보안에도 치명적임. 영지식 아날로그 생체인증 시스템의 특징은 아날로그 상태의 생체 신호를 무작위 압축하여 디지털화 함. 따라서 원본 생체정보를 생성조차 하지 않으므로 생체정보의 유출이 원천 차단함.

- 생체인증 시스템에서 개인 생체정보의 유출을 막기 위한 방안은 이전부터 연구됨. 이전에 제안된 방안들은 암호화 기술을 사용하여 원본 생체 데이터를 숨긴 후 이를 인증에 사용함. 사용된 방안들은 크게 Fuzzy commitment, Fuzzy vault, Fuzzy extraction, 폐기형(Cancalable) 암호화, 동형암호화 등으로 분류됨. 이전 방안들은 공통적으로 디지털화 된 원본 생체정보를 알고리즘을 사용해 숨김. 즉, 생체 아날로그 신호는 이미 원본 그대로 디지털화 되어 컴퓨터의 어딘가에 저장된 적이 있으므로, 해킹 등에 의해 탈취될 위협이 존재함. 따라서 원본 생체 디지털 정보를 사용하지 않으면서 개인 인증을 수행 할 수 있도록 하는 영지식 아날로그 생체인증 시스템이 필요함.

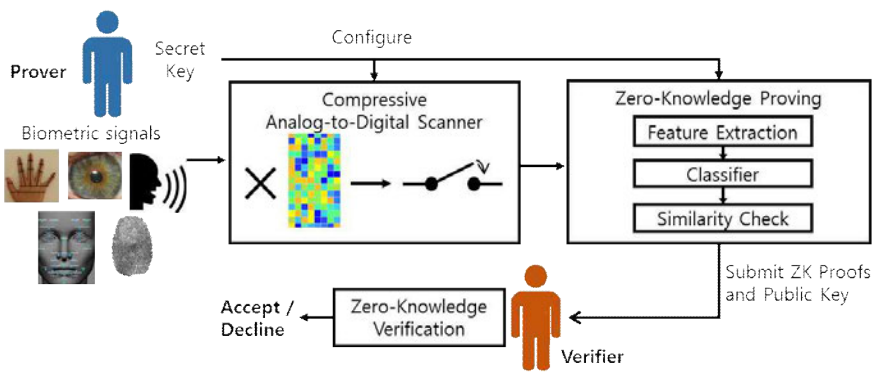


그림 13. 영지식 아날로그 생체인증 시스템 블록도

- 제안하는 영지식 아날로그 생체인증 시스템은 그림 12와 같이 압축 아날로그-디지털 스캐너, 인증 알고리즘(특징추출, 분류, 및 유사도계산 알고리즘), 그리고 영지식 증명 프로토콜 (영지식 증거생성 및 검증 알고리즘)으로 구성됨. 사용자 (Prover)가 자신의 생체 신호를 사용하여 자가 인증 후 그 증거를 생성하고 검증자에게 인증 요청함. 검증자 (Verifier)는 증거가 올바른지 확인함으로써 사용자가 자가 인증을 올바르게 수행하였는지를 확인함.

- 압축 아날로그-디지털 스캐너의 특징은 비가역적이라는 것임. 이러한 스캐너를 설계하기 위해 압축센싱이론이 활용됨. 압축센싱은 희소한 아날로그 신호를 짧은 길이의 디지털 신호로 무손실 압축 및 샘플링이 가능하도록 함. 반면 우리의 입력 신호인 생체 신호는 자연 신호이므로, 불확정성이 매우 크며 희소하지 않음. 따라서 생체 신호를 압축센싱할 경우 원 신호를 되돌릴 수 없음. 즉, 비가역적임. 원본 신호는 되돌릴 수 없지만, 생체 신호의 종류에 따라 특징 성분 (feature)는 희소하게 표현할 수 있음. 이는 곧 특징 성분에 관한 정보는 보존할 수 있다는 의미이며, 개개인을 분류하는 목적을 달성할 수 있음을 의미함.

- 사용자는 아날로그 비밀 키를 사용하여 압축 아날로그-디지털 스캐너의 전달 특성을 설정할 수 있음. 즉, 스캐너의 물리적 특성을 물리적인 방법으로 조절할 수 있음. 아날로그 비밀 키를 아는 사용자만이 압축된 디지털 신호로부터 원본 생체 신호의 특징 성분을 복원 할 수 있음. 그러므로, 아날로그 비밀 키를 알지 못할 경우, 압축된 디지털 신호가 유출되어도 악용될 수 없음.

- 인증 알고리즘은 특징성분 추출 알고리즘, 분류 알고리즘, 그리고 등록된 데이터와의 유사도 계산 알고리즘으로 구성되어 있음. 인증 알고리즘은 이전부터 연구되어 온 성능이 검증된 알고리즘을 사용함. 만약 생체정보를 원본 그대로 사용하여 인증할 경우, 검증자는 사용자의 신원뿐만 아니라 원본 생체정보까지 확보하므로 개인정보 유출 및 악용의 가능성이 있음. 이를 방지하기 위하여 영지식 증명을 사용함. 영지식 증명은 사용자가 스스로 자기 자신을 인증하고, 그 증거만을 검증자에게 건네도록 하는 보안 프로토콜임. 영지식 증명의 증거 생성 알고리즘은 인증 알고리즘과 같은 복잡한 알고리즘에 사용자가 올바른 비밀 입력값을 사용하여 공개된 출력값을 얻는 일련의 계산 과정 전체를 암호화 및 압축하여 증거를 생성함. 이후 검증자는 영지식 검증 알고리즘, 증거, 그리고 공개된 출력값을 사용하여 증거가 올바르게 생성되었다는 사실을 확인함. 결과적으로, 사용자가 자신의 생체정보를 사용하여 스스로 인증 알고리즘을 실행하기 때문에, 검증자는 사용자의 생체정보를 알지 못함.

- 결과적으로, 영지식 아날로그 생체인증 시스템은 압축 아날로그-디지털 스캐너를 사용하므로 원본 생체 신호의 디지털 데이터는 존재하지 않고, 영지식 증명 프로토콜을 사용하여 개인 데이터의 노출 없이 개인 인증이 가능함. 따라서 개인 생체정보의 유출 가능성을 원천 차단함.

○ **블록체인 및 클라우드 상에서 동작하는 연합학습 플랫폼 개발**

- **(초분광 이미지 학습 모델 개발)** 본 연구팀의 H/W적 접근 방식은 기기의 소형화를

달성할 수는 있으나, CMOS 센서 위에 부착하는 분광 필터의 개수가 많아짐에 따라 공간 정보 해상도가 저하되는 문제가 있음. 이를 딥러닝 기술을 이용한 초분광 이미지 개선 알고리즘 및, Few-Shot learning / Transfer Learning을 이용한 초분광 데이터셋 학습 모델, 그리고 Blockchain 환경에서의 Federated Learning dataset을 활용한 Overfitting 방지 기술에 대해 연구함. 또한 본 연구팀은 초분광 이미징 장치의 분광 특성 결과를 활용하여, 배경 잡음 제거 (Dehazing) 알고리즘을 개발하고자 함. 고해상도 파장별 측정 결과를 활용하여, 외부 배경과 목표 대상 물체간의 대비(Contrast)를 높여 측정 이미지의 배경 잡음은 제거하고, 물체의 선명도는 높이고자 함.

- **(딥러닝 기반 화자인증 모델 개발)** 심층 신경망을 이용하여 기존의 특징 추출 방법을 data-driven transformation으로 대체하여 화자인증 모델을 개발하고자 함. 구체적으로, 지수적으로 증가하는 dilation factor를 이용하여 장기 의존성 문제를 해결한 Temporal convolutional networks (TCNs)기반의 화자인증을 위한 특징 모듈로 사용하며 최적의 특징을 탐색하는 연구를 진행함. 추가적으로 Squeeze and network (SENet) 모듈을 각각의 TCNs 내부 블록에 추가하여 channel-wise self-attention을 추가적으로 수행함.

- **(딥러닝 기반 화자인증 모델 고도화)** 실용적으로, 화자인증 기술이 원거리 혹은 잡음 환경에서 입력된 음성 신호에 대해서는 성능이 급격하게 저하된다는 문제점이 있음. 이를 극복하기 위해 딥러닝을 이용하여 공간 정보를 추출하고, 빔포밍과 같은 다채널 신호 처리 기법을 추가적으로 활용하여 입력된 신호에서 잡음과 같은 방해요소를 제거하고 화자의 음성을 강조하는 기술을 개발함. 또한 기존의 화자인증 기술은 입력된 화자의 음성의 길이가 충분히 길지 않을 경우 만족할만한 성능을 기대하기 힘들다는 한계점이 있음. 이를 극복하기 위해, 짧은 음성 입력만으로도 신뢰할만한 화자인증 성능을 확보하기 위한 연구를 진행함. 최종적으로 개발된 화자인증 모델의 성능을 유지하면서 디바이스 탑재 등과 같은 실용적인 애플리케이션을 고려하기 위해 모델의 파라미터 수를 줄이기 위한 연구를 진행함.

② 제2프로젝트 연구계획

가. 연구과제 개요

프로젝트명	국문	Edge 컴퓨팅 기반 블록체인 서비스 인프라 기술		
세부 프로젝트 책임자	성명	김종원	직위	교수
	소속기관	광주과학기술원	부서명(학과명)	AI대학원
	연락처	[redacted]	E-Mail	[redacted]
혁신도약형 과제 여부		X	기술분야 (소분류 기준)	블록체인-블록체인 기반기술

나. 연구내용

○ 본 프로젝트에서는 차세대 디지털 트윈 구축을 위해 필요한 네트워크 서비스 인프라 구축 과정을 연구함. 블록체인 기반 지능형 디지털 트윈 플랫폼을 운용하는 데에 필요한 네트워크 인프라의 자원 활용을 분석하고 신뢰성을 확보할 수 있는 시스템을 구축하는 것이 본 프로젝트의 목표임. 이를 위한 주요 연구 내용은 다음과 같음.

[2-1] Connected DataLake에 기반한 지능형 Edge Infra 기술
[광주과학기술원 김종원 교수 연구팀]

[2-2] 블록체인 플랫폼을 위한 신뢰-정의 네트워크 기술
[고려대학교 김황남 교수 연구팀]

- Connected DataLake에 기반한 지능형 Edge Infra 기술 연구에서는 D-N-A 패러다임에 충실하게 대응하는 유연한 데이터 공유가 가능한 에지 인프라의 구축을 목표로 함. 이때 D-N-A 패러다임에서 요구하는 데이터의 확장성 있는 수집·공유를 위해서는 'Connected DataLake' 개념에 따른 에지 및 코어 클라우드 인프라간의 유기적인 협력에 따른 블록체인 대응 데이터 공유 지원이 매우 중요함.
- 블록체인 플랫폼을 위한 신뢰-정의 네트워크 기술 연구에서는 블록체인 데이터 스트림 전송에 대한 안정성을 보장하고, 데이터 정보에 대한 신뢰성을 향상시킬 수 있는 신뢰 정의 네트워크를 설계, 개발함. 본 연구에서는 블록체인 플랫폼의 지속 가능한 운영성을 보장하기 위한 네트워크 신뢰성 개선 시스템을 구축하는 것을 목표로 함.

○ [2-1] Connected DataLake에 기반한 지능형 Edge Infra 기술

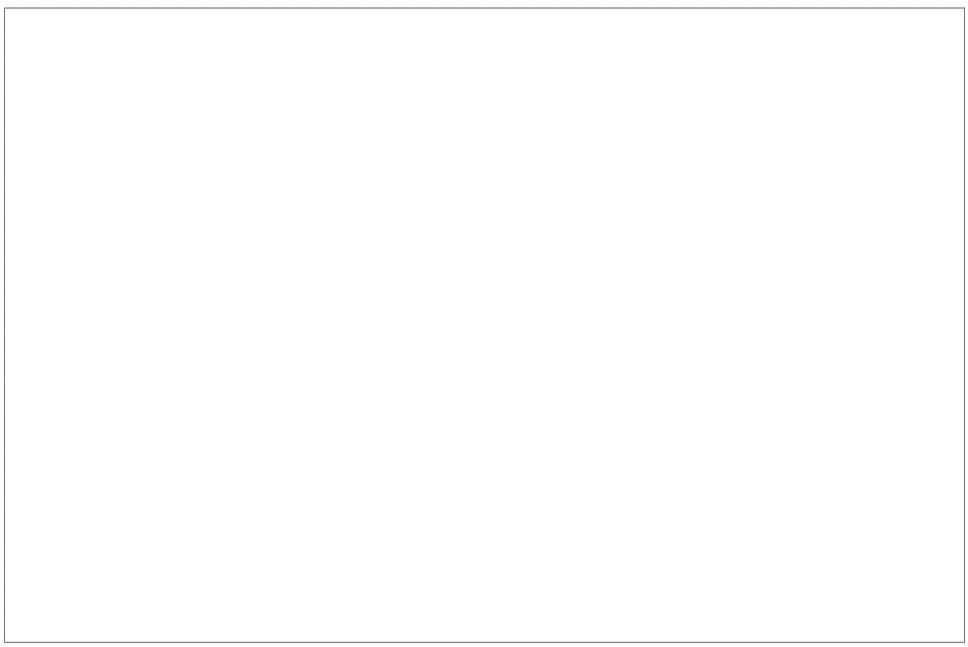


그림 14. Connected DataLake 기반 지능형 Edge Infra 기술

- **블록체인 대응 Cloud-native Connected DataLake 구성 설계**: 블록체인 대응 데이터 공유의 기능 측면 요구사항을 고려하여 Cloud-native Connected DataLake 프레임워크의 하드웨어/소프트웨어 구성을 설계함. 세부적으로 종단으로부터 에지 인프라로 유입되는 대용량의 데이터를 저장하고 이를 정제하여 블록체인 기반으로 검증된 상태로 공유하도록 지원함. 이때 Connected DataLake 인프라를 구성함에 있어서 컨테이너를 활용하는 Cloud-native 클라우드 방식으로 에지에 위치한 DataPond 클러스터의 노드 구성과 각 노드의 컴퓨팅·스토리지·네트워킹 자원 구성을 설계하고, 멀티사이트에 분산된 클러스터 노드 간에 고속의 대용량 데이터 교환이 가능하도록 계획함. 또한 구성된 에지

인프라의 하드웨어 설계를 바탕으로, 분산 클러스터상의 블록체인 대응 데이터 공유가 가능한 Connected DataLake 소프트웨어 구성에 관한 프레임워크를 설계함.

- **블록체인 대응 Cloud-native Connected DataLake PoC 검증**: 블록체인 대응 Cloud-native Connected DataLake 프레임워크의 설계를 바탕으로 광주과기원 내부에 에지 클라우드 기반의 DataPond 클러스터 하드웨어를 시험적으로 구축하고 점진적인 보완을 개시함. 이어서 광주과기원 내부에 DataPond 클러스터를 추가하여 멀티사이트 클러스터의 형태를 갖추면서, 설계된 블록체인 대응 Cloud-native Connected DataLake 프레임워크에 따른 사이트간의 블록체인 대응 데이터 공유에 대한 프로토타입의 가능성을 검증함.

- **멀티사이트 Connected DataLake를 위한 블록체인 대응 체계화된 데이터 저장·교환 기능 구현**: 블록체인 대응 Connected DataLake 프레임워크의 요소 기능인 데이터 태깅(tagging) 기능을 개발함. 즉 DataPond 클러스터에 수합된 다양한 종류의 데이터를 구분하기 위해, 미리 정의된 규격에 따라 데이터에 태그를 자동으로 추가하는 기능을 모색함. 태그에는 데이터 유형, 버전, 생성/업데이트 시간/장소, 관련주체(운영자/사용자) 등의 체계화를 위한 추가 정보를 기입할 수 있음. 그리고 분산된 DataPond 클러스터 상의 데이터에 대한 메타데이터를 병합/분할(merge/split)하고, 이를 분산된 클러스터에 배포하는 과정을 반복하여 Connected DataLake 전체에 걸친 데이터 셋의 메타데이터를 일치시키는 데이터 저장/교환 기능을 구현함.

- **블록체인 대응 체계적인 데이터 저장·교환을 지원하는 멀티사이트 Connected DataLake 구현**: Connected DataLake는 데이터 셋의 체계적인 관리와 물리적인 위치에 상관없는 데이터 접근을 지원해야 하므로, 추가 사이트에 DataPond 클러스터를 배포하여 3개 사이트간 10G 연동이 가능한 Connected DataLake 인프라를 확보함. 그리고 클러스터의 스토리지 자원 통합을 위하여 Ceph, MinIO 등과 같은 오픈소스 기반 Cloud-native 스토리지 소프트웨어를 적용하여 통합관리가 가능한 분산형 스토리지 클러스터로 구성하는 작업을 진행함. 결과적으로 3개 사이트의 DataPond 클러스터들이 상호 연동되어 블록체인 대응 데이터 교환 및 메타데이터 공유를 지원하는 초기 수준의 Connected DataLake 프레임워크를 구현함.

- **블록체인 대응 Connected DataLake를 위한 초고속 데이터 전달 기능 개발**: Connected DataLake의 데이터 공유를 위해 멀티사이트 DataPond 클러스터 간 대용량 데이터 교환이 가능하도록 초고속 데이터 전달 기능을 개발함. SmartNIC 기반 DPU(Data Processing Unit) 등을 활용하여 초고속의 데이터 처리/전송을 SmartNIC/DPU에서 오프로딩하도록 구현함으로써 블록체인 처리를 포함한 데이터 처리+전송 지연을 최소화하도록 시도함. 이를 통하여 데이터 저장·처리·전송 성능을 통합적으로 개선하는 블록체인 대응 Connected DataLake를 위한 오프로딩 기반 초고속 데이터 전달 및 공유 성능 개선을 연구함.

- **다중경로 초고속 데이터 전달 기능 개발을 통한 블록체인 대응 Connected DataLake 개선**: Connected DataLake 인프라의 네트워킹 가속화를 위해 다중경로 패브릭 네트워크(Fabric Network)를 활용하도록 시도함. 즉 분산된 클러스터 노드 간의 데이터 전달시 복수의 RDMA 기반 네트워킹 경로로 병렬 송수신하여 네트워킹 성능을 개선하고 전달

안정성을 개선하는 초고속 데이터 전달 기능을 구현함. 이를 통해 멀티사이트 DataPond 클러스터 간 초고속 데이터 전달을 지원하여 글로벌 네임스페이스(Global Namespace) 개념을 기반한 하나의 논리적인 데이터 스토리지 풀로 동작하도록 Connected DataLake 프레임워크를 개선함.

- 블록체인 대응 Connected DataLake 고도화를 위한 분산된 데이터 셋 동기 기능 개발: 시험 운영중인 Connected DataLake 인프라의 고도화를 위하여 구현된 프레임워크 요소 기능들의 개선을 진행함과 동시에 추가·수정된 데이터 셋을 멀티사이트 클러스터 전체에 신속하게 업데이트하는 동기 기능을 개발함. 그리고 Cloud-native 기반으로 자동화되고 용이한 데이터 저장·조회·활용을 지원하는 개방형 데이터 파이프라인 인터페이스를 개선함으로써 Connected DataLake 고도화를 준비함. 세부적으로는 Connected DataLake에 저장된 데이터 셋은 데이터 지속성과 접근 지연감소를 위해 지리적으로 분산된 다수의 복사본을 가질 수 있으므로, 이를 하나의 통합된 데이터 셋으로 다루기 위한 동기화 기술이 필수적임. 따라서 발생한 변경사항을 Connected DataLake 전체에 실시간으로 반영하는 데이터 셋 동기 기능을 변경치(delta, changed data) 중심의 초고속 업데이트를 통해 실현하고, 이를 블록체인에 대응하여 확정하도록 지원함.

- 분산된 데이터 셋 동기·공유 기능을 통합한 블록체인 대응 Connected DataLake 고도화: 다년간 운영과 보완으로 정비된 블록체인 대응 멀티사이트 Cloud-native Connected DataLake의 고도화를 완성하기 위해서, 개방형 데이터 파이프라인 인터페이스를 통한 초고속 대용량의 블록체인 대응 데이터 셋 동기화와 공유를 통합함. Connected DataLake의 고도화를 통해 블록체인 검증 데이터 또는 이를 활용하는 블록체인 응용들의 물리적인 위치에 관계없이 빠르고 용이한 저장·조회·획득 기반의 초고속 대용량 데이터 공유를 지원하는 고도화된 Connected DataLake 프레임워크를 완성하고 검증함.

○ **[2-2] 블록체인을 위한 신뢰-정의 네트워크 기술**

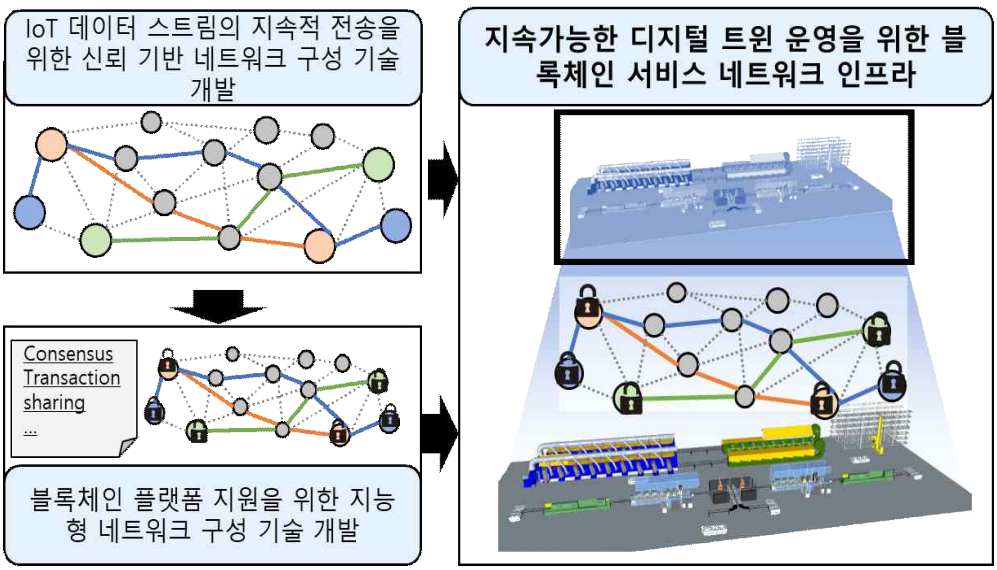


그림 15. 블록체인 플랫폼을 지원하는 네트워크 구성 인프라 구성

- **신뢰 정의 네트워크(Trust-Defined Network) 기술의 정의**: 신뢰 정의 네트워크 기술은, 네트워크를 이루는 각 기기(센서 단말, 라우터 및 스위치 등)의 네트워크 자원 사용 현황을 분석하고, 디지털 트윈 및 블록체인 운용에 필요한 각종 트래픽에 대하여 필요 자원을 보장할 수 있는 네트워크 제어를 수행하는 기술임. 기존의 네트워크 기술은 대체로 저지연성 및 수율(data rate)을 최적화하는 성능 기반 기법을 도출하였으나, 이는 특정 지역의 데이터 집중적으로 네트워크 자원을 점유하거나 일부 데이터에 대해 지연 전송 및 전송 실패를 야기하는 등 블록체인 운용에 있어 부적절한 결과를 도출할 수 있음. 또한, 오직 비용측면에서 최적화된 전송 경로를 설정하게 되어있기 때문에 링크의 안전성을 100% 담보할 수 없음. 따라서 해당 세부 프로젝트에서는 네트워크에 포함된 각 노드 및 네트워크 장비, 링크에 대한 신뢰성을 평가하고, 이를 기반으로 네트워크를 제어함으로써, 주어진 네트워크 요구조건에 대하여 높은 신뢰도를 달성할 수 있는 신뢰 정의 네트워크를 개발함. 또한 이러한 네트워크를 구축하기 위해 공유되는 네트워크 구성요소간의 제어 데이터의 신뢰성을 평가할 수 있는 체계또한 개발함.

- **신뢰 정의 네트워크 기반 고신뢰성 블록체인 네트워크 제어 인프라 구축**: 신뢰 정의 네트워크의 기반 기술을 통해, 본 프로젝트에서는 블록체인 등의 분산 검증 네트워크의 고신뢰성을 확보할 수 있는 네트워크 제어 인프라를 구축하는 것을 목표함. 본 연구 과제는 블록체인 플랫폼 운영 및 이에 관련한 제어 네트워크에 대한 검증 및 보안 기술을 연구/개발하며, 해당 기술을 운용하기 위한 네트워크는 높은 신뢰성 및 자원 효율성을 보장해야 함. 본 세부 프로젝트에서는 블록체인의 각종 네트워킹 요소(블록 배포, 합의 알고리즘 등)를 분석함으로써, 블록체인 운용을 위한 검증자 간 네트워킹의 고신뢰성을 확보할 수 있는 네트워크 제어 시스템을 구축함.

- **신뢰 정의 네트워크 제안 및 데이터 스트림 연결성 평가 모델 설계**: 블록체인 플랫폼에서는 트랜잭션들이 지속적으로 노드들 사이에 전송되어야 하므로 이를 전달하는 매개체인 네트워크의 신뢰성이 매우 중요함. 따라서 각종 라우터 및 스위치에 대한 신뢰도를 다양한 파라미터를 활용해 수식화하고 이를 기반으로 네트워크 연결 구성 및 전송을 진행함. 이와 연계하여 각종 IoT 기기에 대한 데이터 스트림 특성을 파악하며, 연결성을 실시간으로 점검할 수 있는 평가 모델을 설계함. 본 평가 모델에 기반하여 디지털 트윈과 연동되는 각종 네트워크 구성요소 (IoT 디바이스, 게이트웨이, 라우터 등) 에 대한 신뢰성을 확보하고 필요할 경우 추가적인 요구조건을 도출함으로써 네트워크를 지능적으로 제어할 수 있는 기반을 확보함.

- **분산적 신뢰 기반 네트워크 관리 기술 개발**: 블록체인 플랫폼의 규모가 확대될 경우, 중앙화된 TDN 제어 방식으로 모든 데이터 스트림을 관리하기에는 네트워크 및 계산적 비용이 높을 수 있음. 또한, 집중 형태이므로 낮은 장애허용성을 가진다. 이를 개선하기 위해, TDN 컨트롤러를 중복·분산된 형태로 네트워크 토폴로지 상에 배치함. 너무 밀집되게 컨트롤러를 배치할 경우 명령 중복 및 네트워크 운용 측면에서 비용 낭비가 발생하므로 노드 밀집도, 해당 로컬네트워크 특성 등에 따라 계산된 최적화 배치 모델링의 연구가 필요함. 이와 더불어 네트워크의 구성요소를 상황에 맞게 여러 수평/수직 계층으로 구분하고, 계층별로 자원을 관리함으로써 데이터 스트림의 안정성 및 성능을 높일 수 있는 기술을 개발함.

- **지속가능한 분산형 TDN 컨트롤러 운용을 위한 소규모 프라이빗 블록체인 도입**: 기존 소프트웨어 정의 네트워크의 경우 하나의 컨트롤러가 모든 네트워크 상황을 통제하기 때문에 운용에 충돌이 생길 일이 없음. 하지만 지속 가능한 신뢰 정의 네트워크 도입을 위해 분산형 TDN을 도입 시 각 컨트롤러 별로 가지고 있는 네트워크에 대한 정보가 일치하지 않을 수 있고 이는 트래픽 포워딩에 혼란을 초래함. 따라서 TDN을 운영하는 시스템 레벨에서 소규모 프라이빗 블록체인을 도입하여 각 컨트롤러가 해당 블록의 정보를 보고 각자 최신화된 네트워크 정보를 갱신할 수 있도록 함. 합의 과정은 합의 주체인 컨트롤러의 노드 개수가 많지 않기 때문에 외부의 악의적인 공격에 효과적으로 대응할 수 있는 소규모 합의 알고리즘을 개발함.

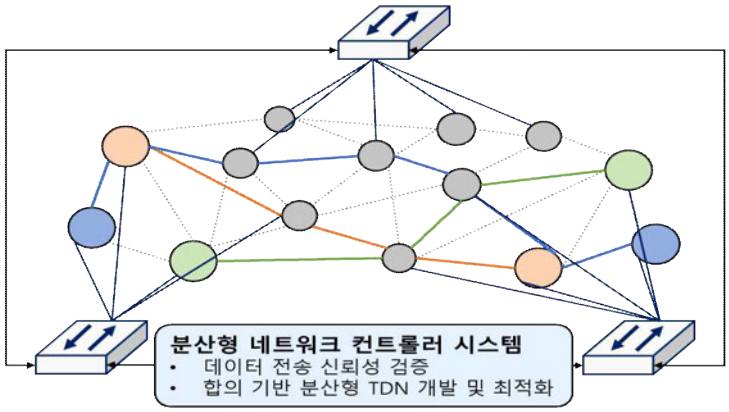


그림 16. 분산형 네트워크 컨트롤러 시스템

- **노드간 멀티캐스트 최적화 기술 개발**: 특정 블록체인 검증자(agent)에 다수의 센서 데이터 스트림이 집중되거나 특정 데이터 스트림의 검증이 지연되는 현상을 방지하기 위해, 디지털 트윈에 필요한 데이터 스트림이 블록체인 네트워크에 확산될 때 소요되는 시간을 단축시키고 신뢰성을 보장할 수 있는 오버레이 네트워크 지원용 멀티캐스트 최적화 기술을 개발함. 또한, 블록체인 플랫폼에서 높은 리소스 비중을 차지하는 트랜잭션 플러딩 과정을 최적화할 수 있는 기술을 개발함.

③ 제3프로젝트 연구계획

가. 연구과제 개요

프로젝트명	국문	DID 블록체인 기반 IoT 인증 및 보안 기술 개발		
세부 프로젝트 책임자	성명	김기형	직위	교수
	소속기관	아주대학교	부서명(학과명)	사이버보안학과
	연락처	[redacted]	E-Mail	[redacted]
혁신도약형 과제 여부		X	기술분야 (소분류 기준)	블록체인-블록체인 기반기술

나. 연구내용

○ 본 세부 프로젝트에서는 DID 블록체인 기반 IoT 인증 및 보안 기술 개발을 목표함. DID 응용 서비스에 대한 안전과 신뢰를 위해서는 IoT 디바이스의 인증·허가 기술이 필수적이며 탈중앙화 및 자기주권 인증 기술이 IoT 기술의 핵심 기반임. 이를 위한 주요

연구내용은 다음과 같음.

[3-1] DID 블록체인 기반의 상호 인증/인가 기술 연구
[아주대학교 김기형 교수 연구팀]

[3-2] 물리계층을 활용한 탈중앙화 블록체인 기반 IoT 기기 인증 및 접근제어 기술 개발
[광주과학기술원 황의석 교수 연구팀]

- [3-1] 자기주권형 탈중앙화 인증/인가 기술 개발을 위하여 IoT 생태계에 최적화된 안전한 데이터 전송 기법과 DID 블록체인 기반의 인증/인가 기술을 연구함.
- [3-2] 물리계층을 활용한 블록체인 기반의 탈중앙화 IoT 기기 인증 및 접근제어 기술 개발 연구에서는 제한된 자원을 가진 IoT의 보안취약점 해결을 위하여 블록체인과 물리계층 기술을 결합한 지능형 보안관리 시스템을 연구함.

○ **[3-1] DID 블록체인 기반 상호 인증/인가 기술 개발**

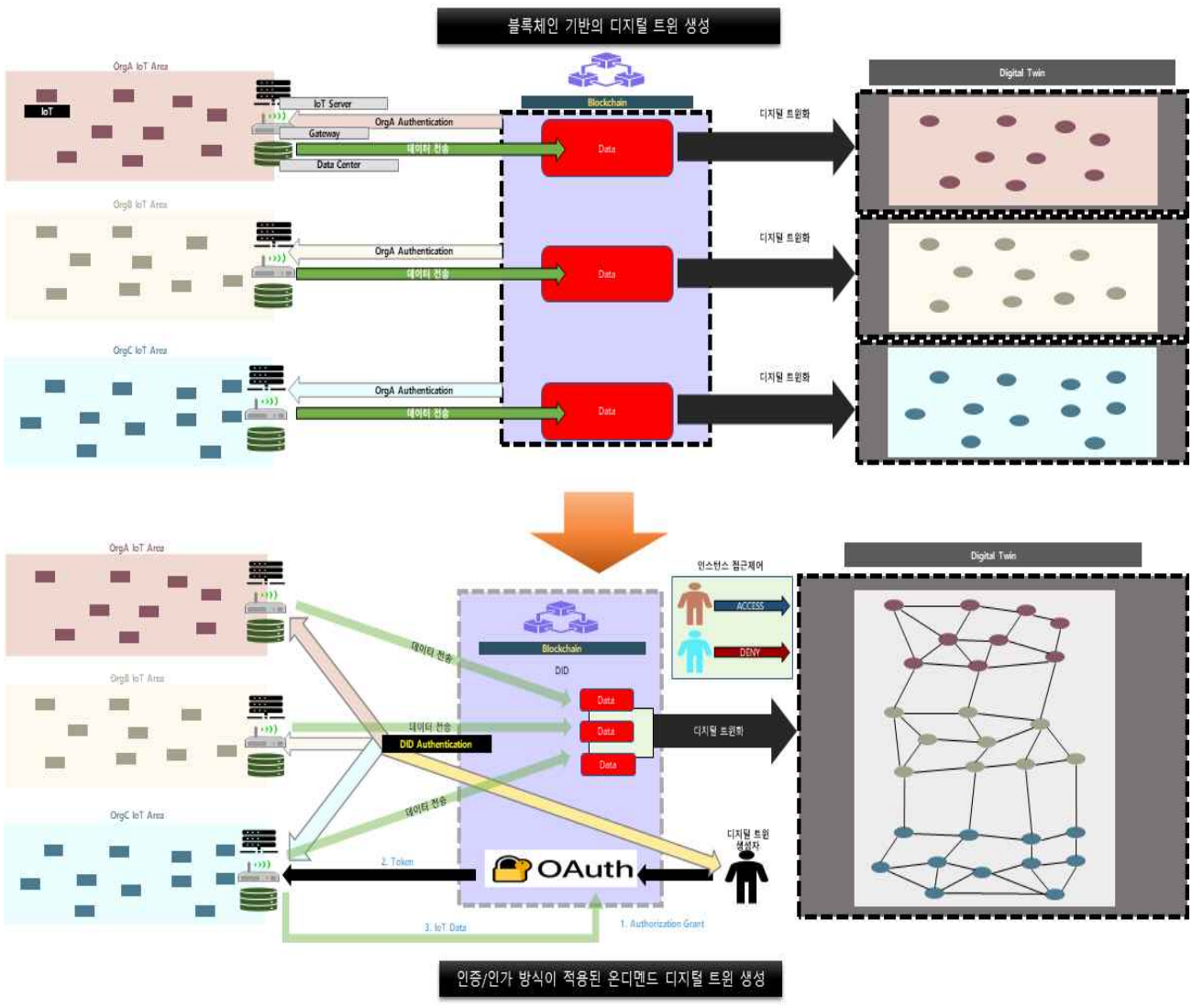


그림 17. 인증/인가 방식이 적용된 온디맨드 디지털 트윈 생성

- **DID 블록체인 기반 인증/인가 서비스 요구사항 분석**: 자기 주권성을 확보하는 DID기반 기술을 적용할 인증/인가 서비스를 조사하고 DID 응용의 단계별 기술 구성 및 특성을 분석하여 서비스 적용 방안을 연구함.

- **IoT에 적합한 DID 블록체인 인증 기법 연구**: 안전한 IoT 생태계를 유지 · 운영하기 위해서는 기존 IoT 네트워크의 구조와 특성을 분석하여 이를 반영한 DID 블록체인 네트워크 구조를 설계함.

- **안전한 IoT Data 전송 방법 연구**: DID 블록체인을 활용한 인증/인가 과정에 IoT 단말에서 전송되는 자료들을 안전하게 보호하고 검증하기 위한 전송 기법을 연구함.

- **DID기반의 자기주권형 액세스 토큰 발행 연구**: OAuth2.0와 DID를 활용하여 액세스 토큰을 발행해 인가를 부여하고 사용자 정보를 안정적으로 가져올 수 있는 액세스 토큰 활용 방법을 분석함. 공개된 원장 블록체인을 활용하여 고유한 DID를 각 객체에 부여함으로써 관리용이성 및 타 플랫폼과의 투명성을 높이는 연구를 수행함.

- **IoT 환경의 DID 블록체인 기반 접근제어 방법 연구**: 분산된 환경의 IoT 환경에서 검증 가능하고 효과적인 접근제어를 통한 자기 주권형 액세스 토큰의 저장 기법을 연구함. DID 블록체인을 이용해 IoT 단말에서 직접 검증할 수 있는 접근제어를 통해 각 IoT 단말을 식별하고 인증/인가 시도를 제어하는 기법을 연구함.

- **영지식 증명 기반의 자격증명 방법 연구**: 블록체인에 저장되는 정보 중 개인정보 · 민감정보를 식별하여 영지식 증명 기술을 통해 이를 보호하는 기술을 연구함.

- **DID 인증/인가 기반 블록체인 응용 시스템 연구/개발**: DID 블록체인 기반 시스템 간에 요구되는 특성들을 분석하여 사용자를 안전하게 인증 및 인가하는 응용 시스템을 연구하고 개발함.

- **DID 인증/인가 기반 블록체인 응용 시스템 성능 측정/평가**: 최종 연구 결과물로 생성될 DID 블록체인 기반 인증/인가 시스템의 성능을 테스트베드를 구성하여 자체적으로 측정, 평가함으로써 자체 목표 달성률 측정 및 피드백을 통해 개선 · 고도화함.

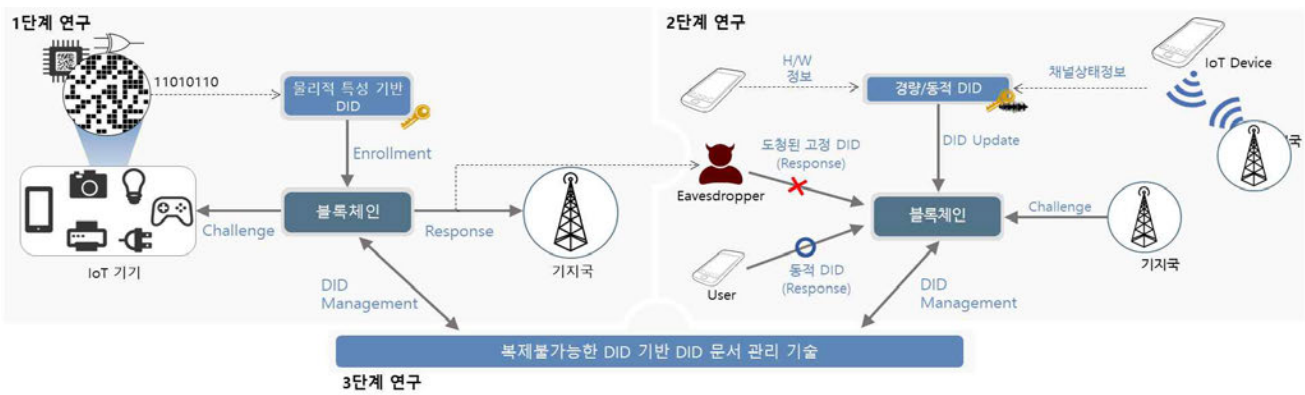


그림 18. 물리계층을 활용한 블록체인 기반 DID 생성 기술 및 인증 시스템 개념도

○ **[3-2] 물리계층을 활용한 탈중앙화 블록체인 기반 IoT 기기 인증 및 접근제어 기술 개발**

- **보안통신을 위한 기존 암호화 방식의 보안취약점 분석**: 기존 중앙형 보안체계의 보안

취약점 및 암호화 방식의 한계점을 분석하고, 이를 보완하기 위한 분산형 인증시스템을 설계함.

- 물리계층 기반 경량 DID 생성 알고리즘 개발: 암호 알고리즘 기반의 인증 방식은 제한된 하드웨어를 가지는 IoT환경에 적합하지 않음. 따라서 IoT 기기의 고유의 물리적 특성을 활용하여 많은 계산자원이 요구되지 않는 경량 DID 기술을 연구함.

- 대규모 IoT 환경에서 신뢰성을 가지는 DID 생성 기술개발: 물리적 특성 기반의 DID의 경우 도청 등의 공격으로 인해 노출되게 되면 DID를 쉽게 변경할 수 없으므로 보안성이 악화될 수 있음. 따라서, 채널 상태 정보 등을 활용해서 환경에 따라 DID가 생성/업데이트 할 수 있는 알고리즘을 연구함.

- 스마트 기지국의 DID 관리 프로토콜 개발: 초기 인증을 시도하는 IoT 기기의 인증, 생성된 DID 관리 그리고 동적 DID 업데이트로 구성된 프로토콜을 개발함.

- 복제불가능한 DID 기반 DID 문서 생성 및 제어기법 개발: 기지국에서 DID 기반 인증 블록 생성 및 기지국 간 공유를 수행하는 프로토콜을 설계함. 기기 인증 시, 트랜잭션 블록 생성을 통해 기지국 간 블록체인을 구성하는 분산원장 기술을 구현함.

- 이종 네트워크 사이의 프로토콜 호환성 연구: 타 플랫폼 또는 이종 네트워크 간 IoT 기기의 인증을 안전하게 구현하는 인증체계 및 접근제어기법을 연구함.

- IoT 기기를 활용한 test-bed 구성 및 프로토콜 테스트: 스마트 기지국의 DID 문서 생성 제어, DID의 인증 적합성, 프로토콜의 적절성을 평가하기 위해, universal software radio peripheral (USRP) 및 IoT 기기를 이용하여 성능을 평가함.

3 연구개발과제의 추진전략·방법 및 추진체계

1) 연구개발과제의 추진전략 · 방법

① 연구개발 과제의 추진전략



그림 19. 총괄 추진전략

○ 본 과제를 성공적으로 수행하기 위하여 세부 팀별 간 긴밀한 협조 체계를 구축하고자 함. 먼저 세부연구 주제를 정립하고 연구 연속성을 확보한 후, 연구목표에 대한 최적의 결과를 도출함. 이후 신규 핵심기술의 확보와 연구결과의 활용에 주력함.

② 팀별, 센터별 구성 및 공동 연구 추진 전략 및 프로세스

그림 20. 센터 연구개발 추진전략 및 프로세스

○ 본 연구는 생체정보 및 영지식 증명 기반 블록체인 플랫폼, Edge 컴퓨팅 기반 서비스 인프라 기술 및 DID 블록체인 기반 IoT 인증 기술에 기반을 두고 있음. 때문에, 관련 연구 개발 경험이 있는 여러 연구팀들이 모여 연관 주제별로 세부 팀별 구성하여 공동 개발하고자 함.

- [1세부] 광주과학기술원 이홍노, 신종원 교수 연구팀은 유니버셜 영지식 증명 기술, 탈중앙화 보안성 ECCPoW 기술 및 화자인식을 위한 1D 신호처리 기반 음성(Speach) 신호처리 기술을 보유함.
- [1세부] 서울대학교 문수묵 교수 연구팀은 가상머신 및 Edge 컴퓨팅 기반 블록체인 최적화 기술을 보유함.
- [2세부] 광주과학기술원 김종원 교수 연구팀은 AI 기반 SmartX 서비스 제공을 위한 Cloud-native 컴퓨팅 기술을 보유함.
- [2세부] 고려대학교 김황남 교수 연구팀은 WLAN 및 군집형 무인이동체 운용을 위한 네트워킹 기술을 보유함.
- [3세부] 아주대학교 김기형 교수 연구팀은 IoT(LoRA) 단말 및 데이터 인증 무결성 확보를 위한 블록체인 기반 보안 프레임워크 (BIoT) 기술을 보유함.
- [3세부] 광주과학기술원 황의석 교수 연구팀은 Cellular IoT 환경 기반 물리계층 보안통신 기술 및 AI 기반 에너지 재난대응 플랫폼 기술을 보유함.

○ 광주과학기술원 이홍노 교수 주관 하에 각 세부 연구팀은 영지식 센싱, 암호인증 및 클라우드 서비스 융합을 위한 개별 연구 주제에 대해 고찰함. 1세부 연구팀에서는 영지식 증명을 통한 사용자 인증 과정을 통해 신뢰성있는 People centricity 블록체인 대해 고찰하며, 2/3세부 연구팀에서는 서비스 인프라 및 IoT 보안 기술을 통해 Location Independence 블록체인에 대해 고민함. 마지막으로 각 세부 연구팀은 주기적 연구 모임을 통한 고찰 및 결과물 공유를 통해, 서로간의 혁신적 기술 융합을 추구함.

- [1세부] 본 과제의 혁신 도약형 연구 기술인 DIDs 및 ZKP 기반 탈 중앙성 블록체인 기술은 주관기관인 광주과학기술원 이홍노 교수 연구팀에서 주도적으로 개발을 수행함.
- [1세부] 광주과학기술원 신종원 교수 연구팀은 1D 신호처리 기반 화자인증 기술을 통해 신뢰성 있는 사용자 인증 기술 개발을 수행함.
- [1세부] 서울대학교 문수묵 교수 연구팀은 데이터 경량화 기술, 비잔틴 저항성 확보 기술 및 신뢰실행환경 기반 블록체인 경량화 기술 개발을 수행함.
- [2세부] 광주과학기술원 김종원 교수 연구팀은 지속가능한 디지털 트윈 실행을 위한 Connected DataLake 기반 지능형 Edge Infra 기술 개발을 수행함.
- [2세부] 고려대학교 김황남 교수 연구팀은 블록체인 플랫폼 지원을 위한 신뢰 정의 네트워크 (Trust-Defined Network, TDN) 기술 개발을 수행함.
- [3세부] 아주대학교 김기형 교수 연구팀은 DIDs 기반 인가 및 시스템 접근 제어 기술 개발을 수행함.

- [3세부] 광주과학기술원 황의석 교수 연구팀은 물리계층 식별자를 활용한 탈 중앙화 IoT 장치 접근제어 기술 개발을 수행함.

③ 영지식 센싱, 암호인증 및 클라우드 서비스 융합 플랫폼 구축 방안

○ 단계별 연구 결과물에 대한 우수성 및 서비스 가능성을 제시함.

- 1단계에서는 블록체인 지능 플랫폼 핵심기술의 검증을 통해 개발 기술의 우수성을 증명
  - 세부 팀별 SCI논문, 기술 표준화 및 특허, 데이터베이스 구성 등을 통한 핵심기술의 우수성 증명
- 2단계에서는 개별 핵심기술 개발 완료 및 DID-ZKP 기반 인증 블록체인 클라우드 플랫폼 프로토타입 개발해 기술의 성능 평가 및 확장성을 제고함.
  - 생체정보 및 ZKP 기반 고 신뢰성 블록체인 개발
  - Edge computing 및 신뢰 정의 네트워크 기반 블록체인 인프라 기술 개발
  - DIDs 및 물리 계층 식별자 기반 블록체인 네트워크 접근 제어 기술 개발
  - 블록체인 기반 디지털 트윈 응용 서비스 개발
  - 탈 중앙화, 초경량 / 초확장성, 신뢰성 및 접근 제어 능력 등 기술 적용에 따른 블록체인 네트워크 성능 검증
- 3단계에서는 핵심 기술 보완 및 인증 블록체인 클라우드 플랫폼 완성도 재고
  - 대규모 노드 기반 디지털 트윈 응용 서비스 실증
  - 개발 기술의 노드 확장에 따른 성능 평가 및 분석
  - 기술 보완에 따른 디지털 트윈 플랫폼 완성도 재고

④ 기관별 적극적 정보 교류를 통한 효율적인 연구 개발 수행

○ 본 과제를 성공적으로 수행하기 위하여 세부 팀별 간 긴밀한 협조 체계를 구축함.

- 산・학・연 관련 전문가 초빙 추진하여 세부 기술 자문 추진(수시)
- 세부 팀별 / 센터별 주기적 연구 세미나 개최를 통한 연구 진행 사항 공유
  - 반기별 1회 [세부팀], 주기별 1회 [센터]
  - 연구 결과물 공유 및 융합 혁신 추구
  - 블록체인 네트워크 플랫폼 및 IoT 센서 등의 H/W 뿐만 아닌 물리 엔진 및 네트워크 시뮬레이션 툴 등의 S/W 등 다양한 연구 인프라 공유 추구
  - Github을 통한 SW 공유로 연구 공동 추진
  - Slack을 통한 File 및 업무프로세스 공유로 연구 협업 추진

○ 국/내외 커뮤니티 참여를 통한 기술 교류 및 외부 전문가 활용

- 개발 환경 구축, 서비스 구축 등 본 과제의 핵심 기술이 아닌 부분은 외부 전문가를 통한 용역 과제로 진행하여, 참여자들이 핵심 분야에 집중할 수 있도록 도모
- 국/내외 학회 참석을 통한 최신 디지털 트윈 / 블록체인 기술 논의 및 인적 네트워킹 추구

2) 연구개발과제의 추진체계

① 연구개발 추진체계

○ 본 과제를 성공적으로 수행하기 위하여 세부 팀별 간 긴밀한 협조 체계를 구축하고자 함. 먼저 세부연구 주제를 정립하고 연구 연속성을 확보한 후, 연구목표에 대한 최적의 결과를 도출함. 이후 신규 핵심기술의 확보와 연구결과의 활용에 주력함

- 국제 학회 등을 참석하여 최신의 블록체인 동향을 분석하고 최신 기술이 설계에 피드백으로 반영되어 원천연구기술에 반영되어 개발될 수 있도록 추진 및 관리 실시

- 특히 새로운 규격이 제시될 경우 이를 수용할 수 있도록 설계 및 구현 시 확장성을 고려할 예정

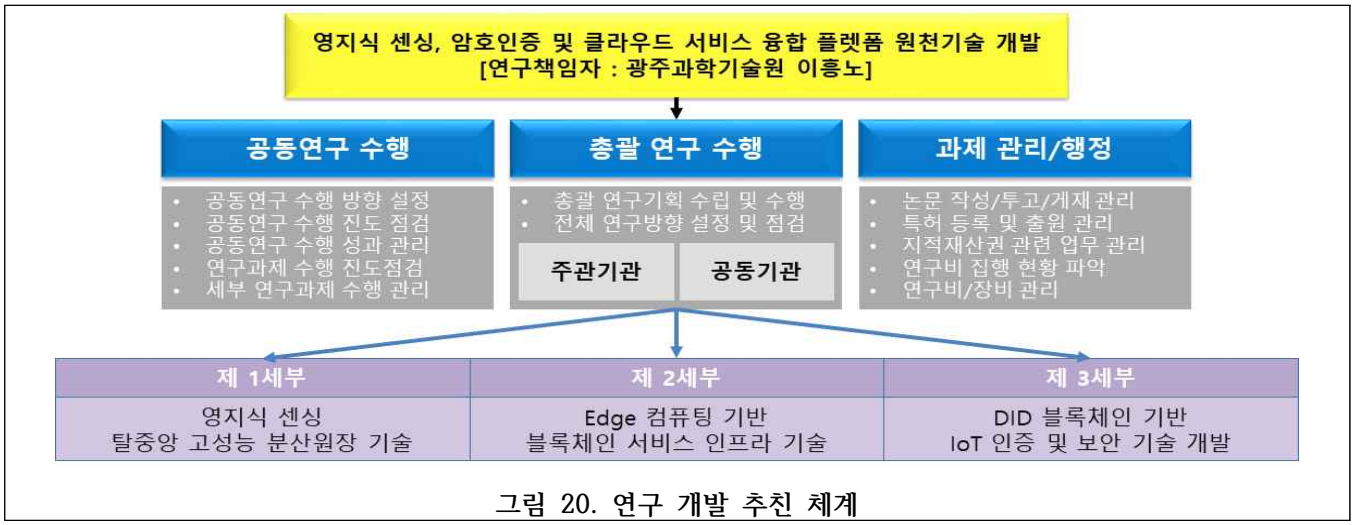


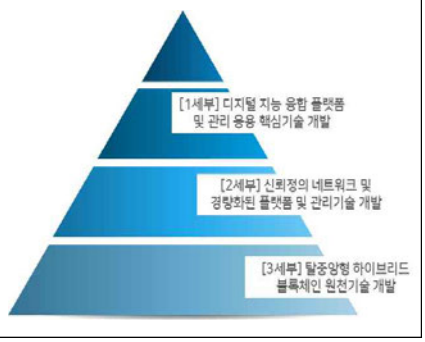
그림 20. 연구 개발 추친 체계

분류		참여기관
산학연관	대학	GIST, 서울대, 고려대, 아주대
	참여기업	**Consensys, 퓨처센스, 시티랩스, 헥슬란트, 엘에스웨어(주)**
	연구소/센터	서울블록체인지원센터, 한국블록체인학회, 블록체인법학회, GIST인공지능연구소, GIST AI대학원, GIST블록체인경제센터, 한국ICT융합협회
	관	광주광역시, 과기정통부

② 연구그룹의 연계성

○ 미래 블록체인 핵심 융합기술을 확보하기 위하여 각 그룹은 다음과 같이 역할을 분담

- 각 그룹 간 긴밀한 연계를 통하여 성공적인 블록체인 지능 융합 핵심기술 확보 및 플랫폼 기술 경쟁력 확보와 국내외 핵심산업 혁신성장을 선도할 것임

그룹	그룹 목표	연계성
1세부	• 영지식 센싱 및 정보 관리 기술에 기반한 분석과 데이터 거래 및 공유 서비스가 가능한 디지털 지능 융합 플랫폼 및 관리 응용 핵심기술 개발	
2세부	• Cloud Native 블록체인 온디멘드 서비스를 위한 신뢰 정의 네트워크 및 경량화된 블록체인 플랫폼 및 관리 기술 개발	
3세부	• PUF 와 DID 등 개인 인증과 허가 에 기반한 IOT 와 데이터 스마트 계약을 위한 탈 중앙형 하이브리드 블록체인 원천기술 개발	

③ 품질 평가 항목 및 내용

○ 본 연구팀은 개발한 영지식 센싱, 암호인증 및 블록체인 기반 클라우드 서비스 융합 기술에 대해, 다음의 품질 평가 항목 및 내용을 적용하고자 함.

대항목	중항목	소항목	설명
품질	효율성	처리 성능	• 처리 성능은 블록을 처리하는 성능으로 블록 생성 시간, 트랜잭션 사이즈 등이 영향을 주며 서비스 별로 처리 성능에 대한 정의 (범위)를 명확히 해야 함 • 블록의 처리 기법에 따른 트레이드-오프가 발생할 수 있어 다른 평가 항목 간의 유기적인 관계를 고려
		네트워크 성능	• 네트워크 성능은 시스템 전체 성능에 가장 많은 영향을 줄 수 있음 • 네트워크 성능에 변화를 줄 수 있는 기법(전용망 설치 등)이 있으나 이에 따른 트레이드-오프가 발생할 수 있어 다른 평가 항목 간의 유기적인 관계를 고려
		블록 확정 성능	• 블록체인시스템의 블록 확정은 합의 방식(PoW, PBFT 등)에 따라 결정 • 합의 방식에 따른 특징, 성능 등의 차이가 있어 적용 시 고려
		참조 성능	• 블록체인에 저장된 거래내역 등을 참조하는데 소요되는 시간을 확인
	호환성	중앙집중형	• 블록체인 시스템이 중앙집중형 시스템 또는 이기종 블록체인 시스템과 연계 될 경우 시스템 간의 호환 여부를 확인
		이기종	
	확장성	처리 성능 향상 가능성	• 사용량/데이터 증가로 데이터 처리 / 전송 속도 문제의 해결을 위해 처리 성능과 네트워크 성능의 향상 가능성을 확인
		네트워크 성능 향상 가능성	
		용량 확장성	• 블록의 용량을 확장할 경우 합의에 참여 가능한 노드의 수가 감소(예: 저성능 노드 등의 제외)할 수 있어 합의의 신뢰성 저하가 우려되므로 이를 고려한 용량 확장성을 평가
		노드 확장 정도	• 노드 수의 증가로 인해 네트워크가 동시에 처리/전송할 수 있는 노드의 수를 초과할 경우 처리/전송되지 않는 노드가 발생할 수 있어 네트워크 성능을 고려한 노드 확장성의 평가
	신뢰성	성숙성	• 명시된 요구사항의 구현정도, 노드의 장애 유무 확인, 부정 공격 방지, 장애회복 시간 등으로 전체 시스템의 신뢰성을 평가 • 분산 환경에서의 신뢰성은 전체 노드의 성능 및 운영 환경에 따라 결정되므로 이에 대한 고려가 요구
		가용성	
		장애 허용성	
		회복성	
	보안	기밀성	• 암호화된 데이터 전송 및 접근 제어 등을 통한 데이터의 기밀성, 무결성 확보에 대한 평가
		무결성	
		부인방지	• 거래에 대한 부인방지와 노드 간 공유된 자원(데이터)의 동기화 여부에 대한 진정성을 평가 • 부인방지와 데이터의 진정성은 합의 방식의 특성에 따라 달라 질 수 있어 사용하는 합의 방식의 특징과 장단점의 • 분석이 요구
		진전성	
	이식성	적응성	• 블록체인 플랫폼 또는 서브시스템이 중앙집중형 시스템으로 적용 가능 여부에 대한 평가
		변경 용이성	• 사용 중인 블록체인 플랫폼 또는 서브시스템이 다른 형태로 대체 가능 여부에 대한 평가

④ 센터 조직체계 및 운영계획

○ 연구 센터 조직 체계

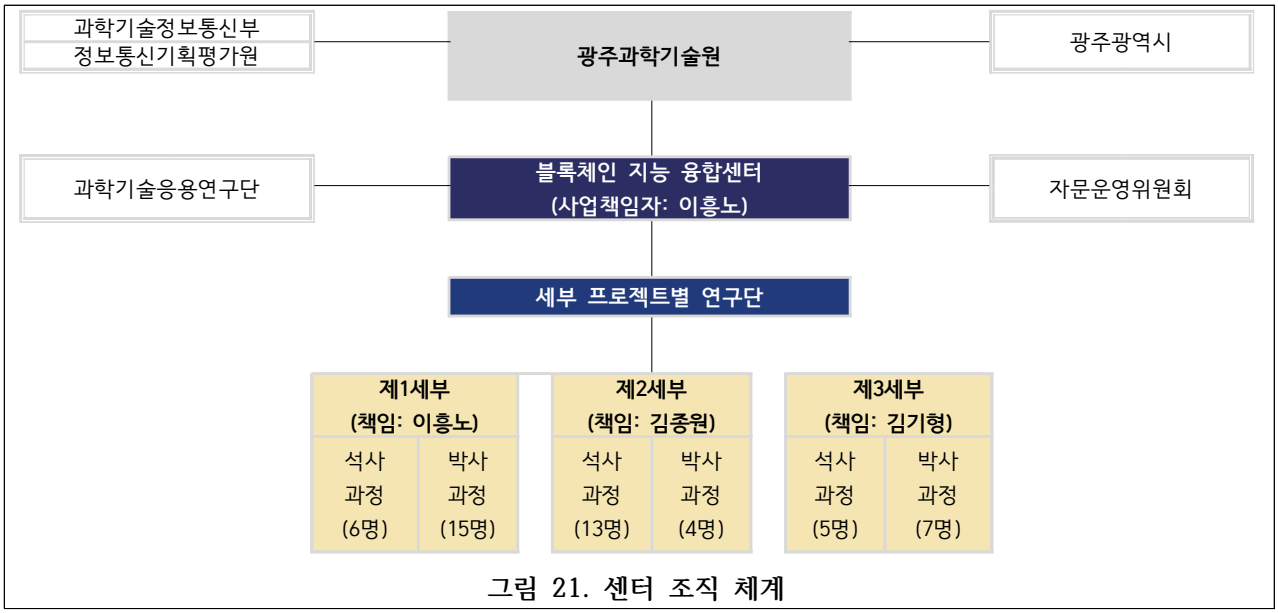


그림 21. 센터 조직 체계

- 광주광역시: 예산 지원

- 자문운영위원회: 기술·정책·법률·비즈니스와 관련된 각 분야 전문가·대기업 종사자·벤처 투자자 대표 등이 참여하며, 세부 프로젝트별 연구진 및 자문위원이 참여한 **프로젝트 발표회를 통해 발전·투자·창업의 결과로 맺을 수 있도록 전문가 자문위원회를 운영함**

  - 각 세부연구과제별 논의 및 협업을 위해 GIST 서울사무소(서초구 소재)의 지리적 이점을 활용, 각 대학별 온오프라인 정기회의 등이 원활하게 이루어질 수 있도록 지원
  - 아래 표는 현재 GIST블록체인인터넷경제센터(이흥노 센터장)를 자문협력하는 위원회 명단임. ITRC사업 수탁시, 블록체인 지능 융합 센터의 자문위원으로 초대하여 위원회를 구성할 예정임.

자문운영위원회 명단		
외부 자문위원		내부 자문위원
- 박광우 KAIST 금융전문대학원 원장 - 김형중 고려대 정보보호대학원 교수 - 박수용 한국블록체인학회 회장·서강대 컴퓨터공학과 교수 - 이영환 ㈜딜라이트체인 대표 - 정상호㈜ 델리오 대표 - 장중혁 ㈜ 아톰릭스컨설팅 이사 - 김종승 SK텔레콤 팀장 - 구태언 법무법인 린 TEK&LAW 변호사	- 이정엽 블록체인법학회 회장/의정부지법 부장판사 - 오정근 한국 ict 융합협회 회장 - 정순형 ㈜ 온더 대표. - 백양순 한국ICT융합협회 대표 - 임명수 서울블록체인지원센터 센터장 - 신동명 엘에스웨어(주) 연구소장 - 안다미 퓨처센스 대표이사	- 임혁 전기전자컴퓨터공학부 교수 - 김상호기초교육학부 교수 - 김희삼 기초교육학부 교수 - 진규호 기초교육학부 교수 - 안창욱 인공지능 대학원 교수 - 전해곤 인공지능 대학원 교수 - 김건우 기초교육학부 교수

○ 연구 센터 운영 방안

- 본 연구센터는 블록체인 지능 융합 연구에서의 국내 핵심 연구센터로 거듭나기 위해, ‘핵심 거점 연구센터 추진’, ‘기술 자생력 확보’ 및 ‘외부 기관과의 협력 등’ 다음과 같이 센터를 운영하고자 함.



그림 23. 블록체인-지능 융합 연구센터 운영방안

항 목	내 용
연구센터 운영계획	• 광주 및 서울을 포함한 전국을 커버하는 핵심 연구센터로 출발하여 관련분야의 연구와 산학협력 활동을 증진시키기 위한 **연구활동 활성화와 다양한 프로그램을 시도**할 예정 • 장기적으로 대외 연구센터 지원 사업 확보를 위한 연구주제 선정, 공동 작업에 따른 소요예산 확보가 필요하며 **기 수행 핵심 과제들을 기반으로 연구** • 미래형 블록체인 지능 융합 신기술에 분야에 관한 세미나의 정례화를 통하여 관련 기술 트랜드를 소개하고, 연구원들의 지속적인 관심을 유도하고, 공동 수행된 융합기술을 바탕으로 전국의 **산업체를 대상으로 한 산업 기술 워크숍 및 세미나 진행**
국내 블록체인 연구를 선도하는 센터로의 발전 방안	• 블록체인산업 분야의 기술개발에 필요한 핵심원천기술 및 응용연구를 수행하고 연구 거점 대학의 융합 연구 센터로 국내 **특성에 알맞은 정보화 혁신사회 실현** • 다학제 권역거점형 연구센터로서 블록체인 및 기반기술의 융합은 물론, 사회 전반에서 요구되는 융합화 추세와 더불어 융합연구 활동에 적극 대처하여 4차산업 핵심기술을 통하여 블록체인 융합분야의 연구 활성화 예정 • 지역의 블록체인 및 핵심 기술과 관련된 연구소 및 산업체들과 활발한 산학연활동을 추진하여 세계 **경제와 기술에 공헌하는 연구센터로 확립**
연구센터 자생력 확보	• 산업화가 가능한 고부가가치 기술들의 개발에 중점을 두어 특허등록과 기술이전을 활성화하여 자립기반을 구축하며, **기술이전 건수를 높여 재정확보** • 외부 연구비 수주를 장려하고, 연구소 연구원의 국제공동연구 활성화 추진 • 대학 연구소 운영지원 보조금/간접비 지원을 증액, 연구소 재정기반 마련
외부기관과의 협력/운영 계획	• 국내외의 다양한 기업들과 공동연구를 진행할 계획에 있으며, 해외 유관기관과의 교류 및 긴밀한 국제공동 블록체인 유관 협동 연구를 진행할 예정 • 지능형 블록체인에 관한 기술은 세계적으로 수요가 예상되는 기술로 미국 및 유럽 선진국의 대학 및 기업들과 국제협력 프로그램 개발 예정
학술대회 및 세미나 개최	• 매년 신기술 관련 워크숍 국내 및 국제 학술대회 지속적인 개최 • 저명 블록체인 관련 국제학술대회에서 블록체인 신기술에 관한 워크숍 개최

⑤ 연구개발팀 편성도



○ 세부과제 1

세부과제책임 연구실	참 여 연 구 자	담당기술내용
광주과학기술원 이흥노 교수	이흥노 교수 외 11명	▪ 신원 인증 키 비가역적 아날로그-디지털 스캐너 ▪ 영지식 증명 검증 프로토콜 ▪ 영지식 아날로그 생체인증 시스템
참여 연구실	참 여 연 구 자	담당 기술개발 내용
서울대학교 문수묵 교수	문수묵 교수 외 5명	▪ 블록체인 스토리지 및 데이터 검증 경량화 기술 개발 ▪ 비잔틴 환경에서의 연합학습 방법론 개발 ▪ IoT 네트워크를 위한 블록체인 기반 디지털 트윈 개발
참여 연구실	참 여 연 구 자	담당 기술개발 내용
광주과학기술원 신종원 교수	신종원 교수 외 10 명	▪ 딥러닝 기반 화자인증 모델 개발 ▪ 화자인증 전처리를 위한 다채널 잡음 제거기술

○ **세부과제 2**

세부과제책임 연구실	참 여 연 구 자	담당기술내용
광주과학기술원 김종원 교수	김종원 교수 외 10명	▪ Connected DataLake 인프라를 위한 멀티사이트 Data Pond 클러스터 설계 및 구축 ▪ 오픈소스 기반의 Connected DataLake 프레임워크 요소 기능 및 통합 소프트웨어 개발
참여 연구실	참 여 연 구 자	담당 기술개발 내용
고려대학교 김황남 교수	김황남 교수 외 9명	▪ 블록체인 데이터 스트림 전송을 위한 신뢰 기반 네트워크 기술 개발 ▪ 블록체인 플랫폼 구축을 위한 오버레이 네트워크 지원 기술 개발

○ 세부과제 3

세부과제책임 연구실	참 여 연 구 자	담당기술내용
아주대학교 김기형 교수	김기형 교수 외 5명	▪ DID기반 블록체인 기반 인증/인가 기술 ▪ OAuth2.0을 활용한 DID 블록체인 기반의 액세스 토큰 활용 기술
참여 연구실	참 여 연 구 자	담당 기술개발 내용
광주과학기술원 황의석 교수	황의석 교수 외 7명	▪ 물리계층을 활용한 DID 생성기술 개발 ▪ DID 기반 블록체인 연계 보안인증 기술 개발 ▪ DID 관리 및 접근 권한 제어 프로토콜 연구

3) 연구책임자 및 참여인력의 역량

① 연구센터장(총괄책임자)의 전문성 및 역량

○ 전문성

총괄책임자	이홍노
연구역량	•NRF-도약연구 과제 책임자 (2010-2024), ADD 전자전 연구센터 사업 (EW-11, EW-33) 세부 과제 책임자 (정보이론, 통신 및 레이더 관련 분야) •레이다, 분광기, 뇌컴퓨터 인터페이스 시스템에 특화된 지능형 신호 복원 및 분류 시스템 연구(2018-2021), •확장가능한 탈중앙화 보안성 ECCPoW 블록체인 과제 책임자(2019) •NRF-핵심연구 과제 책임자, MT-IT 연구센터 세부과제 책임자, ETRI 위탁과제, 미국에서 NSF, DARPA, Boing, Raytheon, Hughes 여러 프로젝트에서, 연구책임자 (무선통신 및 네트워킹 관련 연구진행)
경력 및 수상	**경력** • GIST 인공지능연구소 소장(2021~현재) • 우정사업정보센터 블록체인 자문위원(2018~2019) • GIST 블록체인인터넷경제 연구센터장 (2018~현재) • (사)대한전자공학회 통신소사이어티 회장 (2017~2019) • GIST 연구원장/ 전략연구사업단장 / 연구정책센터장(2016~2017) • GIST 센서지능화 연구센터장(2015~현재) • 미래창조과학부 연구개발특구위원회 위원(2015~2017) • IEEE Technical Program Committees for ICC, Globecom (WC, AHSN분야, 위원회 멤버로 활동 중) • 광주과학기술원 교수(2013~현재) • University of Pittsburgh, ECE Dept., 조교수(2002~2008) **수상** • GIST **연구상(산학협력)** (2020. 11월) • 대한전자공학회 **해동학술상** (2019. 12월) • 한구연구재단 **이달의 과학기술자상** (2014. 1월) • 한국연구재단 **기초연구 우수성과 50선 선정** (2013. 10월) • 미래 창조 과학부 **국가연구 개발 우수성과 100선** 선정(2012. 8월)  <이달의 과학기술자상>  <국가 R&D 우수성과100선>  <기초연구 우수성과 50선>
연구실적 (2008-현재)	• 총 280여편의 논문 (국제논문 176편 (SCI 논문 80여편) 국내논문 104편 (저널 14편) • 국외특허 31건 (등록 14건)과 국내특허 30건 (등록 23건) 및 8권의 Book Chapter를 출판 확정.

- 현재 박사과정 학생 5명, 석사과정 1명의 인력을 양성하고 있으며, 9명의 박사졸업생, 15명의 석사졸업생을 배출하였고, 포스트 닥터 7명과 연구교수 2명을 지도하였음.

- 2013.1. ~2017.2. : IEEE Gwangju Section Chair로 광주 전남 지역 IEEE Membership development 활동, 세계적 석학 Distinguished Lecturer 초빙, 국내외 전문가들을 초청 세미나 개최 등 여러 활동을 계획하고 BoG 위원회를 관리하였음.

- 2009.1. ~ 현재 : GIST 교수 임용 이후, 도약 연구사업 수주(2010.~현재), 핵심 연구 사업 수주(2010.) 그 외에 공동연구로 ADD(2013.~현재), ETRI(2012.) 등 여러 국책 과제를 수주하였음.

- 2002.1. ~ 2008.12. : Pittsburgh 대학 교수 임용 이후 National Science Foundation, U.S. Army, Pennsylvania 주 정부 등으로부터 과제를 받아 연구를 수행함.

- 1999.1. ~ 2002.1. : 세계 최초로 루비 레이져를 개발한 IT 분야의 최첨단 연구소로 널리

알려진 휴즈 리서치 랩 (HRL Lab)에 연구원 (Member of Technical Staff)으로 활동함. Internet traffic modeling, cross-layer networking, wireless ad hoc networks등 분야 연구.

○ **추진전략**

- 분기별로 논문, 특허, 학술대회 발표, 조직기여도 등에 따라서 실적을 분석하고 A(매우우수), B(우수), C(보통), D(미흡), E(매우미흡) 등으로 구분하여 인센티브를 차등 지급. 이를 통해 우수 연구 인력에게 지속적으로 연구 성과를 이끌어 낼 수 있는 연구 환경 제공.
- 해외 출장의 경우 우수한 학술대회를 엄선하여 연구원을 보내고 학술결과를 발표하게 함.

○ **컨소시엄 및 참여기관과의 협력방안**

- 해외연구실과 연구 결과를 공유하고 문제점을 논의하는 등 수평적 관계의 인적네트워크 형성.
- 연구에서 도출된 연구결과를 지적재산권화 하기 위하여 광주과학기술원내에 설치된 과학기술응용 연구소 (GTI)와의 협력 추진.
- 국가 경쟁력을 선도할 수 있는 전파 전문 인력을 양성하기 위해 컨소시엄 및 참여기관과 교육프로그램을 기획하고, 이를 추진하고자 함.
- 집단 지능과 같은 인문·사회적 환경을 고려한 전파 네트워크 시스템을 협력 개발하고자 함.

② 참여교수의의 전문성 및 역량

참여교수	문수묵 (1세부 참여교수)
연구역량	• 임베디드 시스템 소프트웨어, 웹 플랫폼 그리고 블록체인 플랫폼 분야의 국내 최고 수준의 전문성 보유
연구개발 실적	• LG전자, 삼성전자, IBM, Sun, Intel, HP 등의 기업들과 가상머신 및 웹 플랫폼 최적화 연구 진행 • 삼성전자 ‘IoT향 Blockchain Platform 설계 및 개발’을 산학협력 진행 • 뉴럴비씨 ‘고성능 탈중앙화 거래소(DEX)를 위한 2-way relayer 개발’을 주제로 산학협력 진행 • 한국연구재단 ‘탈중앙화 방식의 딥러닝 플래폼’을 주제로 중견연구 과제 수행

참여교수	신종원 (1세부 참여교수)
연구역량	• 음성/음향 신호처리를 중심으로 다양한 신호처리 및 머신러닝 기법의 국내 최고 수준의 전문성 보유
연구개발 실적	• 단일 마이크로폰을 이용한 음성의 통계적 모델링, 음성 검출, 음성 향상, 음성 강화 등 음성통신 시스템의 전/후처리 기술 연구 수행 • 음원 정위, 음성 검출, 잡음 억제, 음성 강화, 에코 제거 등 음성 통신과 음성 인식 전처리 핵심 기술에 대한 연구 수행 • 음성 감정 인식, 화자인증 등 음성/음향 애플리케이션 연구 수행 • 최근 5년간 SCI(E)급 국제 학술지 논문 11편 발표

참여교수	김종원 (2세부 책임교수)
연구역량	• 클라우드 기반 계산/저장 인프라 운영 및 AI기반 HPC/HPDA 서비스 실증을 포함한 Networked Intelligence 분야 국내 최고 수준의 전문성 보유
연구개발 실적	• 데이터 중심 지능형 서비스를 유연하고 자원 효율적으로 운용할 수 있게 설계된 SmartX Intelligence Cluster 개발 • Cloud-native 기반의 다양한 서비스에 고속의 대용량 스토리지 자원의 유용한 제공을 위해 Connected DataLake를 목표로 클라우드 스토리지 클러스터 연구개발 수행 • 데이터의 효율적인 관리를 위한 데이터 파이프라인 기술과 데이터 전송 기술 연구개발 수행 • 국제 SCI급 학술지 96 편, 국내특허 등록 50건, 국제특허 등록 19건

참여교수	김황남 (2세부 참여교수)
연구역량	• Cellular 및 WLAN 무선 네트워크 기술, 군집형 무인이동체 운용과 네트워킹 및 안티드론 기술 분야 국내 최고 수준의 전문성 보유
연구개발 실적	• 애드혹 커뮤니티를 위한 서비스 프레임워크 및 네트워킹 인프라에 관한 연구 수행 • 지능형 전력망에서 전력 공급 차등화 및 보안 프레임워크에 관한 연구 수행 • 다중 무인체 기반 사이버-물리 융합공격 대응 기술 및 자율복원 통신/보안기술 개발 등 사물인터넷(IoT), 네트워크 인프라 구성, 블록체인을 주제로 다양한 연구를 진행해옴 • 국제 SCI급 학술지 67 편, 국내특허 등록 34건, 국제특허 등록 9건

참여교수	김기형 (3세부 책임교수)
연구역량	• 블록체인 기반 빅데이터 수집과 분석 인프라 구축 • IoT(LoRa) 단말과 데이터의 인증 무결성 확보를 위한 블록체인 기반 보안 프레임워크 연구 • 지능형 Internet of Energy(IoE) Data 블록체인 기술 전문성 보유
연구개발 실적	• 블록체인을 활용한 인터넷 사기 정보 수집 및 공유방안 연구 • LPWA 기반 Farm IoT 서비스 지원 시스템 개발 • 기계학습을 통한 LPWA 위치인식 정확도 향상기법 개발 • LoRa 네트워크 노드캡쳐 공격에 대한 방어기법 개발 • IoT를 위한 효율적인 MAC 스케줄링 알고리즘 개발 • 클라우드팜 서비스 지원을 위한 자가망 기반의 LPWA 네트워크 구축 기술 개발

참여교수	황의석 (3세부 참여교수)
연구역량	• 에너지, 인공지능, 기계, 전기컴퓨터 공학 등 다양한 학제 간 연구를 통해 데이터 채널과 관련된 핵심 SoC 코딩과 신호처리 관련 국내 최고 수준의 전문성 보유
연구개발 실적	• Cellular IoT 환경을 위한 물리계층 장치인식 기반 간결한 보안통신 연구과제 수행 • AI 기반의 사전예방 및 빠른 복구가 가능한 에너지 재난 대응 플랫폼 개발과제를 수행 • 데이터 채널 및 에너지 응용 지능형 신호처리 기술 등을 연구하여 국제 저명 학술 저널 (IEEE Trans. Smart Grid, IEEE Trans. Veh. Tech, IEEE Trans. Magn. 등)과 학술 대회에서 다수의 연구논문을 발표함 • 최근 5년간 SCI 논문 18건, Scopus 학회 논문 15건, 미국 등록 특허 11건, 국내 학회 우수발표논문상 3건, 우수포스터상 1건 수상

4) 수행기관의 역량

① 광주과학기술원 기관 역량

○ 광주과학기술원은 도전, 소통, 융합하는 4차 혁신인재 양성에 주력하며 국가 인공지능 연구를 통한 4차 산업혁명 시대에 광주, 전남권 혁신성장을 선도하는 것을 목표로 하고 있음. 좀 더 세부적으로 과학기술 중심사회에 Global Leader를 양성하고 국민, 지역사회와 함께하는 포용적 연구개발을 수행하기 위해서 3가지 전략 목표를 세웠음. 첫 번째로 우수 과학기술인재를 양성하고 인재양성을 위한 인프라 구축. 그리고 연구역량 강화로 과학기술 발전을 선도. 마지막으로 미래선도형 연구 및 성과확산으로 사회문제를 해결하고 성장동력을 발굴.

○ 그 일환으로 블록체인 인터넷 경제 센터, 소프트웨어 교육센터, 인공지능 연구소 및 인공지능 대학원이 출범. 이는 대학정보통신연구센터(ITRC)의 기본 목표인 ICT분야의 고급 인력 양성은 물론 국가적으로 중요한 ICT 원천 기술을 연구 개발하고, 그 결과를 산업체로의 기술 이전하는 등 산학 공동 발전을 도모하기 위함임. 이러한 이유로 본 센터에 장기적으로 집중 지원이 필요.

② 인력

○ **(교수)** 1인당 논문피인용수 세계 5위(QS, 국내 1위) 수준

- 우수한 교수 보유 및 지속 충원 중(2015 대비 18.8% 증가)

- 우수 연구인력 보유 : 대학원 교원 1인당 SCI 주저논문 평균 2.5편, 평균 IF 4.38
- 산업현장 전문가 확보 : 산업체경력 교원 46명, 산학협력 중점교원 4명
- 글로벌 역량 증대 필요 : 외국인 교원 12명(전체 대비 6.3%)으로 글로벌 역량 강화 위한 외국인 교원 증원 필요
- 여성비율 증대 필요 : 여성교원 19명(전체대비 10.1%)

※ 교원 188명 보유

○ **(학생)** 미래 과학기술 리더 및 세계 수준의 고급 연구인력 보유
- 대학원 : 세계 수준의 연구역량을 갖춘 R&D/창업 인재 보유
- 대학 : Liberal Arts College 운영을 통한 창의융합인재 양성 중

※ 학사 943, 석사 382, 박사 576, 통합 344, 계 2,245명 보유

○ **(직원)** 직무별 전문인력 보유 및 정규직 기반 안정적 운영 중
- 최근 IT전공, 박사학위자 등 전문인력/신설부서 중심 인력 충원
- 2018년 83명의 정규직 전환을 통한 지원인력의 안정적 운영

※ 행정직 100명, 일반·관리직 121명 보유

○ **(연구원)** 우수 연구원 및 기관차원 연구전략, 대형연구과제 기획·발굴 위한 연구기획 분야 전담 연구원 보유

※ 연구원 29명 보유

⇒ GIST 구성원의 혁신역량을 강화하기 위한 글로벌 융합리더 양성(학생), 우수교원 양성, 직원역량 강화 필요성 대두

③ 예산

○ **(미래성장 재원 보유)** 정부 차원의 지속적인 구조조정에도 불구, 안정적인 출연금 및 자체수입 확보를 통한 기관 미래성장 재원 확보

※ 2019년 정부출연금 1034억원, 외부수탁연구비 1100억원

- 다만, 주요 건설사업 종료에 따른 시설비 축소로 향후 출연금 축소 불가피(최근 10년간 출연금 내 시설비 비중 평균 26.3%)

○ **(신규사업 역량 보유)** 미래 신규 성장동력 창출을 위한 신규사업 예산 확보 중(지속 확보 필요)
- 최근 5년간 융합기술연구, 기후기술연구, G.I.4.0, 미래재난안전기술개발 등 신규 연구사업 예산 확보

○ **(수탁사업 확보 필요)** 현재 자체수입 대부분(80%)을 인건비로 투자 중으로, 안정적 기관 운영 위해 수탁사업 확대 필요

※ 학사사업비, 인건비, 경상경비, 시설비 등에 자체재원 투입 필요

○ (인건비 지원확보 절실) PBS 적용기관이면서 교육·연구 동시 수행 필요성을 감안하여 정부의 인건비지원 확대 절실(현 60%→70%)

④ 성과 관리

○ **(기관평가 관리)** 2015~2018 임무중심형 기관평가 우수등급 취득

○ **(대학평가 관리)** QS 및 주요언론사 대학평가 최상위권 달성

- QS세계대학평가 교수1인당 논문피인용수 부문 4년 세계 5위(국내 1위)
- QS세계대학평가 개교 50년 미만 대학순위 국내 3위(세계 32위)
- 대학창업지수 2년 연속 국내 1위
- 이공계대학평가 3년 연속 기술이전수입 국내 1위

○ **(성과확산 관리)** 지속적인 산학협력 및 기술이전·창업을 통한 연구성과의 확산 체계 관리

※ 기술이전 역량 : 2018년 기술료 수입 14억여원 확보

※ 창업 역량 : 창업기업 총 68개사 1,122억 원 매출(누적), 교수 창업기업 애니젠(주) 코스닥 상장 (시총 800억 원, 2018.8.)

- 산학협력 촉진을 위한 혁신적인 교원 인사제도 운영(산학경험 교원임용, 산학협력 중점트랙 신설)
- GRI 산하 연구전략팀 신설로 미래 먹거리 발굴역량 확보 및 창업진흥센터 확대개편으로 기관 강점인 창업역량 강화 예정

⇒ 지역기반 융합기술 창업성과 창출 및 전략적 홍보 강화로 기관성과 전파를 통한 모범기관 GIST로서 광주·전남권 혁신성장을 선도

⑤ 조직

○ **(융합역량 강화)** 융합기술원 운영으로 에너지·문화기술 및 의료, 로봇 분야를 중심으로 하는 다학제 융합교육 및 연구 기반 조성

○ **(기초과학 강화)** 물리화학부 설립·운영을 통해 산업융합·응용분야 이외에 기초과학분야 역량을 강화

○ **(4차 산업혁명 대응)** SW교육센터, 인공지능연구소 등 4차산업혁명 대비 AI기반 융합연구·교육 플랫폼을 구축

○ **(연구리소스 관리·집중)** GIST연구원 설립·운영을 통한 연구 전주기에 걸친 자원배분과 성과환류 기반 구축 및 광·에너지·문화기술·생명노화·국제환경·AI 등 주요 연구소 중심 연구역량 결집

○ **(연구기반시설 구축)** 연구안전센터, 실험동물자원센터, 중앙연구기기센터 구축을 통해 연구기반시설의 대형화, 전문화 달성

○ **(성과확산 체계화)** 지식재산센터, 혁신기업가교육센터, 창업진흥센터, 기업경영전략팀 등 체계화된 연구성과 확산 조직 보유

○ **(소통·존중 창구 개설)** 상담·경력개발센터, 인권센터 등 대학 내 소통기능 강화 및 상호 존중 문화를 위한 전담 창구 구축 ⇒ 참여소통의 조직문화 조성 및 연구윤리와 청렴문화 확대를 통환 기관 조직역량 강화

⑥ 육성 계획

○ 본 센터는 무한국제경쟁시대에 능동적이며 역동적으로 대처할 수 있는 고급과학기술 인재를 양성할 것이며, 미래 인류문제 해결을 위한 창의적 블록체인 및 인공지능 핵심 원천기술 개발연구를 선도하고 연구의 국제화를 달성하고자함. 특히, 블록체인과 지능형 네트워크 보안 관리 분야의 핵심원천기술 및 지적재산권을 확보하고자 함.

○ 독창적인 핵심원천기술을 확보하고, 국제적 감각과 글로벌 경쟁력을 겸비한 우수 전문 연구개발 인력을 양성 및 배출함으로써 기술선진국으로서의 위상을 확립하고 더 나아가 국가산업발전 및 일자리 창출에 기여하고자 함.

○ 연구센터/산업기관과의 공동연구 및 지속적인 연구 성과 평가 및 관리체계를 확보하고자 함.

○ 산학공동프로젝트를 통해 이론 및 실무능력을 겸비한 핵심기술 인력양성과 창업 아이템 공모 및 지역특화산업과 연계한 졸업생 창업을 유도, 벤처기업을 육성함.

○ 지적/산업 재산권을 확보 후 본원에 설치된 GTI센터를 통해 보유기술의 산업체 이전 활성화 추진

⑦ 육성 전략

○ 교육연구실적 평가기준 기반의 평가 제도로 교육/연구 장려.

○ SW기반 연구개발 관리 체계를 수립하여 연구개발 관리의 선진화, 체계화를 이룸.

○ 블록체인 센서 디지털 지능 트윈 플랫폼 및 관리 기술, 네트워크 지능 기반 블록체인 서비스, 지능형 네트워크 보안 관리 기술 등을 개발하기 위해 미래 전략적 과학 기술 분야에서 전문성 확보가 가능한 수업 및 교육을 실시할 것이며, 이를 통해 새롭게 변화하는 21세기 블록체인 및 인공지능 분야의 고급인력을 양성하고자 함.

○ 특허 및 창업, 기업가 정신 등과 관련된 교과목을 통해 산업, 사회 현장에서 활용 가능한 실질적인 기술에 대한 교육을 진행하고, 중견/중소기업과 연계하여 협력공동연구를 진행하는 등 산업체에 인력수급이 가능한 전문 인력 양성.

○ 연구 분야에 대한 발표 및 토론이 가능한 수업을 개설하여 자율적이고 자유로운 학습 분위기를 유도하고, 학생들 간 아이디어 교류 및 협력을 통해 창의적 사고가 가능하도록 지도함. 또한 창의자율 과제를 운영하여 박사 또는 교수의 지도하에 자율적으로 창의 아이디어를 연구하고 개발할 수 있는 환경을 조성하고자 함.

○ 특허 및 창업, 기업가 정신 등과 관련된 교과목을 통해 산업, 사회 현장에서 활용 가능한 실질적인 기술에 대한 교육을 진행하고, 중견/중소기업과 연계하여 협력공동연구를 진행하는 등 산업체에 인력수급이 가능한 전문 인력 양성.

○ 연구기관, 산업체, 공공기관과의 주기적인 세미나를 통해 학문적 교류를 함으로써, 다양한 학문과 융합 연구가 가능한 교육 환경을 제공함.

4 인력양성 계획

1) 전주기적 인력양성 계획

인재상 도전정신과 창의력을 갖춘 포용적 ICT인재

목표
- ⊘ 도전정신과 창의력을 갖춘 포용적 블록체인·융합 ICT연구분야 인재 양성
- ⊘ 혁신적·도전적 연구 및 교육수행을 위한 센터 운영 시스템 구축
- ⊘ 산·학·연 협력체계를 통한 기술구현과 비즈니스(창업 및 스케일업) 창출 지원

추진전략

전략 1: 3대 세부연구* 기반 교육-연구-창업 선순환 시스템 구축
① 영지식 센싱 탈중앙 고성능 분산원장 기술 개발
② Edge 컴퓨팅 기반 블록체인 서비스 인프라 기술
③ DID 블록체인 기반 IoT 인증 및 보안 기술 개발

전략 2: "3CIP" 개념의 전주기 대응 집중적 융합인재 교육
Ⓒhallenge, 불확실성에 도전할 수 있는 인재양성을 위해 반증가능성(falsifiability)을 염두에 둔 실험적 교육·연구환경 조성
Ⓒooperation, 동료학습·참여대학 간 다학제적 융합 및 사업화 논의 환경을 조성하여 소통 및 협업을 통한 연구 발전·심화·응용 가능성 제고
Ⓟroblem Solving, 도전이 낳은 검증된 지식(tested knowledge) 축적과 교수진의 체계적 피드백을 통해 학생 스스로 답을 찾을 수 있는 능동적 학습이 가능한 양질의 교육 환경 구현
Ⓒreativity, 창의자율과제 지원, 산학연 멘토링, 실험실 기반의 아이디어 사업(창업)화 지원 등을 통해 우리 사회에 새로운 생산성을 제고하는 창의적 인재 육성

전략 3: 주관-참여기관 ICT인프라를 활용한 역동적 연구생태계 구축
o GIST-고려대-아주대-서울대 간 교육·연구 유동성 확대
o 참여기관 간 정기 자문·토론 협의체 구성을 통한 공동연구 활성화 체계 마련
o 국내외 산·학·연 교류협력을 통한 연구협력 및 취업연계 확대

그림 23. 블록체인 지능 융합센터의 인재상 및 인재 양성 전략

① 인재양성 전략 및 목표

가. 블록체인 지능 융합 연구센터의 인재상

○ COVID-19 사태 이후 디지털 대전환 시대를 빠르게 맞이함에 따라 혁신적으로 발전하는 기술과 다양한 변수의 상호작용으로 인한 예측불가능성과 불확실성이 증대

○ 본 센터에서는 향후 미래세대가 맞이할 불확실한 문제에 도전할 수 있는 도전정신과 실험을 통한 검증된 지식(Tested Knowledge) 기반의 창의력을 갖춘 포용적 ICT인재 양성을 목표로 함. 차세대 ICT연구분야의 핵심기술에 대한 이해력과 창의적 실행능력을 갖춘 인재를 양성하여 지속가능한 미래 사회를 위해 포용적 부가가치를 창출하는데 이바지할 예정

※ 블록체인 지능 융합 연구센터의 인재상 정립 기준

1) 유수 국제기구(2020년 세계경제포럼(WEF) 및 2018년 OECD)**가 제시하는 미래 과학기술인재의 핵심역량**

▲ 유연성·변화대응 역량 : 전공기초·디지털 역량, 자기주도 학습역량 등

▲ 문제발견·해결 역량 : 창의·비판적 사고, 융합역량, 디자인씽킹 등

▲ 기업가적사고 역량 : 공감·소통·협업, 글로벌·개방적 사고, 실행·추진력 등

▲ 사회적책임·인간중심사고 : 윤리의식, 인문학적 소양, 포용적 사고 등

2) 4차 산업혁명시대를 대표하는 기업 CEO(마크 주커버그/페이스북, 비탈릭 부테린/이더리움, 스티브 잡스/애플, 데미스 하사비스/구글딥마인드)**가 보여주는 인재의 공통 역량**

▲ (도전정신) 전문적 지식과 경험을 갖고 새롭고 불확실한 상황에서 끝까지 밀어붙이는 에너지

▲ (창조적 사고) 새로운 지식을 창출하고 다양한 지식을 융합하여 새로운 가치를 창출

▲ (협업능력) 다른 사고방식을 가진 사람을 이해하고 활용해 시너지를 극대화하고 가치를 창출

나. 인재양성 전략 및 목표

○ **(전략의 필요성)** 미래사회 문제를 해결하고 차세대 신기술을 선도하는 인재양성 기관이 되기 위해 100년 전 탄생한 '대학교' 개념을 탈피한 교육과 연구 환경을 제시

- 새로운 혁신리더형 인재양성 시스템을 추진하기 위해 세계적인 혁신 대학으로 주목받는 미국의 미네르바 대학(Minerva School)이나 구글·나사가 투자한 혁신 대학인 싱귤래리티 대학(Singularity University)과 같은 수준별 맞춤교육(Cross Contextual Scaffolding), 완전한 능동적 학습(Fully Active Learning), 체계적인 피드백(Systematic Formative Feedback)으로 시스템을 구축하여 학생 스스로 답을 찾을 수 있는 교육을 구현

구분	사례 소개
(미국) 미네르바 대학	캠퍼스와 강의실이 없는 4년제 대학, 온라인 토론과 세미나 기반 수업 운영
(미국) 싱귤래리티 대학	실리콘밸리 위치 창업대학. 매년 10주 과정 프로그램을 진행하여 창업기업 배출
(미국) 애리조나 주립대	온라인 강의(MOOC) 중심 학사 학위 과정, AI기반 맞춤형 학습 프로그램 도입
(핀란드) 알토 대학	기존 공대·디자인대·경제대를 통합, 산업체와 협력하여 실무중심 교육 제공
(프랑스) 에꼴42	교수·교재·학비(3無)가 없는 IT 교육기관, 팀프로젝트를 기반으로 결과물 도출

○ **(전략 및 목표)** 이에 본 센터는 아래의 단계를 거친 인재를 양성하고자 함

단계	내용
1	**(질문을 던지는 환경 마련)** 당면한 현실과 미래의 불확실성에 도전할 수 있는 반증가능성(falsifiability)에 대해 질문 제시
2	**(검증된 지식의 구축)** 실험중심의 연구를 통해 질문의 답을 찾고 그 과정에서 검증된 지식(Tested Knowledge) 구축
3	**(창의력 배양)** 구축된 지식을 바탕으로 또다른 질문을 던지는 과정에서 새로운 부가가치를 창출하는 창의력 배양

- **(단계 1)** 현실과 불확실성에 대한 질문을 던지기 위해서는 사회에 대한 통찰력과 문제의식이 선행되어야 함. 본 센터는 교육의 첫 단계를 질문을 던질 수 있는 인재로 성장하기 위한 기초 지식과 사회적 동향 파악에 둠

- **(단계 2)** 산학연 협력체계 구축을 통해 본 센터가 지향하는 인재의 3CIP 소양(Challenge, Cooperation, Problem Solving, Creativity)을 갖출 수 있는 실험지원, 세부연구과제 협동수행 및 융합교육, 실험을 통해 학생들의 검증된 지식 확장을 도모

- **(단계 3)** 참여교수진 및 자문위원으로 구성된 정기세미나 등을 통해 각 학생들의 아이디어를 발표, 이를 기술수요 및 산업동향과 매칭해 창의적 연구결과·사업(창업)화로 도출될 수 있도록 지원

다. 석·박사 배출목표

성과지표		1년차	2년차	3년차	4년차	5년차	6년차	7년차	8년차
수혜인원(명)	석사	24	24	24	24	24	24	24	24
	박사	26	26	26	26	26	26	26	26
	합계	50	50	50	50	50	50	50	50
배출인원(명)	석사	6	12	12	12	12	12	12	12
	박사	2	4	4	4	4	4	4	4
	합계	8	16	16	16	16	16	16	16
취업률(%)		100%	100%	100%	100%	100%	100%	100%	100%
창업기업 수	학생	0	1	0	0	0	2	0	1
	교수	0	1	0	0	0	2	0	1
	합계	0	2	0	0	0	4	0	2
기업가정신교육인원		4	10	10	10	10	10	10	10
기타(추가가능)									

라. 인재양성 전주기 모델

a) 인재양성 전주기 모델



그림 24. 우수 인재 양성을 위한 전주기 모델

○ 대학 내 블록체인 인프라 및 산학협력 확대를 통해 연구센터 참여학생의 실무능력 향상

- 블록체인 분야 다양한 필수·전문 융합교과목을 개설하여 블록체인 전략산업에 적합한 체계적인 커리큘럼 구축

- 이를 토대로 매년 20명 이상의 석사과정 신입생을 유치함으로써, 비보안전공계열의 학생 참여 및 컴퓨터·보안정보계열 학생들의 융합과정 이수를 통한 창의적 블록체인 기술분야 인재 양성

○ 전주기적 인재양성 추진전략

- 블록체인 특화형 교육을 통한 이론과 실무적응형 인재육성으로 참여기업 및 산업체 요구에 부응하는 인력을 양성해 취업난 해소

- 산업 특성에 맞는 블록체인 분야 전문지식을 바탕으로 주력 산업기술 및 기업수요에 걸맞는 기술성과물 창출 및 기술이전

- 블록체인 산업분야 적용을 위한 산업의 이해와 융합보안 실무능력을 접목시킨 개방형 블록체인 전문인력 양성

- 참여기업 및 산업체의 수요를 반영한 산학협력 R&D수행으로 블록체인 분야 기술능력 고도화를 통한 전문 인력 양성

b) 학생선발 및 교육방안

○ 블록체인-인공지능 융합대학원 설립을 통한 우수인재 양성

- 블록체인과 인공지능 등 4차 산업혁명 관련 대학원을 신설(2023년)하고 세계 최고 수준의 인재를 양성함으로써 블록체인 핵심 산업을 뒷받침할 인력양성 기반을 마련

전공	구분	내용
블록체인-인공지능 융합대학원	대학원명	블록체인 대학원
	목적	블록체인 핵심산업 뒷받침을 위한 핵심인력 양성
	설립 추진	2021. 7
	운영 시기	2023. 3
	운영형태	고려대, 서울대, 아주대 등과 공동강의 등 협력하여 운영

○ 연차별 모집계획

- 본 센터 연구의 참여학생 혜택, R&D형 인력양성 사업 등 본 센터의 이점 홍보를 통해 우수입학자 조기 선발 및 확보

- 우수학생 선발 시 우대 기준

  - 블록체인 분야 기술 관련 기업가 정신을 가진 자
  - 관련 산업체 근무 및 인턴 경력자
  - 지원분야의 연구 수행능력을 평가하기 위해 출신 학과의 전공에 관련된 기초지식 평가
  - 박사과정 지원자는 본인의 연구 분야에 대한 프레젠테이션

c) 취업・창업・진학 지원 방안

○ 블록체인 전문 스타트업 보육센터 운영 (2023년부터)

- 물리적 환경 제공: 서울블록체인지원센터, GIST서울사무소 등

- 네트워크 활용: 본 연구센터의 산학연관 네트워크를 활용하여 블록체인 기술을 개발하고, 다양한 산업체와의 협업을 통해 실무형 인재를 국내 산업계에 공급

○ 산업체 등 초청강연과 참여교수・학생・산학연관 전문가 매칭을 통한 교육-프로젝트 성과-취업 연계 시스템 마련

- 참여 기업 전문가 초청 정기 세미나를 통한 기술 및 산업동향 파악

- 참여 학생들의 취업의식 고취 및 진로상담 병행

- 학생별 연구아이디어・성과에 대한 전문가 피드백 및 취업 연계 독려

d) 사후관리 방안

○ 배출인력 재교육

- 기술의 급속한 발전으로 퇴화되기 쉬운 산업체 현장기술인력에 대한 신기술 재교육(이론 및 실습 등)을 통해 기업의 생산성 제고

- 생산현장의 당면한 기술인력 부족현상을 집중적 교과과정 개설을 통해 해소

○ 미취업자 지원

- 미취업 석·박사인력의 취업촉진을 위해 중소기업, 연구법인 및 기업부설연구소와의 연계를 통한 고용 추진

○ 만족도 조사

- 배출인력들에 대한 현장(기업, 연구소 등)의 만족도 조사

○ 피드백 반영

- 만족도 조사를 바탕으로 현장의견을 수렴하여 현재 교육과정에 반영

2) 창의·혁신 인재양성 방안

① 학생주도 창의·자율과제 운영계획

○ 학생 스스로 질문과 답을 도출할 수 있는 창의적 교육환경 마련

- 인류가 속한 모든 사회에서 발생하는 현상과 문제는 단일 분야에 국한하여 해석과 해결이 불가능함을 인지, 융복합적 문제 해결을 위한 독특한 창의자율과제를 수행할 수 있도록 지원

  • 지원예산: 팀별 1,000만원 범위 (최소 5개 팀 이상 지원)

- **(팀별 멘토링 제공)** 참여대학 및 협력기관 전문가 멘토로 구성된 멘토링 시스템을 통해 학생들의 연구결과를 정기적으로 공유, 이를 발전시키기 위한 상시적 자문 제공

  • 창의·자율과제가 산업현장과 사회에서 발생하는 문제해결에 대한 부가가치를 창출할 수 있도록 센터 내 전문가 네트워크를 활용, 블록체인·법·융합기술 등 다양한 전문가 그룹을 매칭한 멘토링 제공

  • COVID-19와 같은 글로벌 문제를 포함한 사회 현상에 대해 기술적 해결방안과 관련 산업별 고려사항 및 기반구축 등을 함께 논의하여 솔루션 마련→법적 기반 점검·필요사항 확인→비즈니스 모델 구축→사업화→창업의 과정을 거쳐 성공과 실패를 여러 번 경험하도록 함

  • 여러 번의 주도적 프로젝트 경험을 통해 학생들은 문제해결을 고민하는 과정에서 습득된 기술의 응용력을 증대시킬 수 있으며 실패를 두려워하지 않고 착안을 현실화하고자 노력하는 실제적 의미의 창의력을 배양

② 기업가정신교육 운영계획

○ 센터 참여대학 내 자체적으로 편성·운영되고 있는 기업가정신교육 과정을 활용

○ 광주과학기술원: 교내 기업가정신 아카데미 프로그램 활용

- 2014년 4월 과학기술응용연구단(GTI)과 창업진흥센터의 조직을 개편하여 기업가정신

교육센터(GEEC, GIST Entrepreneurship Education Center)를 설립하여 운영중

- 본 센터의 “New Venture-Tech Venture Academy” 기본과정 이수 활성화

- 주요내용: 기업가정신과 비즈니스 마인드, 기업과 리더, 기술사업화와 비즈니스 마인드, 글로벌 비즈니스 트렌드 , 개인/조직/제품의 혁신 등

○ 고려대학교: “기업가정신 함양” 전공선택강좌 운영 활성화

- 기술경영전문(MOT)대학원과 공과대학 전기전자공학과에 개설된 기업가정신 교육 강좌를 통해서 수강이 가능함

- 현재 “공학인을 위한 기업가 정신”과목 등 교내 기업가정신교육 관련 강좌는 3학점의 전공 선택으로 총 5과목이 개설되어 성공적으로 운영 중에 있음(리더십의 이해, 신사업 발굴 프로세스, 창의경영, R&D 프로젝트 기획 교육 등의 교육과정).

- 주요내용: 리더십의 이해, 신사업 발굴 프로세스, 창의경영, R&D 프로젝트 기획

○ 서울대학교: 교내 벤처경영기업가센터 활용

- 2013년 12월 국내 최초로 학부 과정 벤처경영학 연합 전공을 신설한 뒤 2014년 9월 벤처 경영 기업가 센터를 설립하여 운영중

- 본 센터는 대학 내 창업교육을 통한 기업가 정신 확산과 창업 활성화 및 세계 수준의 창업 교육시스템을 구축하여 우수 연구논문 창출을 통한 창업 선도대학으로 성장을 목표

- “벤처 창업론”, “사회적 기업의 창업”, “기술 트렌드와 사업기회 분석” 등의 교과를 제공하는 연합 전공 벤처 경영학 수업 수강이 가능

- 주요내용: 기업가적 사고 방식, 일반 창업, 사회적 기업 창업, 기술 창업, 벤처 생태계 동향 분석, 전문가 특강 및 법률/회계 자문

○ 아주대학교: 교내 기업가정신과 경영 교과목 이수

- 아주대학교는 경영학과에서 매학기 ‘기업가 정신과 경영’ 과목을 운영 중이며 해당 교과목을 본 센터 내 학생들이 이수하도록 하여 기업가 정신을 기본소양으로 습득할 수 있도록 함

- 단순히 정해진 정답을 맞추기 위한 주입식 교육의 한계점을 보완하기 위하여 스스로 생각하며, 학생간 토론을 진행하여 학생 중심의 자기주도 학습을 수행

- 주요내용: 경영학의 발달과 접근체계, 기업의 본질과 목표, 기업가정신과 혁신, 경영환경 변화, 기업의 사회적 책임 등

③ 배출인력의 고용연계 및 창업지원 방안

구분	특성화 방향 및 독창성
연구센터차별화	• 산학연관 네트워크를 활용하여 창의적 프로그램 개발 및 운영 • 산학협력 프로젝트 과정을 신설하는 등 산학연계 과정을 제도화함으로써 학생 간 창의문화를 조성하고 취·창업으로 연계되는 선순환 구조를 구축 → 산학협력 R&D의 사후관리 강화를 통하여 우수기업과의 연계성 극대화 • 우수연구 성과자의 센터 내 시상을 통하여 학계 및 산업계에서의 지위 격상 • 우수기업 취업지원협의회 활성화를 통한 민관산학 협력 구조 정착 • 상설업무로써 체계적으로 자리매김할 수 있도록 지원 및 자립화의 기반 마련
기업밀착화	• 블록체인 분야 관련 일자리박람회, 산학협력협의회 등 우수기업과의 접점을 지속적으로 확대하여 유관 기업의 애로사항 청취 및 해결 • 산업체 수요조사를 강화하고 기업 맞춤형 트랙의 운영을 활성화하며, 기업위주의 위원회 구성 등 기업과의 접점을 높임
미스매칭해소	• 블록체인 우수기업의 취업요청에 대하여 해당 기업만을 위한 열린 채용회 개최 • 참여대학들과 “블록체인 우수기업 민간공채”를 공동으로 개최하여 우수 블록체인 유관기업과의 접점 확대 및 졸업 후 우수기업 취업 유도 ※ 석사과정 2학기부터 지역 우수 블록체인 유관기업과의 만남의 기회 확대 시행
정부/지자체 정책연계	• 정부 및 지자체 정책과 연계한 핵심기술 연구개발사업 추진 → 정부정책 및 지역산업과의 밀착도에 따라 특성화 인재양성 프로그램 운영 및 우수 연구성과 관리 / 최우수 실적의 경우 추가 인센티브 부여 • 정부정책 및 지자체 정책에 따라 블록체인 신산업 홍보 추진
지역사회 관심유도	• 다양한 블록체인 우수 핵심인력양성 사업을 지자체로부터 수주하여 참여연구원이 최상의 실적을 낼 수 있도록 독려 • 언론보도 등을 통한 전략산업 및 블록체인 분야 인식 계기 제공 • 유관기관과 합동 세미나, 기술이전 등 다양한 행사를 통하여 사업홍보

5 연구개발과제의 활용방안 및 기대효과

1) 연구개발성과의 활용방안

① 활용분야

○ 음성, 초분광, IoT 센서 데이터를 활용한 DID 기반기술과 영지식 증명을 통한 인증기술 개발은 데이터 유출문제를 해결하는데 기여할 것으로 생각됨. 이는 높은 보안성이 필요한 생체인증, 의료 데이터 보호 등 마이데이터 산업의 다양한 애플리케이션에 필수적인 요소 기술로 활용될 수 있음

○ 초경량·초확장 블록체인 원천기술과 대용량 데이터 관리/공유 기술, 신뢰 정의 네트워크 구축 기술 등을 활용하여 다양하고 혁신적인 융합서비스를 창출할 수 있는 기반 환경을 조성할 수 있음

○ DID 기반기술, 블록체인 기술, 기계학습 기술들은 본인인증·자격인증의 핵심기술로서 다양한 산업 분야에서 인가 및 접근제어의 핵심기술로 사용될 것으로 예상됨

○ IoT 기기가 적용되는 다양한 콘텐츠의 요구사항에 맞추어 DID 기술과 프로토콜을 설계하여 저지연 서비스를 제공할 수 있음. 대규모 IoT 기기가 연결되는 환경에 적합한 보안기술을 적용할 수 있음

② 활용방안

○ 본 과제의 최종 목적은 지능적 자원관리와 효율성과 탄력성이 제고된 영지식 센싱, 암호인증, 클라우드 서비스 융합기술을 개발하는 것이 목표임. 완성된 시스템을 구체적 애플리케이션에 활용·적용하기 위한 후속연구를 추진하고, 이를 바탕으로 기술창업 과제 수주, 석·박사급 고급인력 양성, 데이터 산업 인프라 구축에 활용할 예정임

2) 연구개발성과의 기대효과

① 기술적 측면

○ **(DID 기반 기술)** 음성 정보, 초분광 정보를 활용한 DID 기반 기술과 영지식 증명을 통한 인증 기술을 통해 생체정보 데이터의 유출을 원천 차단함. 생체정보 데이터의 유출에 의한 2차 피해 가능성을 제거하여, 생체인증 시스템의 보안성이 향상됨

○ **(IoT 탈중앙화 블록체인)** 국내 초경량·초확장 블록체인 원천기술과 더불어 대용량 데이터 관리 기술의 핵심 원천기술을 확보할 수 있음. 블록체인 인프라 기반 분산형 비즈니스 애플리케이션에 활용 가능하며 보안인증, DID, 금융 등 다양한 서비스 모델에 대한 경쟁력 제고 가능

○ **(데이터 저장/공유 기술)** 블록체인 대응 데이터에 대한 체계적인 저장/공유 기술개발을 통해 신뢰가능한 데이터 공유 인프라 기반기술 확보. 확보된 인프라를 활용하여 체계적 딥러닝 학습이 가능해짐

○ **(신뢰정의 네트워크 기술)** 기계학습을 위해서는 수많은 데이터들이 지속적으로 공급되어야 함. 데이터의 진본성과 보안에 대한 우려로 연결이 제한적임. 신뢰 정의 네트워크

원천기술 확보를 통해 데이터 스트리밍 환경에서 안정적인 연결을 기반으로 데이터 수집과 축적할 수 있으며, 궁극적으로 안정적인 기계학습 인프라를 제공할 수 있음

○ **(IoT 보안 통신 기술)** 물리계층 기반 간결한 DID를 통해 차세대 IoT 보안통신 기술 선점. 높은 보안성을 갖는 저전력, 저지연 보안통신 기술은 스마트그리드, 스마트팩토리등 다양한 분야에 활용될 수 있음

② 경제적, 산업적 측면

○ **(DID 시장 개척)** 세계 DID 시장의 연간 성장률은 26%, 2025년 30조원 규모로 예측되며, 2025년 국내 전자문서 시장은 약 8조원으로 예측되고 있는 바, 본 연구개발을 통해 DID와 영지식 증명 기술을 통해 확장 서비스를 구현하여 연간 수십~백억 원 규모의 시장을 창출할 것으로 기대됨

○ **(데이터 경제 활성화)** 실물 세상에서 발생하는 대규모 데이터를 블록체인에 분산 저장/관리하여 가용성, 무결성, 소유권을 보장하는 블록체인 데이터 저장 원천기술 확보로 누구나 쉽게 블록체인 데이터를 기반 융합서비스 제공 역량 향상, 산업간 융합 IT 서비스 개발 활성화로 4차산업혁명 견인

○ **(글로벌 수준의 네트워크 경쟁력 확보)** 신뢰 정의 네트워크는 다양하고 혁신적인 융합 서비스를 창출할 수 있는 기반 환경을 조성함으로써, 차세대 글로벌 네트워크 전략에서 국가 경쟁력 확보의 발판을 마련하고 4차 산업혁명을 선도하기 위한 글로벌 수준의 네트워크 경쟁력 유지에 이바지할 수 있음

○ **(IoT 보안통신 시장 선점)** IoT 장치를 위한 보안통신 원천기술 확보를 통해 로열티 및 기술 수출로 새로운 시장을 창출 할 수 있을 것으로 기대함. 높은 보안성을 요구하는 산업(원자력, 의료, 전력 등)의 시설을 IoT 장치로 대체할 수 있게 되어 관련 사업의 효율성을 극대화할 것으로 기대됨

③ 사회적 측면

○ **(개인정보 보호)** 기존 ICT 기술은 비대면 디지털 시대에 개인정보 보호에 한계가 있음. 본 과제의 DID 기반 인증과 영지식 증명 기술은 사생활 민감정보, 의료정보, 생체인증 등에 대한 인증·증명 시스템 활용 가능

○ **(신뢰사회 구축)** DID 기술, 블록체인 합의 기술, 신뢰정의 네트워크 기술, IoT 보안통신 기술 등이 개발됨으로써 위변조가 불가능하고 원본의 무결성을 보증하여 투명하고 신뢰받는 사회를 구현할 수 있으며, 특히 정부·공공 서비스에 위와 같은 기술을 활용하여 강력한 신뢰사회를 구축할 수 있음

○ **(참여와 공유 확산)** 신뢰 정의 네트워크는 모든 사람·사물의 데이터가 교환·소통되는 사회시스템의 신경망 임무를 수행함으로써, 시공간을 초월한 자유로운 참여와 공유 확산 효과를 기대할 수 있음

○ **(산업구조 개선)** 암호화폐에 비해 블록체인 기술은 산업·경제화가 매우 부진한 상황임. 본 과제에서는 블록체인의 기술적 한계를 극복하여 국내 블록체인산업의 저변확대는 물론 신생 전문기업의 출현과 같은 산업구조의 변화를 기대함

○ **(전문인력 양성 일자리 창출)** 초고속・대용량의 검증된 데이터 공유 기술을 실증할 수 있는 Connected DataLake와 같은 환경을 활용하여 데이터, 블록체인 기술 전문인재 양성 효과를 기대함. 또한, 블록체인산업 분야의 성공적 모델 제시를 통해 신규 일자리 창출이 가능함

○ **(데이터 주권 확보)** DID 기술, 영지식 증명 기술 등은 개인정보를 보호하고 제공을 최소화하여 데이터를 보호하는데 중요한 기술임. 이러한 기술은 마이데이터 산업 활성화에 역할을 하며 데이터 자기주권 확보

3) 인재양성 파급효과

○ 국제 및 산학연 협력을 통해 블록체인・AI・클라우드 등 ICT 기술 분야의 글로벌 인적 네트워크를 구축하고, 암호학, 보안 등 여러 분야와의 융복합을 통해 전문지식을 갖춘 석・박사급 고급인력 양성

- 디지털 세계에서 다양한 학문 융복합이 이뤄질 것으로 예견되는 미래사회를 위한 성장 동력 확보
- 인공지능 학습환경 및 다양한 빅데이터 기반 시스템에서의 신뢰성 있는 네트워크를 다룰 수 있는 실질적 인재 배출
- 산업현장 진출 인력에 대한 체계적 사후관리 및 수련 중인 학생들과 지속적인 연계를 통해 관련분야 고급기술 인력 간의 인적 관계망 구축

○ 배출된 전문인력을 중심으로 지역 내 대기업 및 중소기업에 수요 맞춤형 인력을 공급하여 기업이 필요로 하는 기술개발의 공동 참여기회 확대 및 신시장 개척

- 블록체인 등 융복합 학문을 산업체에 공급함으로써, '기술개발-상품화-수익창출-전문인력 양성' 의 순환 고리를 연결하는 교육·연구개발 통합클러스터 구축
- 블록체인 IoT 기술을 확보한 연구원들이 스마트그리드, 스마트팩토리, 드론, 첨단 의료 등 다양한 플랫폼에 확장 적용하여 새로운 시장 개척 가능

○ 안전하고 신뢰성 있는 분산형 블록체인 기반 공공 서비스 제공을 통해, 목적별 공공분야 발전 및 국민 편익 증대

○ 분산형 네트워크 컨트롤러를 사용하는 신뢰 정의 네트워크(Trust Defined Network) 구성하고, 이에 따른 네트워크 기술 고도화 및 IoT 데이터 검증기술을 위한 핵심 전공인력 육성과 차세대 네트워크 융합산업의 초석 마련

6 연구개발성과의 사업화 전략 및 계획

1) 블록체인 지능 융합 센터 사업의 SWOT 분석

강점(Strength)	약점(Weakness)
- 주관연구기관 GIST블록체인경제센터는 확장가능한 탈중앙화 보안성 블록체인(ECCPow) 및 공공서비스 블록체인 시스템 등 '18년부터 여러 블록체인 연구과제를 수행해왔으며, 연구 결과물로 ECCPoW, BTCECC, ETHECC를 보유 - 블록체인 지능 융합센터 보유기술 (블록체인 분야) · (서울대) 가상머신 및 Edge 컴퓨팅 기반 블록체인 최적화 기술 · (아주대) 블록체인 기반 보안 프레임워크(BIoT) 기술	- 최근 블록체인 기술은 활성화되고 있으나 폐쇄형과 개방형이 상호운용 가능한 미래형 블록체인이 가능한 기술에 대한 연구 진행은 미비한 실정이며 연구사례가 많지 않음
기회(Opportunity)	**위협(Threat)**
- 정부는 2025년까지 국가 혁신 프로젝트인 '디지털 뉴딜 정책'을 추진, 블록체인 기반 기술로 비대면 경제사회를 확산하고 지능형 정부 및 보안 시범사업 등에서 핵심기술로 활용 - 또한 '20년 블록체인 기술 확산 전략에서 발표한 바 있듯 국민이 체감할 수 있는 7대 분야에 블록체인 전면 도입 및 분산신원증명(DID) 기술을 집중 육성할 계획	- 블록체인 기술 분야에서 선진국과의 기술격차는 美 기준 2.3년으로 나타남 - 거래처리 속도, 대규모 데이터 관리 등 블록체인 기술의 기존 한계를 극복할 원천기술 필요

2) 블록체인 지능 융합 센터 사업의 전략목표 및 시사점

① 블록체인 융합 연구수행 및 성과확산을 기반으로 미래 블록체인 기반 신산업 육성에 기여

○ 산학연 협력을 통한 시장과 수요자 중심의 기술개발 및 사업화 전개로 '기술발굴-비즈니스 모델 개발-기업 애로사항 해결-마케팅' 등 사업화 전주기적 지원과 지속적인 기술공급 체계 구축

② 4차 혁신인재 양성을 주도하는 인재육성 모델 정립

○ 주관기관인 GIST 및 고려대, 아주대, 서울대 등 참여대학들의 우수 연구인력·인프라를 적극 활용하여 과학 기술 중심사회에 융합역량을 갖춘 인재양성

③ 국민·지역사회와 함께하는 포용적 연구개발 수행 및 창업 촉진으로 미래 일자리 창출

○ 광주광역시의 미래 전략산업인 인공지능(AI) 및 3대 특화산업(헬스케어, 자동차, 에너지)과 관련된 핵심원천기술 개발

④ 다양한 특허출원 및 지재권 전략 수립

○ 선진국과의 기술 격차 해소를 위해 블록체인 핵심기술을 개발하고 국내·외 특허를 출원하여 국내 기업의 경쟁력을 높이고 기술자립을 지원

○ 국내 기업과 지속적인 기술교류 및 협력방안을 모색하고, 지식재산권화 및 사업화 통해 신산업 창출 및 기업경쟁력 제고

- 전략적 IP-R&D 연계를 위해 인력확보와 운영방안 마련 필요

3) 블록체인 지능 융합 센터의 기술이전(사업화) 전략

① 기술사업화 지원

○ **(기술사업화 전문인력 활용)** 우수연구 성과가 시장으로 이어질 수 있도록 기술사업화 전담인력 및 전문가를 활용하여 시장 중심형 연구개발 추진

○ **(특허 분석 및 자산실사)** 변리사의 특허 분석 및 특허 자산실사로 특허 활용률 증대

○ **(유망기술 사업화관리 플랫폼, G-BizBoard)** 사업화 유망기술에 대한 발굴에서 기술이전 및 사후관리까지의 전담제와 공동 모니터링을 통한 효율적 기술이전 사업화 추진

② GIST의 창업 프로그램 및 지원 제도 활용

○ **(GIST 창업지원 교육프로그램)** 실용성 강화한 창업교과목 운영 및 교육프로그램 보유

- 모의·실전창업론, 특허개론, 기술사업화, 기업가정신 등 창업에 필요한 교과목 14개 개설 및 운영

- 학생 대상 세계적 창업선도대학인 이스라엘 테크니온공대 '기업가정신 프로그램' 참여 기회 제공

- 3D 모델링, 3D 프린팅, 시제품 제작, 워크샵 공간을 포함한 G팩토리 복합창업공간을 구축하여 교육과정 및 네트워킹 지원

○ **(기술창업 지원)** 다양한 창업 지원 프로그램 운영

- AI 기술창업교육인 Data, AI, Management(D.A.M) 교육, AI 창업경진대회 꿈꾸는 아이, AI&IoT 메이커톤 창업경진대회 등 다양한 프로그램 등을 통한 기술기업 발굴 및 설립지원

- 연구실에서 개발한 혁신 기술을 새로운 서비스와 제품을 통해 부가가치를 발굴하는 실험실 창업 적극 지원

4) 연구개발성과 활용계획

① 연구개발성과의 특허관리, 기술이전, 사업화 시행

○ 광주과학기술원 과학기술응용연구소(GTI)와 연계한 기술이전, 상용화, 시제품개발 시행; 연구성과를 GTI에 제공하고 GTI는 특허관리, 사업 확장, 애로기술지도 등 지원

○ 수익금은 연구 사업에 재투자하여 지속가능한 R&D형 인력양성 사업 기반 마련

※ 광주과학기술원 과학기술응용연구소(GTI): 기초연구역량을 산업에 접목하여 실용화 및 사업화하며, 지적재산권(IP) 관련 특허 업무, 기술이전 업무 및 우수한 기술 실용화 관련 업무를 수행

② 고급인력 배출을 통한 고용촉진 장려

○ 참여 학생들의 지속적 관리를 통한 취업 촉진 및 기업과 연계한 고용 장려 시스템 구축

○ GIST의 지적 인프라를 활용하여 협력업체의 기술개발에 도움을 줄 수 있는 연결고리 역할 수행

5) 연구개발성과 상용화 계획

① 단계별 개발 수준의 상용화

○ 단계별로 개발되는 연구 성과들의 상용화 우선 수행

○ 상용화 전 연구 성과의 타당성, 기술 경쟁력 등을 면밀히 분석하여 상용화를 시행하고, 단계별 상용화 결과를 고려하여 종합 개발 연구 진행에 적용

② 종합 개발 수준의 상용화

○ 표준화 모델 수립, 응용 기기 개발 등 세부 연구별 최종 성과물에 대한 상용화 시행

○ 연구 성과의 성공적인 상용화를 위해 사전 제품 브랜드 개발 및 홍보

③ 장기적 개발 수준의 상용화

○ 최종 연구 성과에 대한 상용화시행→결과물의 상품성 검토→성능, 유연성, 새로운 접근 방식 등을 통한 결과물 개선 수행

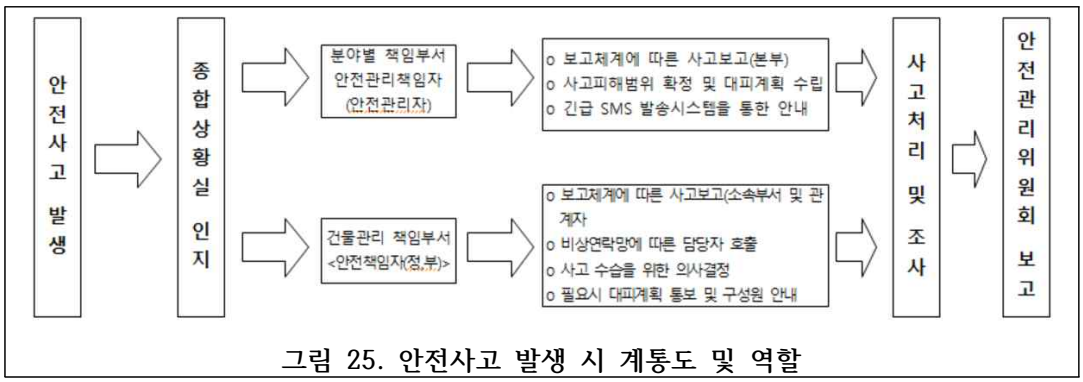
7 연구개발 안전 및 보안조치 이행계획

1) 안전조치 이행계획

① 주관연구개발기관 : 광주과학기술원

구분	세부 계획
안전책임자 지정 여부	O(√), X( )
안전교육 실시 여부	O(√), X( )
안전사고 발생 시 보고 및 조치계획	안전관리규칙에 의거, 안전 관련 업무 관리하는 조직으로 안전팀과 시설운영팀을 두며, 부서별 업무분장이 있음. 구성원이 재해를 당하였을 때, 즉시 재해 정도에 따라 현장에서 필요한 응급조치를 하거나 병원으로 후송조치, 긴급시 전용신고전화(내부2119) 로 안전관리부서에 신속히 통보 후, 안전관리부서는 사고현장에 출두하여 사고보고서를 작성하여 지휘, 보고 절차에 따라 보고, 사고조사를 위해 “사고조사위원회”를 구성할 수 있음.
사고 발생 시 대처방안 및 행동요령	구성원이 재해를 당하였을 때, 즉시 재해 정도에 따라 현장에서 필요한 응급조치를 하거나 병원으로 후송조치, 긴급시 전용신고전화(내부2119) 로 안전관리부서에 신속히 통보 후, 안전관리부서는 사고현장에 출두하여 사고보고서를 작성하여 지휘, 보고 절차에 따라 보고, 사고조사를 위해 “사고조사위원회”를 구성할 수 있음.
정밀안전진단 실시계획	정기적으로 안전보건교육을 실시하며, 근로자 채용 및 업무 변경 시 담당업무와 관련한 안전보건교육을 실시하며, 유해하거나 위험한 작업을 담당하는 경우 추가적인 안전보건교육을 실시하고 있음. 연구실 일상점검 및 매년1회 이상 정기점검 시행, 2년에1회 이상 정기적으로 정밀안전진단을 시행함.

② 참여기관 : 고려대학교, 서울대학교, 아주대학교

구분	세부 계획 (고려대학교)
안전책임자 지정 여부	O(✓), X( )
안전교육 실시 여부	O(✓), X( )
안전사고 발생 시 보고 및 조치계획	근거 : 교내 규정/학칙 - 안전관리규정 - 제6장 사고 조치 및 대책수립 - 제 32조(사고조사 및 보고)에 따름. 내용 : 제 32조 (사고조사 및 보고) ① 안전관리자는 재해 발생 시 사고현장에 출두하여 정확한 사고원인을 별지의 사고보고서에 작성하여 보고하고, 재발 방지를 위한 대처방안을 마련하여야 한다. ② 안전책임자(정.부)는 재해 발생 시 비상연락망에 따라 즉시 보고 및 현장 안전조치를 취해야 하며, 사고원인 조사가 끝날 때까지 현장의 변경 및 훼손 없이 원상태로 보존하고, 사고경위서를 작성하여 주관부서에 보고하여야 한다. ③ 중대 재해가 발생하였거나 원인 규명이 어렵다고 판단될 때에는 외부전문기관에 의뢰 할 수 있다. ④ 사고 발생 시 그 사고의 경중에 따라 교내 구성원들에게 공표할 수도 있다.  그림 25. 안전사고 발생 시 계통도 및 역할 ⑤ 안전사고 발생 시 분야별 사항에 대하여 주관부서 안전관리자는 다음 각호와 같이 보고한다. 안전관리총괄책임자(안전관리위원회)에게 즉시(직접) 보고: 실제 사고 발생으로 관공서(소방서, 경찰서 등)의 공권력이 투입되거나 중대한 인명피해가 발생한 경우 안전관리위원회 종합보고(교내 사건·사고보고): 사고수준이 경미 하여 자체적으로 처리한 경우
사고 발생 시 대처방안 및 행동요령	근거 : 교내 규정/학칙 - 안전관리규정 - 제6장 사고 조치 및 대책수립 - 제 33조(사고 후속 조치)에 따름. 내용 : 제 33조 (사고후속조치) 안전사고의 신속하고 원활한 수습을 위하여 각 행정부서 및 대학은 다음 각호의 사항을 준수하여야 한다. 1. 사고 관련 유관부서(총무부, 건축팀, 시설운영팀) : 신속한 사고수습을 위하여 인력, 장비 등의 지원 협조 2. 공간관리 주관부서(건축팀) 및 각 단과대학 : 사고로 인하여 학사행정 지체 등의 문제가 발생할 경우, 사후 지원 협조

정밀안전진단 실시계획	1) 근거 [연구실안전법 제8조(안전점검의 실시) 및 시행령 제7조(안전점검의 실시 등)] ▪ 연구주체의 장은 연구실의 기능 및 안전을 유지관리하기 위하여 제7조의 규정에 따른 안전점검지침에 따라 소관 연구실에 관한 안전점검을 실시하여야 한다. ▪ 일상점검 : 연구개발활동에 사용되는 기계・기구・전기・약품・병원체 등의 보관상태 및 보호장비의 관리 실태 등을 육안으로 실시하는 점검으로서 연구개발활동을 시작하기 전에 매일 1회 실시 ▪ 정기점검 : 연구개발활동에 사용되는 기계・기구・전기・약품・병원체 등의 보관상태 및 보호장비의 관리 실태 등을 안전점검기기를 이용하여 실시하는 세부적인 점검으로 매년 1회 이상 실시 [산업안전보건법 제42조(작업환경측정 등) 및 시행규칙 제93조(작업환경측정 대상 작업장 등)] ▪ 사업주는 유해인자로부터 근로자의 건강을 보호하고 쾌적한 작업환경을 조성하기 위하여 인체에 해로운 작업을 하는 작업장으로서 고용노동부령으로 정하는 작업장에 대하여 고용노동부령으로 정하는 자격을 가진 자로 하여금 작업환경측정을 하도록 한 후 그 결과를 기록・보존하고 고용노동부령으로 정하는 바에 따라 고용노동부장관에게 보고하여야 한다 2) 목표 및 성과 연구활동종사자가 안전한 환경에서 실험할 수 있도록 안전점검을 통해 연구환경을 개선하여 안전사고를 예방하고 쾌적한 연구환경을 조성하기 위함이다. 3) 내용 가) 일상점검 (1) 점검주체 : 연구(실험)실 (2) 점검시기 : 매일 실험 전 나) 정기점검 (1) 점검주체 : 안전관리팀 연구실안전환경관리자 (2) 점검대상 : 연구(실험)실 약 650개소 (3) 점검시기 : 2021. 04.~ 06.(확인점검:2022.09.~10.) 다) 일상점검 탁상달력 제작 라) 작업환경측정 (1) 측정대상 : 연구(실험)실, 위험물저장소, 폐수처리장, 기계실, 목공실 ※ 고위험 연구(실험)실 10개소 선정 (2) 측정시기 : 2회/년 ※ 최초 측정 시 유해인자가 기준량(소음 90dB) 이하일 때 연 1회 측정

구분	세부 계획 (서울대학교)
안전책임자 지정 여부	O(√), X( )
안전교육 실시 여부	O(√), X( )
안전사고 발생 시 보고 및 조치계획	서울대학교 환경안전연구원 보고 및 재발 방지를 위한 재교육 실시
사고 발생 시 대처방안 및 행동요령	안전교육에서 취급하는 행동요령에 따라 안전한 대피 후 추후 사고 재발 방지를 위한 교육 실시
정밀안전진단 실시계획	실험실 관리현황, 유독물질 저장관리현황, 폐기물 관리현황 등의 점검항목에 대한 철저한 진단 후 자가진단보고서의 정기적 제출

구분	세부 계획 (아주대학교)
안전책임자 지정 여부	O(✓), X( )
안전교육 실시 여부	O(✓), X( )
안전사고 발생 시 보고 및 조치계획	• 안전관리 담당자에게 신속히 연락 • 긴급 상황 발생시 소방서 및 응급실에 연락
사고 발생 시 대처방안 및 행동요령	• 누출처리에 필요한 사람을 제외하고는 안전한 곳으로 대피 • 가능하면 누출된 화학・가스 등이 다른 곳으로 더 이상 흘러나가지 못하도록 조치할 것 • 가능한 화재나 사고를 신속히 진압할 것 • 누출된 화학・가스 제거시 반드시 안전보호구를 착용하고 작업 할 것
정밀안전진단 실시계획	• 1년 단위로 전기안전 및 소방안전 점검 실시

2) 보안조치 이행계획

① 자체보안관리 진단

<자체보안관리진단표: 광주과학기술원>

구분	체크항목	결과 체크 (√표)	비고 (미실시 사유)
보안관리 체계	o 기관 내 보안관리규정을 제정/적용하고 있다	O(√), X( )	
	o 보안관리 조직이 있으며, 자체 보안점검실시 등 잘 운영되고 있다	O(√), X( )	
	o 보안교육을 정기적(1회이상/연)으로 실시하고 있다	O(√), X( )	
	o 보안사고에 대한 방지대책 및 비상시 대응계획이 준비되어 있다	O(√), X( )	
참여연구원 관리	o 참여연구원에 대하여 보안서약서를 받았다	O(√), X( )	
	o 참여연구원에게 보안관리의 중요성 등을 인식시키고 있다	O(√), X( )	
연구개발 내용/결과 관리	o 주요 연구자료 및 결과물의 무단유출 방지대책을 수립하고 있다	O(√), X( )	
	o 보안성 검토 방법 및 절차를 이행하고 있다	O(√), X( )	
	o 기술이전 관련 내부규정 및 절차를 준수하고 있다	O(√), X( )	
연구시설 관리	o 연구시설 보안관련 내부규정 또는 지침을 이행하고 있다	O(√), X( )	
	o 주요 시설에는 보안장비가 설치되어 있다	O(√), X( )	
	o 보호구역이 지정되어 있다	O(√), X( )	
정보통신망 관리	o 정보통신망 보안관련 내부규정 또는 지침이 구비되어 있다	O(√), X( )	
	o 보안관리책임자의 승인 항목이 구분되어 있다	O(√), X( )	
	o 주요 데이터에 대해 백업을 실시하고 있다	O(√), X( )	
	o 개인용 정보통신장비(노트북, USB메모리)에 대하여 인가/관리중이다	O(√), X( )	
	o 전산망 보호를 위한 HW 및 SW 등을 도입하여 적용하고 있다	O(√), X( )	
	o 직책, 임무별 열람 권한을 차등화하여 부여하고 있다	O(√), X( )	
외국인(기관) 보안관리	o 외국기업, 단체 및 외국인의 과제 참여시 보안 절차 등에 대한 내부규정 또는 지침이 구비되어 있다	O(√), X( )	
	o 외부 기관, 단체 및 외국인의 과체 참여시 공동(협동·위탁 포함) 연구협약 시 성과물의 귀속, 자료 제공 및 장비 반납 등에 관한 사전 보안대책이 마련되어 있다	O(√), X( )	
보안과제 관리	o 연구성과물 기술 실시(사용) 계약 시 “제3자 기술 실시(사용)권 금지 협약”이 체결되어 있다	O(√), X( )	
	o 연구시설 출입자에 대한 개인별 출입권한 차등 부여 및 통제	O(√), X( )	
	o 외부방문자 출입 및 연구시설 출입자에 대한 통제 보안이 실시되고 있다	O(√), X( )	
	o 내부망의 연구실별 물리적 또는 논리적(방화벽 등) 분리가 되어 있다	O(√), X( )	
	o 업무용 컴퓨터 자료의 복사 및 외부 전송, 자료 유출 방지 등에 대한 보안 대책이 마련되어 있다	O(√), X( )	

<자체보안관리진단표: 고려대학교>

구분	체크항목	결과 체크 (√표)	비고 (미실시 사유)
보안관리 체계	o 기관 내 보안관리규정을 제정/적용하고 있다	O(√), X( )	
	o 보안관리 조직이 있으며, 자체 보안점검실시 등 잘 운영되고 있다	O(√), X( )	
	o 보안교육을 정기적(1회이상/연)으로 실시하고 있다	O(√), X( )	
	o 보안사고에 대한 방지대책 및 비상시 대응계획이 준비되어 있다	O(√), X( )	
참여연구원 관리	o 참여연구원에 대하여 보안서약서를 받았다	O(√), X( )	
	o 참여연구원에게 보안관리의 중요성 등을 인식시키고 있다	O(√), X( )	
연구개발 내용/결과 관리	o 주요 연구자료 및 결과물의 무단유출 방지대책을 수립하고 있다	O(√), X( )	
	o 보안성 검토 방법 및 절차를 이행하고 있다	O(√), X( )	
	o 기술이전 관련 내부규정 및 절차를 준수하고 있다	O(√), X( )	
연구시설 관리	o 연구시설 보안관련 내부규정 또는 지침을 이행하고 있다	O(√), X( )	
	o 주요 시설에는 보안장비가 설치되어 있다	O(√), X( )	
	o 보호구역이 지정되어 있다	O(√), X( )	
정보통신망 관리	o 정보통신망 보안관련 내부규정 또는 지침이 구비되어 있다	O(√), X( )	
	o 보안관리책임자의 승인 항목이 구분되어 있다	O(√), X( )	
	o 주요 데이터에 대해 백업을 실시하고 있다	O(√), X( )	

구분	체크항목	결과 체크 (√표)	비고 (미실시 사유)
	o 개인용 정보통신장비(노트북, USB메모리)에 대하여 인가/관리중이다	O(√), X( )	
	o 전산망 보호를 위한 HW 및 SW 등을 도입하여 적용하고 있다	O(√), X( )	
	o 직책, 임무별 열람 권한을 차등화하여 부여하고 있다	O(√), X( )	
외국인(기관) 보안관리	o 외국기업, 단체 및 외국인의 과제 참여시 보안 절차 등에 대한 내부규정 또는 지침이 구비되어 있다	O(√), X( )	
	o 외부 기관, 단체 및 외국인의 과체 참여시 공동(협동·위탁 포함) 연구협약 시 성과물의 귀속, 자료 제공 및 장비 반납 등에 관한 사전 보안대책이 마련되어 있다	O(√), X( )	
보안과제 관리	o 연구성과물 기술 실시(사용) 계약 시 “제3자 기술 실시(사용)권 금지 협약”이 체결되어 있다	O(√), X( )	
	o 연구시설 출입자에 대한 개인별 출입권한 차등 부여 및 통제	O(√), X( )	
	o 외부방문자 출입 및 연구시설 출입자에 대한 통제 보안이 실시되고 있다	O(√), X( )	
	o 내부망의 연구실별 물리적 또는 논리적(방화벽 등) 분리가 되어 있다	O(√), X( )	
	o 업무용 컴퓨터 자료의 복사 및 외부 전송, 자료 유출 방지 등에 대한 보안 대책이 마련되어 있다	O(√), X( )	

<자체보안관리진단표: 서울대학교>

구분	체크항목	결과 체크 (√표)	비고 (미실시 사유)
보안관리 체계	o 기관 내 보안관리규정을 제정/적용하고 있다	O(√), X( )	
	o 보안관리 조직이 있으며, 자체 보안점검실시 등 잘 운영되고 있다	O(√), X( )	
	o 보안교육을 정기적(1회이상/연)으로 실시하고 있다	O(√), X( )	
	o 보안사고에 대한 방지대책 및 비상시 대응계획이 준비되어 있다	O(√), X( )	
참여연구원 관리	o 참여연구원에 대하여 보안서약서를 받았다	O(√), X( )	
	o 참여연구원에게 보안관리의 중요성 등을 인식시키고 있다	O(√), X( )	
연구개발 내용/결과 관리	o 주요 연구자료 및 결과물의 무단유출 방지대책을 수립하고 있다	O(√), X( )	
	o 보안성 검토 방법 및 절차를 이행하고 있다	O(√), X( )	
	o 기술이전 관련 내부규정 및 절차를 준수하고 있다	O(√), X( )	
연구시설 관리	o 연구시설 보안관련 내부규정 또는 지침을 이행하고 있다	O(√), X( )	
	o 주요 시설에는 보안장비가 설치되어 있다	O(√), X( )	
	o 보호구역이 지정되어 있다	O(√), X( )	
정보통신망 관리	o 정보통신망 보안관련 내부규정 또는 지침이 구비되어 있다	O(√), X( )	
	o 보안관리책임자의 승인 항목이 구분되어 있다	O(√), X( )	
	o 주요 데이터에 대해 백업을 실시하고 있다	O(√), X( )	
	o 개인용 정보통신장비(노트북, USB메모리)에 대하여 인가/관리중이다	O(√), X( )	
	o 전산망 보호를 위한 HW 및 SW 등을 도입하여 적용하고 있다	O(√), X( )	
	o 직책, 임무별 열람 권한을 차등화하여 부여하고 있다	O(√), X( )	
외국인(기관) 보안관리	o 외국기업, 단체 및 외국인의 과제 참여시 보안 절차 등에 대한 내부규정 또는 지침이 구비되어 있다	O(√), X( )	
	o 외부 기관, 단체 및 외국인의 과체 참여시 공동(협동·위탁 포함) 연구협약 시 성과물의 귀속, 자료 제공 및 장비 반납 등에 관한 사전 보안대책이 마련되어 있다	O(√), X( )	
보안과제 관리	o 연구성과물 기술 실시(사용) 계약 시 “제3자 기술 실시(사용)권 금지 협약”이 체결되어 있다	O(√), X( )	
	o 연구시설 출입자에 대한 개인별 출입권한 차등 부여 및 통제	O(√), X( )	
	o 외부방문자 출입 및 연구시설 출입자에 대한 통제 보안이 실시되고 있다	O(√), X( )	
	o 내부망의 연구실별 물리적 또는 논리적(방화벽 등) 분리가 되어 있다	O(√), X( )	
	o 업무용 컴퓨터 자료의 복사 및 외부 전송, 자료 유출 방지 등에 대한 보안 대책이 마련되어 있다	O(√), X( )	

<자체보안관리진단표: 아주대학교>

구분	체크항목	결과 체크 (√표)	비고 (미실시 사유)
보안관리 체계	o 기관 내 보안관리규정을 제정/적용하고 있다	O(√), X( )	
	o 보안관리 조직이 있으며, 자체 보안점검실시 등 잘 운영되고 있다	O(√), X( )	
	o 보안교육을 정기적(1회이상/연)으로 실시하고 있다	O(√), X( )	
	o 보안사고에 대한 방지대책 및 비상시 대응계획이 준비되어 있다	O(√), X( )	
참여연구원 관리	o 참여연구원에 대하여 보안서약서를 받았다	O(√), X( )	
	o 참여연구원에게 보안관리의 중요성 등을 인식시키고 있다	O(√), X( )	
연구개발 내용/결과 관리	o 주요 연구자료 및 결과물의 무단유출 방지대책을 수립하고 있다	O(√), X( )	
	o 보안성 검토 방법 및 절차를 이행하고 있다	O(√), X( )	
	o 기술이전 관련 내부규정 및 절차를 준수하고 있다	O(√), X( )	
연구시설 관리	o 연구시설 보안관련 내부규정 또는 지침을 이행하고 있다	O(√), X( )	
	o 주요 시설에는 보안장비가 설치되어 있다	O(√), X( )	
	o 보호구역이 지정되어 있다	O(√), X( )	
정보통신망 관리	o 정보통신망 보안관련 내부규정 또는 지침이 구비되어 있다	O(√), X( )	
	o 보안관리책임자의 승인 항목이 구분되어 있다	O(√), X( )	
	o 주요 데이터에 대해 백업을 실시하고 있다	O(√), X( )	
	o 개인용 정보통신장비(노트북, USB메모리)에 대하여 인가/관리중이다	O(√), X( )	
	o 전산망 보호를 위한 HW 및 SW 등을 도입하여 적용하고 있다	O(√), X( )	
	o 직책, 임무별 열람 권한을 차등화하여 부여하고 있다	O(√), X( )	
외국인(기관) 보안관리	o 외국기업, 단체 및 외국인의 과제 참여시 보안 절차 등에 대한 내부규정 또는 지침이 구비되어 있다	O(√), X( )	
	o 외부 기관, 단체 및 외국인의 과제 참여시 공동(협동·위탁 포함) 연구협약 시 성과물의 귀속, 자료 제공 및 장비 반납 등에 관한 사전 보안대책이 마련되어 있다	O(√), X( )	
보안과제 관리	o 연구성과물 기술 실시(사용) 계약 시 “제3자 기술 실시(사용)권 금지 협약”이 체결되어 있다	O(√), X( )	
	o 연구시설 출입자에 대한 개인별 출입권한 차등 부여 및 통제	O(√), X( )	
	o 외부방문자 출입 및 연구시설 출입자에 대한 통제 보안이 실시되고 있다	O(√), X( )	
	o 내부망의 연구실별 물리적 또는 논리적(방화벽 등) 분리가 되어 있다	O(√), X( )	
	o 업무용 컴퓨터 자료의 복사 및 외부 전송, 자료 유출 방지 등에 대한 보안 대책이 마련되어 있다	O(√), X( )	

8 연구개발기관 현황

1) 연구책임자 등 현황

① 주관연구개발기관 연구책임자 (제1세부 프로젝트 책임)

가. 인적사항

개인	국문	이흥노	국적	대한민국
	영문	Heung-No Lee	국가연구자번호	[redacted]
직장	기관명	광주과학기술원	전화번호	[redacted]
	부서	전기전자컴퓨터공학부	휴대전화	[redacted]
	직위	교수	전자우편	[redacted]
	주소	(우:61005) 광주광역시 북구 첨단과기로123 광주과학기술원 전기전자컴퓨터공학부 C-204호		

나. 학력

취득연월	학교명	전공	학위	지도교수
1994.11~1999.12	University of California, Los Angeles	Electrical Eng.	박사	Gregory J. Pottie
1993.09~1994.10	University of California, Los Angeles	Electrical Eng.	석사	
1990.09~1993.05	University of California, Los Angeles	Electrical Eng.	학사	
최종학위 논문명: Adaptive Diversity Combining, Equalization and Sequence Decoding for Time-Varying Dispersive Channels				

다. 경력

기간	기관명	직위	비고
2021.01.18~2023.01.17	광주과학기술원	GIST 인공지능연구소 소장	
2017.01.01~2019.12.31	대한전자공학회	통신소사이어티 회장	
2017.01.01~2018.12.31	대한전자공학회	부회장 (등기이사)	
2016.09.13~2017.11.30	광주과학기술원	GIST 연구원장/ 전략연구사업단장/ 연구정책센터장	

라. 주요 연구개발 실적

중앙행정기관 (전문기관)	세부사업명	연구개발과제명	주관연구 개발기관 / 당시 소속기관	연구개발기간 (참여한 기간)	역할: 연구책임자 /연구자	비고 (신청/수행 중/완료)
한국연구재단	중견연구자지원사업	전파신호식별 및 초분광이미징에 특화된 지능형 신호 복원 및 분류 시스템 연구	광주과학기술원 / 광주과학기술원	21.03.01~24.02.29	연구책임자	수행중
정보통신기획평가원	블록체인융합기술개발	일반 연산을 검증하고 트랜잭션 검증량과 저장 공간을 줄여주는 유니버셜 영지식 증명 서킷 기반 가상머신 개발	(주) 온더 / 광주과학기술원	20.04.01~21.12.31	공동연구원	수행중
한국연구재단	뇌과학원천사업	다중생체정보 기반 치매 중증도 모니터링 기술 개발	광주과학기술원 / 광주과학기술원	16.06.01~21.08.31 (16.06.01~21.08.31)	연구책임자	수행중
한국연구재단	중견연구자지원사업	레이다, 분광기, 뇌컴퓨터 인터페이스 시스템에 특화된 지능형 신호복원 및 분류 시스템 연구	광주과학기술원 / 광주과학기술원	18.03.01~21.02.28	연구책임자	완료
정보통신기획평가원	블록체인융합기술개발	확장가능한 탈중앙화 보안성 ECCPoW 블록체인	광주과학기술원 / 광주과학기술원	19.04.01~19.12.31 (19.04.01~19.12.31)	연구책임자	완료

마. 대표적 논문/저서 실적

구분 (논문/저서)	논문명/저서명	게재지(권, 쪽)	게재연도 (발표연도)	역할	비고(피인용 지수)
논문	Hilbert Transform Design based on Fractional Derivatives and Swarm Optimization	IEEE Transactions on Cybernetics (50(5), 2311-2320	2020	교신저자	IF : 11.079
논문	ECCPoW: Error-Correction Code based Proof-of-Work for ASIC Resistance	SYMMETRY-BASEL (12(6), 988)	2020	교신저자	IF : 2.645
논문	Profitable Double-Spending Attacks	Applied Sciences (10(23), 1-23)	2020	교신저자	IF : 2.474
논문	Blockchain-Based Distributed Patient-Centric Image Management System	Applied Sciences (11(1), 196)	2020	교신저자	IF : 2.474
논문	Fabrication of 2D thin-film filter-array for compressive sensing spectroscopy	Optics and Lasers in Engineering (115(2019), 53-58)	2019	연구책임자	IF : 4.273

바. 지식재산권 출원・등록 실적

구분 (특허/프로그램 등)	지식재산권명	국가명	출원・등록일	출원・등록번호/ 출원・등록자 수	비고
특허	블록체인 전자투표시스템, 그 시스템의 운용방법	대한민국	2020-12-09	2020-0171424/2명	
특허	검표자가 필요 없는 영지식 증명과 스마트 컨트랙트 기반 블록체인 투표 시스템	미국	2020-11-12	63/112723/2명	
특허	블록체인의 거래검증시스템, 및 블록체인이 거래검증방법	PCT	2019-12-12	PCT/KR2019/017571/2명	
특허	부호-암호 화폐 시스템	대한민국	2019-11-22	2019-0151246/1명	
특허	새로운 블록체인 및 암호화폐 작업증명 생성, 증명, 검증 시스템	미국	2018-07-18	62/699751/4명	

사. 대표적 기타 실적

구분	실적명	내용요약	실적연도
수상실적	GIST 연구상	GIST 연구상 수상	2020
수상실적	해동학술상	전자공학회 해동학술상 수상	2019
MOU 체결	GIST블록체인경제센터와 ConsenSys 공동양해각서	블록체인 자문, 홍보, 기술 구현 지원에 관한 업무협약체결	2019
기술실시 계약	차세대 체감형 로봇게임기 및 인공지능 기술 노하우	기술실시 계약 금액 10,000,000 원	2018
토론회 개최	블록체인,AI,빅데이터 융합산업 육성 정책토론회 개최	장소 : 여의도 국회본청 목적 : 블록체인 기반 데이터 산업을 육성, 기업의 자율적인 시장 환경 조성을 위해 개최	2018

② 제2세부 프로젝트 책임자

가. 인적사항

개인	국문	김종원	국적	대한민국
	영문	JongWon Kim	국가연구자번호	[redacted]
직장	기관명	광주과학기술원	전화번호	[redacted]
	부서	AI대학원	휴대전화	[redacted]
	직위	교수	전자우편	[redacted]
	주소	(우:61005) 광주광역시 북구 첨단과기로 123 광주과학기술원 EECS-C 207호		

나. 학력

취득연월	학교명	전공	학위	지도교수
89.03~94.02	서울대학교	제어계측공학	박사	이상욱
87.03~89.02	서울대학교	제어계측공학	석사	이상욱
83.03~87.02	서울대학교	제어계측공학	학사	이상욱
최종학위 논문명: 계층적 가변 블럭 크기를 이용한 움직임 영상 부호화 기법				

다. 경력

기간	기관명	직위	비고
2019 ~ 현재	광주과학기술원 AI대학원	대학원장	
2019년 ~ 현재	FIF (Future Internet Forum)	Chair	
2016년 ~ 현재	ICT Express (Elsevier)	Editorial board member	
2014년 ~ 현재	APAN (Asia-Pacific Advance Network) Community Technology Area & Cloud WG	Area Director & WG Chair	
2008년 ~ 2019	광주과학기술원 전기전자컴퓨터공학부	교수	
2008년 ~ 현재	광주과학기술원 슈퍼컴퓨팅센터 (SCENT: Super Computing CENTer)	센터장	
2005년 ~ 현재	Journal of Visual Communication and Image Representation (Elsevier)	Editorial board member	
2012년 ~ 2012년	광주과학기술원 정보통신공학부	학부장	
2011년 ~ 2019년	Asia Future Internet Forum (AsiaFI)	Steering Group Member	
2011년 ~ 2018년	FIF (Future Internet Forum) Testbed WG	WG Chair	
2011년 ~ 2014년	APAN (Asia-Pacific Advance Network) Community Network Research Group	Group Chair	
2009년 ~ 2015년	IEEE Circuit & System Society - MSATC (Technical Committee on Multimedia Systems and Application)	Voting Member	
2009년 ~ 2010년	NCSU (North Carolina State University) 컴퓨터과학부	Visiting Scholar	
2006년 ~ 2008년	광주과학기술원 정보통신공학과	학과장	

라. 주요 연구개발 실적

중앙행정기관 (전문기관)	세부사업명	연구개발과제명	주관연구 개발기관 / 당시 소속기관	연구개발기간 (참여한 기간)	역할: 연구책임자/연구자	비고 (신청/수행중/완료)
NIPA	인공지능중심 산업융합 집적단지 조성사업	악천후 등 외부환경 대응 가능한 V2X 기반 connected 플랫폼 기술 개발	광주과학기술원 / 광주과학기술원	2020.5.1~2024.12.31 (2020.5.1~2024.12.31)	연구책임자	수행중
IITP	인공지능핵심고급인재 양성	인공지능대학원지원	광주과학기술원 / 광주과학기술원	2019.9.1.~2023.12.31 (2019.9.1.~2023.12.31)	연구책임자	수행중
IITP	SW컴퓨팅산업원천기술개발사업	이종 다수 클라우드 간의 자동화된 SaaS 호환성 지원 기술 개발	광주과학기술원 / 광주과학기술원	2016.4.1~2018.12.31 (2016.4.1~2018.12.31)	연구책임자	완료
NIA	미래네트워크선도시험망(KOREN/TEIN) 실증시험과제	산업인터넷 확산을 위한 KOREN 오픈플랫폼 실증	광주과학기술원 / 광주과학기술원	2017.6.26~2017.11.30 (2017.6.26~2017.11.30)	연구책임자	완료
IITP	정보통신방송연구개발사업	평창 동계올림픽 5G 통신서비스 지원을 위한 차세대 통신네트워크 기술개발	광주과학기술원 / 광주과학기술원	2016.6.1~2018.5.31 (2016.6.1~2018.5.31)	연구자	완료

마. 대표적 논문/저서 실적

구분 (논문/ 저서)	논문명/저서명	게재지(권, 쪽)	게재연도 (발표연도)	역할	비고(피 인용 지수)
논문	Mobile IoT device Summarizer using P2P web search engine and inherent characteristic of contents	Peer-to-Peer Networking and Applications13, 684-693	2019	교신	2.397
논문	SmartX Multi-View Visibility Framework with Flow-Centric Visibility for SDN-Enabled Multisite Cloud Playground	Applied Sciences9(10), 2045	2019	교신	1.484
논문	Interactive Visualization of SDN-Enabled Multisite Cloud Playgrounds Leveraging SmartX MultiView Visibility Framework	The Computer Journal 62(6)	2018	교신	0.98
논문	International Network Performance and Security Testing based on Distributed Abyss Storage Cluster and Draft of Data Lake Framework	Security and Communication Networks 2018	2018	교신	0.904
논문	Global experimental verification of Docker-based secured mVoIP to protect against eavesdropping and DoS attacks	EURASIP Journal on Wireless Communications and Networking 63	2017	교신	1.529

바. 지식재산권 출원・등록 실적

구분 (특허/프로그램 등)	지식재산권명	국가명	출원・등록일	출원・등록번호/ 출원・등록자 수	비고
특허	Integrated network controlling system, network controlling apparatus and method, and a readable media containing the method	미국	2019.02.19	10,212,216	
특허	Method for detecting intrusion in network	미국	2016.06.23	09,876,808B2	
특허	Flow control node for managing throughput fairness of a plurality of flows, transmission node, method for controlling flow, and method for controlling transmission rate	미국	2016.01.26	09,246,813	
특허	클라우드 서비스를 위한 사용자 인터페이스	대한민국	2019.06.03	10-1987018	
특허	클러스터 시각화 장치	대한민국	2019.04.04	10-1967744	

③ 제3세부 프로젝트 책임자

가. 인적사항

개인	국문	김 기 형	국적	대한민국
	영문	Ki-Hyung Kim	국가연구자번호	
직장	기관명	아주대학교	전화번호	
	부서	사이버보안학과	휴대전화	
	직위	교수	전자우편	
	주소	(우: 16499) 경기도 수원시 영통구 월드컵로 206 아주대학교 산학원 516호		

나. 학력

취득연월	학교명	전공	학위	지도교수
1986년 3월 ~ 1990년 2월	한양대학교	전자통신공학	공학사	한영열
1990년 3월 ~ 1992년 2월	한국과학기술원	전자공학 - (컴퓨터공학)	공학석사	박규호
1992년 3월 ~ 1996년 8월	한국과학기술원	전자공학 - (컴퓨터공학)	공학박사	박규호

최종학위 논문명:

Distributed Simulation Methology Based on System Theoretic Formalism : An Asynchronous Approach

시스템 이론적 형식론에 기반한 분산 시뮬레이션 방법론 : 비동기적 접근

다. 경력

기간	기관명	직위	비고
2005.03 ~ 현재	아주대학교	교수	
2011.09 ~ 2012.08	Stony Brook University	교환교수	
2000.12 ~ 2001.08	AdForce Inc.	Senior Engineer	美 캘리포니아 산호세 소재
1997.03 ~ 2005.02	영남대학교	부교수	

라. 주요 연구개발 실적

중앙행정기관 (전문기관)	세부사업명	연구개발과제명	주관연구 개발기관 / 당시 소속기관	연구개발기간 (참여한 기간)	역할: 연구책임자 /연구자	비고 (신청/수행중/완료)
한국연구재단	이공학개인기초기본연구	IoT(LoRa)단말과 데이터의 인증 및 무결성 확보를 위한 블록체인기반 보안 프레임워크 연구(BIoT)	아주대학교 / 아주대학교	2018.06.01. ~ 2021.05.31.	연구책임자	수행중
정보통신긴술진흥센터	대학ICT연구센터지원사업	지능형 Internet of Energy(IoE) Data 연구	국민대학교 / 아주대학교	2018.06.01. ~ 2021.12.31.	연구책임자	수행중
금융감독원	정부용역사업	블록체인 기반 빅데이터 수집과 분석 인프라 구축 방안 연구	아주대학교 / 아주대학교	2018.05.24. ~ 2018.12.15.	연구책임자	완료

마. 대표적 논문/저서 실적

구분 (논문/저서)	논문명/저서명	게재지(권, 쪽)	게재연도 (발표연도)	역할	비고(피인용 지수)
논문	Image Compression using Chain Coding for Electronic Shelf Labels (ESL) Systems	IEEE Access (9)	2021	공동	4.64
논문	CEnsLoc: Infrastructure-Less Indoor Localization Methodology Using GMM Clustering-Based Classification Ensembles.	MOBILE INFORMATION SYSTEMS	2018	교신	0.958
논문	Escalator: An Autonomous Scheduling Scheme for Convergecast in TSCH	SENSORS (18.4)	2018	교신	2.475
논문	A hybrid mode to enhance the downward route performance in routing protocol for low power and lossy networks	International Journal of Distributed Sensor Networks (14.4)	2018	교시	1.787
논문	An Algorithm for Timely Transmission of Solicitation Message in RPL for Energy-Efficient Node Mobility	SENSORS (17.4)	2017	공동	2.033

바. 지식재산권 출원・등록 실적

구분 (특허/프로그램 등)	지식재산권명	국가명	출원・등록일	출원・등록번호/ 출원・등록자 수	비고
특허	ESL 관리 서버, 디스플레이 장치 및 그 제어 방법	대한민국	2020.12.01.	10-2020-0165925	
특허	TSCH에서의 컨비지캐스트를 위한 자율 스케줄링 방법	대한민국	2018.04.20.	10-2018-0046319	

④ 제1세부 프로젝트 참여연구자

가. 인적사항

개인	국문	문수묵	국적	한국
	영문	Soo-Mook Moon	국가연구자번호	[redacted]
직장	기관명	서울대학교	전화번호	[redacted]
	부서	전기정보공학부	휴대전화	[redacted]
	직위	교수	전자우편	[redacted]
	주소	(우: 08826) 서울특별시 관악구 관악로1 서울대학교 공과대학 301동 801호		

나. 학력

취득연월(최근 순으로 작성)	학교명	전공	학위	지도교수
90.9 ~ 93.8	Univ. of Maryland	컴퓨터공학	이학 박사	Ashok Agrawala
88.9 ~ 90.8	Univ. of Maryland	컴퓨터공학	이학 석사	Ashok Agrawala
83.3 ~ 87.2	서울대학교	컴퓨터공학	공학사	김영택
최종학위 논문명: Compile-time Parallelization of Nonnumerical Code VLIW and Superscalar				

다. 경력

기간	기관명	직위	비고
93.07~94.08	Hewlett-Packard Company	A software design engineer	
94.09~현재	서울대학교 전기정보공학부	교수	
01.12~03.02	Sun Microsystems	A visiting professor	
07.08~07.08	IBM T. J. Watson Research Center	A visiting scientist	

라. 주요 연구개발 실적

중앙행정기관 (전문기관)	세부사업명	연구개발과제명	주관연구개발기관 / 당시 소속기관	연구개발기간 (참여한 기간)	역할: 연구책임자/ 연구자	비고 (신청/수행중/ 완료)
LG전자	산학협력	웹 앱의 로딩 타임 개선을 위한 V8 snapshot 최적화 적용	서울대학교 / 서울대학교	18.04.01~18.12.31 (18.04.01~18.12.31)	연구책임자	완료
삼성전자	산학협력	IoT향 Blockchain Platform 설계 및 개발	서울대학교 / 서울대학교	18.12.06~19.06.05 (18.12.06~19.06.05)	연구책임자	완료
뉴럴비씨	산학협력	고성능 탈중앙화 거래소(DEX)를 위한 2-way relayer 개발	서울대학교 / 서울대학교	18.12.01~19.05.31 (18.12.01~19.05.31)	연구책임자	완료
한국연구재단	중견연구	저사양 임베디드 기기에서 오프로딩을 이용한 머신러닝 앱 수행 환경	서울대학교 / 서울대학교	17.03.01~20.02.29 (17.03.01~20.02.29)	연구책임자	완료
한국연구재단	중견연구	탈중앙화 방식의 딥 러닝 플랫폼	서울대학교 / 서울대학교	20.03.02~현재 (20.03.02~현재)	연구책임자	수행중

마. 대표적 논문/저서 실적

구분(논문/저서)	논문명/저서명	게재지(권, 쪽)	게재연도 (발표연도)	역할	비고(피인용 지수)
논문게재	Ethanos: Efficient Bootstrapping for Full Nodes on Account-based Blockchain	EuroSys	2021	공동저자	-
논문게재	PerDNN: Offloading Deep Neural Network Computations to Pervasive Edge Servers	40th IEEE International Conference on Distributed Computing Systems (ICDCS 2020)	2020	공동저자	-
논문게재	Dynamic Offloading of Web Application Execution Using Snapshot	ACM Transactions on the Web (TWEB) 14.4 (2020): 1-24.	2020	공동저자	IF:1.157
논문게재	Snapshot-based Loading Acceleration of Web Apps with Nondeterministic JavaScript Execution	The 2019 World Wide Web Conference (WWW' 19)	2019	공동저자	-
논문게재	Web Application Migration with Closure Reconstruction	The 2017 World Wide Web conference (WWW 2017)	2017	공동저자	-

바. 지식재산권 출원・등록 실적

구분 (특허/프로그램 등)	지식재산권명	국가명	출원・등록일	출원・등록번호/ 출원・등록자 수	비고
특허	Code Rearranger and Method for Rearranging Code for a Virtual Machine	미국	2016	US9262186B2 / 5	
특허	Multimedia data pre-processing on idle core by detecting multimedia data in application	미국	2016	US9372720B2 / 3	

⑤ 제1세부 프로젝트 참여연구자

가. 인적사항

개인	국문	신종원	국적	한국
	영문	Jong Won Shin	국가연구자번호	[redacted]
직장	기관명	광주과학기술원	전화번호	[redacted]
	부서	전기전자컴퓨터공학부	휴대전화	[redacted]
	직위	부교수	전자우편	[redacted]
	주소	(우: 61005) 광주광역시 북구 첨단과기로 123 광주과학기술원 전기전자컴퓨터공학부 c동 503호		

나. 학력

취득연월	학교명	전공	학위	지도교수
02.03~08.08	서울대학교	전기 및 컴퓨터공학	박사	김남수
98.03~02.02	서울대학교	전기공학	학사	김남수
최종학위 논문명: Pre- and Post-processing Algorithms for Speech Coders in Adverse Environment				

다. 경력

기간	기관명	직위	비고
18.03~	광주과학기술원	부교수	
12.09~18.02	광주과학기술원	조교수	
08.12~12.08	Qualcomm Inc.	Senior Engineer	

라. 주요 연구개발 실적

중앙행정기관 (전문기관)	세부사업명	연구개발과제명	주관연구 개발기관 / 당시 소속기관	연구개발기간 (참여한 기간)	역할: 연구책임자/연구자	비고 (신청/수행중/완료)
과학기술정보통신부 (한국전자통신연구원)	초실감 입체공간 미디어・콘텐츠 원천기술 연구	음향 압축 손실 개선을 위한 전/후처리 기술 개발	한국전자통신연구원 / 광주과학기술원	20.03.16~24.12.31 (20.03.16~)	연구책임자	수행중
과학기술정보통신부 (국토교통과학기술진흥원)	혁신성장동력프로젝트	다중 화자간 대화 음성인식 기술개발	한국전자통신연구원 / 광주과학기술원	19.04.01~21.12.31 (19.04.01~)	연구책임자	수행중
과학기술정보통신부 (한국연구재단)	과학기술분야 기초연구사업	무선 이어버즈의 다중 마이크로폰을 활용한 음질 개선 연구	광주과학기술원 / 광주과학기술원	19.04.01~21.12.31 (19.04.01~)	연구책임자	수행중
문화체육관광부 (한국콘텐츠진흥원)	문화기술연구개발지원사업	사회적 소통약자의 접근성을 고려한 지능형 도구 기반 콘텐츠 제작 및 향유 지원 기술 개발	광주과학기술원 / 광주과학기술원	19.08.30~21.12.31 (19.08.30~)	연구책임자	수행중
산업통상자원부 (산업통상자원부)	산업기술혁신사업	로봇용 free-running 임베디드 자연어 대화음성인식을 위한 원천기술 개발	한양대학교 / 광주과학기술원	17.4.1~20.12.31 (17.4.1~20.12.31)	연구책임자	수행완료

마. 대표적 논문/저서 실적

구분 (논문/저서)	논문명/저서명	게재지(권, 쪽)	게재연도 (발표연도)	역할	비고(피인용지수)
발표	Time-Domain Speaker Verification Using Temporal Convolutional Networks	Proceedings of International Conference on Acoustics, Speech, and Signal Processing	2021	교신저자	
논문게재	Affective Latent Representation of Acoustic and Lexical Features for Emotion Recognition	Sensors (20, 2614)	2020	교신저자	
논문게재	Sound Localization Based on Phase Difference Enhancement Using Deep Neural Networks	IEEE/ACM Transactions on Audio, Speech and Language Processing (27, 1335)	2019	교신저자	
발표	DNN-based Emotion Recognition Based on Bottleneck Acoustic Features and Lexical Features	Proceedings of International Conference on Acoustics, Speech, and Signal Processing	2019	교신저자	
논문게재	Multichannel Speech Reinforcement Based on Binaural Unmasking	Signal Processing (139, 165)	2017	교신저자	

바. 지식재산권 출원・등록 실적

구분 (특허/프로그램 등)	지식재산권명	국가명	출원・등록일	출원・등록번호/ 출원・등록자 수	비고
특허	Residual echo estimator to estimate residual echo based on time correlation, non-transitory computer-readable medium storing program code to estimate residual echo, and application processor	미국	20.12.01	10,854,215/5	
특허	Method of enhancing speech using variable power budget	미국	19.03.26	10,242,691/2	
특허	Audio user interaction recognition and application interface	미국	18.09.11	10,073,521/3	
특허	Audio user interaction recognition and context refinement	미국	17.08.15	9,319,510/3	
특허	음성 향상 방법	대한민국	20.03.02	10-2085739/2	

⑥ 제2세부 프로젝트 참여연구자

가. 인적사항

개인	국문	김 황 남	국적	대한민국
	영문	Hwangnam Kim	국가연구자번호	[redacted]
직장	기관명	고려대학교	전화번호	[redacted]
	부서	전기전자공학부	휴대전화	[redacted]
	직위	전임교수	전자우편	[redacted]
	주소	(우:02841) 서울특별시 성북구 안암로 145 고려대학교 창의관 320호		

나. 학력

취득연월	학교명	전공	학위	지도교수
99.08~04.12	University of Illinois at Urbana-Champaign	컴퓨터과학	공학박사	Jennifer C. Hou
92.03~94.02	서울대학교	컴퓨터공학	공학석사	Sukho Lee
88.03~92.02	부산대학교	컴퓨터공학	공학사	
최종학위 논문명: Enabling Theoretical Model Based Techniques for Simulating Large Scale Networks				

다. 경력

기간	기관명	직위	비고
94.01~99.06	LG전자 (구 LG정보통신)	선임연구원	현 주임연구원
00.03~01.02	Bytemobile, Inc., California	SW Engineer	창사멤버
04.12~05.08	University of Illinois at Urbana-Champaign	Post-doctor	
05.09~06.01	삼성전자	책임연구원	
12.01~12.12	University of California at Los Angeles	방문연구원	
06.03~현재	고려대학교	교수	

라. 주요 연구개발 실적

중앙행정기관 (전문기관)	세부사업명	연구개발과제명	주관연구 개발기관 / 당시 소속기관	연구개발기간 (참여한 기간)	역할: 연구책임자 /연구자	비고 (신청/수행중/ 완료)
-	-	Ambient Device Access over Virtualized Network	Google Inc. / 고려대학교	20.11.02~21.11.01 (20.11.02~21.11.01)	연구책임자	수행중
대한민국 방위사업청	국방특화연구실사업	군집형 무인 CPS 특화연구실	국방과학연구소 / 고려대학교	19.12.15~24.12.31 (19.12.15~24.12.31)	연구자	수행중
산업통상자원부	에너지인력양성사업	에너지거래 플랫폼 융합인력양성	산업통상자원부 / 고려대학교	20.04.01~24.12.31 (20.04.01~24.12.31)	공동책임자	수행중
과학기술정보통신부	중견연구자지원사업	광대역 드론 감시, 추적 및 무력화 기반 지능형 안티드론 시스템	한국연구재단 / 고려대학교	20.03.01~25.02.28 (20.03.01~25.02.28)	연구책임자	수행중

마. 대표적 논문/저서 실적

구분 (논문/ 저서)	논문명/저서명	게재지(권, 쪽)	게재연도 (발표연도)	역할	비고(피인용 지수)
논문	Channel Quality Estimation for Improving Awareness of Communication Situation in the 2.4 GHz ISM Band	IEEE Transactions on Mobile Computing (Volume: 17, Page: 2002 - 2013)	2018	교신	IF 5.112
논문	Wi-Fi Seeker: A link and Load Aware AP Selection Algorithm	IEEE Transactions on Mobile Computing (Volume: 16, Page: 2366 - 2378)	2017	교신	IF 5.112
논문	Future of IoT networks: A survey	MDPI Applied Sciences (Appl. Sci. 2017, 7(10), 1072;)	2017	교신	IF 2.474
논문	Preserving privacy and efficiency in data communication and aggregation for AMI network	Journal of Network and Computer Applications (Volume: 59, Page: 333 - 344)	2016	교신	IF 5.57
논문	DAG-Based Distributed Ledger for Low-Latency Smart Grid Network	MDPI Energies (Energies 2019, 12(18), 3570;)	2019	교신	IF 2.702

바. 지식재산권 출원・등록 실적

구분 (특허/프로그램 등)	지식재산권명	국가명	출원・등록일	출원・등록번호/ 출원・등록자 수	비고
특허	METHOD AND APPARATUS OF ALLOCATING RESOURCES FOR TERMINAL IN WIRELESS COMMUNICATION SYSTEM	United States	201803	US 9,913,280 B2 / 7	
특허	METHOD AND APPARATUS FOR SELECTING POWER PREFERENCE INFORMATION IN A WIRELESS COMMUNICATION SYSTEM NETWORK	United States	201804	US 9,949,208 B2 / 7	
특허	APPARATUS AND METHOD FOR DETECTING NETWORK CONGESTION BASED ON DESTINATION NODE	United States	201711	US 9,819,600 B2 / 2	
특허	APPARATUS AND METHOD FOR DETECTING CONGESTION OF WIRELESS NETWORK IN COMMUNICATION SYSTEM	United States	201711	US 9,826,428 B2 / 6	
특허	APPARATUS AND METHOD FOR ATTEMPTING CONNECTION TO AN ACCESS POINT	United States	201707	US 9,706,488 B2 / 6	

사. 대표적 기타 실적

구분	실적명	내용요약	실적연도
수상	Infocom 2020 Distinguished TPC Award	국제 학회 Infocom 2020에서 Distinguished TPC 수상	2020
수상	국내 학술지 우수논문상 수상	국내 학술지 Korea Institute of communications and Information Sciences에서 2020 해동우수논문 장려상 수상	2020
수상	ICT EXPRESS BEST EDITOR AWARD	국제 학술지 ICT Express (SCIE)의 Best Editor Award 수상	2016
운영위원	TPC Chair	국제 학회 ICUFN 2017의 TPC Chair로써 참여	2017
편집위원	Guest Editor of Special Issue	국제 학술지 MDPI Sensors의 Special issue Guest Editor	2020

⑦ 제3세부 프로젝트 참여연구자

가. 인적사항

개인	국문	황의석	국적	대한민국
	영문	Euiseok Hwang	국가연구자번호	[redacted]
직장	기관명	광주과학기술원	전화번호	[redacted]
	부서	전기전자컴퓨터공학부	휴대전화	[redacted]
	직위	부교수	전자우편	[redacted]
	주소	(우: 61005) 광주광역시 북구 첨단과기로 123(오룡동, 광주과학기술원) 신재생에너지공학동 208호		

나. 학력

취득연월	학교명	전공	학위	지도교수
2007.05~2011.08	Carnegie Mellon University	전기컴퓨터공학	박사	B.V.K. Vijaya Kumar
2006.08~2007.05	Carnegie Mellon University	전기컴퓨터공학	석사	B.V.K. Vijaya Kumar
1999.03~2000.02	서울대학교	기계설계학	석사	강연준
1994.03~1998.02	서울대학교	원자핵공학	학사	황일순
최종학위 논문명: Storage channels with write errors: Two-dimensional magnetic recording and advanced memory systems				

다. 경력

기간	기관명	직위	비고
2000.02~2006.07	대우전자	주임/선임/책임 연구원	
2011.09~2014.12	LSI Corp. (now Broadcom)	Senior research staff	

라. 주요 연구개발 실적

중앙행정기관 (전문기관)	세부사업명	연구개발과제명	주관연구 개발기관 / 당시 소속기관	연구개발기간 (참여한 기간)	역할: 연구책임자/연구자	비고 (신청/수행중/완료)
환경부 (한국환경산업기술원)	글로벌탑환경기술개발사업	레이저유도 기술 기반 금속류 자동선별 상용장치 개발	광주과학기술원 / 광주과학기술원	16.08.01~21.04.30 (16.08.01~21.04.30)	연구자	수행중
과학기술정보통신부 (정보통신기술진흥센터)	방송통신산업기술개발	Cellular IoT 환경에서 물리계층장치 인식 기반 간결한 보안 통신 연구	광주과학기술원 / 광주과학기술원	17.04.01~21.06.30 (18.10.01~21.06.30)	연구책임자	수행중
과학기술정보통신부 (정보통신산업진흥원)	인공지능중심산업융합집적단지조성	AI 기반의 사전예방 및 빠른 복구가 가능한 에너지 재난대응 플랫폼 개발	서울대학교 / 광주과학기술원	20.04.01~23.12.31 (20.04.01~23.12.31)	연구자	수행중
과학기술정보통신부 (한국연구재단)	중견연구	대규모 IoT 환경을 위한 심층 군집화 기반 가벼운 전력 데이터 압축 프레임워크	광주과학기술원 / 광주과학기술원	21.03.01~24.02.29 (21.03.01~24.02.29)	연구책임자	수행중
과학기술정보통신부 (한국연구재단)	기본연구	인터레이스 기록 기반 고밀도 다중레벨 자기 정보 저장기기 기술과 적용 가능 비이진 등가 채널 신호처리 기술 연구	광주과학기술원 / 광주과학기술원	17.06.01~20.05.31 (17.06.01~20.05.31)	연구책임자	수행완료

마. 대표적 논문/저서 실적

구분 (논문/ 저서)	논문명/저서명	게재지(권, 쪽)	게재연도 (발표연도)	역할	비고(피인용 지수)
논문	Time—Frequency Mask Estimation based on Deep Neural Network for Flexible Load Disaggregation in Buildings	IEEE Transactions on Smart Grid (Accepted)	2021	교신저자	8.267
논문	On the Error Probability of Support Recovery for Orthogonal Matching Pursuit With a Random Measurement Matrix	IEEE Access (8, 95503-95511)	2020	교신저자	3.745
논문	Hybrid Load Forecasting for Mixed-use Complex based on Characteristic Load Decomposition by Pilot Signals	IEEE Access (7, 12297-12306)	2019	교신저자	3.745
논문	A Unified Approach for Compression and Authentication of Smart Meter Reading in AMI	IEEE Access (7, 34383-34394)	2019	공동저자	3.745
논문	Secure multiple access based on multicarrier CDMA with induced random flipping	IEEE Transactions on Vehicular Technology (66, 5099-5108)	2017	공동저자	5.379

바. 지식재산권 출원・등록 실적

구분 (특허/프로그램 등)	지식재산권명	국가명	출원・등록일	출원・등록번호/ 출원・등록자 수	비고
특허	레이저 유도붕괴 분광분석을 이용한 금속 분류 시스템 및 그것의 동작 방법	대한민국	2020.03.09	10-2089068/4명	
특허	딥러닝 기반 마스크 추정 방법을 이용한 건물 부하 식별 시스템의 동작 방법	대한민국	2020.02.20	10-2081912/2명	
특허	전기자동차 충방전 시스템과 그 방법	대한민국	2020.01.03	10-2064525/2명	
특허	레이저 유도 붕괴 분광법을 이용한 성분 표시 장치	대한민국	2019.09.17	10-2023913/2명	
특허	전력 수요 예측 방법 및 그 시스템	대한민국	2019.01.16	10-1941003/3명	

⑧ 참여연구자 및 연구근접지원인력

가. 참여연구자 현황

구분	대학 교수	산중 교수	전임 연구원	참여학생			산업체	기타 (포닥, 연구교수 등)	합계
				학사	석사	박사			
인원	7명	0명	1명	0명	24명	26명	0명	6명	64명

번호	국적	소속 기관	생년월일 (성별)	학위 및 전공			역할	신규채용구분 *(해당 시 작성)*	국가연구개발사업 인건비계상률 (%)[B]	전체 인건비계상률 [A+B,≦100%]
	성명	직위	국가연구자번호	최종 학위	전공	취득 년도	과제참여기간		**본 과제 인건비계상률 (%) [A]**	**국가연구개발사업 참여과제수*(건)** (A+B,≦5)
1	한국	광주과학기술원	66.11.20 (남)	박사	전자공학	1999	연구책임자	기존	65.18	95.18
	이홍노	교수	10512408				’21.07~’28.12		30	5
2	한국	광주과학기술원	88.06.14 (남)	학사	전기공학	2012	박사과정	기존	30	40
	강주성	통합(박사과정)	11322195				’21.07~’28.12		10	4
3	한국	광주과학기술원	90.12.30 (남)	석사	정보통신	2015	박사과정	기존	10	20
	최해웅	박사과정	11320849				’21.07~’28.12		10	3
4	한국	광주과학기술원	89.09.21	석사	정보통신	2016	박사과정	기존	60	70
	장재혁	박사과정	11427859				’21.07~’28.12		10	4
5	한국	광주과학기술원	93.03.15 (남)	학사	전기전자	2016	박사과정	기존	30	40
	김철순	통합(박사과정)	11580537				’21.07~’28.12		10	4

6	방글라데시	광주과학기술원	87.01.01 (남)	학사	전기전자공학	2009	박사과정	기존	30	40
	RAHMAN S M WAHIDUR	통합(박사과정)	12576865				’21.07~’28.12		10	4
7	한국	광주과학기술원	95.10.17 (여)	학사	컴퓨터공학	2020	석사과정	기존	30	40
	박하영	석사과정	12632239				’21.07~’28.12		10	4
8	한국	광주과학기술원	80.07.02 (여)	학사	유전공학	2002	연구원	기존	90	100
	정우정	위촉연구원	11102546				’21.07~’28.12		10	4
9	한국	광주과학기술원	85.06.22 (여)	학사	Human 생물학	2009	연구원	기존	90	100
	이명은	위촉연구원	12509002				’21.07~’28.12		10	4
10	한국	광주과학기술원	87.08.26 (여)	학사	독일어/영문학	2013	연구원	기존	90	100
	권효민	위촉연구원	12569625				’21.07~’28.12		10	4
11	인도	광주과학기술원	89.09.28 (남)	박사	전기전자컴퓨터	2020	포닥	기존	90	100
	Jabarulla, Mohamed Yaseen	Post-doc	11580512				’21.07~’28.12		10	4
12	채용예정						전임연구원	신규		
							’21.07~’28.12			
13	한국	광주과학기술원	64.06.04 (남)	박사	영상	1994	교수	기존	20	30
	김종원	교수	11195224				’21.07~’28.12		10	3
14	한국	광주과학기술원	70.12.25 (남)	박사	컴공	2004	교수	기존	70	80
	차병래	연구부교수	10123114				’21.07~’28.12		10	4
15	한국	광주과학기술원	72.04.01 (남)	박사	컴공	2007	교수	기존	70	80
	박선	연구부교수	10056048				’21.07~’28.12		10	4
16	한국	광주과학기술원	88.12.30 (남)	석사	정보통신	2016	박사과정	기존	50	60
	신준식	박사과정	11364882				’21.07~’28.12		10	3
17	한국	광주과학기술원	92.03.26 (남)	석사	정보통신	2015	박사과정	기존	50	60
	장한이	석박통합	11559535				’21.07~’28.12		10	3
18	파키스탄	광주과학기술원	80.09.09 (남)	석사	정보통신	2010	박사과정	기존	50	60
	Amard	박사과정	11739063				’21.07~’28.12		10	3
19	한국	광주과학기술원	98.09.09 (남)	학사	정보통신	2020	석사과정	기존	50	60
	구동환	석박통합	11754171				’21.07~’28.12		10	3
20	파키스탄	광주과학기술원	92.01.10 (여)	학사	컴공	2010	석사과정	기존	50	60
	Talaya	석박통합	12690217				’21.07~’28.12		10	3
21	우즈베키스탄	광주과학기술원	98.06.03 (남)	학사	컴공	2020	석사과정	기존	50	60
	Ulugbek	석박통합	12706400				’21.07~’28.12		10	3
22	우즈베키스탄	광주과학기술원	93.01.10 (남)	학사	컴공	2020	석사과정	기존	50	60
	Anvarjon	석박통합	12707253				’21.07~’28.12		10	3

23	한국	광주과학기술원	98.09.04 (남)	학사	컴공	2021	석사과정	기존	50	60
	김호	석박통합	12467210				’21.07~’28.12		10	3
24	한국	광주과학기술원	79.05.21 (남)	박사	전기및컴퓨터공학	2008	연구책임자	기존	50	60
	신종원	부교수	10190181				’21.07~’28.12		10	4
25	한국	광주과학기술원	89.01.24 (남)	석사	정보통신공학	2014	박사과정	기존	25	35
	변재욱	박사과정	11427802				’21.07~’28.12		10	3
26	한국	광주과학기술원	90.01.17 (남)	석사	정보통신공학	2016	박사과정	기존	20	30
	황수중	박사과정	11320038				’21.07~’28.12		10	3
27	한국	광주과학기술원	91.05.10 (남)	학사	전자공학	2014	박사과정	기존	20	30
	김한솔	박사과정	11320046				’21.07~’28.12		10	3
28	한국	광주과학기술원	93.06.03 (남)	석사	전기전자컴퓨터공학	2018	박사과정	기존	20	30
	김민승	박사과정	11563848				’21.07~’28.12		10	3
29	한국	광주과학기술원	94.03.09 (남)	학사	전자전기컴퓨터공학	2017	박사과정	기존	20	30
	이은균	박사과정	11710207				’21.07~’28.12		10	3
30	한국	광주과학기술원	95.05.16 (남)	학사	메카트로닉스공학	2018	박사과정	기존	20	30
	안영도	박사과정	11685507				’21.07~’28.12		10	3
31	한국	광주과학기술원	93.10.27 (남)	석사	전기전자컴퓨터공학	2020	박사과정	기존	20	30
	한상욱	박사과정	11665250				’21.07~’28.12		10	3
32	한국	광주과학기술원	92.08.16 (남)	학사	정보통신공학	2016	박사과정	기존	20	30
	오진우	석사과정	12540053				’21.07~’28.12		10	3
33	한국	광주과학기술원	93.01.16	학사	전기전자공학	2018	석사과정	기존	30	40
	박정원	석사과정	12624417				’’21.07~’28.12		10	4
34	한국	광주과학기술원	98.10.11	학사	전기전자컴퓨터공학	2021	석사과정	기존	10	20
	최영원	석사과정	12589084				’21.07~’28.12		10	2
35	한국	광주과학기술원	76.01.15 (남)	박사	전기컴퓨터	2011	교수	기존	66.81	76.81
	황의석	부교수	11426351				’21.07~’28.12		10	5
36	한국	광주과학기술원	91.12.07 (남)	학사	전기공학	2014	박사과정	기존	30	50
	윤승욱	박사과정	11499206				’21.07~’28.12		20	3
37	한국	광주과학기술원	91.04.17 (남)	학사	전자공학	2015	박사과정	기존	30	50
	송준호	박사과정	11511703				’21.07~’28.12		20	3
38	한국	광주과학기술원	95.09.11 (남)	석사	전기전자컴퓨터	2021	박사과정	기존	30	50
	한승남	박사과정	11773517				’21.07~’28.12		20	2
39	한국	광주과학기술원	96.11.20 (남)	석사	전기전자컴퓨터	2021	박사과정	기존	10	30
	서기업	박사과정	11871037				’21.07~’28.12		20	2
40	한국	광주과학기술원	96.09.06 (남)	학사	기계공학	2019	박사과정	기존	10	30
	김동주	박사과정	12489350				’21.07~’28.12		20	2
41	한국	광주과학	96.10.04	학사	기계	2019	석사과정	기존	20	50

		기술원	(여)		공학					
	김명선	석사과정	11877855				’21.07~’28.12		30	3
42	한국	광주과학기술원	93.12.28 (남)	학사	기계정보공학	2020	석사과정	기존	20	50
	정희훈	석사과정	12598871				’21.07~’28.12		30	3
43	한국	고려대학교	71.02.05 (남)	박사	컴퓨터공학	2004	교수	기존	40	50
	김황남	교수	10138962				’21.07~’28.12		10	3
44	한국	고려대학교	89.01.07 (남)	학사	컴퓨터공학	2012	박사과정	기존	28	43
	이동규	박사과정	11314082				’21.07~’28.12		15	3
45	한국	고려대학교	93.04.15 (남)	학사	항공우주공학	2019	석사과정	기존	36	51
	이재승	석사과정	12437516				’21.07~’28.12		15	3
46	한국	고려대학교	95.02.23 (남)	학사	전자공학	2020	석사과정	기존	41	56
	이상민	석사과정	12552416				’21.07~’28.12		15	3
47	한국	고려대학교	92.04.20 (남)	학사	전자전기컴퓨터공학	2017	석사과정	기존	42	57
	주현태	석박통합	12552419				’21.07~’28.12		15	3
48	한국	고려대학교	94.09.21 (남)	학사	전자공학	2020	석사과정	기존	13	28
	유태훈	석사과정	12655703				’21.07~’28.12		15	2
49	한국	고려대학교	92.05.28 (남)	학사	전자공학	2020	석사과정	기존	61	76
	김기석	석사과정	12728861				’21.07~’28.12		15	2
50	한국	고려대학교	93.03.09 (남)	학사	전기전자공학	2021	석사과정	기존	33	48
	윤남경	석박통합	12635795				’21.07~’28.12		15	2
51	한국	고려대학교	97.01.20 (남)	학사	전자통신공학	2021	석사과정	기존	33	48
	박창민	석사과정	12728315				’21.07~’28.12		15	2
52	한국	고려대학교	96.01.04 (여)	학사	전자공학	2021	석사과정	기존	33	48
	박소연	석박통합	12728310				’21.07~’28.12		15	3
53	한국	서울대학교	65.02.27 (남)	박사	컴퓨터공학	1993	교수	기존	75	85
	문수묵	교수	10105773				’21.07~’28.12		10	4
54	한국	서울대학교	96.08.20 (남)	학사	전기정보공학	2019	박사과정	기존	55	75
	이준모	박사과정	11887123				’21.07~’28.12		20	3
55	한국	서울대학교	95.03.27 (남)	학사	컴퓨터공학	2019	박사과정	기존	55	75
	박상현	박사과정	11750525				’21.07~’28.12		20	3
56	한국	서울대학교	99.08.28 (남)	학사	전기정보공학	2021	석사과정	기존	10	20
	김민재	석사과정	12522847				’21.07~’28.12		10	2
57	한국	서울대학교	98.07.09 (남)	학사	전기및전자공학	2021	석사과정	기존	10	20
	김준호	석사과정	12711489				’21.07~’28.12		10	2
58	한국	서울대학교	95.09.30 (남)	학사	전기전자공학	2021	석사과정	기존	10	20
	조중하	석사과정	12711218				’21.07~’28.12		10	2
59	한국	아주대학교	67.02.06 (남)	박사	전자공학	1996	교수	기존	70	80
	김기형	교수	10057424				’21.07~’28.12		10	4
60	한국	아주대학교	79.10.25 (남)	석사	정보통신	2008	박사과정	기존	45	75

	신운섭	박사과정	11101013				’21.07~’28.12		30	4
61	한국	아주대학교	87.07.29 (남)	학사	컴퓨터공학	2010	박사과정	기존	55	85
	황동엽	박사과정	11234430				’21.07~’28.12		30	4
62	한국	아주대학교	92.02.20 (남)	학사	컴퓨터공학	2018	석사과정	기존	90	100
	김경훈	석사과정	12580625				’21.07~’28.12		10	4
63	한국	아주대학교	93.07.14 (남)	학사	공상관리	2018	석사과정	기존	90	100
	이민형	석사과정	12581669				’21.07~’28.12		10	4
64	한국	아주대학교	94.07.11 (남)	학사	사이버보안	2020	석사과정	기존	65	95
	임승주	석사과정	12513549				’21.07~’28.12		30	4

* 위탁연구개발과제는 참여과제수에서 제외

나. 연구근접지원인력 현황

성명	국적	소속 기관	직위	학위 및 전공			담당 역할	신규채용 구분	시간 선택제 근무 구분	지원연도				총 지원기간
				최종 학위	전공	취득 년도				1단계		2~3단계		
										1년	2년	3~6년	7~8년	
채용예정							과제 수행지원	신규		0	0	0	0	7년 6개월

⑨ 연구개발기관이 아닌 관계 기관

가. 기관명 : 광주광역시청 (연구개발비 부담하나 사용하지 않음)

책임자	성명	국문	이정선	국적	대한민국
		영문			
	기관명		광주광역시청	전화번호	[redacted]
	부서		스마트시티과	휴대전화	[redacted]
	직위		주무관	전자우편	[redacted].kr
주소			(우: 61945) 광주 서구 내방로 111 (치평동)		

※ 광주광역시 시비 2억원 예산 지원 부분 증빙자료는 요청 시, 시에서 발급받아 제출할 예정

나. 기관명: 한국ICT융합협회 (역할: 자문・협력)

책임자	성명	국문	백양순	국적	대한민국
		영문	Baek, Yang-Sun		
	기관명		한국ICT융합협회	전화번호	[redacted]
	부서		-	휴대전화	[redacted]
	[redacted]		[redacted]	[redacted]	[redacted]
주소			(우: 06130) 서울 강남구 강남대로 94길 27-11 투명교정센터 3층		

다. 기관명: 서울블록체인지원센터 (역할: 자문・협력)

책임자	성명	국문	임명수	국적	대한민국
		영문	Im, Myungsu		
	기관명		서울블록체인지원센터	전화번호	[redacted]
	[redacted]		[redacted]	[redacted]	[redacted]
	[redacted]		[redacted]	[redacted]	[redacted]
주소			(우: 04168) 서울 마포구 마포대로 78 자람빌딩 8, 11층		

라. 기관명: 엘에스웨어(주) (역할: 자문・협력)

책임자	성명	국문	신동명	국적	대한민국
		영문	Shin, Dong myung		
	기관명		엘에스웨어(주)	전화번호	[redacted]
	[redacted]		[redacted]	[redacted]	[redacted]
	[redacted]		[redacted]	[redacted]	[redacted]
주소			(우: 08504) 서울 금천구 가산동 서부샛길 606		

마. 기관명: 퓨처센스 (역할: 자문・협력)

책임자	성명	국문	안다미	국적	대한민국
		영문	An, Dami		
	기관명		퓨처센스 주식회사	전화번호	
	부서		-	휴대전화	[redacted]
	[redacted]		[redacted]	[redacted]	[redacted]
주소			(우: 06083) 서울 강남구 영동대로602, 6층 삼성글로비즈 에스지비016호		

바. 기관명: 헥슬란트 (역할: 자문・협력)

책임자	성명	국문	노진우	국적	대한민국
		영문	Noh, Jinwoo		
	기관명		헥슬란트	전화번호	-
	부서		-	휴대전화	[redacted]
	[redacted]		[redacted]	[redacted]	[redacted]
주소			(우: 06242) 서울시 강남구 강남대로 340 경원빌딩 7층		

사. 기관명: 시티랩스 (역할: 자문・협력)

책임자	성명	국문	김항진	국적	대한민국
		영문	Kim, Hangjin Phillip		
	기관명		㈜시티랩스	전화번호	[redacted]
	[redacted]		[redacted]	[redacted]	[redacted]
	[redacted]		[redacted]	[redacted]	[redacted]
주소			(우: 13911) 경기도 안양시 만안구 예술공원로 153-32 3층		

2) 연구개발기관 일반 현황(기업정보 데이터베이스와 연계 가능)

(단위: 천원, 백분율)

순번	구분 \ 기관명		광주과학기술원	서울대학교 산학협력단	고려대학교	아주대학교 산학협력단
1	사업자등록번호		410-82-07550	119-82-03684	209-82-00433	124-82-14602
2	법인등록번호		204271-0000871	114371-0009224	114431-0000084	135871-0005871
3	대표자 성명/국적		총장직무대행 김인수/대한민국	최해천/대한민국	정진택/대한민국	권용진/대한민국
4	기관 유형 (대학, 정부출연연, 중소기업 등)		특정연구기관	대학	대학	대학
5	최대 주주 성명/국적		-	-	-	-
6	설립 연월일		1993년11월17일	-	1905년.04월.03일	2004년05월25일
7	주생산 품목		-	-	-	-
8	상시 종업원 수		-	-	-	-
9	전년도 매출액		-	-	-	-
10	매출액 대비 연구개발비 비율		-	-	-	-
11	부채 비율 (최근 3년 간 결산 기준)	yy년	-	-	-	-
		yy년	-	-	-	-
		yy년	-	-	-	-
12	유동 비율 (최근 3년 간 결산 기준)	yy년	-	-	-	-
		yy년	-	-	-	-
		yy년	-	-	-	-
13	자본잠식 현황 (최근 3년 간 결산 기준)	자본 총계 yy년	-	-	-	-
		자본 총계 yy년	-	-	-	-
		자본 총계 yy년	-	-	-	-
		자본금 yy년	-	-	-	-
		자본금 yy년	-	-	-	-
		자본금 yy년	-	-	-	-
14	이자 보상 비율 (최근 3년 간 결산 기준)	yy년	-	-	-	-
		yy년	-	-	-	-
		yy년	-	-	-	-
15	영업 이익 (최근 3년 간 결산 기준)	yy년	-	-	-	-
		yy년	-	-	-	-
		yy년	-	-	-	-
16	연구개발기관의 연구개발과제 지원 담당자 *(※ 대학의 경우 산학협력단의 연구개발과제 지원 담당을 말하며, 표지의 "실무담당자"와 다름)*	성명	김근영	-	-	-
		부서	연구관리팀	-	-	-
		직위	담당	-	-	-
		직장전화	[redacted]	[redacted]	[redacted]	[redacted]
		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		[redacted]	[redacted]	-	-	-

9 연구개발비 사용에 관한 계획

1) 연구개발비 지원·부담계획

(단위 : 천원)

구분			정부지원 연구개발비	기관부담 연구개발비				그 외 기관 등의 지원금						합 계		
								지방자치단체			기타( )					
단계	연차	연구개발기관명 (기관역할[1])	현금	현금	현물	소계	비고[2]	현금	현물	소계	현금	현물	소계	현금	현물	합계
1	1	*광주과학기술원 (주관)*	[redacted]													
	2	*광주과학기술원 (주관)*														
	소계															
2	1	*광주과학기술원 (주관)*														
	2	*광주과학기술원 (주관)*														
	3	*광주과학기술원 (주관)*														
	4	*광주과학기술원 (주관)*														
	소계															
3	1	*광주과학기술원 (주관)*														
	2	*광주과학기술원 (주관)*														
	소계															
총계																

2) 연구개발비 사용계획

① 연구개발기관별 사용계획

(단위: 천원)

연구개발 기관명		연구개발비													미 포함예 산[5]	미 지급인 건비[6]
		직접비											간접비	합계		
		인건비	학생인건비		연구시설·장비비		연구재료비	위탁연구개발비	국제공동연구개발비	연구개발부담비	연구활동비	연구수당	소계			
			일반[1]	특례[2]	일반[3]	특례[4]										
광주과학기술원 (주관)	현금	[redacted]														
	현물															
	소계															
총계	현금															
	현물															
	합계															

② 연차별 사용계획

(단위: 천원)

연차		연구개발비													미포함 예산	미지급 인건비
		직접비											간접비	합계		
		인건비	학생인건비		연구시설·장비비		연구재료비	위탁연구개발비	국제공동연구개발비	연구개발부담비	연구활동비	연구수당	소계			
			일반	특례	일반	특례										
1년차 (2021년도)	현금															
	현물															
	소계															
2년차 (2022년도)	현금															
	현물															
	소계															
3년차 (2023년도)	현금															
	현물															
	소계															
4년차 (2024년도)	현금															
	현물															
	소계															
5년차 (2025년도)	현금															
	현물															
	소계															
6년차 (2026년도)	현금															
	현물															
	소계															
7년차 (2027년도)	현금															
	현물															
	소계															
8년차 (2028년도)	현금															
	현물															
	소계															
총계	현금															
	현물															
	합계															

③ 연구개발기관별-연차별 사용계획

가. 주관연구개발기관명: 광주과학기술원

(단위: 천원)

연차		연구개발비												미포함 예산	미지급 인건비
		직접비										간접비	합계		
		인건비	학생인건비		연구시설·장비비		연구재료비	위탁연구개발비	국제공동연구개발비	연구개발부담비	연구활동비	연구수당	소계				
			일반	특례	일반	특례											
1년차 (2021년도)	현금														
	현물														
	소계														
2년차 (2022년도)	현금														
	현물														
	소계														
3년차 (2023년도)	현금														
	현물														
	소계														
4년차 (2024년도)	현금														
	현물														
	소계														
5년차 (2025년도)	현금														
	현물														
	소계														
6년차 (2026년도)	현금														
	현물														
	소계														
7년차 (2027년도)	현금														
	현물														
	소계														
8년차 (2028년도)	현금														
	현물														
	소계														
총계	현금														
	현물														
	합계														

3) 연구시설・장비 구축・운영계획

① 연구시설・장비 구축계획

(단위: 천원)

연구개발기관명	연구시설・ 장비명	기존/ 신규 구분	현금/ 현물 구분	구축방식*	규격	수량	구축비용	구축기간	설치장소
광주과학기술원	알고리즘 개발용 PC	신규	현금	구매	SET	2	5,000	*2021*	광주과기원 전컴C-315,317
광주과학기술원	GPU 서버	신규	현금	구매	서버	1	10,000	*2022*	
광주과학기술원	GPU 서버	신규	현금	구매	서버	1	10,000	*2023*	
광주과학기술원	GPU 서버	신규	현금	구매	서버	1	10,000	*2024*	
광주과학기술원	GPU 서버	신규	현금	구매	서버	1	10,000	*2025*	
광주과학기술원	GPU 서버	신규	현금	구매	서버	1	10,000	*2026*	
광주과학기술원	GPU 서버	신규	현금	구매	서버	1	10,000	*2027*	
광주과학기술원	GPU 서버	신규	현금	구매	서버	1	10,000	*2028*	
광주과학기술원	데이터전송 박스	신규	현금	구매	미니서버	1	5,000	*2021*	광주과기원 전컴C-212
광주과학기술원	데이터레이크 박스 (증설)	신규	현금	구매	서버	1	10,000	*2022*	
광주과학기술원	데이터레이크 박스 (증설)	신규	현금	구매	서버	1	10,000	*2023*	
광주과학기술원	데이터레이크 박스 (증설)	신규	현금	구매	서버	1	10,000	*2024*	
광주과학기술원	데이터레이크 박스 (증설)	신규	현금	구매	서버	1	10,000	*2025*	
광주과학기술원	데이터레이크 박스 (증설)	신규	현금	구매	서버	1	10,000	*2026*	
광주과학기술원	데이터레이크 박스 (증설)	신규	현금	구매	서버	1	10,000	*2027*	
광주과학기술원	데이터레이크 박스 (증설)	신규	현금	구매	서버	1	10,000	*2028*	
광주과학기술원	DNN 알고리즘 개발용 노트북	신규	현금	구매	SET	1	1,500	*2021*	광주과기원 전컴C-516
광주과학기술원	DNN 알고리즘 개발용 노트북	신규	현금	구매	SET	1	1,500	*2022*	
광주과학기술원	DNN 알고리즘 개발용 노트북	신규	현금	구매	SET	1	1,500	*2023*	
광주과학기술원	DNN 알고리즘 개발용 노트북	신규	현금	구매	SET	1	1,500	*2024*	
광주과학기술원	DNN 알고리즘 개발용 노트북	신규	현금	구매	SET	1	1,500	*2025*	
광주과학기술원	DNN 알고리즘 개발용 노트북	신규	현금	구매	SET	1	1,500	*2026*	
광주과학기술원	DNN 알고리즘 개발용 노트북	신규	현금	구매	SET	1	1,500	*2027*	
광주과학기술원	DNN 알고리즘 개발용 노트북	신규	현금	구매	SET	1	1,500	*2028*	
광주과학기술원	MEMS 기반 3차원 레이저 센서	기존	현물	-	EA	1	6,655	*2017*	광주과기원 전컴C-317
광주과학기술원	DGX-A100	기존	현물	-	서버	2	230,043	2020	광주과기원 전컴C-108
광주과학기술원	DGX-1V	기존	현물	-	서버	2	124,586	2019	
광주과학기술원	Pure Storage	기존	현물	-	EA	1	99,550	2020	
광주과학기술원	Mellanox QM 8700	기존	현물	-	EA	1	7,730	2020	
광주과학기술원	SuperServer 1029U-TN10RT	기존	현물	-	서버	2	7,200	2020	
광주과학기술원	Susol SS 2029BZ-HNR	기존	현물	-	서버	1	27,200	2020	
광주과학기술원	APC smart-UPS	기존	현물	-	EA	2	8,400	2020	

② 연구시설 운영・활용계획

연구개발기관명	연구시설명	기존/신규 구분	운영기간	연간운영 비용(천원)	전담인력 수	활용계획	설치장소
광주과학기술원	개발용 PC	신규	21-28	-	2	알고리즘 개발	광주과기원 전컴C-315,317
광주과학기술원	GPU 서버	신규	22-28	-	2	블록체인개발 및 시뮬레이션 실증용	광주과기원 전컴C-315,317
광주과학기술원	데이터전송 박스	신규	21-28	-	-	-	광주과기원 전컴C-212
광주과학기술원	데이터레이크 박스 (증설)	신규	22-28	-	-	-	광주과기원 전컴C-212
광주과학기술원	DNN 알고리즘 개발용 노트북	신규	21-28	-	1	DNN 알고리즘 개발	광주과기원 전컴C-516
광주과학기술원	MEMS 기반 3차원 레이저 센서 및 측정 장비	기존	21-28	-	1	IOT 센서 Data 생성 - LiDAR 신호	광주과기원 전컴C-317
광주과학기술원	DGX-A100	기존	21-28	-	4	고성능 알고리즘 학습 - GPU processing	광주과기원 전컴C-108
광주과학기술원	DGX-1V	기존	21-28	-	4	고성능 알고리즘 학습 - GPU processing	광주과기원 전컴C-108
광주과학기술원	Pure Storage	기존	21-28	-	2	디지털 트윈 데이터 저장용	광주과기원 전컴C-108
광주과학기술원	Mellanox QM 8700	기존	21-28	-	1	신뢰성 있는 데이터 송/수신용	광주과기원 전컴C-108
광주과학기술원	SuperServer 1029U-TN10RT	기존	21-28	-	2	고성능 알고리즘 학습 - CPU processing	광주과기원 전컴C-108
광주과학기술원	Susol SS 2029BZ-HNR	기존	21-28	-	2	고성능 알고리즘 학습 - CPU processing	광주과기원 전컴C-108
광주과학기술원	APC smart-UPS	기존	21-28	-	1	서버 신뢰성 확보용	광주과기원 전컴C-108

10 평가기준 및 평가방법

1) 연차별 자율성과 목표

○ 목표치

성과지표			1년차	2년차	3년차	4년차	5년차	6년차	7년차	8년차
기술이전 수입료(천만원)										
특허 등록 (건)	국내	SMART값 (A ~ AAA)		2	5	5	5	5	5	5
		SMART값 (B ~ BBB)								
		SMART값 (C ~ CCC)								
	국제	주요국		1	1	2	2	2	2	2
		주요국외								
특허 출원 (건)	국내			5	5	5	5	5	5	5
	국제			4	4	4	4	4	4	4
표준화 채택 (건)	국내									
	국제									
표준화 승인 (건)	국내									
	국제									
학생논문 IF(평균)				2.5	2.5	2.5	2.5	2.5	2.5	2.5
학생 창의자율과제(건)				5	5	5	5	5	5	5
자체 지표	학생지도 점수	석사 입학		12	12	12	12	12	12	12
		박사 입학		4	4	4	4	4	4	4
		석사 배출		12	12	12	12	12	12	12
		박사 배출		4	4	4	4	4	4	4
	교과목 개설 지표	신규		1	0	0	0	0	0	0
		기존		2	3	3	3	3	3	3

* 특허 주요국 : 미국, 유럽, 일본

○ 점수환산표

성과지표			1년차	2년차	3년차	4년차	5년차	6년차	7년차	8년차
기술이전 수입료			-							
특허 등록	국내	SMART값 (A ~ AAA)	-	2건 (200)	5건 (500)	5건 (500)	5건 (500)	5건 (500)	5건 (500)	5건 (500)
		SMART값 (B ~ BBB)	-							
		SMART값 (C ~ CCC)	-							
	국제	주요국	-	1건 (500)	1건 (500)	2건 (1000)	2건 (1000)	2건 (1000)	2건 (1000)	2건 (1000)
		주요국외	-							
특허 출원	국내		-	5건 (50)	5건 (50)	5건 (50)	5건 (50)	5건 (50)	5건 (50)	5건 (50)
	국제		-	4건 (200)	4건 (200)	4건 (200)	4건 (200)	4건 (200)	4건 (200)	4건 (200)
표준화 채택	국내		-							
	국제		-							
표준화 승인	국내		-							
	국제		-							
학생논문 IF			-	평균 2.5 (500)	평균 2.5 (500)	평균 2.5 (500)	평균 2.5 (500)	평균 2.5 (500)	평균 2.5 (500)	평균 2.5 (500)
학생 창의자율과제			-	5건 (500)	5건 (500)	5건 (500)	5건 (500)	5건 (500)	5건 (500)	5건 (500)
자체 지표	학생지도 지표		-	224	224	224	224	224	224	224
	교과목 개설 지표		-	50	30	30	30	30	30	30
목표점수			-	2200	2500	3000	3000	3000	3000	3000

* 학생지도 지표: 석사생 일인 참여시 4점. 박사생 일인 참여시 20점. 석사 일인 배출시 2점. 박사생 일인 배출시 12점.

* 교과목 개설 지표: 블록체인 지능 융합 관련 교과목 신설 시 30점, 기존 교과목 개설 시 15점. 두 지표의 합이 최고점은 250점.

2) 그 외 주요 성과목표

구 분	성과지표			1년차	2년차	3년차	4년차	5년차	6년차	7년차	8년차
과학 기술적 성과	3극 특허 수			0	1	0	0	1	1	0	1
	논문	SCI(E)	국내	0	0	0	0	0	0	0	0
			국외	7	14	14	14	14	14	14	14
		비SCI	국내	3	7	7	7	7	7	7	7
			국외	7	14	14	14	14	14	14	14
		JCR 상위 10% 논문 수(SCI)		2	4	4	4	4	4	4	4
	SW등록(건)			0	7	0	0	0	7	0	7
인력 양성 성과	수혜인원(명)		석사	24	24	24	24	24	24	24	24
			박사	26	26	26	26	26	26	26	26
			합계	50	50	50	50	50	50	50	50
	배출인원(명)		석사	6	12	12	12	12	12	12	12
			박사	2	4	4	4	4	4	4	4
			합계	8	16	16	16	16	16	16	16
	취업률(%)			100%	100%	100%	100%	100%	100%	100%	100%
	창업기업 수		학생	0	1	0	1	0	1	0	1
			교수	0	1	0	1	0	1	0	1
			합계	0	2	0	2	0	2	0	2
	기업가정신교육인원			4	10	10	10	10	10	10	10
	기타(추가가능)										

3) 기타

- 해당 사항 없음 -

부록

[붙임1] 연구기술 요약문

[붙임2] 연구책임자 대표 연구실적 요약문

[붙임3] 세부 프로젝트 책임자 대표 연구실적 요약문

[붙임4] 공동연구개발기관 보유기술 현황

[붙임5] 연구기반(인프라) 현황

[붙임6] 사업 전용공간 확보계획서

[붙임7] 기관부담연구개발비 출자 납입확약서

[붙임8] 참여기관 현황(총괄표)

[붙임9] 전임연구원 채용 확인서

[붙임10] 신규 참여연구원 채용 확인서

[붙임11] 시설장비 심의요청서

[붙임1]

연구기술 요약문

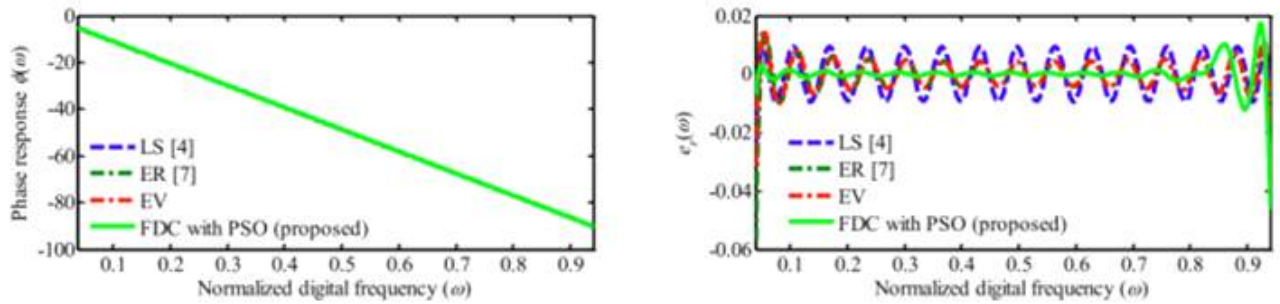
산 업
4차 산업혁명의 핵심기술인 인공지능과 블록체인의 관계는 혁신과 신뢰, 성장과 분배라고 볼수 있음. 인공지능 발달에 있어 데이터 신뢰성 문제가 발생하면 학습 알고리즘이 훌륭해도 의미가 퇴색함. 블록체인은 모든 거래와 정책, 의사 표현을 이해당사자와 참여자 모두에게 공개하면서도 변조나 해킹이 불가능한 기술이기에 초신뢰 경제혁명을 불러올 뿐만 아니라 신뢰성을 획기적으로 신장시킬 역량을 보유함. 즉, 블록체인 기술은 이해관계자 간 민주적 절차를 통해 이용자가 함께 플랫폼 규칙을 설정할 수 있어 개방성과 투명성을 확보할 수 있으며 참여자에게 공정한 분배시스템을 보장함으로써 네트워크 확장과 경제적 성장을 함께 얻을 수 있음.
요소기술
1. 영지식 센싱 탈중앙 고성능 분산원장 기술 개발 - 블록체인 합의 메커니즘 - 블록체인의 기반 연합학습 - 화자인증 기술 - 생체인증을 위한 초분광 이미징 2. Edge 컴퓨팅 기반 블록체인 서비스 인프라 기술 - Cloud-native 컴퓨팅에 기반한 상호호환적인 에지 컴퓨팅 - 인프라와 이에 기반한 블록체인 대응 데이터 수집·정제 및 공유 기능의 필요성 - 블록체인을 통한 신뢰할 수 있는 데이터 스트림 제공 기술 - 중앙화된 네트워크 컨트롤러의 장애에 대한 해결 기술 3. DID 블록체인 기반 IoT 인증 및 보안 기술 개발 - IoT 디바이스 생체인증의 보안 - IoT 디바이스 보안방법 개발 필요
기대효과
블록체인 기반기술 개발을 통하여 다음과 같은 효과를 기대할 수 있음. (1) 기술적 측면 o DID 기반기술, 초경량 · 초확장 블록체인 기술, 대용량 데이터 저장/공유/관리 기술, 신뢰정의 네트워크 기술 확보 (2) 경제적, 산업적 측면 o 블록체인 데이터 저장 원천 기술 확보로 누구나 쉽게 블록체인 데이터 기반 융합 서비스 제공 역량 향상, 산업간 융합 IT 서비스 개발을 활성화하는 등 데이터 경제 활성화를 통해 4차산업혁명 견인 (3) 사회적 측면 o 본 과제의 DID 기반 인증과 영지식 증명 기술로 사생활 민감정보, 의료정보, 생체인증 등에 대한 인증/증명 시스템에서 개인정보 보호능력 제고

[붙임2]

‘연구책임자(이홍노)’ 대표 연구실적 요약문

연구실적 제목	Hilbert Transform Design based on Fractional Derivatives and Swarm Optimization
연구실적 유형	논문(○) 특허( ) 기술이전( ) 표준화( )

연구에서는 Fractional derivative와 군집 최적화를 기반으로 모든 패스 필터를 사용하여 힐베르트 변환을 구현하는 새롭고 효율적인 방법을 제안했다. 제안하는 방법을 통해 설계된 필터와 실제 구현된 필터의 차이가 줄었다. FD 방법은 레퍼런스 주파수의 높은 정확도를 달성하기 위해 이용되었다. 또한 군집 최적화 방식을 통해 FD 와 레퍼런스 주파수에 적합한 값을 찾아 낼 수 있었다. 군집 최적화 방식을 통해 얻어낸 FD 와 레퍼런스 주파수를 통해, Phase 에러, group delay error, total squared group delay error 등을 줄일 수 있었다.



힐베르트 변환의 위상 응답 그래프 (좌), 오차 비교 그래프 (우): 제안하는 방식이 가장 적은 에러를 보임

Hilbert Transform Design Based on Fractional Derivatives and Swarm Optimization

Anil Kumar, *Member, IEEE*, Nikhil Agrawal, *Student Member, IEEE*, Ila Sharma, *Student Member, IEEE*, Seungchan Lee, and Heung-No Lee, *Senior Member, IEEE*

***Abstract*—This paper presents a new efficient method for implementing the Hilbert transform using an all-pass filter, based on fractional derivatives (FDs) and swarm optimization. In the proposed method, the squared error difference between the desired and designed responses of a filter is minimized. FDs are introduced to achieve higher accuracy at the reference frequency (ω_θ), which helps to reduce the overall phase error. In this paper, two approaches are used for finding the appropriate values of the FDs and reference frequencies. In the first approach, these values are estimated from a series of experiments, which require more computation time but produce less accurate results. These experiments, however, justify the behavior of the error function, with respect to the FD and ω_θ, as a multimodal and nonconvex problem. In the second approach, a variant of the swarm-intelligence-based multimodal search space technique, known as the constraint-factor particle swarm optimization, is exploited for finding the suitable values for the FD and ω_θ. The performance of the proposed FD-based method is measured in terms of fidelity aspects, such as the maximum phase error, total squared phase error, maximum group delay error, and total squared group delay error. The FD-based approach is found to reduce the total phase error by 57% by exploiting only two FDs.**

***Index Terms*—Evolutionary technique (ET), fractional derivative (FD), Hilbert transform (HT), particle swarm optimization (PSO).**

I. Introduction

HILBERT transform (HT) is a very important transform in signal processing. It can be used to represent a narrow-band signal in terms of its frequency-domain amplitude at a frequency-modulation point. HT is useful in many applications such as latency analysis in neurophysiological signals [1], fault classification in electrical systems, data compression in communication, and as an data analysis tool [2], [3]. Research on efficient HT design methods has been progressing over the last three decades. In the early stages of research, several works used infinite impulse response (IIR) filters, which satisfied the magnitude and phase specifications simultaneously [4]–[8]. In these techniques, the design problem was formulated as a minimization of the mean squared error between the desired and designed responses, and the solution was obtained by solving a set of linear equations [5]–[8]. Squared error-based methods usually produced a matrix, which was symmetric and positive definite, and the solutions of such problems were computed by either evaluating the eigen vector (EV) corresponding to the least value of the eigen value or multiplying the matrix inverse with the multiplicand. The EV approach was computationally complex, whereas the second approach had $o(n^3)$ complexity. However, this complexity could be reduced to $o(n^2)$ using the Cholesky decomposition or split Levinson algorithms [9]. Liu [1] and Kidambi [7] used a Toeplitz-plus-Hankel matrix to solve a set of linear equations. For reducing the computational complexity further, Su *et al.* [9] proposed a closed-form method for the design of the HT. Literature review on the implementation of HT using all-pass filters (APFs) has corroborated that several methods have been proposed [4]–[8]. However, an APF with smaller number of filter taps and higher accuracy in terms of the degree of approximation, which can produce a desired response at a certain reference frequency point, has not been considered yet for the efficient design of an HT.

Recently, fractional derivatives (FDs) have been used in several engineering problems, such as fractional system identification, edge detection in image processing, electrocardiogram signal R-peak detection, and numerous other engineering applications, owing to their ease of realization and higher efficiency compared to integral-order derivatives [10]–[15]. Several researchers have used FDs for digital filter design [13]–[15]. In these techniques, the determination of optimal orders of FDs was computationally expensive, and the work involved increases with the increase in the order of FD. A possible solution for this problem is to use evolutionary techniques (ETs) for filter-bank design, which is explained in [15]. However, a mechanism to tune the suitable values of the reference frequency in the region of interest, where these FDs are being evaluated, has not been established yet.

Manuscript received September 13, 2018; revised September 26, 2018; accepted October 8, 2018. Date of publication October 29, 2018; date of current version April 15, 2020. This work was supported by the National Research Foundation of Korea (NRF) funded by the Korean Government (MSIP) under Grant NRF-2018R1A2A1A19018665. This paper was recommended by Associate Editor Q. Zhang. *(Corresponding author: Heung-No Lee.)*

A. Kumar is with the Department of Electronics and Communication Engineering, PDPM Indian Institute of Information Technology, Design and Manufacturing Jabalpur, Jabalpur 482005, India, and also with the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea (e-mail: anilkdee@gmail.com).

N. Agrawal and I. Sharma are with the Department of Electronics and Communication Engineering, PDPM Indian Institute of Information Technology, Design and Manufacturing Jabalpur, Jabalpur 482005, India (e-mail: nikhil.agrawal@iiitdmj.ac.in; ilasharma23@gmail.com).

S. Lee and H.-N. Lee are with the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea (e-mail: heungno@gist.ac.kr; seungchan@gist.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCYB.2018.2875540

2168-2267 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on November 03,2020 at 06:40:29 UTC from IEEE Xplore. Restrictions apply.

연구실적 제목	ECCPoW: Error-Correction Code based Proof-of-Work for ASIC Resistance
연구실적 유형	논문(○) 특허( ) 기술이전( ) 표준화( )

비트코인의 채굴은 작업증명(PoW) 에 의해서 이루어진다. 하지만 PoW는 높은 해시 파워를 요구하므로 채굴만을 위한 application-specific integrated circuit (ASIC)을 가진 마이닝 풀과 같은 채굴 연합체가 비트코인 채굴을 독점하게 되는 문제가 있다. 본 연구팀은 이러한 문제를 해결할 수 있도록 ASIC을 사용한 채굴에 저항성을 가지는 새로운 작업증명 방법인 ECCPoW를 제안하였다. 제안된 ECCPoW는 오류정정부호 중 한 종류인 low density parity check (LDPC) 코드 및 해쉬함수를 결합한 방법이다. 본 연구팀은 ECCPoW를 비트코인에 적용하고 ECCPoW를 기존 비트코인의 PoW방법과 채굴중앙화, 보안성, 확장성 측면에서 비교하였다.

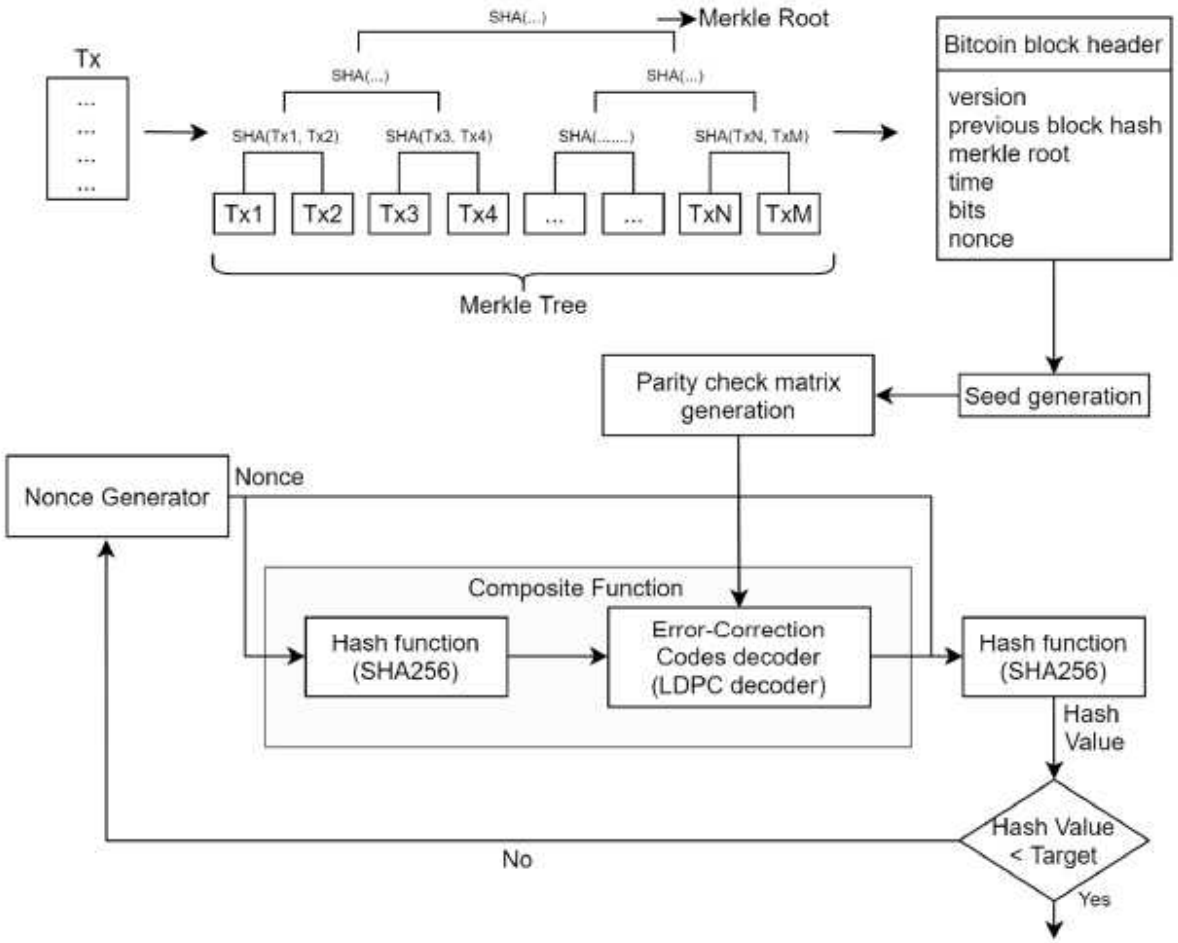


Figure 1. Condition 2 for determining the error-correction codes proof-of-work (ECCPoW) crypto puzzle resolution.

symmetry

MDPI

Article

ECCPoW: Error-Correction Code Based Proof-of-Work for ASIC Resistance

Hyunjun Jung [1] and Heung-No Lee [2,*]

[1] Blockchain Internet Economy Research Center, Gwangju Institute of Science and Technology, Gwangju 61005, Korea; junghj85@gist.ac.kr
[2] School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, Korea
* Correspondence: heungno@gist.ac.kr; Tel.: +82-62-715-2237

Received: 20 May 2020; Accepted: 5 June 2020; Published: 9 June 2020

Abstract: Bitcoin is the first cryptocurrency to participate in a network and receive compensation for online remittance and mining without any intervention from a third party, such as financial institutions. Bitcoin mining is done through proof of work (PoW). Given its characteristics, the higher hash rate results in a higher probability of mining, leading to the emergence of a mining pool, called a mining organization. Unlike central processing units or graphics processing units, high-cost application-specific integrated circuit miners have emerged with performance efficiency. The problem is that the obtained hash rate exposes Bitcoin's mining monopoly and causes the risk of a double-payment attack. To solve this problem, we propose the error-correction code PoW (ECCPoW), combining the low-density parity-check decoder and hash function. The ECCPoW contributes to the phenomenon of symmetry in the proof of work (PoW) blockchain. This paper proposes the implementation of ECCPoW, replacing the PoW in Bitcoin. Finally, we compare the mining centralization, security, and scalability of ECCPoW and Bitcoin.

Keywords: error-correction codes proof-of-work (ECCPoW); proof-of-work (PoW); ECCPoW implementation; ASIC resistance

1. Introduction

We use digital signatures from third-party trust agencies to promote trust in Internet commerce. We warrant proof of data forgery using middlemen. Satoshi Nakamoto proposed an electronic money system without a middleman in a peer-to-peer (P2P) network through a Bitcoin white paper [1]. Bitcoin applies blockchain technology to electronic monetary systems to guarantee transactions without intermediaries (e.g., banks). Blockchain is a ledger management technology based on a distributed computing technology that cannot be arbitrarily modified by storing the transaction content in a chain-based distributed data storage environment in the form of a block [2,3].

The blockchain stores the same ledger on a global network and is designed to pay certain rewards to maintain the block. This is called mining, and mining creates blocks and obtains cryptocurrency by executing a hash function. Miners belong to mining pools because of the probability and convenience of being rewarded for mining cryptocurrency [4,5].

The hash rate is the number of hash values calculated per second as a measure of computational processing power for mining cryptocurrency. The hash rate of cryptocurrency is determined by the total number of the participating nodes. Most miners belong to a hash pool and occupy a high percentage of the hash rate. We should be concerned about the risk of a double-payment attack if the mining pool in the blockchain accounts for a high percentage of the total hash rate of cryptocurrency [6]. A double-payment attack occurs when the mining pool seizes at least 51% of the total hash rate

6. Conclusions

In this paper, we explained ECCPoW and applied the proposed method to Bitcoin. This paper addressed the problem of centralization of mining due to the emergence of ASICs. We proposed a PoW concept based on error-correction codes to solve this problem. The core of the ECCPoW is the connection of the hash function with the LDPC decoder. Blockchain applied with the ECCPoW determines the completion of the POW using the output value of the decoder. In addition, ASIC suppression is possible because ASICs use the LDPC decoder.

This paper contributes to the phenomenon of symmetry in the PoW blockchain. The total block size of the PoW blockchain symmetrically influences the hash rate. The PoW blockchain increases stability as the hash rate increases in size. However, a high hash rate causes the waste of computing power in block generation. Thus, we mitigate the causes of a high hash rate using ASIC resistance.

The ECCPoW offers a method of solving different puzzles in each block to avoid ASICs. This maximizes the benefits of how existing studies use a limited number of hash functions to solve each block of different hash functions. The proposed method offers more effective connections and the use of multiple hash functions. We presented the difficulty control, parity-check matrix generation method, hash vector generation, and code determination methods for implementing ECCPoW. Furthermore, ECCPoW was applied to Bitcoin to verify the proposed method. We assessed the mining centralization, security, and scalability of ECCPoW and Bitcoin. We found that ECCPoW maintained security and scalability, showing 32% higher mining centralization than Bitcoin. Using the proposed method, ECCPoW does not require high hash rates, and miners can compete more fairly.

This study contributes to the previous literature on ASIC resistance in the PoW blockchain. The ASIC resistance study of the PoW blockchain was conducted in response to hardware development by ASIC manufacturers. One of the studies on ASIC resistance induced forced access to memory in the hash function, thereby lowering the performance of ASICs. Other studies have suggested periodically altering the hash function to render ASICs useless. The ECCPoW provides a method of releasing different hash functions for each block, rather than changing the periodic hash function.

The ECCPoW revealed the limit of block generation time in the experiment. For the stable operation of the blockchain, the block generation time of the ECCPoW must be stable and controllable. The results of the block generation time test were unstable in the experiment using one node. Thus, ECCPoW requires further research regarding the difficulty of block generation for stable operation. We also need to increase the number of nodes in the ECCPoW to carry out mining.

Author Contributions: H.J. and H.-N.L. contributed equally to this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean government (MSIP) (NRF-2018R1A2A1A19018665).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nakamoto, S. *Bitcoin: A Peer-to-Peer Electronic Cash System*; **2008**. Available online: https://bitcoin.org/bitcoin.pdf (accessed on June 9 2020)
2. Nofer, M.; Gomber, P.; Hinz, O.; Schiereck, D. Blockchain. *Bus. Inf. Syst. Eng.* **2017**, 59, 183–187, doi:10.1007/s12599-017-0467-3.
3. Saberi, S.; Kouhizadeh, M.; Sarkis, J.; Shen, L. Blockchain Technology and Its Relationships to Sustainable Supply Chain Management. *Int. J. Prod. Res.* **2019**, 57, 2117–2135, doi:10.1080/00207543.2018.1533261.
4. Qin, R.; Yuan, Y.; Wang, F.-Y. Research on the Selection Strategies of Blockchain Mining Pools. *IEEE Trans. Comput. Soc. Syst.* **2018**, 5, 748–757, doi:10.1109/TCSS.2018.2861423.
5. Liu, X.; Wang, W.; Niyato, D.; Zhao, N.; Wang, P. Evolutionary Game for Mining Pool Selection in Blockchain Networks. *IEEE Wirel. Commun. Lett.* **2018**, 7, 760–763, doi:10.1109/LWC.2018.2820009.
6. Karame, G.O.; Androulaki, E.; Capkun, S. Double-Spending Fast Payments in Bitcoin. In Proceedings of the 2012 ACM Conference on Computer and Communications Security, Raleigh, NC, USA, 16–18 October 2012; pp 906–917, doi:10.1145/2382196.2382292.

연구실적 제목	Profitable Double-Spending Attacks
연구실적 유형	논문(○) 특허( ) 기술이전( ) 표준화( )

비트코인의 채굴은 작업증명(PoW) 에 의해서 이루어진다. 하지만 PoW는 높은 해시 파워를 요구하므로 채굴만을 위한 application-specific integrated circuit (ASIC)을 가진 마이닝 풀과 같은 채굴 연합체가 비트코인 채굴을 독점하게 되는 문제가 있다. 본 연구팀은 이러한 문제를 해결할 수 있도록 ASIC을 사용한 채굴에 저항성을 가지는 새로운 작업증명 방법인 ECCPoW를 제안하였다. 제안된 ECCPoW는 오류정정부호 중 한 종류인 low density parity check (LDPC) 코드 및 해쉬함수를 결합한 방법이다. 본 연구팀은 ECCPoW를 비트코인에 적용하고 ECCPoW를 기존 비트코인의 PoW방법과 채굴중앙화, 보안성, 확장성 측면에서 비교하였다.

Table 2. Numerical computations of required resources for profitable DS attacks with $p_A = 0.4$ when $t_{cut} = cN_{BC}\lambda_H^{-1}$ with $c = 4$.

Block Confirmation Number (N_{BC})	1	3	5	7	9
Attack success probability (P_{AS})	0.411	0.419	0.376	0.334	0.297
Expected attack success time ($\mathbf{E}_{T_{AS}}$)(Scaled by $\lambda_H{}^{-1}$)	1.953	5.338	8.434	11.418	14.325
Expected OPEX ($\mathbf{E}_X$)(Scaled by γ)	2.106	6.139	10.436	14.977	19.716
Required value of target transaction ($C_{Sat.}$)(Scaled by γ)	1.302 $\cdot(1-\mu)$ + 3.819	3.559 $\cdot(1-\mu)$ + 11.10	5.622 $\cdot(1-\mu)$ + 22.15	7.612 $\cdot(1-\mu)$ + 37.25	9.550 $\cdot(1-\mu)$ + 56.96

applied sciences

MDPI

Article

Profitable Double-Spending Attacks

Jehyuk Jang and Heung-No Lee *

School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology (GIST), 123 Cheomdangwagi-ro, Buk-gu, Gwangju 61005, Korea; jjh2014@gist.ac.kr

* Correspondence: heungno@gist.ac.kr

Received: 9 November 2020; Accepted: 25 November 2020; Published: 27 November 2020

check for updates

Abstract: Our aim in this paper is to investigate the profitability of double-spending (DS) attacks that manipulate an a priori mined transaction in a blockchain. It was well understood that a successful DS attack is established when the proportion of computing power an attacker possesses is higher than that of the honest network. What is not yet well understood is how threatening a DS attack with less than 50% computing power used can be. Namely, DS attacks at any proportion can be a threat as long as the chance to make a good profit exists. Profit is obtained when the revenue from making a successful DS attack is greater than the cost of carrying out one. We have developed a novel probability theory for calculating a *finite time* attack probability. This can be used to size up attack resources needed to obtain the profit. The results enable us to derive a sufficient and necessary condition on the value of a transaction targeted by a DS attack. Our result is quite surprising: we theoretically show how a DS attack at any proportion of computing power can be made profitable. Given one's transaction value, the results can also be used to assess the risk of a DS attack. An example of profitable DS attack against BitcoinCash is provided.

Keywords: blockchain; double-spending attack; Fraud risk analysis; profitability; time-finite analysis; probability distribution; combinatorics

1. Introduction

A blockchain is a distributed ledger which has originated from the desire to find a novel alternative to centralized ledgers such as transactions through third parties [1]. Besides the role as a ledger, blockchains have been applied to many areas, e.g., managing the access authority to shared data in the cloud network [2] and averting collusion in e-Auction [3]. In a blockchain network based on the proof-of-work (PoW) mechanism, each miner verifies transactions and tries to put them into a block and mold the block to an existing chain by solving a cryptographic puzzle. This series of processes is called mining. However, the success of mining a block is given to only a single miner who solves the cryptographic puzzle for the first time. The reward of minting a certain amount of coins to the winner motivates more miners to join and remain in the network. As a result, blockchains have been designed so that the validity of transactions is confirmed by a lot of decentralized miners in the network.

A consensus mechanism is programmed for decentralized peers in a network to share a common chain. If a full-node succeeds in generating a new block, it has the latest version of the chain. All of the nodes in the network continuously communicate with each other to share the latest chain. A node may run into a situation in which it encounters mutually different chains more than one. In such a case, it utilizes a consensus rule with which it selects a single chain. Satoshi Nakamoto suggested the longest chain consensus for Bitcoin protocol in which the node selects the longest chain among all competing chains [1]. There are also other consensus rules [4,5], but a common goal of consensus rules is to select the single chain by which the most computation resources have been consumed based on the belief that it may have been verified by the largest number of miners.

of future profits can be updated by substituting it into Equation (33). That is to say, at each state, the estimation of profits can be updated and used as the rewards resulting from the continue action.

Goffard [19] and Karame et al. [20] have derived the PDFs of attack success time, but none of their DS attack scenarios matched with ours in Definition 1. In [19], Goffard derived the PDF of catch-up time spent for the fraudulent chain to catch up with the honest chain given that the length of honest chain is initially ahead by several blocks. The author used counting processes such as order statistic point process and renewal process which are more general than PCP, but there was no analytic result similar to what is given in Proposition 1. In [20], Karame et al. derived the PDF of the first attack success time under a fast-payment model which fixed $N_{BC} = 0$. To sum up, the attack success time in neither analysis included the time spent for achieving the first condition: the block confirmation should be realized.

7. Discussion and Conclusions

We showed that DS attacks using 50% or a lower proportion of computing power can be profitable and thus quite threatening. We provided how much quantitative resources are required to make a profitable DS attack. We derive the PDF of attack success time which enables us to figure out the operating time and the expense of mining rigs. We provided MATLAB codes on the website (https://codeocean.com/capsule/2308305/tree) for numerical evaluation of the expected profit function in Equation (33). We also listed an example of the minimum resources required for a profitable DS attack, which is applicable to any blockchain networks by substituting the network parameters γ, β, and λ_H. We also showed a more specific example of the required resources against BitcoinCash network.

Our results quantitatively guide how to set a block confirmation number for a safe transaction. The lower the block confirmation number is, the lower the minimum resource is required for a profitable attack. A solution can be utilized by the network developers to discourage such an attack. On the one hand, given a block confirmation number, we can have the value of any transaction to be set below the required value of making a profitable attack in a given network. On the other hand, given the value of transaction, the network can provide a service to inform the payee with the lowest block confirmation number that leads to negative DS attack profit.

Author Contributions: Conceptualization, J.J. and H.-N.L.; methodology, J.J.; software, J.J.; validation, J.J.; formal analysis, J.J.; investigation, J.J.; resources, J.J.; data curation, J.J.; writing—original draft preparation, J.J.; writing—review and editing, J.J. and H.-N.L.; visualization, J.J. and H.-N.L.; supervision, H.-N.L.; project administration, H.-N.L.; funding acquisition, H.-N.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by Institute of Information & Communications Technology Planning & Evaluation, grant number 2020-0-00958. This work was partially supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (NRF-2018R1A2A1A19018665).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Proof of Lemma 1. For a given sample ω and a given index i, the event $\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}$ is equivalent to the event that there exists an intermediate state index j such that $\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}$. By the mutual exclusiveness of $\mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}$ for integers j, such a state j is unique if it exists. Thus, we can write the probability $p_{DSA,i}$ as follows,

$$\begin{aligned} p_{DSA,i} &= \Pr\left(\exists j \in \mathbb{N}: \omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}\right) \\ &= \sum_{j=N_{BC}}^{\infty} \Pr\left(\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}\right). \end{aligned} \tag{A1}$$

연구실적 제목	Blockchain-Based Distributed Patient-Centric Image Management System
연구실적 유형	논문(○) 특허( ) 기술이전( ) 표준화( )

오늘날 의료 이미지 보관 및 교환을 위해서 클라우드 기반의 중앙화된 데이터센터를 사용한다. 하지만 이러한 방식을 사용하면 방대한 저장공간이 필요하고, 유지보수에 많은 비용이 들며, 네트워크를 통할 때에는 개인정보보호에 관한 문제도 발생한다. 본 연구진은 분산형 환자중심 영상 관리 (PCIM) 시스템을 제안하였다. 제안된 PCIM은 이더리움 블록체인 및 분산형 파일시스템인 IPFS를 활용한다. 영상데이터에 접근시 개인정보 보호를 위한 인증 과정을 위해 이더리움 블록체인 스마트컨트랙트를 사용하였고, IPFS를 분산 데이터 저장을 위해 사용하였다.

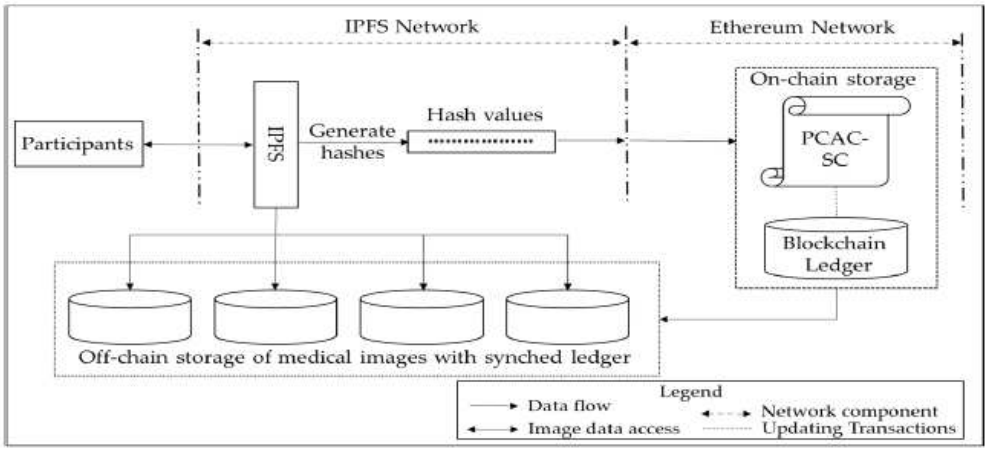


Figure 2. Architecture of patient-centric image management (PCIM) system. The architecture component split into two main decentralized modules: IPFS and Ethereum network.

applied sciences

MDPI

Article

Blockchain-Based Distributed Patient-Centric Image Management System

Mohamed Yaseen Jabarulla and Heung-No Lee *

School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, Korea; yaseen@gm.gist.ac.kr
* Correspondence: heungno@gist.ac.kr

Abstract: In recent years, many researchers have focused on developing a feasible solution for storing and exchanging medical images in the field of health care. Current practices are deployed on cloud-based centralized data centers, which increase maintenance costs, require massive storage space, and raise privacy concerns about sharing information over a network. Therefore, it is important to design a framework to enable sharing and storing of big medical data efficiently within a trustless environment. In the present paper, we propose a novel proof-of-concept design for a distributed patient-centric image management (PCIM) system that is aimed to ensure safety and control of patient private data without using a centralized infrastructure. In this system, we employed an emerging Ethereum blockchain and a distributed file system technology called Inter-Planetary File System (IPFS). Then, we implemented an Ethereum smart contract called the patient-centric access control protocol to enable a distributed and trustworthy access control policy. IPFS provides the means for decentralized storage of medical images with global accessibility. We describe how the PCIM system architecture facilitates the distributed and secured patient-centric data access across multiple entities such as hospitals, patients, and image requestors. Finally, we deployed a smart contract prototype on an Ethereum testnet blockchain and evaluated the proposed framework within the Windows environment. The evaluation results demonstrated that the proposed scheme is efficient and feasible.

Keywords: blockchain; distributed storage; medical image sharing; healthcare system; smart contract; IPFS

check for updates

Citation: Jabarulla, M.Y.; Lee, H.-N. Blockchain-Based Distributed Patient-Centric Image Management System. *Appl. Sci.* **2021**, *11*, 196. https://dx.doi.org/10.3390/app11010196

Received: 24 November 2020
Accepted: 24 December 2020
Published: 28 December 2020

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The transition to electronic management of health records has necessitated practitioners and their patients to make use of several new acronyms such as electronic medical records (EMRs), electronic health records (EHRs), and personal health records (PHRs) [1]. These health records usually contain medical images and patient information, such as physician name, personal statistics (e.g., age and weight), home monitoring device data, and other data processed by practitioners in a text format. Medical images and patient information are stored and maintained by different hospitals, even when being related to the same patient. The Deep Learning Market Report [2] estimated that 90% of medical health records generated in hospitals are in the form of images. In the future, hospitals require more storage space for processing such a large volume of image data, and rapid demand raises for accessing the sensitive images. Hence, it is essential to store and disseminate the medical image securely without interruptions. Current technologies for transferring medical images and patient information are deployed on centralized data centers that are deemed inappropriate due to privacy, accessibility, storage, and security concerns. Over recent decades, medical record data breaches within large medical data centers create additional difficulties for all companies seeking to develop medical image processing applications [3]. Therefore, it is necessary for a patient-centric system, i.e., a patient able

to the patient. Moreover, strict penalties are associated with the unlawful disclosure of sensitive medical images. Thus, healthcare institutions are expected to be hesitant in their adoption of this unconventional architecture. Overall, the practical usefulness of the PCIM system depends on the participant experience. Since blockchain itself does not address broader problems in medical image management. For example, the ability to share access for medical images using the proposed system does not alone ensure image usability unless supporting information about the patient image is not defined by the radiologists. Thus, there must be a good synergy required between the participants of the system. However, Decentralized data management is the groundbreaking of blockchain development, and it is more of a fascinating prototype of what health care technology could look like in the future.

7. Conclusions

Patient medical images are the most valuable asset of any healthcare system's intelligence. Most of the time, these medical images are indeed scattered across different systems, and sharing them is influential for establishing effective and cohesive healthcare. In addition, a centralized hosting location of image data (e.g., cloud-based solution) can be a single point of a security attack. With growing recognition of the distributed nature of health services, attention has been increasingly focused on decentralized architectures and system interoperability. In this paper, we presented the POC design of the proposed PCIM system: an Ethereum blockchain and IPFS-based decentralized framework for storing and sharing access to medical images. Moreover, we introduced a new access management system called PCAC-SC that enables authorized entities to access the relevant blockchain data. The PCIM system facilitates a unique way to improve the right of patients by providing full control over their medical images using PCAC-SC protocol. Patients have complete transparency over their medical images and can grant permission to access or revoke the image for clinical trials or research purposes. We performed the experimental implementation to analyze and evaluate efficiency, rationality and feasibility of the proposed scheme. The proposed system facilitates patient access to an immutable medical database providing higher efficiency, data provenance, and effective audit while sharing access to medical images. The data storage and exchange model is also decentralized; therefore, necessity to involve third-party intermediaries and administrative structures is eliminated.

Our future research goal is to deploy the proposed POC design in the public blockchain using real-world scenarios to form a global PCIM system, and to evaluate policies and regulations to adopt this emerging technology within the healthcare system. Furthermore, we will add a flexible business market model by utilizing a credit mechanism in the SC and artificial intelligence components into our proposed framework. The market model could help the patients receive incentives for providing their medical images for the Deep learning research projects [2], which cannot be accomplished in conventional image management systems. The artificial intelligence component helps clinicians to improve the interpretation of diagnostic images without delays in communication.

Author Contributions: Conceptualized and designed the research work, M.Y.J. and H.-N.L.; performed the implementation and evaluation of the research work, M.Y.J.; wrote the manuscript, M.Y.J.; performed investigation, made revisions to the manuscript and supervised the research work, H.-N.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by Institute of Information & Communications Technology Planning & Evaluation, grant number 2020-0-00958. This work was partially supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) [NRF-2018R1A2A1A19018665].

Data Availability Statement: The Solidity code for implementing and verifying the presented PCAC-SC protocol are available in a publicly accessible GitHub repository. The prototype code can be found here: https://github.com/infonetGIST/PCAC-SC.

Conflicts of Interest: The authors declare no conflict of interest.

연구실적 제목	Fabrication of 2D thin-film filter-array for compressive sensing spectroscopy
연구실적 유형	논문(○) 특허( ) 기술이전( ) 표준화( )

본 논문에서는 2차원 필터 어레이를 활용한 압축센싱 분광기를 제안하였다. 광학적 설계 통해 분광기의 각각의 필터는 서로 다른 투과 특성을 갖도록 디자인되었으며, 다층 박막 공정을 통해 제작되었다. 제안된 방식의 압축센싱 분광기는 적은 수의 샘플로도 원래의 신호를 복원하였으며, 광학적 실험을 통해 제안된 방식의 분광기가 기존의 분광기와 비슷한 성능을 보였다. 제작된 분광기는 한번의 측정을 통해 모든 신호를 획득 하며, 넓은 파장대역에서 안정적인 성능을 보여주었다.

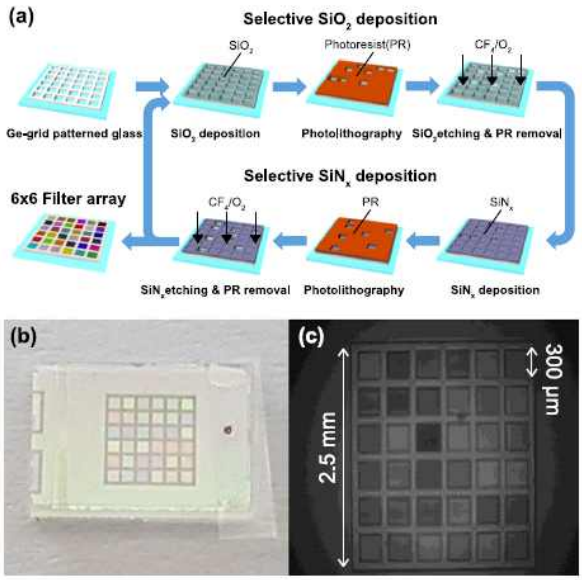


Fig. 2. (a) Schematic of the thin-film filter-array fabrication process. (b) Photograph of a fabricated thin-film filter-array. (c) Monochrome image of the thin-film filter-array taken at a wavelength of 700 nm.

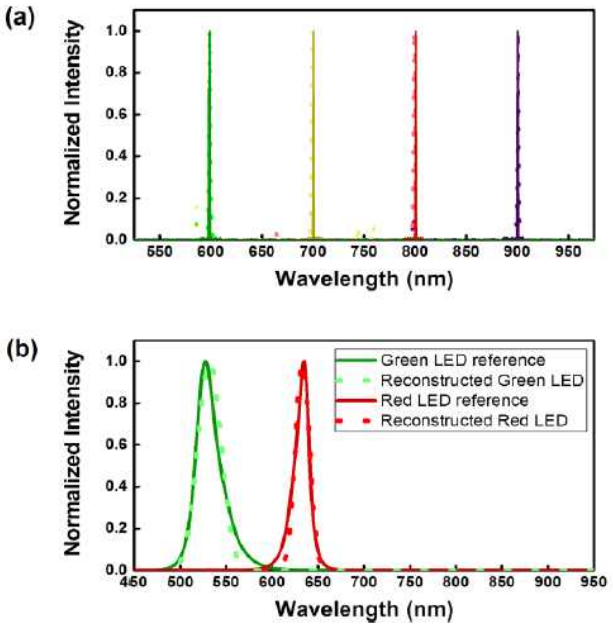


Fig. 6. Spectral reconstructions of several different input light sources. (a) Spectral reconstructions of monochromatic lights (dots) compared with reference spectra (solid lines): 600 nm (green), 700 nm (yellow), 800 nm (red), and 900 nm (purple). (b) Spectral reconstructions of LEDs (dots) compared with reference spectra (solid lines): green LED (527 nm), and red LED (635 nm).

Optics and Lasers in Engineering 115 (2019) 53–58

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

journal homepage: www.elsevier.com/locate/optlaseng

Fabrication of 2D thin-film filter-array for compressive sensing spectroscopy

Cheolsun Kim, Woong-Bi Lee, Soo Kyung Lee, Yong Tak Lee, Heung-No Lee*

School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea

ARTICLE INFO

Keywords:
Spectroscopy
Thin films
Inverse problems
Compressive sensing

ABSTRACT

We demonstrate 2D filter-array compressive sensing spectroscopy based on thin-film technology and a compressive sensing reconstruction algorithm. To obtain different spectral modulations, we fabricate a set of multilayer filters using alternating low- and high-index materials and reconstruct the input spectrum using a small number of measurements. Experimental results show that the fabricated filter-array provides compatible spectral resolution performance with a conventional spectrometer in monochromatic lights and LEDs. In addition, the fabricated filter-array covers a wide range of wavelengths with a single exposure.

1. Introduction

The demand for spectrum information is increasing not only in research and development but also in the private sector. In response to this demand, researchers are trying to make spectrometers that are both small and inexpensive. These spectrometers could be used in various fields, such as medical systems, mobile applications, and remote sensing [1–3]. In particular, optical filter-based spectrometers do not need motorized or dispersive elements, and their filter-array can be directly attached to the detectors so that they can be easily miniaturized. However, there is a trade-off between size (for integrating filters) and spectral resolution with miniaturized spectrometers.

Over the years, numerous approaches to applying compressive sensing (CS) techniques have been proposed to reduce the size of spectroscopes without reducing spectral resolution, or potentially even improving it. These approaches [4–7] include the following: band pass filters [4], random transmittance filters [5], photonic crystal slabs [6], and liquid crystal phase retarders [7]. Recently, Fabry–Perot (FP)-based CS spectroscopy methods have been presented [8,9]. To acquire differently modulated spectral measurements, a 2D array of FP resonators with different cavity depths has been tried [8] as well as a piezo-actuated device that changes the distance between two FP mirrors has been tried [9]. A hundred FP resonators are used to recover the input spectrum in [8], and the operational range of the piezo-actuator imposes mechanical limitations in [9].

The CS framework [10–12] is an efficient sampling and reconstruction scheme that requires fewer samples to reconstruct the signal than that required by conventional sampling. The CS framework can be applied to filter-based spectroscopy, offering the advantage of reducing the number of filters and detectors required and allowing the system to be miniaturized.

In spectroscopy, the relation between the spectral components of the input light source $\mathbf{x} \in \mathbb{R}^{N\times 1}$ and the modulated signal $\mathbf{y} \in \mathbb{R}^{M\times 1}$ can be expressed as follows:

$$\mathbf{y} = \mathbf{T}\mathbf{x}, \tag{1}$$

where $\mathbf{T} \in \mathbb{R}^{M\times N}$ is the sensing matrix. Each row of the sensing matrix is to represent the transmission function (TF) of i-th filter, $\boldsymbol{T}_m \in \mathbb{R}^{1\times N}$ for $m = 1, 2, \ldots, M$. In order to achieve miniaturization of spectroscope without degradation of spectral resolution, the CS framework is utilized in spectroscopy, where the number of filters is set to be smaller than the number of spectral components ($M < N$). Then, Eq. (1) becomes an underdetermined linear system. A sparse signal reconstruction algorithm with $L1$ norm minimization can be used to solve Eq. (1), if the input spectrum is either naturally sparse or can be sparsely represented in some basis $\boldsymbol{\Phi} \in \mathbb{R}^{N\times N}$, i.e., $\mathbf{x} = \boldsymbol{\Phi}\mathbf{s}$, where $\mathbf{s} \in \mathbb{R}^{N\times 1}$ is a sparse vector. Then, Eq. (1) becomes

$$\mathbf{y} = \mathbf{T}\boldsymbol{\Phi}\mathbf{s} \tag{2}$$

The sparse signal $\mathbf{s}$ can be estimated by solving the following $L1$ norm minimization problem:

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \|\mathbf{s}\|_1 \text{ subject to} \|\mathbf{y} - \mathbf{T}\boldsymbol{\Phi}\mathbf{s}\|_2 \le \varepsilon \tag{3}$$

where ε is a small non-negative constant. The reconstructed input spectrum $\hat{\mathbf{x}}$ is then $\boldsymbol{\Phi}\hat{\mathbf{s}}$.

In this paper, we demonstrate 2D filter-array CS spectroscopy. This uses a multilayer thin-film filter-array for spectral modulation, where each filter modulates the input spectrum using different sensing patterns. A CMOS image camera reads out the modulated signals with a

* Corresponding author.
E-mail address: heungno@gist.ac.kr (H.-N. Lee).

https://doi.org/10.1016/j.optlaseng.2018.10.018
Received 22 February 2018; Received in revised form 2 October 2018; Accepted 31 October 2018
Available online 24 November 2018
0143-8166/© 2018 Elsevier Ltd. All rights reserved.

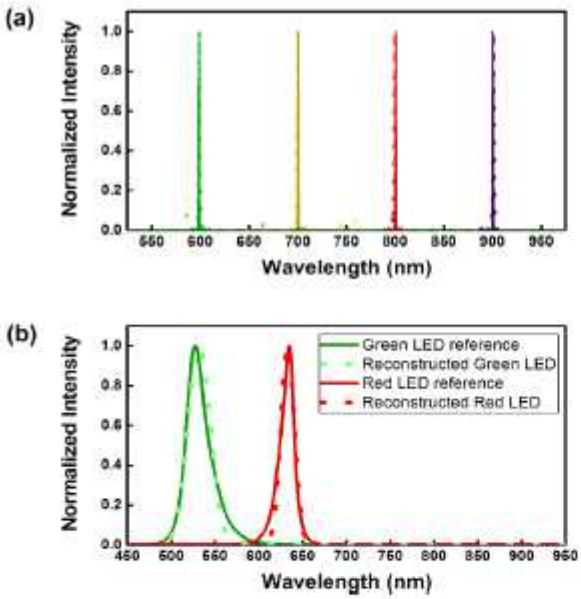


Fig. 6. Spectral reconstructions of several different input light sources. (a) Spectral reconstructions of monochromatic lights (dots) compared with reference spectra (solid lines): 600 nm (green), 700 nm (yellow), 800 nm (red), and 900 nm (purple). (b) Spectral reconstructions of LEDs (dots) compared with reference spectra (solid lines): green LED (527 nm), and red LED (635 nm).

(red LED), and the reconstructed spectral peak locations are 531 nm (green LED) and 633 nm (red LED). The peak signal-to-noise ratios are 28.3 dB (green LED) and 31.7 dB (red LED).

Discussing Fig. 6, the spectra of reconstructed monochromatic lights show several negligible spikes. This is probably due to background noise in the optical experiments. But overall, the reconstruction results of the proposed CS spectroscopy system for monochromatic lights and LEDs are similar to those of the grating spectrometer. Furthermore, the number of modulated signals is significantly small ($M = 36$) that the measurement to wavelength sample ratio is 36:500 (ratio between M and N).

To further explore the performance of the proposed CS spectroscopy, we conducted the computational experiment on the fabricated filter-array using a continuous light source, halogen lamp. For the experiment, we used the measured sensing matrix**T**. The conventionally measured spectrum of the halogen lamp was used as the input spectrum **x**. The modulated signals were generated by numerically multiplying the sensing matrix and the input spectrum. By solving Eq. (3), we reconstructed the continuous spectrum of light. In Fig. 7, we present computational spectral reconstruction of the halogen lamp. The peak signal-to-noise ratio is 43.8 dB. Due to the limitations of our optical components to reject the spectrum of the halogen lamp except for the wavelength range from 500 to 1000 nm, we could not perform the optical experiment on the continuous source. However, the computational reconstruction result of the halogen lamp indicates that the fabricated filter-array can be utilized for recovering the various kinds of spectra in the given wavelength range without limitations of the optical components.

Fabricating the proposed filter-array can be more difficult than fabricating Fabry–Perot structure due to the large number of layers for the proposed filter-array. However, the proposed spectroscope is compact and it does not need motorized components which were used with the Fabry–Perot structure [9]. In addition, thanks to the 2D array structure, the proposed spectroscope captures all measurements in a single exposure. But the Fabry–Perot spectroscope [9] required a number of

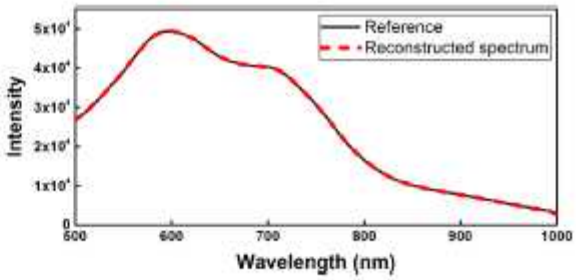


Fig. 7. Computational spectral reconstruction of a halogen lamp (red dash line) compared with the reference spectrum (black solid line) measured by a conventional spectrometer.

exposures as many times as the number of measurements. Compared to Fabry–Perot spectroscope [8], the proposed spectroscope utilizes 36 filters to cover the wavelength range from 500 to 1000 nm, but 100 filters were used in [8] to cover the range from 500 to 750 nm.

4. Conclusion

We have demonstrated a 2D array CS spectroscope based on thin-film technology. A 2D thin-film filter-array is fabricated based on array processing. Using the fabricated filter-array, measurements are obtained to which the CS reconstruction algorithm is applied. Finally, demonstration of input spectrum reconstruction is successfully made. The proposed system is compact, portable, and obtains the necessary measurements in a single exposure thanks to its structural advantages. Moreover, it works over a wide spectral range, from the visible light region to the near-infrared region. Compared with conventional spectrometers (non-CS spectrometers), the proposed system has a high optical throughput and compatible spectral resolution performance in monochromatic lights and LEDs with significantly less number of measurements.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (NRF-2018R1A2A1A19018665).

References

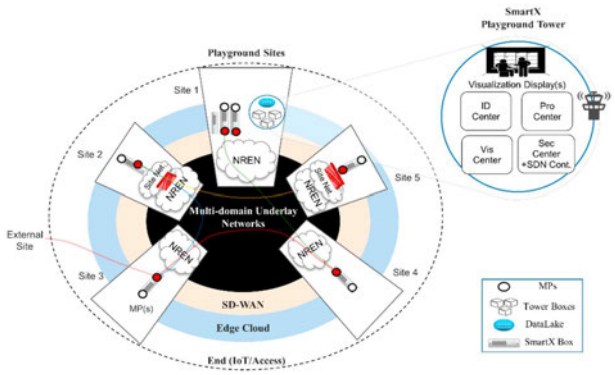
[1] Bacon CP, Mattley Y, DeFrece R. Miniature spectroscopic instrumentation: applications to biology and chemistry. Rev Sci Instrum 2004;75:1–16.
[2] Kim S, Cho D, Kim J, Kim M, Youn S, Jang JE, et al. Smartphone based multispectral imaging: system development and potential for mobile skin diagnosis. Biomed Opt Express 2016;7:5294–307.
[3] Clark RN, Roush TL. Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. J Geophys Res Solid Earth 1984;89:6329–40.
[4] Oliver J, Lee W, Park S, Lee H-N. Improving resolution of miniature spectrometers by exploiting sparse nature of signals. Opt Express 2012;20:2613–25.
[5] Oliver J, Lee W-B, Lee H-N. Filters with random transmittance for improving resolution in filter-array-based spectrometers. Opt Express 2013;21:3969–89.
[6] Wang Z, Yu Z. Spectral analysis based on compressive sensing in nanophotonic structures. Opt Express 2014;22:25608–14.
[7] August Y, Stern A. Compressive sensing spectrometry based on liquid crystal devices. Opt lett 2013;38:4996–9.
[8] Huang E, Ma Q, Liu Z. Etalon array reconstructive spectrometry. Sci Rep 2017;7.
[9] Oiknine Y, August I, Blumberg DG, Stern A. Compressive sensing resonator spectroscopy. Opt lett 2017;42:25–8.
[10] Donoho DL. Compressed sensing. IEEE Trans Inf Theory 2006;52:1289–306.
[11] Baraniuk RG. Compressive sensing [lecture notes]. IEEE Signal Process Mag 2007;24:118–21.
[12] Qaisar S, Bilal RM, Iqbal W, Naureen M, Lee S. Compressive sensing: from theory to applications, a survey. J Commun Netw 2013;15:443–56.

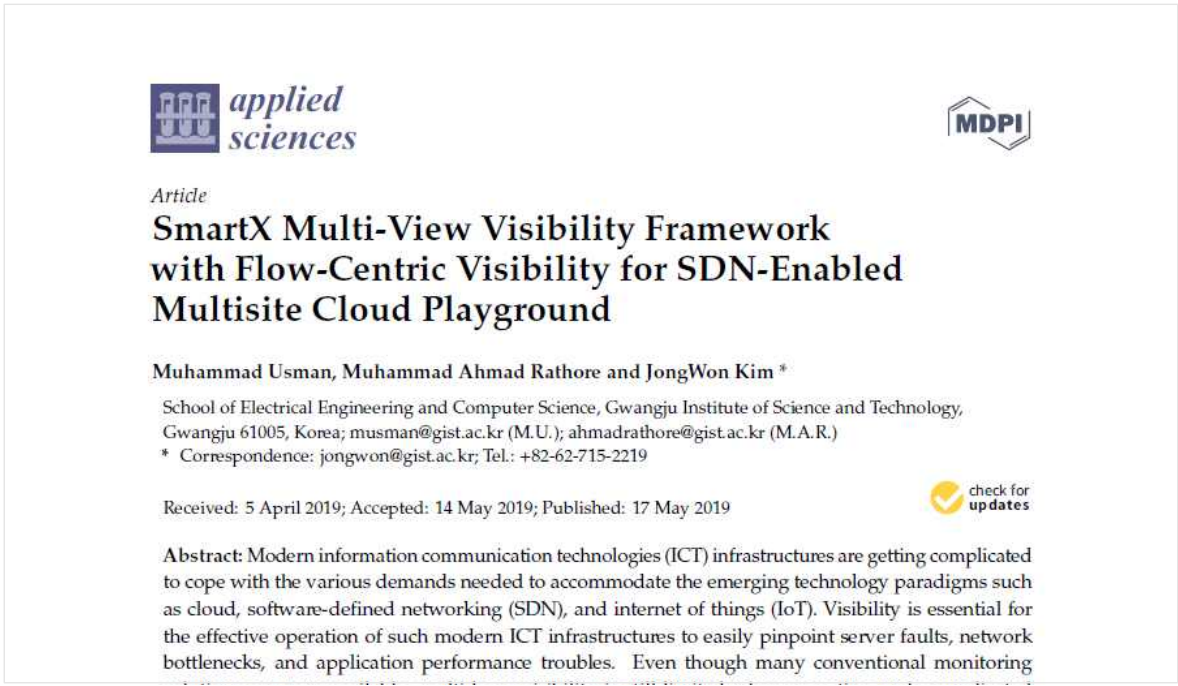
[붙임3]

‘제2세부 프로젝트 책임자(김종원)’ 대표 연구실적 요약문

연구실적 제목	SmartX Multi-View Visibility Framework with Flow-Centric Visibility for SDN-Enabled Multisite Cloud Playground
연구실적 유형	논문(○) 특허( ) 기술이전( ) 표준화( )

○ Visibility is essential for the effective operation of such modern ICT infrastructures to easily pinpoint server faults, network bottlenecks, and application performance troubles. Even though many conventional monitoring solutions are now available, multi-layer visibility is still limited when operating such complicated infrastructures. To address this particular limitation, a futuristic multi-layer visibility framework denoted as SmartX multi-view visibility framework (MVF), is proposed for unifying the visibility of underlay, physical and virtual resources, flow, and workload layers. To unify multi-layer visibility, this paper presents a comprehensive extension of SmartX MVF with flow-centric visibility for simultaneously monitoring physical-virtual resources, flows classification, and visualization to eventually assist secured operation of SDN-enabled multisite cloud infrastructure.





applied sciences

MDPI

Article

SmartX Multi-View Visibility Framework with Flow-Centric Visibility for SDN-Enabled Multisite Cloud Playground

Muhammad Usman, Muhammad Ahmad Rathore and JongWon Kim *

School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, Korea; musman@gist.ac.kr (M.U.); ahmadrathore@gist.ac.kr (M.A.R.)
* Correspondence: jongwon@gist.ac.kr; Tel.: +82-62-715-2219

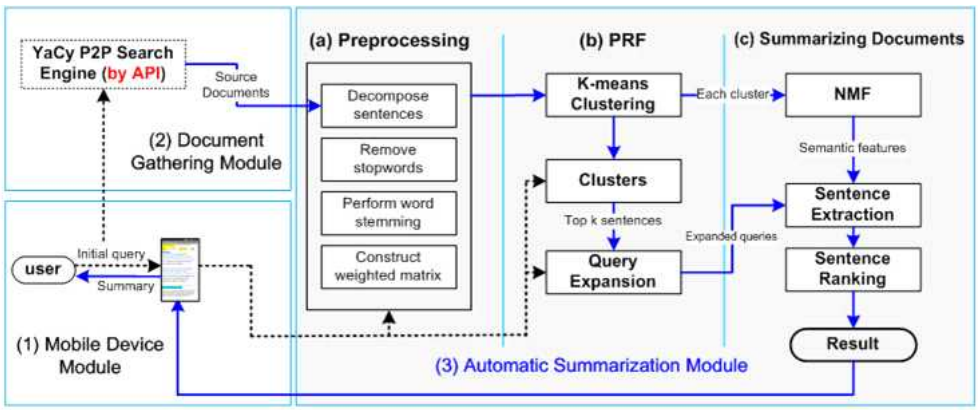
Received: 5 April 2019; Accepted: 14 May 2019; Published: 17 May 2019

check for updates

Abstract: Modern information communication technologies (ICT) infrastructures are getting complicated to cope with the various demands needed to accommodate the emerging technology paradigms such as cloud, software-defined networking (SDN), and internet of things (IoT). Visibility is essential for the effective operation of such modern ICT infrastructures to easily pinpoint server faults, network bottlenecks, and application performance troubles. Even though many conventional monitoring

연구실적 제목	Mobile IoT device summarizer using P2P web search engine and inherent characteristic of contents
연구실적 유형	논문(○) 특허( ) 기술이전( ) 표준화( )

○ A new containerized automatic mobile device summarizer is proposed, which it use pseudo relevance feedback with inherent characteristic of documents by NMF and clustering method. The proposed summarizer consists of mobile device module of user side, Document Gathering module and automatic summarization module regarding server side. The mobile device module is user interface for documents summarization. The document gathering module collects source documents related to the initial query by using the YaCy API. The automatic summarization module performs the document summarization. The document gathering and summarization modules are implemented by using the container for flexible management.



Peer-to-Peer Networking and Applications
https://doi.org/10.1007/s12083-019-00780-w

Check for updates

Mobile IoT device summarizer using P2P web search engine and inherent characteristic of contents

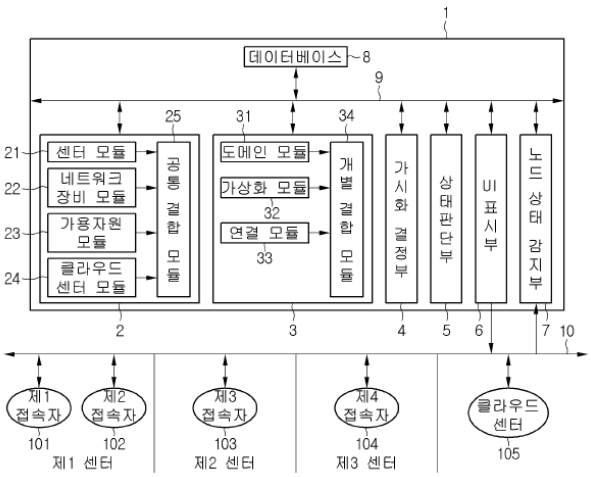
Sun Park[1] · ByungRea Cha[1] · Kyungyong Chung[2] · JongWon Kim[3]

Received: 28 January 2019 / Accepted: 21 June 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

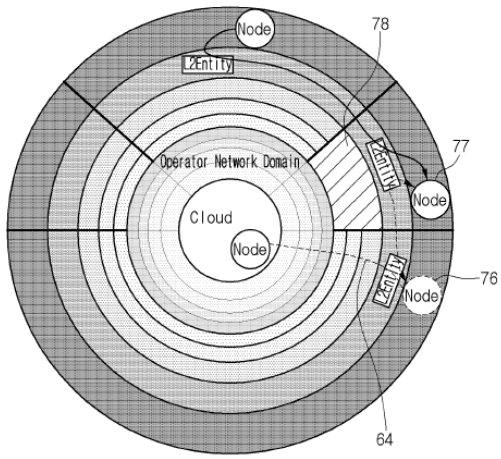
Abstract
Grasping the sense of web pages on small screen of mobile devices is too difficult, so mobile users prefer summarized reports. One popular technique to generate automatic summarized report is based on vector model. The vector model summarization methods are simple and low cost to implement on mobile devices. However, the quality of summarized reports of the model may not be good because there is a semantic difference between manual summarized reports and the machine summarized reports. In addition, the censored documents of the centralized search engine cannot accurately reflect the query of user in the summary results. To overcome this constraints, automatic mobile device summarizer using a semantic feature of pseudo relevance feedback and P2P web search engine is proposed. The proposed method increases the quality of summarized reports by using the latent features of documents and clustering technique. The method uses none censored documents of the P2P web search engine to more accurately reflect user requirements in summarized reports. The automatic summarizer consists of a server-side containerized auto-summary module for flexible management of the summarization functions and a user-side mobile device module to reduce the overload of the mobile device.

연구실적 제목	Integrated network controlling system, network controlling apparatus and method, and a readable media containing the method
연구실적 유형	논문( ) 특허(○) 기술이전( ) 표준화( )

○ 인공지능 서비스 인프라를 위한 네트워크 관리장치는 적어도 두 개의 센터에서 동작하는 적어도 하나의 노드를 감지하는 노드상태감지부, 서로 겹치지 않고 적어도 공통되는 내부 영역를 가지는 적어도 두 개의 폐곡선에 있어서, 적어도 두 개의 폐곡선의 공통되는 내부로 정의되는 제 1 영역과, 두 개의 폐곡선의 사이 공간으로 정의되는 제 2 영역을 가지고, 두 개의 센터는 제 2 영역의 내부에서 두 개의 폐곡선을 연결하는 구획선에 의해서 서로 구획하여 표시하는 센터모듈 있음. 또한 노드들의 연결 상태를 선으로 나타내는 연결선을 표시하는 연결모듈 및 센터모듈 및 상기 연결모듈에 의해서 제공되는 선을 표시하는 UI표시부가 포함됨. 이를 통하여 인공지능 서비스 인프라의 네트워크의 가시성이 향상시킬 수 있음.



[인공지능 서비스를 위한 네트워크 관리장치]



[인공지능 서비스를 위한 인프라 네트워크 UI]

공개특허 10-2016-0065506

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0065506
(43) 공개일자 2016년06월09일

(51) 국제특허분류(Int. Cl.)
H04L 12/24 (2006.01)
(21) 출원번호 10-2014-0169554
(22) 출원일자 2014년12월01일
심사청구일자 없음

(71) 출원인
광주과학기술원
광주광역시 북구 첨단과기로 123 (오룡동)
(72) 발명자
김종원
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 정보통신공학부
나태흠
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 정보기전공학부
(74) 대리인
김기문

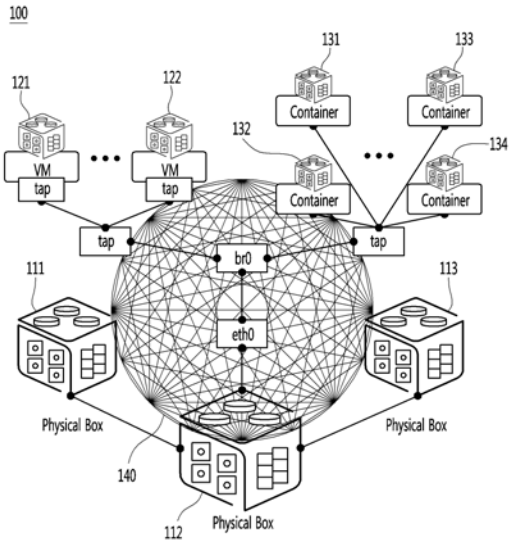
전체 청구항 수 : 총 27 항

(54) 발명의 명칭 통합 망 관리 시스템, 네트워크 관리장치, 네트워크 관리방법, 및 그 방법이 수록되는 기록매체

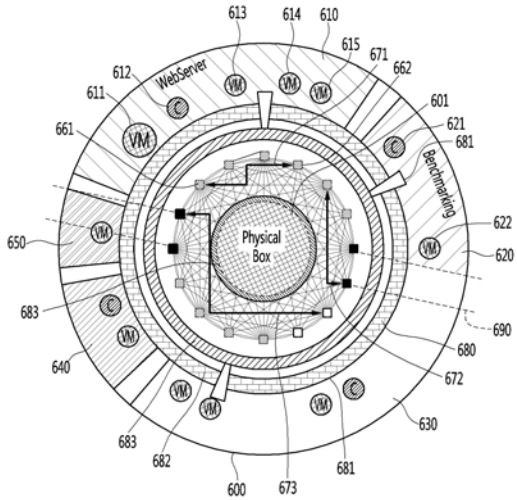
(57) 요 약

연구실적 제목	클러스터 시각화 장치
연구실적 유형	논문( ) 특허(○) 기술이전( ) 표준화( )

○ 현재 클라우드 기반의 인공지능 서비스가 각광받고 있으며, 인공지능 서비스를 위한 클라우드 인프라구성에 최근 컨테이너(Container)와 같은 유연성을 가진 논리적인 자원들의 활용빈도가 증가하고 있는 상황임. 이러한 추세에 따라 물리적인 다수의 박스뿐만 아니라 논리적인 박스들을 포함하여 다수의 박스를 엮어 클러스터 단위로 활용하는 사례가 증가하고 있음. 클러스터를 구성하는 복수의 박스들의 상이한 타입/크기/사용용도/네트워킹 연결 상태/보안 수준을 시각화하는 새로운 방식의 시각화 모델이 필요하게 되었음. 본 기술은 인공지능 서비스를 위한 클라우드 인프라의 통합적 관측을 위한 3차원 원뿔대 모양의 시각화 모델과 이를 통해 클러스터 전체 또는 부분적인 시각화가 가능 함.



[인공지능 서비스 대상 환경(클러스터)의 예]



[인공지능 인프라 클러스터 상태 정보 3차원 가시화]

공개특허 10-2019-0005632

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0005632
(43) 공개일자 2019년01월16일

(51) 국제특허분류(Int. Cl.)
G06F 16/00 (2019.01)
(52) CPC특허분류
G06F 16/00 (2019.01)
(21) 출원번호 10-2017-0086579
(22) 출원일자 2017년07월07일
심사청구일자 2017년07월07일

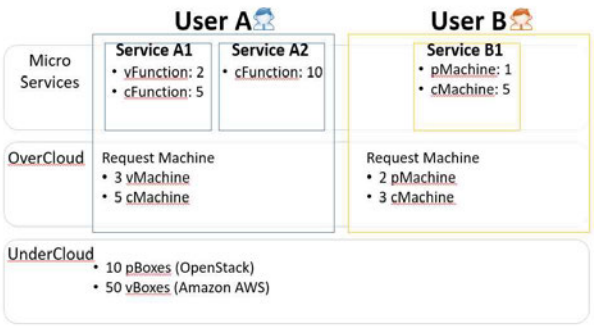
(71) 출원인
광주과학기술원
광주광역시 북구 첨단과기로 123 (오룡동)
(72) 발명자
김종원
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 전기전자컴퓨터공학부
남택호
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 전기전자컴퓨터공학부
(74) 대리인
김기문

전체 청구항 수 : 총 8 항

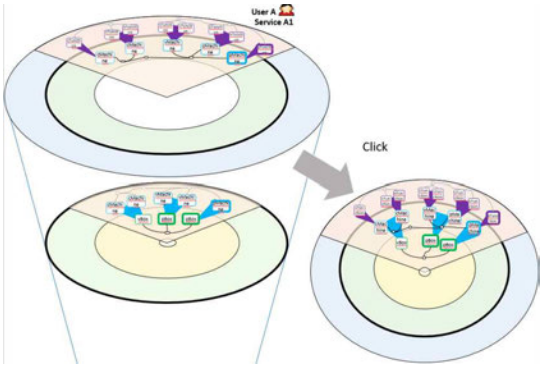
(54) 발명의 명칭 클러스터 시각화 장치

연구실적 제목	클라우드 서비스를 위한 사용자 인터페이스 기술
연구실적 유형	논문( ) 특허(○) 기술이전( ) 표준화( )

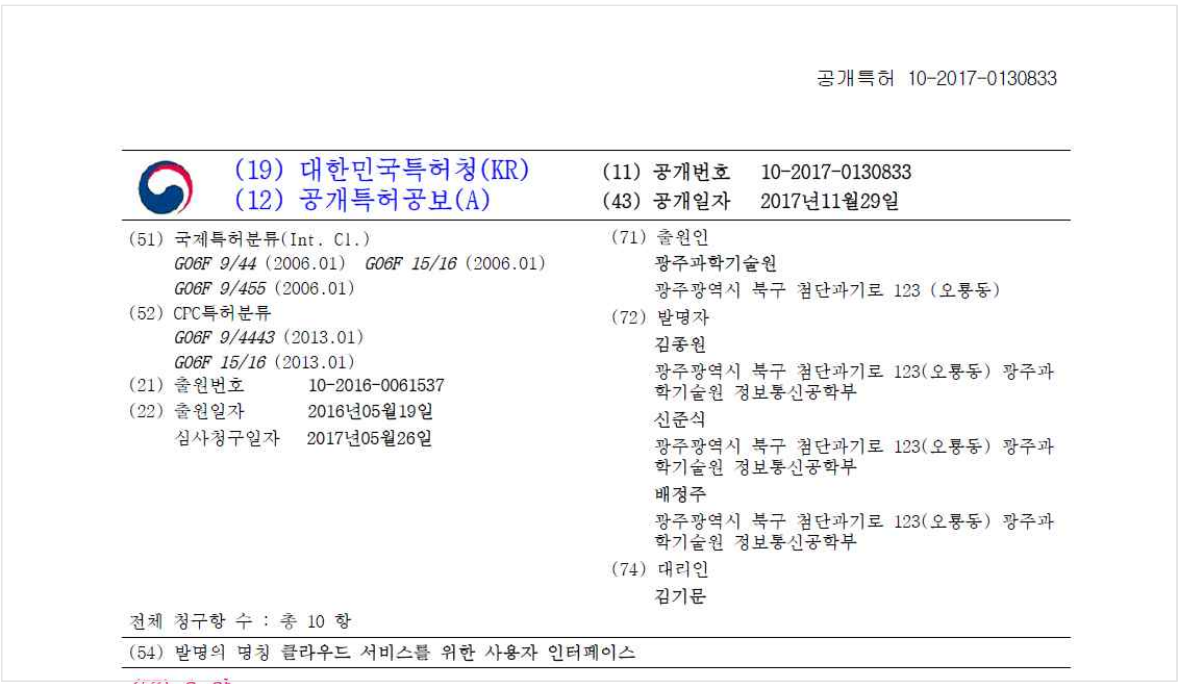
○ 인공지능 서비스를 위한 클라우드 사용자인터페이스를 위해서 실제 자원을 제공하는 언더레이 클라우드의 박스와 오버레이 클라우드의 클러스터 머신 정보를 매칭시켜 하나의 가시성 계층을 제공하는 “언더레이 클라우드-오버레이 클라우드 매핑 계층” X와 오버레이 클라우드의 클러스터와 그 위에서 동작하는 서비스의 요소기능들에 대한 매핑을 통해 통합적인 가시성을 제공하는 “오버레이 클라우드-마이크로서비스 매핑 계층” Y를 하나의 화면상에서 제공함으로써 전체 계층형 클라우드 인프라의 통합적인 가시성을 제공함. 또한 각 계층별로 p(Physical), v(Virtual), c(Container) 요소는 테두리의 모양 및 두께로 구분하여 표기함으로써, 일견에 해당 요소가 물리/가상/컨테이너 중 어느 것에 해당하는 것인지를 식별할 수 있게 하는 인터페이스 기술을 통하여 인공지능 서비스를 지원할 수 있음.



[인공지능 서비스의 가시성 제공 방법에 대한 예제 시나리오]



[인공지능 서비스 인프라 클러스터 상태 정보 3차원 가시화]



공개특허 10-2017-0130833

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0130833
(43) 공개일자 2017년11월29일

(51) 국제특허분류(Int. Cl.)
G06F 9/44 (2006.01) *G06F 15/16* (2006.01)
G06F 9/455 (2006.01)
(52) CPC특허분류
G06F 9/4443 (2013.01)
G06F 15/16 (2013.01)
(21) 출원번호 10-2016-0061537
(22) 출원일자 2016년05월19일
심사청구일자 2017년05월26일

(71) 출원인
광주과학기술원
광주광역시 북구 첨단과기로 123 (오룡동)
(72) 발명자
김종원
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 정보통신공학부
신준식
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 정보통신공학부
배정주
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 정보통신공학부
(74) 대리인
김기문

전체 청구항 수 : 총 10 항

(54) 발명의 명칭 클라우드 서비스를 위한 사용자 인터페이스

‘제3세부 프로젝트 책임자 (김기형)’ 대표 연구실적 요약문

연구실적 제목	Image Compression Using Chain Coding for Electronic Shelf Labels (ESL) Systems
연구실적 유형	논문(O) 특허( ) 기술이전( ) 표준화( )

One of the most important and recurring tasks in managing a store is to provide accurate, up-to-date price information to customers on the shelves. Manual updating of price tags has been a timeconsuming, error-prone task with high labor costs. An Electronic shelf labels (ESL) system is becoming an attractive alternative for this task because of the dynamic-price-updating and customer's productevaluation-display features. A common ESL system conguration in a retail store includes thousands of battery-powered ESL tags that are mostly connected wirelessly in a dense indoor environment. Raising the success ratio of wireless communication is essential for the system's viability due to its limited battery life. Most of the ESL trafc is the image data of goods that appear on the tags, and reducing the amount of the data is one of the most effective ways to enhance communication performance and reduce retransmission. This paper proposes an ESL image compression mechanism based on chain coding that utilizes ESL images' characteristics. The performance results show that the proposed mechanism could compress the ESL images smaller and decompress faster.

IEEE Access

Received December 3, 2020, accepted December 20, 2020, date of publication January 8, 2021, date of current version January 14, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3049868

Image Compression Using Chain Coding for Electronic Shelf Labels (ESL) Systems

YOUNGJUN KIM[1], KI-HYUNG KIM[2], (Senior Member, IEEE), AND WE-DUKE CHO[3]

[1]Department of Computer Engineering, Graduate School, Ajou University, Suwon 16499, South Korea
[2]Department of Cyber Security, Ajou University, Suwon 16499, South Korea
[3]Department of Electrical and Computer Engineering, Ajou University, Suwon 16499, South Korea

Corresponding author: Ki-Hyung Kim (kkim86@ajou.ac.kr)

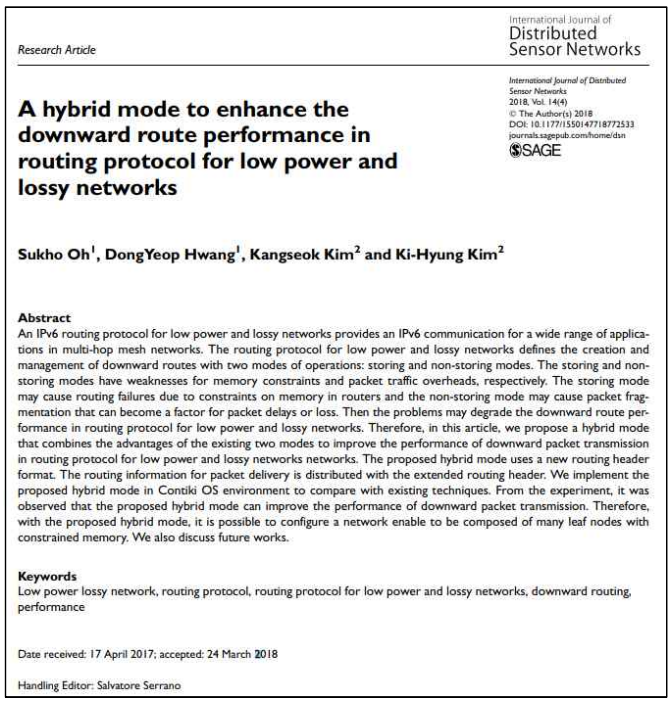
This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology under Grant 2018R1D1A1B07048697, in part by the Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government Ministry of Trade, Industry and Energy (MOTIE) (The Competency Development Program for Industry Specialist) under Grant P0008703, and in part by the Ministry of Science and ICT (MSIT), South Korea, through the Information Technology Research Center (ITRC) Support Program, supervised by the Institute for Information & Communications Technology Planning & Evaluation (IITP), under Grant IITP-2020-2018-0-01396.

ABSTRACT One of the most important and recurring tasks in managing a store is to provide accurate, up-to-date price information to customers on the shelves. Manual updating of price tags has been a time-consuming, error-prone task with high labor costs. An Electronic shelf labels (ESL) system is becoming an attractive alternative for this task because of the dynamic-price-updating and customer's product-evaluation-display features. A common ESL system configuration in a retail store includes thousands of battery-powered ESL tags that are mostly connected wirelessly in a dense indoor environment. Raising the success ratio of wireless communication is essential for the system's viability due to its limited battery life. Most of the ESL traffic is the image data of goods that appear on the tags, and reducing the amount of the data is one of the most effective ways to enhance communication performance and reduce retransmission. This paper proposes an ESL image compression mechanism based on chain coding that utilizes ESL images' characteristics. The performance results show that the proposed mechanism could compress the ESL images smaller and decompress faster.

INDEX TERMS Electronic shelf label, chain code, image compression.

연구실적 제목	A hybrid mode to enhance the downward route performance in routing protocol for low power and lossy networks
연구실적 유형	논문(O) 특허( ) 기술이전( ) 표준화( )

An IPv6 routing protocol for low power and lossy networks provides an IPv6 communication for a wide range of applications in multi-hop mesh networks. The routing protocol for low power and lossy networks defines the creation and management of downward routes with two modes of operations: storing and non-storing modes. The storing and non-storing modes have weaknesses for memory constraints and packet traffic overheads, respectively. The storing mode may cause routing failures due to constraints on memory in routers and the non-storing mode may cause packet fragmentation that can become a factor for packet delays or loss. Then the problems may degrade the downward route performance in routing protocol for low power and lossy networks. Therefore, in this article, we propose a hybrid mode that combines the advantages of the existing two modes to improve the performance of downward packet transmission in routing protocol for low power and lossy networks. The proposed hybrid mode uses a new routing header format. The routing information for packet delivery is distributed with the extended routing header. We implement the proposed hybrid mode in Contiki OS environment to compare with existing techniques. From the experiment, it was observed that the proposed hybrid mode can improve the performance of downward packet transmission. Therefore, with the proposed hybrid mode, it is possible to configure a network enable to be composed of many leaf nodes with constrained memory. We also discuss future works.



Research Article

International Journal of Distributed Sensor Networks

International Journal of Distributed Sensor Networks
2018, Vol. 14(4)
© The Author(s) 2018
DOI: 10.1177/1550147718772533
journals.sagepub.com/home/dsn
SAGE

A hybrid mode to enhance the downward route performance in routing protocol for low power and lossy networks

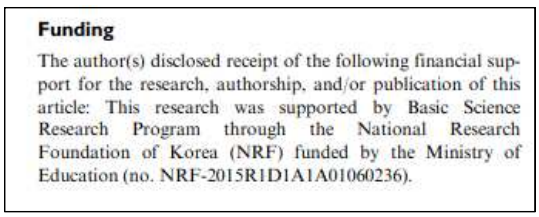
Sukho Oh[1], DongYeop Hwang[1], Kangseok Kim[2] and Ki-Hyung Kim[2]

Abstract
An IPv6 routing protocol for low power and lossy networks provides an IPv6 communication for a wide range of applications in multi-hop mesh networks. The routing protocol for low power and lossy networks defines the creation and management of downward routes with two modes of operations: storing and non-storing modes. The storing and non-storing modes have weaknesses for memory constraints and packet traffic overheads, respectively. The storing mode may cause routing failures due to constraints on memory in routers and the non-storing mode may cause packet fragmentation that can become a factor for packet delays or loss. Then the problems may degrade the downward route performance in routing protocol for low power and lossy networks. Therefore, in this article, we propose a hybrid mode that combines the advantages of the existing two modes to improve the performance of downward packet transmission in routing protocol for low power and lossy networks networks. The proposed hybrid mode uses a new routing header format. The routing information for packet delivery is distributed with the extended routing header. We implement the proposed hybrid mode in Contiki OS environment to compare with existing techniques. From the experiment, it was observed that the proposed hybrid mode can improve the performance of downward packet transmission. Therefore, with the proposed hybrid mode, it is possible to configure a network enable to be composed of many leaf nodes with constrained memory. We also discuss future works.

Keywords
Low power lossy network, routing protocol, routing protocol for low power and lossy networks, downward routing, performance

Date received: 17 April 2017; accepted: 24 March 2018

Handling Editor: Salvatore Serrano

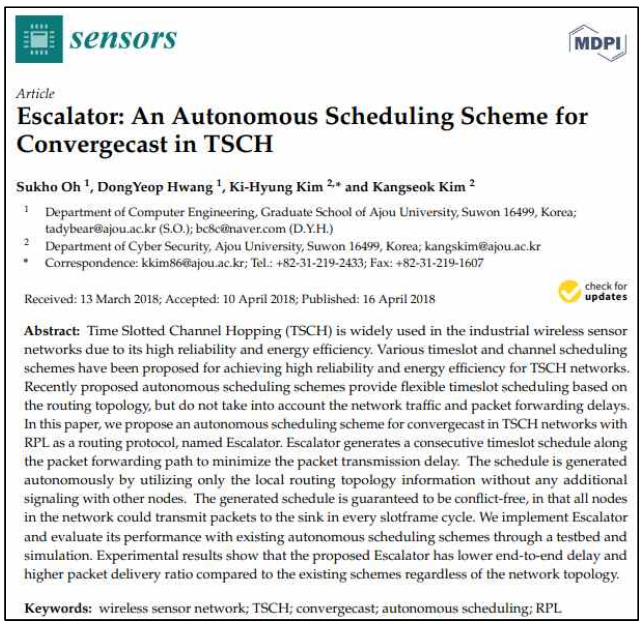


Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (no. NRF-2015R1D1A1A01060236).

연구실적 제목	Escalator: An Autonomous Scheduling Scheme for Convergecast in TSCH
연구실적 유형	논문(O) 특허( ) 기술이전( ) 표준화( )

Time Slotted Channel Hopping (TSCH) is widely used in the industrial wireless sensor networks due to its high reliability and energy efficiency. Various timeslot and channel scheduling schemes have been proposed for achieving high reliability and energy efficiency for TSCH networks. Recently proposed autonomous scheduling schemes provide flexible timeslot scheduling based on the routing topology, but do not take into account the network traffic and packet forwarding delays. In this paper, we propose an autonomous scheduling scheme for convergecast in TSCH networks with RPL as a routing protocol, named Escalator. Escalator generates a consecutive timeslot schedule along the packet forwarding path to minimize the packet transmission delay. The schedule is generated autonomously by utilizing only the local routing topology information without any additional signaling with other nodes. The generated schedule is guaranteed to be conflict-free, in that all nodes in the network could transmit packets to the sink in every slotframe cycle. We implement Escalator and evaluate its performance with existing autonomous scheduling schemes through a testbed and simulation. Experimental results show that the proposed Escalator has lower end-to-end delay and higher packet delivery ratio compared to the existing schemes regardless of the network topology.



sensors

MDPI

Article

Escalator: An Autonomous Scheduling Scheme for Convergecast in TSCH

Sukho Oh [1], DongYeop Hwang [1], Ki-Hyung Kim [2,*] and Kangseok Kim [2]

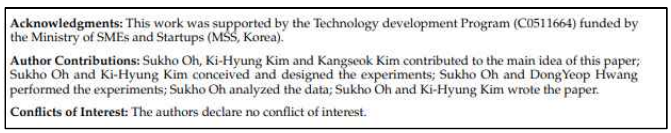
[1] Department of Computer Engineering, Graduate School of Ajou University, Suwon 16499, Korea; tadybear@ajou.ac.kr (S.O.); bc8c@naver.com (D.Y.H.)
[2] Department of Cyber Security, Ajou University, Suwon 16499, Korea; kangskim@ajou.ac.kr
* Correspondence: kkim86@ajou.ac.kr; Tel.: +82-31-219-2433; Fax: +82-31-219-1607

Received: 13 March 2018; Accepted: 10 April 2018; Published: 16 April 2018

check for updates

Abstract: Time Slotted Channel Hopping (TSCH) is widely used in the industrial wireless sensor networks due to its high reliability and energy efficiency. Various timeslot and channel scheduling schemes have been proposed for achieving high reliability and energy efficiency for TSCH networks. Recently proposed autonomous scheduling schemes provide flexible timeslot scheduling based on the routing topology, but do not take into account the network traffic and packet forwarding delays. In this paper, we propose an autonomous scheduling scheme for convergecast in TSCH networks with RPL as a routing protocol, named Escalator. Escalator generates a consecutive timeslot schedule along the packet forwarding path to minimize the packet transmission delay. The schedule is generated autonomously by utilizing only the local routing topology information without any additional signaling with other nodes. The generated schedule is guaranteed to be conflict-free, in that all nodes in the network could transmit packets to the sink in every slotframe cycle. We implement Escalator and evaluate its performance with existing autonomous scheduling schemes through a testbed and simulation. Experimental results show that the proposed Escalator has lower end-to-end delay and higher packet delivery ratio compared to the existing schemes regardless of the network topology.

Keywords: wireless sensor network; TSCH; convergecast; autonomous scheduling; RPL



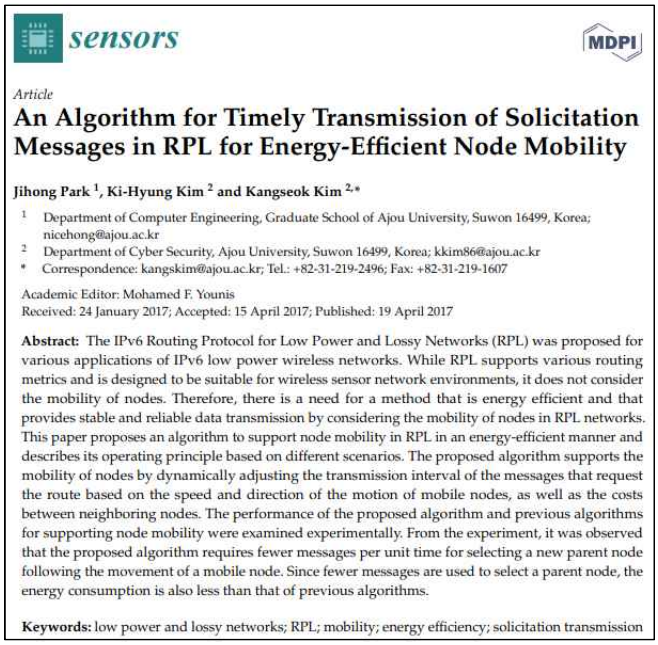
Acknowledgments: This work was supported by the Technology development Program (C0511664) funded by the Ministry of SMEs and Startups (MSS, Korea).

Author Contributions: Sukho Oh, Ki-Hyung Kim and Kangseok Kim contributed to the main idea of this paper; Sukho Oh and Ki-Hyung Kim conceived and designed the experiments; Sukho Oh and DongYeop Hwang performed the experiments; Sukho Oh analyzed the data; Sukho Oh and Ki-Hyung Kim wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

연구실적 제목	An Algorithm for Timely Transmission of Solicitation Message in RPL for Energy-Effcient Node Mobility
연구실적 유형	논문(O) 특허( ) 기술이전( ) 표준화( )

The IPv6 Routing Protocol for Low Power and Lossy Networks (RPL) was proposed for various applications of IPv6 low power wireless networks. While RPL supports various routing metrics and is designed to be suitable for wireless sensor network environments, it does not consider the mobility of nodes. Therefore, there is a need for a method that is energy efficient and that provides stable and reliable data transmission by considering the mobility of nodes in RPL networks. This paper proposes an algorithm to support node mobility in RPL in an energy-efficient manner and describes its operating principle based on different scenarios. The proposed algorithm supports the mobility of nodes by dynamically adjusting the transmission interval of the messages that request the route based on the speed and direction of the motion of mobile nodes, as well as the costs between neighboring nodes. The performance of the proposed algorithm and previous algorithms for supporting node mobility were examined experimentally. From the experiment, it was observed that the proposed algorithm requires fewer messages per unit time for selecting a new parent node following the movement of a mobile node. Since fewer messages are used to select a parent node, the energy consumption is also less than that of previous algorithms.



sensors

MDPI

Article

An Algorithm for Timely Transmission of Solicitation Messages in RPL for Energy-Efficient Node Mobility

Jihong Park [1], Ki-Hyung Kim [2] and Kangseok Kim [2,*]

[1] Department of Computer Engineering, Graduate School of Ajou University, Suwon 16499, Korea; nicehong@ajou.ac.kr
[2] Department of Cyber Security, Ajou University, Suwon 16499, Korea; kkim86@ajou.ac.kr
* Correspondence: kangskim@ajou.ac.kr; Tel.: +82-31-219-2496; Fax: +82-31-219-1607

Academic Editor: Mohamed F. Younis
Received: 24 January 2017; Accepted: 15 April 2017; Published: 19 April 2017

Abstract: The IPv6 Routing Protocol for Low Power and Lossy Networks (RPL) was proposed for various applications of IPv6 low power wireless networks. While RPL supports various routing metrics and is designed to be suitable for wireless sensor network environments, it does not consider the mobility of nodes. Therefore, there is a need for a method that is energy efficient and that provides stable and reliable data transmission by considering the mobility of nodes in RPL networks. This paper proposes an algorithm to support node mobility in RPL in an energy-efficient manner and describes its operating principle based on different scenarios. The proposed algorithm supports the mobility of nodes by dynamically adjusting the transmission interval of the messages that request the route based on the speed and direction of the motion of mobile nodes, as well as the costs between neighboring nodes. The performance of the proposed algorithm and previous algorithms for supporting node mobility were examined experimentally. From the experiment, it was observed that the proposed algorithm requires fewer messages per unit time for selecting a new parent node following the movement of a mobile node. Since fewer messages are used to select a parent node, the energy consumption is also less than that of previous algorithms.

Keywords: low power and lossy networks; RPL; mobility; energy efficiency; solicitation transmission

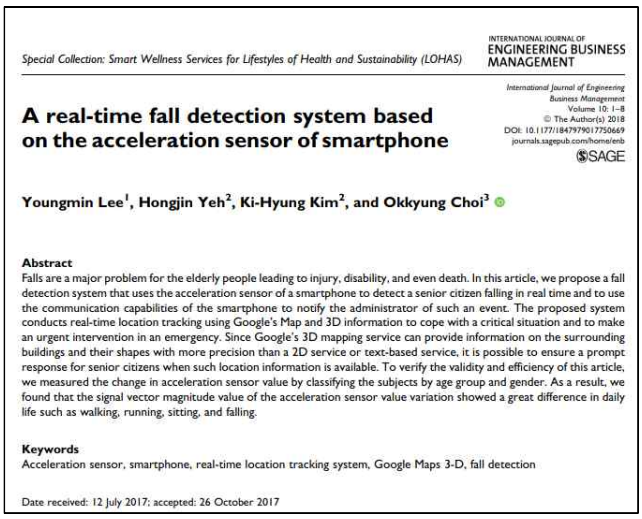
Acknowledgments: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. NRF-2015R1D1A1A01060236).

Author Contributions: Jihong Park, Ki-Hyung Kim and Kangseok Kim contributed to the main idea of this paper; Jihong Park and Ki-Hyung Kim conceived and designed the experiments; Jihong Park performed the experiments; Jihong Park and Kangseok Kim analyzed the data; Jihong Park and Kangseok Kim wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

연구실적 제목	A real-time fall detection system based on the acceleration sensor of smart phone
연구실적 유형	논문(O) 특허( ) 기술이전( ) 표준화( )

Falls are a major problem for the elderly people leading to injury, disability, and even death. In this article, we propose a fall detection system that uses the acceleration sensor of a smartphone to detect a senior citizen falling in real time and to use the communication capabilities of the smartphone to notify the administrator of such an event. The proposed system conducts real-time location tracking using Google's Map and 3D information to cope with a critical situation and to make an urgent intervention in an emergency. Since Google's 3D mapping service can provide information on the surrounding buildings and their shapes with more precision than a 2D service or text-based service, it is possible to ensure a prompt response for senior citizens when such location information is available. To verify the validity and efficiency of this article, we measured the change in acceleration sensor value by classifying the subjects by age group and gender. As a result, we found that the signal vector magnitude value of the acceleration sensor value variation showed a great difference in daily life such as walking, running, sitting, and falling.



Special Collection: Smart Wellness Services for Lifestyles of Health and Sustainability (LOHAS)

INTERNATIONAL JOURNAL OF ENGINEERING BUSINESS MANAGEMENT

International Journal of Engineering Business Management
Volume 10: 1–8
© The Author(s) 2018
DOI: 10.1177/1847979017750669
journals.sagepub.com/home/enb
SAGE

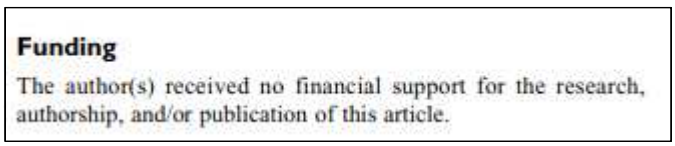
A real-time fall detection system based on the acceleration sensor of smartphone

Youngmin Lee[1], Hongjin Yeh[2], Ki-Hyung Kim[2], and Okkyung Choi[3]

Abstract
Falls are a major problem for the elderly people leading to injury, disability, and even death. In this article, we propose a fall detection system that uses the acceleration sensor of a smartphone to detect a senior citizen falling in real time and to use the communication capabilities of the smartphone to notify the administrator of such an event. The proposed system conducts real-time location tracking using Google's Map and 3D information to cope with a critical situation and to make an urgent intervention in an emergency. Since Google's 3D mapping service can provide information on the surrounding buildings and their shapes with more precision than a 2D service or text-based service, it is possible to ensure a prompt response for senior citizens when such location information is available. To verify the validity and efficiency of this article, we measured the change in acceleration sensor value by classifying the subjects by age group and gender. As a result, we found that the signal vector magnitude value of the acceleration sensor value variation showed a great difference in daily life such as walking, running, sitting, and falling.

Keywords
Acceleration sensor, smartphone, real-time location tracking system, Google Maps 3-D, fall detection

Date received: 12 July 2017; accepted: 26 October 2017



Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

연구실적 제목	모바일 노드들의 그룹 설정 방법 및 그룹 설정 장치(Method and device of configuring mobile nodes) 기술이전
연구실적 유형	논문( ) 특허( ) 기술이전(O) 표준화( )

사용자 파라미터를 기초로 하여 복수의 모바일 노드들을 관리자 노드들과 그룹원 노드들로 빠르고 간편하게 설정할 수 있는 모바일 노드의 그룹 설정 방법이 개시된다. 모바일 노드의 그룹 설정 방법은 하나의 그룹을 형성하고자 하는 복수의 모바일 노드들을 사용자 파라미터를 기초로 하여 제 1 및 제 2 서브 영역으로 분류하는 단계, 상기 모바일 노드들에 동일 그룹 식별자를 전송하여 하나의 그룹을 형성하는 단계, 및 상기 제 1 및 제 2 서브 영역에 위치한 모바일 노드들에 각각 서브 그룹 식별자를 전송하여 상기 모바일 노드들을 관리자 노드들 또는 그룹원 노드들로 설정하는 단계를 포함한다. 따라서 본 발명은 복수의 모바일 노드들을 빠르고 간편하게 관리자 노드들과 그룹원 노드들로 설정할 수 있으므로 시간과 비용을 절약할 수 있다.

기술이전 계약 확인서

□ 확인 대상 계약서

o 계약명 : 기술양도계약
o 계약 대상 : ㈜태진티엔에스
o 계약 내용 : 아주대학교 산학협력단(이하 "갑"이라 한다)과 ㈜태진티엔에스(이하 "을"이라 한다)는 "갑"이 보유한 "모바일 노드들의 그룹 설정 방법 및 그룹 설정 장치" 특허(등록번호 10-0897614호)를 아주대학교 김기형교수를 기술이전 책임자로 하여 "을"에게 이전하는 계약을 체결한다.
o 계약일 : 2019.11.27.
o 기술이전료 : [redacted]

□ 확인 내용

o 2019년 11월 27일 아주대학교 산학협력단과 ㈜태진티엔에스 간에 "모바일 노드들의 그룹 설정 방법 및 그룹 설정 장치" 기술을 이전하는 계약은 다음과 같이 국민대학교 에너지인터넷연구센터 연구결과로 수행된 기술이전 계약임을 확인합니다.

- 다 음 -

지원기관	과학기술정보통신부
전문기관	정보통신기획평가원
사 업 명	정보통신기술인력양성(R&D)
연구과제명	지능형 Internet of Energy(IoE) Data 연구 (과제코드 : 2018-0-01396-002)
연구책임자	국민대학교 창의공과대학 전자공학부 장영민 교수

2019년 12월 19일

확인자 아주대학교 산학협력단장 임재성 (인)

에너지인터넷연구센터장 장영민 귀중

[붙임4]

주관기관 및 참여대학 보유기술 현황

기관명	유형	기 술 명	보유자	등록/인증번호	개발연도	권리관계
광주과학기술원	해외(미국) 특허 등록	Spectrometry apparatus and spectrometry method	이흥노	US 10,458,843 B2	2019	자체보유
광주과학기술원	해외(일본) 특허 등록	Apparatus for Improving Spectral Resolution using Random Transmittance in Optical Spectrometers	이흥노	JP, 6290905	2018	자체보유
광주과학기술원	해외(미국) 특허 등록	Cross Reference to Related Application	이흥노	US 10,080,485 B2	2018	자체보유
광주과학기술원	해외(미국) 특허 등록	METHOD FOR RECONSTRUCTING SPARSE SIGNAL IN FINITE FIELD, APPARATUS FOR RECONSTRUCTING SPARSE SIGNAL IN FINITE FIELD, AND RECORDING MEDIUM FOR RECORDING RECONSTRUCTION METHOD	이흥노	US 9,734,128 B2	2017	자체보유
광주과학기술원	해외(일본) 특허 등록	유한체의 희소신호 복구방법, 유한체의 희소신호 복구장치, 및 이 방법을 기록되는 기록매체	이흥노	JP, 5914755	2016	자체보유
광주과학기술원	해외(미국) 특허 등록	Method and Apparatus for Processing Optical Signal Of Spectrometer Using Sparse Nature of Signals	이흥노	US 9,030,662	2015	자체보유
광주과학기술원	해외(미국) 특허 등록	Signal Acquisition and Method for Distributed Compressive Sensing and Joint Signal Recovery	이흥노	US 8,391,800 B2	2013	자체보유
광주과학기술원	국내 특허 등록	디지털 신호처리를 이용한 필터 기반 분광기의 각도 추정 기법	이흥노	10-2030735	2019	자체보유
광주과학기술원	국내 특허 등록	하이퍼스펙트럴 이미지 장치	이흥노	10-1986998	2019	자체 보유
광주과학기술원	국내 특허 등록	분광장치 및 분광방법	이흥노	10-2030735	2019	자체보유
광주과학기술원	국내 특허 등록	촬상장치, 촬상방법, 거리측정장치, 및 거리측정방법	이흥노	10-1865126	2018	자체보유
광주과학기술원	국내 특허 등록	랜덤필터모듈의 투과율 검출방법	이흥노	10-1572080	2015	자체보유
광주과학기술원	국내 특허 등록	분광계의 광 신호 처리 방법 및 그 장치	이흥노	10-1423964	2014	자체보유
광주과학기술원	해외(미국)	Blockchain e-voting system without trusted	이흥노	500.20.502.19.USP	2020	자체보유

	특허 출원	tallying agency				
광주과학기술원	PCT 특허 출원	블록체인의 거래검증시스템, 및 블록체인의 거래검증방법	이흥노	PCT/KR2019/017571	2019	자체보유
광주과학기술원	PCT 특허 출원	새로운 블록체인 및 암호화폐 작업증명 생성, 증명, 검증 시스템	이흥노	PCT/KR2019/008904	2019	자체보유
광주과학기술원	PCT 특허 출원	Sensing data processing apparatus and data processing method	이흥노	PCT/KR2017/005653	2017	자체보유
광주과학기술원	국내 특허 출원	블록체인 전자투표시스템, 그 시스템의 운용방법	이흥노	10-2020-0171424	2020	자체보유
광주과학기술원	국내 특허 출원	부호-암호 화폐 시스템	이흥노	10-2019-0151246	2019	자체보유
광주과학기술원	국내 특허 출원	Scalable DeSecure ECCPoW Blockchains 난이도 조절 Algorithm	이흥노	10-2019-0099153	2019	자체보유
광주과학기술원	국내 특허 출원	블록체인거버넌스	이흥노	10-2019-0084800	2019	자체보유
광주과학기술원	국내 특허 출원	크로스레이어 인증 프로토콜을 위한 결합 방법 및 시스템	황의석	10-2019-0154998	2019	자체보유
광주과학기술원	국내 특허 출원	물리 레이어 인증을 수행하기 위한 전자 장치 및 방법	황의석	10-2019-0150363	2019	자체보유
고려대학교	국내 특허 등록	분산 장부에 기반한 개인 정보 접근 권한 거래 방법 및 그 방법을 이용한 기록 매체	고려대학교 산학협력단	10-2187700	2019	자체보유
고려대학교	해외 (미국) 특허 등록	METHOD AND APPARATUS OF ALLOCATING RESOURCE OF TERMINAL IN WIRELESS COMMUNICATION SYSTEM	고려대학교 산학협력단	US 9,913,280	2018	자체보유
고려대학교	해외 (미국) 특허 등록	APPARATUS AND METHOD FOR CONGESTION DETECTION OF WIRELESS NETWORK IN A COMMUNICATION SYSTEM	고려대학교 산학협력단	US 9,826,428	2017	자체보유
고려대학교	해외 (미국) 특허 등록	APPARATUS AND METHOD FOR DETECTING NETWORK CONGESTION BASED ON DESTINATION NODE	고려대학교 산학협력단	US 9,819,600	2017	자체보유
고려대학교	해외 (미국) 특허 등록	METHOD AND APPARATUS FOR ESTIMATING CHANNEL QUALITY OF WIRELESS COMMUNICATION TECHNOLOGY	고려대학교 산학협력단	US 10,237,003	2019	자체보유
고려대학교	해외 (미국) 특허 등록	APPARATUS AND METHOD FOR ATTEMPTING CONNECTION TO AN ACCESS POINT	고려대학교 산학협력단	US 9,706,488	2017	자체보유
아주대학교	국내 특허 등록	제1 네트워크에서 제2 네트워크로 전송되는 데이터 패킷의 서비스 품질을 제어하는 방법 및 장치	아주대학교 산학협력단	10-0843645	2008	자체보유

아주대학교	국내 특허 등록	심플 서비스 로케이션 프로토콜 네트워크 시스템 및 이를 구성하는 에이전트	아주대학교 산학협력단	10-0854679	2008	자체보유
아주대학교	국내 특허 등록	무선 네트워크의 보안을 관리하기 위한 방법 및 이를 수행하기 위한 장치	아주대학교 산학협력단	10-0854680	2008	자체보유
아주대학교	국내 특허 등록	인터넷 프로토콜 유비쿼터스 센서 네트워크와 단순네트워크 관리 프로토콜 망을 상호 연동하기 위한 게이트웨이	아주대학교 산학협력단	10-0854681	2008	자체보유
아주대학교	국내 특허 등록	다른 센서 노드들 간의 전송을 지원하는 무선 개인 영역네트워크 시스템	아주대학교 산학협력단	10-0870657	2008	자체보유
아주대학교	국내 특허 등록	게이트웨이와 이를 이용한 패킷 변환 방법	아주대학교 산학협력단	10-0870655	2008	자체보유
아주대학교	국내 특허 등록	센서네트워크에서 계층적 주소할당방법	아주대학교 산학협력단	10-0870651	2008	자체보유
아주대학교	국내 특허 등록	센서 노드	아주대학교 산학협력단	10-0900796	2009	자체보유
아주대학교	국내 특허 등록	센서 포탈 시스템	아주대학교 산학협력단	10-0903983	2009	자체보유
아주대학교	국내 특허 등록	센서 네트워크의 구성 정보를 설정하는 방법과 시스템	아주대학교 산학협력단	10-0924562	2009	자체보유
아주대학교	국내 특허 등록	도서관의 좌석 관리를 위하여 사용 가능한, 복수의 사람에의하여 공동으로 사용되는 물건의 점유 상태를 관리하는방법 및 장치	아주대학교 산학협력단	10-0924703	2009	자체보유
아주대학교	국내 특허 등록	센서 네트워크 내의 센서로부터 출력된 센싱 데이터를 전송하는 방법과 시스템, 및 이를 수행하는 명령어들을 저장하는 저장 매체	아주대학교 산학협력단	10-0926291	2009	자체보유
아주대학교	국내 특허 등록	센서 네트워크 관리 시스템, 센서 네트워크를 관리하는 방법, 이를 수행하는 명령어들을 저장하는 저장 매체	아주대학교 산학협력단	10-0941396	2010	자체보유
아주대학교	국내 특허 등록	로우팬 부트스트래핑 에이전트 장치 및 로우팬 부트스트래핑 시스템	아주대학교 산학협력단	10-0964602	2010	자체보유
아주대학교	국내 특허 등록	센서 네트워크 통신 방법 및 장치 (SENSOR NETWORK COMMUNICATION METHOD AND APPARATUS)	아주대학교 산학협력단	10-0964601	2010	자체보유
아주대학교	국내 특허 등록	로봇 게임 시스템 (ROBOT GAME SYSTEM)	아주대학교 산학협력단	10-0974094	2010	자체보유
아주대학교	국내 특허 등록	적어도 하나 이상의 센서를 포함하는 센서 네트워크를 이용하여 해썹을 관리하는 방법 및 시스템	아주대학교 산학협력단	10-0978120	2010	자체보유
아주대학교	국내 특허 등록	홈 센서 네트워크 서비스 장치 (HOME SENSOR NETWORK SERVICE APPARATUS)	아주대학교 산학협력단	10-1010020	2011	자체보유

아주대학교	국내 특허 등록	홈 센서 네트워크 서버 (HOME SENSOR NETWORK SERVER)	아주대학교 산학협력단	10-1018826	2011	자체보유
아주대학교	국내 특허 등록	저전력 무선 개인 영역 네트워크상에서 계층적 라우팅을 위한 모바일 노드의 주소 할당 방법 및 패킷 전송 방법	아주대학교 산학협력단	10-1039097	2021	자체보유
아주대학교	국내 특허 등록	웹 페이지에 의하여 관리하는 홈 센서 네트워크 시스템	아주대학교 산학협력단	10-1039096	2011	자체보유
아주대학교	국내 특허 등록	릴레이 노드 배치방법, 릴레이 노드 배치 장치 및 컴퓨터로 읽을 수 있는 기록매체	아주대학교 산학협력단	10-1040216	2011	자체보유
아주대학교	국내 특허 등록	아이피 기반 유에스엔과 기존 아이피 네트워크를 연결하기 위한 방법 (Method for connecting IP-based USN with conventional IP network)	아주대학교 산학협력단	10-1043587	2011	자체보유
아주대학교	국내 특허 등록	아이피 기반의 유에스엔 상에서 객체 정보 서비스 시스템(OBJECT DIRECTORY SYSTEM WITH IP-BASED USN)	아주대학교 산학협력단	10-1047875	2011	자체보유
아주대학교	국내 특허 등록	무선 센서 네트워크에서의 라우팅 경로 설정 방법 및 이를 수행하기 위한 장치	아주대학교 산학협력단	10-1055416	2011	자체보유
아주대학교	국내 특허 등록	복수개의 게이트웨이들을 포함하는 센서 네트워크 시스템	아주대학교 산학협력단	10-1059517	2011	자체보유
아주대학교	국내 특허 등록	저전력 무선 개인 영역 네트워크에서, 디렉토리 프록시 에이전트를 동작시키는 방법	아주대학교 산학협력단	10-1064926	2011	자체보유
아주대학교	국내 특허 등록	무선 센서 네트워크에서의 패밀리-기반 주소 부여 방법과, 이를 이용한 계층적 라우팅 경로설정 방법 (METHOD FOR ALLOWING A FAMILY-BASED ADDRESS IN A WIRELESS SENSOR NETWORK, AND METHOD FOR HIERARCHICAL ROUTING A PATH SETTING USING THE SAME)	아주대학교 산학협력단	10-1079137	2011	자체보유
아주대학교	국내 특허 등록	로우맙 시스템과 디로우맙 시스템(LoWMob SYSTEM & DLoWMob SYSTEM)	아주대학교 산학협력단	10-1081038	2011	자체보유
아주대학교	국내 특허 등록	유비쿼터스 센서 네트워크 시스템 및 이를 구성하는 방법 (UBIQUITOUS SENSOR NETWORK SYSTEM AND METHOD OF CONFIGURING THE SAME)	아주대학교 산학협력단	10-1091696	2011	자체보유
아주대학교	해외 (미국) 특허 등록	Method for connecting IP-based USN with conventional IP network	아주대학교 산학협력단	8208468	2012	자체보유
아주대학교	국내 특허	무선 네트워크 상의 혼잡을 줄이는 네트워크 관리 방법 및	아주대학교 산학협력단	10-2012-0002396	2012	자체보유

	출원	시스템				
아주대학교	국내 특허 등록	무선 네트워상의 보더 게이트웨이의 트래픽 혼잡을 줄이는 기법	아주대학교 산학협력단	10-1364705	2014	자체보유
아주대학교	국내 특허 등록	무선 네트워크상의 지연전송을 개선하기 위한 슈퍼프레임 스케줄링 방법	아주대학교 산학협력단	10-1409461	2014	자체보유
아주대학교	해외 (미국) 특허 등록	유비쿼터스 센서 네트워크 시스템 및 이를 구성하는 방법	아주대학교 산학협력단	8756299	2014	자체보유
아주대학교	해외 (미국) 특허 등록	저전역 무선 개인 영역 네트워크에서, 디렉토리 프록시 에이전트를 동작시키는 방법, 특정 서비스 정보를 검색하는 방법 및 이를 수행하는 디렉토리 프록시 에이전트	아주대학교 산학협력단	8886122	2014	자체보유
아주대학교	국내 특허 출원	IoT 기반 아기 요람 장치	아주대학교 산학협력단	10-2015-0023374	2015	자체보유
아주대학교	국내 특허 등록	서비스 품질 기반의 보조그래프 생성 방법 및 그 시스템	아주대학교 산학협력단	10-1505627	2018	자체보유
아주대학교	국내 특허 출원	TSCH에서의 컨버지캐스트를 위한 자율 스케쥴링 기법	아주대학교 산학협력단	10-2018-0046319	2018	자체보유
아주대학교	국내 특허 출원	ESL 관리 서버, 디스플레이 장치 및 그 제어 방법	아주대학교 산학협력단	10-2020-0165925	2020	자체보유

[붙임5]

연구기반(인프라) 현황

1. 주관대학의 연구기반 지원 계획

가. 연구인력 및 연구기기 지원

1) 주관대학의 운영 및 행정 인력지원

ITRC 블록체인 지능 융합 센터장 이흥노 교수는 현재 GIST인공지능연구소 소장 보직을 수행하고 있고 제2세부과제 책임자 김종원 교수는 AI대학원 원장을 맡고 있으며, 연구소 내의 아래 인력을 센터 참여 교수님들의 인공지능 관련 GPU 컴퓨팅 니즈, 협업공간 활용의 장소 제공, 온라인 강의 제작 등을 지원하고자 함. 아래에서는 본 ITRC 센터 사업 선정시, 인공지능연구소에서 상기 내용의 지원을 위한 연구소 내 인력과 장비 현황임.

이름	업무경력	전공
이흥노	22	전자공학
공득조	5	전기전자 컴퓨터
이승윤	5.5	전자컴퓨터공학
이성호	15	전산학
이영환	1	제어계측

2) 주관대학의 장비 보유 현황

○ 광주과학기술원

연번	장비명	단위	수량	구입연도	내구연한	구매방식	단가		설치장소	활용과정	활용목표(시간)	인프라지원여부(자체/지원)	장비유형
							정부지원금	자체부담금					
1	DGX-A100	대	2	2020		조달입찰	230,043,000	–	EECS 108	AI-X 컴퓨팅	1,800	자체	실습용
	DGX-1 V	대	2	2019		조달입찰	124,586,000	–	EECS 108		1,800	자체	실습용
	Pure Storage	대	1	2020		조달입찰	99,550,000	–	EECS 108		1,800	자체	실습용
	Mellanox SN2100	대	3	2019 (1개), 2020 (2개)		조달입찰	11,543,000	–	EECS 108		1,800	자체	실습용
	Mellanox	대	1	2020		조달	7,730,000	–	EECS		1,800	자체	실습용

	QM 8700					입찰			108				
	SuperServer 1029U-TN10RT	대	2	2020		조달입찰	7,200,000	–	EECS 108		1,800	자체	실습용
	Susol SS 2029BZ-HNR	대	1	2020		조달입찰	27,200,000	–	EECS 108		1,800	자체	실습용
	APC smart-UPS	대	2	2020		조달입찰	8,400,000	–	EECS 108		1,800	자체	실습용
2	자율주행자동차KIT	set	10	2020		수의계약	726,000	-	Mobility Studio	모빌리티 스튜디오	1,800	자체	실습용
	비디오 프로젝터	대	1	2020		조달입찰	5,308,000	-	Mobility Studio		1,800	자체	실습용
	스크린	대	1	2020		조달입찰	1,256,750	-	Mobility Studio		1,800	자체	실습용
	자율주행트랙	set	1	2020		조달입찰	23,930,500	-	Mobility Studio		1,800	자체	실습용
	EDU Drone	set	15	2020		수의계약	150,000	-	Mobility Studio		1,800	자체	실습용
	미니PC	대	5	2020		수의계약	700,000	-	Mobility Studio		1,800	자체	실습용
	웹캠	대	15	2020		수의계약	42,000	-	Mobility Studio		1,800	자체	실습용
3	TV	대	1	2020		수의계약	1,590,000	-	Media Studio	미디어 스튜디오	1,800	자체	실습용
	TV거치대	대	1	2020		수의계약	42,000	-	Media Studio		1,800	자체	실습용
	조명	대	3	2020		수의계약	-	-	Media Studio		1,800	자체	실습용
	카메라	대	1	2020		수의계약	4,840,000	-	Media Studio		1,800	자체	실습용
	PC	대	1	2020		조달입찰	8,840,000	-	Media Studio		1,800	자체	실습용
	AV스위쳐	대	1	2020		수의계약	525,000	-	Media Studio		1,800	자체	실습용
	오디오믹서	대	1	2020		수의계약	820,000	-	Media Studio		1,800	자체	실습용
	삼각대	대	3	2020		수의계약	170,000	-	Media Studio		1,800	자체	실습용
	모바일짐벌	대	1	2020		수의계약	-	-	Media Studio		1,800	자체	실습용
	크로마키시스템	set	1	2020		수의계약	-	-	Media Studio		1,800	자체	실습용
	마이크	대	3	2020		수의계약	590,000	-	Media Studio		1,800	자체	실습용
4	2029U-TN24R	EA	2	2020	2030	조달입찰	25,025,000	·	GIST		1,800	자체	VDI기반 AI

	4T												교육용
	MSN2100-CB2F	EA	1	2020	2030	조달입찰	12,770,000	·	교육지원행정동 2층 서버실		1,800	자체	100G 스위치
	Sonic Firewall NSA3650	EA	1	2020	2030	조달입찰	7,000,000	·			1,800	자체	방화벽
	NETGEAR GS752TP	EA	1	2020	2030	직접구매	·	770,000			1,800	자체	관리용 스위치
	2029BZ-HNR	EA	1	2020	2030	조달입찰	40,500,000	·			1,800	자체	AI 교육용 스토리지 서버
	1029U-TN10RT	EA	2	2020	2030	조달입찰	16,170,000	·			1,800	자체	AI 교육용 서버 관리용
	DGX-1 V	EA	1	2017	2030	조달입찰	166,580,000	·			1,800	자체	AI 교육용 고성능 GPU
	DGX-A100	EA	1	2020	2030	조달입찰	205,329,300	·			1,800	자체	AI 교육용 고성능 GPU

○ AI-X 컴퓨팅 클러스터

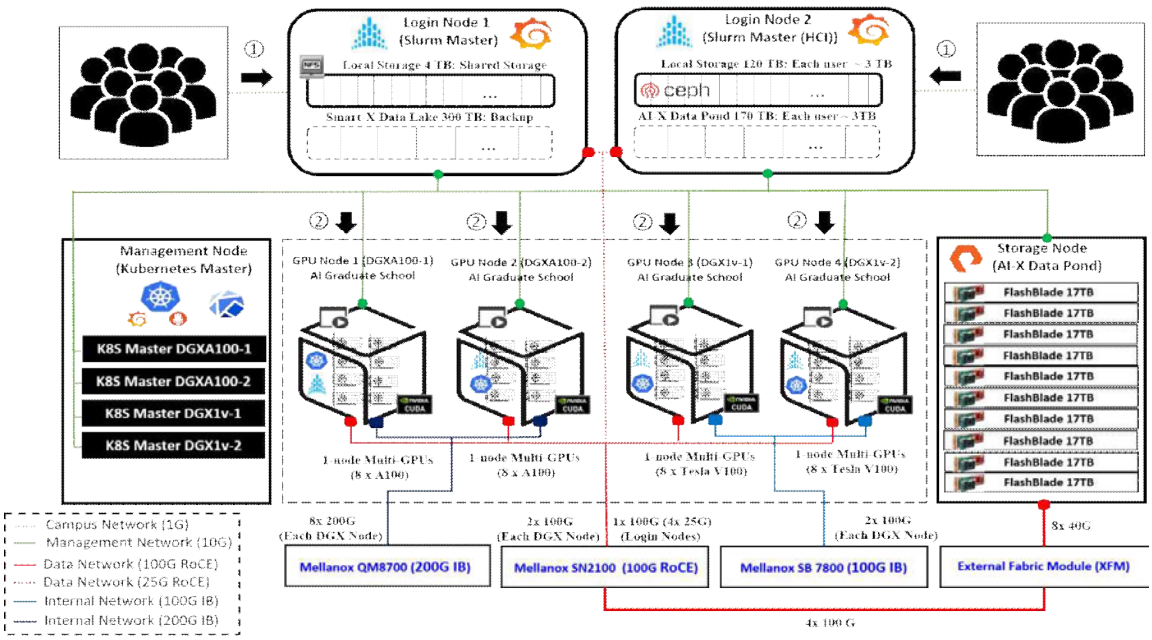


그림 26. AI-X 컴퓨팅 클러스터

- 장비 연번 1을 이용해 그림1에서 나타난 대로 GPU 8장과 200G 및 100G 급의 네트워킹 카드를 보유한 고성능 GPU 서버인 NVIDIA DGX-A100 2대 및

DGX-1V 2대, 일반 CPU 서버 4대, 스토리지 서버 1대, 그리고 고성능 스위치 3대를 이용해 DeepOps 기반 AI 컴퓨팅 클러스터 환경이 광주과학기술원에 구축

- 도입 장비들을 200Gbps 및 100Gbps 지원 고성능 스위치인 Mellanox Infiniband/RoCE 스위치에 연결하여 네트워킹 연동하고, 세계의 여러 슈퍼 컴퓨터에서 사용되는 Slurm 오픈 소스 작업 스케줄러 및 Kubernetes 기반의 클라우드-네이티브 기술을 동시에 지원하는 hybrid 클러스터 형태로 클라우드 환경을 구성, 초고속 데이터 저장 및 처리를 위해 구축한 시스템을 170TB 스토리지 박스 (Pure Storage)와 4x 100G 네트워크로 연동하여 제공

○ **모빌리티 스튜디오**

- 장비 연번 2을 이용해 자율주행 로봇 및 다객체 강화학습 검증을 위한 스튜디오로 자율주행자동차인 DeepRacer 장비를 구매하고 커스터 마이징 트랙을 제작하여 로봇실험교과목 및 창의자율프로젝트를 위한 플랫폼, 서비스, 동작 테스트를 지원
- DeepRacer는 실제 자동차의 1/18 크기로 제작된 자율주행 레이싱 카로, 인텔 Atom 프로세서와 트랙을 확인할 수 있는 400만 화소 카메라를 탑재했으며, Ubuntu OS와 인텔의 컴퓨터 비전 솔루션인 Open VINO 툴킷을 지원

○ **미디어 스튜디오**

- 장비연번 3을 이용해 1인 미디어 콘텐츠 제작 및 비대면 강의 녹화 등의 크리에이터를 위한 맞춤 공간으로 미디어 방송을 위한 1인용 촬영 및 조정 장비 완비 및 연출, 편집을 지원

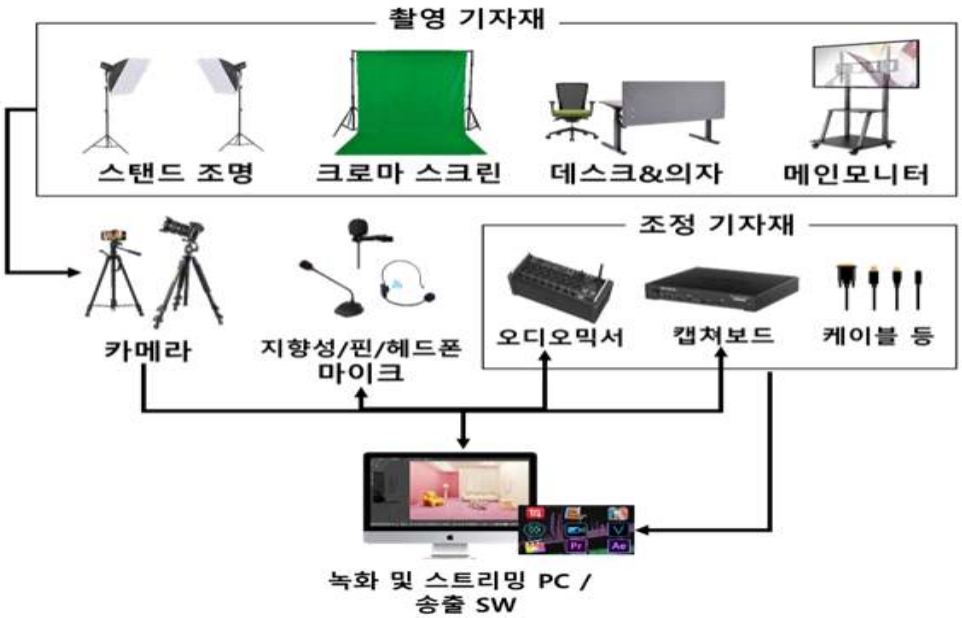


그림 27. 미디어 스튜디오 구성도

나. 센터 전용공간 지원 (센터 설치 후 6개월 이내에 확보)

주관대학에서는 GIST 인공지능연구소, GIST블록체인경제센터, GIST AI대학원 등에서 센터 전용공간을 아래와 같이 지원할 예정임.

구분	면적(m^2)	위치(예정)	공간 확보를 위한 대학차원의 조치사항
연구실/실습실	833.0m^2	AI대학원	
강의/공동기기실	147.4m^2	GIST 금호관 2층	
센터장사무실	23.1m^2	전기전자컴퓨터공학부 C동 205호	
합 계	1003.5m^2		

○ 시설 보유 현황

연번	시설 구분	시설명	지원 일자	정부 지원금	자체 부담금	위치	면적	수용 인원
1	강의실	TED 홀	2020			지스트 AI대학원 교육연구지원동	170m^2	100명
	실습실	AI 스튜디오					167m^2	
	실습실	모빌리티 스튜디오					153m^2	
	실습실	미디어 스튜디오					73m^2	
	회의실	회의실 /기타					270m^2	
	합계						833m^2	
2	강의실 (실습실)	SW 지원센터 1	2017			GIST 금호관 2층	75.6m^2	30명
3	강의실 (실습실)	SW 지원센터 2	2017			GIST 금호관 2층	71.8m^2	30명
합계							980.4m^2	

○ AI대학원 교육연구지원동 시설

- 지스트(GIST) AI대학원 전용공간으로 건평 2,006m^2(연면적 2,244m^2) 단독건물
- 1층 공간에는 TED형 강의실 , AI스튜디오, 미디어스튜디오, 회의실이 구축되어 있으며, 교육 및 실습공간으로 연면적 833m^2 공간 활용



그림 28. AI 대학원 교육연구지원동

○ AI대학원 교육연구지원동 교육·연구 실습 환경

- 지스트 AI대학원 교육연구지원동(S7)에는 창의활동 및 특화교육을 위한 AI 스튜디오, 자율주행 실습 환경인 모빌리티 스튜디오, 1인 미디어를 제작할 수 있는 미디어 스튜디오 등이 조성

- AI 스튜디오

• 본 AI대학원 교육연구지원동 1층에 위치한 AI 스튜디오는 교육생에게 제공되는 공간으로 AI 전문교육이 진행되며, 교육과 실습에 필요한 실습 장비 구축

[표] AI 스튜디오 장비 현황

	일체형 Windows PC 6대	일체형 iMAC 10대
CPU	Intel i5-10210U 1.60GHz	6코어 intel Core i5 3.3GHz
OS	Windows10 64bit	MAC OS (iMAC Retina 5K, 27-inch, 2020)
RAM	8GB	8GB 2667 MHz DDR4
SSD	256GB	1TB
GPU	GTX 1050	AMD Radeon Pro 5300 4GB

- 미디어(Media) 스튜디오

- 1인 미디어 콘텐츠 제작 및 강의 녹화 등의 크리에이터를 위한 맞춤 공간으로 방송을 위한 1인용 촬영 및 조정 장비 완비 및 연출을 지원하는 공간을 제공



그림 29. AI 스튜디오, 미디어 스튜디오

- 모빌리티(Mobility) 스튜디오

- 자율주행 로봇 플랫폼 중 AWS의 DeepRacer를 이용하여 다객체 강화학습 알고리즘을 검증하는 실습 환경 제공
- 창의 자율연구 환경 및 모빌리티 기술 지원 공간을 제공



그림 30. 모빌리티 스튜디오

- TED Hall

• 지식 나눔을 목표로 다양한 이벤트 및 컨퍼런스 행사를 개최할 수 있는 개방형 로비 교육장



그림 31. TED 홀

○ SW 교육센터 교육실

- AI+X SW 교육환경으로는 70여 제곱미터, 30대 규모의 PC 기반 SW 교육실 2개실을 활용 예정.

[표] SW 교육센터 교육실 장비 현황

	PC 60대
CPU	Intel i7-8700 3.20GHz
OS	Windows10 64bit
RAM	8GB
SSD	256GB
HDD	1TB



그림 32. S/W 교육실

○ HPC(High performance computing)센터 (구축 예정)

- 130억 원 규모의 HPC센터 구축 후 센터 전용공간 확보 예정
- 6PF급의 하이엔드 HPC(NVIDIA SuperPOD 2기) 도입 예정(21.09)



그림 33. High Performance Computing 센터 (구축 예정)

다. 연구센터장(연구책임자)에 대한 지원

1) 강의시간 및 보직

구분	강의시간	보직명	업적평가 가산점	비고
현재	3시간(주당)	인공지능연구소 소장	없음	임기 :2021.1.18.~2023.1.17
변경	3시간(주당)	인공지능연구소 소장	가산점 10% 부여	

* 신청일 현재 보직을 맡고 있는 경우, 보직명과 임기를 반드시 기재바람

2) 기타 지원계획

라. 주관대학의 기타지원

* 운영인력(행정요원 등) 지원 등 대학자체에서 설정한 지원 내용

2. 출연금 및 기관부담연구개발비 배분내역(총 연구기간)

구 분			광주과학기술원	비율 (%)	연구개발기관 외 기관의 지원금	비율 (%)	계	비율 (%)
정부출연금								
민 간 부담금	1년차	현금						
		현물						
	2년차	현금						
		현물						
	3년차	현금						
		현물						
	4년차	현금						
		현물						
	5년차	현금						
		현물						
	6년차	현금						
		현물						
	7년차	현금						
		현물						
	8년차	현금						
		현물						
합계								

[붙임6]

[붙임6]

사업 전용공간 확보계획서

사 업 명	대학ICT연구센터육성지원		
센 터 명	블록체인 지능 융합센터		
주 관 기 관	광주과학기술원	연구책임자	이흥노

위 사업수행을 위한 사업계획서의 사업내용에 동의하고, 본 사업의 수행을 위한 전용공간을 관련법령 및 제반사항에 의거하여 제공할 것을 확약합니다.

전 용 공 간	
주소 및 위치	광주과학기술원(광주광역시 북구 첨단과기로 123) AI대학원 교육연구지원동, 전기전자컴퓨터공학부 C-205호 및 금호관 2층 SW교육센터
전용공간 면적	(1003.5m²)

2021년 4월 28일

주관연구개발기관장 총장직무대행 김인수 (인)

정보통신기획평가원장 귀하

[붙임7]

[붙임7]

기관부담연구개발비(주관/공동연구개발기관) 출자 납입약서

□ 과제연왕

(단위 : 천원)

사 업 명	대학ICT연구센터육성지원				
센 터 명	블록체인 지능 융합센터				
주관대학	광주과학기술원	연구책임자	이홍노		
총 연구비	구분	정부	주관대학	기업체	지자체 등	계
	현금	[redacted]				
	현물					
총 계						

□ 기관부담연구개발비 납입 계획

(단위 : 천원)

기 관 명	1년차		2년차		3년차		4년차		5년차		6년차		7년차		8년차		합계	
	현금	현물	현금	현물	현금	현물	현금	현물	현금	현물	현금	현물	현금	현물	현금	현물	현금	현물
광주과학기술원	[redacted]																	
합계																		

위와 같이 정보통신·방송 연구개발 관리규정에 따라 **수행기관이 부담하여야 하는 기관부담연구개발비를 성실히 납부할 것을 확인(확약)**하며, 이를 이행하지 않을 경우 협약의 해약, 국가연구개발사업 참여제한 등의 제재조치를 감수할 것을 확약합니다.

2021년 4월 28일

연구책임자 : 이홍노 (인)

개발기관명 : 총장직무대행 김인수 (인) 직인/인감

정보통신기획평가원장 귀하

[붙임7-1]

[붙임7-1 : 해당시 작성]

연구개발기관 외 기관의 지원금 납입 확약서

□ 과제현황

(단위 : 천원)

사 업 명		대학ICT연구센터육성지원				
센 터 명		블록체인 지능 융합센터				
주관대학		광주과학기술원		연구책임자	이홍노	
총 연구비	구분	정부	주관대학	기업체	지자체 등	계
	현금	[redacted]				
	현물					
총 계						

□ 연구개발기관 외 기관의 지원금 납입 계획

(단위 : 천원)

기 관 명	1년차		2년차		3년차		4년차		5년차		6년차		7년차		8년차		합계	
	현금	현물	현금	현물	현금	현물	현금	현물	현금	현물	현금	현물	현금	현물	현금	현물	현금	현물
광주광역시청	[redacted]																	
합계																		

위와 같이 연구개발기관외 기관의 지원금을 납부할 것을 확인(확약)하며, 이를 이행하지 않을 경우 협약의 해약, 국가연구개발사업 참여제한 등의 제재조치를 감수할 것을 확약합니다.

2021년 4월 28일

연구책임자 : 이홍노 (인)

개발기관명 : 총장직무대행 김인수 (인) - 직인/인감

정보통신기획평가원장 귀하

[붙임8]

참여기관 현황(총괄표)

구 분	구분	기 관 명	위치	기관장명	고용보험가입여부 (사업장 등록번호)
대 학	주관	광주과학기술원	광주광역시	총장직무대행 김인수	가입 (418-02-07550)
	참여	고려대학교	서울특별시	정진택	가입 (209-82-00433)
	참여	아주대학교 산학협력단	경기도 수원시	권용진	가입 (124-82-14602)
	참여	서울대학교 산학협력단	서울특별시	최해천	가입 (119-82-03684)
기 업 체					
연 구 소					
기타기관					
합 계					

[붙임9]

[붙임9]

전임연구원 채용 확약서

사 업 명	대학ICT육성지원		
센 터 명	블록체인 지능 융합센터		
주 관 기 관	광주과학기술원	연구책임자	이흥노

위 사업수행을 위해 전임연구원 채용을 이행할 것을 확약합니다

채용 구분	채용유형 (전임, 비전임)	채용여부			역할	참여율 (%)
		채용	채용 예정	확약일		
의무	전임		O	2021.07.01.	연구실무자 및 학생창의자율과제 전담 지도요원	50
추가						
추가						

증빙. 전임연구원 채용 증명 1부 (재직증명서 등)

2021년 4월 28일

주관연구개발기관장 총장직무대행 김인수 (인)

정보통신기획평가원장 귀하

*

[붙임10]

신규 참여연구원 채용 확인서(확약서)

* 공동연구개발기관용

□ 과제현황

사 업 명			
센 터 명			
주 관 기 관		연구책임자	

□ 신규 참여연구원 인적사항

성 명			(인)
생년월일 (과학기술인등록번호) 예시:101210 (1234 5678)			
소속 및 직위			
담당업무 / 전공		채용일(예정일)	
본 과제 참여역할			
주소 / 연락처	/		
참여기간	20 . . . ~ 20 . .	참 여 율	

상기인은 과학기술정보통신부에서에서 시행하는 대학ICT연구센터육성지원사업 과제 수행을 위하여 신규로 채용하였음을 확인합니다. (채용할 것을 확약합니다.)

20 . . .

공동연구개발기관명 : (인)

공동연구개발기관대표자 : (인)

연구책임자 (인)

정보통신기획평가원장 귀하

별첨 : 고용계약서 사본 1부.

* 채용예정인 경우 성명 및 생년월일은 추후 작성하여 제출

[붙임11]

시설장비 심의요청서 (3천만원이상 ~)

□ 연구시설 · 장비의 개요(1단계 사업기간 내에 해당사항이 있을시 작성)

<table>
<tr><td colspan="2">구 분</td><td colspan="8">내 용</td></tr>
<tr><td colspan="2">과제명</td><td colspan="8"></td></tr>
<tr><td rowspan="2">시설장비명</td><td>한글</td><td colspan="8">※ 연구시설·장비 국문 명칭을 기재</td></tr>
<tr><td>영문</td><td colspan="8">※ 연구시설·장비 영문 명칭을 기재</td></tr>
<tr><td colspan="2" rowspan="2">담당자</td><td colspan="2">소속</td><td colspan="2">이름</td><td colspan="2">연락처</td><td>이메일</td></tr>
<tr><td colspan="2"></td><td colspan="2"></td><td colspan="2"></td><td></td></tr>
<tr><td colspan="2" rowspan="2">제작사 및 모델명
(입찰예정이면 제작사 및 모델명을 2개 이상 작성)</td><td colspan="2">제작국가명</td><td colspan="2">제작사명</td><td colspan="4">모델명</td></tr>
<tr><td colspan="2"></td><td colspan="2"></td><td colspan="4"></td></tr>
<tr><td colspan="2" rowspan="2">취득방법
(해당란에 '○'표시)</td><td>구 매</td><td>임 대</td><td>제작의뢰</td><td>자체제작</td><td colspan="5">기 타(직접 기재)</td></tr>
<tr><td></td><td></td><td></td><td></td><td colspan="5"></td></tr>
<tr><td colspan="2" rowspan="3">구축비용
(단위 : 백만 원)</td><td rowspan="2">단가</td><td rowspan="2">수량</td><td rowspan="2">총금액</td><td rowspan="2">'00년 정부출연금 신청금액</td><td rowspan="2">'00년 자체부담 금액
(매칭펀드로 구축하는 경우)</td><td rowspan="2">적용환율
(외자일 경우)</td><td colspan="2">연도별 분할납부 금액 및 임대료
(분할납부예정 또는 임대일 경우)</td></tr>
<tr><td rowspan="2"></td><td rowspan="2"></td><td rowspan="2"></td><td>'00년</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>'00+1년</td></tr>
<tr><td rowspan="2">클라우드컴퓨팅 서비스 이용검토</td><td>가능</td><td colspan="8">예상이용료(xx백만원/연간)</td></tr>
<tr><td>불가</td><td colspan="8">불가사유기재</td></tr>
<tr><td colspan="2">구축일정</td><td colspan="4">발주예정일</td><td colspan="4">설치예정일</td></tr>
<tr><td colspan="2"></td><td colspan="4">YYYY-MM-DD ~ YYYY-MM-DD</td><td colspan="4">YYYY-MM-DD ~ YYYY-MM-DD</td></tr>
<tr><td colspan="2" rowspan="2">구축장소
(수량별 구축장소가 다른 경우 구분하여 작성)</td><td colspan="2">설치예정 지역명</td><td colspan="2">설치예정 기관명</td><td colspan="4">설치예정 세부 장소(건물명 등)</td></tr>
<tr><td colspan="2"></td><td colspan="2"></td><td colspan="4"></td></tr>
<tr><td colspan="2" rowspan="2">시설장비 용도</td><td colspan="8">○
-
※ 장비의 측정 목적, 피시험물, 취득하고자 하는 결과물 등 자세하게 기재</td></tr>
<tr><td>분석</td><td>시험</td><td>교육</td><td>계측</td><td>생산</td><td colspan="3">기타</td></tr>
<tr><td colspan="2"></td><td>(해당란에 '○'표시)</td><td></td><td></td><td></td><td></td><td colspan="3">※ 직접기재</td></tr>
<tr><td colspan="2">주요사양</td><td colspan="8">○
-
※ 제작사가 제공하는 주요 사양을 5가지 이상 기재
※ 심의위원들이 판단할 수 있게 사양을 구체적으로 자세하게 기재. 품목의 특성 및 성능을 구체적으로 기재</td></tr>
<tr><td colspan="2">외산장비 도입 필요성</td><td colspan="8">○
-
※ 제작사가 외국기업인 경우 작성</td></tr>
</table>

□ 연구시설·장비 구축의 목적 및 내용

<table>
<tr><th>구 분</th><th>내 용</th></tr>
<tr><td>사업(연구)
부합성</td><td>○
-

※ 신청 장비 도입이 본 사업(연구) 내용 중 어떤 부분과 연관성이 있는지 기술
※ 사업(연구) 수행에 반드시 필요한 장비인지 기술</td></tr>
<tr><td>연구장비의
중복성</td><td>○
-

※ 동일기관, 타기관에서 해당장비와 동일하거나 유사한 장비를 이미 보유하고 있는지 여부를 기술</td></tr>
<tr><td>연구장비의
활용성</td><td>○
-

※ 동 사업(연구)에서 활용 계획 및 방법 작성
※ 동 사업(연구)에서 활용도가 높은 장비인지 기술. 해당사업(연구) 종료 후 타 사업(연구)에서도
활용이 가능한 장비인지 기술
※ 구축 후 타기관과의 공동활용이 가능한 장비인지 기술. 가능한 경우 주요활용 기관명(예상)을 작성</td></tr>
<tr><td>연구장비의
적정성</td><td>○
-

※ 연구목적 달성을 위해 적합한 구성(Specifications) 및 성능(Performance)의 장비인지 기술
※ 신청한 연구시설·장비 가격의 적정성에 대하여 기술(기구축 동일 장비 가격, 타 제작사 장비 가격과 비교하는 등)
※ 신청 수량이 2개 이상인 경우 본 연구 관련하여 신청 수량만큼 필요한 타당한 이유를 기술</td></tr>
<tr><td>장비운영의
계획성</td><td>
<table>
<tr><th colspan="6">신청 시설장비의 전문기술인력 확보 현황(계획)</th></tr>
<tr><th>구분
(신규, 기존)</th><th>성명
(채용예정자는 OOO)</th><th>소속부서명</th><th>최종학위
(고졸, 학사, 석사, 박사)</th><th>고용형태
(정규직, 계약직)</th><th>담당장비수
(신청장비 포함)</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>
○
-
※ 신청한 시설장비의 구축과 운영을 위한 설치공간 확보방안을 기술
※ 신청한 시설장비의 운영비(운영인력 인건비, 유지보수비 등) 확보방안을 기술
※ 연구과제(사업) 종료 후의 운영(활용) 계획을 기술</td></tr>
</table>

2021년 대학ICT연구센터육성지원사업

- 자문·협력의사확인서 -

사업명	2021년 대학ICT연구센터육성지원사업		
센터명	블록체인 지능 융합 연구센터		
주관기관	광주과학기술원	연구책임자	이흥노 교수

당사에서는 위와 같이 주관연구기관이 수행하는 2021년 대학ICT연구센터육성지원사업을 성공적으로 수행하기 위해

- 센터 연구방향, 인력양성방안 및 세미나 등의 강사 섭외 등을 위한 적극 자문 등의 노력과,
- 석·박사과정생 인력양성을 위한 교육, 연구, 졸업생 취업 및 창업 연계 등을 위한 자문위원 역할 수행
- 기타 사업성공을 위한 주관기관과의 공동의 노력 등

상기 기술된 내용에 자문 및 협력할 의사가 있음을 확약합니다.

2021년 4월 23일

기업체(단체)명 한국ICT융합협회

작성자 백양순 (인)

2021년 대학ICT연구센터육성지원사업

- 자문 · 협력의사확인서 -

사업명	2021년 대학ICT연구센터육성지원사업		
센터명	블록체인 지능 융합 연구센터		
주관기관	광주과학기술원	연구책임자	이흥노 교수

당사에서는 위와 같이 주관연구기관이 수행하는 2021년 대학ICT연구센터육성지원사업을 성공적으로 수행하기 위해

- 센터 연구방향, 인력양성방안 및 세미나 등의 강사 섭외 등을 위한 적극 자문 등의 노력과,
- 석·박사과정생 인력양성을 위한 교육, 연구, 졸업생 취업 및 창업 연계 등을 위한 자문위원 역할 수행
- 기타 사업성공을 위한 주관기관과의 공동의 노력 등

상기 기술된 내용에 자문 및 협력할 의사가 있음을 확약합니다.

2021년 4월 23일

기업체(단체)명 서울블록체인지원센터

작성자 임 명 수 (인)

2021년 대학ICT연구센터육성지원사업

- 자문 · 협력의사확인서 -

사업명	2021년 대학ICT연구센터육성지원사업		
센터명	블록체인 지능 융합 연구센터		
주관기관	광주과학기술원	연구책임자	이흥노 교수

당사에서는 위와 같이 주관연구기관이 수행하는 2021년 대학 ICT연구센터육성지원사업을 성공적으로 수행하기 위해

- 센터 연구방향, 인력양성방안 및 세미나 등의 강사 섭외 등을 위한 적극 자문 등의 노력과,
- 석·박사과정생 인력양성을 위한 교육, 연구, 졸업생 취업 및 창업 연계 등을 위한 자문위원 역할 수행
- 기타 사업성공을 위한 주관기관과의 공동의 노력 등

상기 기술된 내용에 자문 및 협력할 의사가 있음을 확약합니다.

2021년 4월 23일

엘에스웨어(주) 연구소장 신동명

2021년 대학ICT연구센터육성지원사업

- 자문 · 협력의사확인서 -

사업명	2021년 대학ICT연구센터육성지원사업		
센터명	블록체인 지능 융합 연구센터		
주관기관	광주과학기술원	**연구책임자**	이홍노 교수

당사에서는 위와 같이 주관연구기관이 수행하는 2021년 대학ICT연구센터육성지원사업을 성공적으로 수행하기 위해

- 센터 연구방향, 인력양성방안 및 세미나 등의 강사 섭외 등을 위한 적극 자문 등의 노력과,
- 석·박사과정생 인력양성을 위한 교육, 연구, 졸업생 취업 및 창업 연계 등을 위한 자문위원 역할 수행
- 기타 사업성공을 위한 주관기관과의 공동의 노력 등

상기 기술된 내용에 자문 및 협력할 의사가 있음을 확약합니다.

2021년 4월 23일

기업체(단체)명 퓨처센스

작성자 안 다 미 (인)

2021년 대학ICT연구센터육성지원사업

- 자문·협력의사확인서 -

사업명	2021년 대학ICT연구센터육성지원사업		
센터명	블록체인 지능 융합 연구센터		
주관기관	광주과학기술원	**연구책임자**	이흥노 교수

당사에서는 위와 같이 주관연구기관이 수행하는 2021년 대학ICT연구센터육성지원사업을 성공적으로 수행하기 위해

- 센터 연구방향, 인력양성방안 및 세미나 등의 강사 섭외 등을 위한 적극 자문 등의 노력과,
- 석·박사과정생 인력양성을 위한 교육, 연구, 졸업생 취업 및 창업 연계 등을 위한 자문위원 역할 수행
- 기타 사업성공을 위한 주관기관과의 공동의 노력 등

상기 기술된 내용에 자문 및 협력할 의사가 있음을 확약합니다.

2021년 4월 23일

기업체(단체)명 (주)헥슬란트

작성자 노정두 (인)

2021년 대학ICT연구센터육성지원사업

- 자문 · 협력의사확인서 -

사업명	2021년 대학ICT연구센터육성지원사업		
센터명	블록체인 지능 융합 연구센터		
주관기관	광주과학기술원	연구책임자	이흥노 교수

당사에서는 위와 같이 주관연구기관이 수행하는 2021년 대학ICT연구센터육성지원사업을 성공적으로 수행하기 위해

- 센터 연구방향, 인력양성방안 및 세미나 등의 강사 섭외 등을 위한 적극 자문 등의 노력과,
- 석·박사과정생 인력양성을 위한 교육, 연구, 졸업생 취업 및 창업 연계 등을 위한 자문위원 역할 수행
- 기타 사업성공을 위한 주관기관과의 공동의 노력 등

상기 기술된 내용에 자문 및 협력할 의사가 있음을 확약합니다.

2021년 4월 23일

기업체(단체)명 ㈜시티랩스

작성자 김항진 (인)

스마트 포용도시 시범구현 사업제안(안)

GIST블록체인경제센터
국정기획위원회 제출본

제출자: 이흥노 국정기획위원회 전문위원

2021.3.18.

<table>
<tr><td colspan="2">분야</td><td>디지털、안전망、지역균형</td></tr>
<tr><td colspan="2">책임제안자
실무자</td><td>이흥노 GIST교수, 인공지능연구소장/블록체인경제센터장
이명은 블록체인경제센터 대외협력팀장</td></tr>
<tr><td colspan="2">사업명(과제명)</td><td>스마트 포용사회 실증
제도개선과 인공지능과 블록체인 기술 적용으로 근로가치가 고취되는 스마트 포용도시 건설</td></tr>
<tr><td colspan="2">신설/보완 구분</td><td>신설</td></tr>
<tr><td rowspan="3">내용</td><td>필요성</td><td>○ 극심한 양극화 갈등사회
- 돈의 의미 퇴색과 자산인플레에 대응할 저축수단의 부재, 근로가치 절하, 부의 양극화, 중산층 붕괴, 지방 소멸, 수도권 집중 가속화, 저출산/인구절벽 가속화, 국가 유지비용 증가
○ 지식추구 사회로 가치 전환이 절실함
- 부의 원천은 지식. 즉, 연구·창업·근로를 통한 지식생산활동이 부를 새롭게 생산하는 활동임. 근로 가치 절상, 중산층 증가, 균형 발전 등으로 사회가치를 전환시킬 필요성이 매우 큼.</td></tr>
<tr><td>주요
핵심내용</td><td>○ 포용 스마트도시 건설 실증
- 현재 지방에서 추진되고 있는 스마트도시 조성사업은 인공지능, 자율주행, IoT, 5G 등 자동화 기반 성장과 효율 추구 중심. 상기한 양극화사회 구조속에서 우수인재의 참여미흡과 지역인재의 수도권 유출이 가속화하는 실정임. 지역 내에서는 자동화에 따른 일자리 소멸과 양극화가속 우려 또한 존재.
- 스마트도시 조성 시 포용정책을 내포한 포용전환 기술 개발과 실증·적용 사업을 제안하고자 함. 발전토큰(도시가치연동저축)을 발행하고, 참여자에게 제공하여 우수한 인재의 조기 참여를 촉진하고자함. 발전토큰은 조성 지역내 대지·부동산 지분과 연동함. 토큰 소유자는 자산인플레에 대응하는 헤징수단을 갖게 됨. 일하는 사람이 성장한 스마트도시에서 젠트리피케이션 당하지 않고 주인으로 성장함. 조성초기 불확실성을 뛰어넘고 도시발전을 위해 초기참여 및 헌신할 동기를 제고함. 다양한 인재의 참여와 헌신으로 스마트도시사업의 성공가능성을 제고하고, 지식생산활동과 근로가치를 고취, 국토균형발전을 추동함.</td></tr>
<tr><td>법、제도화
및
거버넌스</td><td>○ 스마트도시조성법, 조성 지역이 속한 지자체의 조례 등을 개정하여 발전토큰 발행총량과 도시지가상승분과 연동성 강화.</td></tr>
</table>

1 사업개요

오늘날 전 세계적인 현상은 부의 양극화임. 선진국의 양적완화가 주요 원인으로 밝혀짐. 4차 산업혁명의 여파로 빠르게 고용구조가 변화하고 있음. 수도권-비수도권 불균형이 악화되고 있음. 고령사회 구조에서, 근로의지마저 약화되고 있음.

나라의 지속성장을 위해서는 스마트 포용사회로 전환이 필요함. 그러나, 스마트(성장 및 효율)은 매우 강조되고 있는 반면 동반성장을 위한 정책개발은 상대적으로 미흡하고 합의를 이루기 어려운 상황임. 현재 논의되는 정책은 일자리나누기, 기본소득, 선별지급, 최저임금 등 주요국의 실험에서도 실효성 논란이 큰 실정이기 때문임.

GIST는 이 문제를 “일하는 자가 부자가 되는 나라”라는 비전을 제시하고 형국의 전환을 시도하고자 함. 특히, 교육과 연구 활동을 중시하고, 새로운 일에 도전하고 사회문제를 지식생산의 힘으로 해결하며, 창업을 통해 제품과 서비스를 만드는 일에 종사하는 분들을 통칭하여 지식생산자라 명하고 이들의 근로 가치를 높게 평가하고 지원하는 정책을 개발하고 집행하고자 함.

지방에 조성되는 스마트도시 사업은 AI/데이터/IoT/센서/5G 등 첨단기술의 집약체임. 계획대로 구현되고 시민이 참여하면, 균형 발전, 안전사회, 친환경 도시, 고령사회 등 여러 문제를 한꺼번에 해결할 수 있음.

문제는 시민의 참여임. 참여를 독려하고, 지식생산자 사회를 촉진하기 위해, 발전토큰과 지대상승분을 연동하여, “일하는 자가 주인이 되는 나라”를 기획, 설계, 실행하고, 성장하면서도 지속 가능한 스마트 포용타운 건설의 시범 사례를 구축하고자 함.

2 배경: 사회 양극화 극심에 따른 국가적 해결방안 모색 필요

□ 양적완화에 따른 경제적 불평등 악화

○ **(양극화 심화)** 유럽중앙은행은 2008년 금융위기 이후 약 10년 간 전개된 양적완화 규모의 절반 이상이 코로나 19 사태 수습에 동원됨에 따라 **절대적인 자원 배분의 비효율성과 양극화 심화를 초래**할 수 있다는 연구결과를 발표. 포스트 코로나 시대, 양적완화에 따른 실물과 시장 간 괴리로 인해 **취약층을 중심으로 근로·사업소득이 감소하는 상황**에서 자산가들의 재산소득은 증가하는 이른바 ***'코로나 디바이드(Corona Divide)'*로 이어질 가능성** 증대

- 대표적 자산시장 중 하나인 부동산 시장의 경우, 이러한 자산시장의 과열을 직접 보여주는 양상으로 특히 서울과 지방권 아파트 가격 격차는 상승하는 반면, 수도권의 아파트 매매가격은 상향 평준화를 이룰 것으로 전망

- 소득과 자산 불평등과 더불어 수도권 인구가 전체 인구의 50%를 돌파하는 등 수도권 집중 현상이 강화되고 있어 균형 발전을 위한 대책 마련 마련이 시급

○ **(중산층 붕괴)** 기획재정부가 '19년도 발표한 「중산층 기준 및 최근 중산층 소득개선 현황」 자료에 따르면 **중위 소득 부근의 인구비중 하락은 OECD 등 주요국에서 동일하게 나타나는 전세계적 현상**이며, 세계화 및 자동화 등에 따른 **중간층 일자리 감소 등으로 저소득층과 고소득층 비중은 증가하고 중간계층 비중은 감소**

- 최근 국립재난안전연구원은 **포스트 코로나 시대엔 중산층이 붕괴할 것으로 전망**한 바 있음. 전례없는 위기로 발생한 **K자 양극화로 산업 간, 계층 간 차이가 더 벌어질 것으로 예측**됨에 따라, 중산층의 붕괴 문제를 해결하고 중산층 기반 확충을 위한 대책 마련이 시급

□ 수도권-비수도권 불균형 악화

○ **(지리적 문제점)** 개발연대 한국 경제를 지배하는 패러다임은 도시화, 산업화를 통한 고도성장의 추구였음. **투기적 도시화로 인한 부의 성장**은 수도권 부동산, 대기업의 탄생, 주식 등 금융자산의 가격 상승으로 이어지며 경제적 성장의 모든 축이 서울을 중심으로 한 수도권에 집중됨. 개발의 효과가 다른 지역으로 파급되지 않고 오히려 역류 효과가 발생하여 지방의 중소 도시나 농어촌 지역은 성장이 침체, 이에 따라 **수도권으로의 인구 이동이 불균형적으로 상승**

- 인구정책 태스크포스는 수도권 쏠림 현상 문제 해결을 위한 지역 소멸 선제 대응방안으로 **수도권 거주 인구와 기업의 지방 이전 촉진을 위한 유인을 강화하고 통합 지원 프로그램을 개발**할 것을 발표

□ 4차 산업혁명과 고용사회 변화

○ **(일자리 충격)** 인공지능 발전과 비대면 및 자동화에 따라 인공지능 서비스 로봇으로 기존 일자리가 대체되고 있으며 기술의 발전에 따라 **일자리 판이 변화**하고 있음. 세계경제포럼(WEF)은 「2020 일자리의 미래」 보고서에서 **2025년까지 행정, 회계, 제조업 등의 분야에서 8,500만 개의 일자리가 기계·기술로 대체**될 것으로 전망함.

- 현재 한국 사회는 저성장, 저출산, 고령화 사회로 나아가고 있어서 일의 생산성을 향상할 수 있는 인공지능의 활용능력에 따라 삶의 질에도 많은 변화가 있을 것으로 전망됨.

3 대응방안: 양극화 해소 · 사회정상화를 위한 포용사회 구현

□ 스마트 포용사회로의 변화 필요

○ UN Habitat Ⅲ은 포용도시를 *"현재 및 미래세대의 모든 거주민이 어떤 종류의 차별도 받지 않고 거주하고 생산할 수 있는 정의롭고, 안전하며, 건강하고, 접근성이 좋고, 회복력 있고, 지속 가능한 도시"*로 규정. **사회가 경제적·사회적 발전을 지속하면서 사회적 신뢰와 협력을 이어가기 위해서는 이러한 포용도시로의 변화가 필요**

○ 스마트시티는 도시의 여러 데이터를 활용해 디지털 인프라와 4차 산업혁명 기술을 도시 설계, 개발, 운영 및 유지관리의 기반으로 둠으로써 물리적 환경과 디지털 환경을 통합한 살아있는 생활환경임. 그러나 **포스트 코로나 시대의 양극화 완화와 생활환경의 안전, 거주자의 경제활동 보장을 위해서는 단순히 물리적 환경과 디지털을 통합하는 것을 넘어 포용성과 회복성을 포함**해야 함. 이를 '스마트 포용사회'로 명명함.

□ 포용적 성장을 위한 혁신 - 프로토콜 경제

○ 코로나 19는 **인공지능(AI), 블록체인(Blockchain), 클라우드(Cloud), 데이터(Data)와 같은 기술이 이끄는 새로운 뉴노멀을 형성**. 코로나 19의 확산으로 모든 산업과 생활양식의 변화는 **비대면 디지털 전환을 초래**했으며 특히 전 세계적으로 디지털 경제활동을 가속화함.

○ 언택트에 기반하여 급성장하기 시작한 플랫폼 경제는 데이터 독점과 폐쇄성 문제 등 여러 사회적 문제를 발생시키기 시작함. 이를 극복할 수 있는 **차세대 경제모델로 프로토콜 경제가 각광**을 받기 시작

- 프로토콜 경제는 플랫폼 경제에 대한 문제의식으로 생겨난 새로운 **경제개념으로 시스템이 데이터를 교환하기 위해 사용하는 규약**을 의미. 2021년 정부의 경제정책 방향에도 프로토콜 경제 발전전략 수립이 과제로 담김.

<참고> 프로토콜 경제 사례

○ 국내 사례

- KT: 블록체인 기반 '할랄 인증 Trust 플랫폼*' 구축

* 할랄 인증 처리와 발급, 검증 등 전 과정을 시스템화하여 공개

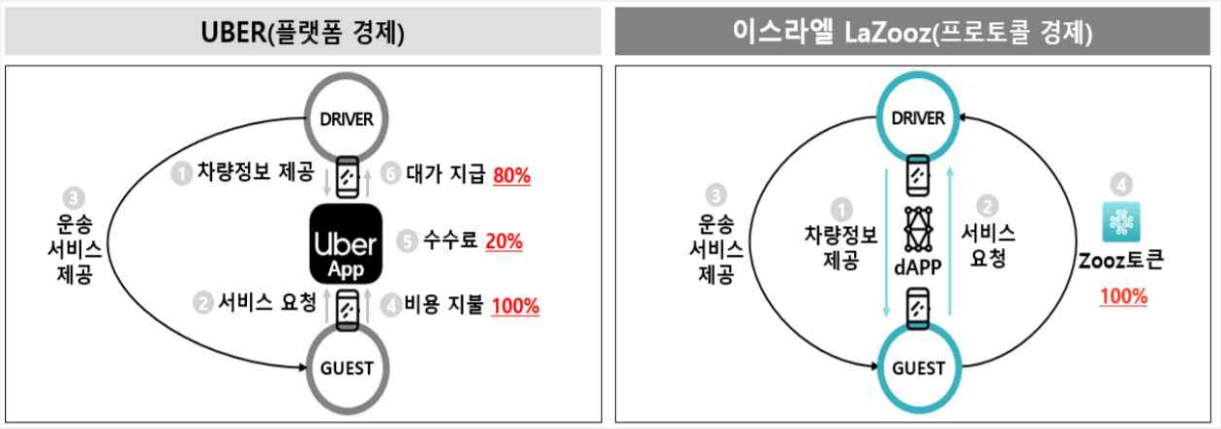
- 쿠팡: 배송인력인 '쿠팡맨' 등 직원들에게 1,000억원 규모 주식 배분 계획 발표
- 코인플러그: 시민참여형 공공안전 제보서비스(앱)를 통해 안전 미션 제보 참여 시 보상획득 및 디지털 바우처(지역화폐)로 교환 가능
- 그린랩스: 클라우드형 스마트팜 서비스로 농업정보 비대면 문제 해결에 착안, 농사에 필요한 정보(날씨, 경락시세, 농사사전, 농산물 거래소, 농산물 상점 등) 제공
- 코드스테이츠: 온라인 코딩 부트캠프로 교육과정에 더해 국내 150여개 기업 채용 연계 서비스 제공. 위-윈(We-Win) 프로그램을 통해 일정 연봉 이상으로 취업 성공 시 기수료한 교육비를 후지불하는 형식으로 소득을 공유하는 제도를 업계 최초로 선보임

○ 해외 사례

- 에어비앤비: IPO과정에서 비의결 주식 920만주를 숙박공유 호스트에 제공 (숙박공유 호스트 기부펀드에 기부)
- Slock.it: '잠글 수 있는' 모든 것을 소유자와 사용자가 직접 거래하며 스마트 계약으로 사용조건을 결정, 사물인터넷(IoT)으로 상품 관리
- Walmart: 블록체인을 통해 식료품 유통과정 전반을 추적하고 공개하여 원산지 정보를 비롯한 세부 유통과정의 정확성을 획기적으로 확보
- Origin Protocol: 블록체인 기반 P2P 마켓 플레이스를 운영하고, 직거래를 위한 무료 온라인 스토어 제공
- Uber: 미국 증권거래위원회(SEC)는 우버가 드라이버에게 연봉의 15%까지 주식으로 지급할 수 있도록 허용
- LaZooz: 차량공유서비스로 dApp*과 Zooz토큰**을 통해 비용 지불(중개 수수료 0%)

* decentralized application: 블록체인을 통해 제작·활용되는 어플리케이션

** 디지털 암호화 토큰으로 거래소에서 현금화 가능



출처: 중소기업연구원, 프로토콜 경제의 현황(2020.11)

□ 블록체인 기술 도입을 통한 4차 산업혁명의 포용성 강화

○ **4차 산업혁명의 핵심기술인 인공지능과 블록체인의 관계는 혁신과 신뢰, 성장과 분배**라고 볼수 있음. 인공지능 발달에 있어 데이터 신뢰성 문제가 발생하면 아무리 알고리즘이 훌륭해도 데이터 신뢰를 못하면 의미가 없음. 블록체인은 모든 거래와 정책, 의사 표현을 참여자 모두에게 공개하면서도 변조나 해킹이 불가능한 기술이기에 **초신뢰 경제혁명을 불러올 뿐만 아니라 신뢰성을 획기적으로 신장**시킴.

○ 상기 기술된 프로토콜 경제는 블록체인 업계에선 익숙한 개념으로 지난해 급성장한 탈중앙화 금융(Defi) 시장이 이러한 형태를 띰. 디파이 성장의 배경에는 플랫폼이 플랫폼 이용 대가로 이용자들에게 지급했던 거버넌스 토큰*이 있음. 즉, **특정 플랫폼을 이용하는 것으로 경제적 이득을 얻을 수 있음을 이용자가 이해할 경우 플랫폼 성장과 이용자의 이익 상승을 동반 실현 가능**

* 거버넌스 토큰: 해당 플랫폼의 운영 정책(프로토콜)을 결정할 때 사용할 수 있는 일종의 투표권으로 플랫폼이 커질수록 가치가 올라가며 타인에게 금전적 양도가 가능

○ 즉, 블록체인 기술은 이해관계자 간 민주적 절차를 통해 이용자가 함께 플랫폼 규칙을 설정할 수 있어 **개방성과 투명성을 확보할 수 있으며 참여자에게 공정한 분배시스템을 보장**함으로써 네트워크 확장과 경제적 성장을 함께 얻을 수 있다는 점에서 **포용적 성장을 위한 필수적 기술**임

4 스마트 포용타운 시범구현 제안(안)

□ GIST블록체인경제센터는 심화하는 양극화 현상과 지역 불균형을 해결하는 실증방안으로 ‘스마트 포용타운 시험구현’을 제안함

□ 스마트 포용타운의 추구방향

ㅇ 지식추구 사회로 가치 전환

- 미국의 유명 경제학자이자 미래학자인 조지 길더(George Gilder)는 그의 저서 「The Scandal of Money」에서 ‘돈’의 의미를 고찰한 바 있음. 부는 화폐를 찍어내 공급함으로써 창출되는 것이 아니라 **지식의 수준이 발전함에 따라 창출**되는 것임을 주장하며 사회의 구성원이 지속해서 학습함에 따라 경제가 성장한다고 설명함. 즉 자본주의란 **인센티브가 아니라 정보생산이 사회번영**을 만든다는 점으로, 화폐는 가치 있는 것의 척도를 제공하고, 돈은 인류의 시간의 가치와 연동됨을 강조함.

- GIST가 제안하는 스마트 포용사회는 블록체인 공동체 안에 규약을 세우고 미래 자산을 공유할 수 있도록 시스템을 제공하여 **지식 생산 활동, 즉 연구·창업·근로를 통해 사회가 진보하고 부의 생산**으로 이어지게 함. 경제적 독립성을 확보할 수 있는 사다리를 제공하여 일하면 누구나 중산층으로의 진입 및 성장을 할 수 있는 사회를 추구함.

□ 스마트 포용타운 구현 개요

ㅇ **시범 구현 지역: GIST 캠퍼스 및 광주시 인공지능 산업융합 집적단지 (광주연구개발특구 첨단3지구)**

- GIST캠퍼스 및 광주시 인공지능 산업융합 집적단지(광주연구개발특구 첨단3지구)

- 스마트 포용타운 파일럿 사례를 통해 교육·연구·네트워크 실증이 가능하며 향후 비수도권지역 스마트도시 건설 방안으로 발전 가능

<참고> 광주연구개발특구 첨단3지구

◆ 첨단3지구는 2020.12.31.부로 실시계획에 대한 승인・고시 완료
◆ GIST는 5년간 광주연구개발특구 첨단3지구에 자립형 인공지능 중심 창업 단지 조성 예정
◆ 100대 국정과제 추진을 위해 R&D특구 내 조성될 예정
◆ 위치 및 면적: 광주광역시 북구 오룡동·대촌동·월출동, 광산구 비아동, 장성군 남면·진원면 일원 (3,616,853㎡)
◆ 사업시행자: 광주광역시도시공사사장 노경수 (공영개발)
◆ 관련부서: 광주광역시 미래산업정책과, 전라남도 연구바이오산업과, 장성군 미래성장개발과

ㅇ 스마트 포용타운 핵심

- 거버넌스: 이해관계자 자본주의 모델 기반 스마트 거버넌스
- 경제: 타운가치 연동 토큰분배를 통한 프로토콜 경제 구현
- 기술: 인공지능・블록체인 기술을 활용한 효율화・포용 사회 인프라 구현

ㅇ 현재의 혁신단지 단계별 개발 체계

- 다양한 이해당사자들이 존재함. 가령, 시청, 기업, 대학, 연구소, 참여자, 토지수용대상자, 입주자, 건설사, 토지주택공사 등

- 단계별 소유권 이전이 되는 과정에서, 최대 수혜자는 누구인가?

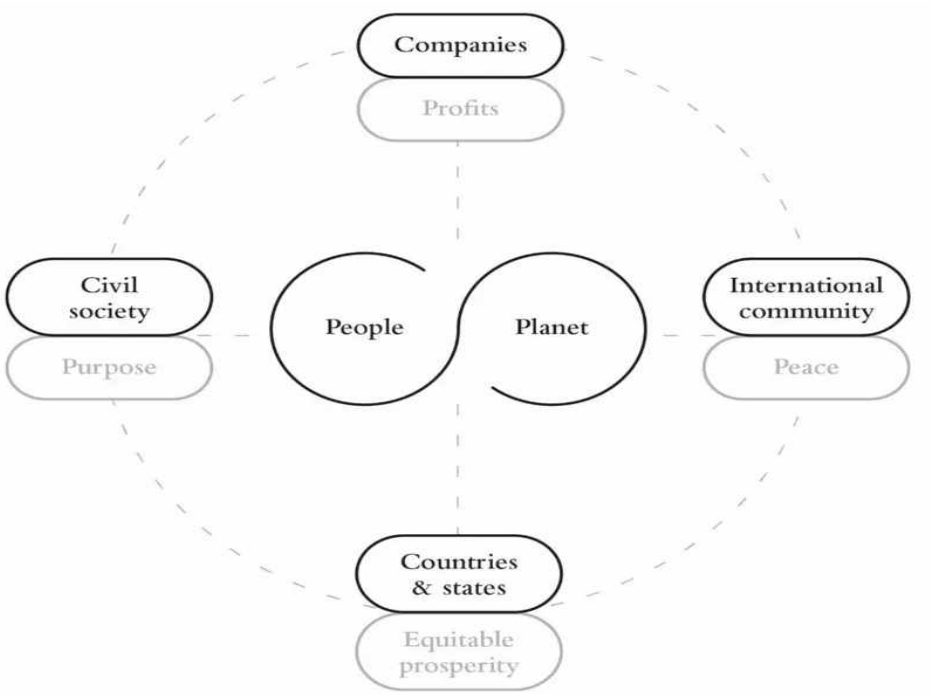
□ 스마트 포용타운 거버넌스

ㅇ 이해관계자 자본주의 기반 스마트사회 거버넌스

- **스마트사회 거버넌스 구축**: 스마트 포용타운은 세계경제포럼 회장인 클라우스 슈밥이 언급한 **이해관계자 자본주의 모델 기반으로 스마트사회 거버넌스를 구축**하고자 함.
- 포용타운에 참여하는 모든 사람은 운영을 위한 의사결정에 영향을

미칠 권리를 가지며 이를 통해 **사회적 이해관계를 균형 있게 맞출 수 있는 거버넌스 구축**이 가능

- **이해관계자 간 최종 합의는 타운 내 블록체인상에 기록되어 투명하게 공개**되며, 합의된 경제활동을 통한 이익은 하기 기술될 토큰발행을 통해 경제·사회의 모든 참여자에게 고루 분배됨. 이러한 순환의 지속을 통해 독자적이고 경제적인 운영모델을 구축할 수 있음.



[그림1] 클라우스 슈밥의 글로벌 이해관계자 자본주의 모델

출처: 「Stakeholder Capitalism」(2021.1), 클라우스 슈밥

□ 스마트 포용타운 경제

ㅇ 타운가치 연동 토큰발행

- **경제적·질적 상승효과를 참여 인력에 환류하는 프로토콜 경제 구현**
  - **스마트타운의 기반이 되는 기술과 시간에 따른 경제효과를 반영할 수 있는 새로운 형태의 부가가치**로서 네트워크 내 블록체인을 통해 사용 가능한 발전토큰*을 발행함. **발행된 토큰 총량은 사업지역 지대상승분과 연동되며, 타운성공에 구심적 역할을 할 초기참여자·창업가 등에게 지급됨.**

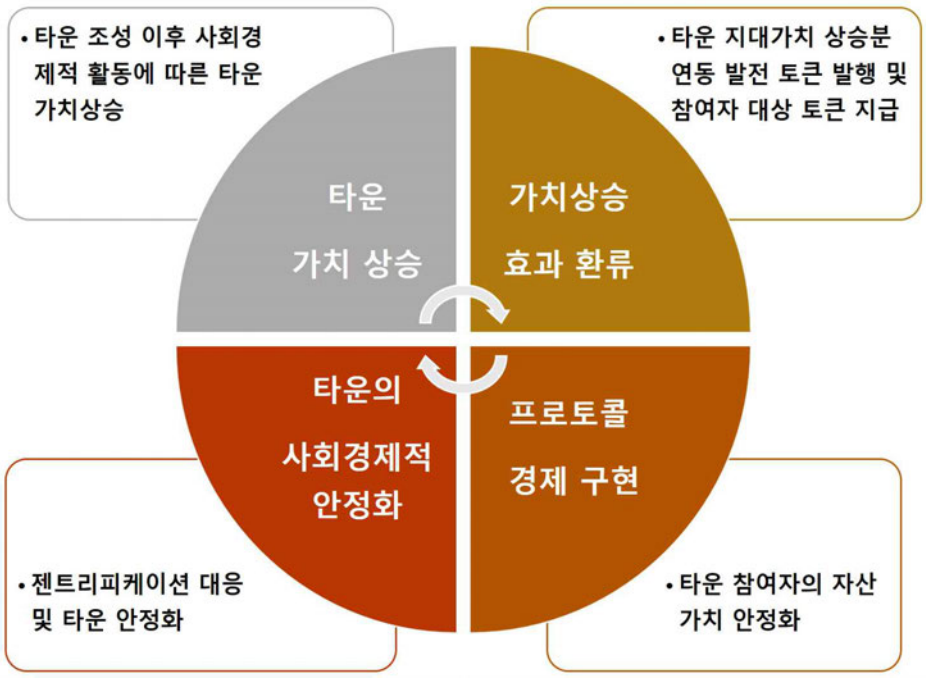
* 발행되는 발전토큰은 스마트타운의 법정화폐가 아니며, 증권형 토큰(STO,

Security Token)이나 자산형 토큰(Asset Token)과 상이하다.

- 스마트타운 구현 후 블록체인에 기반한 타운 내 토큰의 발행 총량은 한정됨. 해당 발전토큰은 타운의 지대상승분과 연동되어 타운 참여자에게 지급되며, 구현 초기 참여자일수록 많은 토큰을 받도록 구현하여 타운 내 창출되는 경제적 가치가 참여자에게 환류되도록 함

- 인재 유입 효과 및 젠트리피케이션 대응 가능

- 토큰은 초기참여자, 우수성과 참여자, 창업가 등에 지급되는 **스마트타운 사업의 지분 혹은 주식과 같은 개념**으로 상호 거래가 가능
- 이는 시간에 따른 **경제적 인플레이션에 뒤처지지 않는 자산수단을 제공함으로써** 참여 인재를 확보할 수 있는 **경제적 유인**이자 타운 구축 후 젠트리피케이션 화에 따른 부동산 가격 급등·인구 공동화를 방지하고 안정적 타운을 구축할 수 있는 **사회경제적 수단**임
- **토큰발행 및 배분은 자산인플레이션에 대한 헤징수단으**로서의 경제적 울타리 효과와 프로토콜 경제 구현을 통한 근로 가치 절상 효과를 발휘
- 이러한 시스템을 통해 스마트타운의 부가가치를 창출하는 정보의 역할과 지역 이전으로의 경제적 인센티브를 제공하고 지속적인 우수 인재의 참여 촉진 효과를 기대할 수 있음.



[그림2] 타운가치 연동 토큰발행의 기대효과

ㅇ 토큰발행을 위한 제도적 기반 마련

- 스마트 포용타운이 소속된 지자체 조례 및 특례 개정 등을 통해 토큰발행의 제도적 기반을 마련하고 토큰 총량과 지대상승분 연동성을 강화할 수 있는 사전 작업 필요
- 실증 후 스마트도시조성법 개정을 통해 지대상승연동 토큰발행을 법으로 보장하는 정책 제안

<참고> 사회적 자산화 추진사례

◆ 부동산 가격 급등 및 젠트리피케이션 등 도시화 부작용의 대응방안으로 사회적 자산화에 관심이 커지며, 주식회사·협동조합 등 다양한 조직이 부동산(건물)을 공동소유하려는 움직임이 나타남. 사회적 자산은 소상공인과 같은 경제행위자가 토지·부동산 소유주의 자산 수익률에 대한 위협을 느끼지 않으면서 이용할 수 있는 자원을 의미

◆ 부동산 이용자가 부동산을 준소유하는 사회적 자산화는 주체에 따라 민간주도형과 공공주도형이 존재하며, 활성화를 위해선 법·제도적 뒷받침과 금융 지원이 필수임. 블록체인 기술을 활용하여 사회적 자산화를 추구할 경우 운영의 투명성과 신뢰 확보가 가능하며, 공정한 이익 분배가 가능함.

◆ 사회적 자산화 사례

- 서울시는 서울시정 4개년 계획(2019~2022)에서 '상생과 균형발전을 위한 사회적 경제 공동체 자산 조성'을 역점 추진과제로 상정
- 영국은 정부와 민간이 소유한 토지와 건물을 지역사회 조직이 활용하는 사회적 자산화 지원을 위해 '공동체 자산 이전'과 '공동체 우선 입찰권' 제도를 운영
- 또한, 해외에선 다양한 방식의 시민자산화*를 진행함. 시민자산화를 위한 협동조합 설립 또는 별도 회사를 세우거나, 민관이 합작하여 회사를 세워 시민자산화 전략을 취함

* 시민자산화: 다수 시민이 공동 소유의 자산을 마련해 사용·운영·관리할 권리를 확보하는 것

□ 스마트 포용타운의 핵심기술

ㅇ 인공지능과 블록체인

- 스마트시티*를 넘어 지속발전하는 스마트사회에는 인공지능과 블록체인 기술 결합하여 **성장과 효율성**(인공지능), 그리고 **포용과 투명성**(블록체인)을 아우르는 기술구현이 필수

* 스마트시티: 도시의 경쟁력과 삶의 질의 향상을 위하여 건설•정보통신기술 등을 융•복합하여 건설된 도시기반 시설을 바탕으로 다양한 도시서비스를 제공하는 지속가능한 도시(스마트도시 조성 및 산업진흥 등에 관한 법률 제2조 제1항)

- **(인공지능) 성장・효율 추구 및 생산성 향상 기여**
  - **인공지능(AI), 블록체인(Blockchain), 클라우드(Cloud), 데이터(Data) 기술이 결합된 ABCD 네트워크를 구성**, 타운 내 발생하는 데이터를 축적하고 해당 데이터를 통해 AI 소프트웨어, 모델 등은 인공지능 자동학습을 통해 업그레이드할 수 있음. 학습을 통해 향상된 사항은 교통신호와 연동된 안전한 도로주행(자동차), 정확한 장소에 배달하는 배달로봇, 효율적인 전력 생산 및 거래, 행정정보 제공 등 시민의 안전과 편의를 극대화할 수 있음.

- **(블록체인) 탈중앙화에 따른 모두의 참여와 프로토콜 설정 합의에 따른 포용과 투명성 보장 가능**
  - 블록체인경제센터는 Blockchain: ETH-ECC, BIT-ECC 네트워크 프로토콜 코어를 보유, 이를 인공지능, 클라우드, 데이터 기술과 합의 프로토콜이 가동되는 네트워크 구성을 테스팅 중. 컴퓨팅 파워를 보유한 본 네트워크는 **클라우드 컴퓨팅 기반의 데이터 스토리지 공간 확보를 통해 실증 대상인 스마트사회의 데이터 축적 가능**

국정기획위원회 제출 후 검토 결과 수신 내용

디지털 뉴딜 신규보완 사업 제안(113건) 검토 내용 회신

(정책기획위) 디지털뉴딜 제안과제 부처 검토내용 회신

보낸날짜:
2021.06.25 06:34:52

사업명: 스마트 포용사회 실증: 재도개선과 인공지능과 블록체인 기술 적용으로 근로 가치가 고취되는 스마트 포용도시 건설

신설/보안: 신설

담당부처: 국토부

담당과 담당자 성명: 국토부 도시경제과 박oo 사무관

분류 1: 정책사업

분류 2: 불수용

분류 3: 불수용

검토의견: 도시와 연동한 자산가치 부여에 대한 관계기관, 지자체, 시민 등과 공감대 형성이 필요한 사항으로 신중한 검토 필요

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0135119
(43) 공개일자 2020년12월02일

(51) 국제특허분류(Int. Cl.)
G06Q 20/06 (2012.01) *G06Q 20/38* (2012.01)
(52) CPC특허분류
G06Q 20/065 (2013.01)
G06Q 20/38 (2020.05)
(21) 출원번호 10-2019-0120655
(22) 출원일자 2019년09월30일
심사청구일자 2019년09월30일
(30) 우선권주장
1020190061493 2019년05월24일 대한민국(KR)

(71) 출원인
광주과학기술원
광주광역시 북구 첨단과기로 123 (오룡동)
(72) 발명자
장재혁
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 전기전자컴퓨터공학부
이홍노
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 전기전자컴퓨터공학부
(74) 대리인
김기문

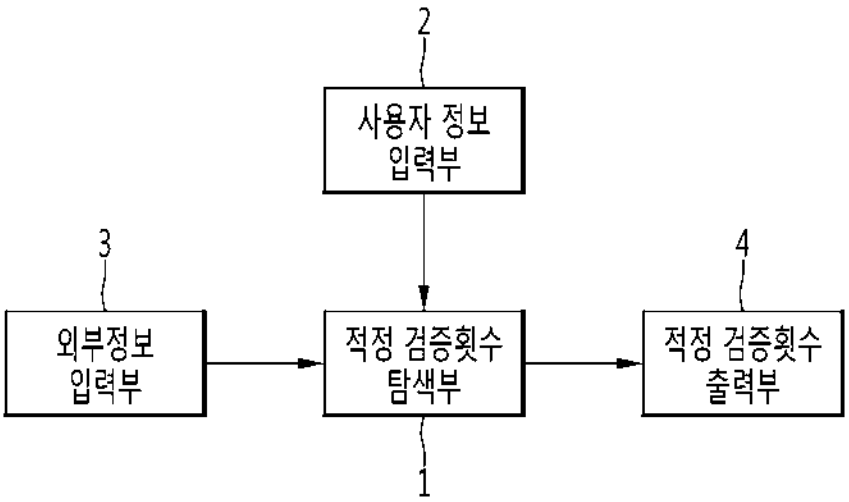
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 **블록체인의 거래검증시스템, 및 블록체인이 거래검증방법**

(57) 요 약

본 발명에 따른 블록체인의 거래검증시스템에는, 사용자 정보가 입력되는 사용자 정보 입력부; 암호화폐시스템과 관련되는 외부정보가 입력되는 외부정보 입력부; 상기 외부정보 및 상기 사용자 정보를 이용하여, 현재 일어나는 현재 거래의 검증에 적정한 블록검증의 수를 탐색하는 적정 검증횟수 탐색부; 및 상기 블록검증의 수를 출력하는 적정 검증횟수 출력부가 포함된다. 본 발명에 따르면, 사용자는 안전하고 신속한 거래를 위하여 필요한 블록검증의 수를 자동으로 알아낼 수 있다.

대 표 도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호	1711093581(NN24010)
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기술진흥센터
연구사업명	융합기술개발
연구과제명	확장가능한 탈중앙화 보안성 ECCPoW 블록체인
기 여 율	1/2
과제수행기관명	광주과학기술원
연구기간	2019.04.01 ~ 2019.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	GK11380
부처명	광주과학기술원
과제관리(전문)기관명	광주과학기술원
연구사업명	실용화연구개발
연구과제명	암호-부호 작업증명(PoW) 비트코인/이더리움 하드포크 1.0개발
기 여 율	1/2
과제수행기관명	광주과학기술원
연구기간	2019.02.01 ~ 2019.12.31

명 세 서

청구범위

청구항 1

사용자 정보가 입력되는 사용자 정보 입력부;

암호화폐시스템과 관련되는 외부정보가 입력되는 외부정보 입력부;

상기 외부정보 및 상기 사용자 정보를 이용하여, 현재 일어나는 현재 거래의 검증에 적정한 블록검증의 수를 탐색하는 적정 검증횟수 탐색부; 및

상기 블록검증의 수를 출력하는 적정 검증횟수 출력부가 포함되는 블록체인의 거래검증시스템.

청구항 2

제 1 항에 있어서,

상기 사용자 정보에는, 적어도 사용자가 입력하는 입력 거래금액이 포함되는 블록체인의 거래검증시스템.

청구항 3

제 1 항에 있어서,

상기 외부정보에는,

-현재 블록 하나를 채굴하면 지급되는 보상금,

-현재 블록 하나를 채굴할 때 소용되는 평균비용, 및

-현재 블록의 채굴속도가,

중의 적어도 하나가 포함되는 블록체인의 거래검증시스템.

청구항 4

제 3 항에 있어서,

상기 외부정보에는,

-블록 하나를 채굴하면 지급되는 보상금,

-현재 블록 하나를 채굴할 때 소용되는 평균비용, 및

-현재 블록의 채굴속도가,

모두 포함되는 블록체인의 거래검증시스템.

청구항 5

제 2 항에 있어서,

상기 적정 검증횟수 탐색부에는,

필요한 정보가 저장되는 메모리;

상기 외부정보를 이용하여, 현재 상태에서 안전한 안전거래 한도액을 추출하는 안전거래 한도액 추출부; 및

상기 안전거래 한도액과 상기 입력 거래금액을 비교하여 안전성을 판별하는 안정성 판별부가 포함되는 블록체인의 거래검증시스템.

청구항 6

제 5 항에 있어서,

상기 안전거래 한도액 추출부에는,

공격자가 기대이윤을 영이 되게 만드는 금액을, 상기 안전거래 한도액으로 추출하는 연산부가 포함되는 블록체인의 거래검증시스템.

청구항 7

제 6 항에 있어서,

상기 공격자의 기대이윤은, 공격자의 기대수익에서 공격자의 기대비용을 감산하여 제공되는 블록체인의 거래검증시스템.

청구항 8

제 6 항에 있어서,

상기 안전거래 한도액(C_{Req})는

$C_{Req.} = (1-P_{AS})/P_{AS}*X(\gamma, t_{cut})+X(\gamma, T_{AS})-R(\beta, T_{AS})$

에 의해서 추출되고,

여기서,

P_{AS}(무차원)는 이중지불공격이 t_{cut} 시간 내에 성공할 확률이고,

$X(\gamma,t)$는 공격자가 t 시간 동안 채굴 할 경우 소요하는 평균 블록 채굴 비용(γ와 t에 대해 증가하는 함수)이고,

γ는 블록 하나를 채굴할 때 소요되는 평균비용이고,

T_{AS}는 공격이 성공하기까지 소요된 시간이고,

β는 블록 하나를 채굴하면 지급되는 보상금이고,

$R(\beta,t)$는 공격자가 t 시간 동안 채굴할 경우 획득하는 평균 블록 채굴 보상금(β와 t에 대해 증가하는 함수)인,

블록체인의 거래검증시스템.

청구항 9

제 5 항에 있어서,

상기 안전거래 한도액 추출부에는 파라미터 최적화부가 포함되고,

상기 파라미터 최적화부에는,

- 공격자의 손절시간, 및

- 공격자의 자원비율 중의 적어도 하나를

최적으로 추측하는 블록체인의 거래검증시스템.

청구항 10

제 5 항에 있어서,

상기 안전성 판별부에 의해서 안전하지 않다고 판별되는 경우에는, 상기 블럭검증의 수를 증가시켜 상기 안전거래 한도액을 다시금 추출하는 블록체인의 거래검증시스템.

청구항 11

암호화폐시스템에 관한 정보인 외부정보 및 사용자가 입력하는 사용자 정보를 입력받는 것;

안전하고 신속한 거래를 위하여 적정한 블록검증의 수인 적정 검증횟수를 탐색하는 것; 및

상기 적정 검증횟수를 출력하는 것이 수행되는 블록체인의 거래검증방법.

청구항 12

제 11 항에 있어서,

상기 적정 검증횟수를 탐색하는 것은,

상기 외부정보 및 임의의 현재 블록검증의 수를 이용하여, 현재 상태에서 안전한 거래금액인 안전거래 한도액을 계산하는 것;

상기 안전거래 한도액과, 상기 사용자 정보로서 입력되는 입력 거래금액을 비교하는 것; 및

상기 안전거래 한도액과 상기 입력 거래금액을 비교하여, 상기 안전거래 한도액이 큰 경우에는 안전한 것으로 판단하고, 상기 현재 블록검증의 수를 출력하는 것이 수행되는 블록체인의 거래검증방법.

청구항 13

제 12 항에 있어서,

상기 안전거래 한도액과 상기 입력 거래금액을 비교하여, 상기 안전거래 한도액이 작은 경우에는 안전하지 않은 것으로 판단하고, 상기 현재 블록검증의 수를 증가시킨 다음에, 상기 안전거래 한도액을 다시 계산하는 것이 수행되는 블록체인의 거래검증방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 블록체인의 거래검증시스템, 및 블록체인의 거래검증방법에 관한 것이다.

배 경 기 술

[0002] 근래, 암호화폐는 상거래에서 화폐의 기능을 수행하는 단계에 이르고 있다. 상기 암호화폐를 통한 거래는 블록검증(block confirmation)을 거쳐서 거래가 완료된다. 상기 블록검증은, 거래자 간의 현재 거래행위가 일어난 이후에, 설정된 블록검증의 수에 상당하는 블록이 생성되는 것을 확인한 다음에, 거래를 완료하는 기술이다. 예를 들어, 상기 블록검증의 수가 6인 경우에는, 거래자 간의 거래가 기록되는 블록 이후에 6개의 블록이 추가로 생성되는 것을 기다린 후에 거래를 완료한다. 상기 거래완료는, 상품을 가진자가 상품을 암호화폐의 제공자에게 발송하는 것을 예시할 수 있다.

[0003] 비특허문헌 1에서는 상기 블록검증의 수가 클수록 거래가 안전하다는 것을 증명하고 있다.

[0004] 상기 블록검증은 이중지불(double spending)을 방어하는 수단으로서 활용될 수 있다. 상기 이중지불은 악의적인 사용자가, 가장 긴 블록체인이 살아남는 블록체인의 원칙을 이용하여, 실제거래의 블록체인보다 긴 악의적인 블록체인을 비밀스럽게 제공하여, 실제거래를 무산시키는 공격행위를 말한다.

[0005] 상기 블록검증의 수가 클수록 이중지불을 당할 확률은 작아진다. 그러나, 상기 블록검증의 수가 클수록 거래완료까지 더 오랜 시간이 소요되는 문제점이 있다. 예를 들어, 비트코인(Bitcoin)의 경우에 평균 블록생성의 주기가 10분이기 때문에, 상기 블록검증의 수가 6일 경우 거래완료시간은 60분이 된다. 이 말은 콜라 한 잔을 사기 위하여 60분을 기다릴 수 있는 것이 된다. 이러한 문제점은 암호화폐의 실제사용에 큰 장애물이 되고 있다.

[0006] 이와 같은 배경 하에서, 현재 암호화폐를 이용하는 거래에서는, 거래자가 자신의 책임하에서 거래금액별로 주어지는 가이드 라인에 따라서, 블록검증의 수를 입력하도록 하는 것이 일반적이다. 거래자는, 자신의 거래에서 블록검증의 수를 얼마나 할지를 알 수가 없고, 검증횟수가 아무리 많아도 이중공격 성공확률이 영이 되지는 않는다.

선행기술문헌

비특허문헌

[0007] (비특허문헌 0001) S Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System”[Online] Available: https://bitcoinorg/bitcoinpdf

발명의 내용

해결하려는 과제

[0008] 본 발명은 현재 조건에 맞추어서 적정의 블록검증의 수를 제안하는 블록체인의 거래검증시스템, 및 거래검증방법을 제안한다.

[0009] 본 발명은 거래자가 입력하는 거래금액뿐만 아니라, 다양한 외부상황에 대응하여 안전하게 신속하게 암호화폐를 통하여 거래를 완료할 수 있는, 블록체인의 거래검증시스템, 및 거래검증방법을 제안한다.

[0010] 본 발명은 암호화폐를 통한 안전한 상거래 및 신속한 거래완료의 두 목적을 함께 달성할 수 있는, 블록체인의 거래검증시스템, 및 거래검증방법을 제안한다.

과제의 해결 수단

[0011] 본 발명에 따른 블록체인의 거래검증시스템에는, 사용자 정보가 입력되는 사용자 정보 입력부; 암호화폐시스템과 관련되는 외부정보가 입력되는 외부정보 입력부; 상기 외부정보 및 상기 사용자 정보를 이용하여, 현재 일어나는 현재 거래의 검증에 적정한 블록검증의 수를 탐색하는 적정 검증횟수 탐색부; 및 상기 블록검증의 수를 출력하는 적정 검증횟수 출력부가 포함된다. 본 발명에 따르면, 사용자는 안전하고 신속한 거래를 위하여 필요한 블록검증의 수를 자동으로 알아낼 수 있다.

[0012] 상기 사용자 정보에는, 적어도 사용자가 입력하는 입력 거래금액이 포함되어, 사용자가 입력하는 입력거래 금액에 대응하여 안전한 거래를 신속하게 수행할 수 있다.

[0013] 상기 외부정보에는, 블록 하나를 채굴하면 지급되는 보상금, 현재 블록 하나를 채굴할 때 소용되는 평균비용, 및 현재 블록의 채굴속도 중의 적어도 하나가 포함되어, 암호화폐시스템을 이용하는 공격자의 공격경향을 파악하여 더 정확한 블록검증의 수를 알아낼 수 있다. 나아가서, 암호화폐시스템의 종류에 따라서 최적의 안전거래 한도액을 알아낼 수 있다.

[0014] 상기 외부정보에는, 블록 하나를 채굴하면 지급되는 보상금, 현재 블록 하나를 채굴할 때 소용되는 평균비용, 및 현재 블록의 채굴속도가 모두 포함되어, 보다 정확한 블록검증의 수를 알아낼 수 있다.

[0015] 상기 적정 검증횟수 탐색부에는, 필요한 정보가 저장되는 메모리; 상기 외부정보를 이용하여, 현재 상태에서 안전한 안전거래 한도액을 추출하는 안전거래 한도액 추출부; 및 상기 안전거래 한도액과 상기 입력 거래금액을 비교하여 안전성을 판별하는 안정성 판별부가 포함된다. 이에 따르면, 현재 상태를 반영하여 실시간으로 변동하는 안전한 금액을 실시간으로 알아낼 수 있다. 나아가서, 암호화폐시스템의 종류에 따라서 최적의 안전거래 한도액을 알아낼 수 있다.

[0016] 상기 안전거래 한도액 추출부에는, 공격자가 기대이윤을 영이 되게 만드는 금액을, 상기 안전거래 한도액으로 추출하는 연산부가 포함된다. 이에 따르면, 공격자의 이익에 기반하기 때문에, 공격자의 공격경향을 예측하기 때문에, 더 정확한 블록검증의 수를 알아낼 수 있다.

[0017] 여기서, 상기 공격자의 기대이윤은, 공격자의 기대수익에서 공격자의 기대비용을 감산하여 제공될 수 있다.

[0018] 여기서, 상기 안전거래 한도액(C_{Req})는, $C_{Req.} = (1-P_{AS})/P_{AS}*X(\gamma, t_{cut})+X(\gamma, T_{AS})-R(\beta, T_{AS})$ 에 의해서 추출될 수 있다.

[0019] 상기 안전거래 한도액 추출부에는 파라미터 최적화부가 포함되고, 상기 파라미터 최적화부에는, 공격자의 손절시간, 및 공격자의 자원비율 중의 적어도 하나를 최적으로 추측할 수 있다. 이에 다르면, 공격자의 선택 파라미터를 최적으로 예상할 수 있어서, 더 안전한 블록검증의 수를 제안할 수 있다.

[0020] 상기 안전성 판별부에 의해서 안전하지 않다고 판별되는 경우에는, 상기 블럭검증의 수를 증가시켜 상기 안전거래 한도액을 다시금 추출하도록 하는 과정이 반복될 수 있다. 이에 의해서 사용자는 가장 안전하면서 최소한 블록검증의 수를 알아낼 수 있다.

[0021] 본 발명에 따른 블록체인이 거래검증방법에는, 암호화폐시스템에 관한 정보인 외부정보 및 사용자가 입력하는 사용자 정보를 입력받는 것; 안전하고 신속한 거래를 위하여 적정한 블록검증의 수인 적정 검증횟수를 탐색하는 것; 및 상기 적정 검증횟수를 출력하는 것이 수행된다. 사용자는 안전하고 신속하게 블록체인을 통한 거래를 수행할 수 있다.

[0022] 상기 적정 검증횟수를 탐색하는 것은, 상기 외부정보 및 임의의 현재 블록검증의 수를 이용하여, 현재 상태에서 안전한 거래금액인 안전거래 한도액을 계산하는 것; 상기 안전거래 한도액과, 상기 사용자 정보로서 입력되는 입력 거래금액을 비교하는 것; 및 상기 안전거래 한도액과 상기 입력 거래금액을 비교하여, 상기 안전거래 한도액이 큰 경우에는 안전한 것으로 판단하고, 상기 현재 블록검증의 수를 출력하는 것이 수행된다. 이에 따르면, 사용자는 안전한 블록검증의 수를 알아낼 수 있거나, 실제로 블록검증의 수가 적용됨으로써, 안전하게 거래가 일어날 수 있다.

[0023] 상기 안전거래 한도액과 상기 입력 거래금액을 비교하여, 상기 안전거래 한도액이 작은 경우에는 안전하지 않은 것으로 판단하고, 상기 현재 블록검증의 수를 증가시킨 다음에, 상기 안전거래 한도액을 다시 계산하는 것이 수행된다. 이와 같은 반복계산에 의해서 실제로 가장 안전하면서 최소한의 블록검증의 수를 알아낼 수 있다.

발명의 효과

[0024] 본 발명에 따르면, 거래자는, 자신이 거래금액만을 입력하는 것으로써, 적정한 블록검증의 수를 알 수 있다. 여기서 적정은, 안전거래 및 신속거래의 두 목적을 합목적적으로 달성할 수 있는 것을 의미한다. 본 문서에서 적정의 의미는 위와 마찬가지로 이해할 수 있다.

[0025] 본 발명에 따르면, 현재 당사자 간의 거래만이 아니라, 당사자 간 거래의 외부에 놓이는 암호화폐전체 시스템의 상황정보(외부정보라고 할 수 있다)을 참조하여, 적정의 블록검증의 수를 제안할 수 있다. 따라서, 공격자의 선택을 예측하여 더 정확하게 상기 블록검증의 수를 제안할 수 있다.

[0026] 그 외의 상세한 구성에 따른 본 발명의 효과는 발명을 실시하기 위한 구체적인 내용에 더 상세하게 제시된다.

도면의 간단한 설명

[0027] 도 1은 실시예에 따른 블록체인의 거래검증시스템을 보이는 도면.

도 2는 적정 검증횟수 탐색부를 상세하게 보이는 도면.

도 3은 적정 검증횟수 탐색부의 상세한 작용을 설명하는 블록체인 거래검증방법의 일부를 설명하는 흐름도.

도 4는 안전거래 한도액 추출부의 자세한 구성을 보이는 도면.

도 5는 공격자가 점유하는 자원비율(p_A)을 변수로 하는 안전거래 한도액의 그래프의 예시도.

도 6은 공격자가 점유하는 자원비율(p_A)을 변수로 손절시간이 무한대일 때 공격성공을 위한 평균소요시간($\mathbb{E}_{T_{AS-ICT}}(p_A;N_{BC})$)의 그래프의 예시도.

도 7은 손절시간을 변수로 안전거래 한도액을 그래프로 예시하는 도면.

발명을 실시하기 위한 구체적인 내용

[0028] 이하에서는 도면을 참조하여 본 발명의 구체적인 실시예를 상세하게 설명한다. 다만, 본 발명의 사상은 이하에 제시되는 실시예에 제한되지 아니하고, 본 발명의 사상을 이해하는 당업자는 동일한 사상의 범위 내에 포함되는 다른 실시예를 구성요소의 부가, 변경, 삭제, 및 추가 등에 의해서 용이하게 제안할 수 있을 것이나, 이 또한 본 발명 사상의 범위 내에 포함된다고 할 것이다.

[0029] 도 1은 실시예에 따른 블록체인의 거래검증시스템을 보이는 도면이다.

[0030] 도 1을 참조하면, 거래검증에 필요한 검증횟수를 탐색하는 적정 검증횟수 탐색부(1), 시스템의 사용자가 입력하

는 정보를 받아 상기 적정 검증횟수 탐색부(1)로 출력하는 사용자 정보 입력부(2), 암호화폐시스템의 상황정보, 즉 외부정보를 입력받아 상기 적정 검증횟수 탐색부(1)로 출력하는 외부정보 입력부(3)가 포함된다. 상기 적정 검증횟수 탐색부(1)는 적정한 블록검증의 수를 탐색하여 적정 검증횟수 출력부(4)로 출력할 수 있다.

[0031] 상기 사용자 정보 입력부(2)는 디스플레이와 입력장치를 가지는 컴퓨터 등의 입력수단이 될 수 있다. 상기 외부정보 입력부(3)는 네트워크에 연결되는 통신수단, 및 네트워크 상의 공개정보를 독출하고 저장하는 컴퓨터가 될 수 있다. 상기 적정 검증횟수 출력부(4)는 디스플레이 또는 통신수단을 가지는 컴퓨터 등의 출력수단이 될 수 있다. 상기 적정 검증횟수 탐색부(1)는 연산장치와 소정의 메모리를 가지는 컴퓨터일 수 있다.

[0032] 상기 외부정보에는, 현재 블록 하나를 채굴하면 지급되는 보상금, 현재 블록 하나를 채굴할 때 소용되는 평균비용, 및 현재 블록의 채굴속도가 포함될 수 있다.

[0033] 상기 사용자 정보에는, 사용자가 현재의 거래를 위하여 입력한 거래금액과 관련되는 정보가 포함될 수 있다.

[0034] 시스템의 사용자는 거래자라고 할 수 있다.

[0035] 실시예의 블록체인의 거래검증시스템은, 상기 사용자 정보 입력부(2)로부터 입력되는 거래금액정보, 및 상기 외부정보 입력부(3)로부터 입력되는 상기 외부정보를 이용하여, 상기 적정 검증횟수 탐색부(1)가 적정한 블록검증의 수를 탐색한다. 탐색된 정보는 적정 검증횟수 출력부(4)를 통하여 출력될 수 있다.

[0036] 사용자는, 상기 적정 검증횟수 출력부(4)에서 출력되는 블록검증의 수를 이용하여 블록검증의 수를 스스로 설정할 수 있다. 블록검증의 수에 따라서 블록검증이 수행된 다음에, 거래가 완료될 수 있다.

[0037] 사용자에 의한 상기 블록검증의 수 설정이 없이, 상기 적정 검증횟수 출력부(4)에서 출력되는 블록검증의 수가 자동으로 적용되어, 블록체인의 거래검증시스템이 운용될 수도 있다.

[0038] 도 1에 제시되는 블록체인의 거래검증시스템은 거래자의 단말기에 탑재될 수 있다.

[0039] 도 1에 제시되는 어느 단위 블럭은 거래자의 단말에, 다른 단위 블럭은 다른 거래자 및 거래소의 단말 중의 적어도 하나에 제공되어, 네트워크를 통하여 연결된 상태로, 블록체인의 거래검증시스템이 동작될 수도 있다.

[0040] 도 1에는 도시되지 않았지만, 상기 블록체인의 거래검증시스템에는, 암호화폐시스템의 임의의 다른 사용자가 블록채굴을 통하여 블록을 체인화하는 네트워크 노드가 포함되는 것은 당연히 이해할 수 있다. 즉, 네트워크 상의 적어도 하나의 다른 노드의 단말기가 협업하여, 현재의 거래를 완료하기 위하여 필요한 블록검증의 수 만큼의 블록을 체인화하여 거래검증을 완료하는 것이 당연히 포함될 수 있다.

[0041] 실시예에 따른 블록체인의 거래검증방법은, 상기 외부정보 및 상기 사용자 정보를 입력받은 다음에, 상기 적정 검증횟수를 탐색한다. 이후에는 탐색된 상기 적정 검증횟수를 출력하는 것에 의해서 수행될 수 있다

[0042] 상기 적정 검증횟수의 출력은, 사용자에게 적정 블록검증의 수가 출력된 다음에, 사용자가 블록검증의 수를 입력하는 것에 의해서 수행될 수 있다. 다른 방법으로, 자동으로 상기 적정 블록검증의 수에 따라서 자동으로 블록검증이 수행되도록 할 수도 있다.

[0043] 도 2는 상기 적정 검증횟수 탐색부를 상세하게 보이는 도면이다.

[0044] 도 2를 참조하면, 상기 적정 검증횟수 탐색부(1)에는, 현재 상태에서 안전한 수준의 검증금액을 추출하는 안전거래 한도액 추출부(10), 상기 안전거래 한도액 추출부(10)에서 추출된 안전거래 한도액과 사용자가 입력한 거래금액을 비교하여 안전성을 판별하는 안전성 판별부(12), 및 상기 안전성 판별부(12) 및 상기 안전거래 한도액 추출부(10)의 동작에 필요한 정보가 저장되는 메모리(11)가 포함될 수 있다.

[0045] 도 3은 상기 적정 검증횟수 탐색부의 상세한 작용을 설명하는 블록체인 거래검증방법의 일부를 설명하는 흐름도이다.

[0046] 도 2 및 도 3을 참조하면, 상기 외부정보 및 상기 사용자 정보가 입력된다(S1). 상기 외부정보에는, 현재 블록 하나를 채굴하면 지급되는 보상금, 현재 블록 하나를 채굴할 때 소용되는 평균비용, 및 현재 블록의 채굴속도가 포함될 수 있다. 상기 사용자 정보에는 사용자가 입력하는 거래금액 정보가 포함될 수 있다. 상기 거래금액에는 각 나라에서 사용되는 화폐 또는 암호화폐의 액수 정보일 수 있다. 예를 들어, 원화 액수 또는 BTC 액수가 정보로서 포함될 수 있다.

[0047] 입력된 정보를 이용하여 안전거래 한도액을 계산한다(S2). 상기 안전거래 한도액은, 상기 외부정보에 의해서 예측되는, 공격자의 비용과 이익을 비교하는 것에 의해서 계산될 수 있다. 상기 안전거래 한도액의 계산은 이후에

더 상세하게 설명한다.

[0048] 상기 안전거래 한도액 계산단계(S2)에서 계산된 상기 안전거래 한도액과, 사용자 정보로서 입력된 입력 거래금액을 비교하여, 현재의 거래가 안전한 지의 여부를 판단한다(S3).

[0049] 상기 안전성 판별 단계(S3)에서, 상기 입력 거래금액과 상기 안전거래 한도액을 비교하여, 상기 입력 거래금액이 작은 경우에는 안전한 것으로 판단하여 현재 검증횟수를 출력한다(S4).

[0050] 반대로 상기 안전성 판별 단계(S3)에서, 상기 입력 거래금액이 큰 경우에는 안전하지 않은 것으로 판단하여, 검증횟수를 증가시키고(S5) 안전거래 한도액 계산단계(S2)를 다시 수행한다. 예를 들어, 현재의 검증횟수가 1인 경우에는 검증횟수를 2로 늘릴 수 있다. 이 후에는 검증횟수가 2가 되고, 현재 시점의 상기 사용자 정보 및 상기 외부정보를 이용하여, 상기 안전거래 한도액을 다시 계산할 수 있다.

[0051] 다시 수행되는 안전거래 한도액 계산단계(S2)에서는 검증횟수가 증가되었으므로, 상기 안전거래 한도액은 올라갈 것으로 예상할 수 있다.

[0052] 상기 안전거래 한도액 계산단계(S2), 및 상기 안전거래 한도액 추출부(10)의 작용은 이하에 상세하게 설명될 것이다.

[0053] 상기 안전거래 한도액은, 공격자가 이중지불공격(본 문서에서 공격이라고 하는 경우에 특별한 설명이 없는 경우에는 이중지불공격을 의미한다)을 하기 위하여 소용되는 비용과, 공격자가 이중지불공격에 성공하는 경우에 얻는 수익을 비교하는 것에 의해서 계산될 수 있다.

[0054] 예를 들어, 공격자가 공격을 하기 위하여 드는 비용에 비하여, 공격자가 공격에 성공하는 경우에 얻는 수익이 작은 경우에는, 공격자가 공격을 하지 않거나, 공격을 하더라도 성공하지 못할 것이기 때문이다.

[0055] 다시 설명하면, '공격자 기대 이윤 = 공격자의 기대 수익 - 공격자의 기대비용'에 의해서 공격자의 기대이윤이 있는 경우에는 공격자가 공격을 할 것이고, 반대로 공격자 기대이윤이 없는 경우에는 공격자가 공격을 하지 않을 것을 예상할 수 있다. 여기서, 상기 공격자 기대이윤이 영이 되는 상태의 거래금액을 상기 안전거래 한도액이라고 정의할 수 있다.

[0056] 상기 공격자의 기대수익은, 공격에 성공 할 경우의 수익(C+R)과, 공격에 실패 할 경우의 수익(영)의 합으로 주어질 수 있다. 간단하게, $P_{AS}*(C+R(\beta,T_{AS}))$로 표시할 수 있다.

[0057] 상기 공격자의 기대비용은, 공격에 성공 할 경우의 일정시간(T_{AS})동안 소요한 비용과, 공격에 실패 할 경우의 손절시간(t_{cut})시간동안 소요한 비용의 합으로 주어질 수 있다. 간단하게, $P_{AS}*X(\gamma,E[T_{AS}]) + (1-P_{AS})*X(\gamma,t_{cut})$로 주어질 수 있다. 여기서 손절시간은 손절시간이 지난 이후에는 공격에 성공할 확률이 희박하다고 판단하고 더 큰 손해를 방지하기 위해 어느 정도의 손해를 입고 공격을 그만두는 시간을 의미할 수 있다.

[0058] 결국, 상기 공격자의 기대이윤은, $P_{AS}*(C+R(\beta,E[T_{AS}])-X(\gamma,E[T_{AS}]))-(1-P_{AS})*X(\gamma,t_{cut})$로 계산될 수 있다.

[0059] 상기 안전거래 한도액C_{Req})은, 공격자의 기대 이윤이 영이 되는 거래금액으로 정의할 수 있다. 따라서, 상기 공격자의 기대이윤이 영이 되도록 거래금액(C)을 좌변에 두고 우변에 나머지 항을 정리하면, $C_{Req.} = (1-P_{AS})/P_{AS}*X(\gamma,t_{cut})+X(\gamma,T_{AS})-R(\beta,T_{AS})$와 같이 정리될 수 있다.

[0060] 상기되는 설명에 제시되는 각 기호, 및 이하의 안전거래 한도액의 실제 계산에 사용되는 다양한 기호의 의미를 표 1로 정리한다.

표 1

[0061]

기호(단위)	설명	비고
C($, BTC)	이중지불 공격대상 거래의 거래금액	사용자입력값
β($, BTC)	블록 하나를 채굴하면 지급되는 보상금	외부입력값
γ($, BTC)	블록 하나를 채굴할 때 소요되는 평균비용	외부입력값
λ_A(No./T)	공격자의 평균 블록채굴 속도(공격자에 따라서 다르고 알려져 있는 것 중의 최고를 이용할 수 있다)(pool에 가입해서 채굴을 하고 있고, 풀 간에 담합할 수 있다)	외부입력값
λ_H(No./T)	일반인의 채굴속도	외부입력값

λ_T(No./T)	전체 채굴속도	외부입력값
t_{cut}(T)	무한정 커질 수 있는 손실을 방지하기 위한 공격제한 시간, 손절시간, cut-time	추측검증값
p_A(%)	공격자가 점유한 계산 자원의 비율	추측검증값
N_{BC}(No.)	거래자가 거래를 완료하기 전 거래를 검증하는 블록검증의 수(block confirmation number)	출력값
E[T_{AS}](T)	공격이 성공하기까지 소요된 평균 시간, t_{cut}보다 작음.	계산값
P_{AS}(무차원)	이중지불공격이 t_{cut} 시간 내에 성공할 확률	계산값
R(β,t)($, BTC)	공격자가 시간 t동안 채굴 할 경우 획득하는 평균 블록 채굴 보상금(β와 t에 대해 증가하는 함수)	계산값
X(γ,t)($, BTC)	공격자가 시간 t동안 채굴 할 경우 소요하는 평균 블록 채굴 비용(γ와 t에 대해 증가하는 함수)	계산값

[0062] 상기 비고란에는 각 기호의 값이 도출되는 방법을 나타내는 것이다.

[0063] 사용자 입력값은, 사용자가 입력하는 사용자 정보로서 여기서는 거래금액을 의미한다.

[0064] 외부입력값은, 각 암호화폐의 시스템에서 공시되고 실시간으로 변하는 값이다. 예를 들어, 상기 외부입력값은, 비트코인의 경우에는 btc.com에, 이더리움의 경우에는 etherscan.io에 공시된다.

[0065] 상기 추측검증값은, 공격자가 선택할 수 있는 값으로서 안전거래 한도액을 계산하는 중에 가장 안전한 값으로 선택될 수 있다. 상기 추측검증값은 추후에 설명될 파리미터 최적화부(16)에 의해서 최적으로 추측될 수 있다.

[0066] 상기 출력값은 본 실시예에서 일 목적으로 출력하는 값으로서, 안전한 검증횟수로서, 블록검증의 수를 나타낸다.

[0067] 상기 계산값은 상기 안전거래 한도액의 계산 중간에 사용되는 값이다.

[0068] 상기 안전거래 한도액을 계산하는 과정, 및 현재 상태에서 안전한 금액을 검출하는 안전거래 한도액 검출부(10)의 구성을 더 상세하게 설명한다.

[0069] 도 4는 상기 안전거래 한도액 추출부의 자세한 구성을 보이는 도면이다.

[0070] 도 4를 참조하면, 상기 안전거래 한도액 추출부(10)에는, 안전거래 한도액을 계산하는 연산부(15), 및 상기 연산부(15)의 연산에 필요한 파라미터를 최적화하는 파라미터 최적화부(16)가 포함된다.

[0071] 상기 파라미터 최적화부(16)는 상기 추측검증값을 최적화하는 블럭으로서, 공격자의 입장에서 공격을 예측할 수 있는 값이다.

[0072] 상기 안전거래 한도액 추출부(10)의 상세한 작용을 다수의 수학식을 참조하여 상세하게 설명한다. 이하의 설명에 있어서 상기 파라미터 최적화부(16)의 작용에 대한 언급이 없는 부분은 모두 상기 연산부(15)에서 수행하는 것으로 이해할 수 있다.

[0073] 하기 수학식 1은 채굴비용/시간의 단위를 채굴비용/블록의 단위로 변환하는 것을 예시한다.

수학식 1

[0074]
$$\gamma = 4.63 \cdot 10^{-22}\ [\mathrm{BTC/hash}] \times 456\mathrm{E}[\mathrm{hashes/block\ mining}] \approx 0.21\ [\mathrm{BTC/block\ mining}].$$

[0075] 상기 수학식 1을 참조하면, 상기 외부정보로서 알려져 있는 채굴비용/해쉬와 해쉬수/블록을 곱하는 것으로서, 블록당 채굴비용을 얻어낼 수 있다. 수학식 1의 결과는 메모리에 저장될 수 있다.

[0076] 하기 수학식 2는 공격자의 채굴속도를 알아내는 것을 예시한다.

수학식 2

$$p_A := \Pr(S_i = n+1 \mid S_{i-1} = n) = \frac{\lambda_A}{\lambda_T},$$

$$p_H := \Pr(S_i = n-1 \mid S_{i-1} = n) = \frac{\lambda_H}{\lambda_T},$$

[0077]

[0078] 상기 수학식 2에서 람다는 속도이고, H(honest)는 일반인이고, T(Total)는 전체를 의미하고, A(Attacker)는 공격자를 의미하고, p_A는 공격자가 가지는 자원의 비율이고, p_H는 일반인이 가지는 자원의 비율이다. p_H에서 소문자 p는 proportion의 두문자로 이해할 수 있다.

[0079] 상기 수학식 2를 통하여 공격자의 채굴속도(λ_T)를 알아낼 수 있다. 상기 수학식 2의 결과는 상기 메모리(11)에 저장될 수 있다. 상기 메모리(11)에는 상기 채굴비용 및 상기 채굴속도뿐만 아니라, 연산에 필요한 정보가 필요에 따라서 저장될 수 있다.

[0080] 하기 수학식 3은 손절시간이 무한대로 공격할 경우에 공격에 성공할 확률로서, 비특허문헌에 인용되는 사토시 나카모토의 논문에 의해서 공지되어 있다.

수학식 3

$$\mathbb{P}_{AS-ICT}(p_A; N_{BC}) = \begin{cases} 1, & p_H \le p_A, \\ 1 - {p_A}^{N_{BC}+1} {p_H}^{N_{BC}} \sum_{j=N_{BC}}^{2N_{BC}} \binom{j-1}{N_{BC}-1} A_j, & p_H > p_A, \end{cases}$$

where

$$A_j \triangleq {p_A}^{j-2N_{BC}-1} - {p_H}^{j-2N_{BC}-1}.$$

[0081]

[0082] 상기 수학식 3에서 P는 확률(probality)를 의미하고, AS(attack success)는 공격성공을 의미하고, ICT(infinite cut time)으로서 손절시간이 무한한 것을 의미하고, N_{BC}는 검증을 위한 블록검증의 수를 의미한다.

[0083] 하기 수학식 4는 상기 수학식 3에 의해서 도출되는 식으로서, 주어진 시간 내에 공격에 성공할 확률에 관한 확률분포를 나타낸다.

수학식 4

$$f_{T^{(1),(2)}}(t)=\frac{p_A\lambda_T e^{-\lambda_T t}\left(p_A p_H(\lambda_T t)^2\right)^{N_{BC}}}{(2N_{BC})!}\cdot\sum_{j=N_{BC}}^{j=2N_{BC}}\binom{j-1}{N_{BC}-1}\,{}_2F_3\left(\mathbf{a};\mathbf{b};p_A p_H(\lambda_T t)^2\right)+\frac{e^{-\lambda_T t}}{t}\frac{(p_H\lambda_T t)^{N_{BC}}}{(N_{BC}-1)!}\left(e^{p_A\lambda_T t}-\sum_{i=0}^{N_{BC}}\frac{(p_A\lambda_T t)^i}{i!}\right)+(1-\mathbb{P}_{AS-ICT})\delta(t-\infty),$$

[0084]

[0085] 상기 수학식 4는 여기서 ${}_pF_q$는 [G Gasper and M Rahman, "Basic Hypergeometric series," in *Basic hypergeometric series*, Second, vol 96, Cambridge University Press, Cambridge, 2004]에 정의되는 일반화 초기하학적 함수(generalized hypergeometric function)이고, a, b는 수학식 5에 정의된다.

수학식 5

$$\mathbf{a}=\begin{bmatrix}N_{BC}+1-j/2\\ N_{BC}+1/2-j/2\end{bmatrix}$$

$$\mathbf{b}=\begin{bmatrix}2N_{BC}+2-j\\ N_{BC}+1\\ N_{BC}+1/2\end{bmatrix}.$$

[0086]

[0087] 상기 수학식 4의 확률분포를 제 1 적분하여, 하기 수학식 6의 손절시간 내에 공격에 성공할 확률을 알아낼 수 있다.

수학식 6

$$\mathbb{P}_{AS}(p_A,t_{cut};N_{BC}):=\Pr\left(T^{(1),(2)}<t_{cut}\right)$$

[0088]

[0089] 상기 수학식 5의 확률분포를 제 2 적분하여, 하기 수학식 7의 손절시간 내에 공격에 성공하는데 소요되는 평균 시간을 알아낼 수 있다.

수학식 7

$$\mathbb{E}_{T_{AS}}\left(p_A, t_{cut}; N_{BC}\right) = \frac{\int_0^{t_{cut}} t f_{T^{(1),(2)}}(t)\,dt}{\mathbb{P}_{AS}\left(p_A, t_{cut}; N_{BC}\right)}$$

[0090]

[0091] 여기서, E(expectation)은 기대값이고, P(probality)는 확률이고, f는 확률론에서의 확률분포이다.

[0092] 상기 수학식 6과 수학식 7을 이용하여, 공격자의 평균공격비용을 수학식 8과 같이 알아낼 수 있다.

수학식 8

$$\begin{aligned}\mathbb{E}_X\left(p_A, t_{cut}; N_{BC}\right) := \\ \mathbb{P}_{AS}\left(p_A, t_{cut}; N_{BC}\right)\mathbb{E}\left[X\left(\lambda_A, T_{AS}\right)\right] \\ +\left(1-\mathbb{P}_{AS}\left(p_A, t_{cut}; N_{BC}\right)\right)X\left(\lambda_A, t_{cut}\right)\end{aligned}$$

[0093]

[0094] 여기서, X는 채굴비용을 나타낸다. 상기 채굴비용 X를 선형함수로 제공하는 것은, 채굴장비를 인터넷 상의 대여소(예를 들어, nicehach.com)를 통하여 대여하는 비용을 선형함수로 예시한 것이다. 만약, 상기 채굴비용이 대여비용이 아닌 장비의 전기사용료 등이 되는 경우와 같이 다양한 경우에 대응하여, 상기 비용함수 X는, n-제곱함수, n-루트함수, 로그함수, 지수함수, 또는 n-루트 함수가 될 수도 있을 것이다.

[0095] 한편, 수학식 3을 이용하여, 손절시간이 무한대일 때 공격성공을 위한 평균소요시간($\mathbb{E}_{T_{AS-ICT}}\left(p_A; N_{BC}\right)$)을 수학식 9와 같이 알아낼 수 있다.

수학식 9

$$\mathbb{E}_{T_{AS-ICT}}\left(p_A; N_{BC}\right) = \frac{\lambda_T^{-1}\left(\sum_{j=N_{BC}}^{2N_{BC}}\binom{j-1}{N_{BC}-1}Z_j + \frac{N_{BC}}{p_H}\right)}{\mathbb{P}_{AS-ICT}\left(p_A; N_{BC}\right)};$$

where

$$Z_j := p_A p_m^{N_{BC}} p_M^{-(N_{BC}-j+1)}\left(\frac{2N_{BC}-2jp_m+1}{p_M-p_m}\right) - jp_A^{-(N_{BC}-j)} p_H^{N_{BC}}.$$

[0096]

[0097] 상기 수학식 9를 이용하여, 공격자가 점유하는 자원을 예측하여 가장 안정된 검증금액을 제시하는, 공격자가 점유하는 자원비율(p_A)을 알아낼 수 있다.

[0098] 구체적으로, 상기 수학식 9를 참조하면, 손절시간이 무한대일 때 공격성공을 위한 평균소요시간($\mathbb{E}_{T_{AS-ICT}}\left(p_A; N_{BC}\right)$)은, 공격자가 점유하는 자원비율($p_A$)을 인자로 하는 함수이다. 이 경우에, 공격자가 점유

하는 자원비율(p_A)을 변수로 손절시간이 무한대일 때 공격성공을 위한 평균소요시간($\mathbb{E}_{T_{AS-ICT}}(p_A;N_{BC})$)을 계산하면, 도 6과 같이 아래로 볼록한 함수로 나타난다.

[0099] 따라서, 도 6의 그래프 상에서, 가장 낮은 곳의 공격자가 점유하는 자원비율(p_A)을 취하여, 공격자가 점유하는 자원비율을 가장 안정적으로 최적화할 수 있다. 공격자는 가장 최소의 비용으로 빠른 공격 효과를 얻어낼 수 있는 점유 자원비율을 선택할 것이기 때문이다.

[0100] 이때 상기 공격자가 점유하는 자원비율에 대한 최적화는 상기 파라미터 최적화부(16)에 의해서 연산부(15)와는 별개의 작용을 통하여 수행될 수 있다. 상기 차라미터 최적화부(16)에서 최적화된, 상기 공격자가 점유하는 자원비율은 상기 연산부(15)로 피드백되어 상기 안전거래 한도액의 획득에 사용될 수 있다.

[0101] 이후에는 상기 안전거래 한도액을 얻기 위하여 하기 수학식 10이 적용된다.

수학식 10

[0102]
$$C_{\text{Req.}}=\frac{\mathbb{E}_X\left(p_A,t_{cut};N_{BC}\right)}{\mathbb{P}_{AS}\left(p_A,t_{cut};N_{BC}\right)}-\mathbb{E}\left[R\left(\lambda_A,T_{AS}\right)\right]$$

[0103] 상기 수학식 10에서 R은 보상금으로서 R은 하기 수학식 11과 같이 주어질 수 있다.

수학식 11

[0104]
$$R\left(\lambda_A,t\right):=\beta\lambda_A t\left(\log_{r_1} r_2\right)^{\lambda_A}\left(\log_{r_3} r_4\right)^{t}$$

[0105] 상기 수학식 11에서 로그텀(log term)은, 보상금이 시간에 대하여 로그함수 또는 지수함수로 가능한 것으로 보이는 것이다. 다만, 시간에 비례하는 함수로 주어지는 경우에는 로그텀들은 모두 1로 주어질 수 있다.

[0106] 다만, 상기 보상금 R도, 채굴비용(X)과 마찬가지로, n-제곱함수, n-루트함수, 로그함수, 지수함수, 또는 n-루트 함수가 될 수도 있을 것이다.

[0107] 다시 수학식 10을 참조하면, 여전히 손절시간(t_{cut})은 알려지지 않았고, 수학식 10은 손절시간을 인자로 하는 함수로 주어질 수 있다. 따라서, 손절시간을 변수로 하여 안전거래 한도액(C_{req})를 최적화할 수 있다.

[0108] 상기 손절시간을 변수로 상기 안전거래 한도액을 그래프로 예시하는 도면이 도 7에 도시된다.

[0109] 도 7을 참조하면, 상기 손절시간에 대하여 상기 안전거래 한도액은 아래로 볼록한 그래프로 나타난다. 가장 안전거래 한도액으로서 가장 낮은 검증금액을 출력할 수 있다. 상기 손절시간에 대응하여 상기 안전거래 한도액의 최적값을 구하는 작용은 상기 파라미터 최적화부(16)를 통하여 수행될 수 있다.

[0110] 상기 파라미터 최적화부(16)는, 현재 상태에서 가장 낮은 상기 안전거래 한도액을, 상기 연산부(15)로 출력할 수 있다.

[0111] 한편, 이상의 설명은 공격자의 자원비율이 50이하인 경우에 적용될 수 있다.

[0112] 반대로 공격자의 자원비율이 50%를 초과하는 경우에는 이미 설명한 바와 같이, 공격자는 손절시간을 무한대로 설정할 수 있다. 왜냐하면, 자원이 50%를 초과하는 경우에는 언젠가는 자원이 우세하므로 공격자가 언젠가는 더 긴 블록체인을 만들 수 있기 때문이다.

[0113] 상기 공격자의 자원비율이 50%이하인 경우(이하 50%공격이라고 한다)와, 상기 공격자의 자원비율이 50%초과하는 경우(이하에서는 51%공격이라고 한다)에, 추측검증값인, 공격자의 손절시간(T_{cut}) 및 공격자의 자원비율(P_A)의 최

적화에 대하여 상세하게 설명한다.

[0114] 상기 공격자의 손절시간(T_{cut}) 및 공격자의 자원비율(P_A)의 최적화는 상기 파라미터 최적화부(16)에서 수행될 수 있다.

[0115] 일반적으로, 51% 공격의 최적 검증횟수가 50% 공격의 최적 검증횟수보다 더 크다. 비트코인과 같은 계산자원의 규모가 거대한 네트워크에서는 특정 집단이 51% 이상의 계산 자원을 확보하기가 매우 어렵다. 따라서 50% 공격에 대한 대비를 하는 것으로도 충분하다. 반대로 계산자원의 규모가 작은 네트워크에서는 51% 공격이 가능할 수 있기 때문에 더욱 철저한 대비가 필요하다.

[0116] 첫번째로, 상기 51% 공격의 상황에서, 공격자에게 최선의 손절시간은 무한대이고, 최적의 공격자원 점유율은 계산가능하다.

[0117] 손절시간이 유한한 경우와 무한한 경우에 대해 이윤을 비교할 수 있다. 예를 들어, 블록 하나를 채굴할 때의 보상금(β)이 블록 하나를 채굴할 때의 비용(γ)보다 큰 경우, 즉 채굴로 얻는 시간당 보상이 채굴로 소요하는 시간당 비용보다 큰 경우에서는, 손절시간이 무한한 경우의 공격 이윤이 가장 크다.

[0118] 반대로, 블록 하나를 채굴할 때의 보상금(β)이 블록 하나를 채굴할 때의 비용(γ)보다 작은 경우에는 손절시간이 무한한 경우의 공격 이윤이 항상 가장 크다고 할 수는 없지만, 대부분의 실질적인 경우에서는 그러하다. 뿐만 아니라 정상적인 네트워크는 시장경제에 의해 β가 γ보다 크도록 형성되어있는 것이 보통이다.

[0119] 51% 공격 상황에서 손절시간이 무한할 때, 상기 안전거래 한도액($C_{Req.}$)은 공격자가 가지는 자원비율(p_A)에 대해 아래로 볼록인 함수이며, 이는 수학적으로 증명된 바가 있고 도 5와 같이 실제로 예시된다.

[0120] 즉, 공격자가 가지는 자원비율(p_A)에 대해 상기 안전거래 한도액(C_{Req})의 하한이 존재한다. 상기 입력 거래금액(C)과 상기 안전한 검증금맥(C_{Req})을 비교하여 $C<C^{*}_{Req.}$를 만족하면, 공격자가 가지는 모든 자원비율에 대해 $C<C_{Req.}$이므로, 거래가 안전함을 알 수 있다. 따라서, 거래자는 안전성을 판단하기 위하여, 공격 자원 점유율이 최악의 경우만 고려하면 된다. 예를 들어, 도 5에서 가장 가장 낮은 상기 안전거래 한도액을 선택하면 된다.

[0121] 두 번째로 50% 공격 상황에서도 공격자에게 최적의 손절시간(t_{cut})과 가장 합리적인 공격자가 가지는 자원비율(p_A)를 고려해 볼 수 있다.

[0122] 상기 50% 공격의 경우에는 공격성공 평균시간이 최소화되는 공격자가 가지는 자원비율(p_A)이 합리적인 공격 자원 점유율이다. 50% 공격인 경우에는 51% 공격과는 다르게 무한한 손절시간을 고려 할 필요가 없는데, 이는 공격자의 기대 손실이 무한히 커지기 때문이다.

[0123] 따라서 공격자는 반드시 유한한 손절시간의 선택이 강제된다. 공격자의 입장에서는 손절시간을 선택할 때 공격이 성공하기까지 소요되는 평균 시간, 즉 공격성공 평균시간을 고려하지 않을 수 없다. 예를 들어 공격성공 평균시간이 알려져 있다면, 공격을 시도 할 때 손절시간을 공격성공 평균시간, 그보다 조금 더 큰 시간, 혹은 그보다 수 배 큰 시간으로 설정 할 수 있다.

[0124] 50% 공격의 경우에는 공격 성공확률이 그리 높지 않으므로, 공격자는 공격을 여러 회 시도하여 이윤을 취할 것이다. 그러므로 매 회의 공격 시도에 걸리는 시간이 길면 길수록 적자에서 흑자로 전환되는데 걸리는 시간이 늦어질 것이다. 신속하고 정확하게 손절시간을 찾기 위해, 공격자는 공격성공 평균시간이 최소화 되는 공격자가 가지는 자원비율(p_A)을 탐색할 것이다.

[0125] 공격성공 평균시간이 도 6과 같이 아래로 볼록인 함수, 즉 최소값이 존재하는 함수임을 계산적으로 확인하였다. 따라서 50% 공격의 경우에는 공격성공 평균 간이 최소화되는 공격자가 가지는 자원비율(p_A)를 합리적인 공격 자원 점유율로 가정할 수 있다.

[0126] 또한, 50% 공격에서 공격자가 가지는 자원비율(p_A)이 고정되어 있을 때, 상기 안전거래 한도액(C_{Req})은 t_{cut}에 대해 아래로 볼록인 함수임을 도 7과 같이 확인 하였다. 즉, 손절시간에 대하여 상기 안전거래 한도액의 하한이 존재한다. 따라서 거래자는 안전성을 판단하기 위해 손절시간이 최악의 경우, 즉 상기 안전거래 한도액이 가장 낮은 경우만 고려하면 된다.

[0127] 공격자의 자원비율이 50%를 초과하는 경우에는 이미 설명한 바와 같이, 공격자는 손절시간을 무한대로 설정할

수 있다.

[0128] 이 경우에는 수학식 9에 제시되는, 손절시간이 무한대일 때 공격성공을 위한 평균소요시간($\mathbb{E}_{T_{AS-ICT}}(p_A;N_{BC})$)을 이용하여, 하기 수학식 12와 같이 상기 안전거래 한도액을 알아낼 수 있다. 다시 말하지만, 이 경우에 손절시간은 무한대이다.

수학식 12

[0129] $$C_{\text{Req.}} = \max\left(0,(\gamma-\beta)\lambda_T p_A \mathbb{E}_{T_{AS-ICT}}(p_A;N_{BC})\right)$$

[0130] 상기 수학식 12를 참조하면, 공격자가 점유하는 자원비율(p_A)은 알려지지 않았고, 수학식 12는 공격자가 점유하는 자원비율(p_A)을 인자로 하는 함수로 주어질 수 있다. 따라서, 공격자가 점유하는 자원비율(p_A)을 변수로 하여 안전거래 한도액(C_{req})를 최적화할 수 있다. 상기 최적화는 상기 파라미터 최적화부(16)에서 수행될 수 있다.

[0131] 공격자가 점유하는 자원비율(p_A)을 변수로, 상기 안전거래 한도액을 그래프로 예시하는 도면이 도 5로 이미 설명된 바가 있다.

[0132] 도 5를 참조하면, 상기 공격자가 점유하는 자원비율(p_A)에 대하여 상기 안전거래 한도액은 아래로 볼록한 그래프로 나타난다. 가장 안전거래 한도액으로서 가장 낮은 검증금액을 출력할 수 있다.

[0133] 본 문서에 기재되는 기술에 다르면, 사용자가 거래금액만을 입력하는 것에 의해서 적정 검증횟수를 출력하고, 상기 적정 검증횟수에 따라서 블록검증의 수를 설정할 수 있다. 이에 따르면, 안전하고 신속한 암호화폐를 이용한 상거래가 가능한 장점이 있다.

[0134] 또한, 거래의 안정성을 판단하기 위한 기준으로서, 상기 안전거래 한도액을 계산하는 과정에 추측되어야 하는 추측검증값이 최적화됨으로써, 사용자가 거래금액만을 입력하더라도 적정한 블록검증의 수를 알아낼 수 있다.

산업상 이용가능성

[0135] 본 발명에 따르면, 암호화폐를 통한 상거래를 안전하고 신속하게 수행할 수 있도록 함으로써, 암호화폐를 이용하는 실물거래를 더욱 촉진할 수 있다.

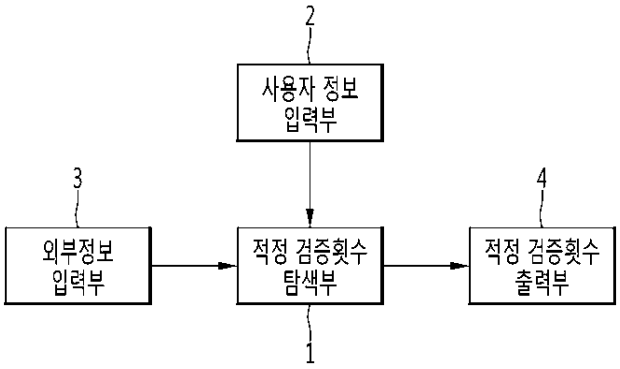
부호의 설명

[0136] 10: 안전거래 한도액 추출부

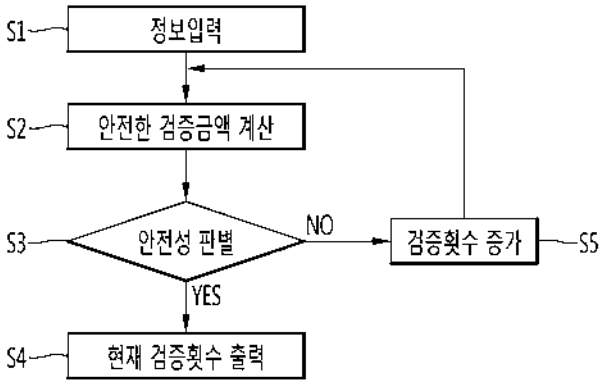
12: 안전성 판별부

도면

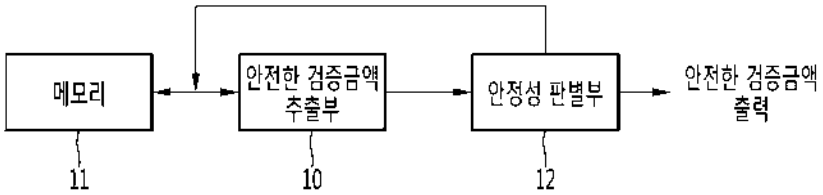
도면1



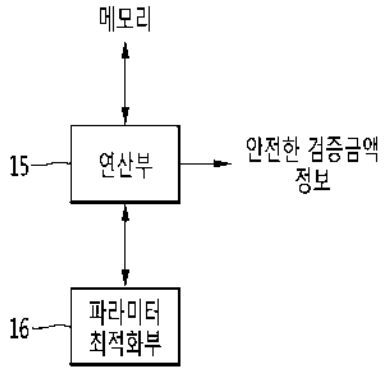
도면2



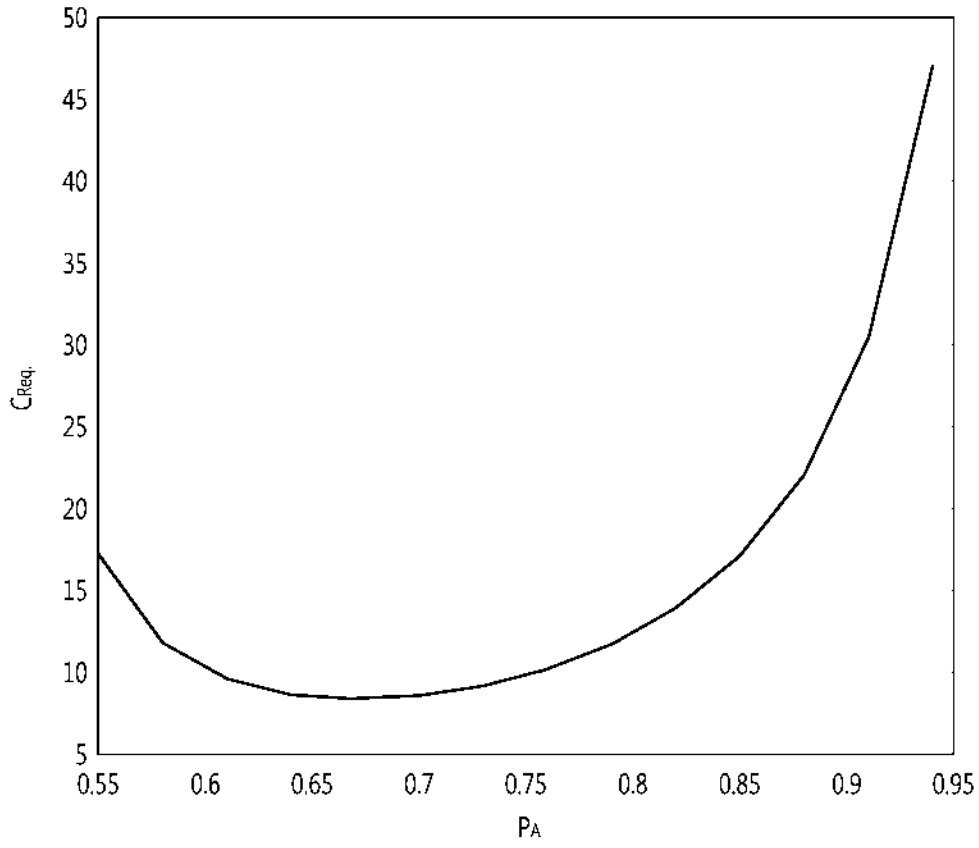
도면3



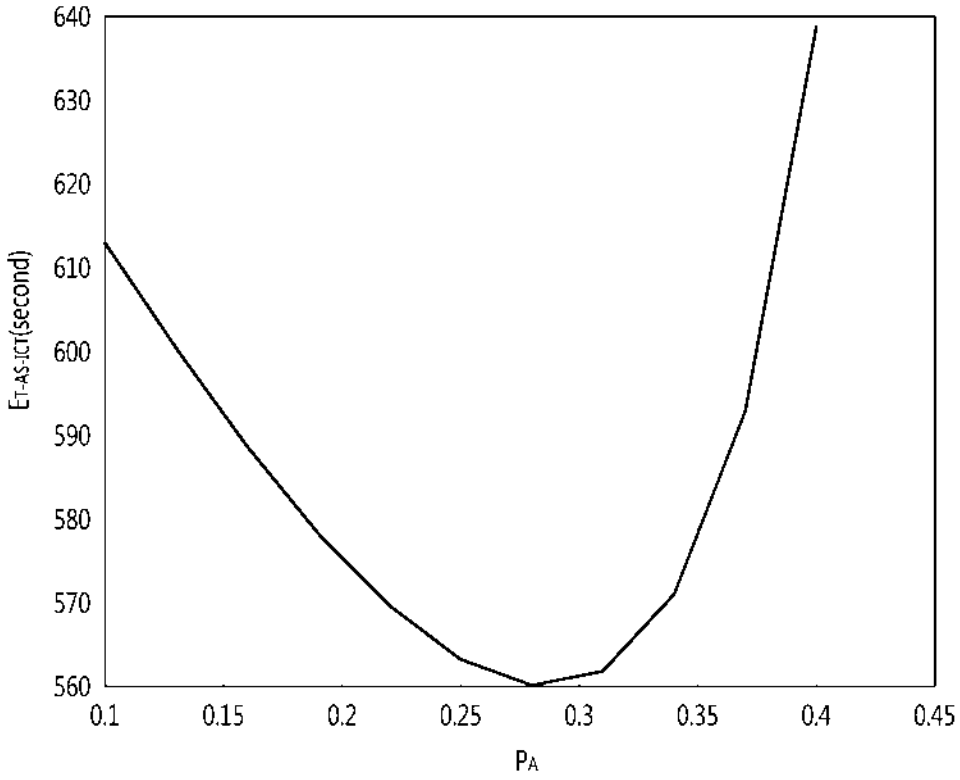
도면4


메모리
15
연산부
안전한 검증금액
정보
16
파라미터
최적화부

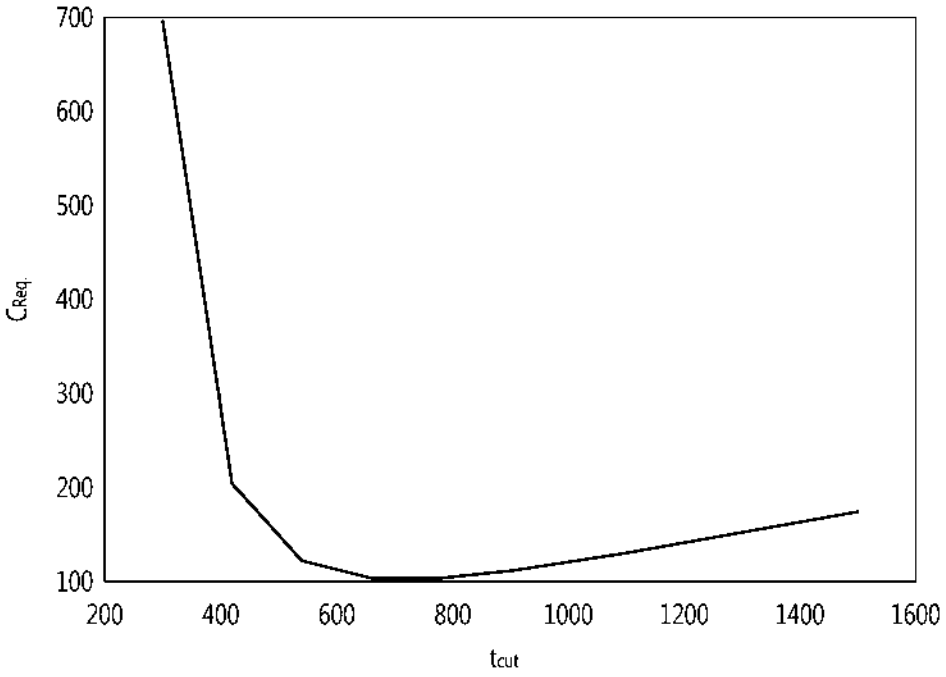
도면5


50
45
40
35
30
25
20
15
10
5
CReq.
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
0.95
PA

도면6


640
630
620
610
600
590
580
570
560
ET-AS-ICT(second)
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
PA

도면7


700
600
500
400
300
200
100
CReq.
200
400
600
800
1000
1200
1400
1600
tcut

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0019922
(43) 공개일자 2021년02월23일

(51) 국제특허분류(Int. Cl.)
H03M 13/11 (2006.01) *H03M 13/00* (2017.01)
H03M 13/45 (2006.01)
(52) CPC특허분류
H03M 13/1111 (2013.01)
H03M 13/455 (2013.01)
(21) 출원번호 10-2019-0151246
(22) 출원일자 2019년11월22일
심사청구일자 2019년11월22일
(30) 우선권주장
1020190099153 2019년08월13일 대한민국(KR)

(71) 출원인
광주과학기술원
광주광역시 북구 첨단과기로 123 (오룡동)
(72) 발명자
이홍노
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 전기전자컴퓨터공학부
(74) 대리인
김기문

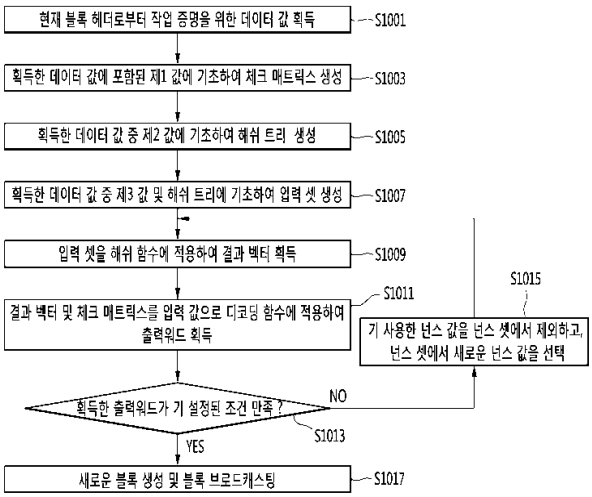
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 **부호-암호 화폐 시스템**

(57) 요 약

본 발명의 일 실시 예에 따른 채굴 장치는, 거래가 발생하면 블록을 생성하고, 작업 증명 합의 메커니즘을 통해 작업 증명을 수행하여 블록 체인을 형성하는 채굴 장치에 있어서, 블록체인 시스템을 모니터링하여, 채굴 난이도를 높이거나 낮추도록 결정하는 채굴 난이도 결정부가 포함되고, 상기 채굴 난이도 결정부에서 결정된 적용 채굴 난이도에 따라서 블록을 생성하여 작업 증명을 완료하는 제어부; 및 상기 생성된 블록을 다른 채굴 장치로 브로드캐스팅하는 네트워크부를 포함하는 것을 특징으로 한다.

대 표 도 - 도2



(52) CPC특허분류
H03M 13/6502 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711093581(NN24010)
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기술진흥센터
연구사업명	융합기술개발
연구과제명	확장가능한 탈중앙화 보안성 ECCPoW 블록체인
기 여 율	1/1
과제수행기관명	광주과학기술원
연구기간	2019.04.01 ~ 2019.12.31

명 세 서

청구범위

청구항 1

거래가 발생하면 블록을 생성하고, 작업 증명 합의 메커니즘을 통해 작업 증명을 수행하여 블록 체인을 형성하는 채굴 장치에 있어서,

블록체인 시스템을 모니터링하여, 채굴 난이도를 높이거나 낮추도록 결정하는 채굴 난이도 결정부가 포함되고, 상기 채굴 난이도 결정부에서 결정된 적용 채굴 난이도에 따라서 블록을 생성하여 작업 증명을 완료하는 제어부; 및

상기 생성된 블록을 다른 채굴 장치로 브로드캐스팅하는 네트워크부를 포함하는 채굴장치.

청구항 2

제 1 항에 있어서,

상기 채굴 난이도 결정부에는,

상기 블록체인 시스템을 모니터링하는 모니터링부;

상기 모니터링부에서 모니터링된 결과에 따라서 상기 적용 채굴 난이도를 결정하는 채굴 난이도 결정부;

상기 적용 채굴 난이도를 구현하기 위하여, 채굴 장치에 사용될 적용 체크 매트릭스를 결정하는 체크 매트릭스 결정부; 및

상기 적용 체크 매트릭스에서 상기 결정 채굴 난이도를 구현하는 해밍 가중치를 결정하는 해밍 가중치 결정부가 포함되고,

상기 체크 매트릭스는, 채굴하고자 하는 현재 블록의 헤더로부터 작업 증명을 위한 데이터 값에 기초하는 제 1 입력 값에 기초하여 구현되는 채굴 장치.

청구항 3

제 2 항에 있어서,

적어도 한 개의 상기 체크 매트릭스 및 상기 체크 매트릭스에 의해서 구현되는 상기 해밍 가중치의 조합에 의해서 구현되는 채굴 난이도와 대응하는 정보가 저장되는, 메모리가 포함되는 채굴 장치.

청구항 4

제 3 항에 있어서,

상기 메모리에는 적어도 두 개의 상기 체크 매트릭스 및 상기 적어도 두 개의 체크 매트릭스에 의해서 구현되는 해밍 가중치에 의해서 구현되는 채굴 난이도와 대응하는 정보가 저장되는 채굴장치.

청구항 5

블록체인 시스템을 모니터링하여 적용 채굴 난이도를 결정하는 단계;

현재 블록 헤더로부터 작업 증명을 위한 데이터 값을 획득하고, 획득한 데이터 값에 포함된 제 1 값에 기초하여 상기 적용 채굴 난이도를 구현할 체크 매트릭스가 결정되는 단계;

획득한 데이터 값에 포함된 제 2 값에 기초하여 해쉬 트리를 생성하는 단계;

획득한 데이터 값에 포함된 제 3 값 및 해쉬 트리에 기초하여 입력 셋을 생성하는 단계;

입력 셋을 해쉬 함수에 적용하여 결과 벡터를 획득하는 단계;

상기 결과 벡터 및 상기 체크 매트릭스에 기초하여 코드워드를 획득하는 단계;

상기 코드워드가 기 설정된 조건 셋을 만족하는지 여부를 판단하는 단계; 및

판단 결과에 기초하여 새로운 블록을 생성하거나, 다른 넌스 값을 사용하여 다시 코드워드를 획득하는 단계를 포함하는 채굴 장치 동작 방법.

청구항 6

제 5 항에 있어서,

상기 코드워드가 기 설정된 조건 셋을 만족하는지 여부를 판단하는 단계는,

작업 증명 과정의 난이도 조정을 위해 기 설정된 조건 셋을 조정하는 것이 포함되고,

상기 기 설정된 조건 셋은 상기 체크 매트릭스 별로 서로 다른, 채굴 장치 동작 방법.

청구항 7

제 6 항에 있어서,

조건 셋을 조정하는 단계는,

코드워드에 포함된 1의 개수를 지칭하는 해밍 가중치의 조합을 결정하는 것에 의해서 조정되는 채굴 장치 동작 방법.

청구항 8

제 6 항에 있어서,

새로운 블록이 생성된 경우 블록을 다른 채굴 장치로 브로드캐스팅하는 단계를 더 포함하고,

상기 블록을 다른 채굴 장치로 브로드캐스팅하는 단계는, 상기 조정된 조건 셋 정보를 블록과 함께 브로드캐스팅하는 단계를 포함하는 채굴 장치 동작 방법.

청구항 9

제 5 항에 있어서,

상기 적용 채굴 난이도를 구현할 체크 매트릭스를 결정하는 단계는, 다수의 체크 매트릭스 중에서, 상기 채굴 난이도와 가장 근접하는 채굴 난이도를 구현할 수 있는 체크 매트릭스로 결정되는 채굴 장치 동작 방법.

청구항 10

제 9 항 중에 있어서,

상기 체크 매트릭스, 및 상기 체크 매트릭스에 의해서 구현되는 해밍 가중치의 조합은 미리 메모리에 저장되는 채굴 장치 동작 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 암호 화폐 시스템에 관한 것이다. 구체적으로 개선된 작업 증명 알고리즘을 통해 블록을 생성하고, 이를 검증하는 채굴 장치를 포함하는 암호 화폐 시스템에 관한 것이다.

배 경 기 술

[0002] 2009년 사토시 나카모토가 개발한 비트코인[1]으로부터 블록 체인 기반의 암호화폐가 등장했다. 블록 체인은 공공 분산 원장을 위한 분산 데이터 처리 기술이다. Peer-to-Peer(P2P) 트랜잭션 데이터는 블록으로 기록되며, 이들 블록은 블록의 링크드 리스트(다시 말해서, 체인)를 형성한다. 네트워크의 각 사용자는 중앙 권한 없이 분산된 방식으로 전체 데이터를 저장한다. 블록 체인의 각 블록이 이전 블록의 해쉬 값을 포함하는 블록 체인을 구성하기 때문에, 블록 체인 기반 암호 화폐에서 트랜잭션을 조작하는 것을 매우 어렵게 만든다.

[0003] 대규모 P2P 네트워크에서 단일 블록 체인을 유지하려면 피어 노드 간에 결과를 상호 검증하여 시스템 무결성을 보장하는 분산 합의 메커니즘이 필요하다. 이 합의 메커니즘은 새로운 블록을 누가 만들었는지, 그리고 체인은 유효한지를 판단하는 것이 가능해야 한다.

[0004] 비트코인은 작업 증명(Proof-of-work, PoW)[2] 알고리즘을 사용하는 대표적인 암호화폐이다. 작업증명은 블록 체인 기반 암호화폐에서 가장 일반적으로 사용되는 합의 알고리즘이다. 요컨대, 특정 요구 사항을 충족하는 해쉬 값을 찾아서 확인하여 합의를 도출하는데 컴퓨팅 능력이 사용된다. 작업 증명은 블록 해쉬 값을 목표 값보다 작게 만드는 넌스(nonce) 값을 찾는 프로세스이다. 넌스는 블록 헤더에 지정된 정보 중 하나이다. 넌스 값의 집합인 넌스 셋이 정의되고, 넌스 셋에서 넌스 값을 하나씩 선택하면서 해쉬 함수에 대입한다. 이때, 한번 사용된 넌스 값은 넌스 셋에서 제외되어 다시 사용되지 않는다. 특정의 넌스 값에서 상기의 요구사항을 만족하는 해쉬 값이 발견되는 경우 해당 넌스 값을 기록하고, 블록 헤더에 삽입하여 블록을 생성한다. 이러한 일련의 과정이 작업 증명이라고 지칭된다. 작업증명 과정에서 목표 값은 넌스 값 계산의 어려움을 나타내며, 평균 10분마다 단일 블록을 생성하도록 조절된다.

선행기술문헌

비특허문헌

[0005] (비특허문헌 0001) [1] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. Consulted, 1:2012, 2008.

(비특허문헌 0002) [2] Dwork, Cynthia; Naor, Moni (1993). "Pricing via Processing, Or, Combatting Junk Mail, Advances in Cryptology". CRYPTO'92: Lecture Notes in Computer Science No. 740. Springer: 139 - 147.

(비특허문헌 0003) [3] Harald Vranken, in Current Opinion in Environmental Sustainability, 2017, 28: 1 - 9

(비특허문헌 0004) [4] https://www.asicminervalue.com/miners/ebang/ebit-e10

(비특허문헌 0005) [5] Henri Gilbert, Helena Handschuh: Security Analysis of SHA-256 and Sisters. en:Selected Areas in Cryptography 2003: pp175-193

발명의 내용

해결하려는 과제

[0006] 본 발명의 일 실시 예에 따른 채굴 장치의 동작 방법은, 기존의 작업 증명 기반의 암호화폐에서 발생하고 있는 재-중앙화(re-centralization)를 해결하고자 한다.

[0007] 더하여, 본 발명의 일 실시 예에 따른 채굴 장치의 동작 방법은 ASIC의 출현으로 소수의 채굴 업자가 해쉬파워를 독점하여 발생할 수 있는 문제를 해결하고자 한다.

[0008] 더하여, 본 발명의 일 실시 예에 따른 채굴 장치의 동작 방법은, 채굴 난이도 조절이 가능하고, 매 블록마다 입력 값이 달라지는 채굴 장치의 동작 방법을 제안하고자 한다.

[0009] 더하여, 본 발명의 일 실시 예에 따르면, 채굴 난이도를 미세하게 조절할 수 있기 때문에, 부호-암호 화폐 시스템을 더욱 견고하게 유지하는 채굴장치 및 채굴장치의 동작방법을 제안한다.

[0010] 더하여, 본 발명의 일 실시 예에 따르면, 부호-암호 화폐 시스템의 상태에 따라 적응적으로 채굴 난이도를 미세하게 조절하여, 악의적인 채굴자를 억제할 수 이는 채굴장치 및 채굴장치의 동작방법을 제안한다.

과제의 해결 수단

[0011] 본 발명의 일 실시 예에 따른 채굴 장치에는, 블록체인 시스템을 모니터링하여, 채굴 난이도를 높이거나 낮추도록 결정하는 채굴 난이도 결정부가 포함되고, 상기 채굴 난이도 결정부에서 결정된 적용 채굴 난이도에 따라서 블록을 생성하여 작업 증명을 완료하는 제어부; 및 상기 생성된 블록을 다른 채굴 장치로 브로드캐스팅하는 네

트워크부를 포함한다.

발명의 효과

[0012] 본 발명의 일 실시 예에 따른 채굴 장치의 동작 방법은 ASIC 칩으로 구현이 거의 불가능한 알고리즘으로서 작업 증명 기반의 암호화폐 시스템에서 발생하고 있는 재-중앙화(re-centralization)를 해결할 수 있다.

[0013] 또한, 본 발명의 일 실시 예에 따른 채굴 장치의 동작 방법은, 재중앙화 해결을 통해 블록 체인 기반의 암호 화폐의 본질을 회복하여 거래 신용 및 화폐 신용을 회복할 수 있다.

[0014] 또한, 본 발명의 일 실시 예에 따른 채굴 장치의 동작 방법은 작업 증명의 난이도를 필요에 따라 조정할 수 있다.

[0015] 또한, 본 발명의 일 실시 예에 따른 채굴 장치의 동작 방법은, 미세한 채굴 난이도 조절이 가능하기 때문에, 채굴되는 암호화폐의 독점을 막고, 악의적인 암호화폐의 채굴을 방지할 수 있다.

도면의 간단한 설명

[0016] 도 1은 일반적인 블록 및 블록들이 연결된 블록 체인을 나타낸다.

도 2는 본 발명의 일 실시 예에 따른 개선된 작업 증명 알고리즘을 나타내는 흐름도이다.

도 3은 앞서 설명한 코드워드와 조건 셋을 나타낸다.

도 4는 본 발명의 일 실시 예에 따른 채굴 장치의 검증 프로세스를 나타내는 흐름도이다.

도 5는 본 발명의 일 실시 예에 따른 작업 증명 알고리즘에 있어서 채굴 난이도를 조절하는 방법을 8 바이 16 체크 매트릭스를 통해 설명하는 도면이다.

도 6은 본 발명의 일 실시 예에 따른 채굴 장치를 나타내는 블록도이다.

도 7은 네 개의 체크 메트릭스를 사용하여 채굴 난이도를 미세하게 조절하는 것을 나타내는 그래프이다.

도 8은 본 발명의 다른 실시 예에 따른 개선된 조건을 만족하는 작업 증명 알고리즘을 나타내는 흐름도이다.

도 9는 다른 실시 예에 따른 채굴 장치를 나타내는 블록도이다.

도 10은 상기 채굴 난이도 결정부를 상세하게 보이는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0017] 이하에서는 도면을 참조하여 본 발명의 구체적인 실시 예를 상세하게 설명한다. 그러나 본 발명의 사상은 이하의 실시 예에 제한되지 아니하며, 본 발명의 사상을 이해하는 당업자는 동일한 사상의 범위 내에 포함되는 다른 실시 예를 구성요소의 부가, 변경, 삭제, 및 추가 등에 의해서 용이하게 제안할 수 있을 것이나, 이 또한 본 발명 사상의 범위 내에 포함된다고 할 것이다.

[0018] 첨부 도면은 발명의 사상을 이해하기 쉽게 표현하기 위하여 전체적인 구조를 설명함에 있어서는 미소한 부분은 구체적으로 표현하지 않을 수도 있고, 미소한 부분을 설명함에 있어서는 전체적인 구조는 구체적으로 반영되지 않을 수도 있다. 또한, 설치 위치 등 구체적인 부분이 다르더라도 그 작용이 동일한 경우에는 동일한 명칭을 부여함으로써, 이해의 편의를 높일 수 있도록 한다. 또한, 동일한 구성이 복수 개가 있을 때에는 어느 하나의 구성에 대해서만 설명하고 다른 구성에 대해서는 동일한 설명이 적용되는 것으로 하고 그 설명을 생략한다.

[0019] 본 발명을 설명하기에 앞서, 블록체인 기반의 암호화폐에서 수행되는 작업 증명을 간단히 설명하고, 관련 용어를 간단히 정의한다.

[0020] 도 1은 일반적인 블록 및 블록들이 연결된 블록 체인을 나타낸다.

[0021] 블록 체인(2)은 다수의 블록(1)이 연결된 형태이다.

[0022] 블록(1)이란 다수의 유효한 거래 정보의 묶음이다.

[0023] 블록 헤더(3)는 블록의 정보를 표시한다. 블록 헤더는 이전 블록의 해쉬 값을 가지고 있다.

[0024] 블록 해쉬(4)는 블록 헤더를 해쉬 함수로 계산한 값이다. 여기에서 해쉬 함수는 임의의 길이의 데이터를 고정된

길이의 데이터로 매핑하는 함수이다.

[0025] 작업 증명 알고리즘에서는 앞서 설명한 바와 같이 블록 해쉬와 목표 값을 비교하여 작업 증명 성공 여부를 판단한다. 구체적으로 작업 증명은 소위 피어 노드에서 수행되는데, 각각의 피어 노드는 헤더 정보 중 넌스 값을 바꿔가면서 해쉬 계산을 하여 목표 값 보다 작은 값이 나오는지 여부를 판단한다.

[0026] 여기에서 피어 노드는 채굴 업자 또는 채굴 장치라고도 지칭되며, 작업 증명을 수행하는 연산 장치 또는 연산 장치의 집합을 의미한다. 예를 들어 연산 장치는 고성능의 그래픽 카드 및 프로세서를 갖춘 컴퓨터가 될 수 있다.

[0027] 피어 노드는 블록 해쉬가 목표 값 보다 작을 때까지 넌스를 바꾸며, 블록 해쉬가 목표 값 보다 작으면 작업 증명 성공으로 판단하고 블록에 포함된 모든 거래를 유효한 거래로 확정한다. 그리고 유효한 블록이 생성되면, 피어 노드는 생성된 블록을 전체 네트워크로 브로드캐스팅하고, 이를 다른 피어 노드들이 승인하고 블록 체인에 추가하면 거래가 완료된다.

[0028] 여기에서, 블록이 여러 피어 노드에서 동시에 생성되어 체인이 분기될 수 있는데, 이 때, 피어 노드들은 정해진 합의 알고리즘을 사용하여 합의에 이르고, 합의된 체인을 확장시켜나간다. 가령, 비트코인 합의메커니즘은 가장 긴 체인만을 옳은 체인으로 간주한다.

[0029] 상술한 작업 증명 기반 암호화폐인 비트코인은 몇 가지 문제가 있다.

[0030] 먼저, 작업 증명을 위해 너무 많은 전력을 소비한다는 것이 있다[3]. 소위 채굴 난이도라고 불리는 암호 퍼즐 난이도가 점차 증가하고 있는데 이는 비트코인 채굴에 참여하는 업자의 수가 늘어나고, 채굴네트워크 전체의 해쉬파워가 증가하기 때문이다. 해쉬파워가 클수록 작업 증명의 속도가 빠르며 결과적으로 블록을 생성하는 속도가 빨라진다. 비트코인 프로토콜은 블록 생성 시간을 평균적으로 10분마다 한 블록씩 채우는 것을 목표로 한다. 이를 위해 비트 코인 프로토콜은 2016 블록마다 해쉬 난이도를 재구성한다[1]. 따라서, 해쉬파워가 커지면 난이도는 증가하게 된다. 이로 인하여, 채굴을 위해 더 많은 전력이 소비가 되는 것이다.

[0031] 암호 화폐에 대한 채굴업자의 수가 많을수록 가능한 공격에 대한 보호 기능이 향상되므로 블록 체인 불변성 및 암호 화폐의 신뢰성 유지에 도움이 된다. 비트 코인 네트워크의 초기 단계(2010 ~ 2013년)에서는 가정용 컴퓨터를 사용하는 일반 사람들도 비트코인의 작업증명을 수행할 수 있었다.

[0032] 그러나, 시간이 갈수록 가정용 컴퓨터의 컴퓨팅 파워만으로는 작업 증명을 빠르게 수행할 수 없었으며, 비트코인 채굴을 위한 컴퓨팅 플랫폼이 GPU, FPGA에서 최근에 ASIC에 이르면서 소수의 채굴업자들이 컴퓨팅 파워를 독점하는 현상이 나타나고 있다. 이를 통해 비트코인 네트워크가 중앙 집중화되어 비트코인 블록체인의 정체성인 탈-중앙화(decentralization)와 멀어지고 있다. 따라서, 압도적인 컴퓨팅 파워를 갖는 소수의 채굴업자에 의해 블록체인이 변조될 가능성이 대두되면서 블록체인에 대한 대중의 신뢰가 붕괴되고 있는 시점이다.

[0033] 이러한 비트코인을 포함한 모든 작업증명(Proof-of-Work) 기반 블록체인 네트워크의 재-중앙화(recentralization) 문제가 유발되는 원인으로, 암호 퍼즐의 제한된 가용성(limited availability)이 있다. 비트코인 합의메커니즘에서의 작업 증명 방식은 고정된 해쉬함수 (e.g. SHA-256) 알고리즘에 따라 이루어져, 결과적으로 비트코인 작업 증명 알고리즘에 최적화된 ASIC 칩을 탄생시켰다[4]. ASIC 칩은 상대적으로 고가이나, 비트코인 작업 증명에 최적화되어 ASIC 칩을 통해 채굴을 수행하는 소수의 채굴업자들이 가진 해쉬파워가, 전체네트워크의 해쉬파워에 큰 부분을 차지하여, 압도적인 채굴 영향력을 갖게 되었다.

[0034] 따라서 이러한 블록체인 네트워크의 재-중앙화를 해소하기 위해서 새로운 작업 증명 알고리즘이 요구되고 있다. 새로운 작업 증명 알고리즘은 아래와 같은 요건을 만족할 것이 요구된다.

[0035] 1) 퍼즐은 풀기 어려워야 하나, 반대로 체크는 쉬워야 한다.

[0036] 2) 퍼즐은 외부 공격으로부터 견고한 저항성을 가져야 한다.

[0037] 3) 풀린 퍼즐의 넌스 값은 재 사용되지 않는다.

[0038] 4) 퍼즐의 난이도는 조절 가능하여야 한다.

[0039] 5) CPU를 갖고 있는 정도라면 누구든지 작업 증명에 참여할 수 있어야 한다.

[0040] 이하에서는 상술한 요구 조건을 만족하는 본 발명의 일 실시 예에 따른 개선된 PoW기반 블록체인 네트워크에서의 작업 증명 알고리즘을 설명한다. 개선된 요구 조건 중 하나는

[0041] 6) 작업증명에 쓰이는 함수가 매 블록마다 변화하여 바뀔 수 있어야 한다.

[0042] 이다.

[0043] 도 2는 본 발명의 일 실시 예에 따른 개선된 조건을 만족하는 작업 증명 알고리즘을 나타내는 흐름도이다.

[0044] 도 2에서 설명하는 작업 증명 알고리즘은 상술한 채굴 장치를 통해 수행될 수 있으며, 구체적으로 채굴 장치에 마련된 CPU와 같은 프로세서에서 수행될 수 있다.

[0045] 채굴 장치는 현재 블록 헤더로부터 작업 증명을 위한 데이터 값을 획득한다(S1001). 여기에서 현재 블록이란, 채굴자가 현재 만들고자 하는 블록으로, 아직 작업증명이 완료되지 않아 체인에 연결되지 않은 블록을 말한다. 블록 체인에 포함된 블록 헤더에는 블록에 관한 하나 이상의 정보를 포함할 수 있다. 블록 헤더가 포함하는 정보는 버전 정보, 난이도 정보, 타임스탬프 정보, 넌스 정보, 이전 블록의 해쉬 값 정보 또는 트랜잭션 셋 정보 중 적어도 하나를 포함할 수 있다.

[0046] 채굴 장치는 획득한 데이터 값에 포함된 제 1 값에 기초하여 체크 매트릭스를 생성한다(S1003). 여기에서 체크 매트릭스를 생성하는데 사용되는 제 1 값은 이전 블록의 해쉬 값일 수 있다. 체크 매트릭스는 Galois Field GF(q) 의 엘레멘트 (q = 2 인 경우는 0 또는 1 바이너리 엘레멘트)로 구성되는 $N_c - N_m$ 바이 Nc 매트릭스일 수 있다. 가령 $N_c=128$ 이고 $N=256$ 이 될 수 있다. 채굴 장치는 현재 블록 인덱스 t (양의정수) 에서 이전 블록의 해쉬 값(h_{t-1})을 이용하여 체크 매트릭스(F_t)를 생성할 수 있다. 여기에서 체크 매트릭스의 생성은 부호이론 분야에서 일반적으로 알려진 것을 사용할 수 있다.

[0047] 채굴 장치는 획득한 데이터 값에 포함된 제 2 값에 기초하여 해쉬 트리를 생성한다(S1005). 여기에서 해쉬 트리를 생성하는데 사용되는 제 2 값은 트랜잭션 셋 정보일 수 있다. 일 실시 예에서, 채굴 장치는 제 2 값에 기초하여 해쉬 트리 값을 생성할 수 있다.

[0048] 채굴 장치는 획득한 데이터 값에 포함된 제 3 값 및 생성된 해쉬 트리에 기초하여 입력 셋(S)을 생성한다(S1007). 여기에서 사용되는 제 3 값은 이전 블록의 버전 정보, 난이도 정보 또는 타임스탬프 정보 중 적어도 하나일 수 있다. 또한, 블록 헤더에 포함된 다른 정보가 추가적으로 제 3 값으로 사용될 수도 있다.

[0049] 채굴 장치는 입력 셋을 해쉬 함수에 적용하여 결과 벡터(r)를 획득한다(S1009). 여기에서 사용되는 해쉬 함수는 SHA 256 [5] 함수일 수 있으며, 그 밖에 보안성이 검증된 다른 함수를 사용할 수도 있다. 이 때 채굴 장치는 특정의 넌스 값에 대한 해쉬 함수 출력인 결과 벡터를 생성한다.

[0050] 채굴 장치는 해쉬 함수의 출력인 결과 벡터(r)와 단계 S1003에서 생성한 체크 매트릭스를 입력 값으로 하여 디코딩 함수에 적용한다(S1011). 채굴 장치는 디코딩 함수의 출력 값으로 출력워드(c^)를 획득할 수 있다.

[0051] 여기에서 디코딩 함수는 Error-Correction Code의 디코더, 통신이론의 Sphere-Decoding 신호수신기, 압축센싱의 Sparse 신호복원 알고리즘 또는 수학의 Inverse 문제의 해를 구하는 알고리즘 등 하나의 입력 값에 대한 유일한 출력값을 만드는 함수가 사용될 수 있다.

[0052] 바람직한 일 실시 예로서, Error correction coding에 관한 디코딩 함수가 사용되는 경우, Error-Correction Code는 Noisy 채널을 사이로 송신기와 수신기가 통신할 때 사용한다. Noisy채널을 통해 수신한 워드는 채널이 발생시키는 에러 때문에 송신한 워드와 다르게 된다. 이 에러를 잡아내고 정정하기 위해 오류정정부호가 사용된다. 송신기는 메시지 벡터를 1-대-1 매핑하여, 즉 인코딩 (Encoding) 하여 코드워드를 생성하고, 메시지 벡터 대신에 생성된 코드워드를 수신기로 전송한다. 수신기는 수신한 에러를 포함한 수신워드를 Decoding 하여 채널에서 발생한 오류를 제거할 수 있다. 수신워드를 입력해주면 Encoding 하는 함수의 역 문제를 풀어 송신된 코드워드를 복원하는 복원함수를 Decoder 또는 디코딩 함수라고 부른다.

[0053] 에러를 스스로 제거 할 수 있은 여유를 만들어 내기 위해서, 코드워드 벡터의 크기 N_c 를 메시지 벡터의 크기 N_m 보다 길게 만든다. 그리고, $R = N_m/N_c$ 을 코드 Rate이라고 칭한다. 코드 Rate이 작으면, 채널 에러에 더욱 강인해 진다; 그러나 Decoding하기위해 보다 많은 컴퓨팅을 필요로 한다.

[0054] 본 발명은 Error-Correction Code에서 쓰는 Decoder를 작업증명에 적용하는 것을 발명의 한 실시 예로 보여준다. Decoder 함수를 SHA-256과 같은 암호함수와 묶어 합성함수를 만드는 것이다. 즉, 암호함수의 출력값을 디코딩 함수의 입력 값의 하나로 넣는 것이다.

[0055] 작업 증명에서는 합성된 함수의 출력워드가 주어진 조건을 만족하는 합성함수의 입력값, 즉 넌스 값을 찾는 것은 매우 어렵지만, 일단 찾은 코드워드 값이 조건에 적합한지 여부를 검증하는 것은 쉬운 경우를 말한다. 그러므로 이렇게 만들어낸 합성함수는 작업 증명 알고리즘의 요건 중 1)을 만족시킨다. 바람직한 실시 예에서, 디코딩 함수로서 graph-decoder 가 사용될 수 있다. 또한, 추가적으로 디코딩 함수는 리니어 graph-decoder 분야에서 가장 빠른 코드워드 매핑이 가능한 알고리즘이 사용될 수 있다.

[0056] 채굴 장치는 획득한 Decoder의 출력워드가 기 설정된 조건을 만족하는지 여부를 판단한다(S1013). 채굴 장치는 매핑된 출력워드에 관한 조건 셋을 미리 가지고 있으며, 단계 S1011에서 획득한 출력워드 값이 코드워드인지 여부와 코드워드인 경우 기 설정된 조건 셋을 만족하는지 여부를 판단한다.

[0057] 도 3은 앞서 설명한 코드워드와 조건 셋을 나타낸다.

[0058] 도 3에서 예를 든 바와 같이, 2^{256}개의 벡터(예를 들어 SHA 함수의 출력)가 있으며, 그 중 1/4 코드 레이트로 2^{64}개 정도의 코드워드가 존재한다. 디코딩 함수(Dec())에 벡터를 입력하는 경우에 하나의 출력워드로 매핑이 되는데, 여기에서 매핑된 출력워드가 조건 셋에 포함되는 코드워드인지 여부를 채굴 장치가 판단한다.

[0059] 다시 도 2로 돌아온다.

[0060] 앞서 설명한 바와 같이, 매핑된 출력워드가 조건 셋을 만족하는 경우, 채굴 장치는 작업 증명이 완료되었다고 판단하여 현재의 넌스 값을 기록하면서 새로운 블록을 생성하고 이를 다른 채굴 장치로 브로드캐스팅한다(S1017). 따라서, 기존 비트코인 작업 증명에서 해쉬 함수 출력 값만으로 작업 증명 완료 여부를 판단하는 것과 달리 디코딩 함수를 이용하여 출력워드 값을 검증하는 과정이 추가되는바 정형화된 알고리즘에서 발생할 수 있는 문제(ASIC을 이용한 재중앙화)를 해결할 수 있다.

[0061] 다시 말해서, 본 발명의 일 실시 예에 따른 합성함수에 의한 작업 증명 알고리즘에서는 Decoder부분의 입력 중 하나인 체크 매트릭스를 매 블록마다 변경할 수 있기 때문에, 매 블록마다 다른 Decoder 퍼즐을 작업 증명으로 수행해야 한다. 이전 블록 해쉬 값에 체크 매트릭스가 의존하는 바 역시 매 블록마다 다른 입력 값이 사용되어 결과적으로 ASIC 칩의 출현을 억제할 수 있다. 즉, ASIC 칩을 사용한 채굴을 힘들게 함으로써, 상대적으로 해쉬 파워가 떨어지는 CPU 로도 누구나 작업증명에 참여할 수 있으므로 새로운 작업 증명 알고리즘의 요건 5) 를 만족한다. 또한, 디코딩에 사용되는 함수 또한 원웨이 특성을 갖는 또 다른 종류의 함수로 변경 가능한 것으로 이 역시 고정화된 알고리즘으로 인한 ASIC 칩의 출현을 억제하는 요소가 된다. 또한 특정 퍼즐문제에 대한 넌스 값을 찾았더라도 매 블록마다 다른 입력 값이 퍼즐을 만드는데 사용되므로, 기존의 넌스 값을 새로운 블록의 넌스 값으로 재사용 할 수 없다. 즉, 새로운 작업 알고리즘의 요건 3)을 만족한다.

[0062] 만약, 매핑된 출력워드가 조건 셋을 만족하지 못하는 경우, 채굴 장치는 기 사용한 넌스 값을 넌스 셋에서 제외하고, 넌스 셋에서 사용되지 않은 새로운 넌스 값을 선택하여 다시 단계 S1009로 돌아간다(S1015). 채굴 장치는 매핑된 출력워드가 조건을 만족할 때까지 새로운 넌스 값을 선택하면서 각 단계를 반복한다.

[0063] 도 4는 본 발명의 일 실시 예에 따른 채굴 장치의 검증 프로세스를 나타내는 흐름도이다.

[0064] 앞서 도 2 내지 3에서 설명한 단계는 채굴 장치의 채굴 프로세스이며, 채굴 장치는 작업 증명이 완료되면 새롭게 생성된 블록을 다른 채굴 장치로 브로드캐스팅한다. 그리고 브로드캐스팅된 블록을 다른 채굴 장치들이 검증하여 정당한 블록임이 인정된다면 채굴 장치들은 해당 블록을 기존의 블록 체인에 더하여 블록 체인 원장을 업데이트 한다.

[0065] 채굴 장치는 블록을 생성한 다른 채굴 장치로부터 브로드캐스팅 된 블록의 블록 헤더 데이터를 획득한다(S2001). 여기에서 채굴 장치가 획득하는 블록 헤더 데이터는 앞서 설명한 것과 동일하다.

[0066] 채굴 장치는 블록 헤더 데이터에 포함된 해쉬 값에 기초하여 체크 매트릭스를 생성한다(S2003). 여기에서 체크 매트릭스를 생성하는 방법은 앞서 설명한 것과 동일하다.

[0067] 채굴 장치는 블록 헤더 데이터에 기초하여 해쉬 함수의 출력 값을 획득한다(S2005). 여기에서 해쉬 함수의 출력 값을 획득하는 단계는 앞서 설명한 것과 동일하다.

[0068] 채굴 장치는 해쉬 함수의 출력 값과 체크 매트릭스를 입력 값으로 하는 디코딩 함수 출력 값으로 출력워드를 획득한다(S2007). 여기에서 사용하는 디코딩 함수는 앞서 설명한 것과 동일하다.

[0069] 채굴 장치는 출력워드가 조건을 만족하는지 여부를 판단한다(S2009). 여기에서 조건은 앞서 채굴 프로세스에서

의 조건과 동일한 조건으로 설정되는 것이다. 검증을 위한 조건 셋은 전체 채굴 장치에 대하여 일괄적으로 설정된 것일 수 있으며, 블록이 브로드캐스팅될 때 함께 전달될 수도 있다.

[0070] 채굴 장치는 판단 결과 출력워드가 조건 셋을 만족하는 경우 브로드캐스팅된 블록을 승인하고 이를 기존의 블록 체인에 연장하여 원장을 업데이트한다(S2011).

[0071] 반면에, 채굴 장치는 판단 결과 출력워드가 조건 셋을 만족하지 못하는 경우 브로드캐스팅된 블록이 정당한 블록이 아니라고 판단하여 브로드캐스팅된 블록 승인을 거절한다(S2013).

[0072] 결과적으로 상술한 검증 방법은 이미 알고 있는(또는 기 설정되어 있는) 디코딩 함수에 전달 받은 값을 넣어 조건 셋에 부합하는지 여부만 검증하면 되는 것으로, 넌스 값을 변경하면서 조건에 맞는 코드워드를 찾아 작업증명을 완료하는 과정에 비해, 상대적으로 매우 간단하여 블록 체인에 적합한 검증 방법이다. 따라서, 외부의 공격자가 블록의 내용을 조작하더라도 검증자가 조작여부를 쉽게 알 수 있으므로 새로운 작업 증명 알고리즘의 요건 2)를 만족한다.

[0073] 도 5는 본 발명의 일 실시 예에 따른 작업 증명 알고리즘에 있어서 채굴 난이도를 조절하는 방법을 8 바이 16 체크 매트릭스를 통해 설명하는 도면이다.

[0074] 도 5(a)는 8 바이 16 체크 매트릭스의 예이다. 실제로는 이것보다 훨씬 큰 크기의 체크 매트릭스를 사용하나 설명의 용이성을 위해 8 바이 16 체크 매트릭스를 예시로 설명한다.

[0075] 도 5(b)는 (a)에서의 체크 매트릭스에 기초하여 코드워드 값을 정리한 표이다. n은 코드워드의 해밍 가중치(Hamming weight)를 나타내는데, 다시 말해서 코드워드에 포함된 0이 아닌 수의 개수를 지칭한다. V_n은 1을 해밍 가중치에서 인디케이팅 하는 수만큼 가지고 있는 조건 셋을 나타낸다. p는 전체 코드워드의 개수 중 V_n 조건 셋에 포함되는 코드워드 개수의 비율을 나타낸다.

[0076] 도 5(b)의 표에서 나타내는 바와 같이 n 개의 1을 포함하는 조건 셋의 크기가 다름을 알 수 있으며, 그에 대한 각각의 점유율을 알 수 있다. 따라서, 조절하고자 하는 암호 퍼즐의 난이도에 따라 해밍 가중치를 결정할 수 있다. 예를 들어 난이도를 낮게 하는 경우에 해밍 가중치를 7 내지 10으로 설정할 수 있다. 반대로 난이도를 높이고자 하는 경우 해밍 가중치를 3으로 설정할 수 있다. 각각의 경우 조건 셋을 만족하는 코드워드가 선택될 확률이 각각 약55%와 2.5%로 차이가 나는바, 후자의 경우 더 많은 연산을 수행하여야 조건 셋을 만족하는 코드워드 값을 획득할 수 있어 결과적으로 작업 증명의 난이도가 상승하게 된다. 즉, 상기한 코드 Rate을 바꾸고, 메이트릭스의 크기를 조절하여 Decoder 퍼즐을 푸는데 필요한 컴퓨팅량을 조절하여 퍼즐의 난이도를 유연하게 조절할 수 있으므로 새로운 작업 증명 알고리즘의 요건 4) 를 만족한다.

[0077] 상기되는 채굴 난이도의 조절은 도 7이하에서 다른 실시예를 통하여 더 구체적으로 설명한다.

[0078] 도 6은 본 발명의 일 실시 예에 따른 채굴 장치를 나타내는 블록도이다.

[0079] 본 발명의 일 실시 예에 따른 채굴 장치는 제어부(110), 네트워크부(120) 및 저장부(130)을 포함한다.

[0080] 제어부(110)는 프로세서를 지칭할 수 있으며, 마이크로 프로세서 또는 제어기를 포함할 수 있다.

[0081] 도 6에 도시된 바와 같이, 제어부(110)는 현재 블록 헤더 데이터 획득부(111), 입력 값 생성부(112), 해쉬 함수 적용부(113), 코드워드 획득부(114) 및 블록생성부(115)를 포함할 수 있다.

[0082] 현재 블록 헤더 데이터 획득부(111)는 채굴하고자 하는 현재 블록의 헤더를 추출하여 헤더에 포함된 데이터를 획득한다. 여기에서 현재 블록 헤더 데이터 획득부(111)가 획득하는 데이터는 버전 정보, 난이도 정보, 타임스탬프 정보, 넌스 정보, 이전 블록의 해쉬 값 정보 또는 트랜잭션 셋 정보 중 적어도 하나를 포함할 수 있다.

[0083] 입력 값 생성부(112)는 현재 블록 헤더 데이터에 기초하여 복수의 입력 값을 생성한다. 일 실시 예에서 입력 값 생성부(112)는 이전 블록 해쉬 값에 기초하여 체크 매트릭스를 생성할 수 있다. 또 다른 실시 예에서 입력 값 생성부(112)는 트랜잭션 셋에 기초하여 해쉬 트리 값을 생성할 수 있다. 또 다른 실시 예에서 입력 값 생성부(112)는 해쉬 트리 값 및 다른 헤더 데이터에 기초하여 해쉬 함수 입력 셋을 생성할 수 있다. 여기에서 다른 헤더 데이터 값이란 예를 들어 버전 정보, 난이도 정보 또는 타임스탬프 정보 중 적어도 어느 하나일 수 있다.

[0084] 해쉬 함수 적용부(113)는, 입력 값 생성부(112)에서 생성된 입력 셋을 입력 값으로 하여 해쉬 함수 출력을 획득한다. 여기에서 해쉬 함수는 SHA 함수일 수 있으며, 다른 보안성이 확보된 함수도 적용 가능하다.

[0085] 출력워드 획득부(114)는 체크 매트릭스와 해쉬 함수 출력 값에 기초하여 출력워드를 획득한다. 일 실시 예에서

출력워드 매핑 알고리즘은 출력워드는 입력 값에서 가까운 하나의 코드워드를 찾는 방법으로 획득될 수 있다. 또 다른 실시 예에서, 출력워드 매핑 알고리즘은 graph codeword mapping이 사용될 수 있다.

[0086] 블록생성부(115)는 획득한 출력워드를 검증하여 검증 결과에 따라 블록을 생성한다. 블록생성부(115)는 넌스 값을 발생시키고, 해당 넌스값에 대응하는 출력워드가 조건 셋을 만족하는지 여부를 판단한다. 블록생성부(115)는 획득한 출력워드가 조건 셋을 만족하는 경우 해당 넌스값을 저장하고 블록을 생성한다. 반면에 블록생성부(115)는 출력워드가 조건 셋을 만족하지 못하는 경우 넌스 값을 바꾸고 다시 출력워드 획득부로부터 새로운 출력워드값을 받아 검증 작업을 수행한다.

[0087] 네트워크부(120)는 유무선 통신 장치이다. 네트워크부(120)는 다른 채굴 장치와 연결을 수행하고, 생성된 블록을 브로드캐스팅한다. 또한 네트워크부(120)는 다른 채굴장치가 생성한 블록을 수신하여 이를 현재 블록 헤더 데이터 획득부(111)에 전달할 수도 있다.

[0088] 저장부(130)는 제어부(110)에서 사용되는 각종 인스트럭션을 저장한다. 또한 저장부(130)는 제어부(110)에서 생성된 블록이 연결된 블록 체인 원장을 저장한다. 저장부(130)는 메모리 장치일 수 있다.

[0089] 이하에서는 상기 채굴 난이도를 미세하게 조절하는 다른 실시예를 제시한다. 이하의 설명은 이미 제시된 설명을 그대로 적용되는 것을 일 예로 하므로, 구체적인 설명이 없는 부분은 이미 제시된 설명이 적용될 수 있다.

[0090] 도 7은 네 개의 체크 메트릭스를 사용하여 채굴 난이도를 미세하게 조절하는 것을 나타내는 그래프이다.

[0091] 도 7을 참조하면, 상기 체크 매트릭스의 열의 수가 64개, 28, 256, 및 512개인 경우에 열의 개수가 증가함에 따라서 채굴 난이도가 상승할 수 있는 것을 나타낸다.

[0092] 도 5에 제시된 바와 같이 체크 매트릭스의 크기가 8 바이 16의 경우에, 상기 해밍 가중치를 선택하는 경우의 수는 해밍 가중치의 조합으로 결정될 수 있다.

[0093] 예를 들어, 상기 채굴 난이도를 극단적으로 높게 하고자 하는 경우에는 점유율이 가장 낮은 단일의 해밍 가중치를 설정할 수 있다. 상기 채굴 난이도를 극단적으로 낮게 하고자 가능한 모든 해밍 가중치를 설정할 수 있다. 이는, 상기 해밍 가중치의 조합(combination)의 문제이다. 상기 해밍 가중치의 조합을 통하여 다양한 경우의 수로, 조건 셋을 만족하는 코드워드가 선택될 확률을 조절할 수 있다. 이를 통하여 채굴 난이도를 조절할 수 있다.

[0094] 상기 채굴 난이도와 대응하는 상기 해밍 가중치 설정의 조합은 서로 대응되어 미리 결정될 수 있다.

[0095] 상기 채굴 난이도에 대응하는 상기 해밍 가중치 설정의 조합은 미리 계산되어 메모리(도 9의 330참조)에 미리 저장되어 있을 수 있다. 도 7에서 오른쪽으로부터 첫 번째 우상향 곡선은, 체크 매트릭스의 크기가 48 바이 64인 경우에, 상기 채굴 난이도가 가장 낮은 해밍 가중치 설정일수록 앞에 위치하고, 상기 채굴 난이도가 가장 높은 해밍 가중치 설정일수록 뒤에 위치하는 것을 도시한다.

[0096] 상기 그래프 정보는, 매칭 테이블 또는 매칭을 위한 함수 정보로서 상기 메모리(330)에 저장될 수 있다.

[0097] 상기 채굴 난이도는, 상기 체크 매트릭스에 따라서 최고 난이도가 결정되어 있다. 예를 들어, 도 5의 경우에는 상기 해밍 가중치가 14인 경우에 조건 셋을 만족하는 코드워드가 선택될 확률이 0.46%로 가장 낮다. 따라서, 더 높은 채굴 난이도를 설정하기 위해서는 더 복잡한 체크 메트릭스를 적용해야 한다.

[0098] 더 복잡한 체크 메트릭스를 구현하기 위하여, 체크 메트릭스의 크기를 증가시킬 수 있다.

[0099] 도 7에서 오른쪽으로부터 두번째 우상향 곡선은, 체크 매트릭스의 크기가 96 바이 128인 경우에, 상기 채굴 난이도가 가장 낮은 해밍 가중치 설정일수록 앞에 위치하고, 상기 채굴 난이도가 가장 높은 해밍 가중치 설정일수록 뒤에 위치하는 것을 도시한다.

[0100] 마찬가지로, 도 7에서 세번째 우상향 곡선은 체크 매트릭스의 크기가 192 바이 256인 경우이고, 도 7에서 네번째 우상향 곡선은 체크 매트릭스의 크기가 384 바이 512인 경우를 나타낸다.

[0101] 이와 같은 방식으로 상기 채굴 난이도는 무한하게 어렵게 조정해 나갈 수 있다. 물론, 어느 때에는 이전보다 높은 채굴 난이도로 변경할 수도 있고, 어느 때에는 이전보다 낮은 채굴 난이도로 변경할 수도 있다. 따라서, 현재 상태에 적응적인 채굴 난이도를 구현할 수 있다.

[0102] 한편, 어느 하나의 체크 매트릭스로부터 획득 가능한 최고의 채굴 난이도를 넘어서는 경우에는, 다른 하나의 체

크 매트릭스로 이동하여 더 높은 채굴 난이도를 구현할 수 있는 것은 당연하다.

[0103] 도 8은 본 발명의 다른 실시 예에 따른 개선된 조건을 만족하는 작업 증명 알고리즘을 나타내는 흐름도이다. 본 실시예는 블록체인 시스템을 모니터링하여 채굴 난이도를 결정하는 것이 추가되는 것이 특징적으로 다르고, 그 외에는 도 2에 제시되는 설명이 그대로 적용될 수 있다.

[0104] 도 8을 참조하면, 채굴 장치는 현재 블록 헤더로부터 작업 증명을 위한 데이터 값을 획득한다(S3001). 여기에서 현재 블록이란, 채굴자가 현재 만들고자 하는 블록으로, 아직 작업증명이 완료되지 않아 체인에 연결되지 않은 블록을 말한다. 블록 체인에 포함된 블록 헤더에는 블록에 관한 하나 이상의 정보를 포함할 수 있다. 블록 헤더가 포함하는 정보는 버전 정보, 난이도 정보, 타임스탬프 정보, 넌스 정보, 이전 블록의 해쉬 값 정보 또는 트랜잭션 셋 정보 중 적어도 하나를 포함할 수 있다.

[0105] 채굴 장치는 풀 블록체인 노드(full blockchain node)와 접속하여 현재 블록체인 시스템의 상태를 파악하고, 그에 따라서 채굴 난이도를 결정할 수 있다(S3003).

[0106] 예를 들어, 채굴장치는, 현재 채굴 난이도가 낮아 다수의 채굴 및 악의적인 채굴이 있는 경우에는, 채굴 난이도를 높이도록 채굴 난이도를 결정할 수 있다. 채굴장치는, 현재 채굴에 참여하는 채굴자의 숫자가 적은 경우 경우에는, 채굴 난이도를 낮추도록 채굴 난이도를 결정할 수 있다.

[0107] 이후에는 결정된 난이도에 따라서, 채굴 장치는 획득한 데이터 값에 포함된 제 1 값에 기초하여 체크 매트릭스를 생성하고(S3005). 및 채굴 장치는 획득한 데이터 값에 포함된 제 2 값에 기초하여 해쉬 트리를 생성하는 것(S3007) 등의 도 2에 제시되는 과정이 동일하게 수행될 수 있다.

[0108] 한편, 상기 코드워드가 기 설정된 조건 셋을 만족하는지 여부를 판단하는 것에 있어서는, 결정된 채굴 난이도를 구현하는 해밍 가중치 조합에 대응하는 조건 셋에 의해서 동작될 수 있다.

[0109] 다시 한번 설명하거니와, 이때, 상기 작업 증명 알고리즘의 동작은 현재 상태에 대응하여 결정된 채굴 난이도를 준수하여 수행된다. 구체적으로, 상기 체크 매트릭스의 선택과 상기 해밍 가중치는, 결정된 상기 채굴 난이도에 대응하는 값으로 결정될 수 있다. 이때 도 7에 제시되는 그래프에 대응하여 수행될 수 있다. 도 7과 관련되는 정보는 메모리(330)에 저장될 수 있다.

[0110] 도 9는 다른 실시 예에 따른 채굴 장치를 나타내는 블록도이다. 본 실시예는 채굴 난이도 결정부가 포함되는 것이 특징적으로 다르다. 따라서 본 도면에서 구체적으로 제시되지 않은 설명은 도 6에 제시되는 채굴 장치의 설명이 적용될 수 있다.

[0111] 도 9를 참조하면, 본 발명의 다른 실시 예에 따른 채굴 장치는 제어부(210), 네트워크부(220) 및 저장부(230)을 포함한다.

[0112] 상기 제어부(210)는, 현재 블록 헤더 데이터 획득부(211), 채굴 난이도 결정부(212), 입력 값 생성부(213), 해쉬 함수 적용부(214), 코드워드 획득부(215) 및 블록생성부(216)를 포함할 수 있다.

[0113] 상기 채굴 난이도 결정부(211)는 현재 블록체인 시스템에 적용할 난이도를 결정하는 블럭이다. 그 외에 제어부에 포함되는 구성요소들은 도 6에 제시되는 설명이 그대로 적용될 수 있다.

[0114] 도 10은 상기 채굴 난이도 결정부를 상세하게 보이는 도면이다.

[0115] 도 10을 참조하면, 상기 채굴 난이도 결정부(211)에는, 블록체인 시스템의 상태를 파악하는 모니터링부(320)가 포함된다. 상기 모니터링부(320)는 풀 블럭체인 노드에 접속하여 현재 및 과거의 채굴상태를 파악할 수 있다.

[0116] 상기 모니터링부(320)에서 감지된 정보에 따라서 채굴 난이도를 결정하는 난이도 결정부(310)가 더 포함될 수 있다.

[0117] 상기 채굴 난이도의 결정은 다양하게 결정되어, 현재 블록생성이 적용될 수 있다. 예를 들어, 현재 채굴 난이도가 낮아 다수의 채굴 및 악의적인 채굴이 있는 경우에는 채굴 난이도를 높이도록 채굴 난이도를 결정할 수 있다. 현재 채굴에 참여하는 채굴 장치의 참여수가 적은 경우에는 채굴 난이도를 낮추도록 채굴 난이도를 결정할 수 있다.

[0118] 상기 채굴 난이도 결정부(310)에서 결정된 상기 채굴 난이도를 구현하기 위하여, 어느 체크 매트릭스를 적용할지를 결정하는 체크 매트릭스 결정부(340)가 더 포함될 수 있다.

[0119] 상기 채굴 난이도 결정부(310)에서 결정된 상기 채굴 난이도를 구현하기 위하여, 어느 해밍 가중치를 적용할 지

를 결정하는 해밍 가중치 결정부(340)가 더 포함될 수 있다.

[0120] 이미 설명된 바와 같이, 체크 매트릭스의 크기가 크면 클수록 더 많은 컴퓨팅이 요구된다. 따라서 낮은 채굴 난이도에 적용하기 위하여 엘레멘트의 수가 많은 체크 매트릭스를 적용할 필요는 없다. 채굴 난이도가 높은 경우에는 엘레멘트의 수가 작은 체크 매트릭스로는 구현할 수가 없다. 이 두 가지 관점을 만족하는 체크 매트릭스를 상기 체크 매트릭스 결정부(340)에서 결정할 수 있다.

[0121] 도 7에서 설명된 바와 같이, 상기 해밍 가중치의 조합을 통하여 동일한 체크 매트릭스를 통하여 미세하게 채굴 난이도를 조정할 수 있다. 도 7에서는 상기 체크 매트릭스 결정부(340)에서 결정된 체크 매트릭스에서 결정된 채굴 난이도를 구현하는 해밍 가중치를 결정할 수 있다.

[0122] 한편, 본 실시예에 따르면, 채굴 난이도를 최적으로 미세하게 조정하는 것이 가능하다.

[0123] 예를 들어, 어느 하나의 제 1 체크 매트릭스에서 구현되는 난이도는 최고 난이도가 10(설명을 위한 숫자로서 높을 수록 난이도가 큰 것을 의미할 뿐으로서, 이하 동일하다)까지이고, 다른 하나의 제 2 체크 매트릭스에서 구현되는 난이도는 최고 난이도가 100까지 일 수 있다.

[0124] 이 경우에, 현재 미세하게 조정하고자 하는 적용 채굴 난이도는 8이라고 하자. 이 경우에 상기 제 1 체크 매트릭스에서, 해밍 가중치를 조정하여 얻을 수 있는 8과 인접하는 채굴 난이도는 7.4, 및 8.1이 있다. 마찬가지로 상기 제 2 체크 매트릭스에서, 해밍 가중치를 조정하여 얻을 수 있는 8과 인접하는 채굴 난이도는 7.95 및 8.2가 있을 수 있다. 이 경우에, 상기 체크 매트릭스 결정부(340)는 상기 제 2 체크 매트릭스를 사용 체크 매트릭스로서 결정할 수 있다.

[0125] 물론, 이 경우에 상기 해밍 가중치 결정부(350)는 제 2 체크 매트릭스의 해밍 가중치 조합을 선택하여 채굴 난이도를 7.95로 결정할 수 있다.

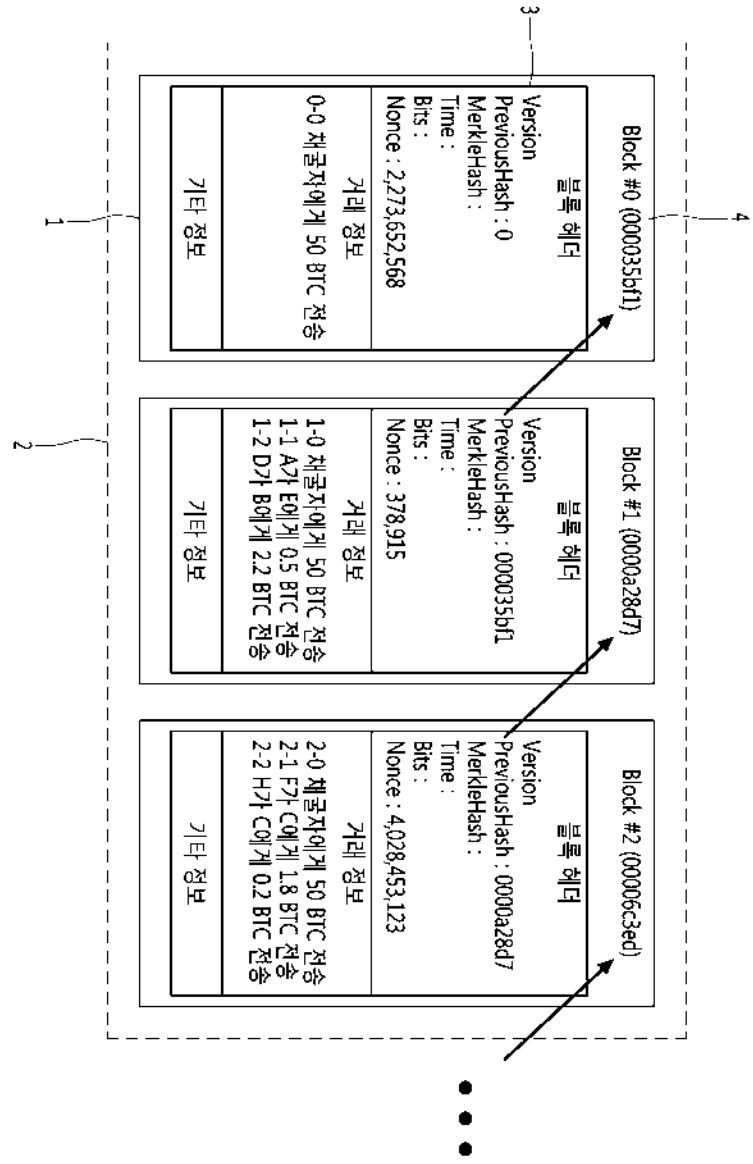
[0126] 위와 같은 방법으로 결정된 체크 매트릭스와 해밍 가중치에 따라서, 도 6과 동일한 과정이 수행될 수 있다.

[0127] 이와 같은 제어에 필요한 정보는 메모리(330)에 미리 저장되어 도출될 수 있다.

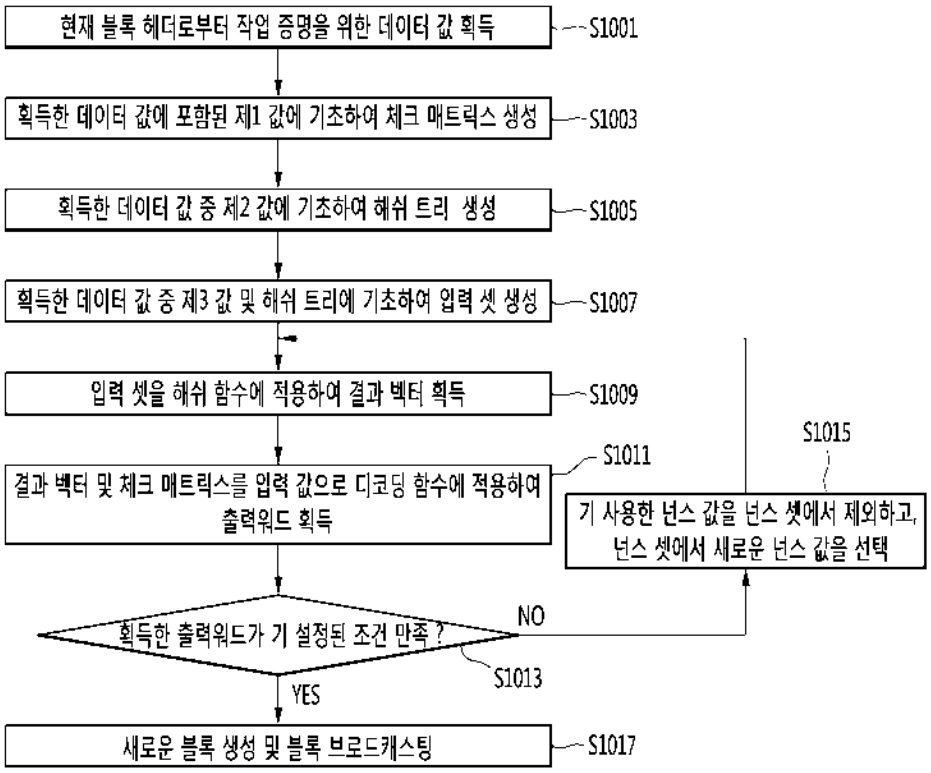
[0128] 전술한 본 발명은, 프로그램이 기록된 매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 매체는, 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 매체의 예로는, HDD(Hard Disk Drive), SSD(Solid State Disk), SDD(Silicon Disk Drive), ROM, RAM, CD-ROM, 자기 테이프, 플로피 디스크, 광 데이터 저장 장치 등이 있다. 따라서, 상기의 상세한 설명은 모든 면에서 제한적으로 해석되어서는 아니되고 예시적인 것으로 고려되어야 한다. 본 발명의 범위는 첨부된 청구항의 합리적 해석에 의해 결정되어야 하고, 본 발명의 등가적 범위 내에서의 모든 변경은 본 발명의 범위에 포함된다.

도면

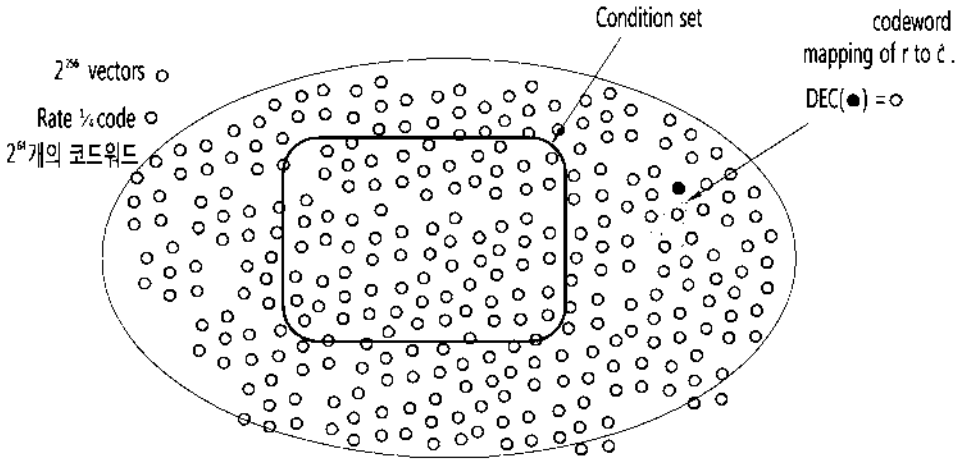
도면1


1
2
3
4
Block #0 (000035bf1)
블록 헤더
Version
PreviousHash : 0
MerkleHash :
Time :
Bits :
Nonce : 2,273,652,568
거래 정보
0-0 채굴자에게 50 BTC 전송
기타 정보
Block #1 (0000a28d7)
블록 헤더
Version
PreviousHash : 000035bf1
MerkleHash :
Time :
Bits :
Nonce : 378,915
거래 정보
1-0 채굴자에게 50 BTC 전송
1-1 A가 E에게 0.5 BTC 전송
1-2 D가 B에게 2.2 BTC 전송
기타 정보
Block #2 (00006c3ed)
블록 헤더
Version
PreviousHash : 0000a28d7
MerkleHash :
Time :
Bits :
Nonce : 4,028,453,123
거래 정보
2-0 채굴자에게 50 BTC 전송
2-1 F가 C에게 1.8 BTC 전송
2-2 H가 C에게 0.2 BTC 전송
기타 정보

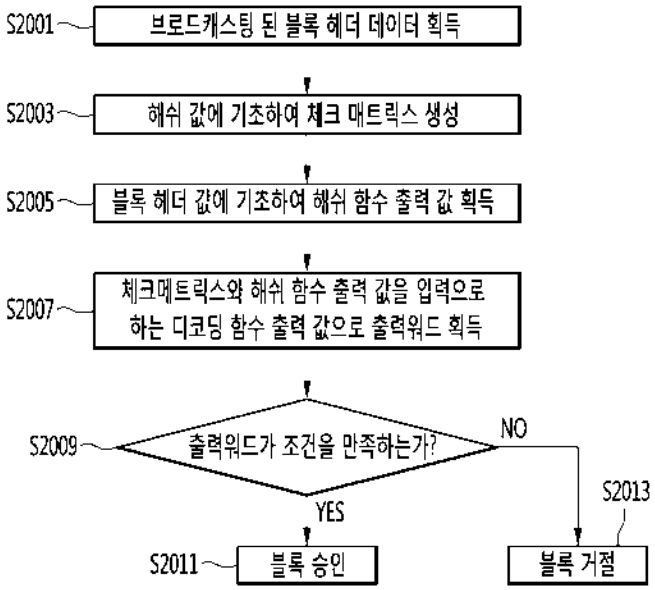
도면2


현재 블록 헤더로부터 작업 증명을 위한 데이터 값 획득
S1001
획득한 데이터 값에 포함된 제1 값에 기초하여 체크 매트릭스 생성
S1003
획득한 데이터 값 중 제2 값에 기초하여 해쉬 트리 생성
S1005
획득한 데이터 값 중 제3 값 및 해쉬 트리에 기초하여 입력 셋 생성
S1007
입력 셋을 해쉬 함수에 적용하여 결과 벡터 획득
S1009
결과 벡터 및 체크 매트릭스를 입력 값으로 디코딩 함수에 적용하여 출력워드 획득
S1011
획득한 출력워드가 기 설정된 조건 만족 ?
S1013
NO
YES
기 사용한 넌스 값을 넌스 셋에서 제외하고, 넌스 셋에서 새로운 넌스 값을 선택
S1015
새로운 블록 생성 및 블록 브로드캐스팅
S1017

도면3


Condition set
codeword
mapping of r to ĉ .
DEC(●) =○
2^256 vectors ○
Rate ½ code ○
2^64개의 코드워드

도면4



도면5

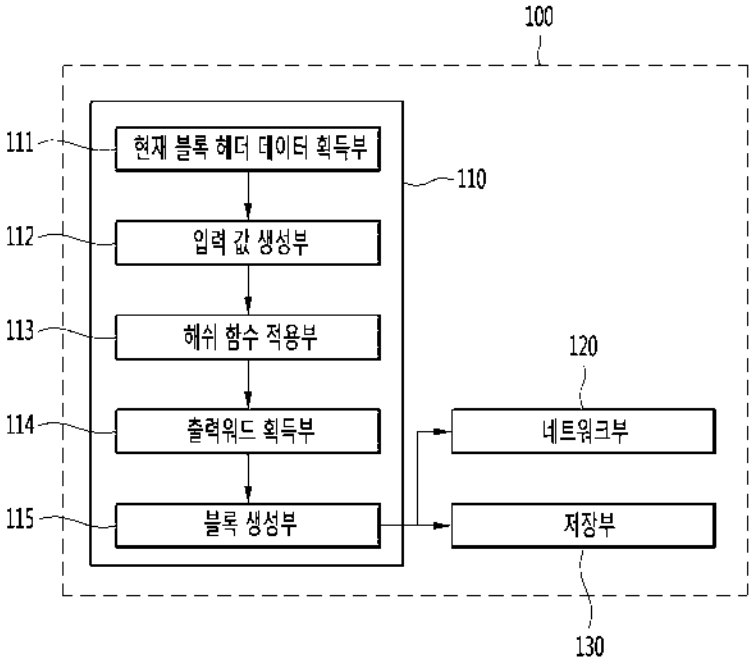
$$F = \begin{bmatrix} 0111001010000000 \\ 0001000001000000 \\ 0000100000100000 \\ 0010000000010000 \\ 0000011000001000 \\ 0000000100000100 \\ 0000100000000010 \\ 0100010100000001 \end{bmatrix}$$

(a)

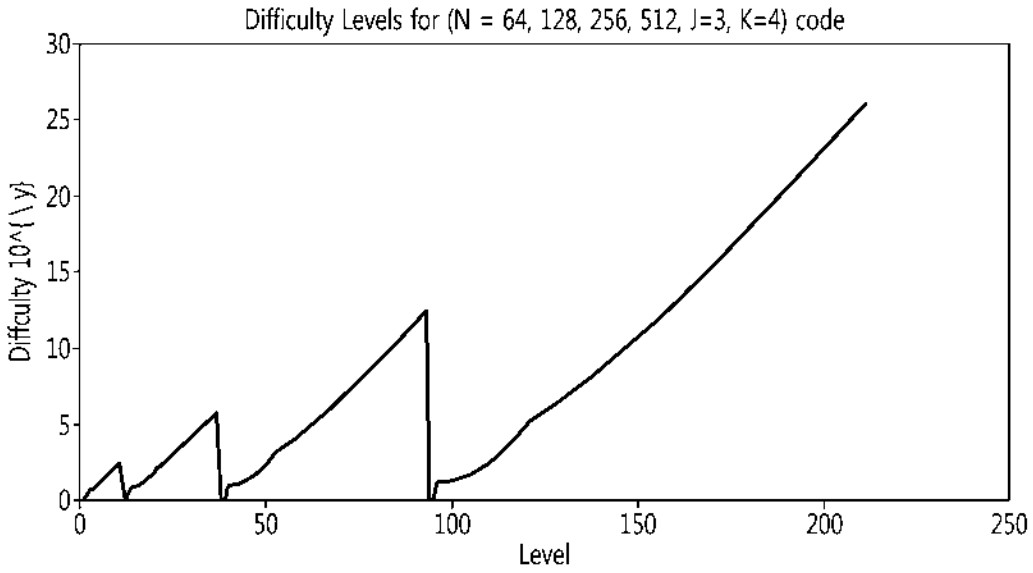
n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\lvert V_n \rvert$	192	0	0	1728	2832	3520	6848	9664	9808	9504	8592	6592	4192	1760	304	0	0
p(%)	0.292969	0	0	2.636719	4.321289	5.371094	10.44922	14.74609	14.96582	14.50195	13.11133	10.05859	6.396484	2.685547	0.463867	0	0

(b)

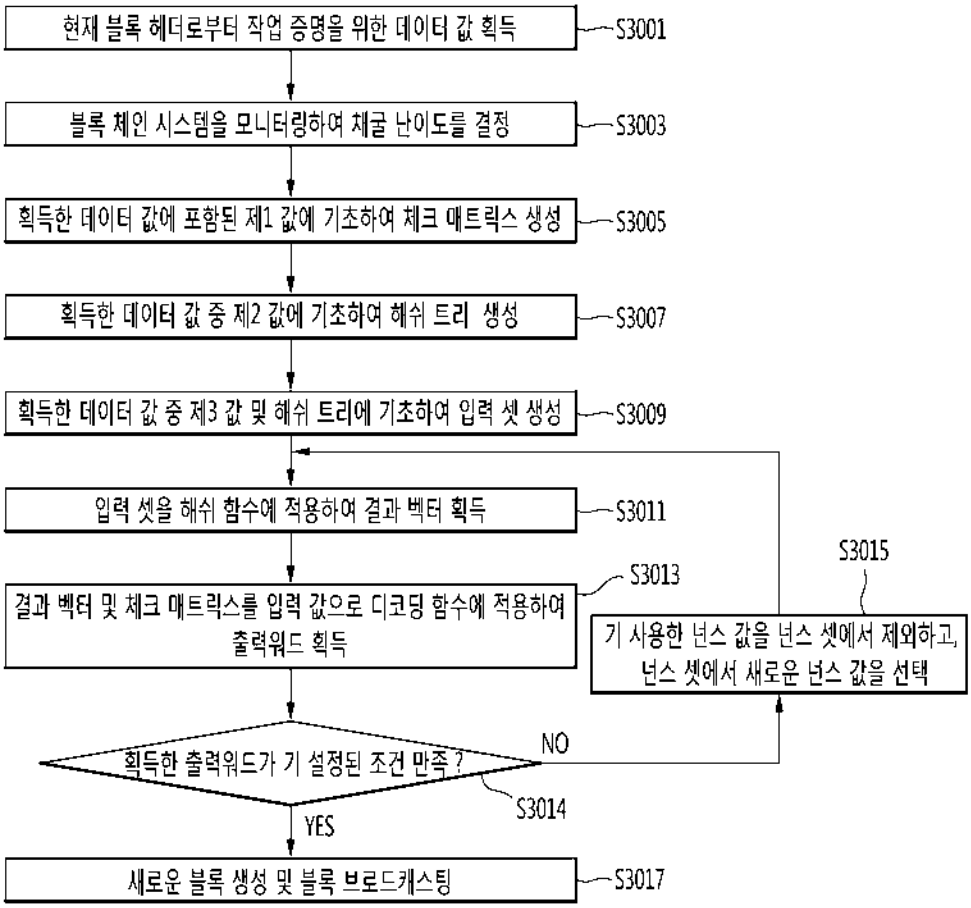
도면6


100
111
현재 블록 헤더 데이터 획득부
110
112
입력 값 생성부
113
해쉬 함수 적용부
120
114
출력워드 획득부
네트워크부
115
블록 생성부
저장부
130

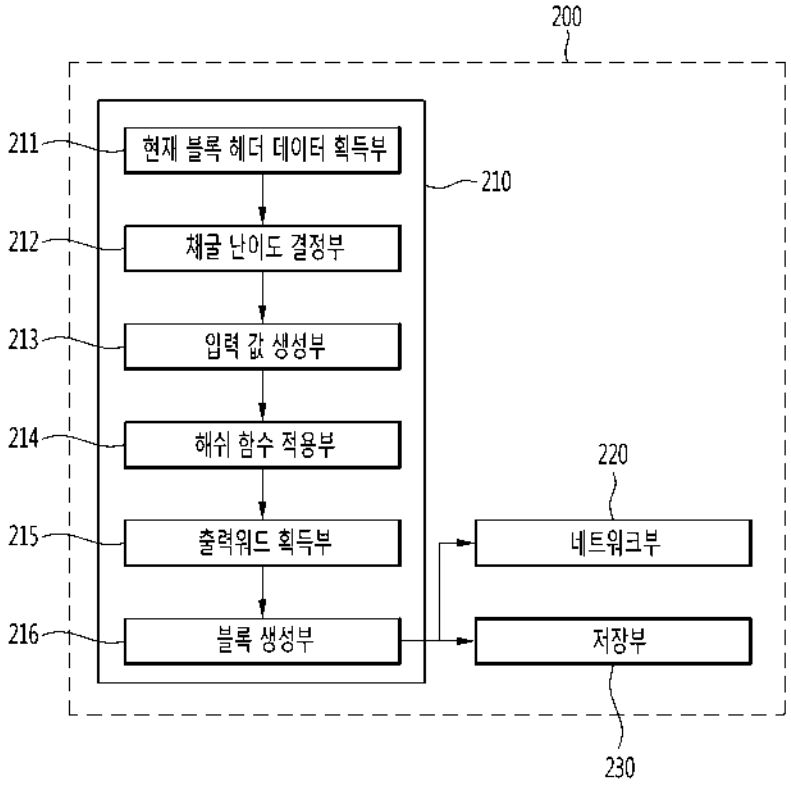
도면7


Difficulty Levels for (N = 64, 128, 256, 512, J=3, K=4) code
Difficulty 10^(-y)
30
25
20
15
10
5
0
0
50
100
150
200
250
Level

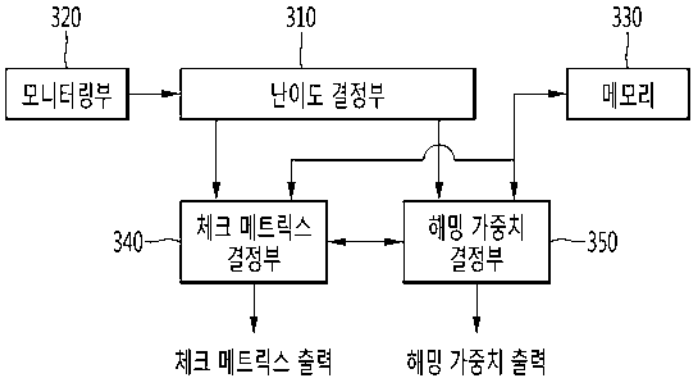
도면8


현재 블록 헤더로부터 작업 증명을 위한 데이터 값 획득
S3001
블록 체인 시스템을 모니터링하여 채굴 난이도를 결정
S3003
획득한 데이터 값에 포함된 제1 값에 기초하여 체크 매트릭스 생성
S3005
획득한 데이터 값 중 제2 값에 기초하여 해쉬 트리 생성
S3007
획득한 데이터 값 중 제3 값 및 해쉬 트리에 기초하여 입력 셋 생성
S3009
입력 셋을 해쉬 함수에 적용하여 결과 벡터 획득
S3011
결과 벡터 및 체크 매트릭스를 입력 값으로 디코딩 함수에 적용하여 출력워드 획득
S3013
획득한 출력워드가 기 설정된 조건 만족 ?
NO
S3014
YES
기 사용한 넌스 값을 넌스 셋에서 제외하고, 넌스 셋에서 새로운 넌스 값을 선택
S3015
새로운 블록 생성 및 블록 브로드캐스팅
S3017

도면9


200
211
현재 블록 헤더 데이터 획득부
210
212
채굴 난이도 결정부
213
입력 값 생성부
214
해쉬 함수 적용부
220
215
출력워드 획득부
네트워크부
216
블록 생성부
저장부
230

도면10


320
310
330
모니터링부
난이도 결정부
메모리
340
체크 메트릭스 결정부
해밍 가중치 결정부
350
체크 메트릭스 출력
해밍 가중치 출력

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0135119
(43) 공개일자 2020년12월02일

(51) 국제특허분류(Int. Cl.)
G06Q 20/06 (2012.01) *G06Q 20/38* (2012.01)
(52) CPC특허분류
G06Q 20/065 (2013.01)
G06Q 20/38 (2020.05)
(21) 출원번호 10-2019-0120655
(22) 출원일자 2019년09월30일
심사청구일자 2019년09월30일
(30) 우선권주장
1020190061493 2019년05월24일 대한민국(KR)

(71) 출원인
광주과학기술원
광주광역시 북구 첨단과기로 123 (오룡동)
(72) 발명자
장재혁
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 전기전자컴퓨터공학부
이홍노
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 전기전자컴퓨터공학부
(74) 대리인
김기문

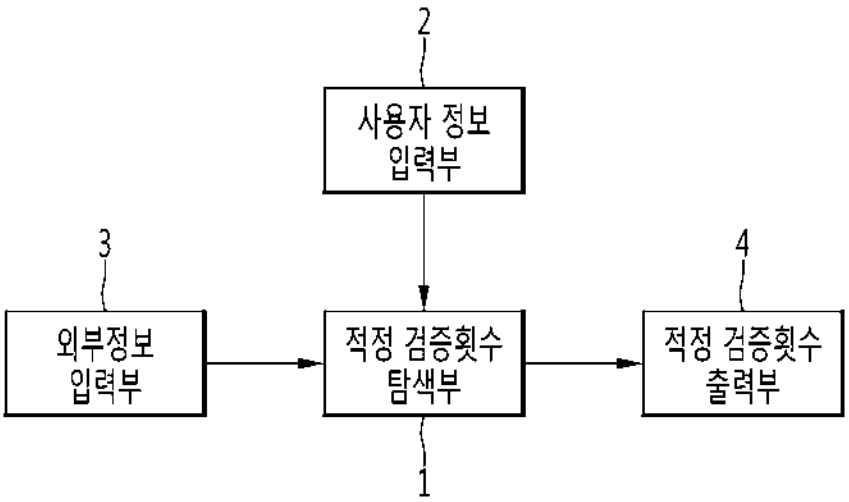
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 **블록체인의 거래검증시스템, 및 블록체인이 거래검증방법**

(57) 요 약

본 발명에 따른 블록체인의 거래검증시스템에는, 사용자 정보가 입력되는 사용자 정보 입력부; 암호화폐시스템과 관련되는 외부정보가 입력되는 외부정보 입력부; 상기 외부정보 및 상기 사용자 정보를 이용하여, 현재 일어나는 현재 거래의 검증에 적정한 블록검증의 수를 탐색하는 적정 검증횟수 탐색부; 및 상기 블록검증의 수를 출력하는 적정 검증횟수 출력부가 포함된다. 본 발명에 따르면, 사용자는 안전하고 신속한 거래를 위하여 필요한 블록검증의 수를 자동으로 알아낼 수 있다.

대 표 도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호	1711093581(NN24010)
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기술진흥센터
연구사업명	융합기술개발
연구과제명	확장가능한 탈중앙화 보안성 ECCPoW 블록체인
기 여 율	1/2
과제수행기관명	광주과학기술원
연구기간	2019.04.01 ~ 2019.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	GK11380
부처명	광주과학기술원
과제관리(전문)기관명	광주과학기술원
연구사업명	실용화연구개발
연구과제명	암호-부호 작업증명(PoW) 비트코인/이더리움 하드포크 1.0개발
기 여 율	1/2
과제수행기관명	광주과학기술원
연구기간	2019.02.01 ~ 2019.12.31

명 세 서

청구범위

청구항 1

사용자 정보가 입력되는 사용자 정보 입력부;

암호화폐시스템과 관련되는 외부정보가 입력되는 외부정보 입력부;

상기 외부정보 및 상기 사용자 정보를 이용하여, 현재 일어나는 현재 거래의 검증에 적정한 블록검증의 수를 탐색하는 적정 검증횟수 탐색부; 및

상기 블록검증의 수를 출력하는 적정 검증횟수 출력부가 포함되는 블록체인의 거래검증시스템.

청구항 2

제 1 항에 있어서,

상기 사용자 정보에는, 적어도 사용자가 입력하는 입력 거래금액이 포함되는 블록체인의 거래검증시스템.

청구항 3

제 1 항에 있어서,

상기 외부정보에는,

-현재 블록 하나를 채굴하면 지급되는 보상금,

-현재 블록 하나를 채굴할 때 소용되는 평균비용, 및

-현재 블록의 채굴속도가,

중의 적어도 하나가 포함되는 블록체인의 거래검증시스템.

청구항 4

제 3 항에 있어서,

상기 외부정보에는,

-블록 하나를 채굴하면 지급되는 보상금,

-현재 블록 하나를 채굴할 때 소용되는 평균비용, 및

-현재 블록의 채굴속도가,

모두 포함되는 블록체인의 거래검증시스템.

청구항 5

제 2 항에 있어서,

상기 적정 검증횟수 탐색부에는,

필요한 정보가 저장되는 메모리;

상기 외부정보를 이용하여, 현재 상태에서 안전한 안전거래 한도액을 추출하는 안전거래 한도액 추출부; 및

상기 안전거래 한도액과 상기 입력 거래금액을 비교하여 안전성을 판별하는 안정성 판별부가 포함되는 블록체인의 거래검증시스템.

청구항 6

제 5 항에 있어서,

상기 안전거래 한도액 추출부에는,

공격자가 기대이윤을 영이 되게 만드는 금액을, 상기 안전거래 한도액으로 추출하는 연산부가 포함되는 블록체인의 거래검증시스템.

청구항 7

제 6 항에 있어서,

상기 공격자의 기대이윤은, 공격자의 기대수익에서 공격자의 기대비용을 감산하여 제공되는 블록체인의 거래검증시스템.

청구항 8

제 6 항에 있어서,

상기 안전거래 한도액(C_{Req})는

$C_{Req.} = (1-P_{AS})/P_{AS}*X(\gamma, t_{cut})+X(\gamma, T_{AS})-R(\beta, T_{AS})$

에 의해서 추출되고,

여기서,

P_{AS}(무차원)는 이중지불공격이 t_{cut} 시간 내에 성공할 확률이고,

$X(\gamma,t)$는 공격자가 t 시간 동안 채굴 할 경우 소요하는 평균 블록 채굴 비용(γ와 t에 대해 증가하는 함수)이고,

γ는 블록 하나를 채굴할 때 소요되는 평균비용이고,

T_{AS}는 공격이 성공하기까지 소요된 시간이고,

β는 블록 하나를 채굴하면 지급되는 보상금이고,

$R(\beta,t)$는 공격자가 t 시간 동안 채굴할 경우 획득하는 평균 블록 채굴 보상금(β와 t에 대해 증가하는 함수)인,

블록체인의 거래검증시스템.

청구항 9

제 5 항에 있어서,

상기 안전거래 한도액 추출부에는 파라미터 최적화부가 포함되고,

상기 파라미터 최적화부에는,

- 공격자의 손절시간, 및

- 공격자의 자원비율 중의 적어도 하나를

최적으로 추측하는 블록체인의 거래검증시스템.

청구항 10

제 5 항에 있어서,

상기 안전성 판별부에 의해서 안전하지 않다고 판별되는 경우에는, 상기 블럭검증의 수를 증가시켜 상기 안전거래 한도액을 다시금 추출하는 블록체인의 거래검증시스템.

청구항 11

암호화폐시스템에 관한 정보인 외부정보 및 사용자가 입력하는 사용자 정보를 입력받는 것;

안전하고 신속한 거래를 위하여 적정한 블록검증의 수인 적정 검증횟수를 탐색하는 것; 및

상기 적정 검증횟수를 출력하는 것이 수행되는 블록체인의 거래검증방법.

청구항 12

제 11 항에 있어서,

상기 적정 검증횟수를 탐색하는 것은,

상기 외부정보 및 임의의 현재 블록검증의 수를 이용하여, 현재 상태에서 안전한 거래금액인 안전거래 한도액을 계산하는 것;

상기 안전거래 한도액과, 상기 사용자 정보로서 입력되는 입력 거래금액을 비교하는 것; 및

상기 안전거래 한도액과 상기 입력 거래금액을 비교하여, 상기 안전거래 한도액이 큰 경우에는 안전한 것으로 판단하고, 상기 현재 블록검증의 수를 출력하는 것이 수행되는 블록체인의 거래검증방법.

청구항 13

제 12 항에 있어서,

상기 안전거래 한도액과 상기 입력 거래금액을 비교하여, 상기 안전거래 한도액이 작은 경우에는 안전하지 않은 것으로 판단하고, 상기 현재 블록검증의 수를 증가시킨 다음에, 상기 안전거래 한도액을 다시 계산하는 것이 수행되는 블록체인의 거래검증방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 블록체인의 거래검증시스템, 및 블록체인의 거래검증방법에 관한 것이다.

배 경 기 술

[0002] 근래, 암호화폐는 상거래에서 화폐의 기능을 수행하는 단계에 이르고 있다. 상기 암호화폐를 통한 거래는 블록검증(block confirmation)을 거쳐서 거래가 완료된다. 상기 블록검증은, 거래자 간의 현재 거래행위가 일어난 이후에, 설정된 블록검증의 수에 상당하는 블록이 생성되는 것을 확인한 다음에, 거래를 완료하는 기술이다. 예를 들어, 상기 블록검증의 수가 6인 경우에는, 거래자 간의 거래가 기록되는 블록 이후에 6개의 블록이 추가로 생성되는 것을 기다린 후에 거래를 완료한다. 상기 거래완료는, 상품을 가진자가 상품을 암호화폐의 제공자에게 발송하는 것을 예시할 수 있다.

[0003] 비특허문헌 1에서는 상기 블록검증의 수가 클수록 거래가 안전하다는 것을 증명하고 있다.

[0004] 상기 블록검증은 이중지불(double spending)을 방어하는 수단으로서 활용될 수 있다. 상기 이중지불은 악의적인 사용자가, 가장 긴 블록체인이 살아남는 블록체인의 원칙을 이용하여, 실제거래의 블록체인보다 긴 악의적인 블록체인을 비밀스럽게 제공하여, 실제거래를 무산시키는 공격행위를 말한다.

[0005] 상기 블록검증의 수가 클수록 이중지불을 당할 확률은 작아진다. 그러나, 상기 블록검증의 수가 클수록 거래완료까지 더 오랜 시간이 소요되는 문제점이 있다. 예를 들어, 비트코인(Bitcoin)의 경우에 평균 블록생성의 주기가 10분이기 때문에, 상기 블록검증의 수가 6일 경우 거래완료시간은 60분이 된다. 이 말은 콜라 한 잔을 사기 위하여 60분을 기다릴 수 있는 것이 된다. 이러한 문제점은 암호화폐의 실제사용에 큰 장애물이 되고 있다.

[0006] 이와 같은 배경 하에서, 현재 암호화폐를 이용하는 거래에서는, 거래자가 자신의 책임하에서 거래금액별로 주어지는 가이드 라인에 따라서, 블록검증의 수를 입력하도록 하는 것이 일반적이다. 거래자는, 자신의 거래에서 블록검증의 수를 얼마나 할지를 알 수가 없고, 검증횟수가 아무리 많아도 이중공격 성공확률이 영이 되지는 않는다.

선행기술문헌

비특허문헌

[0007] (비특허문헌 0001) S Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System”[Online] Available: https://bitcoinorg/bitcoinpdf

발명의 내용

해결하려는 과제

[0008] 본 발명은 현재 조건에 맞추어서 적정의 블록검증의 수를 제안하는 블록체인의 거래검증시스템, 및 거래검증방법을 제안한다.

[0009] 본 발명은 거래자가 입력하는 거래금액뿐만 아니라, 다양한 외부상황에 대응하여 안전하게 신속하게 암호화폐를 통하여 거래를 완료할 수 있는, 블록체인의 거래검증시스템, 및 거래검증방법을 제안한다.

[0010] 본 발명은 암호화폐를 통한 안전한 상거래 및 신속한 거래완료의 두 목적을 함께 달성할 수 있는, 블록체인의 거래검증시스템, 및 거래검증방법을 제안한다.

과제의 해결 수단

[0011] 본 발명에 따른 블록체인의 거래검증시스템에는, 사용자 정보가 입력되는 사용자 정보 입력부; 암호화폐시스템과 관련되는 외부정보가 입력되는 외부정보 입력부; 상기 외부정보 및 상기 사용자 정보를 이용하여, 현재 일어나는 현재 거래의 검증에 적정한 블록검증의 수를 탐색하는 적정 검증횟수 탐색부; 및 상기 블록검증의 수를 출력하는 적정 검증횟수 출력부가 포함된다. 본 발명에 따르면, 사용자는 안전하고 신속한 거래를 위하여 필요한 블록검증의 수를 자동으로 알아낼 수 있다.

[0012] 상기 사용자 정보에는, 적어도 사용자가 입력하는 입력 거래금액이 포함되어, 사용자가 입력하는 입력거래 금액에 대응하여 안전한 거래를 신속하게 수행할 수 있다.

[0013] 상기 외부정보에는, 블록 하나를 채굴하면 지급되는 보상금, 현재 블록 하나를 채굴할 때 소용되는 평균비용, 및 현재 블록의 채굴속도 중의 적어도 하나가 포함되어, 암호화폐시스템을 이용하는 공격자의 공격경향을 파악하여 더 정확한 블록검증의 수를 알아낼 수 있다. 나아가서, 암호화폐시스템의 종류에 따라서 최적의 안전거래 한도액을 알아낼 수 있다.

[0014] 상기 외부정보에는, 블록 하나를 채굴하면 지급되는 보상금, 현재 블록 하나를 채굴할 때 소용되는 평균비용, 및 현재 블록의 채굴속도가 모두 포함되어, 보다 정확한 블록검증의 수를 알아낼 수 있다.

[0015] 상기 적정 검증횟수 탐색부에는, 필요한 정보가 저장되는 메모리; 상기 외부정보를 이용하여, 현재 상태에서 안전한 안전거래 한도액을 추출하는 안전거래 한도액 추출부; 및 상기 안전거래 한도액과 상기 입력 거래금액을 비교하여 안전성을 판별하는 안정성 판별부가 포함된다. 이에 따르면, 현재 상태를 반영하여 실시간으로 변동하는 안전한 금액을 실시간으로 알아낼 수 있다. 나아가서, 암호화폐시스템의 종류에 따라서 최적의 안전거래 한도액을 알아낼 수 있다.

[0016] 상기 안전거래 한도액 추출부에는, 공격자가 기대이윤을 영이 되게 만드는 금액을, 상기 안전거래 한도액으로 추출하는 연산부가 포함된다. 이에 따르면, 공격자의 이익에 기반하기 때문에, 공격자의 공격경향을 예측하기 때문에, 더 정확한 블록검증의 수를 알아낼 수 있다.

[0017] 여기서, 상기 공격자의 기대이윤은, 공격자의 기대수익에서 공격자의 기대비용을 감산하여 제공될 수 있다.

[0018] 여기서, 상기 안전거래 한도액(C_{Req})는, $C_{Req.} = (1-P_{AS})/P_{AS}*X(\gamma, t_{cut})+X(\gamma, T_{AS})-R(\beta, T_{AS})$ 에 의해서 추출될 수 있다.

[0019] 상기 안전거래 한도액 추출부에는 파라미터 최적화부가 포함되고, 상기 파라미터 최적화부에는, 공격자의 손절시간, 및 공격자의 자원비율 중의 적어도 하나를 최적으로 추측할 수 있다. 이에 다르면, 공격자의 선택 파라미터를 최적으로 예상할 수 있어서, 더 안전한 블록검증의 수를 제안할 수 있다.

[0020] 상기 안전성 판별부에 의해서 안전하지 않다고 판별되는 경우에는, 상기 블럭검증의 수를 증가시켜 상기 안전거래 한도액을 다시금 추출하도록 하는 과정이 반복될 수 있다. 이에 의해서 사용자는 가장 안전하면서 최소한 블록검증의 수를 알아낼 수 있다.

[0021] 본 발명에 따른 블록체인이 거래검증방법에는, 암호화폐시스템에 관한 정보인 외부정보 및 사용자가 입력하는 사용자 정보를 입력받는 것; 안전하고 신속한 거래를 위하여 적정한 블록검증의 수인 적정 검증횟수를 탐색하는 것; 및 상기 적정 검증횟수를 출력하는 것이 수행된다. 사용자는 안전하고 신속하게 블록체인을 통한 거래를 수행할 수 있다.

[0022] 상기 적정 검증횟수를 탐색하는 것은, 상기 외부정보 및 임의의 현재 블록검증의 수를 이용하여, 현재 상태에서 안전한 거래금액인 안전거래 한도액을 계산하는 것; 상기 안전거래 한도액과, 상기 사용자 정보로서 입력되는 입력 거래금액을 비교하는 것; 및 상기 안전거래 한도액과 상기 입력 거래금액을 비교하여, 상기 안전거래 한도액이 큰 경우에는 안전한 것으로 판단하고, 상기 현재 블록검증의 수를 출력하는 것이 수행된다. 이에 따르면, 사용자는 안전한 블록검증의 수를 알아낼 수 있거나, 실제로 블록검증의 수가 적용됨으로써, 안전하게 거래가 일어날 수 있다.

[0023] 상기 안전거래 한도액과 상기 입력 거래금액을 비교하여, 상기 안전거래 한도액이 작은 경우에는 안전하지 않은 것으로 판단하고, 상기 현재 블록검증의 수를 증가시킨 다음에, 상기 안전거래 한도액을 다시 계산하는 것이 수행된다. 이와 같은 반복계산에 의해서 실제로 가장 안전하면서 최소한의 블록검증의 수를 알아낼 수 있다.

발명의 효과

[0024] 본 발명에 따르면, 거래자는, 자신이 거래금액만을 입력하는 것으로써, 적정한 블록검증의 수를 알 수 있다. 여기서 적정은, 안전거래 및 신속거래의 두 목적을 합목적적으로 달성할 수 있는 것을 의미한다. 본 문서에서 적정의 의미는 위와 마찬가지로 이해할 수 있다.

[0025] 본 발명에 따르면, 현재 당사자 간의 거래만이 아니라, 당사자 간 거래의 외부에 놓이는 암호화폐전체 시스템의 상황정보(외부정보라고 할 수 있다)을 참조하여, 적정의 블록검증의 수를 제안할 수 있다. 따라서, 공격자의 선택을 예측하여 더 정확하게 상기 블록검증의 수를 제안할 수 있다.

[0026] 그 외의 상세한 구성에 따른 본 발명의 효과는 발명을 실시하기 위한 구체적인 내용에 더 상세하게 제시된다.

도면의 간단한 설명

[0027] 도 1은 실시예에 따른 블록체인의 거래검증시스템을 보이는 도면.

도 2는 적정 검증횟수 탐색부를 상세하게 보이는 도면.

도 3은 적정 검증횟수 탐색부의 상세한 작용을 설명하는 블록체인 거래검증방법의 일부를 설명하는 흐름도.

도 4는 안전거래 한도액 추출부의 자세한 구성을 보이는 도면.

도 5는 공격자가 점유하는 자원비율(p_A)을 변수로 하는 안전거래 한도액의 그래프의 예시도.

도 6은 공격자가 점유하는 자원비율(p_A)을 변수로 손절시간이 무한대일 때 공격성공을 위한 평균소요시간($\mathbb{E}_{T_{AS-ICT}}(p_A;N_{BC})$)의 그래프의 예시도.

도 7은 손절시간을 변수로 안전거래 한도액을 그래프로 예시하는 도면.

발명을 실시하기 위한 구체적인 내용

[0028] 이하에서는 도면을 참조하여 본 발명의 구체적인 실시예를 상세하게 설명한다. 다만, 본 발명의 사상은 이하에 제시되는 실시예에 제한되지 아니하고, 본 발명의 사상을 이해하는 당업자는 동일한 사상의 범위 내에 포함되는 다른 실시예를 구성요소의 부가, 변경, 삭제, 및 추가 등에 의해서 용이하게 제안할 수 있을 것이나, 이 또한 본 발명 사상의 범위 내에 포함된다고 할 것이다.

[0029] 도 1은 실시예에 따른 블록체인의 거래검증시스템을 보이는 도면이다.

[0030] 도 1을 참조하면, 거래검증에 필요한 검증횟수를 탐색하는 적정 검증횟수 탐색부(1), 시스템의 사용자가 입력하

는 정보를 받아 상기 적정 검증횟수 탐색부(1)로 출력하는 사용자 정보 입력부(2), 암호화폐시스템의 상황정보, 즉 외부정보를 입력받아 상기 적정 검증횟수 탐색부(1)로 출력하는 외부정보 입력부(3)가 포함된다. 상기 적정 검증횟수 탐색부(1)는 적정한 블록검증의 수를 탐색하여 적정 검증횟수 출력부(4)로 출력할 수 있다.

[0031] 상기 사용자 정보 입력부(2)는 디스플레이와 입력장치를 가지는 컴퓨터 등의 입력수단이 될 수 있다. 상기 외부정보 입력부(3)는 네트워크에 연결되는 통신수단, 및 네트워크 상의 공개정보를 독출하고 저장하는 컴퓨터가 될 수 있다. 상기 적정 검증횟수 출력부(4)는 디스플레이 또는 통신수단을 가지는 컴퓨터 등의 출력수단이 될 수 있다. 상기 적정 검증횟수 탐색부(1)는 연산장치와 소정의 메모리를 가지는 컴퓨터일 수 있다.

[0032] 상기 외부정보에는, 현재 블록 하나를 채굴하면 지급되는 보상금, 현재 블록 하나를 채굴할 때 소용되는 평균비용, 및 현재 블록의 채굴속도가 포함될 수 있다.

[0033] 상기 사용자 정보에는, 사용자가 현재의 거래를 위하여 입력한 거래금액과 관련되는 정보가 포함될 수 있다.

[0034] 시스템의 사용자는 거래자라고 할 수 있다.

[0035] 실시예의 블록체인의 거래검증시스템은, 상기 사용자 정보 입력부(2)로부터 입력되는 거래금액정보, 및 상기 외부정보 입력부(3)로부터 입력되는 상기 외부정보를 이용하여, 상기 적정 검증횟수 탐색부(1)가 적정한 블록검증의 수를 탐색한다. 탐색된 정보는 적정 검증횟수 출력부(4)를 통하여 출력될 수 있다.

[0036] 사용자는, 상기 적정 검증횟수 출력부(4)에서 출력되는 블록검증의 수를 이용하여 블록검증의 수를 스스로 설정할 수 있다. 블록검증의 수에 따라서 블록검증이 수행된 다음에, 거래가 완료될 수 있다.

[0037] 사용자에 의한 상기 블록검증의 수 설정이 없이, 상기 적정 검증횟수 출력부(4)에서 출력되는 블록검증의 수가 자동으로 적용되어, 블록체인의 거래검증시스템이 운용될 수도 있다.

[0038] 도 1에 제시되는 블록체인의 거래검증시스템은 거래자의 단말기에 탑재될 수 있다.

[0039] 도 1에 제시되는 어느 단위 블럭은 거래자의 단말에, 다른 단위 블럭은 다른 거래자 및 거래소의 단말 중의 적어도 하나에 제공되어, 네트워크를 통하여 연결된 상태로, 블록체인의 거래검증시스템이 동작될 수도 있다.

[0040] 도 1에는 도시되지 않았지만, 상기 블록체인의 거래검증시스템에는, 암호화폐시스템의 임의의 다른 사용자가 블록채굴을 통하여 블록을 체인화하는 네트워크 노드가 포함되는 것은 당연히 이해할 수 있다. 즉, 네트워크 상의 적어도 하나의 다른 노드의 단말기가 협업하여, 현재의 거래를 완료하기 위하여 필요한 블록검증의 수 만큼의 블록을 체인화하여 거래검증을 완료하는 것이 당연히 포함될 수 있다.

[0041] 실시예에 따른 블록체인의 거래검증방법은, 상기 외부정보 및 상기 사용자 정보를 입력받은 다음에, 상기 적정 검증횟수를 탐색한다. 이후에는 탐색된 상기 적정 검증횟수를 출력하는 것에 의해서 수행될 수 있다

[0042] 상기 적정 검증횟수의 출력은, 사용자에게 적정 블록검증의 수가 출력된 다음에, 사용자가 블록검증의 수를 입력하는 것에 의해서 수행될 수 있다. 다른 방법으로, 자동으로 상기 적정 블록검증의 수에 따라서 자동으로 블록검증이 수행되도록 할 수도 있다.

[0043] 도 2는 상기 적정 검증횟수 탐색부를 상세하게 보이는 도면이다.

[0044] 도 2를 참조하면, 상기 적정 검증횟수 탐색부(1)에는, 현재 상태에서 안전한 수준의 검증금액을 추출하는 안전거래 한도액 추출부(10), 상기 안전거래 한도액 추출부(10)에서 추출된 안전거래 한도액과 사용자가 입력한 거래금액을 비교하여 안전성을 판별하는 안전성 판별부(12), 및 상기 안전성 판별부(12) 및 상기 안전거래 한도액 추출부(10)의 동작에 필요한 정보가 저장되는 메모리(11)가 포함될 수 있다.

[0045] 도 3은 상기 적정 검증횟수 탐색부의 상세한 작용을 설명하는 블록체인 거래검증방법의 일부를 설명하는 흐름도이다.

[0046] 도 2 및 도 3을 참조하면, 상기 외부정보 및 상기 사용자 정보가 입력된다(S1). 상기 외부정보에는, 현재 블록 하나를 채굴하면 지급되는 보상금, 현재 블록 하나를 채굴할 때 소용되는 평균비용, 및 현재 블록의 채굴속도가 포함될 수 있다. 상기 사용자 정보에는 사용자가 입력하는 거래금액 정보가 포함될 수 있다. 상기 거래금액에는 각 나라에서 사용되는 화폐 또는 암호화폐의 액수 정보일 수 있다. 예를 들어, 원화 액수 또는 BTC 액수가 정보로서 포함될 수 있다.

[0047] 입력된 정보를 이용하여 안전거래 한도액을 계산한다(S2). 상기 안전거래 한도액은, 상기 외부정보에 의해서 예측되는, 공격자의 비용과 이익을 비교하는 것에 의해서 계산될 수 있다. 상기 안전거래 한도액의 계산은 이후에

더 상세하게 설명한다.

[0048] 상기 안전거래 한도액 계산단계(S2)에서 계산된 상기 안전거래 한도액과, 사용자 정보로서 입력된 입력 거래금액을 비교하여, 현재의 거래가 안전한 지의 여부를 판단한다(S3).

[0049] 상기 안전성 판별 단계(S3)에서, 상기 입력 거래금액과 상기 안전거래 한도액을 비교하여, 상기 입력 거래금액이 작은 경우에는 안전한 것으로 판단하여 현재 검증횟수를 출력한다(S4).

[0050] 반대로 상기 안전성 판별 단계(S3)에서, 상기 입력 거래금액이 큰 경우에는 안전하지 않은 것으로 판단하여, 검증횟수를 증가시키고(S5) 안전거래 한도액 계산단계(S2)를 다시 수행한다. 예를 들어, 현재의 검증횟수가 1인 경우에는 검증횟수를 2로 늘릴 수 있다. 이 후에는 검증횟수가 2가 되고, 현재 시점의 상기 사용자 정보 및 상기 외부정보를 이용하여, 상기 안전거래 한도액을 다시 계산할 수 있다.

[0051] 다시 수행되는 안전거래 한도액 계산단계(S2)에서는 검증횟수가 증가되었으므로, 상기 안전거래 한도액은 올라갈 것으로 예상할 수 있다.

[0052] 상기 안전거래 한도액 계산단계(S2), 및 상기 안전거래 한도액 추출부(10)의 작용은 이하에 상세하게 설명될 것이다.

[0053] 상기 안전거래 한도액은, 공격자가 이중지불공격(본 문서에서 공격이라고 하는 경우에 특별한 설명이 없는 경우에는 이중지불공격을 의미한다)을 하기 위하여 소용되는 비용과, 공격자가 이중지불공격에 성공하는 경우에 얻는 수익을 비교하는 것에 의해서 계산될 수 있다.

[0054] 예를 들어, 공격자가 공격을 하기 위하여 드는 비용에 비하여, 공격자가 공격에 성공하는 경우에 얻는 수익이 작은 경우에는, 공격자가 공격을 하지 않거나, 공격을 하더라도 성공하지 못할 것이기 때문이다.

[0055] 다시 설명하면, '공격자 기대 이윤 = 공격자의 기대 수익 - 공격자의 기대비용'에 의해서 공격자의 기대이윤이 있는 경우에는 공격자가 공격을 할 것이고, 반대로 공격자 기대이윤이 없는 경우에는 공격자가 공격을 하지 않을 것을 예상할 수 있다. 여기서, 상기 공격자 기대이윤이 영이 되는 상태의 거래금액을 상기 안전거래 한도액이라고 정의할 수 있다.

[0056] 상기 공격자의 기대수익은, 공격에 성공 할 경우의 수익(C+R)과, 공격에 실패 할 경우의 수익(영)의 합으로 주어질 수 있다. 간단하게, $P_{AS}*(C+R(\beta,T_{AS}))$로 표시할 수 있다.

[0057] 상기 공격자의 기대비용은, 공격에 성공 할 경우의 일정시간(T_{AS})동안 소요한 비용과, 공격에 실패 할 경우의 손절시간(t_{cut})시간동안 소요한 비용의 합으로 주어질 수 있다. 간단하게, $P_{AS}*X(\gamma,E[T_{AS}]) + (1-P_{AS})*X(\gamma,t_{cut})$로 주어질 수 있다. 여기서 손절시간은 손절시간이 지난 이후에는 공격에 성공할 확률이 희박하다고 판단하고 더 큰 손해를 방지하기 위해 어느 정도의 손해를 입고 공격을 그만두는 시간을 의미할 수 있다.

[0058] 결국, 상기 공격자의 기대이윤은, $P_{AS}*(C+R(\beta,E[T_{AS}])-X(\gamma,E[T_{AS}]))-(1-P_{AS})*X(\gamma,t_{cut})$로 계산될 수 있다.

[0059] 상기 안전거래 한도액C_{Req})은, 공격자의 기대 이윤이 영이 되는 거래금액으로 정의할 수 있다. 따라서, 상기 공격자의 기대이윤이 영이 되도록 거래금액(C)을 좌변에 두고 우변에 나머지 항을 정리하면, $C_{Req.} = (1-P_{AS})/P_{AS}*X(\gamma,t_{cut})+X(\gamma,T_{AS})-R(\beta,T_{AS})$와 같이 정리될 수 있다.

[0060] 상기되는 설명에 제시되는 각 기호, 및 이하의 안전거래 한도액의 실제 계산에 사용되는 다양한 기호의 의미를 표 1로 정리한다.

표 1

[0061]

기호(단위)	설명	비고
C($, BTC)	이중지불 공격대상 거래의 거래금액	사용자입력값
β($, BTC)	블록 하나를 채굴하면 지급되는 보상금	외부입력값
γ($, BTC)	블록 하나를 채굴할 때 소요되는 평균비용	외부입력값
λ_A(No./T)	공격자의 평균 블록채굴 속도(공격자에 따라서 다르고 알려져 있는 것 중의 최고를 이용할 수 있다)(pool에 가입해서 채굴을 하고 있고, 풀 간에 담합할 수 있다)	외부입력값
λ_H(No./T)	일반인의 채굴속도	외부입력값

λ_T(No./T)	전체 채굴속도	외부입력값
t_{cut}(T)	무한정 커질 수 있는 손실을 방지하기 위한 공격제한 시간, 손절시간, cut-time	추측검증값
p_A(%)	공격자가 점유한 계산 자원의 비율	추측검증값
N_{BC}(No.)	거래자가 거래를 완료하기 전 거래를 검증하는 블록검증의 수(block confirmation number)	출력값
E[T_{AS}](T)	공격이 성공하기까지 소요된 평균 시간, t_{cut}보다 작음.	계산값
P_{AS}(무차원)	이중지불공격이 t_{cut} 시간 내에 성공할 확률	계산값
R(β,t)($, BTC)	공격자가 시간 t동안 채굴 할 경우 획득하는 평균 블록 채굴 보상금(β와 t에 대해 증가하는 함수)	계산값
X(γ,t)($, BTC)	공격자가 시간 t동안 채굴 할 경우 소요하는 평균 블록 채굴 비용(γ와 t에 대해 증가하는 함수)	계산값

[0062] 상기 비고란에는 각 기호의 값이 도출되는 방법을 나타내는 것이다.

[0063] 사용자 입력값은, 사용자가 입력하는 사용자 정보로서 여기서는 거래금액을 의미한다.

[0064] 외부입력값은, 각 암호화폐의 시스템에서 공시되고 실시간으로 변하는 값이다. 예를 들어, 상기 외부입력값은, 비트코인의 경우에는 btc.com에, 이더리움의 경우에는 etherscan.io에 공시된다.

[0065] 상기 추측검증값은, 공격자가 선택할 수 있는 값으로서 안전거래 한도액을 계산하는 중에 가장 안전한 값으로 선택될 수 있다. 상기 추측검증값은 추후에 설명될 파리미터 최적화부(16)에 의해서 최적으로 추측될 수 있다.

[0066] 상기 출력값은 본 실시예에서 일 목적으로 출력하는 값으로서, 안전한 검증횟수로서, 블록검증의 수를 나타낸다.

[0067] 상기 계산값은 상기 안전거래 한도액의 계산 중간에 사용되는 값이다.

[0068] 상기 안전거래 한도액을 계산하는 과정, 및 현재 상태에서 안전한 금액을 검출하는 안전거래 한도액 검출부(10)의 구성을 더 상세하게 설명한다.

[0069] 도 4는 상기 안전거래 한도액 추출부의 자세한 구성을 보이는 도면이다.

[0070] 도 4를 참조하면, 상기 안전거래 한도액 추출부(10)에는, 안전거래 한도액을 계산하는 연산부(15), 및 상기 연산부(15)의 연산에 필요한 파라미터를 최적화하는 파라미터 최적화부(16)가 포함된다.

[0071] 상기 파라미터 최적화부(16)는 상기 추측검증값을 최적화하는 블럭으로서, 공격자의 입장에서 공격을 예측할 수 있는 값이다.

[0072] 상기 안전거래 한도액 추출부(10)의 상세한 작용을 다수의 수학식을 참조하여 상세하게 설명한다. 이하의 설명에 있어서 상기 파라미터 최적화부(16)의 작용에 대한 언급이 없는 부분은 모두 상기 연산부(15)에서 수행하는 것으로 이해할 수 있다.

[0073] 하기 수학식 1은 채굴비용/시간의 단위를 채굴비용/블록의 단위로 변환하는 것을 예시한다.

수학식 1

[0074]
$$\gamma = 4.63 \cdot 10^{-22}\ [\mathrm{BTC/hash}] \times 456\mathrm{E}[\mathrm{hashes/block\ mining}] \approx 0.21\ [\mathrm{BTC/block\ mining}].$$

[0075] 상기 수학식 1을 참조하면, 상기 외부정보로서 알려져 있는 채굴비용/해쉬와 해쉬수/블록을 곱하는 것으로서, 블록당 채굴비용을 얻어낼 수 있다. 수학식 1의 결과는 메모리에 저장될 수 있다.

[0076] 하기 수학식 2는 공격자의 채굴속도를 알아내는 것을 예시한다.

수학식 2

$$p_A := \Pr(S_i = n+1 \mid S_{i-1} = n) = \frac{\lambda_A}{\lambda_T},$$

$$p_H := \Pr(S_i = n-1 \mid S_{i-1} = n) = \frac{\lambda_H}{\lambda_T},$$

[0077]

[0078] 상기 수학식 2에서 람다는 속도이고, H(honest)는 일반인이고, T(Total)는 전체를 의미하고, A(Attacker)는 공격자를 의미하고, p_A는 공격자가 가지는 자원의 비율이고, p_H는 일반인이 가지는 자원의 비율이다. p_H에서 소문자 p는 proportion의 두문자로 이해할 수 있다.

[0079] 상기 수학식 2를 통하여 공격자의 채굴속도(λ_T)를 알아낼 수 있다. 상기 수학식 2의 결과는 상기 메모리(11)에 저장될 수 있다. 상기 메모리(11)에는 상기 채굴비용 및 상기 채굴속도뿐만 아니라, 연산에 필요한 정보가 필요에 따라서 저장될 수 있다.

[0080] 하기 수학식 3은 손절시간이 무한대로 공격할 경우에 공격에 성공할 확률로서, 비특허문헌에 인용되는 사토시 나카모토의 논문에 의해서 공지되어 있다.

수학식 3

$$\mathbb{P}_{AS-ICT}(p_A; N_{BC}) = \begin{cases} 1, & p_H \le p_A, \\ 1 - {p_A}^{N_{BC}+1} {p_H}^{N_{BC}} \sum_{j=N_{BC}}^{2N_{BC}} \binom{j-1}{N_{BC}-1} A_j, & p_H > p_A, \end{cases}$$

where

$$A_j \triangleq {p_A}^{j-2N_{BC}-1} - {p_H}^{j-2N_{BC}-1}.$$

[0081]

[0082] 상기 수학식 3에서 P는 확률(probality)를 의미하고, AS(attack success)는 공격성공을 의미하고, ICT(infinite cut time)으로서 손절시간이 무한한 것을 의미하고, N_{BC}는 검증을 위한 블록검증의 수를 의미한다.

[0083] 하기 수학식 4는 상기 수학식 3에 의해서 도출되는 식으로서, 주어진 시간 내에 공격에 성공할 확률에 관한 확률분포를 나타낸다.

수학식 4

$$f_{T^{(1),(2)}}(t)=\frac{p_A\lambda_T e^{-\lambda_T t}\left(p_A p_H(\lambda_T t)^2\right)^{N_{BC}}}{(2N_{BC})!}\cdot\sum_{j=N_{BC}}^{j=2N_{BC}}\binom{j-1}{N_{BC}-1}\,{}_2F_3\left(\mathbf{a};\mathbf{b};p_A p_H(\lambda_T t)^2\right)+\frac{e^{-\lambda_T t}}{t}\frac{(p_H\lambda_T t)^{N_{BC}}}{(N_{BC}-1)!}\left(e^{p_A\lambda_T t}-\sum_{i=0}^{N_{BC}}\frac{(p_A\lambda_T t)^i}{i!}\right)+(1-\mathbb{P}_{AS-ICT})\delta(t-\infty),$$

[0084]

[0085] 상기 수학식 4는 여기서 ${}_pF_q$는 [G Gasper and M Rahman, "Basic Hypergeometric series," in *Basic hypergeometric series*, Second, vol 96, Cambridge University Press, Cambridge, 2004]에 정의되는 일반화 초기하학적 함수(generalized hypergeometric function)이고, a, b는 수학식 5에 정의된다.

수학식 5

$$\mathbf{a}=\begin{bmatrix}N_{BC}+1-j/2\\ N_{BC}+1/2-j/2\end{bmatrix}$$

$$\mathbf{b}=\begin{bmatrix}2N_{BC}+2-j\\ N_{BC}+1\\ N_{BC}+1/2\end{bmatrix}.$$

[0086]

[0087] 상기 수학식 4의 확률분포를 제 1 적분하여, 하기 수학식 6의 손절시간 내에 공격에 성공할 확률을 알아낼 수 있다.

수학식 6

$$\mathbb{P}_{AS}(p_A,t_{cut};N_{BC}):=\Pr\left(T^{(1),(2)}<t_{cut}\right)$$

[0088]

[0089] 상기 수학식 5의 확률분포를 제 2 적분하여, 하기 수학식 7의 손절시간 내에 공격에 성공하는데 소요되는 평균 시간을 알아낼 수 있다.

수학식 7

$$\mathbb{E}_{T_{AS}}\left(p_A, t_{cut}; N_{BC}\right) = \frac{\int_0^{t_{cut}} t f_{T^{(1),(2)}}(t)\,dt}{\mathbb{P}_{AS}\left(p_A, t_{cut}; N_{BC}\right)}$$

[0090]

[0091] 여기서, E(expectation)은 기대값이고, P(probality)는 확률이고, f는 확률론에서의 확률분포이다.

[0092] 상기 수학식 6과 수학식 7을 이용하여, 공격자의 평균공격비용을 수학식 8과 같이 알아낼 수 있다.

수학식 8

$$\begin{aligned}\mathbb{E}_X\left(p_A, t_{cut}; N_{BC}\right) := \\ \mathbb{P}_{AS}\left(p_A, t_{cut}; N_{BC}\right)\mathbb{E}\left[X\left(\lambda_A, T_{AS}\right)\right] \\ +\left(1-\mathbb{P}_{AS}\left(p_A, t_{cut}; N_{BC}\right)\right)X\left(\lambda_A, t_{cut}\right)\end{aligned}$$

[0093]

[0094] 여기서, X는 채굴비용을 나타낸다. 상기 채굴비용 X를 선형함수로 제공하는 것은, 채굴장비를 인터넷 상의 대여소(예를 들어, nicehach.com)를 통하여 대여하는 비용을 선형함수로 예시한 것이다. 만약, 상기 채굴비용이 대여비용이 아닌 장비의 전기사용료 등이 되는 경우와 같이 다양한 경우에 대응하여, 상기 비용함수 X는, n-제곱함수, n-루트함수, 로그함수, 지수함수, 또는 n-루트 함수가 될 수도 있을 것이다.

[0095] 한편, 수학식 3을 이용하여, 손절시간이 무한대일 때 공격성공을 위한 평균소요시간($\mathbb{E}_{T_{AS-ICT}}\left(p_A; N_{BC}\right)$)을 수학식 9와 같이 알아낼 수 있다.

수학식 9

$$\mathbb{E}_{T_{AS-ICT}}\left(p_A; N_{BC}\right) = \frac{\lambda_T^{-1}\left(\sum_{j=N_{BC}}^{2N_{BC}}\binom{j-1}{N_{BC}-1}Z_j + \frac{N_{BC}}{p_H}\right)}{\mathbb{P}_{AS-ICT}\left(p_A; N_{BC}\right)};$$

where

$$Z_j := p_A p_m^{N_{BC}} p_M^{-(N_{BC}-j+1)}\left(\frac{2N_{BC}-2jp_m+1}{p_M-p_m}\right) - jp_A^{-(N_{BC}-j)} p_H^{N_{BC}}.$$

[0096]

[0097] 상기 수학식 9를 이용하여, 공격자가 점유하는 자원을 예측하여 가장 안정된 검증금액을 제시하는, 공격자가 점유하는 자원비율(p_A)을 알아낼 수 있다.

[0098] 구체적으로, 상기 수학식 9를 참조하면, 손절시간이 무한대일 때 공격성공을 위한 평균소요시간($\mathbb{E}_{T_{AS-ICT}}\left(p_A; N_{BC}\right)$)은, 공격자가 점유하는 자원비율($p_A$)을 인자로 하는 함수이다. 이 경우에, 공격자가 점유

하는 자원비율(p_A)을 변수로 손절시간이 무한대일 때 공격성공을 위한 평균소요시간($\mathbb{E}_{T_{AS-ICT}}(p_A;N_{BC})$)을 계산하면, 도 6과 같이 아래로 볼록한 함수로 나타난다.

[0099] 따라서, 도 6의 그래프 상에서, 가장 낮은 곳의 공격자가 점유하는 자원비율(p_A)을 취하여, 공격자가 점유하는 자원비율을 가장 안정적으로 최적화할 수 있다. 공격자는 가장 최소의 비용으로 빠른 공격 효과를 얻어낼 수 있는 점유 자원비율을 선택할 것이기 때문이다.

[0100] 이때 상기 공격자가 점유하는 자원비율에 대한 최적화는 상기 파라미터 최적화부(16)에 의해서 연산부(15)와는 별개의 작용을 통하여 수행될 수 있다. 상기 차라미터 최적화부(16)에서 최적화된, 상기 공격자가 점유하는 자원비율은 상기 연산부(15)로 피드백되어 상기 안전거래 한도액의 획득에 사용될 수 있다.

[0101] 이후에는 상기 안전거래 한도액을 얻기 위하여 하기 수학식 10이 적용된다.

수학식 10

[0102]
$$C_{\text{Req.}}=\frac{\mathbb{E}_X\left(p_A,t_{cut};N_{BC}\right)}{\mathbb{P}_{AS}\left(p_A,t_{cut};N_{BC}\right)}-\mathbb{E}\left[R\left(\lambda_A,T_{AS}\right)\right]$$

[0103] 상기 수학식 10에서 R은 보상금으로서 R은 하기 수학식 11과 같이 주어질 수 있다.

수학식 11

[0104]
$$R\left(\lambda_A,t\right):=\beta\lambda_A t\left(\log_{r_1} r_2\right)^{\lambda_A}\left(\log_{r_3} r_4\right)^{t}$$

[0105] 상기 수학식 11에서 로그텀(log term)은, 보상금이 시간에 대하여 로그함수 또는 지수함수로 가능한 것으로 보이는 것이다. 다만, 시간에 비례하는 함수로 주어지는 경우에는 로그텀들은 모두 1로 주어질 수 있다.

[0106] 다만, 상기 보상금 R도, 채굴비용(X)과 마찬가지로, n-제곱함수, n-루트함수, 로그함수, 지수함수, 또는 n-루트 함수가 될 수도 있을 것이다.

[0107] 다시 수학식 10을 참조하면, 여전히 손절시간(t_{cut})은 알려지지 않았고, 수학식 10은 손절시간을 인자로 하는 함수로 주어질 수 있다. 따라서, 손절시간을 변수로 하여 안전거래 한도액(C_{req})를 최적화할 수 있다.

[0108] 상기 손절시간을 변수로 상기 안전거래 한도액을 그래프로 예시하는 도면이 도 7에 도시된다.

[0109] 도 7을 참조하면, 상기 손절시간에 대하여 상기 안전거래 한도액은 아래로 볼록한 그래프로 나타난다. 가장 안전거래 한도액으로서 가장 낮은 검증금액을 출력할 수 있다. 상기 손절시간에 대응하여 상기 안전거래 한도액의 최적값을 구하는 작용은 상기 파라미터 최적화부(16)를 통하여 수행될 수 있다.

[0110] 상기 파라미터 최적화부(16)는, 현재 상태에서 가장 낮은 상기 안전거래 한도액을, 상기 연산부(15)로 출력할 수 있다.

[0111] 한편, 이상의 설명은 공격자의 자원비율이 50이하인 경우에 적용될 수 있다.

[0112] 반대로 공격자의 자원비율이 50%를 초과하는 경우에는 이미 설명한 바와 같이, 공격자는 손절시간을 무한대로 설정할 수 있다. 왜냐하면, 자원이 50%를 초과하는 경우에는 언젠가는 자원이 우세하므로 공격자가 언젠가는 더 긴 블록체인을 만들 수 있기 때문이다.

[0113] 상기 공격자의 자원비율이 50%이하인 경우(이하 50%공격이라고 한다)와, 상기 공격자의 자원비율이 50%초과하는 경우(이하에서는 51%공격이라고 한다)에, 추측검증값인, 공격자의 손절시간(T_{cut}) 및 공격자의 자원비율(P_A)의 최

적화에 대하여 상세하게 설명한다.

[0114] 상기 공격자의 손절시간(T_{cut}) 및 공격자의 자원비율(P_A)의 최적화는 상기 파라미터 최적화부(16)에서 수행될 수 있다.

[0115] 일반적으로, 51% 공격의 최적 검증횟수가 50% 공격의 최적 검증횟수보다 더 크다. 비트코인과 같은 계산자원의 규모가 거대한 네트워크에서는 특정 집단이 51% 이상의 계산 자원을 확보하기가 매우 어렵다. 따라서 50% 공격에 대한 대비를 하는 것으로도 충분하다. 반대로 계산자원의 규모가 작은 네트워크에서는 51% 공격이 가능할 수 있기 때문에 더욱 철저한 대비가 필요하다.

[0116] 첫번째로, 상기 51% 공격의 상황에서, 공격자에게 최선의 손절시간은 무한대이고, 최적의 공격자원 점유율은 계산가능하다.

[0117] 손절시간이 유한한 경우와 무한한 경우에 대해 이윤을 비교할 수 있다. 예를 들어, 블록 하나를 채굴할 때의 보상금(β)이 블록 하나를 채굴할 때의 비용(γ)보다 큰 경우, 즉 채굴로 얻는 시간당 보상이 채굴로 소요하는 시간당 비용보다 큰 경우에서는, 손절시간이 무한한 경우의 공격 이윤이 가장 크다.

[0118] 반대로, 블록 하나를 채굴할 때의 보상금(β)이 블록 하나를 채굴할 때의 비용(γ)보다 작은 경우에는 손절시간이 무한한 경우의 공격 이윤이 항상 가장 크다고 할 수는 없지만, 대부분의 실질적인 경우에서는 그러하다. 뿐만 아니라 정상적인 네트워크는 시장경제에 의해 β가 γ보다 크도록 형성되어있는 것이 보통이다.

[0119] 51% 공격 상황에서 손절시간이 무한할 때, 상기 안전거래 한도액($C_{Req.}$)은 공격자가 가지는 자원비율(p_A)에 대해 아래로 볼록인 함수이며, 이는 수학적으로 증명된 바가 있고 도 5와 같이 실제로 예시된다.

[0120] 즉, 공격자가 가지는 자원비율(p_A)에 대해 상기 안전거래 한도액(C_{Req})의 하한이 존재한다. 상기 입력 거래금액(C)과 상기 안전한 검증금맥(C_{Req})을 비교하여 $C<C^{*}_{Req.}$를 만족하면, 공격자가 가지는 모든 자원비율에 대해 $C<C_{Req.}$이므로, 거래가 안전함을 알 수 있다. 따라서, 거래자는 안전성을 판단하기 위하여, 공격 자원 점유율이 최악의 경우만 고려하면 된다. 예를 들어, 도 5에서 가장 가장 낮은 상기 안전거래 한도액을 선택하면 된다.

[0121] 두 번째로 50% 공격 상황에서도 공격자에게 최적의 손절시간(t_{cut})과 가장 합리적인 공격자가 가지는 자원비율(p_A)를 고려해 볼 수 있다.

[0122] 상기 50% 공격의 경우에는 공격성공 평균시간이 최소화되는 공격자가 가지는 자원비율(p_A)이 합리적인 공격 자원 점유율이다. 50% 공격인 경우에는 51% 공격과는 다르게 무한한 손절시간을 고려 할 필요가 없는데, 이는 공격자의 기대 손실이 무한히 커지기 때문이다.

[0123] 따라서 공격자는 반드시 유한한 손절시간의 선택이 강제된다. 공격자의 입장에서는 손절시간을 선택할 때 공격이 성공하기까지 소요되는 평균 시간, 즉 공격성공 평균시간을 고려하지 않을 수 없다. 예를 들어 공격성공 평균시간이 알려져 있다면, 공격을 시도 할 때 손절시간을 공격성공 평균시간, 그보다 조금 더 큰 시간, 혹은 그보다 수 배 큰 시간으로 설정 할 수 있다.

[0124] 50% 공격의 경우에는 공격 성공확률이 그리 높지 않으므로, 공격자는 공격을 여러 회 시도하여 이윤을 취할 것이다. 그러므로 매 회의 공격 시도에 걸리는 시간이 길면 길수록 적자에서 흑자로 전환되는데 걸리는 시간이 늦어질 것이다. 신속하고 정확하게 손절시간을 찾기 위해, 공격자는 공격성공 평균시간이 최소화 되는 공격자가 가지는 자원비율(p_A)을 탐색할 것이다.

[0125] 공격성공 평균시간이 도 6과 같이 아래로 볼록인 함수, 즉 최소값이 존재하는 함수임을 계산적으로 확인하였다. 따라서 50% 공격의 경우에는 공격성공 평균 간이 최소화되는 공격자가 가지는 자원비율(p_A)를 합리적인 공격 자원 점유율로 가정할 수 있다.

[0126] 또한, 50% 공격에서 공격자가 가지는 자원비율(p_A)이 고정되어 있을 때, 상기 안전거래 한도액(C_{Req})은 t_{cut}에 대해 아래로 볼록인 함수임을 도 7과 같이 확인 하였다. 즉, 손절시간에 대하여 상기 안전거래 한도액의 하한이 존재한다. 따라서 거래자는 안전성을 판단하기 위해 손절시간이 최악의 경우, 즉 상기 안전거래 한도액이 가장 낮은 경우만 고려하면 된다.

[0127] 공격자의 자원비율이 50%를 초과하는 경우에는 이미 설명한 바와 같이, 공격자는 손절시간을 무한대로 설정할

수 있다.

[0128] 이 경우에는 수학식 9에 제시되는, 손절시간이 무한대일 때 공격성공을 위한 평균소요시간($\mathbb{E}_{T_{AS-ICT}}(p_A;N_{BC})$)을 이용하여, 하기 수학식 12와 같이 상기 안전거래 한도액을 알아낼 수 있다. 다시 말하지만, 이 경우에 손절시간은 무한대이다.

수학식 12

[0129] $$C_{\text{Req.}} = \max\left(0,(\gamma-\beta)\lambda_T p_A \mathbb{E}_{T_{AS-ICT}}(p_A;N_{BC})\right)$$

[0130] 상기 수학식 12를 참조하면, 공격자가 점유하는 자원비율(p_A)은 알려지지 않았고, 수학식 12는 공격자가 점유하는 자원비율(p_A)을 인자로 하는 함수로 주어질 수 있다. 따라서, 공격자가 점유하는 자원비율(p_A)을 변수로 하여 안전거래 한도액(C_{req})를 최적화할 수 있다. 상기 최적화는 상기 파라미터 최적화부(16)에서 수행될 수 있다.

[0131] 공격자가 점유하는 자원비율(p_A)을 변수로, 상기 안전거래 한도액을 그래프로 예시하는 도면이 도 5로 이미 설명된 바가 있다.

[0132] 도 5를 참조하면, 상기 공격자가 점유하는 자원비율(p_A)에 대하여 상기 안전거래 한도액은 아래로 볼록한 그래프로 나타난다. 가장 안전거래 한도액으로서 가장 낮은 검증금액을 출력할 수 있다.

[0133] 본 문서에 기재되는 기술에 다르면, 사용자가 거래금액만을 입력하는 것에 의해서 적정 검증횟수를 출력하고, 상기 적정 검증횟수에 따라서 블록검증의 수를 설정할 수 있다. 이에 따르면, 안전하고 신속한 암호화폐를 이용한 상거래가 가능한 장점이 있다.

[0134] 또한, 거래의 안정성을 판단하기 위한 기준으로서, 상기 안전거래 한도액을 계산하는 과정에 추측되어야 하는 추측검증값이 최적화됨으로써, 사용자가 거래금액만을 입력하더라도 적정한 블록검증의 수를 알아낼 수 있다.

산업상 이용가능성

[0135] 본 발명에 따르면, 암호화폐를 통한 상거래를 안전하고 신속하게 수행할 수 있도록 함으로써, 암호화폐를 이용하는 실물거래를 더욱 촉진할 수 있다.

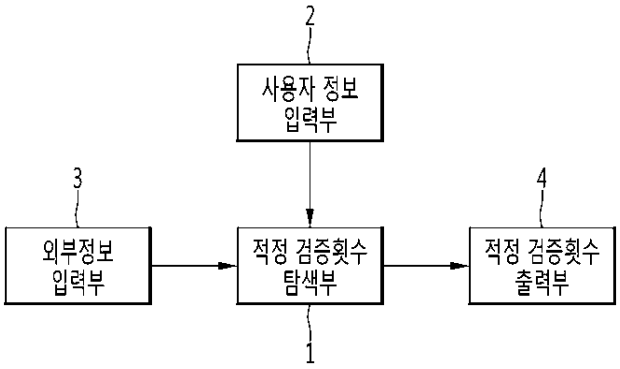
부호의 설명

[0136] 10: 안전거래 한도액 추출부

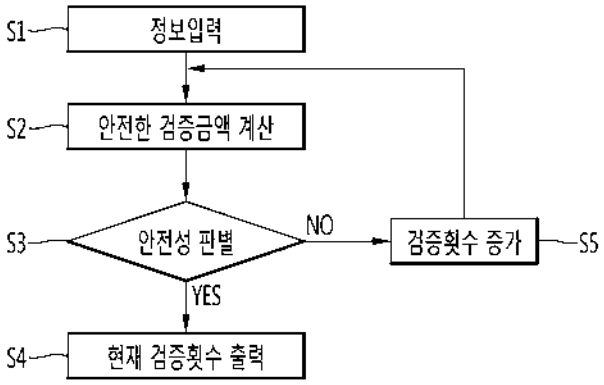
12: 안전성 판별부

도면

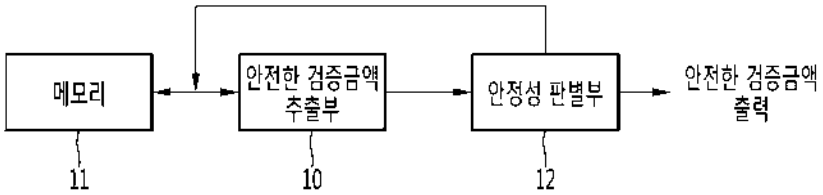
도면1



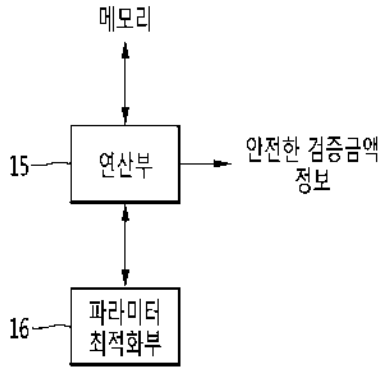
도면2



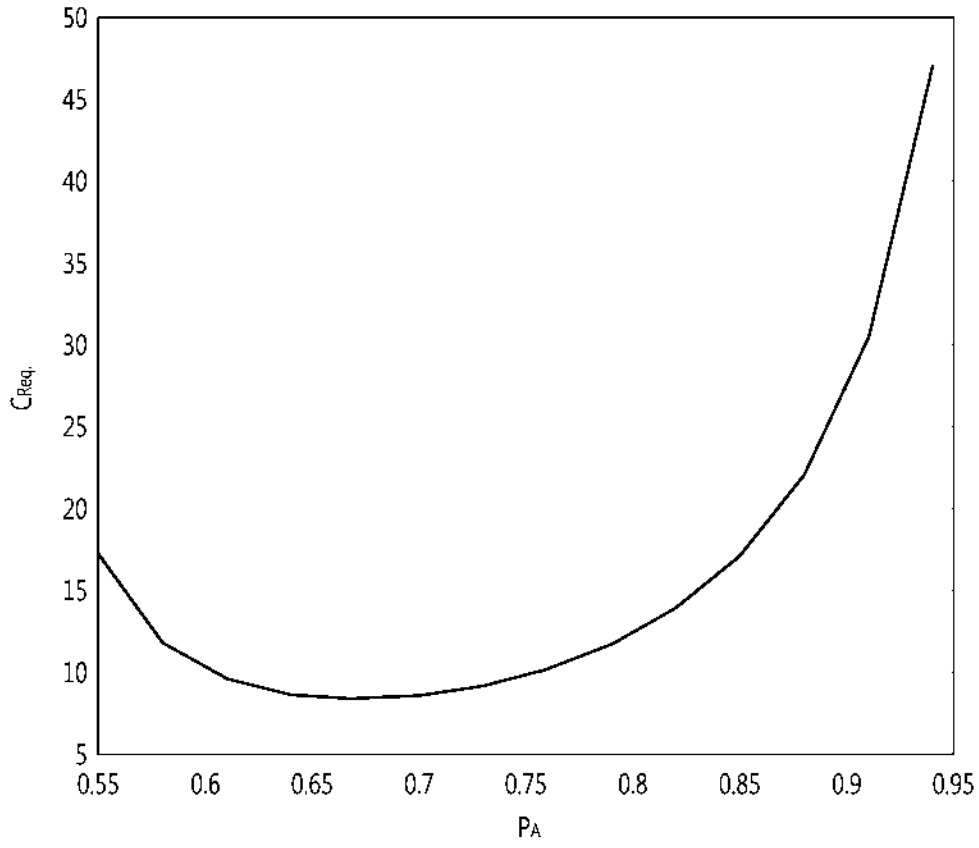
도면3



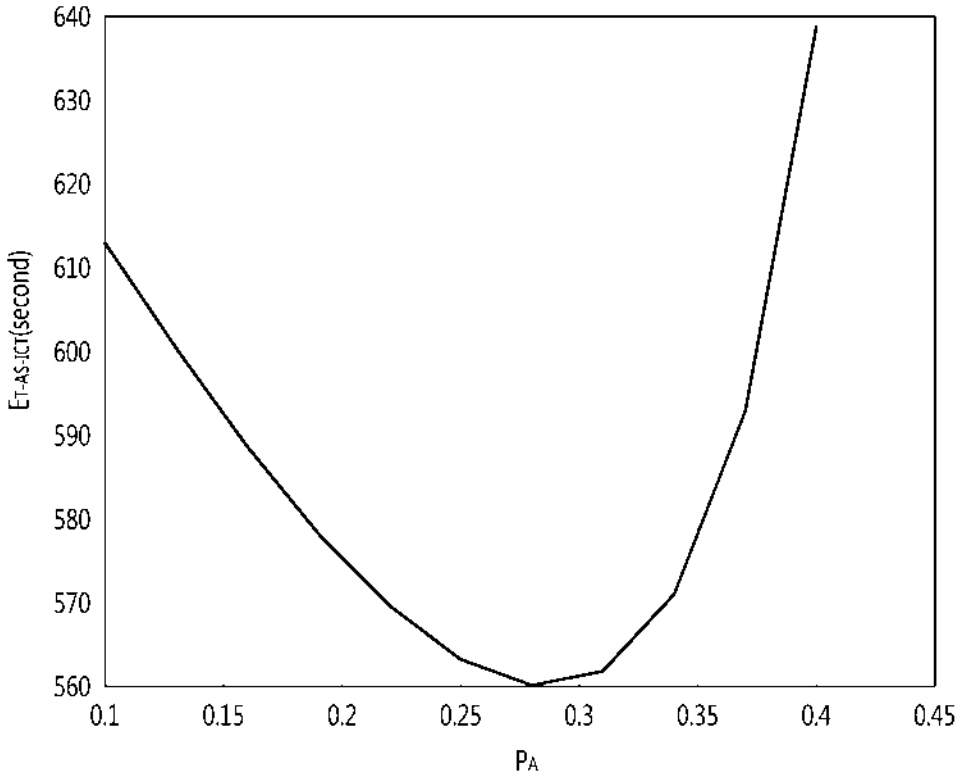
도면4


메모리
15
연산부
안전한 검증금액
정보
16
파라미터
최적화부

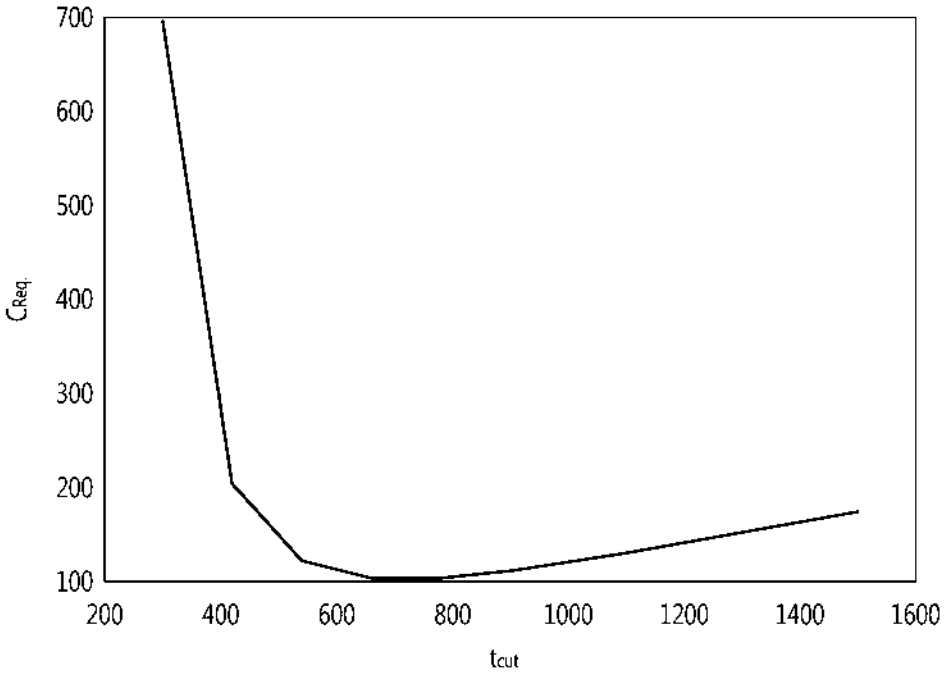
도면5


50
45
40
35
30
25
20
15
10
5
CReq.
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
0.95
PA

도면6


640
630
620
610
600
590
580
570
560
ET-AS-ICT(second)
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
PA

도면7


700
600
500
400
300
200
100
CReq.
200
400
600
800
1000
1200
1400
1600
tcut

(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년01월03일
(11) 등록번호 10-2347174
(24) 등록일자 2021년12월30일

(51) 국제특허분류(Int. Cl.)
G01R 23/16 (2006.01) *G01R 29/08* (2006.01)
G06N 20/20 (2019.01)
(52) CPC특허분류
G01R 23/16 (2021.05)
G01R 29/0892 (2013.01)
(21) 출원번호 10-2021-0095830
(22) 출원일자 2021년07월21일
심사청구일자 2021년07월21일
(56) 선행기술조사문헌
KR1020170115253 A*
KR1020180025721 A*
KR102251381 B1
KR101782493 B1
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
국방과학연구소
대전광역시 유성구 북유성대로488번길 160 (수남동)
광주과학기술원
광주광역시 북구 첨단과기로 123 (오룡동)
(72) 발명자
박진태
대전광역시 유성구 북유성대로488번길 160(수남동)
이홍노
광주광역시 북구 첨단과기로 123(오룡동)
강주성
광주광역시 북구 첨단과기로 123(오룡동)
(74) 대리인
유미특허법인

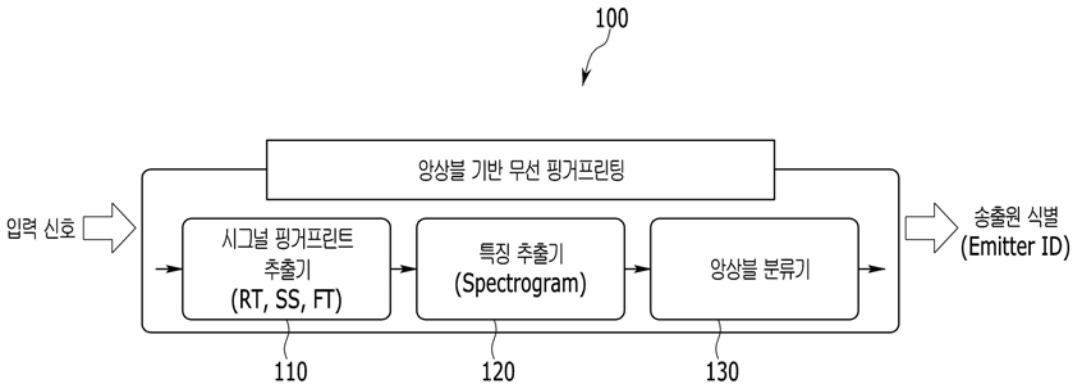
전체 청구항 수 : 총 13 항 심사관 : 오용균

(54) 발명의 명칭 **앙상블 기반 무선 핑거프린팅 장치 및 이를 이용한 송출원 식별 방법**

(57) 요 약

앙상블 기반 무선 핑거프린팅 장치는 입력 신호로부터 상승 천이상태 신호, 정상상태 신호 및 하강 천이상태 신호를 추출하는 시그널 핑거프린트 추출기, 및 상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호 각각의 스펙트로그램을 독립적인 각각의 기본 분류기의 출력으로부터 타겟 송출원이 속할 확률을 계산한 후 확률들 간의 곱에 대한 최대 변수를 확인하여 송출원을 식별하는 앙상블 분류기를 포함한다.

대 표 도



(52) CPC특허분류
G06N 20/20 (2021.08)

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

입력 신호로부터 상승 천이상태 신호, 정상상태 신호 및 하강 천이상태 신호를 추출하는 시그널 핑거프린트 추출기; 및

상기 상승 천이상태 신호의 제1 스펙트로그램을 학습하는 독립적인 제1 분류기의 출력으로부터 타겟 송출원이 속할 제1 확률을 계산하고, 상기 정상상태 신호의 제2 스펙트로그램을 학습하는 독립적인 제2 분류기의 출력으로부터 상기 타겟 송출원이 속할 제2 확률을 계산하고, 상기 하강 천이상태 신호의 제3 스펙트로그램을 학습하는 독립적인 제3 분류기의 출력으로부터 상기 타겟 송출원이 속할 제3 확률을 계산하고, 상기 제1 확률, 상기 제2 확률 및 상기 제3 확률을 모두 곱한 값에 대한 최대 변수를 확인하여 송출원을 식별하는 앙상블 분류기를 포함하는 앙상블 기반 무선 핑거프린팅 장치.

청구항 2

제1 항에 있어서,

상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호 각각의 스펙트로그램을 추출하는 특징 추출기를 더 포함하는 앙상블 기반 무선 핑거프린팅 장치.

청구항 3

제1 항에 있어서,

상기 제1 분류기는 상기 상승 천이상태 신호로부터 추출된 상기 제1 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하고,

상기 제2 분류기는 상기 정상상태 신호로부터 추출된 상기 제2 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하고,

상기 제3 분류기는 상기 하강 천이상태 신호로부터 추출된 상기 제3 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하는 앙상블 기반 무선 핑거프린팅 장치.

청구항 4

제3 항에 있어서,

상기 앙상블 분류기는 상기 제1 분류기, 상기 제2 분류기 및 상기 제3 분류기에서의 결과값을 모두 곱하여 최종 결정 벡터를 구하는 앙상블 기반 무선 핑거프린팅 장치.

청구항 5

물리 계층에서 복조되는 신호의 시간 영역에서의 상승 천이상태 신호, 정상상태 신호 및 하강 천이상태 신호 각각의 스펙트로그램을 추출하는 특징 추출기; 및

상기 상승 천이상태 신호의 제1 스펙트로그램을 학습하는 독립적인 제1 분류기의 출력으로부터 타겟 송출원이 속할 제1 확률을 계산하고, 상기 정상상태 신호의 제2 스펙트로그램을 학습하는 독립적인 제2 분류기의 출력으로부터 상기 타겟 송출원이 속할 제2 확률을 계산하고, 상기 하강 천이상태 신호의 제3 스펙트로그램을 학습하는 독립적인 제3 분류기의 출력으로부터 상기 타겟 송출원이 속할 제3 확률을 계산하고, 상기 제1 확률, 상기 제2 확률 및 상기 제3 확률을 모두 곱한 값에 대한 최대 변수를 확인하여 송출원을 식별하는 앙상블 분류기를 포함하는 앙상블 기반 무선 핑거프린팅 장치.

청구항 6

제5 항에 있어서,

입력 신호의 에너지를 모니터링하여 상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호를 검출하는 시그널 핑거프린트 추출기를 더 포함하는 앙상블 기반 무선 핑거프린팅 장치.

청구항 7

제5 항에 있어서,

상기 제1 분류기는 상기 상승 천이상태 신호로부터 추출된 상기 제1 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하고,

상기 제2 분류기는 상기 정상상태 신호로부터 추출된 상기 제2 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하고,

상기 제3 분류기는 상기 하강 천이상태 신호로부터 추출된 상기 제3 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하는 앙상블 기반 무선 핑거프린팅 장치.

청구항 8

제7 항에 있어서,

상기 앙상블 분류기는 상기 제1 분류기, 상기 제2 분류기 및 상기 제3 분류기에서의 결과값을 모두 곱하여 최종 결정 벡터를 구하는 앙상블 기반 무선 핑거프린팅 장치.

청구항 9

입력 신호로부터 상승 천이상태 신호, 정상상태 신호 및 하강 천이상태 신호를 추출하는 단계;

상기 상승 천이상태 신호의 제1 스펙트로그램을 학습하는 독립적인 제1 분류기의 출력으로부터 타겟 송출원이 속할 제1 확률을 계산하고, 상기 정상상태 신호의 제2 스펙트로그램을 학습하는 독립적인 제2 분류기의 출력으로부터 상기 타겟 송출원이 속할 제2 확률을 계산하고, 상기 하강 천이상태 신호의 제3 스펙트로그램을 학습하는 독립적인 제3 분류기의 출력으로부터 상기 타겟 송출원이 속할 제3 확률을 계산하는 단계; 및

상기 제1 확률, 상기 제2 확률 및 상기 제3 확률을 모두 곱한 값에 대한 최대 변수를 확인하여 송출원을 식별하는 단계를 포함하는 앙상블 기반 무선 핑거프린팅 장치를 이용한 송출원 식별 방법.

청구항 10

제9 항에 있어서,

상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호 각각의 스펙트로그램을 추출하는 단계를 더 포함하는 앙상블 기반 무선 핑거프린팅 장치를 이용한 송출원 식별 방법.

청구항 11

제9 항에 있어서,

상기 상승 천이상태 신호로부터 추출된 상기 제1 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하는 단계;

상기 정상상태 신호로부터 추출된 상기 제2 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하는 단계; 및

상기 하강 천이상태 신호로부터 추출된 상기 제3 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하는 단계를 포함하는 앙상블 기반 무선 핑거프린팅 장치를 이용한 송출원 식별 방법.

청구항 12

제9 항에 있어서,

상기 입력 신호의 에너지를 모니터링하여 상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호를 검출하는 앙상블 기반 무선 핑거프린팅 장치를 이용한 송출원 식별 방법.

청구항 13

제9 항에 있어서,

물리 계층에서 상기 입력 신호를 복조하여 시간 영역에서의 상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호를 추출하는 앙상블 기반 무선 핑거프린팅 장치를 이용한 송출원 식별 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 앙상블 기반 무선 핑거프린팅 장치 및 이를 이용한 송출원 식별 방법에 관한 것으로 더욱 상세하게는 상승 천이상태 신호, 정상상태 신호 및 하강 천이상태 신호를 동시에 사용하는 앙상블 접근 기반의 무선 핑거프린팅 장치 및 이를 이용한 송출원 식별 방법에 관한 것이다.

배 경 기 술

[0002] 무선 핑거프린팅(Radio Frequency(RF) fingerprinting)이란 수신된 전파 신호 속에서 시그널 핑거프린트(Signal Fingerprint, SF)를 찾아 분석하여 수신 신호가 어떤 신호 송출원으로부터 발생되었는지를 식별하는 기술이다. 여기서 시그널 핑거프린트(SF)는 수신된 신호로부터 계산되는 특정 수치적 값으로 계산될 수 있고 송출원 간의 차이를 특정할 수 있는 어떠한 형태로든 정의될 수 있다. 시그널 핑거프린트(SF)는 통신기 내부의 각 부품들이 갖는 비선형적인 특성으로부터 발생될 수 있다. 전력 증폭기(Power amplifier, PA), 주파수 발진기(Frequency Oscillator, FO) 혹은 송신 안테나(transmission antenna) 등 각 부품들은 제조 공정상의 자연적 특성(nature property)으로부터 의도치 않은 출력 오차값을 보이게 된다. 이러한 오차값들이 모여 최종적으로 전파 신호 속에 고유한 편차값을 형성하게 되며, 이 값을 마치 지문과도 같이 사용함으로써 수신된 전파 신호에 대한 송출원을 식별할 수 있다.

[0003] 최근 IoT(Internet of Things) 환경에서의 물리 계층 인증(Physical Layer authentication) 방법으로써 무선 핑거프린팅 기술이 각광 받고 있다. 이에 따라, 다양한 형태의 시그널 핑거프린트(SF) 또한 연구되고 있다. 신호의 시간 영역에서 가능한 시그널 핑거프린트(SF)를 다루는 상승 천이상태(rising transient, RT), 정상상태(steady state, SS) 그리고 하강 천이상태(falling transient, FT)의 신호에 대한 연구가 진행된 바 있다. 그리고 신호의 복조(demodulation) 과정에서 계산 가능한 시그널 핑거프린트(SF)를 의미하는 프리앰블 추출(preamble extraction), I/Q 컨스텔레이션 에러(constellation errors) 등에 대한 연구가 진행된 바 있다. 또한 최근에는 공간 영역에서의 안테나 빔 패턴(antenna beam pattern)을 시그널 핑거프린트(SF)로 사용하기 위한 연구가 보고된 바 있다.

[0004] 상승 천이상태(RT)의 경우, 터키의 Kara 교수 연구팀에서는 전파 신호 속 상승 천이상태(RT) 신호를 추출하여, 힐베르트 스펙트럼(Hilbert spectrum)을 구한 후 13종류의 특징(feature)을 계산하였으며, 선형 SVM(Support Vector Machine) 분류기를 이용해 학습하였다. 이를 통해, 20개의 휴대전화 속 블루투스 장치(Bluetooth device)에 대한 18~23dB SNR 환경에서의 90.5% 정확도를 확보할 수 있었다.

[0005] 정상상태(SS) 관련 최신 연구 사례의 경우, 중국의 한 연구팀에서는 수신된 전파 신호의 힐베르트-황 변환(Hilbert-Huang transform, HHT)을 구한 후, 이를 잔여 네트워크(residual network) 기반 딥러닝 분류기(deep learning classifier)에 학습시킨 바 있다. 이때, 특성 파라미터(parameter)를 바꿔가며 구성한 5개의 송출원에 대해 20dB SNR에서 95% 가까이의 정확도를 보여주었다.

[0006] 하강 천이상태(FT)의 경우, 본 연구팀은 수신된 전파 신호의 하강 천이상태(FT) 신호를 추출하고, 이를 상승 천이상태(RT), 정상상태(SS)와 함께 연결하는 연결 주파수 계수 특징(concatenate frequency coefficient feature)을 정의하고, 이를 희소 묘사 분류기(sparse representation classifier, SRC)를 통해 분류한 바 있다. 이를 통해 8개의 워키토키(walkie-talkie) 송출원에 대한 식별 정확도 99%를 확보하였다.

[0007] 신호의 복조 과정에서 정의할 수 있는 다양한 형태의 시그널 핑거프린트(SF)들이 연구되고 있다. 중국의 Huang 교수 연구팀에서는 프리앰블 신호에 대한 엔트로피(entropy) 계산을 통해 8개의 와이파이(Wi-Fi) 송출원에 대한 96.6% 정확도를 달성한 바 있다.

[0008] Xhi 교수 연구팀에서는 I/Q 신호에 대한 컨스텔레이션 에러를 이용하여, 합성곱 신경망(Convolutional Neural Network, CNN) 및 DBi-LSTM을 통해 16개의 USRP(Universal Software Radio Peripheral) 송출원에 대한 99.5%

정확도를 보여주었다.

[0009] 미국의 Sun 교수 연구팀에서는 수신되는 전파 신호의 공간 영역에서의 빔 패턴(Beam pattern)을 이용하여, 합성곱 신경망(CNN)을 통해 12개의 네트워크 인터페이스 컨트롤러(network interface controller, NIC)에 대한 99%의 식별 정확도를 보여주었다.

[0010] 상기의 기술들은 단 하나의 시그널 핑거프린트(SF)만을 특징으로 고려하고 있는 문제가 있다. 상기의 기술들은 각자의 시그널 핑거프린트(SF)를 정의하고 계산된 특징 벡터(feature vector)를 설계된 분류기(classifier)에 입력하는 방식을 취하고 있다.

발명의 내용

해결하려는 과제

[0011] 각각의 시그널 핑거프린트(SF)에는 서로 다른 독립적인 특징들이 내포되어 있기 때문에 단 하나의 시그널 핑거프린트(SF)만을 특징으로 고려하는 것은 최적의 방법이라고 할 수 없다. 일례로, 상승 천이상태(RT) 신호, 하강 천이상태(FT) 신호 등의 천이 신호의 경우, 목표하는 신호 레벨까지 신호 출력의 증가 및 감소 값을 의미하기에, 전력 증폭기(PA)에 대한 비선형적 요소가 잘 내포되어 있다. 반면에 정상상태(SS) 신호의 경우, 목표한 출력 내에서 정보를 나타내기 위해 신호가 변하기 때문에 주파수 발진기(FO)에 대한 비선형적 요소가 내포되어 있다. 이 경우, 신호 송출원을 식별할 때, 정상상태(SS) 신호 하나만을 보고 결정하는 것이 아닌, 상승 천이상태(RT) 및 하강 천이상태(FT) 신호에서의 특징도 분석하여 종합적으로 결정하는 것이 더 최적의 방법이라 할 수 있다.

[0012] 본 발명이 해결하고자 하는 기술적 과제는 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호를 동시에 사용하는 앙상블 접근 기반의 무선 핑거프린팅 장치 및 이를 이용한 송출원 식별 방법을 제공함에 있다.

과제의 해결 수단

[0013] 본 발명의 일 실시예에 따른 앙상블 기반 무선 핑거프린팅 장치는 입력 신호로부터 상승 천이상태 신호, 정상상태 신호 및 하강 천이상태 신호를 추출하는 시그널 핑거프린트 추출기, 및 상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호 각각의 스펙트로그램을 학습하는 독립적인 각각의 기본 분류기의 출력으로부터 타겟 송출원이 속할 확률을 계산한 후 확률들 간의 곱에 대한 최대 변수를 확인하여 송출원을 식별하는 앙상블 분류기를 포함한다.

[0014] 상기 앙상블 기반 무선 핑거프린팅 장치는 상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호 각각의 스펙트로그램을 추출하는 특징 추출기를 더 포함할 수 있다.

[0015] 상기 앙상블 분류기는, 상기 상승 천이상태 신호로부터 추출된 제1 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하는 제1 분류기, 상기 정상상태 신호로부터 추출된 제2 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하는 제2 분류기, 및 상기 하강 천이상태 신호로부터 추출된 제3 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하는 제3 분류기를 포함할 수 있다.

[0016] 상기 앙상블 분류기는 상기 제1 분류기, 상기 제2 분류기 및 상기 제3 분류기에서의 결과값을 모두 곱하여 최종 결정 벡터를 구할 수 있다.

[0017] 본 발명의 다른 실시예에 따른 앙상블 기반 무선 핑거프린팅 장치는 물리 계층에서 복조되는 신호의 시간 영역에서의 상승 천이상태 신호, 정상상태 신호 및 하강 천이상태 신호 각각의 스펙트로그램을 추출하는 특징 추출기, 및 상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호 각각의 스펙트로그램을 학습하는 독립적인 각각의 기본 분류기의 출력으로부터 타겟 송출원이 속할 확률을 계산한 후 확률들 간의 곱에 대한 최대 변수를 확인하여 송출원을 식별하는 앙상블 분류기를 포함한다.

[0018] 상기 앙상블 기반 무선 핑거프린팅 장치는 입력 신호의 에너지를 모니터링하여 상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호를 검출하는 시그널 핑거프린트 추출기를 더 포함할 수 있다.

[0019] 상기 앙상블 분류기는, 상기 상승 천이상태 신호로부터 추출된 제1 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하는 제1 분류기, 상기 정상상태 신호로부터 추출된 제2 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하는 제2 분류기, 및 상기 하강 천이상태 신호로부터 추출된 제3 스펙트로그램을 입력

샘플로 하여 딥러닝 분류기를 통해 학습하는 제3 분류기를 포함할 수 있다.

[0020] 상기 앙상블 분류기는 상기 제1 분류기, 상기 제2 분류기 및 상기 제3 분류기에서의 결과값을 모두 곱하여 최종 결정 벡터를 구할 수 있다.

[0021] 본 발명의 또 다른 실시예에 따른 앙상블 기반 무선 핑거프린팅 장치를 이용한 송출원 식별 방법은 입력 신호로부터 상승 천이상태 신호, 정상상태 신호 및 하강 천이상태 신호를 추출하는 단계, 상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호 각각의 스펙트로그램을 학습하는 독립적인 각각의 기본 분류기의 출력으로부터 타겟 송출원이 속할 확률을 계산하는 단계, 및 상기 계산된 확률들 간의 곱에 대한 최대 변수를 확인하여 송출원을 식별하는 단계를 포함한다.

[0022] 상기 앙상블 기반 무선 핑거프린팅 장치를 이용한 송출원 식별 방법은 상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호 각각의 스펙트로그램을 추출하는 단계를 더 포함할 수 있다.

[0023] 상기 앙상블 기반 무선 핑거프린팅 장치를 이용한 송출원 식별 방법은 상기 상승 천이상태 신호로부터 추출된 제1 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하는 단계, 상기 정상상태 신호로부터 추출된 제2 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하는 단계, 및 상기 하강 천이상태 신호로부터 추출된 제3 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습하는 단계를 포함할 수 있다.

[0024] 상기 입력 신호의 에너지를 모니터링하여 상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호를 검출할 수 있다.

[0025] 물리 계층에서 상기 입력 신호를 복조하여 시간 영역에서의 상기 상승 천이상태 신호, 상기 정상상태 신호 및 상기 하강 천이상태 신호를 추출할 수 있다.

발명의 효과

[0026] 상승 천이상태 신호, 정상상태 신호 및 하강 천이상태 신호를 동시에 사용하여 종합적으로 결정함으로써 시그널 핑거프린트의 효과, 즉 송출원 식별 정확도를 향상시킬 수 있다.

[0027] 즉, 단일 시그널 핑거프린트가 아닌 다수의 시그널 핑거프린트를 독립적인 분류기를 통해 학습함으로써 송출원 내의 비선형적 특성의 원인을 다양하고 공평하게 분석할 수 있다.

도면의 간단한 설명

[0028] 도 1은 본 발명의 일 실시예에 따른 앙상블 기반 무선 핑거프린팅 장치를 나타내는 블록도이다.

도 2는 본 발명의 일 실시예에 따른 시그널 핑거프린트 추출 과정을 나타내는 예시도이다.

도 3은 본 발명의 일 실시예에 따른 시그널 핑거프린트 추출 과정의 의사 코드(pseudo code)를 나타내는 예시도이다.

도 4는 본 발명의 일 실시예에 따른 특징 추출 과정을 나타내는 예시도이다.

도 5는 본 발명의 일 실시예에 따른 앙상블 분류기를 나타내는 블록도이다.

도 6은 일 비교예에 따른 앙상블 접근법을 나타내는 블록도이다.

도 7은 일 비교예에 따른 무선 핑거프린팅 방법을 나타내는 블록도이다.

도 8은 MAC 계층에서 무선 핑거프린팅을 수행하는 일 예를 나타내는 예시도이다.

도 9는 물리 계층에서 무선 핑거프린팅을 수행하는 일 예를 나타내는 예시도이다.

도 10은 본 발명의 일 실시예에 따른 앙상블 분류기의 식별 정확도를 실험한 결과를 나타내는 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0029] 이하, 첨부한 도면을 참고로 하여 본 발명의 실시예들에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예들에 한정되지 않는다.

[0030] 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 동일 또는

유사한 구성요소에 대해서는 동일한 참조 부호를 붙이도록 한다.

[0031] 또한, 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함" 한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.

[0032] 이하, 도 1 내지 5을 참조하여 본 발명의 일 실시예에 따른 앙상블 기반 무선 핑거프린팅 장치 및 이를 이용한 송출원 식별 방법에 대하여 설명한다.

[0033] 도 1은 본 발명의 일 실시예에 따른 앙상블 기반 무선 핑거프린팅 장치를 나타내는 블록도이다.

[0034] 도 1을 참조하면, 앙상블 기반 무선 핑거프린팅 장치(100)는 시그널 핑거프린트 추출기(110), 특징 추출기(120) 및 앙상블 분류기(130)를 포함한다.

[0035] 시그널 핑거프린트 추출기(110)는 입력 신호로부터 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호를 추출한다. 입력 신호는 송출원에서 송출된 무선 전파 신호일 수 있다. 시그널 핑거프린트 추출기(110)가 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호를 추출하는 과정에 대해서는 도 2 및 3을 참조하여 더욱 상세하게 설명한다.

[0036] 특징 추출기(120)는 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호 각각의 특징을 추출한다. 특징 추출기(120)는 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호 각각의 특징으로 각각의 스펙트로그램(Spectrogram)을 추출할 수 있다. 특징 추출기(120)가 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호 각각의 스펙트로그램을 추출하는 과정에 대해서는 도 4를 참조하여 더욱 상세하게 설명한다.

[0037] 앙상블 분류기(130)는 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호 각각의 스펙트로그램을 독립적인 각각의 기본 분류기(base classifier)를 통해 학습하고, 타겟 송출원에 속할 확률을 각각의 기본 분류기의 출력으로부터 계산한 후 각 확률들 간의 곱에 대한 최대 변수를 확인함으로써 송출원을 식별할 수 있다. 앙상블 분류기(130)가 송출원을 식별하는 과정에 대해서는 도 5를 참조하여 더욱 상세하게 설명한다.

[0038] 도 2는 본 발명의 일 실시예에 따른 시그널 핑거프린트 추출 과정을 나타내는 예시도이다. 도 3은 본 발명의 일 실시예에 따른 시그널 핑거프린트 추출 과정의 의사 코드(pseudo code)를 나타내는 예시도이다.

[0039] 도 2 및 3을 참조하면, 시그널 핑거프린트 추출 과정은 입력 신호(Hop signal)로부터 목표로하는 시그널 핑거프린트를 추출하는 신호처리 과정이다. 즉, 시그널 핑거프린트 추출 과정은 상승 천이상태(RT), 정상상태(SS) 및 하강 천이상태(FT) 중 적어도 하나를 추출하는 신호처리 과정이다. 시그널 핑거프린트 추출 과정의 입력 신호 및 결과 신호는 도 2에 예시되어 있으며, 시그널 핑거프린트 추출 과정의 의사 코드는 도 3에 예시되어 있다.

[0040] 시그널 핑거프린트 추출 과정은 기본적으로 잡음 신호 간 RF 신호의 존재 여부를 확인하는 신호 감지(signal detection) 과정과 밀접하다. 신호 감지 방법에는 상호 상관(cross correlation), 헤더 코드 감지(header code detection), 위상 루프 감지(phase loop detection), 에너지 감지(energy detection) 등 다양한 방법이 존재한다.

[0041] 본 발명의 실시예에서는 빠른 신호의 검출을 위한 l2 norm energy 기반 시그널 핑거프린트(SF) 검출 과정을 사용한다. 자세한 방법은 다음과 같다. 상승 천이상태(RT) 신호는 (잡음 신호로부터) RF 신호에서의 실제 통신이 이뤄지는 정상상태(SS) 신호까지 도달하는 신호 상승 구간으로 정의된다. 이를 반영하여, 입력 신호의 l2 norm energy를 모니터링하고, 일정 구간의 신호 에너지가 이전 구간 대비 10% 이상 상승할 경우를 상승 천이상태(RT) 구간(신호)으로 검출할 수 있다. 하강 천이상태(FT)는 정상상태(SS) 구간을 통해 통신이 종료된 후 (잡음 신호까지) 도달하는 신호 하강 구간으로 정의된다. 이를 반영하여, 입력 신호의 에너지가 이전 구간 대비 10% 이상 하강할 경우를 하강 천이상태(FT) 구간(신호)으로 검출할 수 있다. 정상상태(SS) 신호는 상승 천이상태(RT) 구간과 하강 천이상태(FT) 구간 사이의 통신을 위한 RF 신호 구간으로 정의된다.

[0042] 이러한 정의에 따라 시그널 핑거프린트 추출기(110)는 입력 신호로부터 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호 각각을 추출할 수 있다.

[0043] 도 4는 본 발명의 일 실시예에 따른 특징 추출 과정을 나타내는 예시도이다.

[0044] 도 4를 참조하면, 특징 추출 과정은 시그널 핑거프린트 추출 과정에서 검출된 시그널 핑거프린트로부터 분류 결과를 극대화(높은 분류 정확도를 획득)할 수 있는 특징 도메인(feature domain)으로의 추가적인 계산 과정을 의미한다. 특징 추출 방법으로 통계적 모멘텀(Statistical momentum), 1D 신호 분할(signal segmentation) 등의

다양한 방법이 있다. 본 발명의 실시예에서는 RF 신호의 시간 및 주파수 도메인에서의 전력 밀도 행동(power-density behavior)이 중요하다고 판단하여 특징 추출 방법으로 신호의 스펙트로그램을 계산한다.

[0045] 신호의 스펙트로그램 계산을 위해, 수학식 1과 같이 단시간 푸리에 변환(Short-Time Fourier Transform, STFT)을 사용할 수 있으며, 신호 샘플링 속도(signal sampling rate)에 맞춰 4096 point FFT 및 1024 point window size를 512 point overlap하여 계산할 수 있다.

수학식 1

[0046]
$$\mathbf{x}_{Feature} = spectrogram\{\mathbf{x}_{SF}\} = |STFT\{\mathbf{x}_{SF}\}|^2$$

[0047] 도 5는 본 발명의 일 실시예에 따른 앙상블 분류기를 나타내는 블록도이다.

[0048] 도 5를 참조하면, 앙상블 분류기(130)는 상승 천이상태(RT) 기반의 제1 분류기(131), 정상상태(SS) 기반의 제2 분류기(132) 및 하강 천이상태(FT) 기반의 제3 분류기(133)를 포함한다.

[0049] 제1 분류기(131)는 상승 천이상태(RT) 신호로부터 추출된 제1 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기(deep learning classifier)를 통해 학습한다. 제2 분류기(132)는 정상상태(SS) 신호로부터 추출된 제2 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습한다. 제3 분류기(133)는 하강 천이상태(FT) 신호로부터 추출된 제3 스펙트로그램을 입력 샘플로 하여 딥러닝 분류기를 통해 학습한다. 제1 분류기(131), 제2 분류기(132) 및 제3 분류기(133)는 다양한 형태의 딥러닝 분류기를 적용할 수 있다. 도 5에서는 인셉션 블록(Inception block)에 기반한 딥 인셉션 네트워크(Deep Inception Network) 구조의 기본 분류기를 예시하고 있다.

[0050] 앙상블 분류기(130)는 타겟 송출원 k가 속할 확률 $p(c_k; x_{ss})$를 제1 분류기(131), 제2 분류기(132) 및 제3 분류기(133) 각각의 출력 y로부터 계산할 수 있다. 그리고 앙상블 분류기(130)는 제1 분류기(131), 제2 분류기(132) 및 제3 분류기(133) 각각의 출력으로부터 계산된 확률들 간의 곱에 대한 최대 변수 k를 확인함으로써 송출원을 식별할 수 있다.

[0051] 더욱 상세하게, 앙상블 분류기(130)는 수학식 2 및 3에 따라 최종 결정(송출원 식별)을 내릴 수 있다.

수학식 2

[0052]
$$\arg\max_{k}\{p(c_k;x_{SF})\}$$

수학식 3

[0053]
$$\arg\max_{k}\prod_{SF}p(c_k;x_{SF})$$

[0054] 여기서, SF는 시그널 핑거프린트들(Signal Fingerprints)을 의미하고, 가용 가능한 시그널 핑거프린트들의 세트에 대한 요소이다. 즉, $SF \in \{RT, SS, FT\}$ 이며, 하나의 기본 분류기는 목표하는 시그널 핑거프린트가 정해졌을 때(예를 들어, SF=RT), 주어진 목표 시그널 핑거프린트에 대한 딥 인셉션 분류기(deep inception classifier)를 학습하는 것이다(예를 들어, $\mathbf{y}_{RT} = \mathbf{f}_{Classifier-RT}(\mathbf{x}_{RT})$). 이때, 딥 인셉션 분류기의 출력 벡터 y는 학습한 클래스(Class)의 총 개수만큼의 길이를 갖는 $C \times 1$ 벡터 값을 출력한다(예를 들어, $\mathbf{y} := [y_1, y_2, \ldots, y_C]$). 이 벡터 값을 이용하여 입력된 시그널 핑거프린트의 샘플이 특정 클래스에 포함될 확률

$p(c_k;x_{SF})$를 구할 수 있다(예를 들어, 소프트맥스 함수, $p(c_k;x_{SF})=\frac{e^{y_k}}{\sum_C e^{y_j}}$). 앙상블 분류기(130)는 해당 클래스에 포함될 확률 중 가장 큰 값을 가지는 최대 변수 k를 찾을 수 있다.

[0055] 앙상블 분류기(130)는 기본 분류기인 제1 분류기(131), 제2 분류기(132) 및 제3 분류기(133)에서의 결과값을 모두 곱(product)하여 최종 결정 벡터를 구할 수 있다. 즉, 수학식 3은 $\arg\max_k\{p(c_k;x_{RT})\times p(c_k;x_{SS})\times p(c_k;x_{FT})\}$가 되며, 앙상블 분류기(130)는 개별 목표로 하는 시그널 핑거프린트, 즉 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호를 따로 학습한 제1 분류기(131), 제2 분류기(132) 및 제3 분류기(133)의 결과 벡터에 대한 곱셈을 통해 최종 결정을 구할 수 있다.

[0056] 이하, 도 6을 참조하여 본 발명의 실시예와 다른 앙상블 접근법에 대한 비교예에 대하여 설명한다.

[0057] 도 6은 일 비교예에 따른 앙상블 접근법을 나타내는 블록도이다.

[0058] 도 6을 참조하면, 일반적인 앙상블 접근법(Ensemble approach)은 동일한 데이터 세트에 대한 다수의 약한 추정기(weak estimator)를 이용하여 하나의 추정기를 사용할 때 보다 더 좋은 결과를 이끌어내는 방식을 말한다.

[0059] 비교예에 따른 앙상블 분류기는 동일한 정상상태(SS) 신호로부터 개별 분류기(classifier 1, classifier 2, classifier 3)에 대한 학습을 진행한다. 하지만, 이는 정상상태(SS) 신호에 대한 추정만 여러 번 수행할 뿐이고 하나의 시그널 핑거프린트를 기준으로 설계된 방법이라 할 수 있으며, 본 발명의 실시예에 따른 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호를 종합적으로 분석하여 결정하는 방법이라 할 수 없다.

[0060] 본 발명의 실시예에 따른 앙상블 분류기(130)는 입력된 전파 신호로부터 목표하는 개별 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호를 추출하고, 각각의 스펙트로그램을 계산하여 이를 학습하고, 그 결과를 종합하여 더 좋은 최종 결과를 도출해 낼 수 있다. 이러한 설계를 통해, 다수의 시그널 핑거프린트의 특성을 독립적으로 분석하고, 그 결과를 종합적으로 이용할 수 있다.

[0061] 이하, 도 7을 참조하여 본 발명의 실시예와 다른 무선 핑거프린팅 방법에 대한 비교예에 대하여 설명한다.

[0062] 도 7은 일 비교예에 따른 무선 핑거프린팅 방법을 나타내는 블록도이다.

[0063] 도 7을 참조하면, 비교예에 따른 다수의 특성인자를 이용한 무선 무선 핑거프린팅 방법은 수신된 전파 신호로부터 목표하는 시그널 핑거프린트인 상승 천이상태(RT) 신호 및 정상상태(SS) 신호를 계산하고, 각 신호의 주파수 계수(frequency coefficient)를 구한 후 이를 연관시켜 SRC(Sparse Representation Classifier)라는 분류 기술로 식별하는 방식이다. 이와 같은 접근을 통해, 다수의 시그널 핑거프린트에 대한 분석을 반영하였으며, 다양한 환경에서의 높은 식별 정확도 달성을 목표하고 있다.

[0064] 본 발명의 실시예와 비교예의 무선 핑거프린팅 방법의 차이는 분류기의 구조에서 비롯된다. 비교예에 따른 무선 핑거프린팅 방법으로 분석할 경우, 다수의 시그널 핑거프린트를 반영한다는 점에서 이점을 얻을 수 있으나, 결국에는 이를 분석하여 식별하는 분류기는 단일 구조이다. 이는 시그널 핑거프린트들 간의 신호 세기가 서로 비슷할 경우에는 괜찮으나, 신호 차이가 클 경우에는 주요한 시그널 핑거프린트에만 의존하여 분석이 수행된다는 단점이 존재한다. 즉, 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호를 복합적으로 분석하는 것이 아닌, 셋 중에 가장 신호 세기가 큰 시그널 핑거프린트를 중점으로 최종 결정이 수행되는 것이다.

[0065] 반면에 본 발명의 실시예에 따른 무선 핑거프린팅 방법에서는 분류기의 구조가 완전히 독립되어 있다. 하나의 기본 분류기는 목표하는 시그널 핑거프린트만을 분석하여 송출원에 대한 확률을 추정하고, 이렇게 개별 시그널 핑거프린트별로 추정된 확률을 종합적으로 분석하여 최종 결정을 내림으로써, 시그널 핑거프린트 간의 신호 차이가 클 때도 공평하게(fairness) 시그널 핑거프린트들을 비교할 수 있다.

[0066] 한편, 본 발명의 실시예에 따른 무선 핑거프린팅 방법은 물리 계층에서 수행될 수 있으며, 이에 따라 기존에 비해 빠르게 송출원을 식별할 수 있다. 이에 대하여 도 8 및 9를 참조하여 설명한다.

[0067] 도 8은 MAC 계층에서 무선 핑거프린팅을 수행하는 일 예를 나타내는 예시도이다. 도 9는 물리 계층에서 무선 핑거프린팅을 수행하는 일 예를 나타내는 예시도이다.

[0068] 도 8 및 9를 참조하면, 일반적으로 통신 시스템은 도 8에 예시한 바와 같이 신호 송출원 식별을 위하여 MAC 헤

더(header)에서의 ID 정보를 분석한다. 이에 따라, MAC 계층까지의 복조, 비트 레벨 디코딩(bit level decoding) 및 헤더 해독(header decryption) 과정이 필요하다. 하지만 이는 시긴트(SIGINT, signal intelligence) 목적, 즉 프로토콜(protocol)에 대한 상세 정보를 모르는 공격자 입장에서는 쉽지 않은 일이다.

[0069] 본 발명의 실시예에 따른 무선 핑거프린팅 방법은 도 9에 예시한 바와 같이 디코딩 및 해독 과정이 필요 없는 시간 영역에서의 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호를 활용함으로써 물리 계층에서의 정보만으로 신호 송출원을 더욱 빨리 식별할 수 있다.

[0070] 이하, 도 10을 참조하여 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호 각각의 스펙트로그램을 계산하여 이를 학습하고, 그 결과를 종합하는 본 발명의 실시예에 따른 무선 핑거프린팅 방법의 효과에 대하여 설명한다.

[0071] 도 10은 본 발명의 일 실시예에 따른 앙상블 분류기의 식별 정확도를 실험한 결과를 나타내는 그래프이다.

[0072] 도 10을 참조하면, 본 발명의 실시예에 따른 송출원 식별 방법은 무선 아날로그 신호에서 검출 가능한 시그널 핑거프린트들, 즉 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호를 이용한 앙상블 접근법을 포함한다. 송출원 식별에 사용하는 시그널 핑거프린트들의 조합에 따라 앙상블 분류기(130)의 식별 정확도가 달라질 수 있다.

[0073] 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호의 조합에 따른 식별 정확도를 실험한 결과를 보면, 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호를 모드 사용한 Ensen(초록색), 입력 신호를 바로 사용한 Hop(하늘색), 그리고 정상상태(SS) 신호를 사용한 SS(보라색)의 성능 결과가 높게 형성되는 것을 알 수 있다.

[0074] 이 중에서도 본 발명의 실시예에 따른 앙상블 접근법을 사용한 송출원 식별 방법이 다른 방법에 대비하여 1~2% 이상 식별 정확도가 높은 것을 확인할 수 있다.

[0075] 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호의 조합에 따른 기대 효과는 다음과 같다.

[0076] 상승 천이상태(RT) 신호, 정상상태(SS) 신호 및 하강 천이상태(FT) 신호를 모두 사용 시(Ensen), 송출원 식별 성능이 가장 좋다. 심지어, 시그널 핑거프린트들을 따로 추출하지 않은 입력 신호(또는 홉 신호(Hop signal)) 경우보다 더 나은 성능을 보여준다. 이는 주로 장치 인증(Device Authentication)에 사용될 수 있으며, 아날로그 신호에서의 복제 어려움에 기반하여, IoT 환경에서의 물리 계층에서의 이중 보안 시스템 등에 사용될 수 있다.

[0077] 상승 천이상태(RT) 신호 및 하강 천이상태(FT) 신호 사용 시(Ensen(RTFT))는 신호의 앞부분 및 끝부분만을 사용하여 인증을 수행하는 경우로써, 정상상태(SS) 신호에 추가적인 과정을 수행할 때 사용될 수 있으며, 전자전 환경에서의 재밍(Jamming)을 일예로 들 수 있다. 통신 신호 재밍 시 적군의 정보 공유를 방해하기 위해, 주 통신이 이뤄지는 정상상태(SS) 신호를 재밍하는 것이 중요하다. 이때, 재밍 신호에 대한 ID 확인 과정을 통해 올바르게 재밍이 수행 중인지가 추적될 수 있다.

[0078] 상승 천이상태(RT) 신호 및 정상상태(SS) 신호 사용 시, 또는 정상상태(SS) 신호 및 하강 천이상태(FT) 신호 사용 시는 신호의 앞부분 또는 뒷부분만 활용하여 ID를 확인하는 과정을 의미하며, 마찬가지로 인증 목적으로 사용될 수 있으며, 앞서 설명한 활용도에 다양하게 사용될 수 있다. 다만, 일부 시그널 핑거프린트가 제외됨에 따라 식별 성능이 열화될 수 있으므로, 특수한 경우가 아닌 이상 본 발명의 실시예와 같이 가용 가능한 시그널 핑거프린트를 모두 활용하는 것이 올바른 접근 방법이라 할 수 있다.

[0079] 본 발명의 실시예에 따른 앙상블 기반 무선 핑거프린팅 장치(100) 및 이를 이용한 송출원 식별 방법은 하드웨어나 소프트웨어, 또는 하드웨어와 소프트웨어의 결합으로 구현될 수 있다. 예를 들어, 앙상블 기반 무선 핑거프린팅 장치(100) 및 이를 이용한 송출원 식별 방법의 적어도 일부는 집적회로(IC)와 같이 하드웨어적으로 구현되거나, 컴퓨터 프로그램과 같이 소프트웨어적으로 구현되거나, 또는 컴퓨터 프로그램이 기록된 기록매체와 같이 하드웨어와 소프트웨어의 결합으로 구현될 수 있다.

[0080] 지금까지 참조한 도면과 기재된 발명의 상세한 설명은 단지 본 발명의 예시적인 것으로서, 이는 단지 본 발명을 설명하기 위한 목적에서 사용된 것이지 의미 한정이나 특허청구범위에 기재된 본 발명의 범위를 제한하기 위하여 사용된 것은 아니다. 그러므로 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시 예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 특허청구

범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

[0081] 100: 앙상블 기반 무선 핑거프린팅 장치

110: 시그널 핑거프린트 추출기

120: 특징 추출기

130: 앙상블 분류기

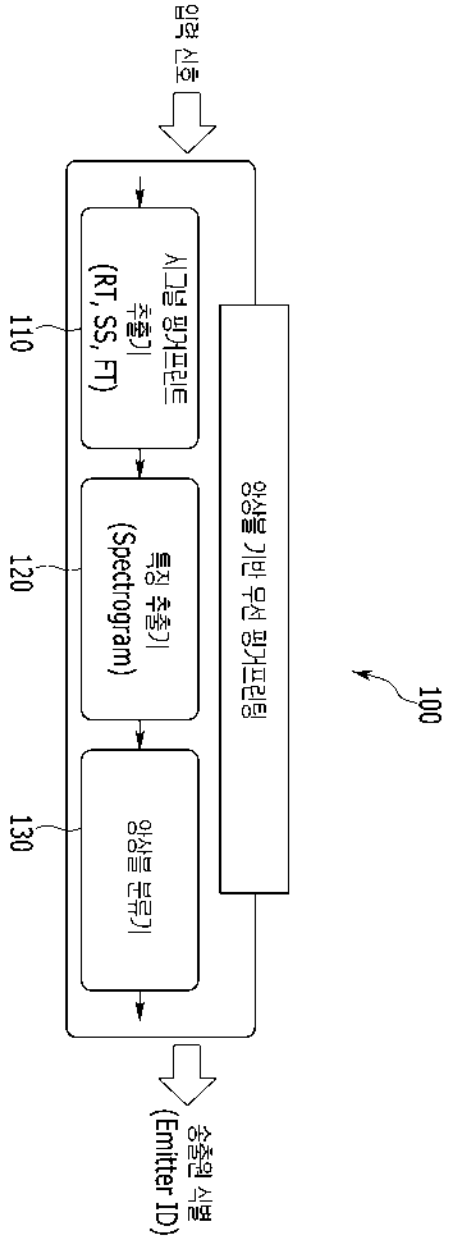
131: 제1 분류기

132: 제2 분류기

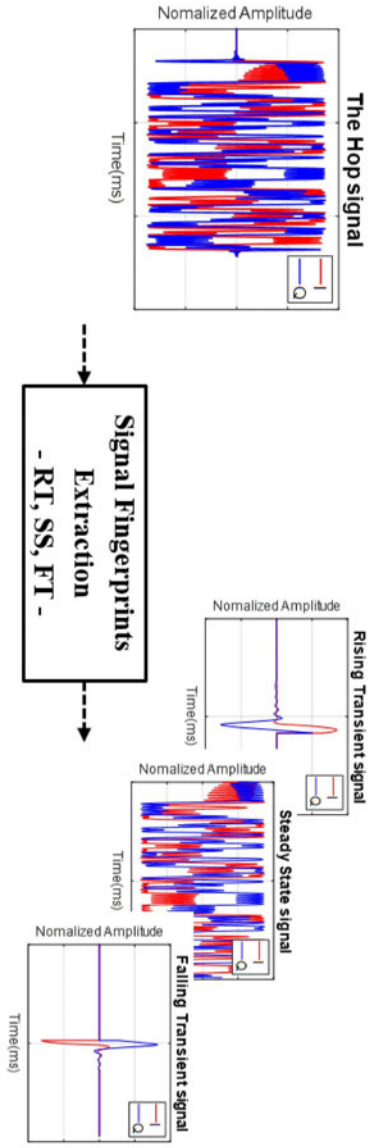
133: 제3 분류기

도면

도면1



도면2


Nomalized Amplitude
Time(ms)
The Hop signal
Signal Fingerprints Extraction
- RT, SS, FT -
Nomalized Amplitude
Time(ms)
Rising Transient signal
Nomalized Amplitude
Time(ms)
Steady State signal
Nomalized Amplitude
Time(ms)
Falling Transient signal

도면3

PSEUDO CODE OF THE SIGNAL FINGERPRINTS EXTRACTIONS

Given: The FH signal x

Initialize: $i = 1$, $T^{RT} = T^{FT} = \emptyset$ for time periods, $Window_{size}$ and Bandwidth of Interest (BOI) BW_{BOI}.

Steps 1: (Extract the target signal fingerprint)

while $i + Window_{size}$ is smaller than length(x) do:

$x_n = x[i : i + window_{size}]$

Calculate the l2 norm energy: $E_n = \|x_n\|_2$

If $((E_n - E_{n-1}) / E_{n-1} \geq 0.1)$: then $T^{RT} \leftarrow [T^{RT}\ i]$

If $((E_n - E_{n-1}) / E_{n-1} \leq -0.1)$: then $T^{FT} \leftarrow [T^{FT}\ i]$

$i = i + 0.5 \times window_{size}$

Set the targets of signal fingerprint x_{SF} by following definitions:

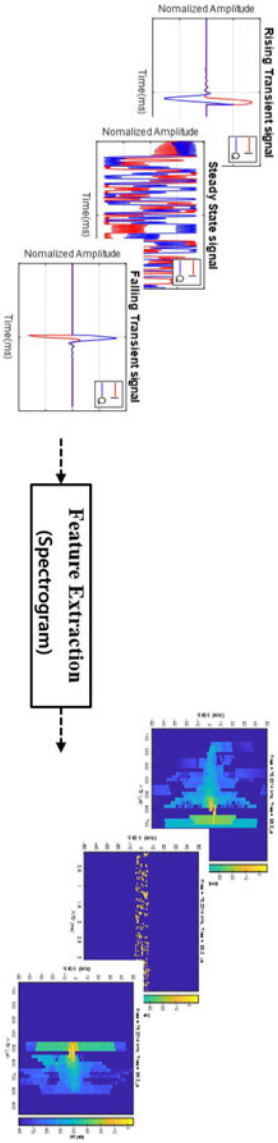
$x_{RT} = x[T^{RT}[1] : T^{RT}[-1]]$,

$x_{SS} = x[T^{RT}[-1] : T^{FT}[1]]$,

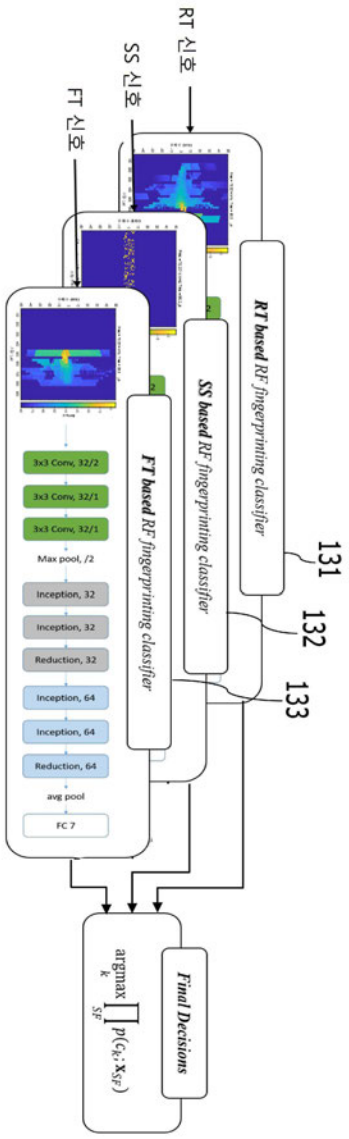
$x_{FT} = x[T^{FT}[1] : T^{FT}[-1]]$,

Outputs: Return the target signal fingerprints, y.

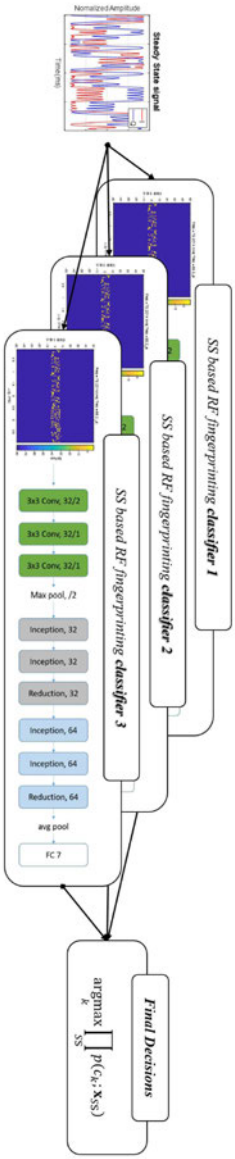
도면4


Normalized Amplitude
Time(ms)
Rising Transient signal
Normalized Amplitude
Time(ms)
Steady State signal
Normalized Amplitude
Time(ms)
Falling Transient signal
Feature Extraction
(Spectrogram)

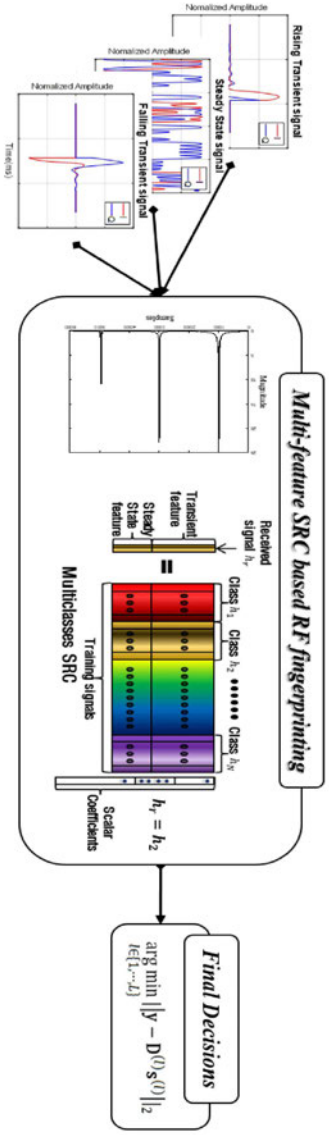
도면5


RT 신호
SS 신호
FT 신호
RT based RF fingerprinting classifier
SS based RF fingerprinting classifier
FT based RF fingerprinting classifier
131
132
133
3x3 Conv, 32/2
3x3 Conv, 32/1
3x3 Conv, 32/1
Max pool, /2
Inception, 32
Inception, 32
Reduction, 32
Inception, 64
Inception, 64
Reduction, 64
avg pool
FC 7
argmax
Final Decisions

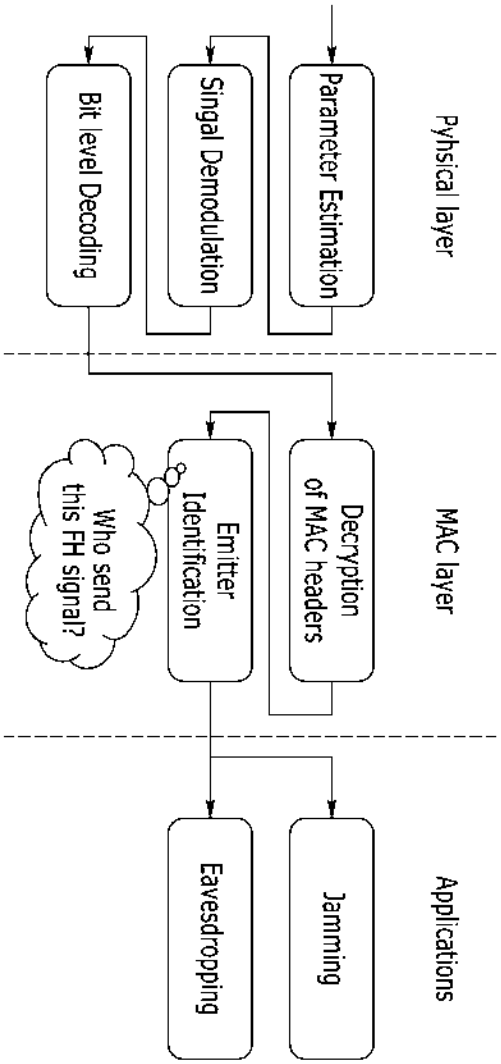
도면6


Normalized Amplitude
Steady State signal
SS based RF fingerprinting classifier 1
SS based RF fingerprinting classifier 2
SS based RF fingerprinting classifier 3
3x3 Conv, 32/2
3x3 Conv, 32/1
3x3 Conv, 32/1
Max pool, /2
Inception, 32
Inception, 32
Reduction, 32
Inception, 64
Inception, 64
Reduction, 64
avg pool
FC 7
Final Decisions

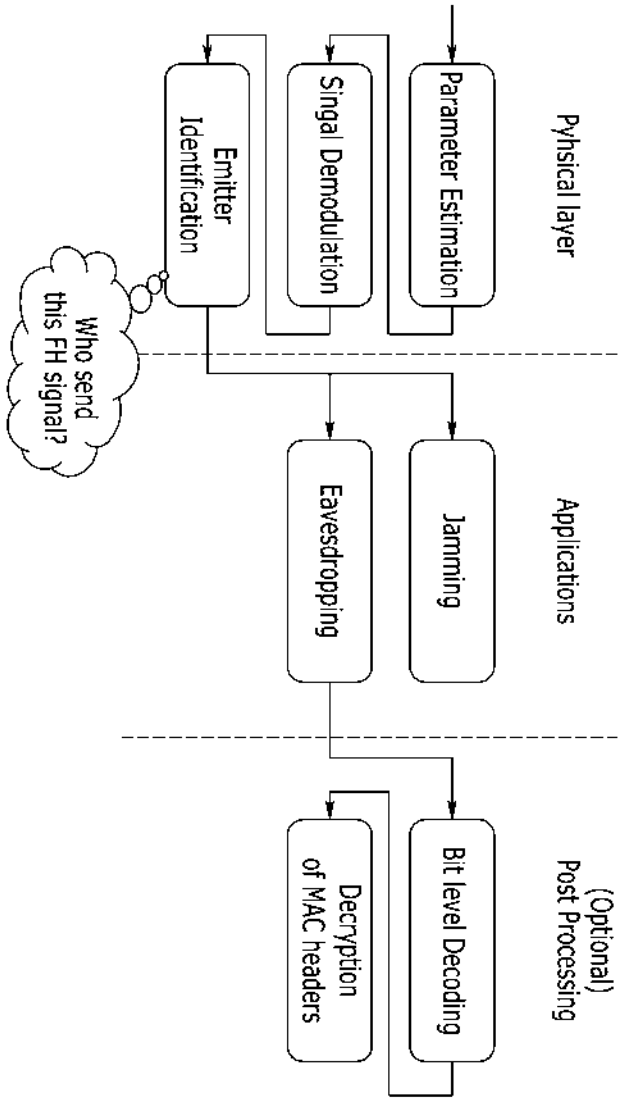
도면7


Normalized Amplitude
Rising Transient signal
Steady State signal
Failing Transient signal
Time (ms)
Multi-feature SRC based RF fingerprinting
Transient feature
Steady State feature
Received signal h_r
Class h_1
Class h_2
Class h_N
Training signals
Multiclasses SRC
Scalar Coefficients
$h_r = h_2$
Final Decisions
$\arg\min_{i\in\{1,\cdots,L\}} \|\mathbf{y} - \mathbf{D}^{(i)}\mathbf{s}^{(i)}\|_2$

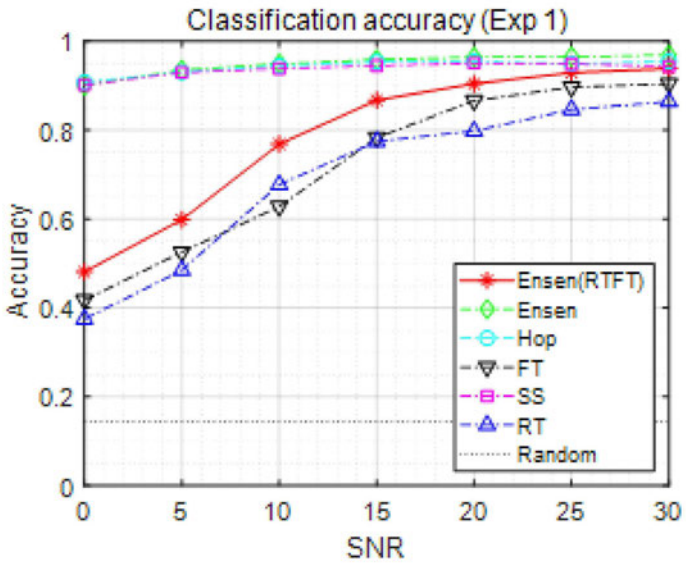
도면8


Pyhsical layer
Parameter Estimation
Singal Demodulation
Bit level Decoding
MAC layer
Decryption of MAC headers
Emitter Identification
Who send this FH signal?
Applications
Jamming
Eavesdropping

도면9


Who send this FH signal?
Emitter Identification
Singal Demodulation
Parameter Estimation
Pyhsical layer
Eavesdropping
Jamming
Applications
Decryption of MAC headers
Bit level Decoding
(Optional) Post Processing

도면10


Classification accuracy (Exp 1)
Accuracy
SNR
0
0.2
0.4
0.6
0.8
1
0
5
10
15
20
25
30
Ensen(RTFT)
Ensen
Hop
FT
SS
RT
Random

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0137795
(43) 공개일자 2018년12월28일

(51) 국제특허분류(Int. Cl.)
G01J 3/45 (2006.01)
(52) CPC특허분류
G01J 3/45 (2013.01)
G01J 2003/452 (2013.01)
(21) 출원번호 10-2017-0077480
(22) 출원일자 2017년06월19일
심사청구일자 2017년06월19일

(71) 출원인
광주과학기술원
광주광역시 북구 첨단과기로 123 (오룡동)
(72) 발명자
김철순
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 전기전자컴퓨터공학부
이홍노
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 전기전자컴퓨터공학부
이웅비
광주광역시 북구 첨단과기로 123(오룡동) 광주과학기술원 전기전자컴퓨터공학부
(74) 대리인
김기문

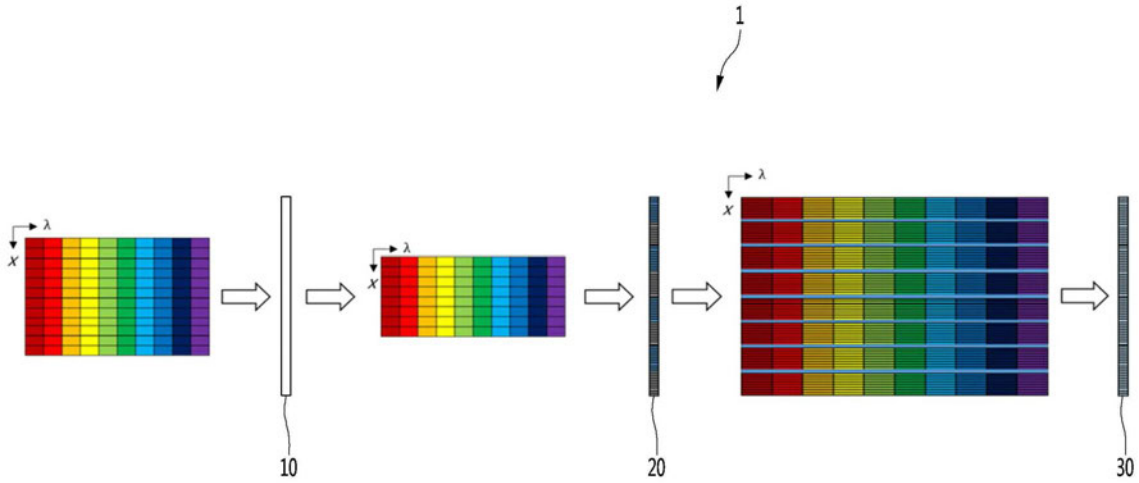
전체 청구항 수 : 총 11 항

(54) 발명의 명칭 **하이퍼스펙트럼 이미지 장치**

(57) 요 약

이미징 장치가 개시된다. 본 발명의 일 실시 예에 따른 이미징 장치는, 입력 소스를 공간 도메인에서 변조하는 멀티 어퍼처, 상기 멀티 어퍼처에서 변조된 입력 소스를 스펙트럼 도메인에서 변조하는 랜덤 필터 어레이 및 상기 변조된 소스를 스펙트럼 범위 감도에 따라 통합하는 포토 센서를 포함한다.

대 표 도



명 세 서

청구범위

청구항 1

입력 소스를 공간 도메인에서 변조하는 멀티 어퍼처;

상기 멀티 어퍼처에서 변조된 입력 소스를 스펙트럼 도메인에서 변조하는 랜덤 필터 어레이; 및

상기 변조된 소스를 스펙트럼 범위 감도에 따라 통합하는 포토 센서를 포함하는

스펙트럼 이미징 장치.

청구항 2

제1항에 있어서,

상기 멀티 어퍼처는 대상 물체에 대한 한번의 스캔으로 다수의 부분에 대한 공간 정보를 압축/변조하는

스펙트럼 이미징 장치.

청구항 3

제2항에 있어서,

상기 멀티 어퍼처는 동일한 평면상에 위치하는 제1 렌즈 및 제2 렌즈를 포함하고,

상기 제1 렌즈가 획득하는 대상 물체의 픽셀에 대한 입력 소스와 상기 제2 렌즈가 획득하는 대상 물체의 픽셀에 대한 입력 소스는 일부 중복되는

스펙트럼 이미징 장치.

청구항 4

제1항에 있어서,

상기 랜덤 필터 어레이는 복수의 층으로 구성되고,

상기 복수의 층은 상대적으로 굴절률이 큰 물질을 포함하는 제1 층과 상대적으로 굴절률이 작은 물질을 포함하는 제2 층을 포함하는

스펙트럼 이미징 장치.

청구항 5

제1항에 있어서,

상기 포토 센서는 상기 멀티 어퍼처 및 랜덤 필터 어레이를 통과한 입력 소스는 포토 센서의 파장 효율과 통합하는

스펙트럼 이미징 장치.

청구항 6

제1항에 있어서,

상기 포토 센서에서 통합된 변조 신호를 재건하는 신호 재건부를 더 포함하는

스펙트럼 이미징 장치.

청구항 7

입력 소스를 공간 도메인에서 변조하는 단계;

상기 공간 도메인에서 변조된 입력 소스를 스펙트럼 도메인에서 변조하는 단계; 및

상기 스펙트럼 도메인에서 변조된 소스를 스펙트럼 범위 감도에 따라 통합하는 단계를 포함하는

스펙트럼 이미징 장치의 동작 방법.

청구항 8

제7항에 있어서,

상기 입력 소스를 공간 도메인에서 변조하는 단계는,

대상 물체에 대한 한번의 스캔으로 다수의 부분에 대한 공간 정보를 압축/변조하는 단계를 포함하는

스펙트럼 이미징 장치의 동작 방법.

청구항 9

제7항에 있어서,

상기 스펙트럼 도메인에서 변조된 소스를 스펙트럼 범위 감도에 따라 통합하는 단계는,

상기 멀티 어퍼쳐 및 랜덤 필터 어레이를 통과한 입력 소스는 포토 센서의 파장 효율과 통합하는 단계를 포함하는

스펙트럼 이미징 장치의 동작 방법.

청구항 10

제1항에 있어서,

상기 통합된 변조 신호를 재건하는 단계를 더 포함하는

스펙트럼 이미징 장치의 동작 방법.

청구항 11

제7항 내지 제10항의 알고리즘을 구현하는 프로그램.

발명의 설명

기 술 분 야

[0001] 본 발명은 하이퍼스펙트럼 이미지 장치에 관한 것이다. 상세하게는 디지털 신호 처리 기반 하이퍼스펙트럼 이미지 장치에 관한 것이다.

배 경 기 술

[0002] 스펙트럼 이미징은 물체의 공간 및 스펙트럼 정보를 얻기 위해 기존의 이미징 및 스펙트럼 방법을 통합하는 기술이다. 스펙트럼 이미징은 Goetz에 의해 1980년대 말에 정의되었고 지구를 원격으로 감지하기 위해 논의 되었다. 스펙트럼 이미징은 스펙트럼 해상도, 밴드 수, 너비 및 밴드 연속성에 따라 멀티 스펙트럼 이미징, 하이퍼 스펙트럼 이미징 및 울트라 스펙트럼 이미징으로 나눌 수 있다.

[0003] 멀티 스펙트럼 영상 시스템은 일반적으로 마이크로 미터 또는 수십 마이크로 미터로 측정되는 상대적으로 인접하지 않은 넓은 스펙트럼 대영에서 데이터를 수집한다. 하이퍼 스펙트럼 이미징은 수백 개의 스펙트럼 밴드를 수집할 수 있지만, 울트라 스펙트럼 이미징 시스템은 하이퍼 스펙트럼 이미징 보다 더 많은 스펙트럼 밴드를 수집한다.

[0004] 도 1은 스펙트럼 이미징의 개념을 나타낸다.

[0005] 도 2는 이미징 시스템들의 기능을 나타낸다.

[0006] 도 1에 도시된 바와 같이, 스펙트럼 이미징 데이터는 그 내부 구조로 인해 3차원 큐브 또는 복수의 2차원 이미지의 스택으로 시각화 될 수 있다. 깊이는 파장(wavelength)의 함수이다. 스펙트럼 이미징은 하나의 이미지의

각 픽셀에 대하여 반사율, 흡수 또는 형광 스펙트럼을 얻을 수 있다.

[0007] 스펙트럼 이미징 기술은 항공기 감시, 또는 위성 이미징과 같은 원격 감지 분야에서 적용되던 기술로서, 광업 및 지질학, 농업, 군사, 환경, 글로벌 변화 연구, 법의학 및 의료 분야에서도 적용되는 기술이다.

선행기술문헌

[0008] 선행기술문헌 [1] Li, Qingli, et al. "Review of spectral imaging technology in biomedical engineering: achievements and challenges." Journal of biomedical optics 18.10 (2013): 100901-100901.

[0009] 선행기술문헌 [2] F. H. Goetz et al., "Imaging spectrometry for Earth remote sensing," Science 228(4704), 1147??1153 (1985).

[0010] 선행기술문헌 [3] A. F. H. Goetz, "Three decades of hyperspectral remote sensing of the Earth: a personal view," Remote Sens. Environ. 113, S5??S16 (2009).

[0011] 선행기술문헌 [4] E. J. Cand붕s, "Compressive sampling," Proc. Int. Congr. Mathematicians 3, 1433??1452 (2006).

[0012] 선행기술문헌 [5]Hagen, Nathan, and Michael W. Kudenov. "Review of snapshot spectral imaging technologies." Optical Engineering 52.9 (2013): 090901-090901.

[0013] 선행기술문헌 [6]Cao, Xun, et al. "Computational Snapshot Multispectral Cameras: Toward dynamic capture of the spectral world." IEEE Signal Processing Magazine 33.5 (2016): 95-108.

[0014] 선행기술문헌 [7]A. Wagadarikar, R. John, R. Willett, and D. J. Brady, "Single disperser design for coded aperture snapshot spectral imaging," Appl. Opt., vol. 47, no. 10, pp. B44??B51, 2008.

[0015] 선행기술문헌 [8] M. E. Gehm, R. John, D. J. Brady, R. M. Willett, and T. J. Schulz, "Singleshot compressive spectral imaging with a dual-disperser architecture," Opt. Express, vol. 15, no. 21, pp. 14013??14027, 2007.

발명의 내용

해결하려는 과제

[0016] 본 발명의 일 실시 예에 따른 스펙트럼 이미징 장치는 일반적인 스펙트럼 이미징 장치보다 물리적 구현에 적합하고 소형화가 가능하다.

과제의 해결 수단

[0017] 본 발명의 일 실시 예에 따른 이미징 장치는, 입력 소스를 공간 도메인에서 변조하는 멀티 어퍼쳐, 상기 멀티 어퍼쳐에서 변조된 입력 소스를 스펙트럼 도메인에서 변조하는 랜덤 필터 어레이 및 상기 변조된 소스를 스펙트럼 범위 감도에 따라 통합하는 포토 센서를 포함한다.

발명의 효과

[0018] 본 발명의 일 실시 예에 따른 스펙트럼 이미징 장치는 일반적인 스펙트럼 이미징 장치보다 물리적 구현에 적합하고 소형화가 가능하다.

도면의 간단한 설명

[0019] 도 1은 스펙트럼 이미징의 개념을 나타낸다.

도 2는 이미징 시스템들의 기능을 나타낸다.

도 3은 스펙트럼 이미징(spectral imagin, 스펙트럼 이미징)에서 일반적으로 사용되는 방식을 나타낸다.

도 4는 본 발명의 일 실시 예에 따른 스펙트럼 이미징 장치가 스펙트럼 이미징을 수행하는 과정을 개념적으로 나타낸다.

도 5는 본 발명의 일 실시 예에 다른 스펙트럼 이미징 장치를 나타낸다.

도 6은 본 발명의 일 실시 예에 따른 멀티 어퍼쳐와 일반적인 멀티 어퍼쳐를 비교하는 도면이다.

도 7은 본 발명의 일 실시 예에 따른 스펙트럼 이미징 방법을 나타내는 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0020] 이하에서는 도면을 참조하여 본 발명의 구체적인 실시예를 상세하게 설명한다. 그러나, 본 발명의 사상은 이하에 제시되는 구체적인 실시예로 제한되지 아니하며, 본 발명의 사상을 이해하는 당업자는 동일한 사상의 범위 내에 포함되는 다른 실시예를 구성요소의 부가, 변경, 삭제, 및 추가 등에 의해서 용이하게 제안할 수 있을 것이나, 이 또한 본 발명의 사상에 포함된다고 할 것이다.

[0021] 도 3은 스펙트럼 이미징(spectral imagin, 스펙트럼 이미징)에서 일반적으로 사용되는 방식을 나타낸다.

[0022] 도 3에 도시된 바와 같이, 스펙트럼 이미징은 whiskbroom, pushbroom, staring 및 snapshot의 방식으로 수행될 수 있다.

[0023] 먼저, whiskbroom 방식을 간단히 설명한다.

[0024] Whiskbroom 이미징 모드는 점 스캐닝 방식이라고도 지칭된다. 단일 지점은 샘플 또는 감지기를 움직여 두 개의 공간 차원(x 및 y)를 따라 스캔된다. 그 다음 반사된 광은 프리즘에 의해 분산되고, 선형 어레이 검출기에 의해 기록된다. 스펙트럼 이미지 데이터 큐브는 2차원 장면을 스캐닝하고 파장 영역에서 분산시켜(dispersing) 획득된다. Whiskbroom 모드는 일반적으로 이미지 수집을 완료하는데 2축 모터 위치 테이블이 필요하여 하드웨어 구성이 복잡하다는 단점이 있다.

[0025] 두번째로, pushbroom 방식을 간단히 설명한다.

[0026] Pushbroom 방식은 라인 스캐닝이라고도 지칭된다. 한번에 한 지점을 스캔하는 whiskbroom 방식과 달리, pushbroom 방식은 스캔 시 슬릿(slit)의 각 공간 지점에 해당하는 공간 정보 뿐만 아니라 각 공간에 해당하는 스펙트럼 정보를 동시에 수집할 수 있는 장점이 있다.

[0027] 세번째로, staring 방식을 설명한다.

[0028] Staring 방식은 밴드 시퀀셜 방식이라고도 지칭된다. Staring 방식은 전체 공간 정보가 포함된 단일 대역 2차원 그레이 스케일 이미지를 한번에 획득하는 스팩트럼 스캐닝 방식이다. Staring 방식에서는 전체 공간 정보가 한번에 하나의 스펙트럼 밴드를 통과하여 초점 평면에 이미징되고, 파장에 따라 이미지가 스태킹되어 3D 데이터를 이룬다. 이것은 whiskbroom 및 pushbroom 방식과 다른 방식이다.

[0029] 네번째로, snapshot 방식을 설명한다.

[0030] Snapshot 방식은 싱글 샷(single-shot) 방식이라고도 지칭된다. Snapshot은 공간 및 스펙트럼 정보를 하나의 노출로 영역 검출기에 기록하는 방식이다. Snapshot은 공간 차원 또는 스펙트럼 차원에서 스캔이 이루어 진다. 따라서, 주기 해상도에 한계를 갖는 whiskbroom, pushbroom 및 staring 모드와 달리, 시간 해상도를 제한하기 때문에, snapshot 모드는 스캔이 전혀 필요없는 이미징 방식이다. Snapshot 모드는 리매핑되고 분산된 이미지 존들을 CCD 검출기에 동시에 직접 이미징함으로써, 완전한 스펙트럼 데이터 큐브를 획득할 수 있다.

[0031] 표 1은 각 방식의 특징을 비교/설명한다.

표 1

[0032]

Point of comparison	Whiskbroom	Pushbroom	Staring	Point of comparison
Scanning	Both spatial scan	One of the spatial scan	Spectral scan	No scan
Spectral dispersion element	Dispersive element	Dispersive element	Tunable filter, Interferometer	Diffractive elements
Wavelength rages	Wide	Wide	Mediate	Wide
Spectral resolution	High	High	Mediate	Low
Band number	Hyperspectral	Hyperspectral	Hyperspectral	Multispectral
Data cube collection time	Long	Long	Short	Fast

[0033] 이하에서는 전산 스펙트럼 이미징 방법(computational spectral imaging)을 설명한다.

[0034] 최근에는 압축 센싱(Compressive Sensing) 이론을 이용하여 스펙트럼 이미징을 수행하고 있다. 압축 센싱은 일반적으로 Nyquist 제한으로 인하여 불충분하게 샘플링된 데이터를 재구성하기 위해 L1-표준 재구성 기술을 사용한다. 압축 센싱에 대한 자세한 내용은 선행기술문헌 4 및 5를 참고한다.

[0035] 첫번째로, 단일 분산 코딩 어퍼처 스냅샷 이미징 시스템(Single dispersive coded aperture snapshot imager)를 간략하게 설명한다.

[0036] 선행문헌 7을 참고하면 SD-SASSI라고 지칭되는 단일 분산기를 사용하여 코딩된 어퍼처 스냅샷 이미징 시스템이 제안되어 있다. 스펙트럼 데이터 큐브는 코딩된 마스크 및 분산에 의해 변조된다. 상이한 파장의 광선은 조리개 코드에 의해 변조된 후 분산 소자에 의해 다르게 오프셋 된다.

[0037] 두번째로, 이중 분산 코딩 어퍼처 스냅샷 이미징 시스템(Dual dispersive coded aperture snapshot imager)를 간략하게 설명한다.

[0038] 일반적인 CASSI 시스템은 깎인(sheared) 3차원 스펙트럼 데이터 큐브를 형성하기 때문에, 관찰된 스냅샷은 분산에 의해 흐려진다. 이러한 흐려짐 효과를 극복하기 위해 Gehm은 코딩된 어퍼처의 양 면에 대칭적으로 배치되어 Replicated slanted 코드를 가지는 깎이지 않은(unsheared) 스펙트럼 코드를 생성하는 듀얼 분산기 구조를 제안한다. 이에 관련된 자세한 설명은 선행기술문헌 8을 참고한다.

[0039] 세번째로, 공간-스펙트럼 코딩 압축 스펙트럼 이미징 시스템(Spatial-spectral coded compressive spectral imager)에 대해 간략하게 설명한다.

[0040] 일반적인 CASSI 및 비공간 변조된 이미징 시스템을 위한 실현 가능한 코드는 그들의 물리적 변조 능력에 의해 제한된다. 이론적으로 공간-스펙트럼 데이터 큐브를 무작위로 부호화하는 3-D-CASSI는 일반적인 CASSI 보다 실현 가능한 코드 및 더 높은 성능을 달성할 수 있다.

[0041] 이하에서는 본 발명의 일 실시 예에 따른 스펙트럼 이미지 장치를 설명한다.

[0042] 최근, 많은 분야에서 스펙트럼 이미징에 대한 요구 및 적용이 증가하고 있다. 그러나 일반적인 스펙트럼 이미징 장치는 3차원 데이터 큐브를 생성하는데 많은 시간이 필요하다. 또한, 스펙트럼 이미지 장치의 가격이 고가이다.

[0043] 상기의 단점을 극복하기 위해 제안된 시스템이 CASSI이나, CASSI는 장치의 크기를 최소화하는데 한계가 있다. 또한, CASSI의 물리적인 구현이 어렵다는 문제가 있다. 따라서, 이를 해결하기 위한 본 발명의 일 실시 예를 이하에서 설명한다. 본 발명의 일 실시 예에 따른 스펙트럼 이미지 장치는 물리적 구현에 적합하며 스펙트럼 이미지 장치를 소형화할 수 있다.

[0044] 도 4는 본 발명의 일 실시 예에 따른 스펙트럼 이미징 장치가 스펙트럼 이미징을 수행하는 과정을 개념적으로 나타낸다. 도 5는 본 발명의 일 실시 예에 다른 스펙트럼 이미징 장치를 나타낸다.

[0045] 도 4 및 도 5에 도시된 바와 같이, 본 발명의 일 실시 예에 따른 스펙트럼 이미징 장치(1)는 멀티 어퍼처(10), 랜덤 필터 어레이(20) 및 CMOS 센서(30)을 포함한다.

[0046] 일반적으로 멀티 어퍼처 방식이란, 각각의 어퍼처가 하나의 물체에 대하여 서로 다른 부분을 바라보는 방식이다. 예를 들어 4개의 픽셀로 된 물체는 멀티 어퍼처가 바라본다고 가정하자. 이때, 첫번째 어퍼처는 첫번째 픽셀을 바라보고, 두번째 어퍼처는 두번째 픽셀, 세번째 어퍼처는 세번째 픽셀, 네번째 어퍼처는 네번째 픽셀만을 보여준다. 그리고, 각각의 어퍼처를 통해 획득된 이미지가 모여 전체 상을 구성한다.

[0047] 한편, 본 발명의 일 실시 예에 따른 하나의 어퍼처가 복수의 픽셀에 대한 이미지를 획득하여 공간 정보를 압축하거나, 변조한다. 예를 들어, 본 발명의 멀티 어퍼처는, 첫번째 어퍼처는 1~2 픽셀을 바라고보, 2번째 어퍼처는 2~4 픽셀을 바라고보, 3번째 어퍼서는 3~4 픽셀을 바라보도록 구성될 수 있다. 여기서 바라본다는 기재의 의미는 해당 어퍼처가 바라보는 대상의 픽셀에 관한 입력 소스를 획득함을 의미한다.

[0048] 또 다른 실시 예에서, 본 발명의 일 실시 예에 따른 멀티 어퍼처는 일반적인 멀티 어퍼처 방식을 사용할 수도 있다. 상술한 일반적인 멀티 어퍼처와 본 발명의 일 실시 예에 따른 멀티 어퍼처의 동작을 도 6에서 나타낸다.

[0049] 도 6은 본 발명의 일 실시 예에 따른 멀티 어퍼처와 일반적인 멀티 어퍼처를 비교하는 도면이다.

[0050] (a)는 일반적인 멀티 어퍼처의 동작을 나타내고, (b)는 본 발명의 일 실시 예에 따른 멀티 어퍼처의 동작을 나타낸다. (b)에서 도시하는 바와 같이, 모든 렌즈가 각각 다른 픽셀에 대한 입력 소스를 획득하는 것이 아니라, 멀티 어퍼처를 구성하는 렌즈들이 대상 물체의 픽셀에 대한 입력 소스를 중복하여 획득한다.

[0051] 다시 도 4 내지 도 5로 돌아온다.

[0052] 멀티 어퍼처(10)는 입력 신호를 수신하여 이를 변조/압축 한다. 구체적으로 멀티 어퍼처(10)는 단일 스냅 샷(single snapshot)에서의 장면의 공간 정보를 압축하거나, 변조한다. 본 발명의 일 실시 예에 따른 스펙트럼 이미지 장치(1)는 단일의 어퍼처를 포함하는 것이 아닌 멀티 어퍼처(10)를 포함함으로써, 한번의 스캔으로 다수의 지점에 대한 장면의 공간 정보를 압축/변조할 수 있다. 도 5를 참조하면, 마이크로 렌즈가 도시되어 있으며, 마이크로 렌즈의 집합을 멀티 어퍼처라고 할 수 있다.

[0053] 랜덤 필터 어레이(20)는 멀티 어퍼처(10)에서 압축/변조된 입력 신호의 스펙트럼 정보를 변조한다. 이때, 일 실시 예에서 랜덤 필터 어레이(20)는 기판상에 증착된 고 굴절률 물질 및 저 굴절률 물질의 다중 층으로 구성된 박막 광학 필터를 포함할 수 있다. 도 5를 참조하면 분광 필터부가 도시되어 있으며, 분광 필터부의 집합을 랜덤 필터 어레이라고 할 수 있다.

[0054] CMOS 센서(30)는 랜덤 필터 어레이(20)에서 변조된 신호를 센싱한다. 다시 말해서 렌덤 필터 어레이(20)에서 변조된 입력 신호가 CMOS 센서(30)에 프로젝션된다. 도 5를 참조하면 포토 센서부가 도시되어 있으며, 포토 센서부의 집합을 CMOS 센서라고 할 수 있다.

[0055] 각각의 랜덤 필터를 통과한 입력 신호는 파장에 대하여 급격하게 변동한다. 따라서, 본 발명의 일 실시 예에 따른 스펙트럼 이미징 장치는 변조기를 사용하지 않고도 컴팩트한 설계가 가능하여 일반적인 스펙트럼 이미징 장치에 비해 변조가 용이하다. 도 7은 본 발명의 일 실시 예에 따른 스펙트럼 이미징 방법을 나타내는 흐름도이다.

[0056] 본 발명의 일 실시 예에 따른 스펙트럼 이미징 방법은, 먼저 입력 소스가 멀티 어퍼처를 통해 공간 도메인에서 변조된다(S101).

[0057] 공간 변조(spatial modulation)에서, 단일 어퍼처 및 입력 신호(X)와의 관계는 수학식 1과 같이 표현된다.

수학식 1

[0058]

[0059] 여기에서, d_i는 공간적으로 변조된 출력이며, a_i는 단일 어퍼처, X는 하이퍼스펙트럼 이미지 입력을 나타낸다.

[0060] 그리고 공간 도메인에서 변조된 입력 소스가 랜덤 필터 어레이를 통해 스펙트럼 도메인에서 변조된다(S103).

[0061] 스펙트럼 변조(spectral modulation)에서, 랜덤 필터 어레이(20)를 통해 변조된 출력 d_i는 수학식 2와 같이 표현된다.

수학식 2

[0062] $$y_i = Td_i$$

[0063] 여기에서, y_i는 i 번째 렌즈에서 랜덤 필터 어레이의 출력 샘플이다. T는 랜덤 필터의 전송 셋이다. i 번째 렌즈란, 멀티 어퍼처에서 i 번째 어퍼처를 구성하는 렌즈를 말한다.

[0064] 과 같이 행렬 Y가 표기될 때, 압축 스펙트럴 이미징 장치는 수학식 3과 같이 모델링 된다.

수학식 3

[0065] $$Y = TXA^T$$

[0066] 여기에서 는 입력 소스이고, A는 멀티 어퍼처 매트릭스이고, T는 필터 어레이의 전송 매트릭스이다. X는 입력소스를 나타내는 N ㅅ P매트릭스이다. N은 파장의 범위이고, P는 픽셀의 개수이다. A는 공간 변조 현상을 설명하는 L ㅅ P매트릭스 이다. L은 어퍼처의 개수이고, P는 픽셀의 개수이다. i번째 어퍼처에 대하여 j번째 픽셀이 보이지 않으며, 행렬 A내 컴포넌트 a_{ij}가 0이다. j번째 픽셀이 i 번째 어퍼처에 대해 모두 보이면 행렬 A내 컴포넌트 a_{ij}가 1이다. 그렇지 않으면 a_{ij}는 0과 1사이에 있다. T는 M ㅅ N의 랜덤 필터 어레이의 전송 매트릭스이다. M은 필터의 개수이고, N은 파장의 범위이다. 전송 매트릭스 T는 수학식 4와 같이 표현된다.

수학식 4

[0067]

[0068] 결론적으로 Y의 차원은 M ㅅ L이다.

[0069] 변조된 소스를 CMOS 센서(30)의 스펙트럼 범위 감도를 따라 통합한다(S105). 구체적으로 멀티 어퍼처를 통과하고 랜덤 필터를 통과한 광원이 CMOS 센서의 파장 효율과 통합(integral)된다.

[0070] 센싱된 변조 신호가 신호 재건부(미도시)에서 재건된다(S107).

[0071] 이하에서는 본 발명의 일 실시 예에 따라 변조된 소스를 재건(reconstruction)하는 과정을 설명한다.

[0072] 압축 센싱 이론은 센싱 매트릭스가 미확인된 경우 인버스 문제를 해결하기 위해 이용된다. 이는 수학식 5 및 수학식 6과 같이 표현된다.

수학식 5

[0073] $$D = XA^T$$

수학식 6

[0074] $$Y = TD$$

[0075] 여기에서 는 N ㅅ L매트릭스이다.

[0076] 매트릭스 T에서, 행의 크기(M)는 열의 크기(N)보다 작다. 수학식 6을 열방향으로 표현하면 수학식 7과 같다.

수학식 7

[0077] $$y = Td$$

[0078] 여기에서 d는 베이시스 Ψ 및 스파스(sparse) 벡터를 이용하여 수학식 8과 같이 표현된다.

수학식 8

[0079] $$d = \Psi s$$

[0080] 그리고 수학식 8을 이용하여 수학식 7은 수학식 9과 같이 다시 표현된다.

수학식 9

[0081] $$y = T\Psi s$$

[0082] 압축 센싱 재건 알고리즘을 통해 하나의 어퍼처의 공간 정보의 스파스 근사를 찾을 수 있다. 이는 최적화 문제로 수학식 10과 같이 표현된다.

수학식 10

[0083]

[0084] 그리고, $\hat{d} = \Psi\hat{s}$를 추정할 수 있다. 그리고, D를 라고 하면, 수학식 5는 수학식 11로 다시 표현된다.

수학식 11

[0085] $$\hat{D} = XA^{T}$$

[0086] 그리고 수학식 11의 양변을 트랜스포즈하면 수학식 12와 같이 표현된다.

수학식 12

[0087]

[0088] 매트릭스 A에 대하여 행의 크기(L)은 열을 크기(P)보다 작다. 따라서, 수학식 12를 열방향으로 수학식 13과 같이 표현할 수 있다.

수학식 13

[0089] $$\hat{d}^T = Ax^T$$

[0090] 그리고 여기에 최적화 문제를 적용하면 수학식 14와 같이 표현된다.

수학식 14

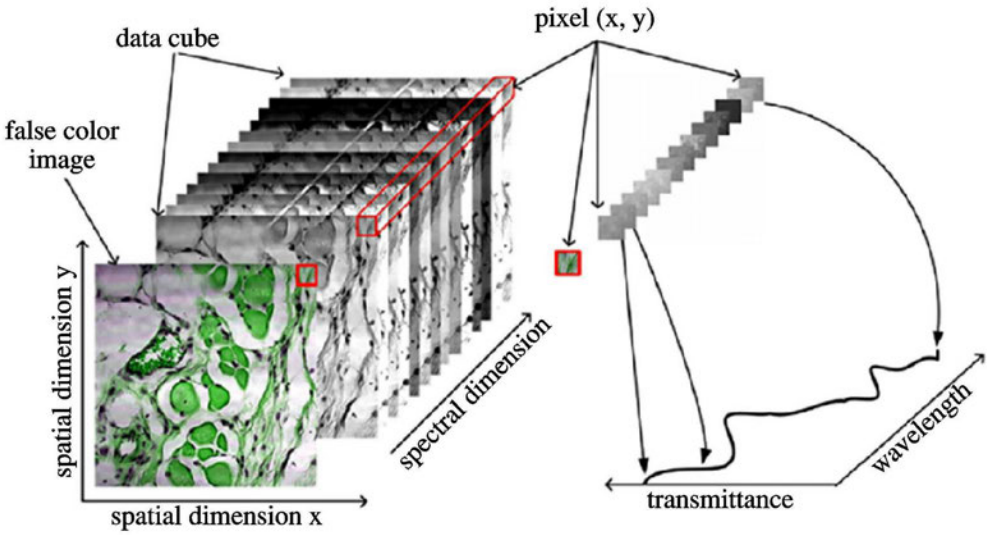
[0091]

[0092] 그리고, 열방향으로 를 트랜스포즈하면 변조전 최초 입력 장면인 $\hat{x}$를 재건할 수 있다.

[0093] 지금까지, 본 발명의 이해를 돕기 위하여 스펙트럼 이미지 장치에 대한 예시적인 실시 예가 설명되고 첨부된 도면에 도시되었다. 그러나 이러한 실시 예는 단지 본 발명을 예시 하기 위한 것이고, 이를 제한하지 않는다는 점이 이해되어야 할 것이다. 그리고 본 발명을 도시되고 설명된 내용에 국한되지 않는다는 점이 이해되어야 할 것이다. 이는 다양한 변형이 본 기술 분야에서 통상의 지식을 가진 자에게 일어날 수 있기 때문이다.

도면

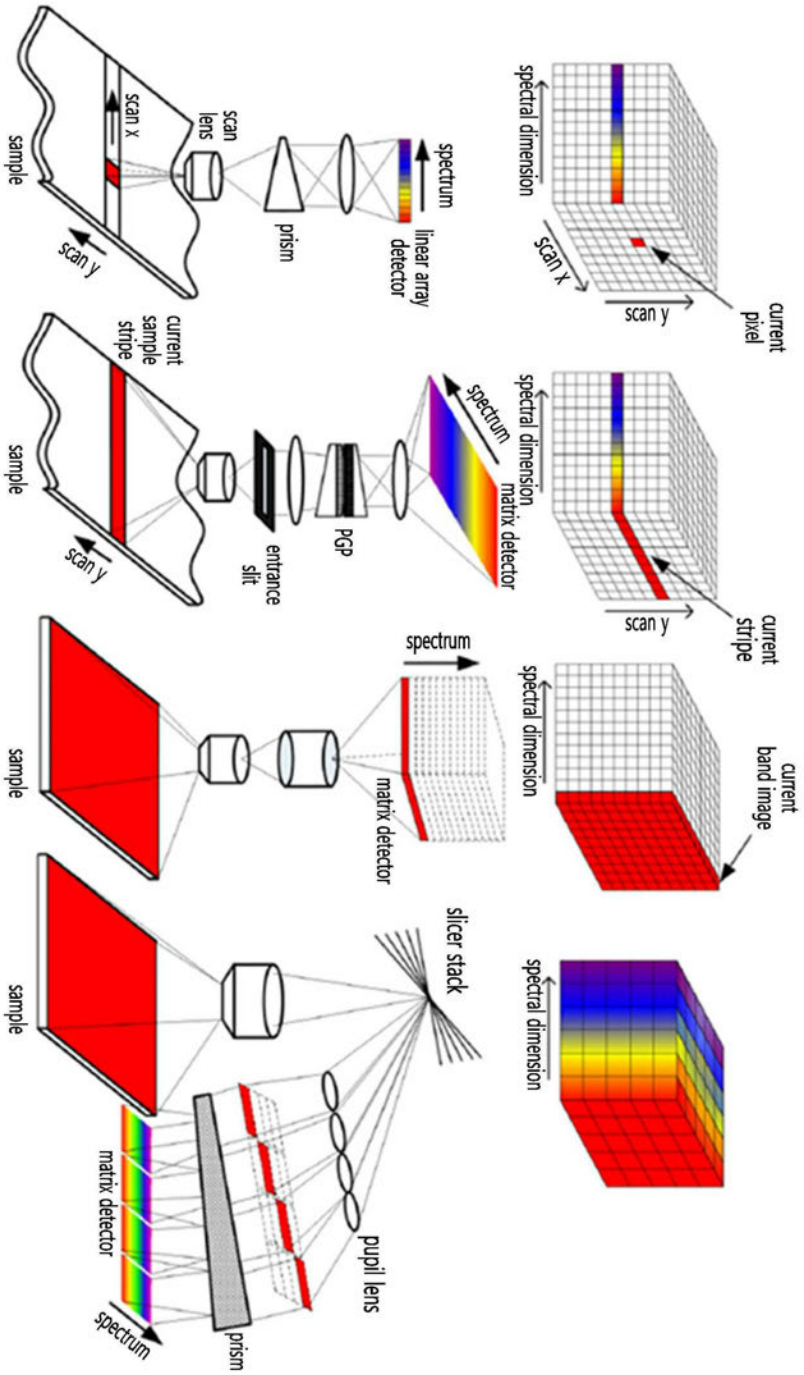
도면1



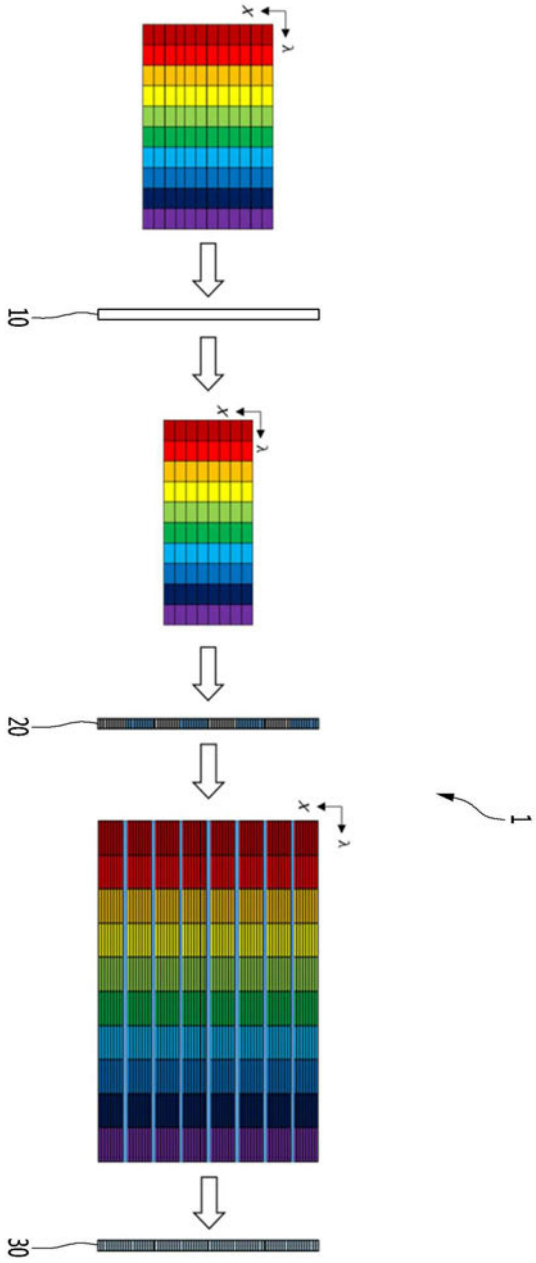
도면2

Feature	Monochrome	RGB	Spectroscopy	Multispectral	Hyperspectral
Spatial information	Yes	Yes	No	Yes	Yes
Band numbers	1	3	From several dozens to hundreds	3 to 10	From several dozens to hundreds
Spectral information	No	No	Yes	Limited	Yes
Multiconstituent information	No	Limited	Yes	Limited	Yes
Sensitivity to minor components	No	No	No	Limited	Yes

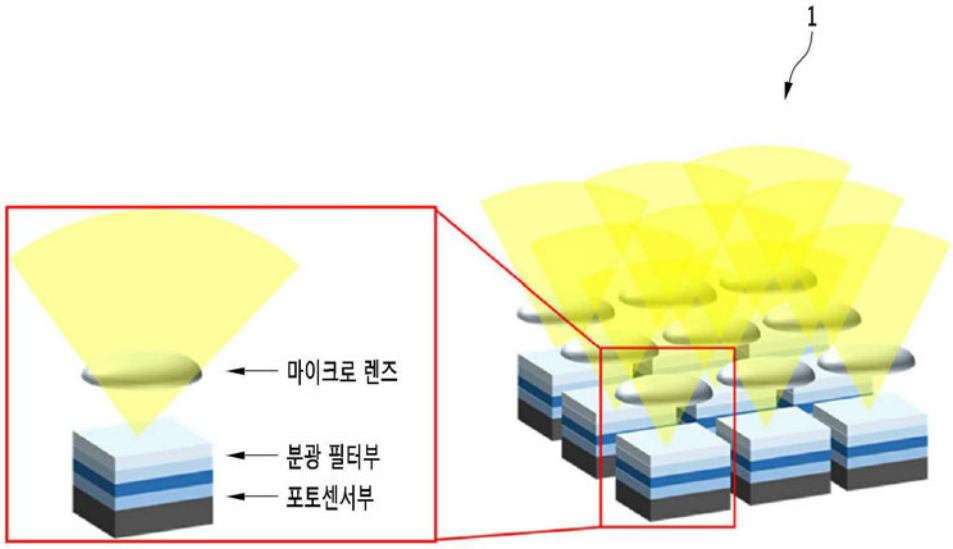
도면3


sample
scan x
scan y
scan lens
prism
spectrum
linear array detector
spectral dimension
scan x
scan y
current pixel
current sample stripe
entrance slit
PGP
matrix detector
current stripe
current band image
slicer stack
pupil lens
prism

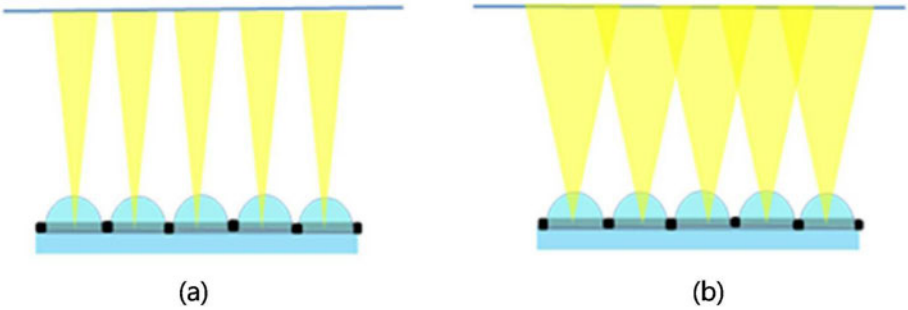
도면4



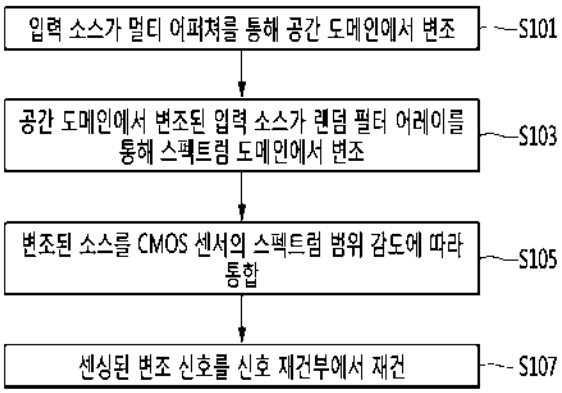
도면5


1
마이크로 렌즈
분광 필터부
포토센서부

도면6


(a)
(b)

도면7


입력 소스가 멀티 어퍼쳐를 통해 공간 도메인에서 변조
S101
공간 도메인에서 변조된 입력 소스가 랜덤 필터 어레이를 통해 스펙트럼 도메인에서 변조
S103
변조된 소스를 CMOS 센서의 스펙트럼 범위 감도에 따라 통합
S105
센싱된 변조 신호를 신호 재건부에서 재건
S107

영지식 센싱 초분광 생체인증

장재혁, 이흥노
GIST

발명의 배경

생체인증 시스템 정의

- 생체인증 시스템
  - 생체인증 시스템은 생체정보를 사용하여 사용자의 신원을 인증하는 시스템임
- 일반적인 생체인증 시스템은 적어도 다음의 요소를 포함함
  - 아날로그 상태의 생체정보를 디지털로 변환 및 저장하는 아날로그-디지털 스캐너
  - 신원이 등록된 사용자들의 생체정보가 저장되어있는 데이터베이스
  - 신원 인증을 요청하는 생체정보의 등록유무를 판별하는 신호 분류기
  - 사용자와 검증자간의 통신 프로토콜

생체인증 시스템의 문제와 종래 기술

- 생체정보의 유출 가능성
  - 생체정보를 디지털 데이터로 변환하여 사용하기 때문에 외부로 유출될 가능성이 존재함
  - 생체정보는 변경이 불가능하므로 한번 악용될 경우 피해가 지속됨
- 종래 기술 동향
  - 이미 디지털화 된 생체정보를 숨기기(hiding) 위해, 해시함수, 공개키 암호화, 비가역적 변형함수, 혹은 동형암호화와 같은 암호화 알고리즘을 사용하였음
  - 공통적으로, 종래 기술들은 디지털 생체 데이터를 데이터베이스에 등록하는 과정 혹은 사용자-검증자간의 통신 과정에서의 정보 유출을 방지하였음

문제 정의

- 생체정보 보호
  - 종래 기술은 이미 디지털화된 생체정보를 숨김
  - 그러나 디지털 생체정보 데이터의 존재 자체가 데이터 유출의 가능성을 내포함
  - 따라서 원본 생체 디지털 데이터를 사용하지 않는 개인인증 시스템이 필요함

- 발명기술의 목표
  - 영지식증명과 압축센싱을 활용한 새로운 생체인증 시스템 개발
  - 인증에 사용되는 인증 데이터로부터 원본 생체정보를 복원 할 수 없어야 함
  - 등록된 사용자는 인증 데이터를 사용하여 검증자를 납득시킬 수 있어야 함
  - 등록되지 않은 사용자의 허위 인증 데이터는 검증자를 납득시킬 수 없어야 함
  - 한번 사용된 인증 데이터를 악의적인 사용자가 재사용 할 수 없어야 함

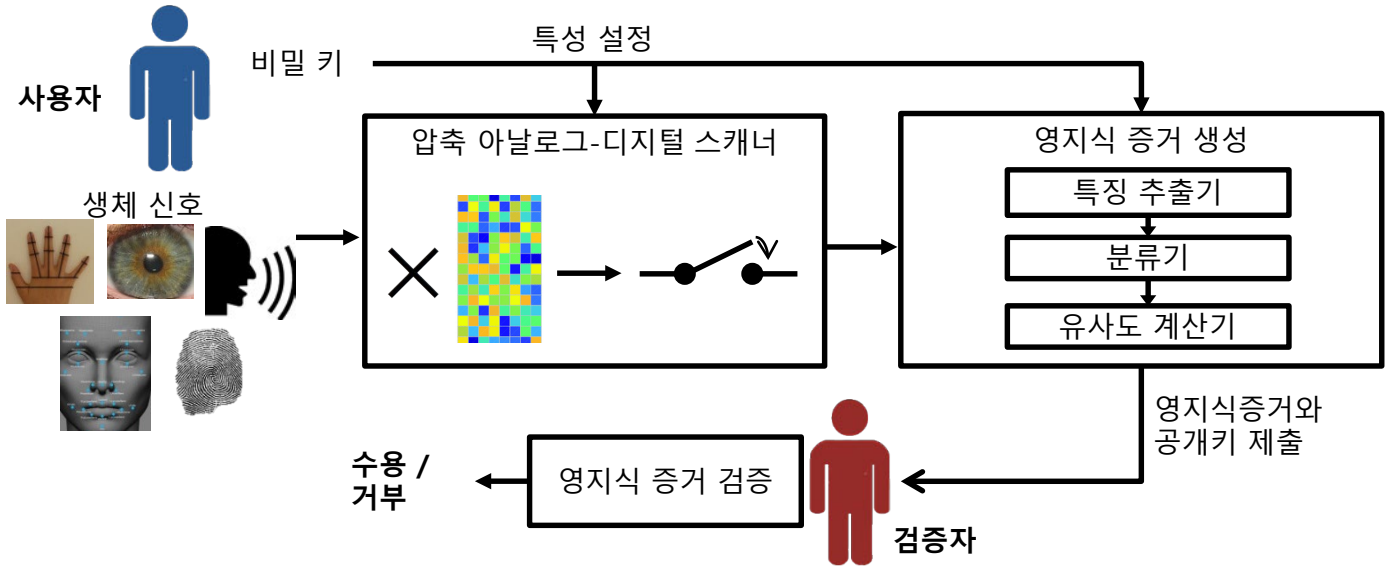
발명: 영지식 센싱 초분광 생체인증 시스템

발명 구성

- 영지식 센싱 생체인증 시스템
  - 생체정보 유출 문제를 해결하는 새로운 생체인증 시스템
  - 압축센싱을 활용하여 생체정보를 숨김
  - 영지식증명을 활용하여 사용자가 검증자를 납득시킴

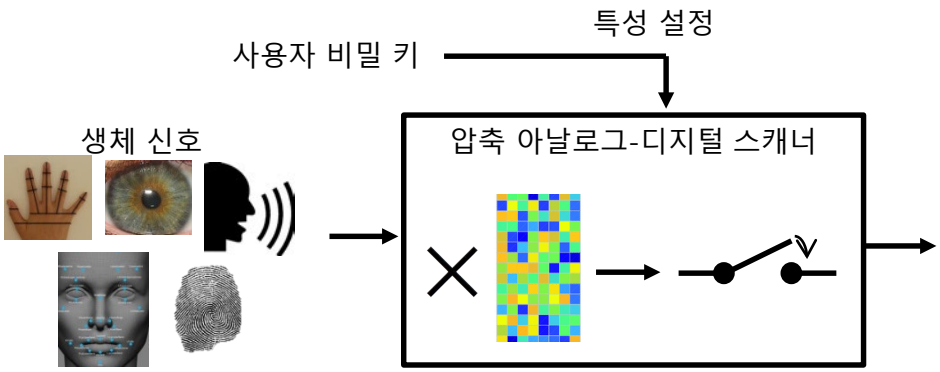
- 초분광 생체정보 기반 DID 기술
  - 초분광 이미지에 대한 딥러닝 기반 신호 분류기
  - 초분광 이미지를 생체인증 수단으로 사용

영지식 센싱 생체인증 시스템



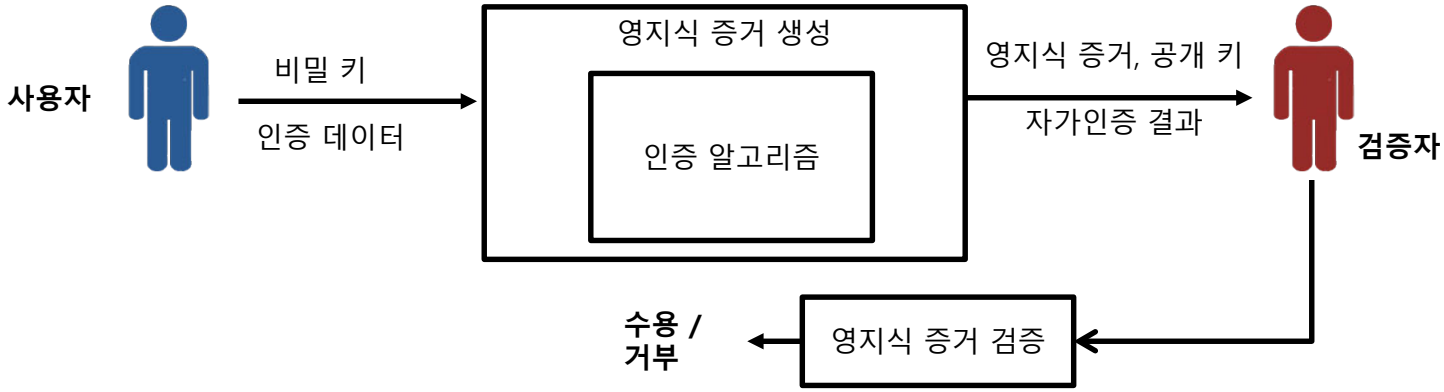
- 시스템 구성
  - 압축 아날로그-디지털 스캐너
  - 인증 알고리즘 (특징 추출기, 분류기, 유사도 계산기)
  - 영지식증명 프로토콜 (영지식 증거 생성 및 검증 알고리즘)
  - 비밀-공개키

압축 아날로그-디지털 스캐너



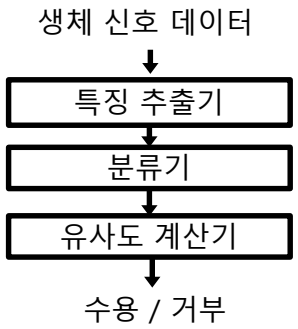
- 압축 아날로그-디지털 스캐너의 특징
  - 특수제작된 하드웨어를 사용하여 아날로그 생체신호를 압축한 후 디지털화함 (생체신호 → 인증 데이터)
  - 비가역적인 전달함수 (인증 데이터 → 생체신호 복원 X)
  - 신호분류에 필요한 최소한의 특징정보만을 보존함
  - 물리적 전달함수를 비밀키를 사용해 설정 가능
  - 사용자 생체정보의 유출을 원천 차단함

영지식증명 프로토콜



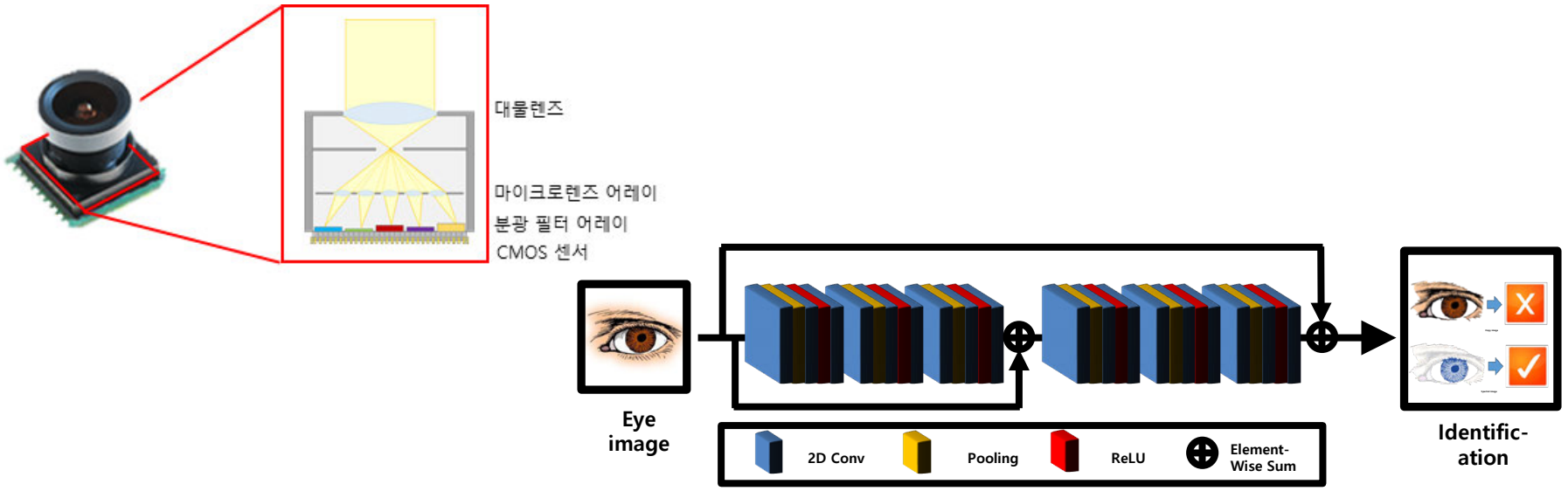
- 영지식증명 프로토콜의 특징
  - 사용자가 스스로 인증 알고리즘을 실행 후 결과를 얻음
  - 공인된 인증 알고리즘을 올바르게 실행하였다는 증거를 함께 생성함
  - 검증자는 증거와 일부 공개된 데이터만을 사용하여 사용자의 자가검증의 유효성을 확인함
  - 사용자 인증 데이터의 유출을 원천 차단함

인증 알고리즘



- 인증 알고리즘
  - 딥러닝을 활용한 알고리즘을 사용
  - 생체 신호 데이터에 따라 인증 알고리즘이 결정됨
  - 생체신호는 신체부위의 이미지, 음성, 초분광 이미지 등이 될 수 있음
  - 우리의 발명에서는 초분광 이미지에 대한 인증 알고리즘을 제안함

초분광 생체정보 기반 DID 기술



- 초분광 생체정보 기반 DID 기술의 구성
  - 초분광 압축 이미지-디지털 스캐너
    - 독자적인 랜덤 압축 센싱 필터 사용
  - 딥 뉴럴 네트워크기반 특징정보 추출 및 분류 알고리즘

발명의 효과

발명의 효과

- 생체정보의 유출이 원천 차단된 생체인증 시스템 개발
  - 생체정보 유출 문제에 의해 저지된 생체인증 시스템 시장의 성장을 가속화 할 것으로 기대
- 생체정보 유출 2단계 차단
  - 압축 아날로그-디지털 스캐너에 의한 생체정보 유출 1차 차단
  - 영지식증명 프로토콜에 의한 인증 데이터 유출 2차 차단
- 초분광 이미지를 생체정보로 활용하는 인증시스템
  - 초분광 이미지의 방대한 정보량을 활용
  - 높은 정확도의 분류성능 제공

감사합니다

Gwangju Institute of
Science and Technology

School of Information and Communications



WorldLand™, My AI Network™

핵심 기술 연구 결과

2024. 01. 08.

Department : EECS

Professor : Heung-No Lee

Research Team : INFONET, LiberVance

MY AI Network Project Progress Report

LiberVance Co. Ltd.
2024. 1. 9.(Tue)

□ Project Overview

○ Project Title: MY AI Network

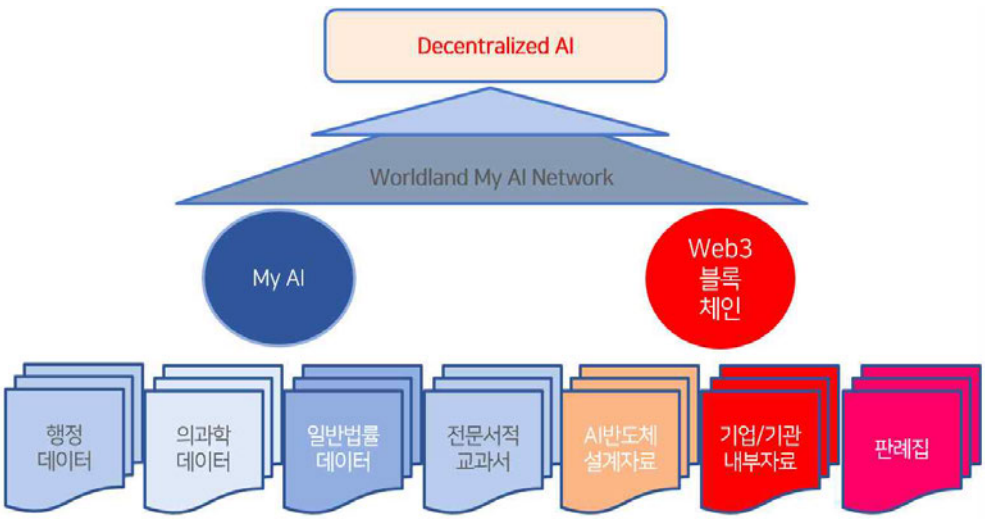
○ Goal: To address the challenges in AI development, which include securing high-quality training data and resolving collaborative learning issues*, this project aims to improve data value and model performance through mutual cooperation while mitigating issues. This project is implemented based on decentralized AI, where individuals own their AI models**.

* **Chat GPT model, which are representative of centralized AI, can cause several issues coming from their centralized nature, leading to potential social problems.**
(i.e. bias, inaccuracy, security, and difficulties in protecting copyright rights)

** **Through the application of Web3 blockchain, we can expect enhanced transparency, security, and trustworthiness, as well as the digitization and economic improvement of digital content creation through smart contracts and tokens. Furthermore, optimization of AI utilization governance processes can be expected within DAOs.**

□ Service Overview

○ Providing Decentralized AI*** services that encourage mutual cooperation for individual AI owners to secure data sovereignty, productivity, and competitiveness.



*** Decentralized AI: Ensuring distributed ownership and control of AI models, individual data protection and safety, autonomous collaboration and interaction, and automated transactions with transparency in governance through smart contracts.

□ Current Project Progress

○ (Technological Differentiation) Secured papers/patents on applications, specialization, and foundational technologies**** (such as **generative AI training, hyperspectral identity authentication, Verifiable Coin Toss, ECCPoW, DeSecure blockchain, zero-knowledge proofs, quantum-resistant cryptography, etc.**) to ensure expertise and technological differentiation for service implementation.

○ **Launch of Web3 network (WorldLand)** capable of implementing the service.

- Selected for the Ministry of SMEs and Startups' Technology Innovation Development Program (TIPS) and ongoing technology development ('22.7 - '24.6).
- Global mainnet launch on August 8, 2023 (current average block creation time 10.4 seconds, 1.23 million generated blocks, 673 connected wallets, 143 daily nodes).
- Mainnet launch event on August 10, 2023, attended by over 60 relevant experts from academia, industry, and government (including speeches by Congressman Yang Hyang-ja, Chairman of the Korea ICT Convergence Association, DSRV CEO, Blockchain Factory CEO, Blockchain Today CEO, etc.).

**** Secured 2 domestic patents, with 3 patent technology transfers planned (including knowledge SNARK proof system and method), 300+ domestic and international research papers from INFONET Lab, 9 international patent applications with 7 registrations, 7 domestic patent applications with 22 registrations, and 3 technology transfers.

□ Future Works

○ (DAO) Performing the role of the WorldLand Ecosystem Growth Foundation through WorldLand DAO, which makes important decisions through stake proof.

○ (WorldLand Eco) Providing decentralized services to various ecosystem players, including DAO foundations, investors, AI network developers, node operators, users, dApp operators, etc., through a decentralized ecosystem.

WorldLand™ My AI Network™ 기술 연구 결과

작성일 : 20240108

작성자 : 강주성

목적:

- 연구실 My AI Network 및 WorldLand 관련 논문 목록 조사 및 수합

핵심 기술 키워드 및 연구 결과 (논문) :

1. **WorldLand™**

   i. **Heung-No Lee, Young-Sik Kim, Dilbag Singh, and Manjit Kaur**, "Green Bitcoin: Global Sound Money," The Journal of Digital Assets," Vol.1, No. 1, pg. 33-47, October, 2022. (The Inaugural Issue of the Journal of Digital Assets, No Impact Factor) (PDF)

   ii. **Hyoungsung Kim, Jehyuk Jang, Sangjun Park, and Heung-No Lee*,** "Error-Correction Code Proof-of-Work on Ethereum"**,** IEEE Access, Vol. 9, pp. 135942-135952, Sep 2021. doi: https://doi.org/10.1109/ACCESS.2021.3113522 (Impact Factor: 3.367, IITP & Do-Yak project) Paper: (Open Access)

   iii. **Sangjun Park, Nam Yul Yu and Heung-No Lee,** "An Information-Theoretic Study for Joint Sparsity Pattern Recovery with Different Sensing Matrices", IEEE Transactions on Information Theory, Vol. 63, no. 9, pp. 5559-5571, May. 2017. (Impact Factor: 2.679, Do-Yak Project) (PDF) DOI: 10.1109/TIT.2017.2704111

   iv. **Sangjun Park, Haeung Choi and Heung-No Lee,** "Time-Variant Proof-of-Work Using Error-Correction Codes," preprint on arXiv, June 2020, https://arxiv.org/abs/2006.12306 .

   v. **Jehyuk Jang and Heung-No Lee,** "Profitable Double-Spending Attacks," Applied Sciences, 10, 8477, .Nov. 2020. doi: https://doi.org/10.3390/app10238477. (IF: 2.474, Do-Yak and IITP) Paper: (Open access)

   vi. **Hyunjun Jung, Heung-No Lee*,** "ECCPoW: Error-Correction Code Based Proof-of-Work for ASIC Resistance", *Symmetry*, June. 2020, 12(6), 988 (impact Factor: 2.143, Do-Yak project) (PDF)

   vii. **Seungmin Kim, Gyeongdeok Maeng, Heung-No Lee,** "Smart Contract-Based Checkpoint for Initial PoW Network Security," 2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN 2023), July 2023, pp.89-93.

   viii. **맹경덕, 김승민, 이흥노, "지속 가능한 경제와 암호화폐," 한국핀테크학회 2023 동계학술대회, Feb. 2023.**

2. My AI Network™

i. **Jian Su; Zhenlong Liao, Zhengguo Sheng, Alex X. Liu, Dilbag Singh, and Heung-No Lee,** "Human Activity Recognition Using Self-powered Sensors Based on Multilayer Bi-directional Long Short-Term Memory Networks" *IEEE Sensors Journal*, (Impact factor: 4.325, Category Instruments & Instrumentation, JCR Quartile: Q1: JIF Percentile: 78.91, Do-Yak project). *(Link)*

ii. **Kiwon Yang, Jusung Kang, Jehyuk Jang and Heung-No Lee**, "Multimodal Sparse Representation-Based Classification Scheme for RF Fingerprinting," *IEEE Communications Letters, Vol. 23, Issue 5, pp. 867-870, May 2019. (Impact Factor: 2.723, Doyak Project) (PDF) (Used DB) (Source Code)*

iii. **Manjit Kaur, Dilbag Singh, Vijay Kumar, Heung-No Lee**, "MLNet: Metaheuristics-based Lightweight Deep Learning Network for Cervical Cancer Diagnosis", IEEE Journal of Biomedical and Health Informatics, (Impact factor: 7.021, Category Mathematical & Computational Biology, JCR Quartile: Q1: JIF Percentile: 93.86). *(Link)*

iv. **Dilbag Singh, Manjit Kaur, Jazem Mutared Alanazi, Ahmad Ali AlZubi, Heung-No Lee**, "Efficient evolving deep ensemble medical image captioning model", IEEE Journal of Biomedical and Health Informatics, (Impact factor: 7.021, Category Mathematical & Computational Biology, JCR Quartile: Q1: JIF Percentile: 93.86). *(Link)*

v. **Dilbag Singh, Manjit Kaur, Vijay Kumar, Mohamed Yaseen Jabarulla, and Heung-No Lee**, "Artificial Intelligence-based Cyber-Physical System for Severity Classification of Chikungunya Disease", IEEE Journal of Translational Engineering in Health and Medicine, (Impact factor: 3.316, Category: Engineering, Biomedical, JCR Quartile: Q2: JCR Percentile: 51.12). *(Open Access)*

vi. **Dilbag Singh, Aravinda C.V, Manjit Kaur, Meng Lin, Jyoti Shety, Vikram Raju Reddicherla, and Heung-No Lee,** "DKNet: Deep Kuzushiji Characters Recognition Network" IEEE Access, (Impact factor: 3.476, Category: Engineering, Electrical & Electronic, JCR Quartile: Q2: JCI Percentile: 62.14, IITP & Do-Yak project). (Open Access)

vii. **Jin-Taek Seong and Heung-No Lee,** "Predicting the Performance of Cooperative Wireless Networking Schemes with Random Network Coding," IEEE Transactions on Communications, vol. 62, no. 8, pp. 2951-2964, Aug. 2014. (Imfact Factor: 1.979, Do-Yak Project) (pdf)

viii. **Sang-Seon Byun, Ilangko Balasingham, Athanasios V. Vasilakos, and Heung-No Lee** "Computation of an Equilibrium in Spectrum Markets for Cognitive Radio Networks," to appear IEEE Transactions on Computers, Published on-line first on Aug. 30th, 2012. (Impact Factor: 1.473, Do-Yak Project, Haek-Sim Project)(pdf)

ix. **Cheng-Chun Chang, Zhi-Hong Mao, and Heung-No Lee,** "MB iterative decoding algorithm on systematic LDGM codes: Performance evaluation," Signal Processing, vol. 90, issue. 1, pp. 373-377, 2010.(Impact Factor: 2.238) (pdf)

x. **Rahman S M Wahidur, Ishmam Tashdeed, Manjit Kaur, and Heung-No Lee,** "Enhancing Zero-Shot Crypto Sentiment with Fine-tuned Language Model and Prompt Engineering", in IEEE Access, Early Access, doi:

10.1109/ACCESS.2024.3350638. (URL)

3. **Hyper Spectrometer Imaging (HSI) system.**

   i. **Cheolsun Kim, Pavel Ni, Kang Ryeol Lee, and Heung-No Lee***, "Mass production-enabled computational spectrometers based on multilayer thin films", Accepted in Scientific Reports, (Impact Factor: 4.380, Category: Multidisciplinary Sciences, JCI Quartile: Q1: JCI Percentile: 85.55, Do-Yak project).

   ii. **Cheolsun Kim, Dongju Park, and Heung-No Lee***, "Compressive sensing spectroscopy using a residual convolutional neural network", *MDPI Sensors*, Vol. 20(3), Jan. 2020. (Impact Factor: 3.031, Do-Yak project) (PDF)

   iii. **Cheolsun Kim. Woong-Bi Lee, Soo Kyung Lee, Yong Tak Lee, and Heung-No Lee***, "Fabrication of 2D thin-film filter-array for compressive sensing spectroscopy", *Optics and Lasers in Engineering*, Vol. 115, pp. 53-58, Apr. 2019. (PDF) (Impact Factor: 3.388 Doyak project)

   iv. **J. Oliver, WoongBi Lee, and Heung-No Lee, "**Filters with random transmittance for improving resolution in filter-array-based spectrometers**,"** Optics Express, Vol. 21, No. 4, pp. 3969–3989, Feb. 2013. (Impact Factor: 3.587; Do-Yak Project). (pdf)

기술 연구 결과 리스트:

[기술 관련 연구 결과, 핵심 기술 연구 결과]

(AI + Blockchain)

1. **Mohamed Yaseen Jabarulla and Heung-No Lee***, "Blockchain and Artificial Intelligence-based Patient-Centric Healthcare System for Combating COVID-19 Pandemic: Opportunities and Applications," *Healthcare,* 2021, 9(8), 1019. doi: https://doi.org/10.3390/healthcare9081019 (Impact Factor: 2.645, Do-Yak project). Paper: (Journal Page) (Cover Story) (Cover Image).

2. **Mohamed Yaseen Jabarulla and Heung-No Lee***, "Blockchain-Based Distributed Patient-Centric Image Management System", *Applied Sciences, 11(1)*, 196, Dec. 2020. doi: https://doi.org/10.3390/app11010196. (IF: 2.679, Do-Yak and IITP) Paper: (Journal Page). PCAC-SC Solidity Code: https://github.com/infonetGIST/PCAC-SC

(Green Bitcoin)

1. **Heung-No Lee, Young-Sik Kim, Dilbag Singh, and Manjit Kaur**, "Green Bitcoin: Global Sound Money," The Journal of Digital Assets," Vol.1, No. 1, pg. 33-47, October, 2022. (The Inaugural Issue of the Journal of Digital Assets, No Impact Factor) (PDF)

(ECCPoW)

1. **Hyoungsung Kim, Jehyuk Jang, Sangjun Park, and Heung-No Lee*,** "Error-Correction Code Proof-of-Work on Ethereum", IEEE Access, Vol. 9, pp. 135942-135952, Sep 2021. doi: https://doi.org/10.1109/ACCESS.2021.3113522 (Impact Factor: 3.367, IITP & Do-Yak project) Paper: (Open Access)

2. **Haeung Choi, Sangjun Park and Heung-No Lee**, "Covert Anti-Jamming Communication Based on Gaussian Coded Modulation", *Applied Sciences, 11(9),* April 2021, doi: https://doi.org/10.3390/app11093759 (EW33 project), (Impact factor: 2.679, Category: Engineering, Multidisciplinary, JCI Quartile: Q2, JCI Percentile: 67.94) Paper: (Open Access)

3. **Jehyuk Jang and Heung-No Lee,** "Profitable Double-Spending Attacks," *Applied Sciences, 10*, 8477, .Nov. 2020. doi: https://doi.org/10.3390/app10238477. (IF: 2.474, Do-Yak and IITP) Paper: (Open access)

4. **Hyunjun Jung, Heung-No Lee*,** "ECCPoW: Error-Correction Code Based Proof-of-Work for ASIC Resistance", *Symmetry*, June. 2020, 12(6), 988 (impact Factor: 2.143, Do-Yak project) (PDF)

5. **Sangjun Park, Nam Yul Yu, Heung-No Lee,** "An Information-Theoretic Study for Joint Sparsity Pattern Recovery with Different Sensing Matrices", IEEE Transactions on Information Theory, Vol. 63, no. 9, pp. 5559-5571, May. 2017. (Impact Factor: 2.679, Do-Yak Project) (PDF) DOI: 10.1109/TIT.2017.2704111

6. **J. Zhang and Heung-No Lee**, "Performance Analysis on LDPC-Coded System over Quasi-Static (MIMO) Fading Channels," *IEEE Transactions on Communications,* vol. 56, no. 12, pp. 2080-2093, Dec. 2008 . (pdf)

7. **J. Wu and Heung-No Lee**, "Performance Analysis for LDPC-Coded Modulation in MIMO Multi-Access Systems," *IEEE Transactions on Communications*, vol. 55, no. 7, pp.1417-1426, Jul, 2007. (pdf)

8. **J. Zhang and Heung-No Lee**, "Performance Analysis of LDPC-Coded Space-Time Modulation over MIMO Fading Channels," *IEEE Communications Letters*, vol. 11, no. 3, pp. 234 – 236, Mar. 2007. (pdf)

9. **Heung-No Lee and X. Hu**, "Robust Iterative Tree-Pruning Detection and LDPCC Decoding," *IEEE Journal on Selected Areas in Communications,* vol. 23, no.5, pp. 1013-1025, May 2005. (pdf)

10. **Sangjun Park, Haeung Choi and Heung-No Lee,** "Time-Variant Proof-of-Work Using Error-Correction Codes," preprint on arXiv, June 2020, https://arxiv.org/abs/2006.12306 .

11. **Hyunjun Jung, Heung-No Lee*,** "ECCPoW: Error-Correction Code Based Proof-of-Work for ASIC Resistance", *Symmetry*, June. 2020, 12(6), 988 (impact Factor: 2.143, Do-Yak project) (PDF)

12. **Seungmin Kim, Gyeongdeok Maeng, Heung-No Lee,** "Smart Contract-Based Checkpoint for Initial PoW Network Security," 2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN 2023), July 2023, pp.89-93.

(Smart Contract)

1. **Satpal Singh Kushwaha, Sandeep Joshi, Dilbag Singh, Manjit Kaur, and Heung-No**

Lee, "Ethereum Smart Contract Analysis Tools: A Systematic Review" n *IEEE Access*, (Impact factor: 3.367, Category: Engineering, Electrical & Electronic, JCR Quartile: Q2: JCI Percentile: 66.75, IITP & Do-Yak project). *(Open Access)*

2. **Satpal Singh Kushwaha, Sandeep Joshi, Dilbag Singh, Manjit Kaur, And Heung-No Lee**, "Systematic Review of Security Vulnerabilities in Ethereum Blockchain Smart Contract", *IEEE Access, Vol. 10, pp. 6605-6621, doi: 10.1109/ACCESS.2021.3140091 (Impact factor: 3.367, Category: Engineering, Electrical & Electronic, JCR Quartile: Q2: JCI Percentile: 66.75, IITP & Do-Yak project). Paper: (Open Access)*

(Cryptography)

1. **M. Kaur et al**., "EGCrypto: A Low-Complexity Elliptic Galois Cryptography Model for Secure Data Transmission in IoT," in IEEE Access, vol. 11, pp. 90739-90748, 2023, doi: 10.1109/ACCESS.2023.3305271.

2. **M. Kaur, A. A. AlZubi, D. Singh, V. Kumar and Heung-No Lee**, "Lightweight Biomedical Image Encryption Approach," in IEEE Access, vol. 11, pp. 74048-74057, 2023, doi: 10.1109/ACCESS.2023.3294570.

3. **Varsha Himthani, Vijaypal Singh Dhaka, Manjit Kaur, Geeta Rani, Meet Oza, and Heung-No Lee,** "Comparative Performance Assessment of Deep Learning based Image Steganography Techniques", *Scientific Reports*, (Impact factor: 4.996, Category Multidisciplinary Sciences, JCR Quartile: Q2: JIF Percentile: 74.66).

4. **Varsha Himthani, Vijaypal Singh Dhaka, Manjit Kaur, Dilbag Singh and Heung-No Lee,** " Systematic Survey on Visually Meaningful Image Encryption Techniques" *IEEE Access*, (Impact factor: 3.476, Category: Engineering, Electrical & Electronic, JCR Quartile: Q2: JCI Percentile: 62.14, IITP & Do-Yak project). *(Open Access)*

5. **Jusung Kang, Young-Sik Kim, and Heung-No Lee**, "Radio Frequency Public Key Generator for Digital Application", *IEEE Access,* Vol. 11, Dec. 2023, doi: 10.1109/ACCESS.2023.3340305, (ITRC project), (Impact factor: 3.9, Category: Engineering, Electrical & Electronic, JCI Quartile: Q2, JCI Percentile: 69.12). Paper: (Link).

6. **Sharanpreet Kaur, Surender Singh, Manjit Kaur, and Heung-No Lee**, "A systematic review of computational image steganography approaches", Archives of Computational Methods in Engineering, Springer, (Impact factor: 7.302, Category: Mathematics, Interdisciplinary Applications, JCR Quartile: Q1: JCR Percentile: 99.54, IITP & Do-Yak project). *(PDF)*

7. **Praveen Bondada, Debabrata Samanta, Manjit Kaur, and Heung-No Lee*,** "Data Security based routing in MANETs using key management mechanism", Applied Sciences, MDPI, *12*(3), 20 Jan. 2022, doi: 10.3390/app12031041, (Impact factor: 2.679, Category: Engineering, Multidisciplinary, JCI Quartile: Q2, JCI Percentile: 67.94, IITP & Do-Yak project). *(Open Access)*

8. **Dilbag Singh, Vijay Kumar, Manjit Kaur, Mohamed Yaseen Jabarulla, and Heung-No Lee***, "Screening of COVID-19 Suspected Subjects using Multi-Crossover Genetic Algorithm based Dense Convolutional Neural Network", IEEE Access, Vol. 9, pp. 142566-142580, Oct 2021. doi: https://doi.org/10.1109/ACCESS.2021.3120717 (Impact factor: 3.367, Category: Engineering, Electrical & Electronic, JCR Quartile: Q2: JCI Percentile: 66.75, IITP & Do-Yak

project) Paper: (Open Access)

(Semantic Analysis)

1. **Shashank S. Singh, Vishal Srivastava, Ajay Kumar. Shailendra Tiwari, Dilbag Singh, and Heung-No Lee**, "Social Network Analysis: A Survey on Measure, Structure, Language Information Analysis, Privacy, and Applications" Accepted in ACM Transactions on Asian and Low-Resource Language Information Processing, (Impact factor: 3.316, Category: Computer Science, Artificial Intelligence, JCR Quartile: Q4: JCR Percentile: 18.35, IITP & Do-Yak project).
2. **Apeksha Aggarwal, Akshat Srivastava, Ajay Agarwal, Nidhi Chahal, Dilbag Singh, Abeer Alnuaim, Aseel Alhadlaq and Heung-No Lee***, "Two-Way Feature Extraction for Speech Emotion Recognition Using Deep Learning", MPDI Sensors (Impact Factor: 3.576, Category: Instruments & Instrumentation, JCI Quartile: Q1: JCI Percentile: 78.91, Do-Yak & IITP project). *(Open Access)*
3. **Rahman S M Wahidur, Ishmam Tashdeed, Manjit Kaur, and Heung-No Lee,** "Enhancing Zero-Shot Crypto Sentiment with Fine-tuned Language Model and Prompt Engineering", in IEEE Access, Early Access, doi: 10.1109/ACCESS.2024.3350638. (URL)

(HSI Classification / regression)

1. **Cheolsun Kim, Pavel Ni, Kang Ryeol Lee, and Heung-No Lee***, "Mass production-enabled computational spectrometers based on multilayer thin films"**,** Accepted in Scientific Reports, (Impact Factor: 4.380, Category: Multidisciplinary Sciences, JCI Quartile: Q1: JCI Percentile: 85.55, Do-Yak project).
2. **Cheolsun Kim, Dongju Park, and Heung-No Lee***, "Compressive sensing spectroscopy using a residual convolutional neural network", *MDPI Sensors*, Vol. 20(3), Jan. 2020. (Impact Factor: 3.031, Do-Yak project) (PDF)
3. **Cheolsun Kim. Woong-Bi Lee, Soo Kyung Lee, Yong Tak Lee, and Heung-No Lee***, "Fabrication of 2D thin-film filter-array for compressive sensing spectroscopy", *Optics and Lasers in Engineering*, Vol. 115, pp. 53-58, Apr. 2019. (PDF) (Impact Factor: 3.388 Doyak project)
4. **J. Oliver, WoongBi Lee, and Heung-No Lee, "**Filters with random transmittance for improving resolution in filter-array-based spectrometers**,"** Optics Express, Vol. 21, No. 4, pp. 3969–3989, Feb. 2013. (Impact Factor: 3.587; Do-Yak Project). (pdf)

(Radio Frequency IoT AI)

1. **Jian Su; Zhenlong Liao, Zhengguo Sheng, Alex X. Liu, Dilbag Singh, and Heung-No Lee,** " Human Activity Recognition Using Self-powered Sensors Based on Multilayer Bi-directional Long Short-Term Memory Networks" *IEEE Sensors Journal*, (Impact factor: 4.325, Category Instruments & Instrumentation, JCR Quartile: Q1: JIF Percentile: 78.91, Do-Yak project). *(Link)*
2. **Harpreet Singh, Suchita Sharma, Manju Khurnana, Manjit Kaur, and Heung-No**

Lee, "Binary drone squadron optimization approaches for feature selection" *IEEE Access*, (Impact factor: 3.476, Category: Engineering, Electrical & Electronic, JCR Quartile: Q2: JCI Percentile: 62.14, IITP & Do-Yak project). *(Open Access)*

3. **Jusung Kang, Younghak Shin, Hyunku Lee, Jintae Park, and Heung-No Lee**, "Radio Frequency Fingerprinting for Frequency Hopping Emitter Identification", *Applied Sciences, 11(22),* Nov. 2021, doi: 10.3390/app112210812 (LIG Nex1 project), (Impact factor: 2.679, Category: Engineering, Multidisciplinary, JCI Quartile: Q2, JCI Percentile: 67.94). Paper: (Open Access) (Paper) (Explanatory Material)

4. **Kiwon Yang, Jusung Kang, Jehyuk Jang and Heung-No Lee**, "Multimodal Sparse Representation-Based Classification Scheme for RF Fingerprinting," *IEEE Communications Letters, Vol. 23, Issue 5, pp. 867-870, May 2019. (Impact Factor: 2.723, Doyak Project) (PDF) (Used DB) (Source Code)*

5. **Jin-Taek Seong and Heung-No Lee,** "Predicting the Performance of Cooperative Wireless Networking Schemes with Random Network Coding," IEEE Transactions on Communications, vol. 62, no. 8, pp. 2951-2964, Aug. 2014. (Imfact Factor: 1.979, Do-Yak Project) (pdf)

6. **Sang-Seon Byun, Ilangko Balasingham, Athanasios V. Vasilakos, and Heung-No Lee** "Computation of an Equilibrium in Spectrum Markets for Cognitive Radio Networks," to appear IEEE Transactions on Computers, Published on-line first on Aug. 30th, 2012. (Impact Factor: 1.473, Do-Yak Project, Haek-Sim Project)(pdf)

7. **Cheng-Chun Chang, Zhi-Hong Mao, and Heung-No Lee,** "MB iterative decoding algorithm on systematic LDGM codes: Performance evaluation," Signal Processing, vol. 90, issue. 1, pp. 373-377, 2010.(Impact Factor: 2.238) (pdf)

(Medical AI)

1. **Manjit Kaur, Dilbag Singh, Vijay Kumar, Heung-No Lee**, "MLNet: Metaheuristics-based Lightweight Deep Learning Network for Cervical Cancer Diagnosis", IEEE Journal of Biomedical and Health Informatics, (Impact factor: 7.021, Category Mathematical & Computational Biology, JCR Quartile: Q1: JIF Percentile: 93.86). *(Link)*

2. **Dilbag Singh, Manjit Kaur, Jazem Mutared Alanazi, Ahmad Ali AlZubi, Heung-No Lee**, "Efficient evolving deep ensemble medical image captioning model", IEEE Journal of Biomedical and Health Informatics, (Impact factor: 7.021, Category Mathematical & Computational Biology, JCR Quartile: Q1: JIF Percentile: 93.86). *(Link)*

3. **Rohini Raina, Naveen Kumar Gondhi, Chaahat, Dilbag Singh, Manjit Kaur, and Heung-No Lee,** "A Systematic Review on Acute Leukemia Detection Using Deep Learning Techniques" Archives of Computational Methods in Engineering, Springer, (Impact factor: 7.302, Category: Mathematics, Interdisciplinary Applications, JCR Quartile: Q1: JCR Percentile: 99.54, IITP & Do-Yak project). *(Link)*

4. **Dilbag Singh, Manjit Kaur, Vijay Kumar, Mohamed Yaseen Jabarulla, and Heung-No Lee**, "Artificial Intelligence-based Cyber-Physical System for Severity Classification of Chikungunya Disease", IEEE Journal of Translational Engineering in Health and Medicine, (Impact factor: 3.316, Category: Engineering, Biomedical, JCR Quartile: Q2: JCR Percentile: 51.12). *(Open Access)*

5. **Jayashree Piri, Puspanjali Mohapatra, Debabrata Singh, Debabrata Samanta, Dilbag**

Singh, Manjit Kaur and Heung-No Lee*, "Mining and Interpretation of Critical Aspects of Infant Health Status Using Multi-Objective Evolutionary Feature Selection Approaches", Accepted in *IEEE Access*, 10.1109/ACCESS.2022.3161154, (Impact factor: 3.367, Category: Engineering, Electrical & Electronic, JCR Quartile: Q2: JCI Percentile: 66.75, IITP & Do-Yak project). *(Open Access)*

6. **Sushama Tanwar, S. Vijayalakshmi, Munish Sabharwal, Manjit Kaur, Ahmad Ali AlZubi, and Heung-No Lee***, "Detection and Classification of Colorectal Polyp using Deep Learning," BioMed Research International, Mar. 2022, (Impact Factor: 3.411, Category: BIOTECHNOLOGY & APPLIED MICROBIOLOGY, JCI Quartile: Q3: JCI Percentile: 43.94, Do-Yak & IITP project). *(Open Access)*

7. **Priyanka Malhotra, Sheifali Gupta, Deepika Koundal, Atef Zaguia, Manjit Kaur and Heung-No Lee***, "Deep Learning Based Computer Aided Pneumothorax Detection Using Chest X-Ray Images", MPDI Sensors (Impact Factor: 3.576, Category: Instruments & Instrumentation, JCI Quartile: Q1: JCI Percentile: 78.91, Do-Yak & IITP project). *(Open Access)*

8. **Abhinav Mishra, Ganapathiraju Dharahas, Shilpa Gite, Ketan Kotecha, Deepika Koundal, Atef Zaguia, Manjit Kaur, and Heung-No Lee***, "**ECG Data Analysis with Denoising Approach and Customized CNNs**", Sensors, MPDI, 22(5), 1928, 2022 (Impact Factor: 3.576, Category: Instruments & Instrumentation, JCI Quartile: Q1: JCI Percentile: 78.91, Do-Yak & IITP project). *(Open Access)*

(Agricultural AI)

1. **Harshit Kaushik, Anvi Khanna, Dilbag Singh, Manjit Kaur, Heung-No Lee,** " TomFusioNet: A tomato crop analysis framework for mobile applications using the multi-objective optimization based late fusion of deep models and background elimination" (PDF)

2. **Dilbag Singh, Yavuz Selim Taspinar, Ramazan Kursun, Ilkay Cinar, Murat Koklu, Ilker Ali Ozkan, and Heung-No Lee***, "Classification and Analysis of Pistachio Species with Pre-Trained Deep Learning Models", MPDI Electronics (Impact Factor: 2.397, Category: Physics and Applied, JCI Quartile: Q2: JCI Percentile: 57.60, Do-Yak & IITP project). *(Open Access)*

(Character recognition AI)

1. **Dilbag Singh, Aravinda C.V, Manjit Kaur, Meng Lin, Jyoti Shety, Vikram Raju Reddicherla, and Heung-No Lee,** "DKNet: Deep Kuzushiji Characters Recognition Network" *IEEE Access*, (Impact factor: 3.476, Category: Engineering, Electrical & Electronic, JCR Quartile: Q2: JCI Percentile: 62.14, IITP & Do-Yak project). *(Open Access)*

(Air quality prediction)

1. **Manjit Kaur, Dilbag Singh, Mohamed Yaseen Jabarulla, Vijay Kumar, Jusung Kang, and Heung-No Lee.** 2023. Computational deep air quality prediction techniques: a systematic review. Artificial Intelligence Review (2023), 1–46

Gwangju Institute of
Science and Technology

School of Information and Communications



WorldLand™

- WorldLand, My AI Network 핵심 기술 연구 결과–

Green Bitcoin: Global Sound Money

Heung-No Lee, Young-Sik Kim, Dilbag Singh, and Manjit Kaur

Abstract

Modern societies have adopted government-issued fiat currencies many of which exist today mainly in the form of digits in credit and bank accounts. Fiat currencies are controlled by central banks for economic stimulation and stabilization. Boom-and-bust cycles are created. The volatility of the cycle has become increasingly extreme. Social inequality due to the concentration of wealth is prevalent worldwide. As such, restoring sound money, which provides stored value over time, has become a pressing issue. Currently, cryptocurrencies such as Bitcoin are in their infancy and may someday qualify as sound money. Bitcoin today is considered as a digital asset for storing value. But Bitcoin has problems. The first issue of the current Bitcoin network is its high energy consumption consensus mechanism. The second is the cryptographic primitives which are unsafe against post-quantum (PQ) attacks. We aim to propose Green Bitcoin which addresses both issues. To save energy in consensus mechanism, we introduce a post-quantum secure (self-election) verifiable coin-toss function and novel PQ secure proof-of-computation primitives. It is expected to reduce the rate of energy consumption more than 90 percent of the current Bitcoin network. The elliptic curve cryptography will be replaced with PQ-safe versions. The Green Bitcoin protocol will help Bitcoin evolve into a post-quantum secure network. In addition, it improves the properties of Bitcoin's hash PoW while addressing environmental concerns.

Keywords: Bitcoin, energy consumption, error-correction codes, post-quantum security, sound money, verifiable random function

I. INTRODUCTION

WE the global citizens not by our choice live in a world of boom-and-bust cycles created by the Federal Reserve Board (FED) of the United States. In the boom phase of a cycle, the FED supplies debt-monetized US dollars to the world, and supplied USDs are used to purchase real items, goods, and services from developing countries. Such supplies prop up bubble markets, such as the US housing, stock, and derivatives markets. Global elite financial institutions, such as investment banks and hedge funds, benefit the most from making risky investments. When the bust part of the cycle comes, the working class worldwide suffers from

This work was supported in part by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (NRF-2021R1A2B5B03002118) and the Ministry of Science and ICT (MSIT), Korea, under the Information Technology Research Center (ITRC) support program (IITP-2021-0-01835). This work was also supported by the Tech Incubator Program for Startup (TIPS) program (S3306777) awarded in June 2022.

Heung-No Lee is with Gwangju Institute of Science and Technology, Gwangju, 61005, South Korea and affiliated with LiberVance,Co., Ltd., Gwangju, South, Korea.

Dilbag Singh and Manjit Kaur are with Gwangju Institute of Science and Technology, Gwangju, 61005, South Korea.

Young-Sik Kim is with Chosun University, Gwangju, 61452, South Korea, and affiliated with LiberVance, Co., Ltd., Gwangju, 61005, South Korea.

Corresponding author: Heung-No Lee (e-mail: heungno@gist.ac.kr).

Received: September 15, 2022; Revised: October 20, 2022; Published: October 31, 2022.

inflation and market crashes. In the aftermath of a market crash, governments bail out financial institutions to prevent chain bankruptcies. Such centralized planning has disrupted societies globally [21][29]. The boom-and-bust volatility peaked with the Dow to gold ratio has become increasingly extreme [6]. Inequality is a prevalent condition worldwide. Work ethics fade. Growing are speculative markets. Keeping this system in its current form does not help advance humanity to the next level.

Throughout history, money has taken many forms, including gold, silver, copper, salt, and seashells [21]. As modern society developed, government-issued fiat currencies became normal, many of which exist in the form of digits in banks and credit accounts today. The soundness of money is determined by its stability; if it is stable, it can function as a medium of exchange, unit of account, and store of value. Central banks know this and seek to stabilize their currency by controlling money supply as the economic condition changes. Today, however, this flexibility is often misused and overused, especially by a new government that needs to satisfy voters and make way for its political agenda. The boom comes from stimulation, and the bust comes back from monetary tightening. It seems clear that innovative measures are needed to address the current situation, which depends entirely on central bank policy decisions.

Today, central banks are often tightly coupled with political powers. When new presidents come into the office, they are tempted to use the central bank's power in money creation capability to fulfill their political agenda. The stability of money, and thus the market order, is broken in the name of economic growth and stabilization. Owing to the manipulation of money, monetized debts, inflation of asset prices, and growing inequality are rampant [4]. Thus, there is a need to restore sound money whose supply is independent of governments' control, which is left alone to a free market and its self-regulation mechanisms.

Bitcoin was humanity's first success in creating decentralized money [21][41]. It realized a sound money Fredrick Hayek proposed in his book, *The Denationalization of Money* [22]. Its success was possible because of cryptographic technologies, such as SHA, digital signature algorithms, and elliptic curve cryptography. Currently, Bitcoin is one of the soundest currencies, but it still does not fully satisfy itself as sound money. For example, it does not work as a means of payment, as blockchain networks incur transaction costs that are too high for daily spending. A lightning protocol (along with other second-layer solutions) [53] facilitates off-chain transactions at high speeds. Unfortunately, these solutions are not secure against post-quantum (PQ) attacks.

These cryptographic primitives must be upgraded regularly. Otherwise, their use is limited and useless if not upgraded. IBM announced its plan to introduce quantum computers with more than four thousand qubits by 2023 [26]. With such advances, cryptographic algorithms used in Bitcoin are on the verge of breaking.

The main contributions of this paper are as follows. This paper reviews Hayek's sound money and discusses the pressing need to restore it in our society. The advantages of Bitcoin concerning its robust performance, such as simplicity in consensus and time-energy-borne wealth characteristics, are emphasized. Subsequently, a post-quantum (PQ) safe Green Bitcoin protocol is proposed. The proposed protocol can help achieve the majority of sound money properties. Green Bitcoin comprises two major parts: a PQ secure verifiable (self-election) coin-toss (VCT) function and a novel PQ secure proof-of-computation (PoC) primitive. The PoC part is built based on a newly published finding known as the error-correction code anti-ASIC proof-of-work (ECCPoW) [30][31][33][45]. PoC primitives will make PQ safer than the ECCPoW. Environmental concerns can be addressed with the VCT function, aided by the PoolResistantComp protocol. This can be used to control the network's energy consumption efficiency. Critical components of opcodes, such as digital signature algorithm, will be enhanced, and elliptic curve cryptography will be replaced with PQ safe versions. The Green Bitcoin protocol will help Bitcoin evolve into a post-quantum secure network. In addition, it helps achieve good properties of Bitcoin's hash PoW while alleviating environmental concerns.

The remainder of this paper is organized as follows: Section II discusses the significance of cryptocurrency and blockchain. Section III discusses the benefits of Bitcoin, sound money, and problems. Section IV presents the Green Bitcoin proposal. Novel PQ secure primitives are discussed in Section V, and a Green Bitcoin testbed and discussion are presented in Section VI. Section VII concludes this paper.

II. CRYPTOCURRENCY AND BLOCKCHAIN

Cryptocurrencies are digital currencies created through encryption algorithms that can be used as a form of payment. Using encryption technologies, cryptocurrencies can serve as currency and virtual accounting systems. Digital wallets are used for cryptocurrency trades. Blockchain networks were developed to provide decentralized requirements for cryptocurrencies. Consensus, virtual machines, and peer-to-peer (p2p) networking are the three primary components of a blockchain. One of the most pressing demands is to 1) update the consensus mechanism that allows a new PQ secure and decentralized blockchain network and 2) upgrade the cryptographic primitives used in consensus and virtual machines to be PQ safe.

Currently, blockchains are not PQ secure, there are environmental concerns and scalability issues. Several major projects have implemented Proof of Stake (PoS) paired with a Byzantine agreement (BA) algorithm to resolve these concerns. The PoS and BA algorithms are perhaps good for a fast-computing platform, but they are not secure enough for a global monetary-grade blockchain. The ideas are not new; the BA algorithm (developed in the 1980s [12][32]) relies on communication across committee nodes to reach a consensus, rendering it subject to various assaults, including DDoS (distributed denial of service) and network partition attacks. To scale it well in terms of the number of nodes, the committee size cannot but remain small; hence, decentralization is compromised.

Blockchains are a dear but expensive solution. The blocks are stored redundantly within every consensus-participating node, and all nodes perform the same work. Each new block is made with an effort, an enormous amount of time and energy, by the entire network. This is the source of the immutability of the records stored in the blockchain; the network must be decentralized to the maximum possible extent. The greater the number of individual nodes participating, the more secure the network is regarding censorship resistance, thwarting Sybil, and double-spending attacks [27][41]. Consider Bitcoin; each block contains a massive amount of computational energy stored in each block. If the block needs to be forged again, it requires the same amount of energy to be stored. On the one hand, the large redundancy and numerous independently working nodes doing the same work can be viewed as a source of security. On the other hand, it can be viewed as a waste of resources and a waste of energy. The blockchain trilemma [5][7][14]—it is difficult to achieve the three blockchain properties simultaneously such as scalability, decentralization, and security—represents a scalability challenge caused by inefficient resource consumption. The complaint about energy issues leads to environmental concerns.

Blockchain technology requires a simple protocol to withstand attacks and perform robustly for years to come [33]. The consensus mechanism should be able to maximize the resilience of numerous unknown attack vectors. How can we make it simple while accommodating many p2p nodes working together to reach a consensus? The participating nodes must make timely judgments and choose one block from several candidate blocks to be the new block attached to the status quo chain. A timely consensus decision should be distributed among numerous independent working nodes over the internet. Consequently, an agreement must be reached with as few contacts as possible among the p2p nodes.

Frequent network delays and partitions occur on the internet. They might be caused by momentary router failures, traffic congestion, or purposeful antagonistic activities. Therefore, a global blockchain network must be resilient against possible attacks and losses.

A consensus mechanism using hash function-based PoW [41] has been shown to provide the most decentralized and secure operations. Thus, it is challenging to design a consensus mechanism that achieves all these needs: a large number of p2p nodes, making timely decisions while working independently, with minimal inter-node communication.

III. BITCOIN, SOUND MONEY, PROBLEMS

A. Bitcoin

Bitcoin [41] is a newly developed type of money known as “cryptocurrency.” Some in the Bitcoin community believe that Bitcoin already represents sound money [23][24]. Sound money is defined as money that has a purchasing power determined by markets, independent of governments and political parties. There would be a vast difference between a world with Bitcoin and one without Bitcoin. While governments continue to print

cheap money, ordinary people can resort to Bitcoin. Thus, the wealth gap can be significantly reduced. We will come closer to a produce-first-and-spend economy rather than a debt-monetized spending-first economy. Bitcoin encourages savings and autonomous management of money; it does not allow bailouts. Governments are not held accountable for policy failure. Tax-paying individuals and corporations are the entities who are held accountable for the economic setbacks caused by policy failures. The use of the power to mint the currency held by governments (non-responsible entities) should be limited. That way, central banks will no longer be able to punish diligent savers by sprinkling them with debt-based currencies. With a stable store of value, people can confidently plan for the future. Zombie companies and bubble markets that thrive on government budgets and easy monetary policies will find it difficult to survive.

An organism's code of life is written at its conception. Bitcoin's DNA, or genetic code, was carefully crafted by Satoshi as the soundest money ever created. Bitcoin's genetic code can be considered a set of instructions designed to incentivize the coordination and organization of cellular functions.

The genetic code of Bitcoin [41] is as follows:

- Bitcoin had to be ignited to become real, so Satoshi coded a fixed supply (21 million Bitcoins) into its DNA. Sound money comes from this fixed supply. Due to the increase in users, miners and developers, Bitcoin has become more expensive over time. Thus, the feedback loop has become self-reinforcing.
- The mining function, that is, the one utilizing a hash PoW, is the metabolism and defense mechanism of Bitcoin. For instance, Bitcoin consumes a lot of electric energy to make new blocks and create virtual walls to protect the network from hackers. The anti-fragility of Bitcoin is attributed to PoW, which makes it more resistant to attacks as it grows.
- The Bitcoin network generates a new block every 10 min on average. This rate is carefully determined for the robust operation of the nodes. Bitcoin nodes are scattered all over the globe; thus, they are separated over long distances. However, they can still communicate and coordinate effectively and have never stopped making new blocks since their inception in January 2009.

B. What Is Sound Money?

The 1974 Nobel Prize in Economics Laureate Friedrich Hayek wrote the *Denationalization of Money* (1976) [22]. He said privately issued money, which continues to develop through competition, will inevitably be superior to fiat money, which does not evolve because the government has exclusive power to supply it.

> "I am more convinced than ever that if we ever again have sound money, it will not come from the government. This is issued by a private enterprise," Friedrich Hayek, 1977

Hayek emphasizes sound money in his book. Literally translated, it can be interpreted as honest money. However, one may wonder what sound money is and why he emphasized it. What will happen if money is sound, and what will happen if it is not? As a firm definition is not given in his book, the answers to these questions can be examined via the statements he provides in his book.

Today, we live in the age of fiat currencies. For example, the US dollar is a fiat currency. Fiat means "by government decree." By law, the U.S. government declares the dollar to be a currency. Because the United States has the most powerful military and the largest economy in the world, it has gained worldwide trust, and the US dollar has become the world reserve currency.

Individuals earn sound money through solid value-creation processes. To produce more gold, one has to mine deeper in the gold mines. Gold has been utilized for thousands of years as a means to store value for future use. No government can print a sound money. No government can devalue the money and weaken the savings. To earn other people's money, one must provide equivalent time and energy for others. Governments should not be exceptions. They must also be prudent and considerate. They must win the hearts of their peoples with reliable, agile and considerate services, and a clear vision. They shall be limited to printing new money.

C. What Is the Problem If Money Is Not Sound?

The US dollar has caused serious social problems, as revealed by several studies [23][25]. The dollar has been overused by the US government and the central bank FED in the name of "stimulating the economy," "resolving the financial crisis," and "fighting the pandemic." Whenever a bust phase occurred, dollars were

printed and supplied to society. Recent examples include quantitative easing measures to address the financial crisis caused by subprime mortgages in 2008. The US government issues treasury bonds and the FED provides dollars to society by buying them. The newly issued dollars spread to the world society. Major Western advanced countries will also have to expand the supply of their currencies to resolve the negative side effects of the large-scale supply of dollars.

This creates a situation where debt increases in major countries worldwide, and an oversupply of money is widespread. The additional money supplied each year causes real estate prices in major cities to skyrocket. The value of financial assets, such as the S&P 500 ETF in the US, continues to grow in a right-upward way by more than 15% annually. There is a huge wealth gap between those who can own these assets and those who cannot. Speculative markets in which unearned income can be obtained as much as the amount of oversupplied fiat currency will grow. However, the value of honest work based on diligence, knowledge, and skills decreases. Knowledge-seeking and corporate activities weaken. What labor activity, what knowledge-seeking activity, or what company's business model could produce growth rates higher than that of the S&P 500?

In a society where honest money reigns, bubbles weaken, and honest work is promoted. Companies invest in new knowledge- and skill-creation projects and develop new products and services. Honest labor and knowledge-seeking activities are vibrantly carried out, the workforce participating in production activities increases, and the economy grows robustly. Because there is no oversupply of money, no speculative bubble markets grow. As the speculative market disappears, more people can find a stable life through honest work.

Gold has been used as honest money for thousands of years [21][22]. Honest money is very important to humans because it performs useful roles such as a means of exchange, a measure of value, and store of value. However, gold is not easily produced. Producing even a single gram of new gold requires someone to work very hard over a long period. The only way to produce gold is to mine it from the ground. All gold that exists on the earth's surface has already been mined. Therefore, to mine gold today, you have to dig deeper. However, the production of gold has become increasingly difficult. This rarity is a key characteristic of gold. Because of these properties, gold was used as sound money.

The United States adopted the gold standard through the Bretton Woods Agreement of 1944. An ounce of gold was pegged at 35 US Dollars. Currencies in the Western world, such as the Franc (France), Pound (United Kingdom), Mark(Germany), and Yen(Japan), were pegged to the US Dollar, and currencies in developed countries were pegged to gold. Therefore, these currencies of major countries around the world was honest. On August 15, 1971, the 37th President of the United States, Nixon, announced [55] that he would temporarily withdraw from the gold standard, shocking the world. Since then, there have been many reports that easy monetary policy has created a polarization of wealth [25].

D. Is Bitcoin Sound Money?

Bitcoin has already acquired the status of an asset (a means of saving and investing) in many countries. El Salvador became the first country in the world to recognize Bitcoin as its legal tender on September 7, 2021. The US Securities and Exchange Commission (SEC) approved the Bitcoin Futures ETF on October 16, 2021 [9].

Is Bitcoin already the sound money Hayek was talking about? This is believed to be so among many Bitcoin enthusiasts [23][24]. According to Hayek's proposition of sound money, however, it should evolve through competition. Therefore, many Bitcoin hard forks should emerge in the future and continue competing to provide better services. As numerous new Bitcoins compete, the quality of monetary services constantly evolves. Fiat currency dominance will not disappear, as both the U.S. government and the FED issuing fiscal and monetary policies will continue to exist. In such a world, ever-evolving sound money protocols will provide a window through which citizens can protect themselves from the side effects of recklessly issued fiat currencies.

E. What Are Good Properties to Inherit from Bitcoin?

(**Simplicity in consensus**) The consensus among many p2p nodes is made every 10 min over the best-effort service and often-times hostile internet. How can robust operations become possible? It is interesting to focus on the Bitcoin consensus. The nodes in the Bitcoin network are not divided according to their job. Each node performs the necessary work, confirms transactions, groups them into a block, adds the proof of work, and publishes the block as quickly as feasible. Consensus is reached as the result of each node simply performing

its work for its own benefit. No node is forced to work in a time-division schedule. Each node does not need to adapt to the progress of the other nodes (therefore requiring no contact with each other) to obtain a consensus. Every node creates blocks, and every node validates the blocks. The protocol is simple and plain. This simplicity results in a robust performance. Hence, blocks cannot but be kept on produced. Rewards are provided as incentives to nodes. The more effort a node makes the more opportunities it will earn rewards. Node righteousness is not required, and no punishment that exists in a PoS-based consensus algorithm [14] is required.

(**BA algorithms are not decentralized.**) Jobs are divided under a BA algorithm. A set of nodes under the proposer's name creates blocks. The proposed blocks are validated by another group of nodes, such as *attesters*. They vote for each candidate block. Cast votes are collected and counted. A block with a satisfying number (i.e., supermajority) of votes was selected and connected to the *status quo* chain. To complete this, each node must fit into a tight schedule. Consequently, the time axis must be separated into slots and epochs. The number of nodes engaged in consensus must be limited to a few hundred to operate properly across a hostile environment such as the internet.

(**Time-energy-borne wealth: Bitcoin is a stored wealth transformed from time and energy.**) Each block header contains information on time and energy consumption. Take any blocks from the past. The block header contains the time stamp when the block was created. The difficulty level of the puzzle is also specified. Using this information, one can calculate the amount of computation (hash cycles) required to solve the hash puzzle. It would have taken more than three years for a single node working alone to solve the hash challenge. But it takes only 10 min for the entire network of operating nodes distributed independently worldwide. See the alone impossible together possible (Al-Im-To-Po) theory [33]. Thus, it indicates that hundreds of millions of computing nodes have to work together at that moment to produce that block. Given the energy efficiency of a mining node, the amount of total energy expended to create a block can be calculated. A predetermined number of new bitcoins are minted on the block. It can be said that the blockchain network has transformed the energy spent on creating the block into the Bitcoins issued from that block. The miners invested time and energy, thus are deserved to get the Bitcoins reward.

In this approach, each block generation process may be considered a new way of storing wealth. It is a transformation of the most basic resources of time and energy. Time and energy are the most valuable resources humans have and are the most fundamental types of wealth. Thus, the coins created in each block are assigned a monetary value from birth. The effort and time are not squandered; they were converted into something valuable: Bitcoin. Others have used their time and energy to create important commodities and services such as food and building houses. One can give Bitcoin to purchase goods and services that others have produced and offered.

F. What Are the Issues That Require Attention?

There are two problems. The first is the security risk in cryptography used in Bitcoin owing to advances in quantum computing technology [1][17]. The second is the huge energy consumption issue of the Bitcoin network.

As mentioned earlier, IBM is close to building a quantum computer that is known to break well-established cryptographic algorithms, such as elliptic curve cryptography, digital signature algorithm, and RSA algorithm. President Biden has signed two executive orders [50]. The first is to address how US leadership can be maintained in quantum computing technology. The second is to outline the U.S. government's strategy for mitigating the risks to vulnerable cryptographic systems due to advances in quantum technology. We address this further in Section V.

Bitcoin's energy consumption is enormous; its annual energy consumption has grown to 138 TWh in early 2022, which is more than the power consumption of a country such as Norway. Its annual carbon dioxide emission reached 114 million tons, which is comparable to that of Belgium [49][54].

Bitcoin advocates claim:

- The energy usage is still a very tiny proportion of global electricity consumption.
- Miners with self-interest move to places where surplus gas, solar panel farms, hydro and wind power plants, low-carbon nuclear plants, and geothermal energy are available.

But many still express concerns [50][54]:

- Chinese ban on Bitcoin mining was a response to a power deficit.
- Kazakhstan has imposed a limit on it owing to energy shortages.
- Sweden has called for the Europe-wide ban, blaming it for slowing climate transition.
- Tesla has withdrawn from its plan to accept Bitcoin as a form of payment, citing environmental concerns.

G. Alternatives miss the merits of Bitcoin

Detractors claim that energy-intensive PoWs are the culprit. Rival currencies that use variants based on proof-of-stake (PoS), such as Ethereum and Solana, are gaining popularity. Ethereum [14] announced its plan to reduce its estimated energy consumption by up to 99% with the completion of the transition to its PoS variant system [5][7][8][14]. Ethereum completed the Merge on September 15, 2022.

However, PoS is known to have many obvious concerns [7][10][35]. The PoS is being introduced to address energy concerns and increase the transactions per second (TPS). However, the penalty may be significant because it is neither decentralized nor secure. The PoS is not a technological advancement; it lacks the advantages of the PoW. It resorts to a sociopolitical solution: Plutocratic politics and the time-energy-born wealth property is lost. Because the PoS does not retain any energy on a block, the blocks may be readily rewritten; hence, it is insecure. Advocates of PoS put forward a policy for bad actors, such as "your stakes will be confiscated if you act badly," This is a "fixing a barn after losing a cow" approach. The richest few can make confidential agreements off-chain and take control of the blockchain. These off-chain conspiring operations leave no on-chain trace; thus, no one can become aware of them. Hence, it is sensitive to bribery and conspiracy [38]. There is a nothing-at-stake risk [35]. There is a risk of a grinding attack if the random function for selecting a block creation node is unfair or predictable [10].

A PoS-based alternative may serve as a global computing platform but may not be decentralized, nor secure and thus not suitable for a global monetary grade network such as Bitcoin.

IV. GREEN BITCOIN

In this section, we discuss Green Bitcoin.

A. Key Performance Metrics

A set of key performance metrics of Green Bitcoin are:

M1. PQ secure cryptography and consensus
M2. Energy consumption efficiency (ECE)
M3. Byzantine fault tolerance (1/2)
M4. The mining schedule (21 million Bitcoin)
M5. The block generation time (10 min average), and
M6: The block size (1 mb)

Note that M4, M5, and M6 are the same as those of Bitcoin. M1 and M2 are our major focus. We aim to develop a PQ secure cryptography and a new energy-efficient consensus mechanism. For M3, as later shown in this paper, Green Bitcoin supports a BFT of up to 50%.

B. Green Bitcoin Consensus

The Green Bitcoin consensus comprises two new major parts: a verifiable (self-election) coin-toss function (VCT) and a novel proof-of-computation (PoC) primitive. A simple PoolResistantComp algorithm can be used to aid these two parts. Green Bitcoin will base its PoC part on a recently published anti-ASIC technology known as the error-correction code proof of work (ECCPoW) [30][31][45]. A more detailed discussion on ECCPoW is provided in Section V.F. A critical component of the virtual machine will also be enhanced; in particular, elliptic curve cryptography will be replaced by our new PQ safe cryptography.

We aim to address the PQ security and environmental concerns while maintaining decentralization. We will let computational challenge decreased. More nodes will be able to join the network as computational difficulty decreases. As a result, security will be bolstered via enhanced decentralization. The time-energy-borne wealth property of Bitcoin will be left de-emphasized. We aim to achieve the sound money property

without its price relative to fiat soaring.

Such goals will be met through the development of new on-chain means. Each node performs a simple task independently; the system's simplicity allows the system to run steadily even with a large body of participant nodes.

The consensus protocol is simple. Each participating node can easily obey the rules, and each node performs the same simple job. No job and time are divided. They need no inter-node communications to reach a consensus; each node performs its work independently. The same procedure is repeated for each new block; consensus for the current block is completed when some nodes announce the valid next block. The only announcement that each participant needs to keep a vigil for is the announcement of this valid new block. This simplicity lowers the entry barrier and invites more participating nodes.

Thus, Green Bitcoin enables the construction of a decentralized, scalable, and secure consensus solution for an extremely large network of participating nodes. It works well even if the number of participating p2p nodes exceeds one million.

(**The base-set of p2p nodes**) The base set of p2p nodes is defined as all the nodes participating in transaction validation and block formation. The protocol is set to perform well with base-set sizes greater than one million.

(**Green Bitcoin consensus**) Like the hash PoW, all nodes in the base set collaborate and contribute to creating each new block. Each node self-selects as a validator, attaches a proof for it, validates transactions, forms a block, and attaches a proof of the solution. Each node repeats this procedure for each new block. The benefit of this is the simplicity of the algorithm. Finality is determined by the longest chain or the amount of energy stored (measurable by each block difficulty level) in the blockchain. If there are two blockchains, each node for its own benefit will select the one with the most energy stored and connect a new block into it.

(**Verifiable coin-toss function**) Let us consider a coin toss game. Every node has its own unique (secret key) coin. Each node tosses its coin. A coin toss features a single output that may be either a pass or fail. Verifiable coin-toss function (VCT) is a verifiable random function (VRF). It has two inputs. One input is the secret key of the node, and the other is the previous block header. Thus, each node cannot but toss this VCT once and only once for each block. The purpose of VCT is to provide a means to turn off a certain portion of the base-set nodes, thus saving energy while allowing a large number of nodes to participate. If the probability of failure is set to 90%, 90% of network nodes are put to rest, and thus energy savings of 90% is achieved. To work on a new next block, each node tosses coin again.

We attempt to design the VCT function so that the odd of pass can be controlled. The probability of pass is a critical parameter that the network designer can use to change the amount of energy saving, given the size of the base set. For example, when the number of nodes participating in the network is small, it can be set to 100% to maximize security.

For those nodes that have been self-selected to perform the computing work, there are three types of computations: VeriComp, SolComp, and PoolResistantComp, each of which has to be performed.

(**VeriComp**) validates transactions and compiles them into a new block. Each node sets out to validate the new block upon receiving a new block announcement. If the block is valid, the node sets out to begin extending a new block next to the validated block.

(**SolComp**) is the computation required to solve the crypto puzzle. Each round presents a completely fresh crypto-puzzle. The puzzle problem is not predictable in advance, but it is determined if the preceding block header is fixed. Each node in the self-elected set starts the race to solve the crypto-puzzle as quickly as possible. A node with a proof of solution inserts the proof into the block header and broadcasts the new block instantly.

(**Coin is tossed before any energy is spent on SolComp.**) In the Green Bitcoin consensus, we observed that all nodes participate in the same simple routine. Each node performs a random turn to form a new block and attaches the proof of solution. The new VCT function makes it possible for them to take random turns. Each node progresses into the energy hungry SolComp routine if its coin toss result is a pass.

C. What If Each Selected Node Forms a Mining Pool?

It is possible that the nodes selected by tossing coins can promise a reward distribution to the unselected nodes and request collaboration. Selected nodes would enjoy such collaboration because they can use them to stay ahead in competition among the selected nodes and increase the probability of winning the block reward. Unless there is a deterrent to prevent selected nodes from making these choices, the proposition that VCT can

reduce energy consumption may turn out to be untrue.

The key question then is: Can we find an on-chain means to discourage such cooperation? Can we find a way to reduce the incentive for a selected node to form a pool by itself? Such a concern has its roots in pool mining practice in current PoW mining networks such as Bitcoin and Ethereum Classic. In a mining pool, miners subscribe to a mining pool server, perform mining tasks, and share rewards by submitting proper solutions.

To quickly grasp the idea, we can delineate our proposed solution for the pool-mining case. Here, the pool mining server is compared with one of the *selected* nodes in the Green Bitcoin consensus mechanism; the miners subscribing to the pool mining server can be compared with the *unselected* nodes. Each selected node can solicit cooperation from unselected nodes and increase its probability of winning the block reward. Unselected nodes can increase their chance of winning a share of the block reward by helping the selected node. Because both parties may benefit from such cooperation, unselected nodes continue to spend energy doing the SolComp work even if they were unselected in the first place by coin-toss. This would erase the energy reduction effect. In a pool mining protocol [46], five major steps are required for miners to obtain shared rewards:

1) register itself as a miner to a Stratum server
2) get the block header information
3) work until a mining success is announced
4) submit the mining results, and
5) get shared rewards from the server.

The key idea in breaking apart the cooperation between the selected and the unselected nodes is to create a conflict between them. One such scheme can be designed such that the selected node is asked to have its private key at risk revealed in order to delegate SolComp work to the unselected nodes.

(PoolResistantComp) There is a method proposed in the literature [15] that aims to break the cooperative tie between the pool server (vis-a-vis the selected node in our case) and the miners (the unselected nodes), thus discouraging pool mining. This is known as a two-phase PoW system. The first PoW is the routine PoW part, which is the same as the existing Bitcoin crypto-puzzle. Miners seek to find the hash of the header that is smaller than the published difficulty parameter. The second PoW is PoolResistantComp: it is the second puzzle in which the miner is asked to include a critical piece of information as an input to the hash function. Namely, the miner is asked to have the header signed with the private key of the coin-base transaction, that is, SHA256 (sign (header, privkey)), and the second puzzle is resolved if a node successfully presents a hash of that signature that is smaller than a second difficulty parameter.

Note here that the second puzzle cannot be relegated to unselected nodes unless each selected node takes the risk of revealing its private key to the unselected nodes.

In our case, we do not need such a two-phase protocol. The PoolResistantComp routine can be incorporated right into the SolComp stage with minimal effort. The PoolResistantComp part is a routine asking for the inclusion of the private key corresponding to the coin-base transaction. ECCPOW (see Fig. 1 in Ref. [45]) is a routine in which a decoder finds a codeword from the hash output, that is, OUT = SHA256(current block header). To embed PoolResistantComp into it, we can replace the hash output routine to include the private key's signature, i.e., OUT = SHA256(sign(current block header, privkey)). Such a single-line update is sufficient to regulate the nodes and force them to rest if they are not selected. This update does not lose the characteristics of *the simple* and *plain* routine discussed in IV.B. Each node simply does its routine, regardless of the states of the other nodes.

V. NOVEL PQ SECURE PRIMITIVES

We now discuss the novel postquantum-ready and suitable cryptographic primitives. They are the novel key generation, sign, and verification functions; it also has a new Green Bitcoin, VRF, VCT, and VC.

Quantum computers are known to break well-established cryptographic algorithms such as the elliptic curve cryptography, digital signature algorithm, and RSA algorithm. In particular, these methods are based on integer factorization and discrete logarithm problems, which are known not quantum-safe. In contrast,

code-based cryptography issues are known to be quantum-safe.

(**A brief history of early code-based cryptography**) McEliece first presented code-based cryptosystems using binary Goppa codes in 1978 [37]. In 1986, Niederreiter proposed a knapsack-type public-key cryptosystem based on error-correction codes using GRS codes [42]. Subsequently, the Niederreiter method was demonstrated to be as secure as the McElice cryptosystem. Sidel'nikov and Shestakov demonstrated in 1992 that Niederreiter's plan to employ GRS codes was insecure [48]. Various methods have been proposed to minimize the public key size by employing different codes, such as the Gabidulin code [18][19], algebraic geometry code [20][28], and Reed-Muller code [47]. However, all of these approaches ultimately proved to be unstable [34][44].

A. Recent PQ Secure Signature Primitives

This section covers recent advances in PQ cryptography and selects a set of suitable PQ secure algorithms. They can be used to meet our goal of developing a PQ secure signature and a PQ secure VRF for Green Bitcoin.

A digital signature (DS) algorithm comprises three parts. The first is the KeyGen component, which produces a public and private key pair. The second is the Sign part. Given the message and private key, it generates a signature. The third is the Verify component, which generates a binary pass or fail output based on a message and signature.

A VRF is similar to the DS algorithm in that it comprises three functions: a KeyGen function that generates a private and public key pair, a VRF function that outputs a signature (proof), and a random number, given the input of a private key and a message, and a Verify function that generates an output of pass or fail, given the input of the public key, message, random number, and proof.

We conducted preliminary research and discovered that Dilithium [36], Falcon [52], and Durandal [2] are good candidates for PQ-safe signature algorithms suitable for Green Bitcoin: the key metrics we used to select them are the size of the keys, the size of the signatures, the time it takes to complete a sign, and the time it takes to verify.

Let us now compare them. The time unit is msec. Dilithium requires 1.4, 6.2, and 1.5 for KeyGen, Sign, and Verify, respectively. Falcon, however, was 197.8, 38.1, and 0.5. Durandal received four, four, and five points, respectively. Durandal, therefore, has the quickest signature time compared with rival methods. The signature time corresponds to the VRF generation time. Thus, Durandal could serve as the first candidate for building a fast PQ-safe VRF. It offers a code-based DS algorithm for ranking metrics. While rendering a quantum-safe signature, it is sufficiently concise. The signature is 4 kb (kilo byte) long, whereas the public key is approximately 20 kb. The signature and verification processes took only 4 ms and 5 ms, respectively. It is powerful, quick, and concise enough to be considered a candidate for a worldwide public monetary-grade blockchain, such as Green Bitcoin.

We proceed to carefully study these candidate PQ secure algorithms and select the best one that satisfies all the key performance metrics of Green Bitcoin. The selected signature algorithm will be implemented using C++, and we will replace elliptic curve DS cryptography.

B. Novel PQ Secure VRF, and VCT Functions

We aimed to discuss how to create a novel PQ secure VCT function. To this end, we first need to create a good PQ secure VRF function. We then use it to create the VCT function.

A VRF is a function that generates a unique random number with a unique signature (proof) attached to it, given a private key and message. It is distinguished from an ordinary random number generator because it also offers a verification procedure. Thus, any verifier can check whether the random number is properly calculated. The quality of the random number generated must be high; given the size of the keys, the entropy of the random number is maximized.

A VRF is similar to a DS scheme, as mentioned in Section V.A. However, a significant difference was observed. It is the signature. The signature for a DS method needs to be non-unique stochastic by design. Stochastic signatures can increase security. However, for a VRF, the signature must be unique to each fixed input. We recall our goal of using VRF to save energy in the SolComp stage. Hence, the VRF should be designed to execute just once and only once every block; otherwise, the node would abuse it by running the VRF as many times as possible until the node produces a suitable output; no energy savings are realized as a result. Such enforcement is achieved if VRF generates a unique signature for a given fixed input message.

The input can be designed to be a piece of public information that existed before the VRF was performed, such as the block header of previous block. The same applies to the public keys. The public key should have been already posted somewhere in the blockchain before running the VRF. Therefore, the private key associated to the fixed public key is fixated. As the result, each node cannot but run it once and only once per each block.

C. Our Approach

Creating a new routine within a cryptographic algorithm is generally difficult because it changes the method's security output. However, it appears that there are a few options for this scenario. Any modification to make a unique proof reduces the degree of freedom (DoF) in the proof part. The system can remain secure if the same amount of DoF is increased elsewhere. To illustrate, suppose that the DoF in the proof is reduced so that the proof is made unique for a given fixed input. We now aim to discuss two possible directions. First, we add the same degree of DoF to the random value; the VCT function may be made to accommodate this modification, which slightly increases the size of the random value. Second, we may take the approach of increasing the same degree of DoF in the private key and this will slightly increase the size of the private key. A right balance shall be found to ensure that such a change does not compromise security.

D. Cryptanalysis for PQ secure VRF

The security of the proposed VRF can be evaluated using routine methods [3][4] such as (a) uniqueness analysis, (b) collision resistance analysis, and (c) pseudo-randomness analysis. Our proposed VRF candidate's performance can be measured by analyzing the time required for making the proposed signature, followed by the hashing time and the overall time, which includes key generation, sign, proof, verification, and building a block [13][51]. Other factors, such as the size of private and public keys, length of signature and hash, computational complexity, and energy consumption, can all be included to evaluate efficiency [16].

E. How to Make a VCT function from the VRF Output?

Given a new good VRF function, we aim to utilize it to create the VTC function. The VCT function accepts the result of the VRF function as input and produces a binary output, pass or fail. Namely, the output space of VRF can be partitioned into two. VRF output is a number. To create a VCT, all we need is a threshold. If VRF output is a number larger than or equal to the threshold, then VCT is set to give a pass output. The probability of pass can then be controlled by raising or lowering the threshold. The probability of pass can be set for any particular energy efficiency goal.

In addition, the probability of pass can be weighted based on the option the network designer can choose. For example, the designer can choose to enable a PoS option. A node's probability of passing can be determined by the stake it has made on chain.

F. PQ Safe Error-Correction Codes PoW

For the VC part, we aim to use ECCPoW for its ASIC resistance, simplistic, time-varying, PQ ready properties. ECCPoW is a new VC method the first author has developed and published [30][31][45]. It is based on an error-correction code called low-density parity check (LDPC) code. It works like an error-correction code (ECC) based cryptosystem. The resistance characteristic of ECC cryptosystems to quantum Fourier sampling attacks has been demonstrated [11]. Faster and more secure ECC cryptosystems are still under study [1]. Durandal, for example, is a lightweight and secure rank-metric code-based cryptosystem [2].

In summary, we aim to extend ECCPoW in two ways. The first is to make it a Green Bitcoin suitable (see M1–M6). The second is to make it PQ safer using medium-density codes. We also aim to ensure that these extensions are safe from well-known security threats.

(**Difficulty Control Algorithm**) The tradeoff relationship between the amount of verifiable computation and energy expenditure can be precisely determined. This tradeoff relationship can be used to devise a difficulty control (DC) algorithm. The DC algorithm seeks to create blocks in a defined regular (average) interval while reacting to variation in the total number of participating p2p nodes over time.

Here, we discuss a way to make ECCPoW PQ safer. In ECCPoW, LDPC codes are used to generate time-varying crypto puzzles. We aim to replace it with moderate-density parity check (MDPC) codes and make ECCPoW PQ safer. LDPC codes do not possess any algebraic structure but only a simple combinatorial

property, i.e., sparsity in the parity-check matrix; this makes them postquantum secure. There have been various suggestions to make a McEliece scheme using LDPC codes [3][4][39][40]. The low-weight parity-check rows in the party-check matrix correspond to low-weight codewords in dual codes. As such, sparsity can be utilized to draw cryptographic attacks. Consequently, MDPC codes have been proposed in which the density is increased about ten times. Furthermore, a quasi-cyclic structure was devised for shorter public and private keys [43]. One famous example of this cryptosystem is BIKE, one of the third-round algorithms in NIST Post-Quantum cryptography (PQC) standardization [1].

G. BFT and Energy Consumption Efficiency of Green Bitcoin

Each node runs the VCT, sees the outcome either as pass or fail, and advances itself to the verifiable computation stage if it is passed. Green Bitcoin, therefore, supports a BFT of 1/2. Suppose a base set of a certain size for Green Bitcoin nodes. To launch a 51% double-spending attack, the attacker must hold 51% of the seats on the SolComp Committee. Using VCT only decreases the overall size of the committee but does not affect its proportion. This necessitates the adversary to have 51% presence in the base set to launch a 51% attack. This remains valid regardless of VCT's pass probability.

Green Bitcoin can choose a certain pass probability (PP) to determine its network's energy consumption efficiency (ECE). An ECE of 90% can be achieved when PP is set to 10%.

VI. TESTBED AND DISCUSSION

A. Testbed

The Green Bitcoin protocol suite will be developed on an existing open-source Bitcoin suite. The opcode table will be upgraded with the developed Green Bitcoin cryptography, replacing elliptic curve encryption for sign and verification. We will present emulation results on a proof-of-concept (PoC) network with a larger number of p2p computers. The Amazon Web Services will be used. Nodes will be scattered around the globe. We aim to employ more than a thousand nodes. Quantum attacks [18] will be used to assess the security of the PoC network.

Similar to our ECCPoW implementation [30][45], we will use the C++ as the developer language. A new genesis block will be created. To build a client, Green Bitcoin will be selected as the consensus algorithm. After defining the chain ID and exporting the genesis, the test Green Bitcoin network will be launched.

B. Safeguarding against Profitable Double-Spending Attacks

The most secure PoW protocols are still vulnerable. Double-spending attacks are still possible to occur [27][33]. This problem is exacerbated when the network's computational power is small. By borrowing computational resources from a mining rig lending site, an attacker can launch a DS assault. If there is a profit-taking opportunity, an attack is possible. Such an opportunity opens up as long as profits overwhelm the costs. The key new finding in [28] is that profitable double-spending (PDS) attacks can occur even if honest nodes have more than 50% of the computational resources of the network. The attacker can attempt to double-spend a transaction whose stake exceeds the cost of leasing mining machines from a lending service. Green Bitcoin aims to safeguard networks against PDS attacks. Such assaults cannot be completely forbidden, but they can be discouraged by lowering the profit the attacker can make and increasing the cost the attacker must bear. Micropayments are not affected by this, but large transactions require attention. We will develop new APIs based on [28] and use them to secure large transactions. These will be made available to the global research community.

C. Discussion

Bitcoin improves self-sovereignty of individuals. One can move along with one's stored wealth in Bitcoin anywhere in the world. No powerful entity can confiscate the portable wealth stored in Bitcoins. One can memorize the pass praise, move to a new country, and restore one's wallet. One does not need to worry about bandits on the travel route or a government's confiscation at the entry point of a port. One can make international payments anywhere in the world using Bitcoin. Currently, Bitcoin faces energy concerns. Green Bitcoin resolves this with its VCT and PoolResistantComp mechanism. Bitcoin faces quantum computer risk. Green Bitcoin provides PQ secure computations. Green Bitcoin is a new protocol designed to retain the merits of Bitcoin and addresses the two pressing concerns of Bitcoin.

VII. CONCLUSION

Imagine a society where sound money is restored and prevails; suppose there is no cheap money. In such a society, when a man has money in his savings account, it means that he has done something beneficial to others in the past. Because this was the only way for him to earn sound money. He earned it. He was paid by the payer. It means he satisfied the payer. To satisfy the payer, he must have worked hard to produce useful goods and services. He reached his silver age. He no longer can work and produce. But he has kept sound money in his wallet. It has retained good value because it is a sound money. He can use it to purchase the valuable goods and services he needs. He does not need a government for that. He could purchase a house in a nice neighborhood and put food on the table for his family. Unearned income disappears in a world dominated by sound money, and honest work prevails.

We made two technical proposals to enhance Bitcoin. The new bitcoin (Green Bitcoin) addresses two issues: reduction of energy consumption and post-quantum computer risk. In the future, we plan to complete the Green Bitcoin protocol and bring forth a Green Bitcoin network to life. Green Bitcoin is a proposal to upgrade Bitcoin, the best sound cryptocurrency, into its quantum-safe and energy-efficient version.

REFERENCES

[1] G. Alagic, J. Alperin-Sheriff, D. Apon, D. Cooper, Q. Dang, J. Kelsey, Y. K. Liu, C. Miller, D. Moody, R. Peralta, and R. Perlner, Status Report on the Second Round of the NIST Post-Quantum Cryptography Standardization Process, National Institute of Standards and Technology Interagency or Internal Report 8309 (NISTIR 8309), 2020.

[2] N. Aragon, O. Blazy, P. Gaborit, A. Hauteville, and G. Zémor, "Durandal: A rank metric based signature scheme," *Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 728-758, Springer, 2019.

[3] M. Baldi, M. Bodrato, and F. Chiaraluce, "A new analysis of the McEliece cryptosystem based on QC-LDPC codes," *Proceedings of the 6th International Conference on Security and Cryptography for Networks*, Springer, pp. 246–262, 2008.

[4] M. Baldi and F. Chiaraluce, "Cryptanalysis of a new instance of McEliece cryptosystem based on QC-LDPC codes, *Proceedings of the IEEE Int. Symposium on Information Theory*, pp. 2591 –2595, 2007.

[5] C. Beekhuizen, "Ethereum's energy usage will soon decrease by ~99.95%," *Ethereum Foundation Blog*, May 18, 2021. https://blog.ethereum.org/2021/05/18/country-power-no-more/.

[6] T. Carlisle, "Chart of the DJIA priced in gold: Buffett and gold redux," November 18, 2009. https://greenbackd.com/2009/11/18/chart-of-the-djia-priced-in-gold/.

[7] A. Castor, "Why Ethereum is switching to proof of stake and how it will work," March 4, 2022. https://www.technologyreview.com/2022/03/04/1046636/ethereum-blockchain-proof-of-stake/.

[8] T. Dar, "Lighthouse confirms prototype merge between Ethereum and ETH2.0 with 99.98% less energy consumption," March 26, 2021.https://crypto8.com/2021/03/26/lighthouse-confirms-prototype-merge-between-ethereum-and-eth-2-0-with-99-98-less-energy-consumption/.

[9] N. De and D. Nelson, "SEC approves Bitcoin futures ETF, Opening crypto to wider investor base," *CoinDesk*, October 16, 2021.

[10] E. Deirmentzoglou, G. Papakyriakopoulos, and C. Patsakis, "A survey on long-range attacks for proof-of-stake protocols," *IEEE Access*, vol. 7, pp. 28712-28725, 2019.

[11] H. Dinh, C. Moore, and A. Russell, "McEliece and Niederreiter cryptosystems that resist quantum Fourier sampling attacks," *Proceedings of the 31st Annual Conference on Advances in Cryptology*, Springer-Verlag, Berlin, Heidelberg, pp. 761–779, 2011.

[12] D. Dolev, and H. R. Strong, "Authenticated algorithms for Byzantine agreement," *SIAM Journal on Computing*, vol. 12, no. 4, pp. 656-666, 1983.

[13] M. F. Esgin, V. Kuchta, A. Sakzad, R. Steinfeld, Z. Zhang, S. Sun, and S. Chu, "Practical post-quantum few-time verifiable random function with applications to Algorand," *Proceedings of the International Conference on Financial Cryptography and Data Security*, pp. 560-578. Springer, Berlin, Heidelberg, 2021.

[14] Ethereum Foundation, "The Merge," July 28, 2022. https://ethereum.org/en/upgrades/merge/.

[15] I. Eyal and E. G. Sirer, "How to disincentivize large Bitcoin mining pools," June 18, 2014. https://hackingdistributed.com/2014/06/18/how-to-disincentivize-large-bitcoin-mining-pools/.
[16] M. S. Ferdous, M. J. M. Chowdhury, M. A. Hoque, "A survey of consensus algorithms in public blockchain systems for crypto-currencies," *Journal of Network and Computer Applications*, vol. 182, 2021.
[17] T. M. Fernández-Caramès and P. Fraga-Lamas, "Towards post-quantum blockchain: A review on blockchain cryptography resistant to quantum computing attacks," *IEEE Access*, vol. 8, pp. 21091-21116, 2020.
[18] E. Gabidulin, A. Ourivski, B. Honary, and B. Ammar, "Reducible rank codes and their applications to cryptography," *IEEE Transactions on Information Theory*, vol. 49, no. 12, pp. 3289–3293, 2003.
[19] E. Gabidulin, A. Paramonov, and O. Tretjakov, "Ideals over a non-commutative ring and their applications to cryptography," *Lecture Notes in Computer Science*, vol. 547, Springer Verlag, 1991.
[20] P. Gaborit, "Shorter keys for code based cryptography," *Proceedings of the 2005 International Workshop on Coding and Cryptography*, pp. 81–90, 2005.
[21] G. Gilder, Scandal of Money: Why Wall Street Recovers but the Economy Never Does, Regnery Publishing, 2016.
[22] F. A. Hayek, "Denationalization of Money, The Institute of Economic Affairs," 1976. https://iea.org.uk/wp-content/uploads/2016/07/Denationalisation%20of%20Money.pdf.
[23] D. Held, "Planting Bitcoin–species: Sound money (sanum pecuniam)," October 26, 2018. https://danhedl.medium.com/planting-bitcoin-sound-money-72e80e40ff62.
[24] D. Held, "Bitcoin, the purchasing power preserver," Bitcoin Magazine," July 11, 2022. https://Bitcoinmagazine.com/markets/Bitcoin-purchasing-power-preserver.
[25] M. E. Herradi and A. Leroy, "Monetary Policy and the Top 1%: Evidence from a Century of Modern Economic History," *International Journal of Central Banking*, vol. 18, no. 5, pp. 237-277, 2021.
[26] IBM Newsroom, "IBM Unveils New Roadmap to Practical Quantum Computing Era; Plans to Deliver 4,000+ Qubit System," May 10, 2022. https://newsroom.ibm.com/2022-05-10-IBM-Unveils-New-Roadmap-to-Practical-Quantum-Computing-Era-Plans-to-Deliver-4,000-Qubit-System.
[27] J. Jang and H.-N. Lee, "Profitable double-spending attacks," *Applied Sciences*, vol. 10, p. 8477, 2020.
[28] H. Janwa and O. Moreno, "McEliece public key cryptosystems using algebraic geometric codes," *Designs, Codes and Cryptography*, vol. 8, pp. 293–307, 1996.
[29] T. J. Jordan, "What constitutes sound money?," *Economic Conference, Progress Foundation*, Swiss National Bank, Zurich, October 8, 2020, https://www.snb.ch/en/mmr/speeches/id/ref_20201008_tjn.
[30] H. Jung, and H.-N. Lee, "ECCPoW: Error-correction code based proof-of-work for ASIC resistance," *Symmetry*, vol. 12, no. 6, pp. 988, 2020.
[31] H. Kim, J. Jang, S. Park, and H.-N. Lee, "Error-correction code proof-of-work on Ethereum," *IEEE Access*, vol. 9, pp. 135942-135952, 2021.
[32] L. Lamport, R. Shostak, and M. Pease, "The Byzantine generals problem," *ACM Transactions on Programming Languages and Systems*, vol. 4, no. 3, pp. 387–389, 1982.
[33] H.-N. Lee, "Blockchain and future society," *4th Lecture Module: PoW Success Probability and Alone-Impossible-Together-Possible Theory*, October 31, 2019. Massive Open Online Course Lecture Video https://gist.edwith.org/Bitcoin-tech and lecture note https://infonet.gist.ac.kr/?page_id=8954.
[34] P. Lee and E. Brickell, "An observation on the security of McEliece's public key cryptosystem," *Lecture Notes in Computer Science*, vol. 330, LNCS, pp. 275–280, Springer Verlag, 1989.
[35] W. Li, S. Andreina, J.-M. Bohli, and G. Karame, "Securing proof-of-stake blockchain protocols," *Data Privacy Management, Cryptocurrencies and Blockchain Technology*, pp. 297-315. Springer, 2017.
[36] V. Lyubashevsky, L. Ducas, E. Kiltz, T. Lepoint, P. Schwabe, G. Seiler, D. Stehlé, and S. Bai, "Crystals-dilithium," *NIST Post-Quantum Cryptography Standardization*, 2017.
[37] R. McEliece, "A public key cryptosystem based on algebraic coding theory," *DSN Progress Report*, pp. 42-44:114–116, 1978.
[38] B. McElrath, "What's wrong with proof-of-stake?" April 13, 2022. https://medium.com/@BobMcElrath/whats-wrong-with-proof-of-stake-77d4f370be15.
[39] R. Misoczki, J.-P. Tillich, N. Sendrier, and P. L. S. M. Barreto, "MDPC-McEliece: New McEliece variants from moderate density parity-check codes," *Proceedings of the IEEE International Symposium on Information Theory*, pp. 2069–2073, Istambul, Turkey, 2013.

[40] C. Monico, J. Rosenthal, and A. Shokrollahi, “Using low density parity check codes in the McEliece cryptosystem,” *Proceedings of the IEEE International Symposium on Information Theory*, p. 215, 2000.
[41] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System,” 2008. https://bitcoin.org/bitcoin.pdf.
[42] H. Niederreiter, “Knapsack-type cryptosystems and algebraic coding theory,” *Problems of Control and Information Theory*, pp. 15:19–34, 1986.
[43] A. Otmani, J. Tillich, and L. Dallot, “Cryptanalysis of two McEliece cryptosystems based on quasi-cyclic codes,” *Mathematics in Computer Science*, vol. 3, no. 2, pp. 129–140, 2010.
[44] R. Overbeck, “A new structural attack for GPT and variants,” *Lecture Notes in Computer Science*, vol. 3715, pp. 50–63. Springer Verlag, 2005.
[45] S. Park, H. Choi, H.-N. Lee, “Time-variant proof-of-work using error-correction codes,” June 22, 2020. https://arxiv.org/abs/2006.12306.
[46] M. Rosenfeld, “Analysis of bitcoin pooled mining reward systems,” 2011. https://arxiv.org/abs/1112.4980.
[47] V. Sidelnikov, “A public-key cryptosystem based on binary Reed-Muller codes,” *Discrete Mathematics and Applications*, vol. 4, no. 3, pp. 191-207, 1994.
[48] V. Sidelnikov, Michilovich, and S. O. Shestakov, “On insecurity of cryptosystems based on generalized Reed-Solomon codes,” pp. 439-444, 1992.
[49] J. Starn and J. Saul, “Why Bitcoin’s Environmental Problems Are So Hard to Fix,” *Washington Post*, March 29, 2022. https://www.washingtonpost.com/business/energy/why-bitcoins-environmental-problems-are-so-hard-to-fix/2022/03/16/b71e1d22-a4df-11ec-8628-3da4fa8f8714_story.html.
[50] The White House, “National Security Memorandum on Promoting United States Leadership in Quantum Computing While Mitigating Risks to Vulnerable Cryptographic Systems,” May 4, 2022. https://www. whitehouse.gov/briefing-room/statements-releases/2022/05/04/national-security-memorandum-on-promoting-united-states-leadership-in-quantum-computing-while-mitigating-risks-to-vulnerable-cryptographic-systems/.
[51] L. Zengpeng, T. G. Tan, P. Szalachowski, V. Sharma, and J. Zhou, “Post-quantum VRF and its applications in future-proof blockchain system,” 2021. https://arxiv.org/abs/2109.02012.
[52] X. Zhang, W. Wu, S. Yang, and X. Wang, “Falcon: a blockchain-based edge service migration framework in MEC,” *Mobile Information Systems*, 2020. https://doi.org/10.1155/2020/8820507.
[53] “Lightning network,” July 1, 2022. https://lightning.network/.
[54] “Mining: Bitcoin takes more energy than mining gold,” Nature, November 6, 2018. https://doi-org.libproxy.gist.ac.kr/10.1038/d41586-018-07283-3.
[55] “Nixon Shock”, August 12, 2022. https://en.wikipedia.org/wiki/Nixon_shock.
[56] “Why sharding is great: demystifying the technical properties,” July 28, 2022. https://vitalik.ca/general/2021/04/07/sharding.html.

Heung-No Lee received the B.S., M.S., and Ph.D. degrees from the University of California, Los Angeles, CA, USA, in 1993, 1994, and 1999, respectively, all in electrical engineering. He worked at HRL Laboratories, LLC, Malibu, CA, USA, from 1999 to 2002 as a Research Staff Member. From 2002 to 2008, he was Assistant Professor with the University of Pittsburgh, PA, USA. Since 2009, he has been with the School of Electrical Engineering and Computer Science, GIST, South Korea, where he is a full professor. His technical works are in information theory, signal processing theory, communications/networking theory, and their applications to Wireless Communications, Networking, Medical Imaging, Brain-Computer Interfaces, Spectroscopy, Cryptocurrencies and Decentralized Finance.

Young-Sik Kim received the B.S., M.S., and Ph.D. degrees in electrical engineering and computer science from Seoul National University, in 2001, 2003, and 2007, respectively. He joined the Semiconductor Division, Samsung Electronics, where he worked in the research and development of security hardware IPs for various embedded systems, including modular exponentiation hardware accelerator (called Tornado 2MX2) for RSA and elliptic-curve cryptography in smartcard products and mobile application processors, until 2010. He is currently a professor with Chosun University, Gwangju, South Korea. He is also a submitter for two candidate algorithms (McNie and pqsigRM) in the first round for the NIST Post Quantum Cryptography Standardization and a submitter for three candidate algorithms for Korean Post-Quantum Cryptography Competition (KpqC).

Dilbag Singh received a Ph.D. degree in computer science and engineering from the Thapar Institute of Engineering and Technology, Patiala, India, in 2019. He was an Assistant Professor at Chandigarh University, Mohali, India; Manipal University Jaipur, Jaipur, India; and Bennett University, Greater Noida, India. In 2021, he joined the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju, South Korea, where he is currently affiliated. His research interests include image processing, computer vision, deep learning, metaheuristic techniques, and information security. He was in the top 2% list issues by "World Ranking of Top 2% Scientists" in 2021.

Manjit Kaur received the M.E. degree in information technology from Punjab University, Chandigarh, India, in 2011, and the Ph.D. degree from the Thapar Institute of Engineering and Technology, Patiala, India, in 2019. She was an Assistant Professor at Chandigarh University, Mohali, India; Manipal University Jaipur, Jaipur, India; and Bennett University, Greater Noida, India. In 2021, she joined the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju, South Korea, where she is currently affiliated. Her research interests include wireless sensor networks, digital image processing, and metaheuristic techniques. She was in the top 2% list issues by "World Ranking of Top 2% Scientists" in 2021.

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 25, 2021, accepted September 7, 2021, date of publication September 16, 2021, date of current version October 11, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3113522

Error-Correction Code Proof-of-Work on Ethereum

HYOUNGSUNG KIM[1], JEHYUK JANG[1], SANGJUN PARK[2], AND HEUNG-NO LEE[1], (Senior Member, IEEE)

[1]School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea
[2]Electronics and Telecommunications Research Institute (ETRI), Gwangju 500-712, South Korea

Corresponding author: Heung-No Lee (heungno@gist.ac.kr)

This research was supported in part by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2021-0-01835) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation), in the part by the IITP Grant through Korean Government MSIT under Grant 2020-0-00958, and in part by the National Research Foundation of Korea (NRF) Grant through Korean Government MSIP under Grant NRF-2021R1A2B5B03002118.

ABSTRACT The error-correction code proof-of-work (ECCPoW) algorithm is based on a low-density parity-check (LDPC) code. ECCPoW can impede the advent of mining application-specific integrated circuits (ASICs) with its time-varying puzzle generation capability. Previous research studies on ECCPoW algorithm have presented its theory and implementation on Bitcoin. In this study, we have not only designed ECCPoW for Ethereum, called ETH-ECC, but have also implemented, simulated, and validated it. In the implementation, we have explained how ECCPoW algorithm has been integrated into Ethereum 1.0 as a new consensus algorithm. Furthermore, we have devised and implemented a new method for controlling the difficulty level in ETH-ECC. In the simulation, we have tested the performance of ETH-ECC using a large number of node tests and demonstrated that the ECCPoW Ethereum works well with automatic difficulty-level change capability in real-world experimental settings. In addition, we discuss how stable the block generation time (BGT) of ETH-ECC is. Specifically, one key issue we intend to investigate is the finiteness of the mean of ETH-ECC BGT. Owing to a time-varying cryptographic puzzle generation system in ECCPoW algorithm, BGT in the algorithm may lead to a long-tailed distribution. Thus, simulation tests have been performed to determine whether BGT distribution is not heavy-tailed and has a finite mean. If the distribution is heavy-tailed, stable transaction confirmation cannot be guaranteed. In the validation, we have presented statistical analysis results based on the two-sample Anderson–Darling test and discussed how the BGT distribution follows an exponential distribution which has a finite mean. Our implementation is available for download at https://github.com/cryptoecc/ETH-ECC.

INDEX TERMS Anderson–Darling test, ASIC-resistant, blockchain, error-correction codes, Ethereum, hypothesis test, LDPC, proof-of-work, simulation, statistical analysis.

I. INTRODUCTION

Blockchain is a peer-to-peer (P2P) network that consists of trustless nodes. In a reliable P2P network, no peers (nodes) would intentionally send wrong information to others. In contrast, in an unreliable P2P network (e.g., a group of trustless nodes), the possibility that some peers may send false information to others should be considered. For example, a node may spread wrong or fake information to others. To address these issues in an unreliable P2P network, Nakamoto proposed using blocks and chaining these blocks with a novel consensus algorithm [1].

In a blockchain, a peer sends a new block containing transactions to other peers. These peers validate the received block and link it to the previous block when there is no problem in the received block, i.e., when the authenticity of the block has been verified. A consensus algorithm is used to accomplish this verification task. If a peer has sent false information to others, such information is detected by the consensus algorithm as there is no collusion among the peers. A generated block contains information about previous blocks, i.e., all blocks are chained; thus, if someone wants to change one block in a chain, all previous blocks of the block to be changed must also be changed. Therefore, unless the network is centralized within a particular group, sending fake information about previous blocks to new peers is impossible.

The associate editor coordinating the review of this manuscript and approving it for publication was Giambattista Gruosso.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

Therefore, to prevent collusion, an unreliable network should avoid centralization.

Nakamoto proposed a proof-of-work (PoW) system for a consensus algorithm. In the PoW system, peers repeat a type of work to solve a cryptographic puzzle using a hash function (e.g., SHA256 [1] and Keccak [2]). When a peer successfully solves a cryptographic puzzle, the peer generates a block. In addition, the peer gets an incentive as a reward for the work done. In an ideal PoW system, new nodes can join to work and receive as much reward as they complete work. However, with an increase in the price of reward, attempts have been made to centralize the network to monopolize incentives.

Centralization is a phenomenon that occurs in PoW-based blockchain networks. In blockchains using PoW as a consensus algorithm, an oligarchy of miners with a disproportionate share of computation resources can monopolize block generation. Such centralization negatively impacts the credibility of a blockchain. For example, in a centralized network, a group of dominant nodes can selectively filter out some transactions belonging to others for their benefit. New nodes will find it difficult to earn trust and join the network in the fear of possible unfair treatment [3], [4].

The emergence of application-specific integrated circuits (ASICs) has accelerated the centralization of PoW. As more nodes use ASICs in generating blocks, the computation complexity in block generation increases. Thus, it has become difficult to generate blocks using general-purpose units, such as a central processing unit (CPU) and a graphics processing unit (GPU). As a result, a few groups equipped with powerful ASICs have surfaced and centralized the blockchain networks. To avoid centralization, researchers have proposed the use of ASIC-resistant PoW (e.g., Ethash of [2], X11 of [12], and Random X of [24]) and alternative consensus algorithms (e.g., proof-of-stake, delegated proof-of-stake, and Byzantium fault tolerance [25]). Networks using alternative algorithms have presented lesser decentralization effects than those have using ASIC-resistant PoW [25]. Specifically, in networks using alternative algorithms, only limited participants can generate blocks, but ASIC-resistant PoW has no limit on the number of participants. Thus, ASIC-resistant PoW presents a more decentralized network than do alternative algorithms.

For an ASIC-resistant PoW, an error-correction code proof-of-work (ECCPoW) algorithm was proposed [6], [7]. In ECCPoW algorithms, a hash value of a previous block generates a varying parity-check matrix (PCM) for error correction. This varying PCM works as a cryptographic puzzle in ECCPoW. These time-varying cryptographic puzzles make ECCPoW ASIC resistant. It is possible to use an ASIC for a specific fixed cryptographic puzzle. In ECCPoW, every newly created puzzle differs from all the previously created puzzles. As a result, if there was an ASIC for ECCPoW, such an ASIC must cover a wide range of cryptographic puzzle generation systems. Such a system, however, would incur huge chip space and cost [10], [11].

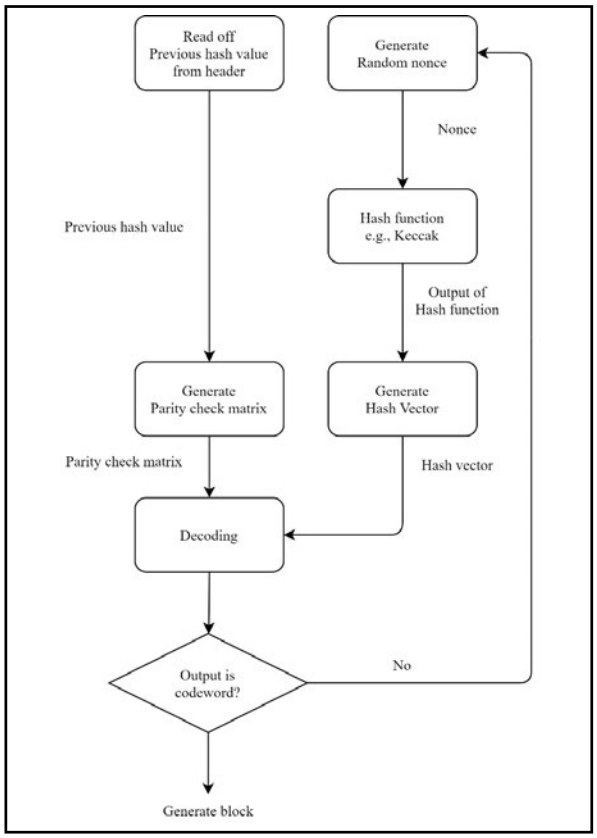


FIGURE 1. Flowchart of ECCPoW Ethereum. Every miner who generates blocks can construct a parity check matrix using a previous hash value. A generated nonce becomes an input of a hash function. A hash vector used for decoding can be generated using the output of a hash function. If decoding is successful, the block is generated; otherwise, a miner generates a new nonce to make a new hash vector for decoding.

In [7], the authors have reported that the time-varying puzzle system may generate large block generation time (BGT), i.e., outliers, for ECCPoW implemented on Bitcoin. If outliers occur frequently, it is of our interest to see whether or not the distribution of BGT is heavy-tailed with an infinite mean [15], [26]. As a result, the proposition made in [6] that BGT has a finite mean needs to be challenged. Previous works on ECCPoW [6], [7] did not include sufficient real-world experiments to conclude that BGT has a finite mean. If BGT does not have a finite mean, ECCPoW cannot be used as a consensus algorithm. In this paper, we aim to study the distribution of BGT of ECCPoW implemented on Ethereum (ETH-ECC). Our experimental results show that the BGT distribution is not heavy-tailed and has a finite mean.

The contributions of our work are as follows:

- We show how ECCPoW is implemented on Ethereum.
- We present a method for controlling the difficulty level in ETH-ECC and report the results of automatic difficulty level change with real-world experiments of ETH-ECC.
- We present a goodness-of-fit result using the Anderson–Darling (AD) test for distribution validation and discuss

the BGT distribution follows the exponential distribution which has a finite mean.

The remainder of this paper is organized as follows. Section II provides a background of the requirements of an ASIC-resistant PoW. Section III demonstrates the implementation of ETH-ECC. Section IV discusses the formulation of the problem. Section V provides the experimental result of the implementated ETH-ECC. Finally, Section VI summarizes our work and concludes the paper.

II. BACKGROUND

We introduce three approaches that can be used to avoid centralization problems in PoW. The first is an intentional bottleneck between an arithmetic logic unit (ALU) and memory, which is used by Ethash of Ethereum [2], [5]. It is also termed a memory-hard technique. The second is the *high complexity of ASIC design* used by Dash [12], Raven [13], and our method, ECCPoW. The third is *hybrid methods* of two methods; Random X of Monero uses *hybrid methods* [24].

A. INTENTIONAL BOTTLENECK

The most known PoW of the intentional bottleneck is Ethash of Ethereum [2], [5]. This method uses the difference between the throughput of ALU and the bandwidth of the memory. If there is a bottleneck between the ALU and memory, it is impossible to use the entire throughput of ALU. Specifically, if a miner needs to obtain data from memory to generate a block, the number of block generation attempts is determined by memory bandwidth. Ethash uses a directed acyclic graph (DAG), which is a set of randomly generated data for the bottleneck. The DAG is a huge dataset that cannot be stored in a cache memory; therefore, the DAG is stored in memory. To generate a block using Ethash, a miner must mix a part of the DAG that is stored in the memory. Owing to this procedure, the miner cannot avoid the bottleneck because of limited memory bandwidth. This method has been ASIC resistant for a long time; however, Bitmain released ASIC for Ethash in 2018.

B. HIGH COMPLEXITY OF ASIC DESIGN

Because of the high complexity of ASIC design, ASICs are less efficient. For example, if ASICs are less efficient than a general-purpose unit such as CPU or GPU, there is no reason to design ASIC. X11 of Dash [12] and X16R of Raven [13] use this method. Unlike PoW of Bitcoin, which uses only one hash function (SHA-256), X11 uses 11 hash functions consecutively: BLAKE, BMW, Grosetl, JH, Keccak, Skein, Luffa, Cubehash, SHAvite-3, SIMD, and ECHO. The BLAKE, which is the first hash function of X11, uses a block header with nonce as inputs; its output becomes the input of the next hash function. Similarly, the next hash function uses the output of the previous hash function. This procedure is repeated until a result is obtained for the last hash function. Miners determine whether they have found a valid nonce using the output of the final hash function.

Designing an ASIC for X11 was expensive; therefore, X11 was ASIC resistant. However, Bitmain released an ASIC for X11 in 2016. There are a few PoW algorithms that extend X11 (e.g., X13, X14, and X15); however, the ASICs for these have been released. X16R of Raven is an extended version of X11 of Dash. In X16R, unlike the previous extension of X11, the sequence of 16 hash functions is randomly changed. Therefore, it is costly to design an ASIC for X16R. However, T. Black, who designed X16R, mentioned that there is some evidence that ASICs for X16R exist [23]. Our ECCPoW employs a *time-varying puzzle generation system* to make ASCI design difficult. ECCPoW can make ASIC powerless as the puzzle generation system changes from block to block. We explain this further in Section III.

C. HYBRID METHODS

Random X of Monero combines the above two methods. Random X uses memory-hard techniques for the bottleneck with random code execution; Random X is optimized for CPU mining [24]. In [24], they mentioned that mining can be performed using a field-programmable gate array; however, it will be much less efficient than CPU mining. It implies that efficient mining hardware can be developed when the cost of developing chipsets is low in comparison to the mining reward. With the proposed ECCPoW, attempts in developing efficient mining hardware can be made when the reward-to-cost ratio increases. However, such attempts can be easily evaded since the parameters of ECCPoW can be easily changed, such as increasing the length of code and the code rate. The next section illustrates further the ASCI-resistance characteristic of ECCPoW.

III. ECCPoW IMPLEMENTED ON ETHEREUM

In this section, we briefly introduce ECCPoW and present how ECCPoW has been implemented on Ethereum using Fig. 1. Furthermore, we present how the difficulty level of ETH-ECC is automatically controlled.

A. OVERVIEW OF ECCPoW

In a blockchain employing the PoW consensus algorithm, a node solves cryptographic puzzles to publish a block. For a given puzzle, the node who solves the puzzle first obtains the authority to publish a block. For example, in the PoW of Bitcoin, the first node that finds a specific output of the secure hash algorithm (SHA) obtains the authority to publish a block. The PoW of Ethereum uses Keccak instead of SHA. The ECCPoW algorithm proposed in [6] is a PoW consensus algorithm that uses error-correction code, which comprises the low-density parity-check (LDPC) code [8], as a cryptographic puzzle. The ECCPoW algorithm consists of a pseudo-random puzzle generator (PRPG) and an ECC puzzle solver. Fig. 1 presents the flowchart of the ECCPoW algorithm. For every block, the PRPG generates a new pseudo-random LDPC matrix. A new LDPC matrix is distinct from the other previously generated matrices. Such a pseudo-random LDPC matrix takes the role of issuing an

independently announced cryptographic puzzle. The ECC puzzle solver uses the LDPC decoder to solve the given announced puzzle. Specifically, to publish a block, a node is required to run through an input header until the LDPC decoder hits a satisfying result; for instance, the output of the decoder is an LDPC codeword (with a certain Hamming weight). In the next subsection, we will discuss ECCPoW implementation on Ethereum with the flowchart presented in Fig. 1.

B. COMPARISON OF ETHASH AND ECCPoW

Ethereum uses `Ethash` for ASIC resistance, and `ETH-ECC` uses ECCPoW for ASIC resistance. In this subsection, we present how `Ethash` and `ETH-ECC` apply ASIC-resistance property to PoW with pseudo-codes.

`Ethash` uses a DAG for ASIC resistance. The DAG is a large size of data and is typically stored in a random access memory (RAM), not in cache memory. It implies that a miner must access the RAM to get the DAG data. Although the miner could be equipped with a high-throughput ALU, the bandwidth access from the RAM to the ALU is limited. That is, the bottleneck is the limited bandwidth of reading DAG information from the RAM; thus, any fast ALU, e.g., an ASIC implementation of `keccak512`, exceeding this bottleneck is of no use. This makes Ethash ASIC resistant.

When a miner reads DAG data from the RAM, the location where the data are read varies. The location of data reading is selected by the "mix"; the mix is a 128-byte hash value generated by the block header and a nonce. The mix is updated using the Fowler-Noll-Vo (FNV) hash function. The miner repeats this process 64 times. After updating the mix, the miner compresses the mix; for compression, the FNV hash is used again. The miner returns a hash value of the result of concatenating the compressed mix and the seed.

If this hash value is less than the desired target, the nonce is validated, and a new block is linked to the previous block. **Algorithm 1** denotes the pseudo-code of `Ethash`.

Algorithm 1 `Ethash`

Require: block header (BH), nonce, DAG

1: Initialize seed: seed = `keccak512`(BH, nonce)
2: Initialize 128 bytes mix:
   mix = `concatenate`(seed, seed)
3: **for** i = 0, 1, 2, . . . ,63:
4: Get data from DAG using mix:
   data = `DAG_lookup`(DAG, mix, i)
5: update mix: mix = `FNV_hash`(mix, data)
6: **end for**
7: **for** i = 0, 4, 8, . . . , `length`(mix):
8: Compress mix: cmix = `compress_mix`
   (mix, i)
9: **end for**
10: return `keccak256(concatenate`(seed, cmix))

`Ethash` uses *the intentional bottleneck* for ASIC resistance, but `ETH-ECC` aims to use *a time-varying puzzle generation system* for ASIC resistance. In `ETH-ECC`, two factors make the design of ASICs very difficult. One is flexible code lengths and randomly generated PCMs. The `ECC_puzzle_solver` generates a hash vector of length-n (subsection C) using a nonce; this n determines the code length. The development of an ASIC for a PCM with length n cannot be realized, as the `ETH-ECC` network changes n and the PCM from one block to another block. The `PRPG` creates a PCM $\mathbf{H}$. A PCM uses a BH as a seed; thus, it is randomly generated. All miners that work to extend the same previous block use the same PCM to solve the ECCPoW puzzle. Thus, it is highly expensive, if not impossible, to implement an ASIC that can handle a time-varying PCM [10], [11]. After generating a hash vector and a PCM, a miner works out how to generate an output word. If this output word satisfies a specific condition, the miner is successful at completing ECCPoW; e.g., the output word can be a codeword, and then, a new block is linked to the previous block. **Algorithm 2** denotes the pseudo-code of `ETH-ECC`. In our implementation, we have replaced `Ethash` and all its relevant peripheral systems with ECCPoW; thus, it has the same requirement as `Ethash` except for the DAG. We present more details about `ETH-ECC` in the following subsections.

Algorithm 2 `ETH-ECC`

Require: block header (BH), nonce

1: Generate hash vector:
   hash_vector = `ECC_puzzle_solver`(nonce)
2: Generate parity check matrix: PCM = `PRPG` (header)
3: output_word = `decoder`(PCM, hash_vector)
4: return output_word

C. ECCPoW ON ETHEREUM

In this subsection, we present how the error-correction process is applied to ETH-ECC using Fig. 1.

$$C := \{\mathbf{c}|\mathbf{Hc} = \mathbf{0} \cap \mathbf{c} \in \{0, 1\}^{n\times 1}\} \tag{1}$$

when a PCM $\mathbf{H}$ is given, a code $\mathbf{c}$, satisfying (1), is referred to as an LDPC code. The goal of the ECCPoW algorithm is to find an LDPC code $\mathbf{c}$ using the PCM $\mathbf{H}$, which is derived by PRPG, and a hash vector $\mathbf{r}$, which is obtained using the ECC puzzle solver. For the PRPG, we employ the previous hash value; the previous hash value, known as the parent hash in the Ethereum block header, randomly generates a PCM. Specifically, we use Gallagher's method to create random PCM [9]; we use the previous hash value as a seed of randomness. Thus, PCMs are changed for every block; because every node has the same seed, they use the same PCM until a block is generated [6].

1) ECC PUZZLE SOLVER ON ECCPoW ETHEREUM

Here, we introduce the ECC puzzle solver process in ETH-ECC. Our definitions are based on [6]. The equations below follow the right-hand side of Fig. 1.

Definition 1 (ECC Puzzle Solver): Hash vector $\mathbf{r}$ in which the size of n can be obtained as follows:

$$s_1 := Keccak(nonce) \in \{0, 1\}^{256} \tag{2}$$

where *Keccak* denotes the hash function used in Ethash of Ethereum [5]. We generate a *nonce* in the same way that Ethereum does. Furthermore, for a longer length of a hash vector, we use $s_u := Keccak(s_1) \in \{0, 1\}^{256}$ with $u = 2, 3, \ldots, l + 1$. We slice or concatenate the result of *Keccak* to generate a flexible length hash vector $\mathbf{r}$:

$$\mathbf{r} := \begin{cases} s_1[1:n] & \text{if } n \leq 256 \\ [s_1 \cdots s_l s_{l+1}[1:j]] & \text{if } n > 256 \end{cases} \tag{3}$$

where $l = \lfloor n/256 \rfloor$ and $j = n - 256 \times l$. For example, when n is less than 256, $\mathbf{r}$ obtains the same length as n, whereas when n is not less than 256, $\mathbf{r}$ concatenates the results of *Keccak*. This flexible length hash vector is used for ASIC resistance.

2) PoW OF THE LDPC DECODER

The goal of the LDPC decoder is to find a hash vector $\mathbf{c}$ that satisfies $\mathbf{Hc} = 0$. The definition below explains the decoding presented in Fig. 1.

Definition 2 (Decoder): Given a PCM $\mathbf{H}$, which is the size of $m \times n$, and hash vector $\mathbf{r}$, which is the size of n, are given, the LDPC decoder uses $\mathbf{H}$ and $\mathbf{r}$ as inputs and obtains output $\mathbf{c}$ using the message-passing algorithm [6], [14]. When $\mathbf{c}$ satisfies (1), $\mathbf{c}$ becomes an LDPC code, and a miner completes LDPC decoding.

$$D_{np} : \{\mathbf{r}, \mathbf{H}\} \mapsto \mathbf{c} \in \{0, 1\}^{n \times 1} \tag{4}$$

A PCM $\mathbf{H}$ is randomly generated; however, all miners use the same previous hash value, which is derived from the previous block. Therefore, predicting the next PCM to mine a block in advance is impossible. In the PoW of Ethereum, miners change a nonce when they obtain a wrong output. We follow the same procedure as Ethereum to obtain a hash value from *Keccak* with a *nonce*, but ETH-ECC uses one more step (3) to generate a hash vector for decoding. When the code derived by (4) does not satisfy (1), the miner generates a new *nonce* and repeats all steps.

Our method is based on the *high complexity of ASIC design* in Section II for an ASIC-resistant PoW. However, unlike the mentioned method in Section II, ECCPoW generates varying cryptographic puzzles of *high complexity*. Specifically, ECCPoW uses two factors to achieve *high complexity*: flexible length LDPC code $\mathbf{c}$ and randomly generated PCM $\mathbf{H}$. ASICs can be released for the n length of code. However, extending the length of code (e.g., $n + 1$) makes ASICs powerless. Furthermore, in [10], [11], it has been proven that implementing an ASIC that can handle variable PCMs is expensive and occupies a lot of space. If developing an ASIC costs more than buying a CPU or GPU, there is no incentive to develop an ASIC. In other words, the ECCPoW algorithm is ASIC resistant as implementing an ASIC that can handle various lengths of changing codes and randomly generated PCMs is inefficient.

D. DIFFICULTY-LEVEL CONTROL OF ETH-ECC

In this subsection, we demonstrate the implementation of ETH-ECC's difficulty-level control. Bitcoin [1] and Ethereum [2] have different difficulty-level control methods. Furthermore, we present one way to add fine difficulty control.

In Bitcoin, the Bitcoin network changes the difficulty level every 2016 block; the desired BGT is 10 min for a block. If miners generate a block every 10 min, generating 2016 blocks takes precisely 2 weeks. Thus, if generating 2016 blocks takes more than 2 weeks, the difficulty level decreases; otherwise, the difficulty level increases. Unlike Bitcoin, the Ethereum network changes the difficulty level every block. Ethereum network allows for a block to be generated between 9 and 18 s. If a block is generated within 9 s, then the difficulty level increases. If it exceeds 18 s, then the difficulty level decreases. Because of this difference between Bitcoin and Ethereum, ECCPoW-based Bitcoin (BIT-ECC) and ETH-ECC also have different difficulty-level control methods. Thus, ETH-ECC cannot use BIT-ECC's method. Because of the need for a new method, we demonstrate the implementation of ETH-ECC's difficulty level control with a difference from Ethereum's method.

Ethereum uses the number of attempts to generate a block per second, termed hash rate, and a probability of block generation. Similarly, ETH-ECC uses the hash rate but considers a probability of decoding success. In [5], the difficulty of Ethereum is defined by the probability of block generation. The difficulty is as follows:

$$n \leq \frac{2^{256}}{Diff} \tag{5}$$

It indicates that

$$Diff \leq \frac{2^{256}}{n} \tag{6}$$

where n denotes the result of PoW and *Diff* denotes the difficulty of Ethereum. Thus, (6) means that when the difficulty level increases, the number of n that satisfies (6) decreases. Furthermore, we can consider that the reciprocal of difficulty is a probability of block generation. Ethereum uses this probability and hash rate to control BGT. For example, without replacement, when the probability of block generation is 1/150 and hash rate is 10 hash per second, brute force takes 15 s. If the hash rate increases, such as 20 hash per second, Ethereum's method adjusts the probability of BGT to 1/300. Thus, brute force takes 15 s even though the hash rate increases.

For ECCPoW, if we can calculate a probability of decoding success, it is possible to control the difficulty level similar to the process in Ethereum. Thus, it is important to know the probability of a successful LDPC decoding according to the LDPC parameter. We use the pseudo-probability of a successful LDCP decoding according to the parameters to test the difficulty level change using the BGT [7]. That is, ETH-ECC uses the probability of decoding success and

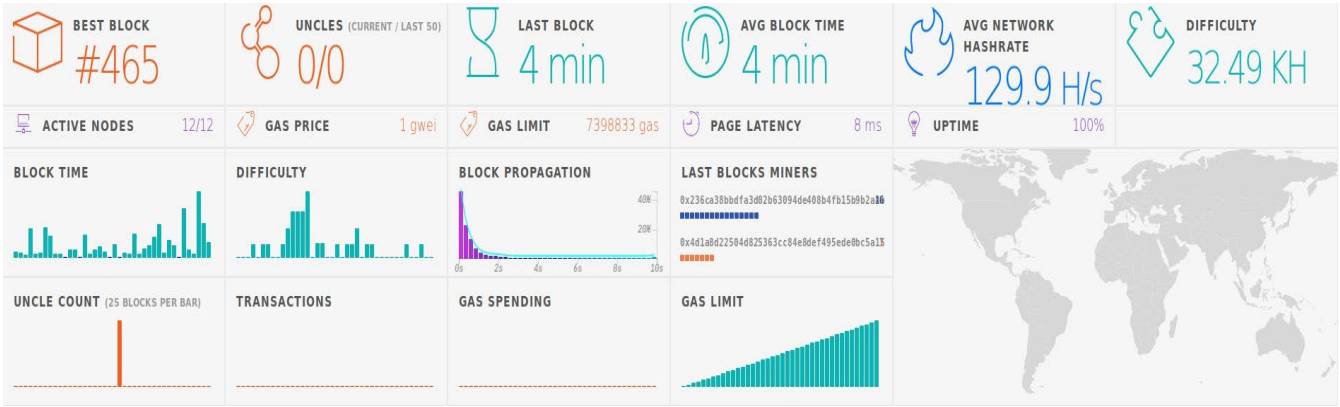


FIGURE 2. This figure shows the simulation results of ECCPoW Ethereum on Amazon Web Services AWS). Twelv nodes are used in the simulation. The tw nodes are bootnodes that help connect th nodes, and the other 10 nodes are sealnodes that participate i block generation. We use the m5.xlarge of AWS EC2 for the simulation. In the charts, BLOCK TIME shows the block generation times for the last 40 blocks, and DIFFICULTY shows the difficulty levels of the last 40 blocks. BLOCK PROPAGATION shows th percentage of the block propagation time corresponding to time.

hash rate to control the difficulty level. For example, without replacement, when the probability of decoding success is 1/150 and the hash rate is 10 hash per second, it takes 15 s, as in the above example of Ethereum's method. However, unlike Ethereum, when the hash rate increase, ETH-ECC tunes parameters of LDPC to adjust the probability of decoding success. By tuning parameters, ECCPoW achieves both difficulty-level control and ASIC resistance. These parameters can be found at https://github.com/cryptoecc/ETH-ECC/blob/master/consensus/eccpow/LDPCDifficulty_utils.go#L65. In Fig. 2, the difficulty of ETH-ECC is 32.49 KH, indicating that the probability of block generation is 1 of 32,490 hash.

One Way to Add Fine Difficulty Control. ECCPoW controls difficulty using integer and discrete variable n. Thus, it may look inappropriate to manage difficulty precisely. However, as the number of blocks increase, block generation time (BGT) converges to the ideal BGT time, which is suitable for a network. For example, when there exist two difficulties: n and $n+1$, we can define average BGT of each difficulty as t_n and t_{n+1}. Thus, we can define the average BGT:

$$averageBGT = \frac{\alpha t_n + \beta t_{n+1}}{k} \tag{7}$$

where α denotes the number of generated blocks with difficulty t_n, β denotes the number of generated blocks with difficulty t_{n+1}, and k denotes the total number of generated blocks (TNGB). Thus, α can be replaced as $\alpha = k - \beta$. As a result, equation (7) is:

$$averageBGT = \frac{(k-\beta)t_n + \beta t_{n+1}}{k} \tag{8}$$

When TNGB k is kept constant, the average BGT is determined by the number of generated blocks β in Eq. (8). Thus, the ideal average BGT, which is suitable for the number of nodes in a network, depends on β. In other words, when TNGB $k$is low, the average BGT cannot meet the ideal average BGT because there are not enough blocks of each difficulty. However, as TNGB k increases, the number of blocks corresponding to the difficulty, such as β, getting closer to the proportion that fits the probability of block generation. As a result, average BGT converges to the ideal average BGT; this convergence confirms our proposition that the network can control difficulty precisely.

IV. PROBLEM FORMULATION

In PoW, there is a case that nodes generate blocks at the same time. Bitcoin allows only one block to be generated at a time; Ethereum allows three blocks to generate at the same time. However, in Ethereum, only one block can be canonical; the other blocks cannot. Blocks that cannot be canonical are called uncle blocks. In Ethereum, nodes rollback transactions of uncle blocks [5]. Therefore, the transaction's participants must wait for block confirmation to prevent a rollback. That is, in the blockchain using PoW, the BGT must have a finite mean for the block confirmation time. For example, if the BGT has an infinite mean, the waiting time for the confirmation of transactions cannot be determined. Therefore, to apply the ECCPoW algorithm in a real network, the BGT must have a finite mean.

In [6], the authors presented the definition of the block generation of the ECCPoW algorithm using a hash rate with a geometric distribution. That is, they assumed that nodes generate a block with specific block generation attempts. However, if the BGT has an infinite mean, there is no guarantee that nodes generate a block with specific attempts. In [7], the authors presented a practical experiment using the ECCPoW algorithm. However, they only mentioned that the BGT of ECCPoW is "unstable." That is, they mentioned that the BGT of ECCPoW has outliers; however, they did not present a discussion on the BGT. Thus, in this study, we present a discussion on the BGT. Specifically, our experimental result presents evidence that the exponential distribution describes the distribution of the BGT of ECCPoW.

V. EXPERIMENT ON ETH-ECC

In this section, we conduct experiments using ETH-ECC. First, we simulate the difficulty level change using multinode networks. Second, we conduct a goodness-of-fit experiment using the AD test [16]–[18] to discuss the distribution of the BGT with a fixed difficulty level.

A. SIMULATION OF THE DIFFICULTY CHANGE

We simulate the difficulty-level change employing Amazon Web Services (AWS) using 12 nodes. Two nodes are *bootnodes* that help connect the nodes, and the other 10 nodes are *sealnodes* that participate in block generation. In the charts presented in Fig. 2, *BLOCK TIME* presents the BGT of the last 40 blocks, and *DIFFICULTY* shows the difficulty level of the last 40 generated blocks. *BLOCK TIME* and *DIFFICULTY* show that because of the large standard deviation, a block is gradually generated despite the low difficulty level, as mentioned in [7]; in the next subsection, we discuss the BGT. In the charts presented in Fig. 2, *LAST BLOCK* shows the BGT of the previous block, and *AVG BLOCK TIME* shows the average of the BGT. In addition, *AVG NETWORK HASHRATE* shows the average hash rate of all miners. *BLOCK PROPAGATION* shows the block propagation time from a miner who generated a block to other miners. We used two different regions: Seoul and US East for *sealnodes*. Specifically, 3 of the 10 *sealnodes* are in the US East region, whereas the rest are in the Seoul region. *BLOCK PROPAGATION* also shows the percentage of blocks that are propagated at corresponding times. *BLOCK PROPAGATION* indicates that the propagation of approximately all blocks between Seoul and US East regions takes less than 2 s. The block propagation is the same method as that of Ethereum.

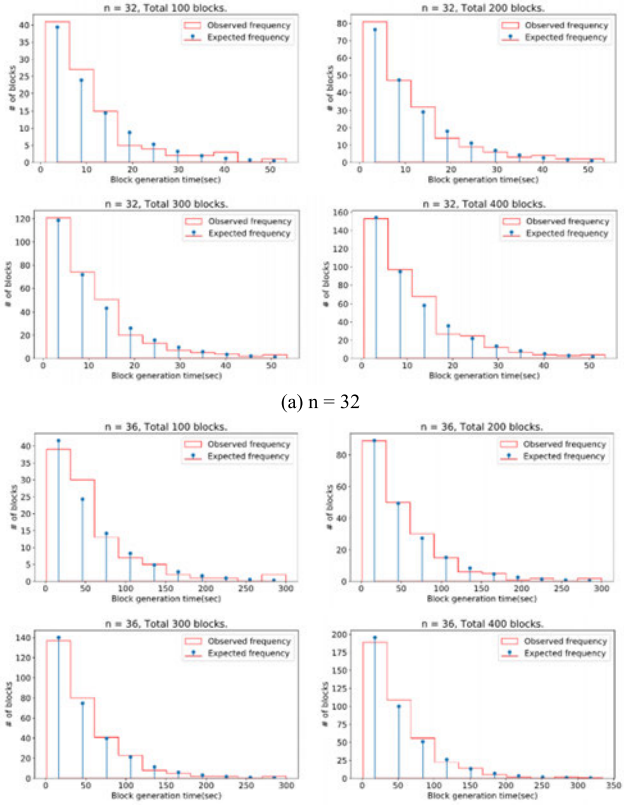
B. STABILITY OF THE BLOCK GENERATION TIME

Fig. 2 demonstrates the importance of determining whether varying puzzles may result in outliers. That is, in *BLOCK TIME* and *DIFFICULTY* of Fig. 2, slow block generations are observed despite the low difficulty level. In other words, the observation of BGT shows outliers. If the outliers are uncontrollable, the BGT distribution has an infinite mean similar to the heavy-tailed distribution. An infinite mean cannot guarantee transaction confirmation. Thus, to achieve a stable BGT that can guarantee transaction confirmation, the BGT must have a finite mean.

We obtain the BGT of ECCPoW Ethereum with a fixed difficulty level to observe the type of distribution with a finite mean the BGT follows. Specifically, if BGT follows an exponential distribution, it has a finite mean. However, if the BGT follows a heavy-tailed distribution, it has an infinite mean [15]. Thus, through the goodness-of-fit experiment, we aimed to discuss what type of distribution the BGT follows. For the goodness-of-fit experiment, we set a null hypothesis H_0 and an alternative hypothesis H_A:

H_0: BGT has the exponential distribution

H_A: BGT does not have the exponential distribution



(a) n = 32

(b) n = 36

FIGURE 3. We did experiments for 100, 200, 300, and 400 blocks to observe the distribution over the number of blocks. As the number of blocks increases, the standard error decrease. That is, when the number of blocks increases, sample distribution reflects an actual distribution of sample distribution. In these figures, experiment results show the tendency; the distributions of observed frequency, known as sample distribution, follow the distribution of expected frequency.

For the goodness-of-fit experiment, we use the AD test [16]–[18]. Other available tests can be used in the goodness-of-fit experiment, such as the chi-square test [19], Kolmogorov–Smirnov test [20], and AD test [16]. The chi-square test has a restrictive assumption that all expected frequencies should be greater or equal to 5 [21]. However, there is no guarantee that our samples will achieve this assumption. If we collect more samples, the chi-square test can be used. However, the p-values used to validate the hypotheses are affected by the number of samples. When the number of samples increased in the chi-square test, the p-values tend to decrease. Therefore, the assumption of the chi-square test is inappropriate for verifying our distributions. The Kolmogorov–Smirnov test is unaffected by sample sizes; however, it is more sensitive to the center of the distribution rather than the tail [22]. We must consider verifying the tail of the distribution to cover all possibilities. Therefore, we have chosen to use the AD test [16], which gives more weight to the tail than does the Kolmogorov–Smirnov test.

C. AD TESTS

In this subsection, we discuss the AD test and verify its usage using test examples. The AD test is used to verify if a sample

TABLE 1. Example of the Anderson–Darling test results.

The number of samples	Standardized A^2_{MN}	p-value
10	-0.59	$p \geq 0.25$
20	0.44	$p = 0.21$
30	0.69	$p = 0.17$

(a) F ~ Exp(1), G ~ Normal(1, 1)

10	1.20	$p = 0.11$
20	3.57	$p = 0.02$
30	4.67	$p = 0.01$

(b) F ~ Exp(1), G ~ Exp(2)

10	1.11	$p = 0.12$
20	-0.41	$p \geq 0.25$
30	-0.08	$p \geq 0.25$

(c) F ~ Exp(1), G ~ Exp(1)

follows a specific distribution. We discuss one-sample and two-sample AD tests. In our work, we use the two-sample AD test; however, to clearly present our contribution, we briefly introduce the one-sample AD test first.

1) ONE-SAMPLE AD TEST

The one-sample AD test is suitable to verify a hypothesis that a sample set comes from a population. The one-sample AD test is as follows. When the cumulative distribution function (CDF) of the population distribution is $F(x)$, and the CDF of the empirical distribution is $F_M(x)$, the one-sample AD test [18] is used as follows:

$$A^2_M = M \int_{-\infty}^{\infty} (F_M(x) - F(x))^2 \, w(x) dF(x) \tag{9}$$

and

$$w(x) = [F(x)(1 - F(x))]^{-1} \tag{10}$$

where M denotes the number of samples and A^2_M denotes the results of the one-sample AD test. Intuitively, in (9), if $F_M(x) - F(x)$ is 0 for all x, A^2_M is 0. This means that when A^2_M is small, the empirical distribution $F_M(x)$ is close to the population distribution $F(x)$. As we have noted, we focus on the tail of the distribution; it can be accomplished using (10). The one-sample AD test result A^2_M can be used to verify if a given sample comes from a population with a specific distribution.

2) TWO-SAMPLE AD TEST

In our work, we want to verify that two-sample sets come from the same unknown population. The two-sample AD test is appropriate for such verification. The two-sample AD test [17], [18] is as follows. There are two-sample empirical distributions $F_M(x)$ and $G_N(x)$. The $F_M(x)$ is an empirical distribution derived from the set $\mathcal{F}$ with a cardinality of the sample set $M = |\mathcal{F}|$. The $G_N(x)$ is also an empirical distribution derived from the set $\mathcal{G}$ with a cardinality of the sample set $N = |\mathcal{G}|$. $F_M(x)$ and $G_N(x)$ are the respective sample sets

TABLE 2. The observed frequency is calculated using the histogram in Fig. 4, and the expected frequency is calculated using the CDF of the exponential distribution derived from the mean in Fig. 4.

Interval(%)	Observed frequency	Expected frequency
[0, 10)	107	118.70
[10, 20)	82	71.73
[20, 30)	56	43.35
[30, 40)	20	26.19
[40, 50)	14	15.83
[50, 60)	7	9.56
[60, 70)	5	5.78
[70, 80)	4	3.49
[80, 90)	2	2.11
[90, 100]	3	1.27

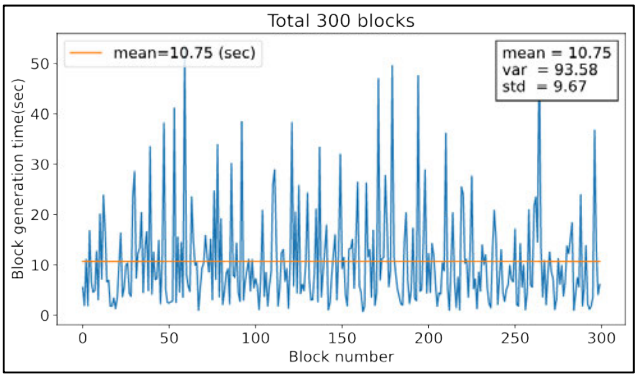


FIGURE 4. This figure presents block generation time of 300 blocks when n is 32. The mean block generation time of 300 blocks is 10.75 s, and it is presented as a horizontal line. Such a result is converted to a histogram. The observed frequency of Table 2 denotes the histogram of Fig. 4. The legend at the top right shows the mean, variance, and standard deviation of the BGT.

independently obtained from two different testing locations. The two-sample AD test can be used to determine whether both sample distributions come from the same distribution. In [17], [18], the two-sample

version is defined as follows:

$$A^2_{MN} = \frac{MN}{K} \int_{-\infty}^{\infty} \frac{(F_M(x) - G_N(x))^2}{H_K(x)(1 - H_K(x))} dH_K(x) \tag{11}$$

where $H_K(x) = (MF_M(x) + NG_N(x))/K$ with $K = M + N$. A^2_{MN} is standardized to remove the dependencies derived by the number of samples. This standardized form is used to calculate the p-value [17], [18]. The p-value evidences the hypothesis test.

The two-sample AD test is suitable to verify a hypothesis that two-sample sets come from the same population. As a null hypothesis H_0 for the two-sample AD test, we set $F_M(x)$ to have the same population as $G_N(x)$. In addition, we set $G_N(x)$ as an exponential distribution. Thus, if $F_M(x)$ and $G_N(x)$ comes from same the population, that is, H_0 is true, we may consider that $F_M(x)$ is the exponential distribution. If the p-value of the AD test is sufficiently large, it proves that H_0 is true.

The p-value is the false positive probability under the assumption that the null hypothesis is true. A low p-value indicates that a test result provides evidence against the null hypothesis; a large p-value does not. That is, a large p-value denotes the probability of a true negative is low. The p-value is determined from the observation of the sample data. Thus, before observing the data, we first set the threshold significance level (*TSL*), $TSL \in [0, 1]$. The *TSL* can be used to determine the critical value. Given a *TSL* and the number of samples that are used in the AD test, the *TSL* table in [18] is used to read off a value corresponding to the TSL and the number of samples. This read-off value is called the critical value. If the standardized A^2_{MN} is smaller than the critical value, this result indicates that the p-value is larger than the predefined *TSL*. In the *TSL* table of [18], the maximum *TSL* is 0.25. Thus, when standardized A^2_{MN} is less than the critical value corresponding to the 0.25 *TSL*, the p-value is capped at 0.25.

3) VERIFICATION OF THE AD TEST

In this subsection, we verify the two-sample AD test method. Verification is performed under the assumption that the input distributions are *a priori* known. This will clearly illustrate how we will use the AD test and interpret its test results.

In Table 1, we present three examples to give an insight into the p-value of the AD test; in this example, we use true distributions for $F_M(x)$ and $G_N(x)$. In Table 1, Exp(θ) indicates the exponential distribution with mean θ and Normal(μ, σ) indicates the normal distribution with mean μ and standard deviation σ. That is, $\mathcal{F} \sim \text{Exp}(\theta)$ denotes the sample set $\mathcal{F}$ of $F_M(x)$; samples are derived from the exponential distribution with mean θ. In Table 1 (a), we use the exponential distribution for $F_M(x)$ and the normal distribution for $G_N(x)$; these distributions have the same mean. This example shows that as the number of samples increases, the p-value tends to decrease if samples are drawn from different distributions. In Table 1 (b), we set both $F_M(x)$ and $G_N(x)$ as the exponential distribution, but each with different mean values. This example shows that as the number of samples increases even though samples are drawn from the same exponential distribution, the p-value tends to decrease if the means of distributions are different. In Table 1 (c), we set both $F_M(x)$ and $G_N(x)$ to be exactly the same exponential distribution. That is, the two-sample sets $\mathcal{F} \sim F_M(x)$ and $\mathcal{G} \sim G_N(x)$ come from the same population. This example shows that, as the number of samples increases, the p-value tends to increase when two-sample sets are drawn from the same population. From these examples in Table 1, we note that the closer the two distributions $F_M(x)$ and $G_N(x)$ are to each other, the larger p-value is obtained.

We determine whether the AD test result of our experiments indicates that $F_M(x)$ is sufficiently close to $G_N(x)$. That is, given there are two-sample sets, one of $F_M(x)$ and the other of the exponential $G_N(x)$, we want to determine whether we can make a quality statement about how close the two-sample sets are to each other according to the AD test.

TABLE 3. Anderson-Darling test result. The test result presents a large p-value. It means that if we reject the null hypothesis, the probability of a true negative is low.

n	# of blocks	Observed mean(sec)	std	Standardized A^2_{MN}	p-value
32	100	10.86	9.84	-1.12	$p \geq 0.25$
32	200	11.24	10.16	-1.20	$p \geq 0.25$
32	300	10.74	9.67	-1.18	$p \geq 0.25$
32	400	11.08	9.84	-1.09	$p \geq 0.25$
32	500	10.91	9.62	-1.11	$p \geq 0.25$
32	600	10.87	9.48	-0.80	$p \geq 0.25$
32	700	10.84	9.41	-0.36	$p \geq 0.25$
32	800	10.76	9.40	-0.36	$p \geq 0.25$
36	100	56.00	55.20	-1.11	$p \geq 0.25$
36	200	51.04	49.71	-1.19	$p \geq 0.25$
36	300	47.84	45.49	-1.12	$p \geq 0.25$
36	400	49.97	47.80	-1.19	$p \geq 0.25$
36	500	49.24	46.95	-1.11	$p \geq 0.25$
36	600	48.23	46.96	-1.18	$p \geq 0.25$
36	700	48.36	47.68	-1.18	$p \geq 0.25$
36	800	48.03	46.89	-1.18	$p \geq 0.25$

The AD test result presents a significant p-value, i.e., $p \geq 0.25$; it is a necessary condition but not a sufficient one for the case that the two distributions are the same. In other words, if a decision is made to reject the null hypothesis, that is, the distribution $F_M(x)$ is not close to the exponential distribution $G_N(x)$, such a decision will result in an error with a probability greater than 0.25.

D. APPLICATION OF AD TEST TO BGT DISTRIBUTION

In this subsection, we use the AD test to determine the distribution of the BGT of ETH-ECC. For this experiment, 90 threads were used to generate a block. We experimented using a fixed code length to observe the BGT without changing the difficulty level. In the test, two kinds of code length n are used: 32 and 36. These are the two lowest types of code length n in our pseudo-difficulty table used in the simulation. We divided the BGT into 10 intervals between the minimum BGT and maximum BGT for a histogram. For example, when the minimum BGT is 10 and the maximum BGT is 20, there are 10 intervals, i.e., [10,11], [11,12], . . . , [19,20]. Using these intervals, we count the observed frequency of the BGT data. We set $F_M(x)$ using the observed frequency and set $G_N(x)$ using the mean of the BGT data. The mean in Fig. 4 is used for the expected frequency of $G_N(x)$ in Table 2. That is, the mean in Fig. 4 is used as $1/\lambda$ for the CDF of the exponential distribution $G_N(x)$:

$$G_N(x) = 1 - e^{-\lambda x} \tag{12}$$

The expected frequency of Table 1 is calculated using the integral of $G_N(x)$ corresponding to the interval time. Because $G_N(x)$ is the exponential distribution, if $F_M(x)$ is close to $G_N(x)$ we may consider $F_M(x)$ is an exponential distribution.

E. DISCUSSION ON AD TEST RESULTS

Fig. 4 shows the example result of the BGT over different blocks. Each block denotes the trial to obtain the BGT. We converted the test results, such as those in Fig. 4, to a distribution over time to analyze the BGT. These converted distributions are presented in Fig. 3. Fig. 3 presents the plots of the distribution of the observed and expected frequencies. These frequencies are calculated using the method described in Section V-D.

When we obtain a distribution using a sample set, there is a standard error; the standard error is high when the number of samples in the set is small. The standard error is expressed as

$$\frac{\sigma}{\sqrt{N}}$$

where σ denotes the standard deviation of a population and N denotes the cardinality of the sample set. The standard error decreases as the number of samples increases. Thus, the sample distribution becomes closer to the actual distribution of the observed samples. If the sample distribution, which reflects the actual distribution, differs from the expected distribution, we can observe that the sample distribution differs from the expected distribution. To observe the tendency of distribution over some blocks, we experimented with 100, 200, 300, and 400 blocks. Fig. 3 shows that the distribution of the observed frequency tends to follow the distribution of the expected frequency. In addition, Table 3 shows that the observed mean and standard deviation tend to converge as the number of blocks increases.

Furthermore, for the quantitative analysis, we use the AD test. Table 3 presents the AD test results to discuss hypotheses H_0 and H_A. These results show a similar result in Table 1 (c). In Table 1 (c), we drew samples from the same true distribution; the results present the largest possible p-value. All p-values in Table 3 are larger than or equal to 0.25, regardless of the number of blocks. In other words, if the null hypothesis is rejected, this decision will cause an error with a probability greater than 0.25. That is, the decision that the BGT distribution $F_M(x)$ does not follow the exponential distribution could be made with a high decision error.

VI. DISCUSSION

The purpose of ECCPoW is not to replace the current PoW of Ethereum. We propose our algorithm to present as one of the options for the Ethereum network. Ethereum can be utilized, for example, not only in a large-scale network but also in local-scale networks. To support a local-scale network, Ethereum provides PoW and PoA(Proof-of-Authority) as consensus algorithms. These algorithms have limitations for the local-scale network. For instance, PoW based local network has a risk of a double-spending attack by ASIC miners; PoA based network has a limitation of a participant because the time complexity of a PoA increases exponentially when the number of participants increases. Our algorithm, ECCPoW, can be utilized in such cases for the benefit of offering a novel PoW that allows numerous participants with deterrence to ASIC-borne attacks. In addition, our novel ECCPoW may open up for an expected use and thus untraveled future to Ethereum.

Extensive Simulation Set up at AWS. We have recruited twelve instances on Amazon Web Service (AWS) EC2; each instance of EC2 instances works as a node in a blockchain network. The cost of using AWS EC2 increases rapidly because PoW utilizes all the resources of instances. We were able to confirm that this scale of the experiment was good enough to achieve our main goal, which is aimed at verifying the stability of the block generation time of ECCPoW Ethereum. AWS simulation was done to obtain the trace data of block generation times. The twelve nodes employed in our simulation were divided into two different kinds of nodes. One kind is *bootnodes* which help the nodes connected. Nodes that want to join a network are connected to *bootnodes* first. After connection, *bootnodes* relay nodes to other nodes. In Ehtereum, *bootnodes* addresses are hardcoded on source codes, but it is possible to set *bootnodes* addresses manually for private networks. We have chosen two *bootnodes*. The other kind of nodes are *sealnodes* that participate in block generation as a miner in the PoW network. We have chosen the number of *sealnodes* to be 10. We use the m5.xlarge of AWS EC2, which has conventional node specification: four virtual CPUs and 16 GB memory for the real-world simulation. All nodes are deployed by Docker according to the guidance of Ethereum. Thus, all of our simulation results, which are shown in Fig. 2, are reproducible.

VII. CONCLUSION

In this work, we present the implementation, simulation, and validation of ETH-ECC. In the implementation, we showed how Ethereum can be updated with ECCPoW as its new consensus algorithm. In the simulation, we conducted a multinode experiment using AWS EC2. The results showed that ETH-ECC with its adaptive difficulty-level controllability is successfully implemented. In the validation, we showed statistical results in which the necessary condition for a finite mean BGT is satisfied such that the distribution of the ECCPoW block generation time is exponential.

REFERENCES

[1] S. Nakamoto. *Bitcoin: A Peer-to-Peer Electronic Cash System*. Accessed: Nov. 3, 2021. [Online]. Available: https://bitcoin.org/bitcoin.pdf

[2] V. Buterin. *A Next-Generation Smart Contract and Decentralized Application Platform*. Accessed: Nov. 3, 2021. [Online]. Available: https://ethereum.org/en/whitepaper/

[3] M. Rosenfeld, "Analysis of hashrate-based double spending," 2014, *arXiv:1402.2009*. [Online]. Available: http://arxiv.org/abs/1402.2009

[4] J. Jang and H.-N. Lee, "Profitable double-spending attacks," 2019, *arXiv:1903.01711*. [Online]. Available: http://arxiv.org/abs/1903.01711

[5] G. Wood, "Ethereum: A secure decentralised generalised transaction ledger," *Ethereum Project Yellow Paper*, vol. 151, pp. 1–32, Apr. 2014.

[6] S. Park, H. Choi, and H. Lee, "Time-variant proof-of-work using error correction codes," Jun. 2020, *arXiv:2006.12306*. [Online]. Available: https://arxiv.org/abs/2006.12306

[7] H. Jung and H. Lee, "Error-correction code based proof-of-work for ASIC resistance," *Symmetry*, vol. 12, no. 6, p. 988, Jun. 2020.

[8] W. Ryan and S. Lin, "Low-density parity-check codes," in *Channel Codes: Classical and Modern*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[9] R. G. Gallager, *Low Density Parity Check Cdes*. Cambridge, MA, USA: MIT Press, 1963.
[10] S. Shao, P. Hailes, T.-Y. Wang, J.-Y. Wu, R. G. Maunder, B. M. Al-Hashimi, and L. Hanzo, "Survey of turbo, LDPC, and polar decoder ASIC implementations," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2309–2333, 3rd Quart., 2019.
[11] Y.-L. Ueng, B.-J. Yang, C.-J. Yang, H.-C. Lee, and J.-D. Yang, "An efficient multi-standard LDPC decoder design using hardware-friendly shuffled decoding," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 60, no. 3, pp. 743–756, Mar. 2013.
[12] E. Duffield and D. Diaz. *Dash: A Payments-Focused Cryptocurrency*. Accessed: Nov. 3, 2021. [Online]. Available: https://github.com/dashpay/dash/wiki/Whitepaper
[13] T. Black and J. Weight. *X16R ASIC Resistant by Design*. Accessed: Nov. 3, 2021. [Online]. Available: https://ravencoin.org/assets/documents/X16R-Whitepaper.pdf
[14] W. E. Ryan and S. Lin, *Channel Codes—Classical and Modern*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[15] S. Foss, D. Korshunov, and S. Zachary, *An Introduction to Heavy-Tailed and Subexponential Distributions*. New York, NY, USA: Springer, 2011.
[16] T. W. Anderson and D. A. Darling, "Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes," *Ann. Math. Statist.*, vol. 23, no. 2, pp. 193–212, Jun. 1952.
[17] A. N. Pettitt, "A two-sample anderson-darling rank statistic," *Biometrika*, vol. 63, no. 1, pp. 161–168, 1976.
[18] F. Scholz and M. Stephens, "K-sample anderson-darling tests," *J. Amer. Stat. Assoc.*, vol. 82, no. 399, pp. 918–924, Sep. 1987.
[19] W. G. Cochran, "The X^2 test of goodness of fit," *Ann. Math. Statist.*, vol. 23, no. 3, pp. 315–345, Sep. 1952.
[20] J. L. Hodges, "The significance probability of the Smirnov two-sample test," *Arkiv Matematik*, vol. 3, no. 5, pp. 469–486, Jan. 1958.
[21] W. G. Cochran, "Some methods for strengthening the common tests," *Biometrics*, vol. 10, pp. 417–451, Dec. 1954.
[22] J. J. Filliben. *1.3.5.16. Kolmogorov–Smirnov Goodness-of-Fit Test*. NIST/SEMATECH e-Handbook of Statistical Methods. Accessed: Nov. 3, 2021. [Online]. Available: http://www.itl.nist.gov/div898/handbook/
[23] T. Black. *Ravencoin–ASIC Thoughts—Round Two*. Accessed: Nov. 3, 2021. [Online]. Available: https://medium.com/@tronblack/ravencoin-asic-thoughts-round-two-f4f743942656
[24] Tevador. *Random X*. Accessed: Nov. 3, 2021. [Online]. Available: https://github.com/tevador/RandomX
[25] M. Belotti, N. Bozic, G. Pujolle, and S. Secci, "A vademecum on blockchain technologies: When, which, and how," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3796–3838, 4th Quart., 2019.
[26] S. Asmussen, "Steady-state properties of of GI/G/1," in *Applied Probability and Queues*. New York, NY, USA: Springer, 2003, pp. 266–301, doi: 10.1007/0-387-21525-5_10.



HYOUNGSUNG KIM received the B.S. degree in electronics and computer engineering from Chonnam National University, Gwangju, South Korea, in 2019, and the M.S. degree from the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju. Since 2021, he has working with Korea Electronics Technology Institute. His research interests include blockchain, distributed systems, and statistical analysis.



JEHYUK JANG received the B.S. degree in electronic engineering from the Kumoh National Institute of Technology, Gumi, South Korea, in 2014, and the M.S. degree in information and communication engineering from GIST, South Korea, where he is currently pursuing the Ph.D. degree with the School of Electrical Engineering and Computer Science. His research interests include sub-Nyquist sampling and compressed sensing.



SANGJUN PARK received the B.S. degree in computer engineering from Chungnam National University, Daejeon, South Korea, in 2009, and the Ph.D. degree in electrical engineering and computer science from Gwangju Institute of Science and Technology, Gwangju, South Korea, in 2019. Since 2020, he has been working with the Electronics and Telecommunications Research Institute. His research interests include information theory, numerical optimization, compressed sensing, blockchain, deep-neural networks, and finite state machine.



HEUNG-NO LEE (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees from the University of California, Los Angeles, CA, USA, in 1993, 1994, and 1999, respectively, all in electrical engineering. From 1999 to 2002, he was a Research Staff Member with HRL Laboratories, LLC, Malibu, CA, USA. From 2002 to 2008, he was an Assistant Professor with the University of Pittsburgh, PA, USA. In 2009, he was with the School of Electrical Engineering and Computer Science, GIST, South Korea, where he is currently affiliated. His research interests include information theory, signal processing theory, blockchain communications/networking theory, and their application to wireless communications and networking, compressive sensing, future internet, and brain–computer interface. He was a recipient of several prestigious national awards, including the Top 100 National Research and Development Award, in 2012; the Top 50 Achievements of Fundamental Researches Award, in 2013; and the Science/Engineer of the Month, in January 2014.

An Information-Theoretic Study for Joint Sparsity Pattern Recovery With Different Sensing Matrices

Sangjun Park, Nam Yul Yu, *Member, IEEE*, and Heung-No Lee, *Senior Member, IEEE*

***Abstract*— In this paper, we study a support set reconstruction problem for multiple measurement vectors (MMV) with different sensing matrices, where the signals of interest are assumed to be jointly sparse and each signal is sampled by its own sensing matrix in the presence of noise. Using mathematical tools, we develop upper and lower bounds of the failure probability of the support set reconstruction in terms of the sparsity, the ambient dimension, the minimum signal-to-noise ratio, the number of measurement vectors, and the number of measurements. These bounds can be used to provide guidelines for determining the system parameters for various compressed sensing applications with noisy MMV with different sensing matrices. Based on the bounds, we develop necessary and sufficient conditions for reliable support set reconstruction. We interpret these conditions to provide theoretical explanations regarding the benefits of taking more measurement vectors. We then compare our sufficient condition with the existing results for noisy MMV with the same sensing matrix. As a result, we show that noisy MMV with different sensing matrices may require fewer measurements for reliable support set reconstruction, under a sublinear sparsity regime in a low noise-level scenario.**

***Index Terms*— Compressed sensing, support set reconstruction, joint sparsity structure, multiple measurement vectors model.**

I. Introduction

CONVENTIONALLY, signals sensed from sensors such as microphones and imaging devices are sampled following the Shannon and Nyquist sampling theory [1] at a rate higher than twice the maximum frequency for signal reconstruction. As the number of samples decided by this theory is often large, the samples go through a compression stage before being stored. Therefore, taking numerous samples, where most of them will be discarded in this stage, is inefficient. Because compressed sensing (CS) [2]–[7] removes the inefficiency, CS has been applied in various areas such as wireless communications [8]–[11], spectrometers [12], multiple input multiple output (MIMO) radars [13], magnetic resonance imaging [14], and imaging/signal processing [15]–[17].

The CS theory states that signals that are sparsely representable in a certain basis are compressively sampled and reconstructed from what we thought is incomplete in information. Let $\mathbf{x} \in \mathbb{R}^N$ be a *K-sparse* vector with a support set $\mathcal{I} := \{i \,|x\,(i) \neq 0\}$ whose indices indicate the positions of the nonzero coefficients of $\mathbf{x}$. It is compressively sampled by a model called *single measurement vector* (*SMV*) as follows:

$$\mathbf{y} = \mathbf{F}\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^M$ is a (noisy) measurement vector, $\mathbf{F} \in \mathbb{R}^{M \times N}$ is a sensing matrix, and $\mathbf{n} \in \mathbb{R}^M$ is a noise vector, whose elements are independent and identically distributed (i.i.d) Gaussian with a zero mean and a σ^2 variance. Once the support set is correctly reconstructed, then (1) can be well-posed, which allows us to obtain an accurate estimate of $\mathbf{x}$ using the least square approach. We thus aim to focus on the support set reconstruction problem.

A. Information-Theoretic Works for CS With SMV

Works [18]–[23] have studied the support set reconstruction problem from an information-theoretic perspective. For reliable support set reconstruction, sufficient and necessary conditions were established in the linear and sublinear sparsity regimes.

For support set reconstruction, Wainwright [18] used the union bound to establish a sufficient condition on the number of measurements M for a maximum likelihood (ML) decoder and used Fano's inequality [24] to obtain a necessary condition on M. This ML decoder was analyzed by Fletcher *et al.* [19] to establish a necessary condition on M. Aeron *et al.* [20] used Fano's inequality to form necessary conditions on both M and σ^2. Then, they used the union bound to obtain sufficient conditions on both M and σ^2 for their sub-optimal decoder. Akcakaya and Tarokh [21] used the union and the large deviation bounds based on empirical entropies to get sufficient conditions on M for their joint typical decoder. They used the converse of the channel coding theorem to get necessary conditions on M. Scarlett *et al.* [22] extended this decoder [21] with the assumption that the distribution of the support set is provided. For a uniform distribution case, their necessary and sufficient conditions are equivalent to those of [21]. However, they are better for a non-uniform distribution case. Scarlett and Cevher [23] linked the support set reconstruction with the problem of coding over a mixed channel, where information spectrum methods were used to obtain necessary and sufficient conditions on M.

B. Information-Theoretic Works for CS With MMV

CS has many applications in wireless sensor networks (WSNs) [8]–[11] and MIMO radars [13]. In these

Manuscript received April 1, 2016; revised November 14, 2016; accepted January 17, 2017. Date of publication May 15, 2017; date of current version August 16, 2017. This work was supported by the National Research Foundation of Korea through the South Korean Government under Grant NRF-2015R1A2A1A05001826. This paper was presented at the 2012 International Symposium on Information Theory.

The authors are with the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea (e-mail: sjpark1@gist.ac.kr; nyyu@gist.ac.kr; heungno@gist.ac.kr).

Communicated by H. Rauhut, Associate Editor for Signal Processing.
Digital Object Identifier 10.1109/TIT.2017.2704111

0018-9448 © 2017 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:47:30 UTC from IEEE Xplore. Restrictions apply.

applications, the signals of interest $\mathbf{x}^s \in \mathbb{R}^N,\ s = 1, 2, \cdots, S$ are often modeled as *jointly K-sparse vectors*, implying that $\mathcal{I} = \mathcal{I}^1 = \mathcal{I}^2 \cdots = \mathcal{I}^S$, where $\mathcal{I}^s$ is the support set of $\mathbf{x}^s$ and $|\mathcal{I}| = K$, which is referred to as a *joint sparsity structure*.

There are two models for sampling jointly *K-sparse* vectors. The first model is called *multiple measurement vectors* (MMV) with *the same sensing matrix* [25], in which they are sampled by the same sensing matrix. The second model is named as MMV with *different sensing matrices* [8], [9], in which each one is sampled by its own sensing matrix.

The authors of [26]–[28] have conducted information-theoretic research to obtain conditions under which the support set of both the models was reconstructed with a high probability. In noisy MMV with the same sensing matrix, Tang and Nehorai [26] used the hypothesis theory to obtain necessary and sufficient conditions on both the number of measurements M and the number of measurement vectors S, and proved that the success probability of the support set reconstruction increases with S, if $M = \Omega\left(K \log \frac{N}{K}\right)$. Jin and Rao [27] exploited the communication theory to establish necessary and sufficient conditions on M and demonstrated the benefits of the joint sparsity structure based on their conditions. A detailed comparison between the results of our paper and those of [27] will be presented in Section IV. Finally, Duarte *et al.* [28] studied noiseless MMV with different sensing matrices, and formed necessary and sufficient conditions on M. However, it is difficult to apply the conditions to noisy MMV with different sensing matrices.

Meanwhile, works [8], [29], [30] have presented conditions of practical algorithms for a reliable support set reconstruction. In noiseless MMV with the same sensing matrix, Blanchard and Davies [30] obtained conditions for a reliable reconstruction from rank aware orthogonal matching pursuit (OMP). In noisy MMV with the same sensing matrix, Kim *et al.* [29] created compressive MUSIC, and presented its sufficient condition. In noiseless MMV with different sensing matrices, Baron *et al.* [8] produced trivial pursuit (TP) and distributed compressed sensing-simultaneous OMP (DCS-SOMP). By analyzing TP with the assumption that each sensing matrix contains i.i.d. Gaussian elements and that the nonzero values of each sparse vector are i.i.d. Gaussian variables, they demonstrated that with $M \geq 1$, TP reconstructs the support set as S is sufficiently large. They conjectured that $M \geq K + 1$ suffice for DCS-SOMP to reconstruct the support set as S is sufficiently large, based on its empirical results.

To the best of our knowledge, no information-theoretic study has been published to get necessary and sufficient conditions for reliable support set reconstruction in noisy MMV with different sensing matrices. Besides, these conditions have not been provided from the practical recovery algorithms for CS with noisy MMV with different sensing matrices.

C. Motivations of This Paper

CS with noisy MMV with different sensing matrices has been applied in many applications and the benefits facilitated by the joint sparsity structure have been empirically reported in [10] and [14]. In WSNs, Caione *et al.* [10] used the joint sparsity structure to reduce the number of transmitted bits per sensor and reported that each sensor can reduce its transmission cost. In magnetic resonance imaging (MRI), Wu *et al.* [14] modeled multiple diffusion tensor images (DTIs) as jointly sparse vectors. They exploited the joint sparsity structure to reduce the number of samples per DTI, while retaining the reconstruction quality. Using the joint sparsity structure, they also empirically reported that the reconstruction quality of each DTI can be improved for a fixed number of samples per DTI.

To theoretically explain the above empirical benefits facilitated by the joint sparsity structure, theoretical tools are required to measure the performance of CS with noisy MMV with different sensing matrices. Such tools can be useful as guidelines for determining the system parameters in various CS applications with noisy MMV with different sensing matrices. For example, if the number of samples per DTI is fixed in the MRI [14], the theoretical tools may enable us to determine the number of DTIs required for achieving a given reconstruction quality. Thus, the first motivation of this paper is to provide the theoretical tools by establishing sufficient and necessary conditions for reliable support set reconstruction.

Next, for noiseless MMV with the same sensing matrix, let $\mathbf{Y}_{\mathrm{A}} = \mathbf{F} \times \left[\mathbf{x}^1\ \mathbf{x}^2\ \cdots\ \mathbf{x}^S\right] \in \mathbb{R}^{M \times S}$. Also, for noiseless MMV with different sensing matrices, let $\mathbf{Y}_{\mathrm{B}} = \left[\mathbf{F}^1\mathbf{x}^1\ \mathbf{F}^2\mathbf{x}^2\ \cdots\ \mathbf{F}^S\mathbf{x}^S\right] \in \mathbb{R}^{M \times S}$. Then, all the elements of $\mathbf{Y}_{\mathrm{B}}$ are uncorrelated because all the sensing matrices are independent. In contrast, those of $\mathbf{Y}_{\mathrm{A}}$ are correlated because they are taken from the same sensing matrix. Now, we consider a case where we set $S > K$ and $M > K$. Then, it is clear that $\mathrm{rank}\,(\mathbf{Y}_{\mathrm{B}}) = \min(S, M)$ with a high probability and $\mathrm{rank}\,(\mathbf{Y}_{\mathrm{A}}) \leq K$. Therefore, for this case, we conclude that $\mathrm{rank}\,(\mathbf{Y}_{\mathrm{B}}) > \mathrm{rank}\,(\mathbf{Y}_{\mathrm{A}})$. This implies that a more reliable support set reconstruction can be expected in noiseless MMV with different sensing matrices for this case. Thus, the second motivation is to verify this perception in the presence of noise, by comparing our results with the existing ones in noisy MMV with the same sensing matrix [27].

D. Contributions of This Paper

The contributions of this paper are as follows: First, we derive upper and lower bounds of a failure probability of the support set reconstruction from Lemmas 1 and 2, by exploiting Fano's inequality [24] and the Chernoff bound [31]. These bounds are used for measuring the performance of CS with noisy MMV with different sensing matrices.

Next, we develop necessary and sufficient conditions for reliable support set reconstruction. Theorem 1 states that

$$M > K\left(1 + \frac{1}{Sf\left(\mathrm{SNR}_{\min}\right)}\right)$$

suffices to achieve reliable support set reconstruction in the *linear sparsity* regime, i.e., $\lim_{N\to\infty} \frac{K}{N} = \beta \in (0, 1/2)$, and it also states that

$$M > K\left(1 + \frac{1}{Sf(\mathrm{SNR}_{\min})} \log \frac{N}{K}\right)$$

suffices to achieve reliable support set reconstruction in the*sublinear sparsity* regime, i.e., $\lim_{N\to\infty} \frac{K}{N} = 0$, where

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:47:30 UTC from IEEE Xplore. Restrictions apply.

$f(\mathrm{SNR}_{\min})$ is an increasing function with respect to the minimum signal-to-noise ratio $\mathrm{SNR}_{\min}$ defined in (4). Next, for a finite S, N, K, and $\mathrm{SNR}_{\min}$, Theorem 3 states that

$$M < \frac{2K \log \frac{N}{K} - 2 \log 2}{S \log (1 + K \times \mathrm{SNR}_{\min})}$$

is necessary for reliable support set reconstruction. The necessary and sufficient conditions can be useful as guidelines to determine the system parameters of CS applications with noisy MMV with different sensing matrices. Corollaries 1 and 2 indicate that reliable support set reconstruction is possible as sufficiently many measurement vectors S for a fixed M are taken at a low $\mathrm{SNR}_{\min}$. For a fixed N and K, Theorem 2 shows that $M \geq K + 1$ measurements suffice for reconstructing the support set, as S is sufficiently large. Then, for a fixed N, K, and $M = K + 1$, Corollary 3 provides a sufficient condition on S for reliable support set reconstruction. We provide theoretical explanations of the benefits of the joint sparsity structure, which conform with the empirical results of CS applications with noisy MMV with different sensing matrices [10], [14]. Finally, we compare the sufficient condition (11) with the known one (26) for noisy MMV with the same sensing matrix [27]. Therefore, we demonstrate that if $S \geq K$, noisy MMV with different sensing matrices may require fewer measurements M for reliable support set reconstruction than noisy MMV with the same sensing matrix under a low noise-level scenario. It confirms the superiority of MMV with different sensing matrices.

II. Notations, System Model & Problem Formulation

A. Notations

The following notations will be used in the whole paper.

1. $\mathbb{P}$, $\mathbb{E}$ and $\mathbb{V}$ denote the probability, expectation and (co)variance, respectively.
2. A small (capital) bold letter $\mathbf{f}$ ($\mathbf{F}$) is a vector (matrix).
3. A sub-vector (sub-matrix) formed by the elements (columns) of $\mathbf{f}$ ($\mathbf{F}$) indexed by a set $\mathcal{I}$ is denoted by $\mathbf{f}_{\mathcal{I}}$ ($\mathbf{F}_{\mathcal{I}}$).
4. For a given matrix $\mathbf{F}$, its inversion, transpose, trace and the ith eigenvalue are denoted by $\mathbf{F}^{-1}$, $\mathbf{F}^T$, $\mathrm{tr}[\mathbf{F}]$ and $\lambda_i(\mathbf{F})$, respectively. Also, its orthogonal projection matrix is defined by
$$\mathbf{Q}(\mathbf{F}) := \mathbf{I}_M - \mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T \tag{2}$$
where $\mathbf{Q}(\mathbf{F})$ maps an arbitrary vector to the space orthogonal onto the space spanned by the columns of $\mathbf{F}$.
5. For given sets $\mathcal{I}$ and $\mathcal{J}$, the relative complements of $\mathcal{J}$ in $\mathcal{I}$ is denoted as $\mathcal{J} \setminus \mathcal{I}$. The cardinality of a set $\mathcal{I}$ is denoted by $|\mathcal{I}|$.
6. For a given function $f(x)$, its nth derivation with respect to x is denoted by $f^n(x)$.
7. The *linear sparsity regime* is defined by $\lim_{N\to\infty} \frac{K}{N} = \beta \in (0, 1/2)$.
8. The *sublinear sparsity regime* is defined by $\lim_{N\to\infty} \frac{K}{N} = 0$.
9. The expression $f(x) = \Omega(g(x))$ denotes $|f(x)| \geq c|g(x)|$ as $x \to \infty$ for a constant $c > 0$.

B. System Model

Let $\mathbf{x}^1, \mathbf{x}^2, \cdots, \mathbf{x}^S$ be jointly *K-sparse* vectors with a support set $\mathcal{I}$ that belongs to

$$\mathcal{S} := \{\mathcal{H} | \mathcal{H} \subset \{1, 2, \cdots, N\}, |\mathcal{H}| = K\}.$$

Thus, the number of nonzero coefficients of each sparse vector is K, the indices of the nonzero coefficients of all the sparse vectors are the same and the indices belong to the support set.

In noisy MMV with different sensing matrices, each sparse vector is sampled by its own sensing matrix, i.e.,

$$\mathbf{y}^s = \mathbf{F}^s\mathbf{x}^s + \mathbf{n}^s \quad s = 1, 2, \cdots S \tag{3}$$

where all the sensing matrices have i.i.d. Gaussian elements with a zero mean and a unit variance, and all the noise vectors have i.i.d. Gaussian elements with a zero mean and a σ^2 variance. We assume that all the noise vectors and all the sensing matrices are mutually independent. Then, we let $x_{\min}$ be the smallest nonzero magnitude of all the sparse vectors and $\mathrm{SNR}_{\min}$ be the minimum signal-to-noise ratio given by

$$\mathrm{SNR}_{\min} := x_{\min}^2/\sigma^2. \tag{4}$$

C. Problem Formulation

We extend Akcakaya and Tarokh [21]'s decoder for noisy MMV with different sensing matrices. It takes all the measurement vectors as its input and yields a support set decision as its output

$$d : \left\{\forall_s \left(\mathbf{y}^s, \mathbf{F}^s\right)\right\} \mapsto \hat{\mathcal{I}} \in \mathcal{S}, \quad s = 1, 2, \cdots, S.$$

Its decision rules are given in Definition 1.

Definition 1: All the measurement vectors $\left\{\mathbf{y}^1, \mathbf{y}^2, \cdots, \mathbf{y}^S\right\}$ and a set $\mathcal{J} \in \mathcal{S}$ are δ *jointly typical* if the rank of $\mathbf{F}^s_{\mathcal{J}}$, $s = 1, \ldots, S$, is K and

$$\left|\left(\sum\nolimits_{s=1}^{S} \left\|\mathbf{Q}\left(\mathbf{F}^s_{\mathcal{J}}\right)\mathbf{y}^s\right\|_2^2\right) - S(M-K)\sigma^2\right| < SM\delta. \tag{5}$$

As each sensing matrix contains i.i.d. Gaussian elements, the rank of each $\mathbf{F}^s_{\mathcal{J}}$, $s = 1, \ldots, S$, is K with a high probability. The decision rule is to find sets that satisfy (5) for all the given measurement vectors and $\delta > 0$. In the entire paper, the support set is denoted by $\mathcal{I}$ and any incorrect support set is denoted by $\mathcal{J}$, where their cardinalities are K, i.e., $|\mathcal{I}| = |\mathcal{J}| = K$.

We define the failure events, wherein the joint typical decoder fails to reconstruct the correct support set. First,

$$\mathcal{E}^c_{\mathcal{I}} := \left\{\left|\left(\sum\nolimits_{s=1}^{S} \left\|\mathbf{Q}\left(\mathbf{F}^s_{\mathcal{I}}\right)\mathbf{y}^s\right\|_2^2\right) - S(M-K)\sigma^2\right| \geq SM\delta\right\} \tag{6}$$

implies that the correct support set is not δ jointly typical with all the measurement vectors. Next, for any $\mathcal{J} \in \mathcal{S} \setminus \mathcal{I}$,

$$\mathcal{E}_{\mathcal{J}} := \left\{\left|\left(\sum\nolimits_{s=1}^{S} \left\|\mathbf{Q}\left(\mathbf{F}^s_{\mathcal{J}}\right)\mathbf{y}^s\right\|_2^2\right) - S(M-K)\sigma^2\right| < SM\delta\right\} \tag{7}$$

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:47:30 UTC from IEEE Xplore. Restrictions apply.

implies that an incorrect support set is δ jointly typical with all the measurement vectors. Based on these failure events, we define a failure probability and give its upper bound as follows:

$$\begin{aligned} p_{err} &:= \mathbb{P}\left\{\hat{\mathcal{I}} \neq \mathcal{I} \middle| \mathbf{x}^1, \cdots, \mathbf{x}^S\right\} \\ &= \mathbb{P}\left\{\mathcal{E}_{\mathcal{I}}^c \bigcup_{\mathcal{J}\in\mathcal{S}\setminus\mathcal{I}} \mathcal{E}_{\mathcal{J}}\right\} \\ &\leq \mathbb{P}\left\{\mathcal{E}_{\mathcal{I}}^c\right\} + \sum_{\mathcal{J}\in\mathcal{S}\setminus\mathcal{I}} \mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\} \end{aligned} \tag{8}$$

where $\mathbb{P}\left\{\mathcal{E}_{\mathcal{I}}^c\right\}$ is taken with respect to all the noise vectors and $\mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\}$ is taken with respect to all the noise vectors and all the sensing matrices. We establish Lemmas 1 and 2 given in Appendix A to give upper bounds of the probabilities of the failure events. Combining these lemmas with (8) yields

$$\begin{aligned} p_{err} &\leq \mathbb{P}\left\{\mathcal{E}_{\mathcal{I}}^c\right\} + \sum_{\mathcal{J}\in\mathcal{S}\setminus\mathcal{I}} \mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\} \\ &\leq 2p\left(d_1\right) + \binom{N}{K} p\left(d_{2,\alpha^*} - 1\right) \end{aligned}$$

where p is defined in (31), $d_1 = \frac{M\delta}{(M-K)\sigma^2}$, $d_{2,\alpha^*} = \frac{(M-K)\sigma^2+M\delta}{(M-K)\alpha^*}$, and $\alpha^* = \sigma^2 + x_{\min}^2$.

It is of interest to examine why $\mathbb{P}\left\{\mathcal{E}_{\mathcal{I}}^c\right\}$ depends only on the noise vectors. As shown in Lemma 3, the random variable to define the event $\mathcal{E}_{\mathcal{I}}^c$ in (6) is $\sum_{s=1}^{S} \left\|\mathbf{Q}\left(\mathbf{F}_{\mathcal{I}}^s\right)\mathbf{y}^s\right\|_2^2/\sigma^2$, where the measurement vector in (3) consists of the two parts: the noise part $\mathbf{n}^s$ and the signal part $\mathbf{F}_{\mathcal{I}}^s\mathbf{x}_{\mathcal{I}}^s$. The signal part belongs to the space spanned by the columns of $\mathbf{F}_{\mathcal{I}}^s$. Then, as specified in (2), the orthogonal projection matrix $\mathbf{Q}\left(\mathbf{F}_{\mathcal{I}}^s\right)$ maps the measurement vector to the space orthogonal onto the space spanned by the columns of $\mathbf{F}_{\mathcal{I}}^s$. Thus, the random variable is a function of the noise vectors only.

III. Main Results

As the main contribution of this paper, this section presents sufficient and necessary conditions on M for reliable support set reconstruction, i.e., p_{err} converges to zero, in noisy MMV with different sensing matrices. We then interpret the conditions to demonstrate the benefits facilitated by the joint sparsity structure.

A. Sufficient Conditions on M

In [18] and [21], the authors have shown that fewer measurements M for a reliable support set reconstruction are required for noisy SMV in the linear sparsity regime, compared to the sublinear sparsity regime. Based on the results of [18] and [21], we are motivated to examine if the same result can be observed in noisy MMV with different sensing matrices.

Theorem 1: For any$\rho > 1$, we let $\delta = \rho^{-1}\left(1 - K/M\right)x_{\min}^2$. If the number of measurements satisfies

$$M > K + \upsilon_1 \frac{K}{S} \tag{9}$$

then the failure probability p_{err} defined in (8) converges to zero in the linear sparsity regime, i.e., $\lim_{N\to\infty} \frac{K}{N} = \beta \in (0, 1/2)$, where

$$\upsilon_1 = -\frac{2\left(1 - \log\beta\right)}{\log\left(1 - \frac{1-\rho^{-1}}{1+\mathrm{SNR}_{\min}^{-1}}\right) + \frac{1-\rho^{-1}}{1+\mathrm{SNR}_{\min}^{-1}}} > 0. \tag{10}$$

Also, under the same conditions on ρ and δ, if the number of measurements satisfies

$$M > K + \upsilon_2 \frac{K}{S} \log \frac{N}{K} \tag{11}$$

then the failure probability p_{err} defined in (8) converges to zero in the sublinear sparsity regime, i.e., $\lim_{N\to\infty} \frac{K}{N} = 0$, where

$$\upsilon_2 = -\frac{2}{\log\left(1 - \frac{1-\rho^{-1}}{1+\mathrm{SNR}_{\min}^{-1}}\right) + \frac{1-\rho^{-1}}{1+\mathrm{SNR}_{\min}^{-1}}} > 0. \tag{12}$$

Proof: The proof is given in Appendix C.

In terms of N, K, and S, the asymptotic order of the sufficient condition on M for the linear sparsity regime is $\Omega\left(K + \frac{K}{S}\right)$, whereas the order for the sublinear sparsity regime is $\Omega\left(\frac{K}{S}\log\frac{N}{K}\right)$. It confirms that fewer measurements are required in the linear sparsity regime, compared to the sublinear sparsity regime. Next, from the sufficient conditions, we observe an inverse relationship between M and S, owing to the joint sparsity structure. This relationship implies that taking more measurement vectors S reduces the number of required measurements M for reliable support set reconstruction. Then, the relationship can be used for explaining the empirical results of Caione *et al.* [10] and Wu *et al.* [14]. In [10], the authors have reported that the number of transmitted bits per sensor could be inversely reduced by the number of sensors, which implies that the transmission cost of each sensor could be saved. The result can be confirmed by our inverse relationship by considering S and M as the number of sensors and the number of transmitted bits per sensor, respectively. In [14], S and M are considered as the number of DTIs and the number of samples of each DTI, respectively. Again, it has been observed from [14] that the joint sparsity structure enabled the number of samples of each DTI to be inversely reduced by the number of DTIs, reducing the acquisition time for each DTI. These results can be confirmed by our inverse relationship.

Theorem 2: For any $\rho > 1$, we let $\delta = \rho^{-1}\left(1 - K/M\right)x_{\min}^2$, N and K be fixed. If the number of measurements satisfies $M \geq K + 1$, the failure probability p_{err} defined in (8) converges to zero as the number of measurement vectors is increased to the infinity.

Proof: The proof is given in Appendix C.

Theorem 2 suggests that with $M \geq K + 1$, reliable support set reconstruction for noisy MMV with different sensing matrices is possible when a large number of measurement vectors is available. The sufficient conditions in Theorem 1, i.e., (9) and (11) have $\mathrm{SNR}_{\min}$ values as shown in (10) and (12). They disappear in the sufficient condition of Theorem 2, i.e., $M \geq K + 1$. The support set reconstruction problem becomes

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:47:30 UTC from IEEE Xplore. Restrictions apply.

robust against noise when the number of measurement vectors is large.

B. Discussions on the Sufficient Conditions

We now examine the effect of $\mathrm{SNR}_{\min}$ on the sufficient conditions of Theorem 1. The aim is to determine the relationship among S, M and $\mathrm{SNR}_{\min}$ for reliable support set reconstruction.

Corollary 1: For any $\rho > 1$, we let $\delta = \rho^{-1}(1 - K/M)\, x_{\min}^2$. The sufficient conditions of Theorem 1 are rewritten as

$$M > K + \left(\frac{\sqrt[-1]{S} + \left(\sqrt{S} \times \mathrm{SNR}_{\min}\right)^{-1}}{1 - \rho^{-1}} \right)^2 4K \log \frac{N}{K} \quad (13)$$

in the sublinear sparsity regime, i.e., $\lim_{N\to\infty} \frac{K}{N} = 0$, and

$$M > K + \left(\frac{\sqrt[-1]{S} + \left(\sqrt{S} \times \mathrm{SNR}_{\min}\right)^{-1}}{1 - \rho^{-1}} \right)^2 4K (1 - \log \beta) \quad (14)$$

in the linear sparsity regime, i.e., $\lim_{N\to\infty} \frac{K}{N} = \beta \in (0, 1/2)$.

Proof: The proof is given in Appendix D.

Corollary 1 suggests that for a fixed M, reliable support set reconstruction is possible as the number of measurement vectors S is increased to infinity, although $\mathrm{SNR}_{\min}$ is low. Namely, we observe a noise reduction effect, which shows that using the joint sparsity structure leads to an increase in $\mathrm{SNR}_{\min}$ or a decrease in σ^2 by the square root of S. This effect can explain the improvement in the reconstruction quality of the DTIs, as empirically reported in [14].

We then improve our noise reduction effect by considering that $\mathrm{SNR}_{\min}$ is larger than a certain value.

Corollary 2: For any $\rho > 3$, we let $\delta = \rho^{-1}(1 - K/M)\, x_{\min}^2$ and $\alpha = 2/3$. If

$$\mathrm{SNR}_{\min} \geq \frac{\alpha}{1 - \rho^{-1} - \alpha} = \frac{2\rho}{\rho - 3}, \quad (15)$$

the sufficient conditions of Theorem 1 are rewritten as

$$M > K + \frac{S^{-1} + (S \times \mathrm{SNR}_{\min})^{-1}}{1 - \rho^{-1}} 4K \log \frac{N}{K} \quad (16)$$

in the sublinear sparsity regime, i.e., $\lim_{N\to\infty} \frac{K}{N} = 0$, and

$$M > K + \frac{S^{-1} + (S \times \mathrm{SNR}_{\min})^{-1}}{1 - \rho^{-1}} 4K (1 - \log \beta) \quad (17)$$

in the linear sparsity regime, i.e., $\lim_{N\to\infty} \frac{K}{N} = \beta \in (0, 1/2)$.

Proof: The proof is given in Appendix D.

First of all, Corollary 2 requires $\rho > 3$ to ensure that the lower bound in (15) is positive. A simple computation shows that Corollary 2 requires fewer measurements in both the regimes compared to Corollary 1 because

$$\begin{aligned} \left(\frac{\sqrt[-1]{S} + \left(\sqrt{S} \times \mathrm{SNR}_{\min}\right)^{-1}}{1 - \rho^{-1}} \right)^2 &= S^{-1} \left(\frac{1 + \mathrm{SNR}_{\min}^{-1}}{1 - \rho^{-1}} \right)^2 \\ &\geq S^{-1} \left(\frac{1 + \mathrm{SNR}_{\min}^{-1}}{1 - \rho^{-1}} \right) \\ &= \frac{S^{-1} + (S \times \mathrm{SNR}_{\min})^{-1}}{1 - \rho^{-1}} \end{aligned}$$

where the second inequality is owing to $\frac{1+\mathrm{SNR}_{\min}^{-1}}{1-\rho^{-1}} = \frac{1}{t} > 1$ for any $\rho > 3$ and t defined in (61). Besides, Corollary 2 improves the noise reduction effect observed in Corollary 1 by showing that $\mathrm{SNR}_{\min}$ is increased by S for the region of $\mathrm{SNR}_{\min}$ in (15).

Theorem 2 suggests, it is to be noted, that $M = K+1$ is sufficient for reliable support set reconstruction if S is sufficiently large with a fixed N and K. Then, it would be interesting to determine how large S should be required for achieving the minimum number of measurements at each sensor, i.e., $M = K + 1$. In wireless sensor networks [34], energy sources used in sensors are very limited due to limitation of sensor sizes. Thus, minimizing the energy used for transmission of data at each sensor which often leads to extending the lifetime of the sensor battery is a value of importance. This point is noted in Caione *et al.* [10] as an advantage of using distributed compressed sensing on joint sparse model-2 signal ensembles (see Section V there). Corollary 3 which aims to provide a sufficient condition on S for achieving $M = K + 1$ thus is motivated.

Corollary 3: Let N and K be fixed and finite. For any $\rho > 1$, we let $\delta = \rho^{-1}(K+1)^{-1} x_{\min}^2$ and $M = K+1$. If the number of measurement vectors satisfies

$$S > \underbrace{\left(\log\left(\binom{N}{K} + 2\right) - \log \varepsilon\right) \times \max\left[\left|\frac{1}{\log \mu_{\mathcal{I}}}\right|, \left|\frac{1}{\log \mu_{\mathcal{J}}}\right|\right]}_{:=S^*} \quad (18)$$

reliable support set reconstruction is possible, i.e., $p_{err} < \varepsilon$ for sufficiently small $\varepsilon \in (0, 1)$, where $\log \mu_{\mathcal{I}}$ and $\log \mu_{\mathcal{J}}$ are defined in (63) and (65), respectively. The sufficient condition on S is decreasing with respect to $\mathrm{SNR}_{\min}$.

Proof: The proof is given in Appendix D.

To the best of our knowledge, the sufficient conditions on S for a reliable support set reconstruction have not yet been developed. A similar result has been reported by Tang and Nehorai [26], in which they reported that $M = \Omega\left(K \log \frac{N}{K}\right)$ and $S = \frac{\log N}{\log \log N}$ suffice for a reliable support set reconstruction in noisy MMV with the same sensing matrix, as N is sufficiently large.

It is of interest to examine whether the sufficient condition S^* in (18) is good. For this, we implement the joint typical decoder in (5) and conduct experiments for different values of $\mathrm{SNR}_{\min}$ and K, for a fixed $N = 50$. We count the number of failure occurrences, wherein the joint typical

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:47:30 UTC from IEEE Xplore. Restrictions apply.

decoder fails to reconstruct the support set. We obtain the smallest S^{emp} such that the ratio of the failure occurrences is smaller than $\varepsilon = 0.01$. By comparing S^{emp} with S^* in (18), we see that S^* approaches S^{emp}, as $\mathrm{SNR}_{\min}$ is sufficiently large. For example, we see that $S^{emp} = 8$ and $S^* = 12$ at $\mathrm{SNR}_{\min} = 20$ [dB], $K = 2$, and $S^{emp} = 5$ and $S^* = 6$ at $\mathrm{SNR}_{\min} = 30$ [dB], $K = 2$. A similar trend is observed with a bigger K, i.e., at $K = 5$. For example, we see that $S^{emp} = 12$ and $S^* = 19$ at $\mathrm{SNR}_{\min} = 20$ [dB], and $S^{emp} = 7$ and $S^* = 10$ at $\mathrm{SNR}_{\min} = 30$ [dB].

Fletcher *et al.* [19] have reported that the ML decoder requires $M = K + 1$ measurements for a reliable support set reconstruction in noisy SMV, when the signal-to-noise ratio is sufficiently large. This result can be observed from Corollary 3. Specifically, we assume that $\mathrm{SNR}_{\min}$ is sufficiently large for a fixed N and K. Then, from (63) and (65), it is easy to see that

$$\lim_{\mathrm{SNR}_{\min}\to\infty} \log \mu_{\mathcal{I}} = -\infty,$$
$$\lim_{\mathrm{SNR}_{\min}\to\infty} \log \mu_{\mathcal{J}} = 2^{-1}\left(1 - \rho^{-1} - \log \rho\right).$$

Hence, (18) is simplified to

$$S > \left(\log\left(\binom{N}{K} + 2\right) - \log \varepsilon\right) \times \left|2\left(1 - \rho^{-1} - \log \rho\right)^{-1}\right|. \tag{19}$$

Note that N, K, and ε are fixed. Thus, for a large ρ, we have

$$\left|1 - \rho^{-1} - \log \rho\right| \gg 2\left(\log\left(\binom{N}{K} + 2\right) - \log \varepsilon\right), \tag{20}$$

which leads to $S \geq 1$. This result suggests that the joint typical decoder requires $M = K+1$ measurements for reliable support set reconstruction in noisy SMV, whenever $\mathrm{SNR}_{\min}$ is sufficiently large and ρ satisfies (20).

C. Necessary Condition on M

We specify a necessary condition that must be satified by a decoder for reliable support set reconstruction in noisy MMV with different sensing matrices. Unlike the sufficient conditions of Theorem 1, the necessary condition is presented for a finite N and K.

We begin by transforming (3) into

$$\underbrace{\begin{bmatrix}\mathbf{y}^1\\ \vdots \\ \mathbf{y}^S\end{bmatrix}}_{=:\mathbf{y}\in\mathbb{R}^{SM}} = \underbrace{\begin{bmatrix}\mathbf{F}^1 & & \\ & \ddots & \\ & & \mathbf{F}^S\end{bmatrix}}_{=:\tilde{\mathbf{F}}\in\mathbb{R}^{SM\times SN}} \underbrace{\begin{bmatrix}\mathbf{x}^1\\ \vdots \\ \mathbf{x}^S\end{bmatrix}}_{=:\mathbf{x}\in\mathbb{R}^{SN}} + \underbrace{\begin{bmatrix}\mathbf{n}^1\\ \vdots \\ \mathbf{n}^S\end{bmatrix}}_{=:\mathbf{n}\in\mathbb{R}^{SM}} \tag{21}$$

where $\mathbf{x}$ is an *SK-sparse* vector belonging to an infinite set

$$\mathcal{X}_{x_{\min}} := \left\{\mathbf{x} \in \mathbb{R}^{SN} \,\middle|\, |x(i)| \geq x_{\min}, \forall i \in \mathcal{I}, |\mathcal{I}| = SK\right\}$$

where $x(i)$ is the ith element of $\mathbf{x}$ and $\mathcal{I}$ is the support set of $\mathbf{x}$. Owing to the joint sparsity structure, the number of possible support sets is $\binom{N}{K}$. Then, we define a failure probability as:

$$p_{err} := \mathbb{E}_{\tilde{\mathbf{F}}} \sup_{\mathbf{x}\in\mathcal{X}_{x_{\min}}} \mathbb{P}\left\{\hat{\mathcal{I}} \neq \mathcal{I} \,\middle|\, \mathbf{x}, \tilde{\mathbf{F}}\right\} \tag{22}$$

where $\hat{\mathcal{I}}$ is an estimate of the support set based on $\mathbf{y}$ and $\tilde{\mathbf{F}}$ in (21). Then, Lemma *III*-3 of [20] yields

$$\sup_{\mathbf{x}\in\mathcal{X}_{x_{\min}}} \mathbb{P}\left\{\hat{\mathcal{I}} \neq \mathcal{I} \,\middle|\, \mathbf{x}, \tilde{\mathbf{F}}\right\} \geq \min_{\hat{\mathbf{x}}\in\mathcal{X}_{\{x_{\min}\}}} \max_{\mathbf{x}\in\mathcal{X}_{\{x_{\min}\}}} \mathbb{P}\left\{\hat{\mathbf{x}} \neq \mathbf{x} \,\middle|\, \mathbf{x}, \tilde{\mathbf{F}}\right\} \tag{23}$$

where $\hat{\mathbf{x}}$ is an estimate for $\mathbf{x}$ based on $\mathbf{y}$ and $\tilde{\mathbf{F}}$ in (21) and

$$\mathcal{X}_{\{x_{\min}\}} := \left\{\mathbf{x} \in \mathbb{R}^{SN} \,\middle|\, x(i) = x_{\min}, \forall i \in \mathcal{I}, |\mathcal{I}| = SK\right\}$$

which is a finite set. Assume that $\mathbf{x}$ is uniformly distributed over this finite set. Applying Fano's inequality [24] to (23) yields

$$\max_{\mathbf{x}\in\mathcal{X}_{\{x_{\min}\}}} \mathbb{P}\left\{\hat{\mathbf{x}} \neq \mathbf{x} \,\middle|\, \mathbf{x}, \tilde{\mathbf{F}}\right\} \geq \mathbb{P}\left\{\hat{\mathbf{x}} \neq \mathbf{x} \,\middle|\, \tilde{\mathbf{F}}\right\} \geq 1 - \frac{\mathbb{I}\left(\mathbf{x}; \mathbf{y} \middle| \tilde{\mathbf{F}}\right) + \log 2}{\log\left(\left|\mathcal{X}_{\{x_{\min}\}}\right| - 1\right)} \tag{24}$$

where $\mathbf{x}$ and $\hat{\mathbf{x}}$ belong to the finite set $\mathcal{X}_{\{x_{\min}\}}$ and $\mathbb{I}(\mathbf{x};\ \mathbf{y})$ is the mutual information between $\mathbf{x}$ and$\mathbf{y}$. We get a necessary condition on M to ensure that the lower bound in (24) is bounded away from zero, as follows:

Theorem 3: Let N and K are fixed and finite. In (21), if the number of measurements satisfies

$$M < \frac{2K \log \frac{N}{K} - 2 \log 2}{S \log\left(1 + K \times \mathrm{SNR}_{\min}\right)} \tag{25}$$

then the failure probability p_{err} defined in (22) is bounded away from zero.

Proof: The proof is given in Appendix C.

IV. Relations to the Existing Information-Theoretic Results

A. Relations to Noisy MMV With the Same Sensing Matrix [27]

Jin and Rao [27] have exploited the Chernoff bound to obtain a tight sufficient condition on M for a reliable support set reconstruction for noisy MMV with the same sensing matrix in the sublinear sparsity regime. Owing to the complicated form of their sufficient condition, they could not clearly show the benefits facilitated by the joint sparsity structure. Thus, they simplified their condition under scenarios such as: *i*) a low noise-level scenario and *ii*) a scenario with S identical sparse vectors. In Table I, we summarize our sufficient conditions on M, and compare them to that of [27] under the low noise-level scenario in the sublinear sparsity regime.

First, in a low noise-level scenario, as shown in Table I, the sufficient condition [27] for noisy MMV with the same sensing matrix is

$$M = \Omega\left(\frac{K \log N}{\min(K, S)}\right). \tag{26}$$

If $S < K$, the sufficient conditions (11) and (26) have the same order, implying that there is no significant performance gap in the support set reconstruction between the models. However, if $S > K$, (26) is $M = \Omega(\log N)$, whereas (11) is

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:47:30 UTC from IEEE Xplore. Restrictions apply.

TABLE I
SUFFICIENT CONDITIONS ON M FOR SUPPORT SET RECONSTRUCTION

	This paper	Yuzhe and Rao [27]
Linear sparsity regime $\lim_{N\to\infty} \frac{K}{N} = \beta \in (0, 1/2)$	$M = \Omega\left(K + \frac{K}{S}\right)$	Not presented
Sublinear sparsity regime $\lim_{N\to\infty} \frac{K}{N} = 0$	$M = \Omega\left(\frac{K}{S}\log\frac{N}{K}\right)$	$M = \Omega\left(\frac{K\log N}{\min(K,S)}\right)$
N and K are finite ($\mathrm{SNR}_{\min} \to \infty$ or $S \to \infty$)	$M \geq K+1$	Not presented

$M = \Omega\left(\frac{K}{S}\log N\right)$. It implies that noisy MMV with different sensing matrices is superior to noisy MMV with the same sensing matrix or $S > K$, with respect to M for reliable support set reconstruction. The result of this comparison supports the perception presented in Section I-C, wherein a more reliable support set reconstruction could be expected in a noiseless MMV with different sensing matrices owing to the linear independency of the measurement vectors. Moreover, it validates the perception, even in the presence of noise.

Second, we consider a scenario with S identical sparse vectors. Then, the sufficient condition of [27] is

$$M = \Omega\left(\frac{K\log N}{\log\left(1 + S\,\|\mathbf{x}\|_2^2/\sigma^2\right)}\right). \tag{27}$$

From (27), we observe that σ^2 is reduced by a factor of S. However, the noise reduction effect for noisy MMV with the same sensing matrix requires a restriction, where all the sparse vectors should be identical, which can be hardly achieved in practice. In contrast, the noise reduction effect for noisy MMV with different sensing matrices does not require this restriction, as shown in Corollaries 1 and 2.

B. Relations to Noisy SMV [21]

Akcakaya and Tarokh [21] have used the joint typical decoder to establish the sufficient conditions on M for a reliable support set reconstruction in noisy SMV. They exploited the exponential inequalities [32] to obtain the upper bounds on the sum of the weighted chi-square random variables. In this subsection, we demonstrate that the approaches developed in this paper are superior to the use of the exponential inequalities. Thus, we use the exponential inequalities to generalize their bounds for noisy MMV with different sensing matrices. We give Propositions 1 and 2 to prove that the generalized bounds are worse than the bounds of Lemmas 1 and 2.

Proposition 1: For any positive δ, we have

$$\mathbb{P}\left\{\mathcal{E}_{\mathcal{I}}^c\right\} \leq 2p\left(d_1\right) \leq 2p_{1,\exp}$$

where both $p\left(d_1\right)$ and d_1 are given in Lemma 1, and

$$p_{1,\exp} := \exp\left(-\frac{S\delta^2}{4\sigma^4}\frac{M^2}{M - K + 2\delta M/\sigma^2}\right). \tag{28}$$

Proof: The proof is given in Appendix E.

Proposition 2: For any $\mathcal{J} \in \mathcal{S}\setminus\mathcal{I}$ and any $\delta > 0$ such that

$$0 < \delta < (1 - K/M)\,x_{\min,\mathcal{J}}^2, \tag{29}$$

we have

$$\mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\} \leq p\left(d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})} - 1\right) \leq p_{2,\mathcal{J},\exp}$$

where both $p\left(d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})} - 1\right)$ and $d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})}$ are given in Lemma 2 and

$$p_{2,\mathcal{J},\exp} := \exp\left(-\frac{S^2\,(M-K)}{4\sum_{s=1}^{S}\alpha_{\mathcal{J},s}^2}\left(x_{\min,\mathcal{J}}^2 - \frac{M\delta}{M-K}\right)^2\right) \tag{30}$$

and $\alpha_{\mathcal{J},s}$ is defined in (39) and $x_{\min,\mathcal{J}}^2$ is defined in (43).

Proof: The proof is given in Appendix E.

If $S = 1$, we can see that $p_{1,\exp}$ and $p_{2,\mathcal{J},\exp}$ are equivalent to the bounds of Akcakaya and Tarokh [21]. Propositions 1 and 2 state that the bounds on the failure probability of Lemmas 1 and 2 are tighter than the bounds of [21] for noisy SMV.

V. CONCLUSIONS

We have studied a support set reconstruction problem for CS with noisy MMV with different sensing matrices. The union and Chernoff bounds have been used to obtain the upper bound of the failure probability of the support set reconstruction, and Fano's inequality has been used to obtain the lower bound of this failure probability. As we have obtained the upper bound by analyzing an exhaustive search decoder, the bound is used to measure the performance of CS with noisy MMV with different sensing matrices. We have then developed the necessary and sufficient conditions in terms of the sparsity K, the ambient dimension N, the number of measurements M, the number of measurement vectors S, and the minimum signal-to-noise ratio $\mathrm{SNR}_{\min}$. They can be useful as guidelines to determing the system parameters in various CS applications with noisy MMV with different sensing matrices.

The conditions are interpreted to provide theoretical explanations for the benefits facilitated by the joint sparsity structure in noisy MMV with different sensing matrices:

i. From the sufficient conditions of Theorem 1, we have observed an inverse relationship between M and S. Due to the inverse relation, we take fewer measurements M per each measurement vector for reliable support set reconstruction by taking more measurement vectors S.

ii. From the sufficient conditions of Corollaries 1 and 2, we have observed a noise reduction effect, which shows that the usage of the joint sparsity structure results in an increase in $\mathrm{SNR}_{\min}$ or a decrease in σ^2 by a factor of S. Therefore, the support set reconstruction can be robust against noise as the number of measurement vectors is increased to infinity.

iii. From Theorem 2, we have shown that $M = K + 1$ is achieved for a fixed N and K, as S is sufficiently large. From Corollary 3, we have provided the sufficient condition on S to reconstruct the support set for a fixed N, K, and $M = K + 1$.

The theoretical explanations confirm the benefits of the joint sparsity structure, as empirically shown in CS applications with noisy MMV with different sensing matrices [10], [14].

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:47:30 UTC from IEEE Xplore. Restrictions apply.

We have compared our sufficient conditions for noisy MMV with different sensing matrices with the other existing results [27] for noisy MMV with the same sensing matrix. For a low-level noise scenario with $S \geq K$, we have shown that the number of measurements for reliable support set reconstruction for noisy MMV with different sensing matrices is lesser than that for noisy MMV with the same sensing matrix. Also, [27] has shown the noise reduction effect. This was achieved under a rather restricted condition though, i.e., all sparse vectors are the same. While such a restricted condition is not required in the noisy MMV problem with *different* sensing matrices studied in this paper, the noise reduction effect has also been observed, which is a significant improvement.

APPENDIX A
LEMMAS 1 AND 2

This section presents Lemmas 1 and 2, which give upper bounds of the probabilities of the failure events defined in (6) and (7), respectively. Also, for simplicity, we define

$$p(x) = \exp\left(-\frac{S(M-K)}{2}x\right)(1+x)^{\frac{S(M-K)}{2}}. \tag{31}$$

Lemma 1: For any positive δ, we have

$$\begin{aligned}\mathbb{P}\left\{\mathcal{E}_{\mathcal{I}}^{c}\right\} &\leq 2\exp\left(-\frac{S(M-K)}{2}d_1\right)(1+d_1)^{\frac{S(M-K)}{2}} \\ &= 2p(d_1)\end{aligned} \tag{32}$$

where the function p is defined in (31), and

$$d_1 := \frac{M\delta}{(M-K)\sigma^2} > 0. \tag{33}$$

Proof: From (6), we have

$$\mathbb{P}\left\{\mathcal{E}_{\mathcal{I}}^{c}\right\} = \mathbb{P}\left\{\mathrm{Z}_{\mathcal{I}} \leq W_1\right\} + \mathbb{P}\left\{\mathrm{Z}_{\mathcal{I}} \geq W_2\right\} \tag{34}$$

where $\mathrm{Z}_{\mathcal{I}}$ is defined in Lemma 3, and

$$W_i = S(M-K) + (-1)^i SM\delta/\sigma^2, \quad i = 1,\ 2.$$

Applying the Chernoff bound [31] to (34) yields

$$\begin{aligned}\mathbb{P}\left\{\mathcal{E}_{\mathcal{I}}^{c}\right\} &\leq \sum\nolimits_{i=1}^{2} \exp(-t_i W_i)\,\mathbb{E}\left[\exp(t_i \mathrm{Z}_{\mathcal{I}})\right] \\ &= \sum\nolimits_{i=1}^{2} \underbrace{\exp(-t_i W_i)(1-2t_i)^{-S(M-K)/2}}_{=:f(t_i;W_i)}\end{aligned} \tag{35}$$

where the equality is from Lemma 3, $t_1 < 0$ and $t_2 \in \left(0, \frac{1}{2}\right)$. As each $f(t_i; W_i)$ is convex, $t_i = t_i^*$at $f^{(1)}(t_i; W_i) = 0$ yields the minimizer of $f(t_i; W_i)$, where

$$t_i^* = 2^{-1}\left(1 - W_i^{-1}S(M-K)\right), \quad i = 1,\ 2.$$

Thus, $f(t_i; W_i) \geq f\left(t_i^*; W_i\right)$ for each i. If $W_1 \leq 0$, it is clear that $\mathbb{P}\{\mathrm{Z}_{\mathcal{I}} \leq W_1\} = 0$ because $\mathrm{Z}_{\mathcal{I}}$ is quadratic. Thus,

$$\mathbb{P}\left\{\mathcal{E}_{\mathcal{I}}^{c}\right\} = \mathbb{P}\left\{\mathrm{Z}_{\mathcal{I}} \geq W_2\right\} \leq f\left(t_2^*; W_2\right) = p(d_1) \tag{36}$$

where $p(d_1)$ and d_1 are defined in (32) and (33), respectively. If $W_1 > 0$ then $f\left(t_1^*; W_1\right) \leq f\left(t_2^*; W_2\right)$ because

$$\begin{aligned}&\log f\left(t_1^*; W_1\right) - \log f\left(t_2^*; W_2\right) \\ &= S(M-K)\left[d_1 + 2\log(1-d_1) - 2\log(1+d_1)\right] < 0.\end{aligned}$$

Thus,

$$\begin{aligned}\mathbb{P}\left\{\mathcal{E}_{\mathcal{I}}^{c}\right\} &= f\left(t_1^*; W_1\right) + f\left(t_2^*; W_2\right) \leq 2f\left(t_2^*; W_2\right) \\ &= 2\exp\left(-\frac{S(M-K)}{2}d_1\right)(1+d_1)^{\frac{S(M-K)}{2}}.\end{aligned} \tag{37}$$

Finally, combining (36) and (37) leads to (32). ■

Lemma 2: Let $\mathcal{J} \in \mathcal{S} \setminus \mathcal{I}$ and a matrix $\mathbf{R}_{\mathcal{J}}$ be

$$\mathbf{R}_{\mathcal{J}} = \begin{bmatrix} \alpha_{\mathcal{J},1}\mathbf{I}_{M-K} & & \\ & \ddots & \\ & & \alpha_{\mathcal{J},S}\mathbf{I}_{M-K} \end{bmatrix} \tag{38}$$

where

$$\alpha_{\mathcal{J},s} := \sigma^2 + \left\|\mathbf{x}_{\mathcal{I}\setminus\mathcal{J}}^{s}\right\|_2^2 > 0. \tag{39}$$

Consider any positive δ such that

$$0 < \delta < (1 - K/M)\left(\lambda_{\min}\left(\mathbf{R}_{\mathcal{J}}\right) - \sigma^2\right)$$

where $\lambda_{\min}\left(\mathbf{R}_{\mathcal{J}}\right)$ is the smallest eigenvalue of $\mathbf{R}_{\mathcal{J}}$. Then,

$$\begin{aligned}\mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\} &\leq \exp\left(-\frac{S(M-K)}{2}\left(d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})} - 1\right)\right) d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})}^{\frac{S(M-K)}{2}} \\ &= p\left(d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})} - 1\right) \\ &\leq p\left(d_{2,\alpha^*} - 1\right)\end{aligned} \tag{40}$$

where the function p is defined in (31),

$$d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})} := \frac{(M-K)\sigma^2 + M\delta}{(M-K)\lambda_{\min}\left(\mathbf{R}_{\mathcal{J}}\right)} \in (0,1), \tag{41}$$

$$\alpha^* := \sigma^2 + x_{\min}^2, \tag{42}$$

and

$$x_{\min}^2 = \min_{\mathcal{J}\in\mathcal{S}\setminus\mathcal{I}} \underbrace{\min_{s\in\{1,2,\cdots,S\}} \left\|\mathbf{x}_{\mathcal{I}\setminus\mathcal{J}}^{s}\right\|_2^2}_{=:x_{\min,\mathcal{J}}^2}. \tag{43}$$

Proof: From (7), we have

$$\mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\} = \mathbb{P}\left\{\mathrm{Z}_{\mathcal{J}} < W_1\right\} - \mathbb{P}\left\{\mathrm{Z}_{\mathcal{J}} < W_2\right\} \leq \mathbb{P}\left\{\mathrm{Z}_{\mathcal{J}} < W_1\right\} \tag{44}$$

where $\mathrm{Z}_{\mathcal{J}}$ is defined in Lemma 4, and

$$W_i = S(M-K)\sigma^2 - (-1)^i SM\delta, \quad i = 1, 2. \tag{45}$$

Applying the Chernoff bound [31] to (44) yields for $t < 0$,

$$\begin{aligned}\mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\} &\leq \exp(-tW_1)\,\mathbb{E}\left[\exp\left(t\mathrm{Z}_{\mathcal{J}}\right)\right] \\ &= \exp(-tW_1)\prod\nolimits_{i=1}^{S(M-K)}\left(1 - 2t\lambda_i\left(\mathbf{R}_{\mathcal{J}}\right)\right)^{-1/2} \\ &\leq \exp(-tW_1)\left(1 - 2t\lambda_{\min}\left(\mathbf{R}_{\mathcal{J}}\right)\right)^{-S(M-K)/2} \\ &=: f(t; W_1)\end{aligned} \tag{46}$$

where the equality is from Lemma 4 and the second inequality is due to that all the eigenvalues are positive. We then define a function $h(t) := \log f(t; W_1)$. Then,

$$h^{(2)}(t) = 2S(M-K)\lambda_{\min}^2\left(\mathbf{R}_{\mathcal{J}}\right)\left(1 - 2t\lambda_{\min}\left(\mathbf{R}_{\mathcal{J}}\right)\right)^{-2} > 0$$

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:47:30 UTC from IEEE Xplore. Restrictions apply.

which implies that h is convex with respect to t. It leads to that f in (46) is logarithmically convex. Thus $t = t^*$ at $f^{(1)}(t; W_1) = 0$ yields the minimizer of $f(t; W_1)$ where

$$t^* = 2^{-1}\left(\lambda_{\min}^{-1}\left(\mathbf{R}_{\mathcal{J}}\right) - W_1^{-1} S(M-K)\right) < 0.$$

Substituting t^* in (46) yields

$$\begin{aligned}\mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\} &\le f\left(t^*; W_1\right)\\ &= \exp\left(-\frac{S(M-K)}{2}\left(d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})} - 1\right)\right) d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})}^{\frac{S(M-K)}{2}}\\ &= p\left(d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})} - 1\right) \end{aligned} \quad (47)$$

where $d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})}$ is defined in (41) and p is defined in (31).

Next, let $\beta = 2^{-1}S(M-K)$ and $x = d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})}$ in the upper bound (47). Then, we have $p(x-1) = x^{\beta}\exp(-\beta(x-1))$, where

$$\frac{\partial p(x-1)}{\partial x} = \beta x^{\beta}\exp(-\beta(x-1))\left(x^{-1} - 1\right) > 0 \quad (48)$$

and

$$\frac{\partial x}{\partial \lambda_{\min}\left(\mathbf{R}_{\mathcal{J}}\right)} = -x < 0. \quad (49)$$

Due to (48) and (49),

$$\begin{aligned}\frac{\partial p(x-1)}{\partial \lambda_{\min}\left(\mathbf{R}_{\mathcal{J}}\right)} &= \frac{\partial p(x-1)}{\partial x}\frac{\partial x}{\partial \lambda_{\min}\left(\mathbf{R}_{\mathcal{J}}\right)}\\ &= -\beta x^{\beta-1}\exp(-\beta(x-1))\left(x^{-1} - 1\right) < 0\end{aligned}$$

which shows that the upper bound in (47) is decreasing with respect to $\lambda_{\min}\left(\mathbf{R}_{\mathcal{J}}\right)$. Then, remind that the matrix in (38) is the covariance matrix of a multivariate Gaussian vector $\mathbf{b}$ in (58). Then for any incorrect support set, its smallest eigenvalue can be easily computed and lower bounded by

$$\lambda_{\min}\left(\mathbf{R}_{\mathcal{J}}\right) = \min_{s\in\{1,2,\cdots,S\}} \alpha_{\mathcal{J},s} = \sigma^2 + x_{\min,\mathcal{J}}^2 \ge \alpha^* \quad (50)$$

where $x_{\min,\mathcal{J}}^2$ is defined in (43) and α^* is defined in (42). Thus, for any incorrect support set $\mathcal{J} \in \mathcal{S}\setminus\mathcal{I}$, we conclude that

$$\mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\} \le p\left(d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})} - 1\right) \le p\left(d_{2,\alpha^*} - 1\right)$$

which completes the proof. ■

APPENDIX B
LEMMAS 3 AND 4

First of all, we give the Scharf's theorem [33] to compute the moment generating function of a quadratic random variable. We then make Lemmas 3 and 4 to give the moment generating functions of the random variables of $\mathcal{E}_{\mathcal{I}}^c$ and $\mathcal{E}_{\mathcal{J}}$ that were used in the proofs of Lemmas 1 and 2, respectively.

Scharf's Theorem [33, p. 64]: Let $\mathbf{b} \in \mathbb{R}^N$ be a multivariate Gaussian vector with a mean $\mathbf{m}$ and a covariance $\mathbf{R}$. Then a random variable $Q \triangleq (\mathbf{b}-\mathbf{m})^T(\mathbf{b}-\mathbf{m})$ is quadratic with $\mathbb{E}[Q] = \mathrm{tr}[\mathbf{R}]$, $\mathbb{V}[Q] = 2\mathrm{tr}\left[\mathbf{R}^T\mathbf{R}\right]$ and for any t

$$\mathbb{E}\left[\exp(tQ)\right] = \prod_{i=1}^{N}(1-2t\lambda_i(\mathbf{R}))^{-1/2}.$$

Lemma 3: In (6), define a quadratic random variable

$$\mathrm{Z}_{\mathcal{I}} := \sum_{s=1}^{S}\left\|\mathbf{Q}\left(\mathbf{F}_{\mathcal{I}}^s\right)\mathbf{y}^s\right\|_2^2/\sigma^2. \quad (51)$$

Then, $\mathbb{E}[\mathrm{Z}_{\mathcal{I}}] = S(M-K)$, $\mathbb{V}[\mathrm{Z}_{\mathcal{I}}] = 2S(M-K)$ and for any $0 < t < 0.5$,

$$\mathbb{E}\left[\exp(t\mathrm{Z}_{\mathcal{I}})\right] = (1-2t)^{-S(M-K)/2}. \quad (52)$$

Proof: The orthogonal projection matrix is decomposed as

$$\mathbf{Q}\left(\mathbf{F}_{\mathcal{I}}^s\right) = \mathbf{U}_{\mathcal{I}}^s\mathbf{D}^s\left(\mathbf{U}_{\mathcal{I}}^s\right)^T$$

where $\mathbf{D}^s$ is a diagonal matrix, whose first $M-K$ diagonals are ones and the remains are zeros, and $\mathbf{U}_{\mathcal{I}}^s$ is a unitary matrix. Then,

$$\begin{aligned}Z_{\mathcal{I}} &= \sum_{s=1}^{S}\left\|\mathbf{Q}\left(\mathbf{F}_{\mathcal{I}}^s\right)\mathbf{y}^s\right\|_2^2/\sigma^2 = \sum_{s=1}^{S}\left\|\mathbf{Q}\left(\mathbf{F}_{\mathcal{I}}^s\right)\mathbf{n}^s\right\|_2^2/\sigma^2\\ &= \sum_{s=1}^{S}\left\|\mathbf{D}^s\underbrace{\left(\mathbf{U}_{\mathcal{I}}^s\right)^T\mathbf{n}^s/\sigma^2}_{=:\mathbf{w}^s}\right\|_2^2 = \sum_{s=1}^{S}\left\|\mathbf{D}^s\mathbf{w}^s\right\|_2^2\end{aligned} \quad (53)$$

where $\mathbf{w}^s$ is a multivariate Gaussian vector with mean $\mathbf{0}_M$ and covariance $\mathbf{I}_M$. Since the first $M-K$ diagonal elements of each diagonal matrix are ones, we have

$$\begin{aligned}Z_{\mathcal{I}} &= \sum_{s=1}^{S}\left\|\mathbf{D}^s\mathbf{w}^s\right\|_2^2 = \sum_{s=1}^{S}\sum_{i=1}^{M-K}\left|w^s(i)\right|^2\\ &= \sum_{s=1}^{S}\left(\mathbf{w}_{\mathcal{P}}^s\right)^T\mathbf{w}_{\mathcal{P}}^s = \mathbf{w}^T\mathbf{w}\end{aligned} \quad (54)$$

which is quadratic, where

$$\mathbf{w}_{\mathcal{P}}^s = \begin{bmatrix} w^s(1) & w^s(2) & \cdots & w^s(M-K)\end{bmatrix}^T$$

and

$$\mathbf{w} = \begin{bmatrix}\left(\mathbf{w}_{\mathcal{P}}^1\right)^T & \left(\mathbf{w}_{\mathcal{P}}^2\right)^T & \cdots & \left(\mathbf{w}_{\mathcal{P}}^S\right)^T\end{bmatrix}^T. \quad (55)$$

In (53), $\mathbf{w}^s$ is determined by $\mathbf{U}_{\mathcal{I}}^s$ and $\mathbf{n}^s$. Since the elements of $\mathbf{U}_{\mathcal{I}}^s$ and $\mathbf{n}^s$ are independent, $\mathbf{w}^i$ and $\mathbf{w}^j$ are mutually independent for any $1 \le i \ne j \le S$. The covariance matrix of $\mathbf{w}$ is an identity matrix. Thus, applying the Scharf's theorem to $Z_{\mathcal{I}}$ completes the proof. ■

Lemma 4: In (7), for any $\mathcal{J} \in \mathcal{S}\setminus\mathcal{I}$, define a quadratic random variable

$$Z_{\mathcal{J}} := \sum_{s=1}^{S}\left\|\mathbf{Q}\left(\mathbf{F}_{\mathcal{J}}^s\right)\mathbf{y}^s\right\|_2^2. \quad (56)$$

Then, $\mathbb{E}\left[\mathrm{Z}_{\mathcal{J}}\right] = \mathrm{tr}\left[\mathbf{R}_{\mathcal{J}}\right]$, $\mathbb{V}\left[\mathrm{Z}_{\mathcal{J}}\right] = 2\mathrm{tr}\left[\mathbf{R}_{\mathcal{J}}^T\mathbf{R}_{\mathcal{J}}\right]$ and for any t,

$$\mathbb{E}\left[\exp(t\mathrm{Z}_{\mathcal{J}})\right] = \prod_{i=1}^{S(M-K)}\left(1-2t\lambda_i\left(\mathbf{R}_{\mathcal{J}}\right)\right)^{-1/2},$$

where $\mathbf{R}_{\mathcal{J}}$ is given in (38).

Proof: Similar to the proof of Lemma 3,

$$\mathbf{Q}\left(\mathbf{F}_{\mathcal{J}}^s\right) = \mathbf{U}_{\mathcal{J}}^s\mathbf{D}^s\left(\mathbf{U}_{\mathcal{J}}^s\right)^T$$

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:47:30 UTC from IEEE Xplore. Restrictions apply.

where $\mathbf{D}^s$ is a diagonal matrix, whose first $M-K$ diagonals are ones and the remains are zeros, and $\mathbf{U}^s_{\mathcal{J}}$ is a unitary matrix. Then,

$$Z_{\mathcal{J}} = \sum_{s=1}^{S} \left\| \mathbf{Q}\left(\mathbf{F}^s_{\mathcal{J}}\right) \mathbf{y}^s \right\|_2^2 = \sum_{s=1}^{S} \left\| \mathbf{Q}\left(\mathbf{F}^s_{\mathcal{J}}\right) \mathbf{c}^s \right\|_2^2 = \sum_{s=1}^{S} \left\| \mathbf{D}^s \underbrace{\left(\mathbf{U}^s_{\mathcal{J}}\right)^T \mathbf{c}^s}_{=:\mathbf{b}^s} \right\|_2^2 = \sum_{s=1}^{S} \left\| \mathbf{D}^s \ \mathbf{b}^s \right\|_2^2 \quad (57)$$

where $\mathbf{b}^s$ is a multivariate Gaussian vector with mean $\mathbf{0}_M$ and

$$\mathbb{V}\left[\mathbf{b}^s\right] = \left(\sigma^2 + \left\| \mathbf{x}^s_{\mathcal{I}\setminus\mathcal{J}} \right\|_2^2\right) \mathbf{I}_M$$

and $\mathbf{c}^s = \mathbf{n}^s + \sum_{u\in\mathcal{I}\setminus\mathcal{J}} \mathbf{f}^s_u x^s(u)$. Since the first $M-K$ diagonal elements of each diagonal matrix are ones, we have

$$\mathrm{Z}_{\mathcal{J}} = \sum_{s=1}^{S} \left\| \mathbf{D}^s \ \mathbf{b}^s \right\|_2^2 = \sum_{s=1}^{S} \sum_{i=1}^{M-K} \left| b^s(i) \right|^2 = \sum_{s=1}^{S} \left(\mathbf{b}^s_{\mathcal{P}}\right)^T \mathbf{b}^s_{\mathcal{P}} = \mathbf{b}^T \mathbf{b} \quad (58)$$

which is quadratic, where

$$\mathbf{b}^s_{\mathcal{P}} = \left[b^s(1) \quad b^s(2) \quad \cdots \quad b^s(M-K) \right]^T$$

and

$$\mathbf{b} = \left[\left(\mathbf{b}^1_{\mathcal{P}}\right)^T \quad \left(\mathbf{b}^2_{\mathcal{P}}\right)^T \quad \cdots \quad \left(\mathbf{b}^S_{\mathcal{P}}\right)^T \right]^T.$$

In (57), $\mathbf{b}^s$ is determined by $\mathbf{U}^s_{\mathcal{J}}$, $\mathbf{n}^s$ and $\left\{\mathbf{f}^s_u : u \in \mathcal{I}\setminus\mathcal{J}\right\}$. Since the elements of $\mathbf{U}^s_{\mathcal{J}}$, $\mathbf{n}^s$ and $\left\{\mathbf{f}^s_u : u \in \mathcal{I}\setminus\mathcal{J}\right\}$ are independent, $\mathbf{b}^i$ and $\mathbf{b}^j$ are mutually independent for any $1 \le i \ne j \le S$. The covariance matrix of $\mathbf{b}$ is diagonal as shown in (38). Thus, applying the Scharf's theorem to $Z_{\mathcal{J}}$ completes the proof. ■

APPENDIX C
PROOFS OF THEOREMS 1, 2 AND 3

A. Proof of Theorem 1

It is clear that K goes to infinity as N goes to infinity in the linear sparsity regime. Then, let $M = cK$ where $c > 1$. From (32),

$$\log \mathbb{P}\left\{\mathcal{E}^c_{\mathcal{I}}\right\} \le 2^{-1} SK(c-1) \underbrace{\left(\log(1+d_1) - d_1\right)}_{=:A} + \log 2$$

where $A < 0$ due to (33). Thus,

$$\lim_{N\to\infty} \mathbb{P}\left\{\mathcal{E}^c_{\mathcal{I}}\right\} \le \lim_{K\to\infty} \exp\left(2^{-1} SK(c-1)A + \log 2\right) = 0$$

implying that the probability that the correct support set is not δ jointly typical with all the measurement vectors vanishes.

Next, from (40),

$$\begin{aligned} \log \sum_{\mathcal{J}\in\mathcal{S}\setminus\mathcal{I}} \mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\} &\le \log\left(\binom{N}{K} p\left(d_{2,\alpha^*} - 1\right)\right) \\ &= \log\binom{N}{K} + 2^{-1} SK(c-1) \underbrace{\left(\log(1-t) + t\right)}_{=:\gamma} \\ &\le K \underbrace{\left(1 + \log\frac{N}{K} + 2^{-1} S(c_1 - 1)\gamma\right)}_{=:\eta} \end{aligned} \quad (59)$$

where the last inequality is due to

$$\binom{N}{K} \le \exp\left(K \log\frac{Ne}{K}\right). \quad (60)$$

In (59), $\gamma < 0$ for any t where

$$t = \frac{1 - \rho^{-1}}{1 + \mathrm{SNR}^{-1}_{\min}} \in (0,1). \quad (61)$$

If $c > 1 + S^{-1}\upsilon_1$, then $\eta < 0$, which yields

$$\lim_{N\to\infty} \sum_{\mathcal{J}\in\mathcal{S}\setminus\mathcal{I}} \mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\} \le \lim_{K\to\infty} \exp(K\eta) = 0$$

implying that the probability that all incorrect support sets are δ jointly typical with all the measurement vectors vanishes. Thus the failure probability p_{err} defined in (8) converges to zero if M satisfies (9).

Next, the remain is to derive (11) in the sublinear sparsity regime. Similarly, let $M = K + cK\log\frac{N}{K}$ wher $c > 1$. From (32),

$$\log \mathbb{P}\left\{\mathcal{E}^c_{\mathcal{I}}\right\} \le 2^{-1} ScK \log\frac{N}{K} \underbrace{\left(\log(1+d_1) - d_1\right)}_{=:A} + \log 2$$

where $A < 0$ due to (33). Thus,

$$\lim_{N\to\infty} \mathbb{P}\left\{\mathcal{E}^c_{\mathcal{I}}\right\} \le \lim_{N\to\infty} \exp\left(2^{-1} ScKA \log\frac{N}{K} + \log 2\right) = 0$$

implying that the probability that the correct support set is not δ jointly typical with all the measurement vectors vanishes.

Then, from (40),

$$\begin{aligned} \log \sum_{\mathcal{J}\in\mathcal{S}\setminus\mathcal{I}} \mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\} &\le \log\left(\binom{N}{K} p\left(d_{2,\alpha^*} - 1\right)\right) \\ &= \log\binom{N}{K} + 2^{-1} ScK \underbrace{\left(\log(1-t) + t\right)}_{=:\gamma} \log\frac{N}{K} \\ &\le K \underbrace{\left(1 + 2^{-1} Sc\gamma\right)}_{=:\eta} \log\frac{N}{K} + K \end{aligned}$$

where the last inequality is due to the bound in (60) and $\gamma < 0$ for any t in (61). If $c > S^{-1}\upsilon_2$, then $\eta < 0$, which yields

$$\lim_{N\to\infty} \sum_{\mathcal{J}\in\mathcal{S}\setminus\mathcal{I}} \mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\} \le \lim_{N\to\infty} \exp\left(K\eta \log\frac{N}{K} + K\right) = 0$$

implying that the probability that all incorrect support sets are δ jointly typical with all the measurement vectors vanishes. Thus, the failure probability p_{err} defined in (8) converges to zero if M satisfies (11), which completes the proof. ■

B. Proof of Theorem 2

From Lemma 1,

$$\mathbb{P}\left\{\mathcal{E}^c_{\mathcal{I}}\right\} \le 2 \left(\underbrace{\exp\left(-\frac{M-K}{2} d_1\right)(1+d_1)^{\frac{M-K}{2}}}_{=:\mu_{\mathcal{I}}} \right)^S. \quad (62)$$

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:47:30 UTC from IEEE Xplore. Restrictions apply.

If $M \geq K+1$, we have

$$\log \mu_{\mathcal{I}} = 2^{-1}(M-K)(\log(1+d_1)-d_1) < 0 \quad (63)$$

due to (33), which implies $\mu_{\mathcal{I}} < 1$. From Lemma 2,

$$\mathbb{P}\{\mathcal{E}_{\mathcal{J}}\} \leq \left(\underbrace{\exp\left(-\frac{M-K}{2}\left(d_{2,\alpha^*}-1\right)\right) d_{2,\alpha^*}^{\frac{M-K}{2}}}_{=:\mu_{\mathcal{J}}} \right)^S. \quad (64)$$

Similarly, if $M \geq K+1$, we have

$$\log \mu_{\mathcal{J}} = 2^{-1}(M-K)(\log(1-t)+t) < 0 \quad (65)$$

due to (61), which implies $\mu_{\mathcal{J}} < 1$. Thus, we conclude

$$\lim_{S\to\infty} p_{err} \leq 2 \lim_{S\to\infty} \mu_{\mathcal{I}}^S + \binom{N}{K} \lim_{S\to\infty} \mu_{\mathcal{J}}^S = 0$$

for $M \geq K+1$ which completes the proof. ■

C. Proof of Theorem 3

The mutual information in (24) is bounded by

$$\begin{aligned}\mathbb{I}\left(\mathbf{x};\, \mathbf{y}|\,\tilde{\mathbf{F}}\right) &= h\left(\mathbf{y}|\,\tilde{\mathbf{F}}\right) - h\left(\mathbf{y}|\,\mathbf{x},\tilde{\mathbf{F}}\right) \leq h(\mathbf{y}) - h(\mathbf{n}) \\ &\leq \sum\nolimits_{i=1}^{SM} h(y_i) - h(\mathbf{n}) \\ &\leq 2^{-1}SM\left(\log\left(2\pi e\left(K x_{\min}^2+\sigma^2\right)\right) - \log\left(2\pi e\sigma^2\right)\right) \\ &= 2^{-1}SM\log(1+K\times \mathrm{SNR}_{\min})\end{aligned}$$

where $h(\mathbf{x})$ is the differential entropy of $\mathbf{x}$, and $h(\mathbf{x}|\,\mathbf{y})$ is the conditional entropy of $\mathbf{x}$ given $\mathbf{y}$. The last inequality is due to that the Gaussian distribution maximizes the differential entropy. The denominator in (24) is bounded by

$$\log\left(\left|\mathcal{X}_{\{x_{\min}\}}\right|-1\right) = \log\left(\binom{N}{K}-1\right) > K\log\frac{N}{K}$$

for sufficiently large N. Then,

$$\begin{aligned}p_{err} &= \mathbb{E}_{\tilde{\mathbf{F}}} \sup_{\mathbf{x}\in\mathcal{X}_{x_{\min}}} \mathbb{P}\left\{\hat{\mathcal{I}} \neq \mathcal{I} \middle|\, \mathbf{x}, \tilde{\mathbf{F}}\right\} \\ &\geq \mathbb{E}_{\tilde{\mathbf{F}}} \min_{\hat{\mathbf{x}}\in\mathcal{X}_{\{x_{\min}\}}} \max_{\mathbf{x}\in\mathcal{X}_{\{x_{\min}\}}} \mathbb{P}\left\{\hat{\mathbf{x}} \neq \mathbf{x} \middle|\, \mathbf{x},\ \tilde{\mathbf{F}}\right\} \\ &> 1 - \frac{2^{-1}SM\log(1+K\times\mathrm{SNR}_{\min})+\log 2}{K\log\frac{N}{K}}. \quad (66)\end{aligned}$$

From (66), the failure probability is bounded away from zero by zero if (25) is satisfied, which completes the proof. ■

Appendix D
Proofs of Corollaries 1, 2 and 3

A. Proof of Corollary 1

From the inequality $\log(1+x) \leq \frac{2x}{2+x}$ for $x \in (-1, 0]$,

$$\upsilon_2 = -\frac{2}{\log(1-t)+t} < \frac{4-2t}{t^2} < \frac{4}{t^2} \quad (67)$$

where t is defined in (61). Then,

$$\frac{\upsilon_2}{S} < \frac{4}{St^2}. \quad (68)$$

From (61),

$$\sqrt{S}t = \frac{1-\rho^{-1}}{\sqrt[-1]{S} + \left(\sqrt{S}\times \mathrm{SNR}_{\min}\right)^{-1}}. \quad (69)$$

Combining (11), (68) and (69) leads to (13). This approach is used to get (14) using the following equality

$$\upsilon_1 = \upsilon_2(1-\log\beta) \quad (70)$$

where $\lim_{N\to\infty} \frac{K}{N} = \beta \in (0, 1/2)$, which completes the proof. ■

B. Proof of Corollary 2

Substituting $\alpha = \frac{2}{3}$ in (15), and rearranging the result with respect to t can yield $\frac{2}{3} \leq t < 1$, where t is defined in (61). Then from (67), a simple computation yields that

$$\upsilon_2 < \frac{4-2t}{t^2} \leq \frac{4}{t}$$

which immediately yields that

$$\frac{\upsilon_2}{S} < \frac{4}{St}. \quad (71)$$

where

$$St = \frac{1-\rho^{-1}}{S^{-1}+(S\times\mathrm{SNR}_{\min})^{-1}}. \quad (72)$$

Combining (11), (71) and (72) leads to (16). This approach is used to get (17) using (70), which completes the proof. ■

C. Proof of Corollary 3

We assume that $\mu_{\mathcal{I}} \geq \mu_{\mathcal{J}}$ and

$$p_{err} \leq \mathbb{P}\{\mathcal{E}_{\mathcal{I}}^c\} + \sum_{\mathcal{J}\in\mathcal{S}\setminus\mathcal{I}} \mathbb{P}\{\mathcal{E}_{\mathcal{J}}\} \leq \left(\binom{N}{K}+2\right)\mu_{\mathcal{I}}^S < \varepsilon < 1. \quad (73)$$

Then, if the number of measurement vectors satisfies

$$S > \frac{\log\varepsilon - \log\left(\binom{N}{K}+2\right)}{\log\mu_{\mathcal{I}}} > 0, \quad (74)$$

(73) is achieved for small ε, and hence, reliable support set reconstruction is possible. If $\mu_{\mathcal{I}} < \mu_{\mathcal{J}}$, we obtain inequalities similar to (73) and (74) by replacing $\mu_{\mathcal{I}}$ by $\mu_{\mathcal{J}}$, where

$$S > \frac{\log\varepsilon - \log\left(\binom{N}{K}+2\right)}{\log\mu_{\mathcal{J}}} > 0. \quad (75)$$

Combining (74) and (75) yields (18).

Next, a simple computation yields that for any d_1 in (33),

$$\frac{\partial \log\mu_{\mathcal{I}}}{\partial d_1} = -\frac{d_1}{2(1+d_1)} < 0$$

where $\log\mu_{\mathcal{I}}$ is given in (63). From (33), we see $d_1 \propto \mathrm{SNR}_{\min}$ that leads to $\log\mu_{\mathcal{I}} \propto \mathrm{SNR}_{\min}^{-1}$. Also, for any t in (61),

$$\frac{\partial \log\mu_{\mathcal{J}}}{\partial t} = -\frac{t}{2(1-t)} < 0$$

where $\log\mu_{\mathcal{J}}$ is given in (65). From (61), we see $t \propto \mathrm{SNR}_{\min}$ that leads to $\log\mu_{\mathcal{J}} \propto \mathrm{SNR}_{\min}^{-1}$. Hence, the sufficient condition on S in (18) turns out to be a decreasing function with respect to $\mathrm{SNR}_{\min}$, which completes the proof. ■

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:47:30 UTC from IEEE Xplore. Restrictions apply.

APPENDIX E
PROOFS OF PROPOSITIONS 1 AND 2

First of all, we introduce the exponential inequalities [32], and use them in the proofs of Propositions 1 and 2.

A. The Exponential Inequalities [32]

Let Y_i, $i = 1, 2, \cdots, D$ be i.i.d. Gaussian variables with a zero mean and a unit variance. Then, let α_i, $i = 1, 2 \cdots, D$ be non-negative. We set

$$|\alpha|_\infty = \sup |\alpha_i|, \ |\alpha|_2^2 = \textstyle\sum_{i=1}^D \alpha_i^2$$

and let

$$Y = \sum\nolimits_{i=1}^D \alpha_i \left(Y_i^2 - 1\right). \tag{76}$$

Then, the following inequalities hold for any positive x

$$\mathbb{P}\left\{Y \geq 2|\alpha|_2 \sqrt{x} + 2|\alpha|_\infty x\right\} \leq \exp(-x) \tag{77}$$

$$\mathbb{P}\left\{Y \leq -2|\alpha|_2 \sqrt{x}\right\} \leq \exp(-x). \tag{78}$$

B. Proof of Proposition 1

In the proof of Lemma 3, $\mathrm{Z}_{\mathcal{I}}$ is represented by

$$\mathrm{Z}_{\mathcal{I}} = \sum\nolimits_{s=1}^S \sum\nolimits_{i=1}^{M-K} w^s(i)^2$$

where $w^s(i)$ is Gaussian with a zero mean and a unit variance. Define a random variable Y as

$$\begin{aligned} Y &= \mathrm{Z}_{\mathcal{I}} - S(M-K) \\ &= \sum\nolimits_{s=1}^S \sum\nolimits_{i=1}^{M-K} \left(w^s(i)^2 - 1\right) \end{aligned}$$

which is of the form of (76). Then,

$$\mathbb{P}\left\{\mathcal{E}_{\mathcal{I}}^c\right\} = \underbrace{\mathbb{P}\left\{Y \leq -SM\delta/\sigma^2\right\}}_{=:A} + \underbrace{\mathbb{P}\left\{Y \geq SM\delta/\sigma^2\right\}}_{=:B}.$$

Combining A with (78) gives

$$\begin{aligned} \mathbb{P}\left\{Y \leq -SM\delta/\sigma^2\right\} &= \mathbb{P}\left\{Y \leq -2\sqrt{S(M-K)x}\right\} \\ &\leq \underbrace{\exp\left(-\frac{SM^2\delta^2}{4(M-K)\sigma^4}\right)}_{=:C} \end{aligned}$$

and combining B with (77) gives

$$\begin{aligned} \mathbb{P}\left\{Y \geq SM\delta/\sigma^2\right\} &= \mathbb{P}\left\{Y \geq 2\sqrt{S(M-K)x} + 2x\right\} \\ &\leq p_{1,\exp} \end{aligned}$$

where $p_{1,\exp}$ is defined in (28). It is readily seen that $p_{1,\exp} \geq C$, which leads to $\mathbb{P}\left\{\mathcal{E}_{\mathcal{I}}^c\right\} \leq 2p_{1,\exp}$.

Next, from (32) and (28),

$$\log p(d_1) = 2^{-1} S(M-K)\left(\log(1+d_1) - d_1\right)$$

and

$$\log p_{1,\exp} = -2^{-1} S(M-K) d_1^2 (2+4d_1)^{-1}$$

where $d_1 > 0$ is defined in (33). Then, we have

$$\log \frac{p(d_1)}{p_{1,\exp}} = \frac{S(M-K)}{2} \underbrace{\left(\log(1+d_1) - d_1 + d_1^2(2+4d_1)^{-1}\right)}_{=:g(d_1)}.$$

For any $d_1 > 0$, $\frac{\partial g(d_1)}{\partial d_1} = -d_1^2(2+3d_1)(1+d_1)^{-1}(1+2d_1)^{-2} < 0$ and $\max_{d_1>0} g(d_1) = 0$. Thus, we conclude $\log \frac{p(d_1)}{p_{1,\exp}} \leq 0$, which completes the proof. ■

C. Proof of Proposition 2

In the proof of Lemma 4, $\mathrm{Z}_{\mathcal{J}}$ is represented by

$$\begin{aligned} \mathrm{Z}_{\mathcal{J}} &= \sum\nolimits_{s=1}^S \sum\nolimits_{i=1}^{M-K} b^s(i)^2 \\ &= \sum\nolimits_{s=1}^S \sum\nolimits_{i=1}^{M-K} \alpha_{\mathcal{J},s} g^s(i)^2 \end{aligned}$$

where $\alpha_{\mathcal{J},s}$ is defined in (39) and $g^s(i)$ is Gaussian with a zero mean and a unit variance. Define a new random variable Y as

$$\begin{aligned} Y &= \mathrm{Z}_{\mathcal{J}} - S(M-K) \\ &= \sum\nolimits_{s=1}^S \sum\nolimits_{i=1}^{M-K} \alpha_{\mathcal{J},s} \left(g^s(i)^2 - 1\right) \end{aligned}$$

which is of the form of (76). Then, from (44)

$$\begin{aligned} \mathbb{P}\left\{\mathcal{E}_{\mathcal{J}}\right\} &\leq \mathbb{P}\left\{Y < SM\delta - (M-K)\sum\nolimits_{s=1}^S \left\|\mathbf{x}_{\mathcal{I}\setminus\mathcal{J}}^s\right\|_2^2\right\} \\ &\leq \mathbb{P}\left\{Y < \underbrace{SM\delta - S(M-K)x_{\min,\mathcal{J}}^2}_{=:A}\right\} \\ &\leq p_{2,\mathcal{J},\exp} \end{aligned} \tag{79}$$

where $p_{2,\mathcal{J},\exp}$ is defined in (30), the last inequality is due to (78). Due to (29), A is negative. Thus the exponential inequality of (78) gives the upper bound $p_{2,\mathcal{J},\exp}$.

Next, from (40) and (30),

$$\log p\left(d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})} - 1\right) = 2^{-1} S(M-K)\left(t + \log(1-t)\right)$$

and

$$\begin{aligned} \log p_{2,\mathcal{J},\exp} &\geq -\frac{S(M-K)}{4}\left(\frac{x_{\min,\mathcal{J}}^2 - \frac{M\delta}{M-K}}{x_{\min,\mathcal{J}}^2 + \sigma^2}\right)^2 \\ &= -4^{-1} S(M-K) t^2 \end{aligned}$$

where $t \in (0,1)$, is defined in (61) and the inequality is due to (50). Then,

$$\log \frac{p\left(d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})} - 1\right)}{p_{2,\mathcal{J},\exp}} \leq \frac{S(M-K)}{4} \underbrace{\left(t^2 + 2t + 2\log(1-t)\right)}_{=:g(t)}.$$

For any $t \in (0,1)$, $\frac{\partial g(t)}{\partial t} = -2t^2(1-t)^{-1} < 0$ and $\max g(t) = 0$. We conclude $\log \frac{p\left(d_{2,\lambda_{\min}(\mathbf{R}_{\mathcal{J}})} - 1\right)}{p_{2,\mathcal{J},\exp}} \leq 0$. It completes the proof. ■

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[3] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:47:30 UTC from IEEE Xplore. Restrictions apply.

[4] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
[5] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
[6] E. J. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.
[7] E. J. Candès and M. B. Wakin, "An introduction to compressive sensing," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
[8] D. Baron, M. B. Wakin, M. F. Duarte, S. Sarvotham, and R. G. Baraniuk. (2009). "Distributed compressed sensing." [Online]. Available: https://arxiv.org/pdf/0901.3403.pdf
[9] M. F. Duarte, S. Sarvotham, D. Baron, M. B. Wakin, and R. G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2005, pp. 1537–1541.
[10] C. Caione, D. Brunelli, and L. Benini, "Compressive sensing optimization for signal ensembles in WSNs," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 382–392, Feb. 2014.
[11] W. Chen, R. D. Rodrigues, and I. J. Wassell, "Distributed compressive sensing reconstruction via common support discovery," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kyoto, Japan, Jun. 2011, pp. 1–5.
[12] J. Oliver, W.-B. Lee, and H.-N. Lee, "Filters with random transmittance for improving resolution in filter-array-based spectrometers," *Opt. Exp.*, vol. 21, no. 4, pp. 3969–3989, Feb. 2013.
[13] S. Goginneni and A. Nehorai, "Target estimation using sparse modeling for distributed MIMO radar," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5315–5325, Nov. 2011.
[14] Y. Wu *et al.*, "Accelerated MR diffusion tensor imaging using distributed compressed sensing," *Magn. Reson. Med.*, vol. 71, no. 2, pp. 763–772, Feb. 2014.
[15] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York, NY, USA: Springer, 2010.
[16] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, May 2012.
[17] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge, U.K.: Cambridge Univ. Press, Apr. 2015.
[18] M. J. Wainwright, "Information-theoretic limits on sparsity recovery in the high-dimensional and noisy setting," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5728–5741, Dec. 2009.
[19] A. K. Fletcher, S. Rangan, and V. K. Goyal, "Necessary and sufficient conditions for sparsity pattern recovery," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5758–5772, Dec. 2009.
[20] S. Aeron, V. Saligrama, and M. Zhao, "Information theoretic bounds for compressed sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 5111–5130, Oct. 2010.
[21] M. Akçakaya and V. Tarokh, "Shannon-theoretic limits on noisy compressive sampling," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 492–504, Jan. 2010.
[22] J. Scarlett, J. S. Evans, and S. Dey, "Compressed sensing with prior information: Information-theoretic limits and practical decoders," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 427–439, Jan. 2013.
[23] J. Scarlett and V. Cevher, "Limits on support recovery with probabilistic models: An information-theoretic framework," *IEEE Trans. Inf. Theory*, vol. 63, no. 1, pp. 593–620, Jan. 2017.
[24] T. Cover and J. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 2006.
[25] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Trans. Signal Process.*, vol. 56, no. 12, pp. 4634–4643, Dec. 2006.
[26] G. Tang and A. Nehorai, "Performance analysis for sparse support recovery," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1383–1399, Mar. 2010.
[27] Y. Jin and B. D. Rao, "Support recovery of sparse signals in the presence of multiple measurement Vectors," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 3139–3157, May 2013.
[28] M. F. Duarte, M. B. Wakin, D. Braon, S. Sarvotham, and R. G. Baraniuk, "Measurement bounds for sparse signal ensembles via graphical models," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4280–4289, Jul. 2013.
[29] J. Kim, O. Lee, and J. Ye, "Compressive MUSIC: Revisiting the link between compressive sensing and array signal processing," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 278–301, Jan. 2012.
[30] J. D. Blanchard and M. E. Davies, "Recovery guarantees for rank aware pursuits," *IEEE Signal Process. Lett.*, vol. 19, no. 7, pp. 427–430, Jul. 2012.
[31] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*. New York, NY, USA: Wiley, 2001.
[32] B. Laurent and P. Massart, "Adaptive estimation of a quadratic functional by model selection," *Ann. Statist.*, vol. 28, no. 5, pp. 1303–1338, Oct. 2000.
[33] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA, USA: Addison-Wesley, 1991.
[34] R. Rajagopalan and P. K. Varsheny, "Data-aggregation techniques in sensor networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 8, no. 4, pp. 48–63, 4th Quart., 2006.

Sangjun Park received the B.S. degree in computer engineering from the Chungnam National University, Daejeon, Korea, in 2009. He is pursuing a Ph.D degree at the School of Electrical Engineering and Computer Science in the Gwangju Institute of Science and Technology (GIST), Gwangju, Korea. His research interests include information theory, numerical optimization and compressed sensing.

Nam Yul Yu (M'07) received the B.S. degree in electronics engineering from the Seoul National University, Seoul, Korea, in 1995 and the M.S. degree in electronic and electrical engineering from the Pohang University of Science and Technology (POSTECH), Pohang, Korea, in 2000. He received the Ph. D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, Ontario, Canada, in 2007. From 2000 to 2003, he was with Telecommunication Research and Development Center, Samsung Electronics, Korea, where he worked on channel coding schemes for wireless communication systems. In 2007, he was a senior research engineer in LG Electronics, Korea, working on the standardization of the 3GPP-LTE. From 2008 to 2014, he has been working as an Assistant/Associate Professor in the Department of Electrical Engineering at the Lakehead University, Thunder Bay, Ontario, Canada. In 2014, he joined the Gwangju Institute of Science and Technology (GIST), Gwangju, Korea, and is currently working as an Associate Professor in the School of Electrical Engineering and Computer Science. From 2009 to 2011, he served as an Associate Editor for Sequences in the IEEE Transactions on Information Theory. His research interests include sequence design, compressed sensing, and security for wireless communications.

Heung-No Lee (SM'13) received the B.S., M.S., and Ph.D. degrees from the University of California, Los Angeles, CA, USA, in 1993, 1994, and 1999, respectively, all in electrical engineering. He then worked at HRL Laboratories, LLC, Malibu, CA, USA, as a Research Staff Member from 1999 to 2002. From 2002 to 2008, he worked as an Assistant Professor at the University of Pittsburgh, PA, USA. In 2009, he then moved to the School of Electrical Engineering and Computer Science, GIST, Korea, where he is currently affiliated. His areas of research include information theory, signal processing theory, communications/networking theory, and their application to wireless communications and networking, compressive sensing, future internet, and brain-computer interface. He has received several prestigious national awards, including the Top 100 National Research and Development Award in 2012, the Top 50 Achievements of Fundamental Researches Award in 2013, and the Science/Engineer of the Month (January 2014). He has written more than fifty international journal publications and a hundred international conferences and workshop papers. He was the Director of Electrical Engineering and Computer Science track within GIST College in 2014. In March 2015, he was appointed as the Dean of Research at GIST.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:47:30 UTC from IEEE Xplore. Restrictions apply.

Time-Variant Proof-of-Work Using Error-Correction Codes

Sangjun Park, Haeung Choi, and *Heung-No Lee, *Senior Member, IEEE*

***Abstract*— The protocol for cryptocurrencies can be divided into three parts, namely consensus, wallet, and networking overlay. The aim of the consensus part is to bring trustless rational peer-to-peer nodes to an agreement to the current status of the blockchain. The status must be updated through valid transactions. A proof-of-work (PoW) based consensus mechanism has been proven to be secure and robust owing to its simple rule and has served as a firm foundation for cryptocurrencies such as Bitcoin and Ethereum. Specialized mining devices have emerged, as rational miners aim to maximize profit, and caused two problems: *i*) the re-centralization of a mining market and *ii*) the huge energy spending in mining. In this paper, we aim to propose a new PoW called Error-Correction Codes PoW (ECCPoW) where the error-correction codes and their decoder can be utilized for PoW. In ECCPoW, puzzles can be intentionally generated to vary from block to block, leading to a time-variant puzzle generation mechanism. This mechanism is useful in repressing the emergence of the specialized mining devices. It can serve as a solution to the two problems of recentralization and energy spending.**

***Index Terms*— Consensus, Cryptocurrency, Blockchain, Proof-of-Work, Error-Correction Codes, Hash Functions**

I. Introduction

In cryptocurrencies, the consensus part plays a role in leading an agreement among trustless nodes without any communications. This part is the most innovative because it can prevent the double spending attack [1] in a peer-to-peer network in the absence of trusted parties. In *Bitcoin* [2], as an example, more than ten thousand of nodes randomly scattered across the world aim to reach a consensus in each block time. The Internet is the only way to connect them; communication packets are delayed and sometimes dropped though the Internet that is designed to provide the best effort service. Cyberattacks frequently happen, making transactions over the Internet insecure. Nevertheless, Bitcoin has shown secure peer-to-peer transactions over the past 10 years. With the help of proof-of-work (*PoW*) which is fundamental to the consensus part, this can be possible.

In Bitcoin, each node does competitive work, called *mining*, to forge a block. The node which wins this competition has the right to mint a specified number of coins as this mining reward. If a node was re-forging all the blocks alone, it could spend the total amount of works done to all the mined blocks.

Without PoW, anybody with a computer can alter the content of the blockchain, implying unauthorized changes in any mined blocks can be possible. If PoW is attached to each mined block, attackers cannot make any unauthorized modifications without redoing all the works. No node can alone alter any mined block, meaning an immutability property.

In Bitcoin, miners make rational decisions to maximize their profits by following a two stage process in which *i*) the miners select a blockchain whose length is the longest and *ii*) they extend this longest one by adding a newly mined block. Suppose there are two blockchains where one is longer than the other one in terms of the length. Since the longer chain has the more accumulated works, altering it is more difficult. This longer chain shall be treated the more trustable and preferable by the miners. Thus, they select the longer chain. Making such a selection is rational for the sake of keeping the mining rewards. The mining reward is a delayed conditional payment, i.e., if a miner mines a block at a given time point t_1, the reward is delayed until the future moment t_2 of time. This time from t_1 to t_2 is measured in terms of the number of blocks, say 100 blocks. If this mined block was not a part of the longest chain at the future time point t_2, the reward vanishes. Thus, rational miners select the longest chain.

In Bitcoin, miners spend computational resources to forge a block by solving a puzzle carved in a bitcoin program as an on-chain policy. This puzzle is made using the secure hash algorithm 256 (SHA256) [3]. To solve the puzzle, the miners have to repeatedly execute SHA256 by varying an input to SHA256 until a good hash is given. This input is the header of the block, i.e., block header, including six fields such as version, previous hash, difficulty, timestamp and nonce, Merkle tree value. The version is fixed. Given a block at a certain height, i.e., the lth block, the previous hash and the difficulty are obtained from its previous block, i.e., $(l–1)^{\text{th}}$ block. They are constant. The rest varies until SHA256 returns a good hash. A good hash can be spotted since it shall possess a certain number of leading zero bits reflecting the difficulty. The block header of a mined block can serve as the proof that a given puzzle is solved without any falsehood.

Satoshi [2] intended for miners to execute SHA256 using a central processing unit (CPU). But, faster computing machines based on application-specific integrated circuit (ASIC) became available. As a result, the miners have chosen to exploit them to maximize their profits. To date, the miners equipped with ASIC

S. Park is with the Electronics and Telecommunications Research Institute (ETRI), South Korea (e-mail: sjpark86@etri.re.kr)

H. Choi and H.-N. Lee are with the School of Electrical Engineering and Computer Science (EECS), Gwangju Institute of Science and Technology (GIST), South Korea (e-mail: haeung@gist.ac.kr, heungno@gist.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean government (MSIP) (NRF-2018R1A2A1A19018665). The corresponding authors: Heung-No Lee (heungno@gist.ac.kr)

mining devices have dominated the mining business [4], leading to two problems:

M1. The mining markets have become re-centralized [5].
M2. The electrical energy spent to mine blocks is huge [6].

First, the miners have a large portion of the total hash power, implying that the plight of the blockchain is left to a handful of these influential miners. It can be possible to modify any mined blocks on their own rights, leading to shattered trust. Namely, they can break the immutability property. Second, new models of ASIC mining devices can surpass old models with respect to the hash power, which is measured as a hash rate. Each miner is forced to buy newer models to win the mining competition. As the new models are widely used, the total hash power inevitably grows. The difficulty level to solve the puzzle must increase to keep a certain predefined range of block generation time that is expectedly consumed to mine a single block. Today, this level has gotten huge, i.e., $O(10^{20})$ hash rate per second. As a result, using CPUs in mining has gotten no longer profitable. Besides, as the total hash power increases, the miners spend more and more electrical energy to mine blocks.

If we prevent the usage of ASIC mining, we can alleviate the problems M1 and M2. To this end, we use the error-correction codes and their decoder. In general, the aim of using the codes in modern communication systems is to combat errors occurring over noisy channels in which the errors introduced over a noisy channel can be corrected by running a decoding algorithm. The codes have a rich history where there are numerous classes of good codes available. They have been used to define both a good public-key crypto system and a good hash function.

The first result can be traced back to the late 1970. McEliece [8] used Goppa codes to make a McEliece cryptosystem where a message is encoded using a public key $\mathbf{A} := \mathbf{SGP}$ where $\mathbf{G}$ is the generator matrix of a binary Goppa code, $\mathbf{S}$ is a nonsingular random matrix and $\mathbf{P}$ is a permutation matrix. The hash, i.e., the encoded result, of a provided message is made as follows: *a*) a word is made by multiplying the message with the public key and *b*) adding a binary random word whose number of ones is at most t to this word[1] is to get the hash. Peters *et al.* [9] extended this system using non-binary Goppa codes to reduce the size of its public key. Even the size reduces, this system can achieve still the same security level as much as that of [8] could. Other codes such as low density generator matrix codes [10], low density parity check (LDPC) codes [11], [12], Reed-Solomon codes [13] and Reed-Muller codes [14] have been used to replace the Goppa codes. The aim of using them is to reduce the size of the public key.

Aside from the applications of the error-correction codes into the McEliece cryptosystem, the codes are used to construct new hash functions. Preneel [15] proposed a method to make new hash functions and proved that their hash functions can provide strong collision resistant properties. The codes in [15] are either the maximum distance separable codes or the Hamming codes. Selman *et al.* [16] used LDPC codes to make a hash function and proved this function as an average universal hash function defined in [17]. These results can motivate us to exploit error-correction codes in designing a new PoW framework.

The contributions of this paper are three folds. First, we propose a new PoW framework which we name as ECCPoW. As the name implies, we add the error-correction code part into PoW. To the best of our knowledge, this is the first work in which the error-correction codes are applied to blockchains. We then explain how we make puzzles, which we call ECC puzzles, and give routines to solve them.

Second, we conduct a probabilistic study to examine a random variable called First Success Hash Cycle (FSHC) representing the number of hash cycles required in solving a given ECC puzzle. Based on this study, we get the following results:

- FSHC follows a geometric distribution with a parameter in (21) that depends on the number of miners M and the code length n.
- The expected value of FSHC is a decreasing function with respect to the number of miners M.
- The expected value of FSHC is an increasing function with respect to the code length n.

Third, we define five properties for a good PoW and explain how ECCPoW satisfies these properties. We shall note that the most innovative property is the time-variant property, making ECCPoW suitable to resolve the problems of recentralization and energy spending.

We organize the rest of this paper as follows. Section II gives literature surveys regarding SHAs and PoWs. Section III elucidates LDPC codes and a decoder. Section IV addresses how ECCPoW works and gives its pseudo codes Section V presents theoretical results of ECCPoW. Section VI discusses properties of ECCPoW. Section VII presents the conclusions of this paper.

II. Literature On Blockchain Cryptographies

A. Secure hash standard and functions

The secure hash standard was formulated by NIST [3]. The purpose of this standard is to offer the specifications of SHAs that yield a hash of a given message. Even the message changes slightly, the hash of the changed message comes out completely different from that of the original message. Thus, a hash can be used to detect whether an original message was altered or not. SHAs with such a property can be used for the generation and the verification of digital signatures as well as for the message authentication.

A secure hash function takes an arbitrarily sized message and produces a fixed-size hash. Let a function h be

$$h : \mathcal{X} \to \mathcal{Y}$$

which is said to be a cryptographically secure hash function if it satisfies the three requirements defined in [17] below:

(***One-way function***) Given any hash y to which a corresponding message is not known, it is computationally infeasible to find a message x such that $h(x) = y$.

[1] For any Goppa code, there is a construction method to guarantee that the minimum distance d of that code is greater than a given positive integer. Thus, the value of t can be known in advance using Theorem 1 [32].

TABLE I. The routine of bitcoin

Inputs: $\mathcal{S}_0$ (block header except for nonce) and L	
Step 1:	for nonce = 0, 1, 2, … $2^{32} - 1$
Step 2:	$\mathbf{e} = \mathrm{SHA256}(\mathcal{S}_0 \cup \{\mathrm{nonce}\})$
Step 3:	If **e** begins with L zero bits, then go to Step 5.
Step 4:	end
Step 5:	Block generation & broadcast

(***Weak collision resistance***) Given an arbitrary message x, it is computationally infeasible to find any message x' which results in the same hash, i.e., $h(x) = h(x')$.
(***Strong collision resistance***) It is computationally infeasible to find any two different messages x and x' which make the same hashes $h(x) = h(x')$.

NIST [3] has proposed a family of SHAs including SHA1, SHA224, SHA256, SHA384, and SHA512. A message of any size less than 2^{64} bits can be given as an input for SHA1, SHA224, and SHA256, while that of less than 2^{128} bits for SHA384 and SHA512. The size of a hash ranges from 160 to 512 bits, depending on the algorithm.

B. PoW of Bitcoin

In Table I, we define a puzzle in Bitcoin and give routines to solve this puzzle. In Step 2, a miner puts a given block header with a selected nonce to SHA256 and obtain a hash of 256-bits. In Step 3, this puzzle is declared to be solved if a hash is smaller than a specified target value, or in a simpler argument it begins with L zero bits, where the target value or the value of L is given as the difficulty level of the PoW puzzle. The miner repeats the routines from Step 2 to Step 3 by varying the nonce. However, there exists a chance in which the miner can fail to hit a good hash even though the whole range of nonces, i.e., $0 \sim 2^{32} - 1$, is used. In such a case, the miner updates the block header, and repeats the routines from Step 2 and Step 3. There are two methods to update the block header. The first one is to update the timestamp field. The second one is to update the Merkle tree value by modifying the list of transactions being included in a block which the miner aims to construct.

C. PoW of Ethereum

Ethash [18] was created for the purpose of preventing the advent of ASIC mining devices in Ethereum. In Ethash, there is a memory structure called directed acyclic graph (DAG) where its data are randomly re-generated every 30,000 block.

Table II shows routines of Ethash. As we have shown in Step 2, the current block header with a nonce is taken by SHA3 to get a hash. This hash is taken by a predefined function to yield mix0 that is random. In Step 4, mix0 is used to determine which data from DAG are fetched. No one predicts which data shall be fetched from DAG because mix0 is random. The mixer takes both the fetched data and mix0 to get a random value in Step 5. In Step 6, mix0 is updated using the random value. The routines from Step 2 to Step 6 are repeated 63 times. Last, the decision is made using this final mix0, as we have shown in Step 8.

The ASIC resistant property in Ethash is originated from the fact that the operation time for the mixer is shorter than that of

TABLE II. The routine of Ethash

Inputs: $\mathcal{S}_0$ (block header except for nonce), L and DAG	
Step 1:	for nonce = 0, 1, 2, … $2^{32} - 1$
Step 2:	$\mathrm{mix0} = f\left(\mathrm{SHA3}(\mathcal{S}_0 \cup \{\mathrm{nonce}\})\right)$
Step 3:	for i = 1, 2, …, 63
Step 4:	$\mathrm{data1} = \mathrm{Fetch}(\mathrm{DAG}, \mathrm{mix0})$
Step 5:	$\mathrm{tmp} = \mathrm{Mixing}(\mathrm{mix0}, \mathrm{data1})$
Step 6:	$\mathrm{mix0} = f(\mathrm{tmp})$
Step 7:	end
Step 8:	If mix0 begins with L zero bits, then go to Step 10.
Step 9:	end
Step 10:	Block generation & broadcast

where f is a predefined function. Details on f is given in [18].

the fetch operation. To be specific, let A_i be the time duration (TD) to conduct the mixing operation in which the subscript i denotes a chip to run this mixing operation. It is clear that the mixing operation TD based on an ASIC chip is significantly less than that based on any CPU chip because the clock speed of an ASIC chip is much faster than that of any CPU chip, i.e.,

$$\mathrm{A}_{\mathrm{ASIC}} \ll \mathrm{A}_{\mathrm{CPU}}.$$

Next, let B_i be the TD to conduct the fetch operation. Unlike to the mixing operation TD, this TD depends on the communications bandwidth between the memory and the cache of the CPU in which the fetched data from DAG are passed through. In other words, the fetch operation time is connected mainly to the bandwidth but not to the clock speed. The purpose to use ASIC chips is to increase the processing speed; is not for obtaining a higher communications bandwidth. The fetch operation TD for CPU and ASIC are thus similar, i.e.,

$$\mathrm{B}_{\mathrm{ASIC}} \approx \mathrm{B}_{\mathrm{CPU}}.$$

We consider the inner routines of Ethash, i.e., Step 4 and Step 5. The mixing operation is conducted after the fetching operation is done. This operation TD at Step 4 can be significantly reduced using ASIC chips; but the fetching operation TD at Step 5 is not reduced even if ASCI chips are used. As a result, there is a bottleneck between Step 4 and Step 5. This bottleneck has the ASIC resistant property enabled.

Recently, a programmatic PoW (ProgPoW), which is planning to be used to replace Ethash, has been proposed to further improve the ASIC resistant property. This improvement is done by changing parameters related to DAG randomly from block to block. Such modifications can make the fetch operation time increased. But, the development of ProgPoW is not completed and ProgPoW is not proven to be secure at the time of writing this manuscript.

D. PoW of Dash

X11 was proposed in 2014 by Duffield [19]. In Table III, we give routines of X11 which consists of hash functions below:

Blake, Bmw, Groestl, Jh, Keccak, Skein, Luffa, Cubehash, Shavite, Simd and Echo.

TABLE III. The routine of X11

Inputs: $\mathcal{S}_0$ (block header except for nonce) and L	
Step 1:	for nonce = 0, 1, 2, … $2^{32} - 1$
Step 2:	$\mathbf{e} = \mathtt{Blake}(\mathcal{S}_0 \cup \{\text{nonce}\})$
Step 3:	$\mathbf{e} = \mathtt{Bmw}(\mathbf{e})$
	….
Step 12:	$\mathbf{e} = \text{Echo}(\mathbf{e})$
Step 13:	If **e** begins with L zero bits, then go to Step 15.
Step 14:	end
Step 15:	Block generation & broadcast

Blake first takes a given set of the current the block header with a selected nonce to get its hash. Next, Bmw takes this hash as its input to yield a hash. The same procedures are repeated until Echo, the last hash function, yields its hash. The decision is made using this last hash, as we have shown in Step 13.

However, the order of using the 11 hash functions is always fixed. This fixed order makes the development of ASIC mining devices possible and in fact an easy task. The development of ASIC mining solution can be done when these hash functions are implemented in a single device. Logic gates to sequentially connect the hash functions can be implemented. The first ASIC mining device targeting X11 was developed in 2016.

The idea behind X11 has been extended to other PoWs such as X13, X15, and X17. As the names suggest, they consists of 13, 15, and 17 hash functions, respectively. To date, there a set of ASIC mining devices for both X13 and X15 while there is no ASIC mining device yet for X17.

E. PoW of Raven

In 2018, a new extension of X11 was proposed in [20]. This is called X16r. It uses multiple hash functions given in Table IV to get the last hash like to other extensions of X11 that we have discussed in Section II D. But, unlike the others, the sequence of the hash functions in X16r can be made to vary from block to block. This variation seems to be a role for preventing the advent of ASIC mining devices for X16r.

We provide an example to address how X16r operates. The sequence X16r is determined upon the last 16 bytes of a previous hash. Let this previous hash be

0x0000...04def2c3eff6da11542ffcdabce.

The last 16 bytes are 6da11542ffcdabce. Then, the sequence is decided on the basis of Table IV below:

Luffa → Shabal → Echo → Bmw → Bmw → Skein → Keccak → Groestl → Sha512 → Sha512 → Fugue → Shabal → Echo → Hamsi → Fugue → Whirlpool.

A miner puts a given set of the block header with a selected nonce through Luffa to yield a hash. Shabal takes this hash as its input to yield its hash. This routine is repeated until the last hash is yielded. In the above example, the last hash is obtained through Whirlpool.

At the time of writing this manuscript, it seems, no one has officially succeeded in implementing ASIC mining devices for X16r, i.e., there is no announced commercial product. However,

TABLE IV. The map for X16r

Value	Hash	Value	Hash
0	Blake	8	Shavite
1	Bmw	9	Simd
2	Groestl	a	Echo
3	Jh	b	Hamsi
4	Keccak	c	Fugue
5	Skein	d	Shabal
6	Luffa	e	Whirlpool
7	Cubehash	f	Sha512

Black and Weight, the developers of X16r, in [20] stated that reordering the sequence cannot make the development of ASIC mining devices impossible. Recently, in [33] at Nov., 2019, Whitefire990 reported a simulation result which indicates the probability of k time-repetition, such that the same hash function is used at least 5 times consecutively when k is 5, exponentially decreases in k. Insisted further is that k greater than 5 can be ignored in designing of ASIC mining devices. As such, what claimed there is that the ASIC-resistant property of Raven can be broken by the said ASIC designing method. It shall be noted, however, that all these claims provided in [33] have not yet been carefully verified through a peer-review system.

F. Short summary from C to E

From the subsection II.C to II.E, we have reviewed the existing ASIC-resistant PoWs categorized as follows:

a. The usage of intentional memory access.
b. The usage of multiple hash functions.

Ethash and ProgPoW can belong to the first class while X11 and its variants such as X13, X15, X17 and X16r can belong to the second class. The basic idea of the first class is to use the bottleneck intentionally caused by randomly fetching data from a memory. The basic idea of the second one is to use the multiple hash functions, which can make the development costs of ASIC mining devices expensive.

At the time of writing this manuscript, ASIC mining devices for Ethash, X11, X13 and X15 are available. The development of ProgPoW is not yet available. X17 and X16r can be resistant to ASIC mining devices. But, as the ASIC-resistant property of the PoWs such as X11, X13 and X15 are broken, that of X17 can be cracked in the near future when the hardware development technology is improved. As mentioned in the subsection II.E, there is a claim that the anti-ASIC property of X16r could be broken; but to date no commercial ASIC mining device has been announced.

III. Literature Surveys on LDPC

An LDPC decoder plays an important role in ECCPoW. That is, the decoder is utilized to generate an unpredictable random output that can be later on used to give a proof whether a puzzle is solved or not. We prepare this section to give a quick summary regarding the LDPC codes and their decoders.

In 1963, LDPC codes were proposed by Gallager in his thesis [7]. But, the codes received no attention back then because computers were not sufficiently fast enough to check the per-

formance of a decoder. Mackay and Neal [22] reported in 1997 that the codes could achieve the Shannon limit [21] closely with a message passing decoder that uses a kind of belief propagation algorithms. Since then, numerous studies on the codes have been made. They are categorized as follows: *i*) constructing the codes to approach the Shannon limit [21] and *ii*) implementing the fast decoders based on either ASIC [23]–[26] or field programmable gate array (FPGA) [27]–[28] to support real-time decoding purpose.

A. LDPC codes

LDPC codes can be generalized to a non-binary alphabet for improving its error-correction capability. But, for the purpose of using these codes in ECCPoW to provide a new time-varying anti-ASIC PoW system, it hence is suffice to consider the binary alphabet version only.

An (n, k) LDPC code is a linear code constructed by supplementing each message $\mathbf{m}$ of size k with parity bits to get a codeword of size n. This code is often defined with respect to a parity check matrix $\mathbf{H}$ of size $m \times n$ such that each element is binary either zero or one and the number of ones is very small, where m is the number of parity bits, i.e., $m = n - k$.

For a given parity check matrix $\mathbf{H}$, its corresponding LDPC code can be either *regular* or *irregular*. If $\mathbf{H}$ contains a constant number w_c, called the column degree, of ones in each column and a constant number w_r, called the row degree, of 1s in each row, the code is called *regular*. For a given *regular* LDPC code, the parameters such as n, k, w_c, and w_r satisfy the following:

$$nw_c = (n-k)w_r = mw_r. \tag{1}$$

If $\mathbf{H}$ contains a different number of 1s in both each column and each row, the code is called *irregular*. In the perspective of the error-correction capability, irregular codes are better than regular codes. To serve our purpose of anti-ASIC PoW mechanism, we aim to consider the *regular* LDPC codes because

i. it is much easier to implement a decoder of regular LDPC codes and
ii. the aim of using this decoder is not to correct errors but to yield an unpredictable random output.

A bipartite graph is often used to represent an LDPC code, as we have shown in Fig. 1. The lower and upper nodes are called the variable nodes and the check nodes, respectively. Each edge shows the adjacency of the ith variable node and the jth check node and corresponds to a nonzero $(i, j)^{\text{th}}$ element in $\mathbf{H}$.

For a given LDPC code, its error-correction capability relies on the minimum (Hamming) distance d. This distance is given by solving an optimization problem in which we consider any pair of $2^k - 1$ different codewords below:

$$d = \min_{\mathbf{u} \in \{\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_{2^k}\} \setminus \mathbf{0}_n} \|\mathbf{u}\|_h \tag{2}$$

which is NP-complete, where $\|\mathbf{x}\|_h$ is the number of 1s in $\mathbf{x}$.

Thus far, studies on the computation of a good approximation to minimum distance for a given fixed $\mathbf{H}$ with reasonable size have been reported futile and as such it has remained as an open problem. Keha and Duman [29] proposed a branch and cut algorithm to obtain the minimum distance of LDPC codes. But, this algorithm requires a large amount of time and memory; it is thus only useful if n is small. Then, Hashemi and Banihashemi [30] proposed a method to find lower and upper bounds of the minimum distance of LDPC codes and obtained both of the bounds even when $n > 64{,}000$.

For regular LDPC codes with a particular pair of w_c and w_r, upper and lower bounds for a relative minimum distance which is the ratio of the minimum distance d to the code length n, are given in [31] and [7], respectively. We use them for our purpose in this paper in Section IV. Once the minimum distance is given, the number of correctable errors can be computed as follows:

Theorem 1 [32]: Let a linear code be defined as a given parity check matrix $\mathbf{H}$ which has the minimum distance d. Then, the number of correctable errors is

$$t = \lfloor (d-1)/2 \rfloor \tag{3}$$

where $\lfloor x \rfloor$ denotes the integer part of x.

We explain how to encode a message $\mathbf{m}$ for a given $\mathbf{H}$ of size $m \times n$. To this end, we build a generator matrix $\mathbf{G}$ of size $n \times k$ whose column space is orthogonal to the row space of $\mathbf{H}$ below:

Step 1: Conduct the Gaussian elimination to rewrite $\mathbf{H}$ as follows:

$$\mathbf{H} = \begin{bmatrix} \mathbf{A}^T & \mathbf{I}_{n-k} \end{bmatrix}$$

where $\mathbf{I}_{n-k}$ is the identity matrix of size $(n-k) \times (n-k)$.
Step 2: Form $\mathbf{G}$ of size $n \times k$ as follows:

$$\mathbf{G} = \begin{bmatrix} \mathbf{I}_k \\ \mathbf{A} \end{bmatrix}.$$

It is noted that $\mathbf{HG} = \mathbf{0}_{n-k,k}$ where $\mathbf{0}_{n-k,k}$ is the zero matrix of size $(n-k) \times k$. Note again $m = n - k$. The message $\mathbf{m}$ is encoded to produce a codeword $\mathbf{c}$ of size $n \times 1$ via $\mathbf{c} = \mathbf{Gm}$. Then, because of the definition of $\mathbf{G}$, it is always seen that the result of multiplying $\mathbf{H}$ with $\mathbf{c}$ is the zero vector of size m, i.e.,

$$\mathbf{Hc} = \mathbf{HGm} = \mathbf{0}_m.$$

A decoder takes both the parity check matrix $\mathbf{H}$ and the corrupted word $\mathbf{r}$, which is $\mathbf{r} = \mathbf{c} + \mathbf{e}$, where $\mathbf{e}$ is an error pattern. The decoder runs a message-passing algorithm [32] shown to be the standard decoding algorithm to remove $\mathbf{e}$.

The principle behind the message-passing algorithm is to iteratively propagate probabilistic information among the variable and check nodes. The iterations are terminated if either the number of iterations exceeds a given number or an output is a codeword. Detailed explanations on how this algorithm operates are provided in [32], i.e., Algorithm 5.1 on page 220. The algorithm takes parameters such as $\mathbf{H}$, $\mathbf{r}$, maxIter, and ε, where maxIter is the number of maximum iterations, and ε is the cross error probability that is used to determine the initial value of the algorithm.

The error-correction performance of the algorithm depends on both maxIter and the crossover error probability. If maxIter is small, the algorithm fails to obtain a converged solution. If it is large, the algorithm may take a considerably long computational time to obtain its solution. In the literature, maxIter is set from 10 to 20 in general. Next, the crossover error probability is set if the transition probability of a binary symmetric channel is either given or estimated. If this is improperly set, the decoding performance is degraded, leading to the poor error-correction capability. For the purpose in correcting errors, this parameter must be carefully considered.

The aim of using the decoder in ECCPoW, please make note of the fact that, is not to correct errors. Thus, there is no need to set maxIter and ε strictly. One condition that we shall aim to satisfy is that all the miners in ECCPoW system have to use the same values for these parameters. This condition can be easily satisfied by letting them to be published fixed constants in the proposed implemented program. As the miners verify a newly published block before accepting it, there is no benefit not to follow and use different ones for these published parameters. That is, any proof obtained from arbitrary parameters other than the published ones must be rejected.

We will give details on how to construct the other parameters, such as **H** and **r**, in Section IV.

B. FPGA and ASIC Implementation

LDPC decoders based on ASIC, a.k.a. ASIC-LDPC decoders are implemented to achieve low power consumption and fast processing. In the decoders, the check and variable nodes have to be physically connected using logical gates for a given parity check matrix. Fixed connections are to imply limited flexibility on the designs on the ASIC-LDPC decoders, making them only to support either a set of pre-defined parity check matrices or structured parity check matrices. We give our surveys related to existing ASIC-LDPC decoders as follows.

First, the ASIC-LDPC decoder in [23] supports quasi-cyclic parity check matrices decomposed into cyclic-shifted identity and zero matrices. These matrices have the same structure that is used in implementing the decoder. Second, the ASIC-LDPC decoder in [24] supports the parity check matrices included in the IEEE 802.16e system. These matrices are fixed; not change. Third, in [25], Hanzo *et al.*, reviewed the state-of-the-art of ASIC-LDPC decoders and stated that these decoders must take a bank of hardware to support many random parity check matrices. Namely, additional components such as memories, controllers and switchable interconnections are required, resulting in that these components occupy the most chip area in the decoders. They supported their statement by providing an example of [26] in which the ASIC-LDPC decoder supports about 100 parity check matrices, but its additional components occupy 75% of the total area of the decoder. This shows that there are no practical implementations on ASIC-LDPC decoders to support an infinite number of random parity check matrices.

There are FPGA-LDPC decoders that are the LDPC decoders implemented on FPGA chips. FPGA-LDPC decoders consume more power rather than ASIC-LPDC decoders do. But, it is much easier to reprogram the FPGA-LDPC decoders, which implies that they can achieve the more flexibility on the designs compared to ASIC-LDCP decoders. The FPGA-LDPC decoder in [27] supports parity check matrices up to n = 65,000. But, it is required to load a parity check matrix onto this decoder when it has to be changed, requiring additional time. In [28], Hanzo et al., stated that FPGA-LDPC decoders require additional routing and processing devices to support many parity check matrices. But, as they pointed, the use of these additional devices can lead to complex designs, increasing the cost of the decoders.

IV. Error-Correction Codes Proof of Work

In this section, we give details on ECCPoW. For simplicity, we organize this section into four subsections. In the first subsection, we list fields belonging to a block header of ECCPoW and provide their simple explanations. In the second subsection, we illustrate an overall structure of ECCPoW and its explanations. In the third subsection, we explain how we construct two inputs that appear as we use a decoder of error-correction codes. In the last subsection, we give the definition of an ECC puzzle generation function and present how to define an ECC puzzle using this ECC function. We end the last subsection by giving routines for solving this ECC puzzle.

A. Block header in ECCPoW

The block header of ECCPoW is defined to be a data structure that has eight fields such as timestamp, Merkle tree value, previous hash value, nonce, version, code length, row degree and column degree.

We use the fields such as version, timestamp, previous hash value, Merkle tree value and nonce to achieve in the purpose of guaranteeing the immutability property similar to Bitcoin.

We use the remains such as the code length, the row degree and the column degree to assign the size of hash vector and parity check matrix (PCM), which appears due to the usage of a decoder of a family of LDPC codes. As we will show in Section V, we change the difficulty level of a puzzle by varying the code length if the column degree and the row degree are fixed.

B. Overall structure of ECCPoW

Fig. 1 is prepared to present an overall structure of ECCPoW. This structure consists of three parts such as *i*) the hash vector generation (HVG) part, *ii*) the LDPC decoder part and *iii*) the decision part. We explain each part as follows.

In the HVG part, we randomly generate a hash vector of size n using a series of SHA256s taking the CBH with a given nonce generated by the nonce generator. The details how to generate this hash vector from the knowledge of the CBH will be given in the subsection IV.C.

In the LDPC decoder part, there is a decoder from a family of LDPC codes. This decoder takes the above hash vector and runs the message-passing algorithm [32] to yield a binary word **c**. It is noted that this decoder takes a parity check matrix (PCM) **H** determining the relation between an input and its corresponding output. The details how to construct this PCM will be given in the subsection IV.C.

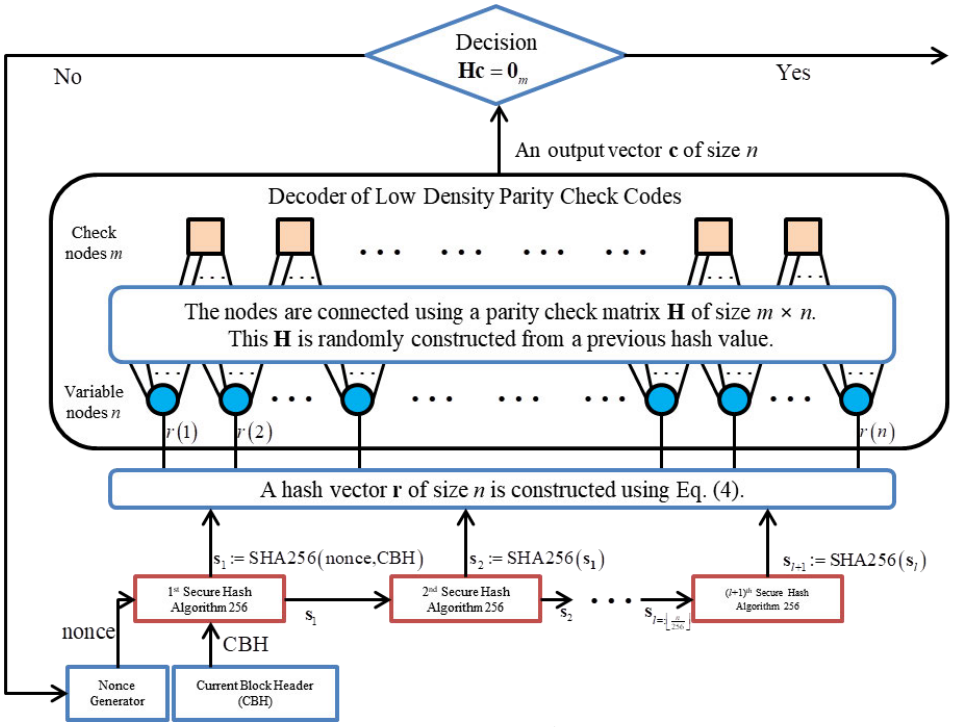


Fig. 1. An overall scheme of ECCPoW.

Last, in the decision part, the decision is made on the basis of the output provided by the decoder, as we have shown in Fig. 1.

C. Construction of hash vector and parity check matrix in ECCPoW

First, we provide the definition of hash vector **r** and give how we construct this hash vector using the current block header.

Definition 1 – *Hash Vector*: A hash vector **r**, which is a vector of concatenating outputs of SHA256s, of size n is defined to be:

$$\mathbf{r} := \begin{cases} \mathbf{s}_1[1:n] & \text{if } n \le 256 \\ \left[\mathbf{s}_1 \quad \cdots \quad \mathbf{s}_l \quad \mathbf{s}_{l+1}[1:j]\right] & \text{if } n > 256 \end{cases} \tag{4}$$

where $l = \lfloor n/256 \rfloor$, $j = n - 256 \times l$,

$$\mathbf{s}_1 := \mathrm{SHA256}(\mathrm{CBH}) \in \{0,1\}^{256} \tag{5}$$

and

$$\mathbf{s}_u := \mathrm{SHA256}(\mathbf{s}_1) \in \{0,1\}^{256} \tag{6}$$

where $u = 2, 3, \ldots, l + 1$ and CBH is the current block header.

The current block header represented to be CBH in (5) is the on-chain information that is stored in the Internet. Anyone thus can access this on-chain information, leading to that anyone can make the same hash vector that a miner made during his mining competition work.

Second, we give a construction method of PCM for a given previous hash. This construction method has to be designed to satisfy two conditions:

C1. Any verifier can reconstruct the PCM using on-chain information that the miner has used.

C2. Formation of a PCM can vary from block to block.

First, suppose that C1 is not met. One choice a miner can opt out is to include a constructed PCM in a block for making any verifier to check the validity of this block. This may result in, under the condition that the block size is fixed, reducing the number of transactions stored in the block. Second, suppose C2 is not met. We then remind the literature survey given in Section III. This survey is to indicate that developing ASIC-LDPC decoders for supporting a single PCM can be possible. Thus, if C2 is not satisfy, there is a possibility to develop ASIC mining devices. Thus, C1 and C2 must be satisfied simultaneously.

Now, we begin to explain the proposed construction method that can satisfy C1 and C2. First, we let you know that the PCM parameters such as the code length n, the row degree w_r and the column degree w_c are essentially required to construct a PCM **H**. More concretely, if they are provided, we can assign the value of m, the number of rows of **H** in (1). These parameters are included in the block header, as we have stated in the second paragraph in this section.

The proposed method is based on the method of Gallager [7]. This aims to construct **H** that can be decomposed into a set of sub-matrices of size $w_c \times n$ as follows:

$$\mathbf{H} = \begin{bmatrix} \mathbf{A} \\ \pi_1(\mathbf{A}) \\ \vdots \\ \pi_{w_c-1}(\mathbf{A}) \end{bmatrix} \in \{0,1\}^{\frac{nw_c}{w_r} \times n} \tag{7}$$

where $\pi_i(\mathbf{A})$ is the ith submatrix constructed by random per-

mutation of the columns of **A**, π_i is the ith permutation order, and

$$\mathbf{A} := \begin{bmatrix} \mathbf{1}_{w_r} & & \\ & \ddots & \\ & & \mathbf{1}_{w_r} \end{bmatrix} \in \{0,1\}^{w_c \times n}$$

whose the ith row has w_c 1s in a row from $(i-1)\times w_r$ to $i\times w_r$ and

$$\mathbf{1}_{w_r} := [1 \quad 1 \quad \cdots \quad 1] \in \mathbf{1}^{1\times w_r}.$$

Let w_c, w_r and n are fixed. Different PCM can be constructed by varying the permutation orders, leading to the serial creation of different PCMs. To come up with different random permutation from block to block, we use a previous hash. Namely, we let PBHV be assumed to be an array of 32 bytes to represent the given previous hash. We generate an initial seed value S below

$$S := \text{PBHV}[0] + \text{PBHV}[1] + \cdots + \text{PBHV}[31] \tag{8}$$

where PBHV[i] is the ith element of PBHV. The first permutation order is generated using a seed value S. The ith permutation order is then generated using $S - i + 1$. The pseudo code of this construction method of PCM is given in Table VI.

We explain how the pseudo code given in Table VI can satisfy the conditions C1 and C2 mentioned earlier. In Step 4, the ith permutation order is constructed using $S - 1 + i$. Any verifier can get the same value S without any communications because PBHVs are on-chain information and thus available within the chain. Thus, the verifiers can easily reconstruct what the miner has constructed in the past, which confirms that the proposed method satisfies C1. Second, PBHVs are hashes and thus possess the characteristics of random numbers; i.e., the initial seed value is a random number. All of the permutation orders are provided using the initial seed value. Thus, the orders vary from block to block, which confirms that the proposed method satisfies C2 as well.

D. Construction of ECC puzzle generation functions and ECC puzzles

Let the current block header (CBH) except nonce be given. We define an ECC puzzle generation function by concatenating the HVG part and the decoder part.

Definition 3 – *ECC puzzle generation function*: Let the current block header (CBH) except nonce be given. For this given CBH, an ECC puzzle generation function (ECCPGF) is defined to be a composite function as follows:

$$h_{\text{CBH}} : \{\text{nonce}\} \mapsto \mathbf{c} \in \{0,1\}^{n\times 1} \tag{9}$$

where **c** is the output of a decoder defined in Definition 4.

Definition 4 – *Decoder*: A decoder takes both the hash vector **r** in (4) and the PCM **H** in (7) as its inputs and runs the message passing algorithm given in [32] to yield a vector **c** of size n:

$$\mathcal{D}_{MP} : \{\mathbf{r}, \mathbf{H}\} \mapsto \mathbf{c} \in \{0,1\}^{n\times 1}. \tag{10}$$

TABLE V. The pseudo codes for ECCPoW

Inputs: CBH and PCM **H**	
Step 1:	A nonce is uniformly chosen from [0, $2^{32} - 1$]
Step 2:	Construct a HV **r** using (4) with a chosen nonce and the given CBH.
Step 3:	Obtain a vector **c** using (10) with the given PCM **H**.
Step 4:	If the constraint in (11) is satisfied, then go to Step 5.
Step 5:	Block generation & broadcast

TABLE VI. The pseudo codes to construct PCM

Inputs: n, w_c, w_r and BHV	
Output: **H**	
Step 1:	Construct S using (8).
Step 2:	Construct **A** by following the statements below (7) and **H** = **A**.
Step 3:	for i = 2 to $w_c - 1$
Step 4:	Construct π_i with the seed value $S - i + 1$.
Step 5:	$\mathbf{H} = \left[\mathbf{H} \quad \pi_i(\mathbf{A})^T\right]$.
Step 6:	end
Step 7:	$\mathbf{H} = \mathbf{H}^T$.

When the CBH with a selected nonce is given, we can make a hash vector **r** according to (4). Using the previous hash value included in the CBH, we can make a PCM **H** as well. The decoder then takes both of them to yield the binary word, as we have shown in (10). This binary word is the output of ECCPGF, as we have shown in (9)

A mapping rule in the LDPC decoder depends on the form of **H**. As we have discussed in the subsection IV.C, we can choose to vary **H** from block to block. As blocks are mined endlessly, infinitely many PCMs can be made. Thus, as we have mentioned in the subsection III.B, this can deter the development of an ASIC-LDPC decoder. Making ASIC chips to function as an ECCPGF becomes extremely difficult, if not impossible. Thus, we use this ECCPGF to define an ECC puzzle as follows.

Definition 5 – *ECC Puzzle*: An ECC puzzle constructed using a given ECCPGF defined in (9) is defined to be

$$\text{find} \quad \textit{nonce} \quad \text{subject to} \quad \mathbf{H}h_{CBH}(\textit{nonce}) = \mathbf{0}_m. \tag{11}$$

Namely, this puzzle is a problem where we aim to find a nonce satisfying the constraint given in (11).

We shall note that this puzzle can be time-variant from block to block and resistant to ASIC chips. We will provide details in Section VI.

Last, we provide codes to solve a puzzle in Table V. In Step 1, a nonce is selected from 0 to $2^{32} - 1$. We construct a hash vector **r** using (4), as we have shown in Step 2. In Step 3, we execute the decoder to give an output by taking both **r** and **H**. The decision is done using this output in Step 4. If the output is not a codeword, we repeat the routines from Step 1 to Step 4. Similar to Bitcoin, there is a case in which we cannot find the solution even we consider the whole nonces. In such a case, we modify the fields such as timestamp and Merkle tree value of the current block header. This modification can lead to the variation of the hash vector (4). We then repeat the whole routines with this modified block header.

V. Analysis on Error-Correction Codes Proof of Work

To solve an ECC puzzle, we repeat the routines from Step 1 to Step 4 many times until finding a nonce that satisfies the constraint in (11). A simple question such that what is the number of trials for finding this nonce naturally arises. To this end, we conduct a probabilistic study for giving answers to the following three questions:

Q1. What is the number of hash cycles needed to solve a given puzzle?
Q2. Which are the parameters which affect the number of hash cycles needed?
Q3. How does the number of miners affect the number of hash cycles needed?

We define the hash cycle, the success event, and the mining game, respectively.

Definition 6 – *Hash Cycle*: The single execution of the whole routines from Step 1 to Step 4 given in Table V is defined to be single hash cycle.

Definition 7 – *Success Event*: A success event occurs if a nonce such that the decoder defined in (10) can return a codeword is found, i.e., the constraint in (11) is satisfied.

Definition 8 – *Mining Game*: Let both a PCM $\mathbf{H}$ of size $m \times n$ and a CBH except for nonce be given. There are M miners, and each use a single computer with the same computing capacity. A mining game (MG)

$$\mathrm{MG}\left\{\mathbf{H}, \mathrm{CBH}, M, p\right\} \tag{12}$$

is defined that the miners compete with each other in a race to hit the success event first. We let p be the decoding success (DS) probability of the decoder for this given PCM $\mathbf{H}$, i.e.,

$$p \coloneqq \Pr\left\{\mathbf{r} : \mathbf{H}\mathbf{c} = \mathbf{0}_m\right\} \tag{13}$$

where $\mathbf{c}$ is the output of the decoder defined in (10).

For a given PCM $\mathbf{H}$ of size $m \times n$, there are 2^k codewords:

$$\left\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \cdots, \mathbf{c}_{2^k}\right\}.$$

Then, we define a sphere set for the given i^{th} codeword

$$\mathcal{A}\left(\mathbf{c}_i, l\right) \triangleq \left\{\mathbf{r} : \left\|\mathbf{r} - \mathbf{c}_i\right\|_h \le l \cap \mathbf{r} \in \left\{0,1\right\}^n\right\}$$

whose cardinality is

$$\left|\mathcal{A}\left(\mathbf{c}_i, l\right)\right| = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{l} = \sum_{l=0}^{s} \binom{n}{l}$$

where s is a positive integer and $\left\|\mathbf{x}\right\|_h$ is the number of 1s in $\mathbf{x}$.

We assume that the decoder defined in (10) is optimal, implying that it can correct up to t errors where t is obtained by (3). The decoder always yields the ith codeword when it takes an input belonging to the ith sphere, i.e.,

$$\mathcal{D}_{MP} : \left\{\mathbf{r}, \mathbf{H}\right\} \mapsto \mathbf{c}_i$$

where $\mathbf{r} \in \mathcal{A}\left(\mathbf{c}_i, t\right)$. Then, we have

$$\begin{aligned} p &= \sum_{i=1}^{2^k} \Pr\left\{\mathbf{c} = \mathbf{c}_i\right\} = \sum_{i=1}^{2^k} \sum_{l=0}^{t} \Pr\left\{\left\|\mathbf{r} - \mathbf{c}_i\right\|_h = l\right\} \\ &= \sum_{i=1}^{2^k} \Pr\left\{\mathbf{r} \in \mathcal{A}\left(\mathbf{c}_i, t\right)\right\} \end{aligned} \tag{14}$$

where $\mathbf{c}$ is an output of the decoder which takes a hash vector $\mathbf{r}$. Since the number of inputs that can be mapped into one of the codewords is

$$\sum_{i=1}^{2^k} \left|\mathcal{A}\left(\mathbf{c}_i, t\right)\right|,$$

the DS probability p can be expressed as follows:

$$p = 2^{-n} \sum_{i=1}^{2^k} \left|\mathcal{A}\left(\mathbf{c}_i, t\right)\right| = 2^{k-n} \left|\mathcal{A}\left(\mathbf{c}_i, t\right)\right| = 2^{k-n} \sum_{l=0}^{\left\lfloor \frac{d-1}{2} \right\rfloor} \binom{n}{l} \tag{15}$$

where the third equality comes from (3).

Intuitively, the number of trials increases as p decreases. It is natural to find which parameters make effects on p. To this end, we establish Proposition 1 to provide the behavior of p in terms of the code length n under the assumption that the row degree and the column degree are fixed.

Proposition 1 – Let $w_c \ge 3$, $w_r > w_c$ be fixed constants and their ratio be a fixed constant as well

$$w_c / w_r =: \alpha \in \left(0,1\right). \tag{16}$$

Let the size of a given PCM $\mathbf{H}$ be $m \times n$. For any $0 < \delta < 1/2$, we have

$$2^{-n\alpha} \le p \le 2^{-n\left(\alpha - H\left(\delta/2\right)\right)} \tag{17}$$

where $H(x)$ is the binary entropy function defined as follows:

$$H\left(x\right) = -x \log_2 x - \left(1-x\right) \log_2 \left(1-x\right).$$

Indeed, let the ratio further satisfy the following:

$$\alpha \in \left(H\left(0.25\right), 1\right). \tag{18}$$

Then, the DS probability p vanishes with increase in n.

Proof: From (1) with (16), we infer that

$$k - n = -n\alpha. \tag{19}$$

The results [7] state that the minimum distance of any regular LDPC code with constant w_r and $w_c \ge 3$ can linearly increase with increase in n. This statement indicates that the distance can be expressed as for any $0 < \delta < 1/2$,

$$d = \left\lfloor \delta n \right\rfloor.$$

Substituting (19) into (15) leads to an upper bound to p

$$p \le 2^{k-n} \sum_{l=0}^{\left\lfloor \frac{\delta n}{2} \right\rfloor} \binom{n}{l} \le 2^{k-n} 2^{nH\left(\frac{\left\lfloor \delta n \right\rfloor}{2n}\right)} \le 2^{-n\left(\alpha - H\left(\delta/2\right)\right)} \tag{20}$$

where the second inequality comes from the fact that any integers $n \geq k \geq 1$ with $k/n \leq 0.5$

$$\sum_{l=0}^{k} \binom{n}{l} \leq 2^{nH(k/n)}.$$

By assuming $t = 0$ in (14), the lower bound is obtained below:

$$p = 2^{k-n} \left| \mathcal{A}(\mathbf{c}_i, t) \right| \geq 2^{k-n} \left| \mathcal{A}(\mathbf{c}_i, 0) \right| = 2^{k-n} = 2^{-n\alpha}.$$

Last, let the ratio satisfy (18). Then, the term of the exponent in the upper bound in (17) is negative because for any $0 < \delta < 1/2$,

$$H(\delta/2) < H(0.25).$$

Thus, increasing n makes the upper bound on p reduced. This leads to that p vanishes as n goes to infinity. ■

We provide quick discussions on Proposition 1. First, for the fixed constant α, decreasing the code length n makes the lower bound on the DS probability p grow. Second, if we select any pairs of w_c and w_r which can satisfy (18), then p vanishes as n increases. This means that no one solves an ECC puzzle using these pairs for sufficiently large n. We thus have to avoid selecting such pairs for preventing this critical problem.

We define a random variable to represent the number of hash cycles required to end a MG and provide its statistical properties, respectively.

Definition 8 – *Random Variable*: For a given MG, X_M is defined as a random variable that represents the number of hash cycles to end this given MG, where the subscript M denotes the number of miners forging simultaneously and independently. We call this random variable First Success Hash Cycle (FSHC).

Theorem 2 – Let an MG{**H**, CBH, M, p} be given. Then the distribution that FSHC occurs at the lth hash cycle is

$$\Pr\{\mathrm{X}_M = l\} = p_{\mathrm{f,a}}^{l-1}\left(1 - p_{\mathrm{f,a}}\right)$$

which is a geometric distribution with a parameter

$$\left(1 - p_{\mathrm{f,a}}\right) \tag{21}$$

where $p_{\mathrm{f,a}} \coloneqq (1-p)^M$ is the probability that all of the miners fail to succeed in solving a given puzzle. Then, we have

$$\mathbb{E}[\mathrm{X}_M] = \left(1 - p_{\mathrm{f,a}}\right)^{-1} \tag{22}$$

and

$$\mathbb{V}[\mathrm{X}_M] = p_{\mathrm{f,a}}\left(1 - p_{\mathrm{f,a}}\right)^{-2}.$$

Proof: The proof is clear, as X_M follows the geometric distribution with (21). We thus omit it. ■

For the constant α defined in (16), Proposition 1 shows that the DS probability p grows with decrease in the code length n. When p grows, the parameter (21) reduces and converges to a real positive number. Thus, the expected value gets reduced and converges to a number, implying that an ECC puzzle becomes easier as n gets smaller.

TABLE VII. The lower bound given in (23)

$w_c = 4$ and $w_r = 5$	$M = 1$	$M = 5$	$M = 20$
Lower bounds, i.e., $\delta_1 = 0.3238$ in (23)			
$n = 80$, $k = 12$	1.58×10^{4}	0.31×10^{4}	0.08×10^{4}
$n = 120$, $k = 24$	6.03×10^{7}	1.20×10^{7}	0.30×10^{7}
$n = 160$, $k = 32$	2.46×10^{9}	0.49×10^{9}	0.12×10^{9}

Next, we consider a case in which n grows. In this case, the upper bound given in (17) is too loose to be considered unless α satisfies (18). Thus, we require another upper bound on p which is tighter than the previous upper bound. To this end, we invoke a table of [31]. For a certain pair of the column degree w_c and the row degree w_r, this table was obtained and given in the form of an upper bound δ_1 to the relative minimum distance δ_0, the ratio of the minimum distance d to the code length n. For $w_c = 4$ and $w_r = 5$, for example, the upper bound is 0.3238. For $w_c = 4$ and $w_r = 8$, it is 0.1765. The result was obtained from an asymptotic analysis; i.e., by letting n go to infinity.

We use the upper bound δ_1 to obtain a tighter bound on the DS probability p as follows:

$$p \leq g(n, k, \delta_1) \tag{22}$$

where δ_1 is given in the table of [31] for a certain pair of w_c and w_r and

$$g(n, k, \delta) \coloneqq 2^{k-n} \sum_{l=0}^{\left\lfloor \frac{\lfloor n\delta \rfloor - 1}{2} \right\rfloor} \binom{n}{l}.$$

The bound in (22) is obtained by simply replacing the minimum distance with the upper bound, respectively:

$$d \coloneqq n\delta_0 \leq \lfloor n\delta_1 \rfloor.$$

Once the upper bound on p has been obtained, we can use it to find a lower bound to the expected value of FSHC below:

$$\frac{1}{1 - \left(1 - g(n, k, \delta_1)\right)^M} \leq \mathbb{E}[\mathrm{X}_M] \tag{23}$$

We can examine the behavior of (22) with respect to the code length n and the number of miners M. In Table VII, we provide the lower bounds to the expected value by varying either n or M. They are obtained for a case in which the column degree w_c and the row degree w_r are 4 and 5, respectively. We can see that the lower bound increases with increase in n for the fixed M. That is, it increases from 1.58×10^4 to 2.46×10^4 when n is increased from 80 to 160. This result implies that increasing n makes an ECC puzzle more difficult to solve. For the other pairs of w_c and w_r given in [31], the same result is observed.

Now, for the fixed code length n, we consider the behavior of the expected value of FSHC by growing the number of miners M. Intuitively, as more miners are involved in solving a given puzzle, this puzzle has to end early. In addition, if an infinite number of miners work, any MG has to end at the 1st hash cycle. These intuitions can be confirmed by Corollary 1 given below.

Corollary 1: Let a MG{**H**, CBH, *M*, *p*} be given. The expected value of a FSHC given in (22) decreases with increase in the number of miners *M*. In particular, this value can converge to 1 as *M* goes to infinity.

Proof: It is immediately seen that

$$\frac{d\mathbb{E}[\mathrm{X}_M]}{dM} = -\frac{\log p_{\mathrm{f,a}}}{p_{\mathrm{f,a}}} \le 0$$

implying that the expected value given in (22) is a decreasing function of *M*. As *M* goes to infinity, the parameter defined in (21) goes to one. Thus, the expected value converges to one. ■

The decoding process, i.e., Step 3 in Table V, can occupy the most computational time in a single hash cycle. In this decoding process, matrix-vector products are required, implying that the computational cost to run a single hash cycle can be modeled as O(*mn*). Each miner uses the same single computer in a given MG. Thus, we can assume that each miner only runs τ operations per second. This assumption makes us define an expected value of a block generation time as follows.

Definition 10 – *Block Generation Time*: A MG{**H**, CBH, *M*, *p*} is given. Each miner is assumed to run τ operations per second. Then, the block generation time *T* can be defined as

$$T \coloneqq \tau^{-1}\mathbb{E}[\mathrm{X}_M]O(mn). \tag{24}$$

Both Proposition 1 and Theorem 2 can indicate that the expected value in (24) is an increasing function of the code length *n*. We thus immediately conclude that the block generation time *T* is an increase function of *n*.

VI. Discussions on ECCPoW

We prepare this section to provide plentiful discussions on ECCPoW. To this end, we begin to define general properties of PoWs and explain how ECCPoW can have them. We introduce a new property that ECCPoW can only have. This new property makes ECCPoW become a solution to the problems such as M1 and M2 which we have stated in Section I.

We now begin to define properties below.

P1. A puzzle has to be time-consuming, but it is easy to check whether a given solution is correct or not.
P2. Any previous solution cannot be used to find a current solution.
P3. A puzzle can be solved with overwhelming probability 1 if and only if miners follow the routines of PoW.
P4. The difficulty of a puzzle can change.
P5. A puzzle can be time-variant from block to block.

The existing PoWs have the properties from P1 to P4. Let us begin to consider how bitcoin satisfies them. First, each puzzle is expected to be solved per 10 minutes. In contrast, validating a given solution can be instantly done. Second, SHA256 takes the block header to get its hash. The block header in Bitcoin has the timestamp field. Due to this timestamp field, the contexts of the current block header can be different to those of any previously mined blocks. Thus, any solutions given in these mined blocks are useless in finding a current solution. Thus, P2 holds. Third, the number of possible hashes is 2^{256} while the number of solutions is about $2^{(256-L)}$ where *L* is a pre-defined value according to the difficulty of the provided block header. Thus, a possibility that a provided nonce is a solution is 2^{L}. In the 567,657th block of bitcoin, *L* is 72. This is to imply that for this block, the probability that a randomly given nonce can be a solution is 2^{-72}. Thus, all the miners have to follow the routines in Table I to solve a puzzle, meaning P3. Next, whenever 2016 blocks are mined, the difficulty periodically changes, implying P4. Last, existing PoW does not hold P5. The reasons for this issue will be given after explaining how ECCPoW satisfies P5.

Now, we begin to explain how ECCPoW can satisfy all the properties using the results established in the previous section.

Corollary 2: ECCPoW can satisfy the first property P1.

Proof: For a given nonce, we complete the verification whether this nonce is a solution or not by conducting the steps from Step 1 to Step 4 in Table V. This verification requires a single construction of a hash vector and a single execution of the decoder (10). In contrast, to solve an ECC puzzle, one has to do work by repeating them many times, as we have stated in Section V. ■

Corollary 3: ECCPoW can satisfy the second property P2.

Proof: As we have stated in Definition 5, we construct an ECC puzzle using an ECCPGF in Definition 3. This ECCPGF takes a hash vector as its input and produces a binary word as its output. The decision is made based on this binary word. To make this hash vector, we put a given set of the block header with a nonce through SHA256. Similar to Bitcoin block header, our block header includes the timestamp field. Thereby, one cannot effortlessly create a particular hash vector even if one has the full knowledge of all the previous solutions, i.e., all the collection of the nonces each of which was a solution to mine a block in the past. This makes ECCPoW satisfy P2. ■

Let us assume that there exists a knowledgeable miner who can solve ECC puzzles by referring to an input-output mapping table of ECCPoW without actually carrying out the decoding work. This miner can then mine blocks much faster than other honest miners can. To show the non-existence of such a miner in ECCPoW, we prepare Corollary 4. This corollary is to imply that all miners cannot but have to run the decoder to solve ECC puzzles.

Corollary 4: ECCPoW satisfies the third property P3.

Proof: Such a malicious miner can appear under an assumption that this miner has a mapping table that maps a given input to its corresponding output without actually carrying out the decoding work (10). That is, he puts each hash vector into the decoder to find what this decoder returns. He then uses this known information to construct the mapping table. However, the number of hash vectors is 2^{n}. For sufficiently large *n*, i.e. $2^{256} \sim 10^{77}$ for $n = 256$, we can easily see that this construction is impossible. Besides, this table depends on a PCM **H**, meaning that the mapping table is updated newly whenever **H** varies. As we have stated in the subsection IV.C, **H** can be set to vary from block to block. Thereby, whenever he aims to solve a new ECC puzzle, he has to construct a new mapping table. The above

assumption there exists the mapping table, we thus conclude, is invalid. Thereby, any miners in ECCPoW cannot but have to obey the routines in Table V to solve puzzles; this leads to the conclusion that ECCPoW satisfies P3. ■

Now, we remind the discussions regarding Theorem 2. In the discussions, we have proven that an ECC puzzle can become easier as the code length n gets smaller. We then have empirically shown that the increase in n makes this puzzle more difficult to solve. They are to indicate that ECCPoW holds P4.

Corollary 5: ECCPoW can satisfy the fourth property P4.
Proof: It is clear; we thus omit it. ■

A PCM **H** defines a mapping function of the decoder in (10). This decoder is used to define an ECCPGF in (9) which is used to construct an ECC puzzle. We conclude that this ECC puzzle can be a function of **H**. As we have mentioned in the subsection IV.C, we construct **H** using a previous hash, i.e., the ith PCM is made using a hash of the $(i–1)^{\text{th}}$ block. Such a construction can make **H** time-variant from block to block, leading to that the ECC puzzle can also be time-variant from block to block. This discussion leads to a conclusion that ECCPoW holds P5 below.

Corollary 6: ECCPoW can satisfy the fifth property P5.
Proof: It is clear; we thus omit it. ■

We now prepare to provide discussions how ECCPoW can be a solution to the problems such as *i*) M1 the re-centralization of the mining markets and *ii*) M2 the huge energy consumption for mining. As miners mine new blocks continuously, an infinite number of PCMs are constructed. Hence, we cannot count how many the number of PCMs is required in advance. Also, we cannot expect what PCMs will be generated. We remind the example given in [25], showing that there is no ASIC-LDPC decoder to support the infinite number of PCMs. By combining this example with the fact that there is an infinite number of PCMs are required, we thus conclude that the LDPC decoder in ECCPoW is operated by either graphical processing units or CPUs. This is the fundamental reason that ECCPoW becomes the solution remediating the aforementioned problems M1 and M2 which we have stated in Section I. This is the most remarkable contribution of ECCPoW.

We end this section by providing a reason that SHA based PoW does not hold P5. We consider SHA256 as a puzzle generation function (PGF) in Bitcoin because it is used to generate a puzzle. This PGF is fixed regardless of the height of a block that miners aim to mine. The puzzle made using a fixed PGF is not time-variant at all, leading to that such a puzzle cannot hold P5.

VII. Conclusions

PoW is fundamental to public blockchains, as it can be used to prohibit an unauthorized modification of mined blocks. For existing PoWs, ASIC mining devices have been introduced and used to mine blocks. The usage of such devices can cause the problems such as M1 and M2 that we have stated in Section I.

In this paper, as a solution to these problems, we proposed a new proof-of-work using error-correction codes which we call ECCPoW. To the best of our knowledge, this is the first study in which a decoder of LDPC codes is applied to the consensus part of blockchain. Specifically, we combined this decoder with SHA256 to construct a composite function named as ECCPGF (9). We used ECCPGF to define a corresponding ECC puzzle (11) and provided the routines to solve a given ECC puzzle.

We also studied the behavior of the expected value of the number of hash cycles for solving a given ECC puzzle. We showed that this value can be either increased or decreased as varying the code length, the size of a hash vector taken by the decoder, and the number of miners. Indeed, we discussed how ECCPoW can satisfy the five properties defined in Section VI, which shows the value of ECCPoW as a general PoW.

As we have reviewed in Section III, there is no ASIC decoder to support an infinite number of LDPC codes. By motivated this survey, we intended to vary the codes from block to block. As a result, we made ECCPGF time-variant, meaning that its mapping function can vary from block to block. This leads to the time-variant property P5 defined in Section VI. This is the most innovative part of ECCPoW in repressing the advent of ASICs, implying that the problems caused from the usage of ASICs can be solved using our ECCPoW.

We have implemented ECCPoW and forked two blockchains such as Bitcoin and Ethereum by replacing their consensus with the implemented ECCPoW. We name these forked versions as BTC-ECC and ETH-ECC, respectively. All of the source codes including ECCPoW, BTC-ECC and ETH-ECC can be available in a GitHub site [34]. We have also provided manuals that shows how to install them, how to compile them and how to run. We believe that anyone can easily operate their own BTC-ECC or ETH-ECC by following these manuals. We believe that this site can be the first repository that makes people in the error-correction codes community get involved in the blockchain community.

References

[1] U. W. Chohan, "The double spending problem and cryptocurrencies", SSRN Electronic Journal, Jan. 2019

[2] Nakamoto, S. *Bitcoin: A Peer-to-Peer Electronic Cash System*. Retrieved from: https://bitcoin.org/bitcoin.pdf

[3] Quynh H. Dang, "Secure Hash Standard FIPS PUB 180-4," U.S. Dept. of Commerce and NIST, Aug. 4th 2015. https://www.nist.gov/publications/secure-hash-standard, Accessed Sept. 25th, 2018.

[4] Jordan Tuwiner, "Bitcoin mining hardware ASICs," https://www.buybitcoinworldwide.com/mining/hardware/, June 30th, 2018. Accessed Sept. 26th, 2018.

[5] A. Gervais, G. Karame, V. Capkun and S. Capkun, "Is bitcoin a decentralized currency?", *IEEE Security and Privacy Magazine*, Vol. 12, No. 3, pp. 54 – 60, May, 2014

[6] Digiconomist. (2018). *Bitcoin Energy Consumption Index*. Retrieved April 3, 2018 from: https://digiconomist.net/bitcoin-energy-consumption

[7] R. G. Gallager, *Low Density Parity Check Codes*, Monograph, M.I.T. Press, 1963

[8] R. J. McEliece, "A public-key cryptosystem based on algebraic coding theory", DSN Progress Report, pp. 114-116, Jan.~Feb. 1978

[9] D. J. Bernstein, T. Lange and C. Peters, "Wild McEliece", in *Proc. Sel. Areas Cryptography*, vol. 6544, pp. 143 – 158, 2011

[10] M. Baldi, M. Bianchi, F. Chiaraluce, J. Rosenthal and D. Schipani, "Using LDGM codes and sparse syndromes to achieve digital signatures", in *Post-Quantum Cryptography 2013, ser. Lecture Notes in Comput. Sci.*, vol. 7931, Springer, pp. 1 – 15, 2013.

[11] Marco Baldi, "LDPC codes in the McEliece cryptosystem: attacks and countermeasures", NATO Science for Peace and Security – D: Information and Communication Security, vol. 23, pp. 160 – 174,

[12] M. Baldi, M. Bodrato and F. Chiaraluce, "A new analysis of the McEliece cryptosystem based on QC-LDPC codes", in 6th Int. Conf. on Sec. and Cryptography for Networks. Springer, pp. 246 – 262, 2008

[13] T. P. Berger, P-L. Cayerel, P. Gaborit and A. Otmani, "Reducing key length of the McEliece cryptosystem", in Progress in Cryptology, Africacrypt'2009, ser. LNCS, vol. 5580. Springer, pp. 77 – 97, 2009

[14] V. M. Sidelnikov, "A public-key cryptosystem based on binary Reed–Muller codes," Discrete Mathematics and Applications, vol. 4 no. 3, 1994.

[15] L. R. Knudsen and B. Preneel, "Construction of secure and fast hash functions using nonbinary error correcting codes", IEEE Transactions on Information Theory, vol. 48, no. 9, pp. 2254 – 2539, Sept., 2002.

[16] S. Ermon, C. P. Gomes, A. Sabharwal and B. Selman, "Low-Density Parity Constraints for Hasing-Based Discrete Integration", Proceedings of the 31st International Conference on Machine Learning, PMLR 32(1):271-279, 2014.

[17] Vadhan, S., *Pseudorandomness.* Foundations and Trends in Theoretical Computer Science, 2011.

[18] Vitalik Buterin. Dagger: *A memory-hard to compute, memory-easy to verify scrypt alternative*, Tech Report, hashcash.org website, 2013.

[19] https://github.com/dashpay/dash/wiki/Whitepaper

[20] T. Black and J. Weight, X16R ASIC resistant by design, 2018. https://ravencoin.org/assets/documents/X16R-Whitepaper.pdf

[21] C. E. Shannon, "A mathematical theory of communication", The Bell System Technical Journal, Vol. 27, 99. 379 – 423, 623 – 656, Jul., Oct., 1948

[22] D. J. C. Mackay and R. M. Neal, "Near Shannon limit performance of low density parity check codes", Electronics Letters, Vol. 33, No. 6, pp. 457 – 458, 1997

[23] M. Awais and C. Condo, "Flexible LPDC decoder architectures", VLSI Design, Vol. 2012, Doi: 10.1155/2012/730835

[24] C-H. Liu, S-W. Yen, C-L. Chen, H-C. Chang, C-Y. Lee and S-J. Jou, "An LDPC decoder chip based on self-routing network for IEEE 802.16e Applications" IEEE Journal of Solid-State Circuits, Vol. 43, No. 3, pp. 684 – 694, Mar. 2008

[25] S. Shao, P. Hailes, T-Y. Wang, J-Y. Wu, R. G. Maunder, B. M Al-Hashimi and L. Hanzo, "Survey of Turbo, LDPC and Polar Decoder ASIC Implementation", IEEE Communications Surveys & Tutorials 2019

[26] Y. L. Ueng, B. J. Yang, C. J. Yang, H. C. Lee, and J. D. Yang, "An Efficient Multi-Standard LDPC Decoder Design Using Hardware-Friendly Shuffled Decoding," IEEE Trans. Circuits Syst. I, vol. 60, no. 3, pp. 743–756, March 2013.

[27] C. Beuschel and H-J. Pfleiderer, "FPGA implantation of a flexible decoder for long LDPC codes", Proc. Int. Conf. Field Program. Logic Appl., Sep. 2008, pp. 185–190

[28] P. Hailes, R. G. Maunder and L. Hanzo, "A survey of FPGA-based LDPC decoders", IEEE Communications Surveys & Tutorials, Dec. 2015.

[29] A. B. Keha and T. M. Duman, "Minimum distance computation of LDPC codes using a branch and cut algorithm", IEEE Transactions on Communication, Vol. 58, No. 4, pp. 1072 – 1079, Apr., 2010

[30] Y. Hashemi and A. H. Banihashemi, "Tight Lower and Upper Bounds on the Minimum Distance of LDPC Codes", IEEE Communications Letters, Vol. 22, No. 1, pp. 33 – 36, Jan., 2018

[31] Y. Ben-Haim and S. Litsyn, "Upper bounds on the rate of LDPC codes as a function of minimum distance", IEEE Transactions on Information Theory, vol. 52, no. 5, pp. 2092 – 2100, May., 2006

[32] W. E. Ryan and S. Lin, *Channel Codes Classical and modern*, Cambridge, 2009

[33] Whitefire990, https://drive.google.com/file/d/14ZYwpI-t6rnQ3I_SEZHvvDwe2yQUnTJ5/view, Nov., 2019

[34] DeSecure Blockchain,https://github.com/cryptoecc/ETH-ECC, 2020

applied sciences

MDPI

Article

Profitable Double-Spending Attacks

Jehyuk Jang and Heung-No Lee *

School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology (GIST), 123 Cheomdangwagi-ro, Buk-gu, Gwangju 61005, Korea; jjh2014@gist.ac.kr

* Correspondence: heungno@gist.ac.kr

Received: 9 November 2020; Accepted: 25 November 2020; Published: 27 November 2020; Corrected: 15 September 2022



Abstract: Our aim in this paper is to investigate the profitability of double-spending (DS) attacks that manipulate an a priori mined transaction in a blockchain. It was well understood that a successful DS attack is established when the proportion of computing power an attacker possesses is higher than that of the honest network. What is not yet well understood is how threatening a DS attack with less than 50% computing power used can be. Namely, DS attacks at any proportion can be a threat as long as the chance to make a good profit exists. Profit is obtained when the revenue from making a successful DS attack is greater than the cost of carrying out one. We have developed a novel probability theory for calculating a *finite time* attack probability. This can be used to size up attack resources needed to obtain the profit. The results enable us to derive a sufficient and necessary condition on the value of a transaction targeted by a DS attack. Our result is quite surprising: we theoretically show how a DS attack at any proportion of computing power can be made profitable. Given one's transaction value, the results can also be used to assess the risk of a DS attack. An example of profitable DS attack against BitcoinCash is provided.

Keywords: blockchain; double-spending attack; Fraud risk analysis; profitability; time-finite analysis; probability distribution; combinatorics

1. Introduction

A blockchain is a distributed ledger which has originated from the desire to find a novel alternative to centralized ledgers such as transactions through third parties [1]. Besides the role as a ledger, blockchains have been applied to many areas, e.g., managing the access authority to shared data in the cloud network [2] and averting collusion in e-Auction [3]. In a blockchain network based on the proof-of-work (PoW) mechanism, each miner verifies transactions and tries to put them into a block and mold the block to an existing chain by solving a cryptographic puzzle. This series of processes is called mining. However, the success of mining a block is given to only a single miner who solves the cryptographic puzzle for the first time. The reward of minting a certain amount of coins to the winner motivates more miners to join and remain in the network. As a result, blockchains have been designed so that the validity of transactions is confirmed by a lot of decentralized miners in the network.

A consensus mechanism is programmed for decentralized peers in a network to share a common chain. If a full-node succeeds in generating a new block, it has the latest version of the chain. All of the nodes in the network continuously communicate with each other to share the latest chain. A node may run into a situation in which it encounters mutually different chains more than one. In such a case, it utilizes a consensus rule with which it selects a single chain. Satoshi Nakamoto suggested the longest chain consensus for Bitcoin protocol in which the node selects the longest chain among all competing chains [1]. There are also other consensus rules [4,5], but a common goal of consensus rules is to select the single chain by which the most computation resources have been consumed based on the belief that it may have been verified by the largest number of miners.

A double-spending (DS) attack aims to double-spend a cryptocurrency for the worth of which a corresponding delivery of goods or services has already been completed. The records of payment are written in transactions and shared in a network via the status-quo chain. Thus, to double spend, attackers need to replace the status-quo chain in the network with their new one, after taking the goods or services. For example, under the longest chain consensus, this attack will be possible if an attacker builds a longer chain than the status-quo. Nakamoto [1] and Rosenfeld [6] have shown that the higher computing power is employed, the higher probability to make a DS attack successful is. In addition, if an attacker invests more computing power than that invested by a network, a success of DS attack is guaranteed. Such attacks are called the 51% attack.

In the last few years, unfortunately, blockchain networks have been recentralized [7,8], which make them vulnerable to DS attacks. To increase the chance of mining blocks, some nodes may form a pool of computing chips. The problem arises when a limited number of pools occupy a major proportion of the computing power in the network. For example, the pie chart (date accessed from BTC.com on November 24, 2020) shown in Figure 1 illustrates the proportion of computing power in the Bitcoin network as of January 2020. In the chart, five pools such as F2Pool, BTC.com, Poolin, and Huobi.pool occupy more than 50% of the total computing power of Bitcoin. In a recentralized network, since most computing resources are concentrated on a small number of pools, it could be not difficult for them to conspire to alter the block content for their own benefits, if aiming to double-spend. Indeed, there have been a number of reports in 2018 and 2019 in which cryptocurrencies such as Verge, BitcoinGold, Ethereum Classic, Feathercoin, and Vertcoin suffered from DS attacks and millions of US dollars have been lost [9].

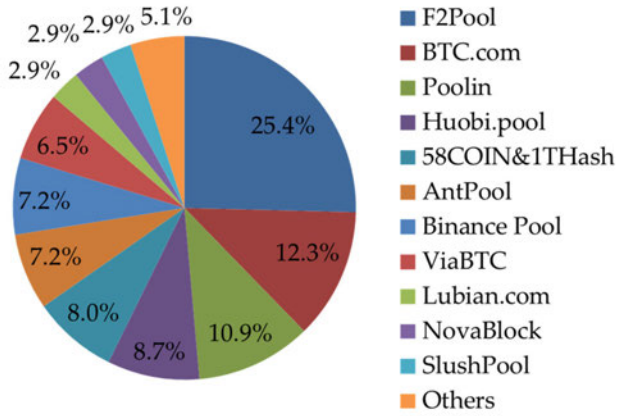


Figure 1. Computation power distribution among the largest mining pools.

In addition to the recentralization, the advent of rental services which lend the computing resources can be a concern as well [10]. Rental services such as nicehash.com which provide a brokerage service between the suppliers and the consumers have indeed become available. The rental service can be misused for making DS attacks easier. The presence of such computing resource rental services significantly reduce the cost of making a profit from double spending. This is because renting a required computing power for a few hours is much cheaper than building such a computing network. Indeed, nicehash.com attracts DS attackers to use their service by posting one-hour fees for renting 51% of the total computing power against dozens of blockchain networks on their website crypto51.app (accessed on 26 November 2020).

Success by making DS attacks is possible but is believed to be difficult for a public blockchain with a large pool of mining network support. By the results in [1,6], 51% attack has been considered as the requirement for a successful DS attack [11]. This conclusion however shall be reconsidered given our result in the sequel that there are significant chances of making a good profit from DS attacks

regardless of the proportion of computing power. The problem to consider, therefore, is to analyze the profitability of such attacks.

The analysis of attack profitability requires the ability to predict the time an attack will take, since the profit would be a function of time. Studies in [12–20] provided DS attack profitability analyses, but their time predictions were not accurate. Specifically, to make the time prediction easier, they either added impractical assumptions to the DS attack model defined by Nakamoto [1] and Rosenfeld [6] or oversimplified the time prediction formula (see Section 6 for details). Whereas, we follow the definition of DS attack in [1,6], and therefore we need to develop a new set of mathematical tools for precise analysis of attack profitability that we aim to report in this paper.

1.1. Contributions

We study the profitability of DS attacks. The concept of *cut-time* is introduced. Cut-time is defined to be the duration of time, from the start time to the end time of an attack. For each DS attempt, the attacker needs to pay for the cost to run his mining rig. A rational attacker would not, therefore, continue an attack indefinitely especially when operating within the regime of less than 50% computing power. To reduce the cost, the attacker needs to figure out how his attack success probability rolls out to be as the time progresses. We define that a DS attack is profitable if and only if the expected profit, the difference between revenue and cost (see Equation (33)), is positive. Our contributions are summarized into two folds:

First, we theoretically show that DS attacks can be profitable not only in the regime of 51% attack but also in the sub-50% regime where the computing power invested by the attacker is smaller than that invested by the target network. Specifically, a sufficient and necessary condition is derived for profitable DS attacks on the minimum value of target transaction. In the sub-50% regime, we also show that profitable DS attacks necessitate setting a finite cut-time.

Second, we derive novel mathematical results that are useful for an analysis of the attack success time. Specifically, the probability distribution function and the first moment expectation of the attack success time have been derived. They enable us to estimate the expected profit of a DS attack for a given cut-time. All mathematical results are numerically-calculable. All numerical examples of the theoretical results given in this paper are reproducible in our web-site (https://codeocean.com/capsule/2308305/tree).

1.2. Organization of the Paper

In Section 2, we define DS attack scenario and sufficient and necessary conditions required for successful DS attacks. Also, we define random variables that are useful in analyzing the attack profits. Section 3 comprises the analytic results of stochastics of the time-finite attack success. In Section 4, we define the profit function of DS attacks, followed by new theoretical results about the conditions for making them profitable. In Section 5, an example analysis of DS attack profitability in sub-50% regime against BitcoinCash network is given. Section 6 compares our results with related works. Finally, Section 7 concludes the paper with a summary.

2. The Attack Model

We define DS attack that we consider throughout this paper. We also define DS attack achieving (DSA) time, which is the least time spent for an occurrence of double-spending. The DSA time is a random variable derived from a random walk of Poisson counting processes (PCP).

2.1. Attack Scenario

We extend a DS attack scenario which has been considered by Nakamoto [1] and Rosenfeld [6]. Specifically, we add a time-finite attack scenario. There are two groups of miners, the normal group of honest miners and a single attacker. The normal group tends the honest chain.

When the attacker decides to launch a DS attack, he/she makes a target transaction for the payment of goods or services. In the target transaction, the transfer of cryptocurrency ownership from the attacker to a victim is written. We denote $t = 0$ as the time at which the last block of the honest chain has been generated. At time $t = 0$, the attacker announces the target transaction to normal group so that normal group starts to put it into the honest chain. At the same time $t = 0$, the attacker makes a fork of the honest chain which stems from the last block and builds it in secret. We refer to this secret fork as fraudulent chain. In the fraudulent chain, a fraudulent transaction is contained which alters the target transaction in a way that deceives the victim and benefits the attacker.

Before shipping goods or providing services to the attacker, the victim will obviously choose to wait for a few more blocks on the honest chain in addition to the block on which the his/her transaction has been entered, i.e., so-called block confirmation. Karame et al. [21] showed the importance of block confirmation: attackers are able to double-spend against zero block-confirmation even without mining a single block on the fraudulent chain at all. The number of blocks the victim chooses to wait for is referred to as the block confirmation number $N_{BC} \in \mathbb{N}$, which includes the block on which the target transaction is entered.

The attacker chooses to make the fraudulent chain public if his/her attack was successful. An attack is successful if the fraudulent chain is longer than the honest chain after the moment the block confirmation is satisfied. We define two necessary conditions $\mathcal{G}^{(1)}$, $\mathcal{G}^{(2)}$, for a success of DS attack:

Definition 1. *A DS attack succeeds only if there exists a DS attack achieving (DSA) time $T_{DSA} \in (0, \infty)$ such that*

1. *$\mathcal{G}^{(1)}$: (block confirmation) the length of the honest chain for the duration of time T_{DSA} has grown greater than or equal to N_{BC}, and*
2. *$\mathcal{G}^{(2)}$: (success in PoW competition) the length of the fraudulent chain for the duration of time T_{DSA} has grown longer than that of the honest chain.*

Rational attackers will not wait for his success indefinitely since growing the attacker's chain incurs the expense per time spent for operating the computing power. The attack thus shall put a limit to the end time to cut loss. We refer to this end time as the cut-time $t_{cut} \in \mathbb{R}^+$. A sufficient condition for the success of DS attack can be defined with applying the cut-time t_{cut}:

Definition 2. *For a given cut-time $t_{cut} \in \mathbb{R}^+$, the success of DS attack is declared if, and only if, there exists a DSA time $T_{DSA} \in (0, t_{cut})$ at which $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ in Definition 1 have been achieved.*

2.2. Stochastic Model

We model the conditions in Definition 2 with a stochastic model. We fit the block generation process using the PCP [22] with a given block generation rate λ (blocks per second). Including Nakamoto [1] and Rosenfeld [6], it has been most conventional to analyze the block generation process of a blockchain using PCP. A rationale why the block generation process is modeled as PCP is given in Bowden et al. [23], where experiments show the fitness of PCP model to real data samples from a live network.

We denote the lengths of the honest chain and the fraudulent chain over time $t \in (0, \infty]$ by two independent PCPs, $H(t) \in \mathbb{N}_0$ with the block generation rate λ_H (blocks per second) and $A(t) \in \mathbb{N}_0$ with the block generation rate λ_A, respectively. Both processes start at the time origin $t = 0$ (at which the DS attack is launched) at which the both chains are at the zero states, i.e., $H(0) = A(0) = 0$. Each chain independently increases at most by 1 at a time point. An increment of 1 in the counting process occurs when the pertinent network adds a new block to its chain.

We represent the difference between $A(t)$ and $H(t)$ in a discrete-time domain as a random walk $S_i \in \mathbb{Z}$ for $i \in \mathbb{N}$. For this purpose, we first define two continuous stochastic processes $M(t)$ and $S(t)$, which are respectively defined as

$$M(t) := H(t) + A(t), \tag{1}$$

and

$$S(t) := H(t) - A(t). \tag{2}$$

The first process $M(t)$ is also a PCP [22] with the rate

$$\lambda_T := \lambda_A + \lambda_H. \tag{3}$$

The second process $S(t)$ is the continuous-time analog of the random walk $S_i \in \mathbb{Z}$ for $i \in \mathbb{N}$ such that

$$S_i := S(T_i), \tag{4}$$

where T_i is the state progression time defined by

$$T_i := \inf\Big\{ t \in \mathbb{R}^+ : M(t) = i \Big\}, \tag{5}$$

which increases as i increases. Random walk S_i is a stationary Markov chain starting from $S_0 = 0$. The state transition probabilities [22] are given by

$$p_A := \Pr(S_i = n - 1 | S_{i-1} = n) = \frac{\lambda_A}{\lambda_T}, \tag{6}$$

and

$$p_H := \Pr(S_i = n + 1 | S_{i-1} = n) = \frac{\lambda_H}{\lambda_T}, \tag{7}$$

for all $i \in \mathbb{N}$ and $n \in \mathbb{Z}$. The state transition probabilities p_H and p_A are the proportions of computing power occupied by the normal miners and that by the attacker, respectively.

We define independent and identically distributed (i.i.d.) state transition random variables $\Delta_i \in \{\pm 1\} \sim \text{Bernoulli}(p_H)$ as

$$\Delta_i := S_i - S_{i-1}, \tag{8}$$

for $i \in \mathbb{N}$. Note that $S_i = \sum_{k=0}^{i} \Delta_k$.

Definition 3. *A DS attack* $\text{DS}(p_A, t_{cut}; N_{BC})$ *is a random experiment that picks a sample* $\omega \in \Omega$*. Each element* ω *is an infinite-length sequence of pairs of* T_i *and* Δ_i *in Equations (5) and (8) for all* $i \in \mathbb{N}$, i.e.,

$$\omega := ((T_1, \Delta_1), (T_2, \Delta_2), \cdots, (T_\infty, \Delta_\infty)). \tag{9}$$

The set Ω is the universal set of all possible ω, i.e.,

$$\Omega := \Big\{ \omega \in \Big\{ \mathbb{R}^+ \times \{\pm 1\} \Big\}^\infty \Big\}. \tag{10}$$

For given a DS sample $\omega \in \Omega$ and a state index $i \in \mathbb{N}$, we denote projections

$$\pi_{T_i}(\omega) := T_i \tag{11}$$

and

$$\pi_{\Delta_i}(\omega) := \Delta_i \tag{12}$$

that retrieve the progression time T_i and the transition Δ_i of the i-th state, respectively.

2.3. DS Attack Achieving Time

Definition 4. *For a DS sample ω of $\mathrm{DS}(p_A, t_{cut}; N_{BC})$, we define the DSA time T_{DSA} which measures the least one among the state progression times $\pi_{T_i}(\omega)$ of state indices i at which ω achieves the necessary conditions $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ in Definition 1.*

To express T_{DSA} as a random variable, we construct event sets $\mathcal{D}_j^{(1)} \subset \Omega$ and $\mathcal{D}_{i,j}^{(2)} \subset \Omega$. The sets $\mathcal{D}_j^{(1)}$ for $j \in \{N_{BC}, N_{BC}+1, \cdots, \infty\}$ consist of DS samples ω which achieves the block confirmation $\mathcal{G}^{(1)}$ at state j for the first time. The sets $\mathcal{D}_{i,j}^{(2)}$ for $i \in \{j, j+1, \cdots, \infty\}$ and $j \in \{N_{BC}, N_{BC}+1, \cdots, \infty\}$ consists of ω which achieves the success in the PoW competition $\mathcal{G}^{(2)}$ at state i for the first time, given that $\mathcal{G}^{(1)}$ has been already achieved at state j. Subsequently, we aim for the samples $\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}$ to achieve the two conditions in Definition 1 at a state pair (i, j) for the first time.

Formally, we first construct a set $\mathcal{D}_j^{(1)}$ focusing only on the first j transitions Δ_k for $k = 1, \cdots, j$ of DS samples $\omega \in \Omega$ with two requirements; one is that they must have N_{BC} number of $+1$'s and $j - N_{BC}$ number of -1's; and the other is that the j-th transition Δ_j must be $+1$ to guarantee that they have never been achieved in any states prior to the state j. The former requirement implies that all $\omega \in \mathcal{D}_j^{(1)}$ hold $S_j = \sum_{k=1}^{j} \pi_{\Delta_i}(\omega) = 2N_{BC} - j$. For example, when $N_{BC} = 2$ and $j = 5$, a sequence $(+1, -1, -1, -1, +1, \cdots)$ of state transitions satisfies the first requirement, and also satisfies $S_j = 2N_{BC} - j$.

We next construct a set $\mathcal{D}_{i,j}^{(2)} \subset \Omega$ which does not care about the first j transitions Δ_k for $k = 1, \cdots, j$, but only focuses on the interim transitions Δ_m for $m = j+1, \cdots, i$. By the definition, all sequences $\omega \in \mathcal{D}_{i,j}^{(2)}$ must achieve $\mathcal{G}^{(1)}$ before the j-th state, which implies that they must hold $S_j = 2N_{BC} - j$. The rest requirement for each $\omega \in \mathcal{D}_{i,j}^{(2)}$ is that the state changes from starting state $S_j = 2N_{BC} - j$ to state $S_i = -1$, while any interim states S_k remain non-negative; i.e., $S_k \geq 0$ for each $k = j+1, \cdots, i-1$.

The sets $\mathcal{D}_j^{(1)}$ for all j are mutually exclusive as each of them represents the first satisfaction of the block confirmation condition exactly at the j-th state. For example, if $\omega \in \mathcal{D}_5^{(1)}$ then $\omega \notin \mathcal{D}_6^{(1)}$ since ω already has achieved the block confirmation at the 5-th state for the first time before reaching the 6-th state. The sets $\mathcal{D}_{i,j}^{(2)}$ for all (i, j) are also mutually exclusive for the same reason. Thus, their intersections $\mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}$ for all (i, j) are also mutually exclusive.

By Definition 4, the attack achieving time T_{DSA} can be measured if there exist index pairs (i, j) such that $\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}$. By the mutual exclusivity of $\mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}$, if there exists such a pair (i, j), it must be unique. In addition, if $\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}$, T_{DSA} equals $\pi_{T_i}(\omega)$, since the state progression time T_k is non-decreasing as k increases. As the result, T_{DSA} can be rewritten as follows,

$$T_{DSA} = \begin{cases} \pi_{T_i}(\omega), & if\ \exists (i,j) \in \mathbb{N}^2 \colon \omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}, \\ \infty, & otherwise. \end{cases} \tag{13}$$

3. The Attack Probabilities

We aim to calculate the probability distribution function (PDF) of the DSA time T_{DSA}. Using this, the success probability of DS attack with a given cut-time t_{cut} can be figured out as the probability that $T_{DSA} < t_{cut}$. Also, the expectation of attack success time can be calculated. The expected attack success time will be used in Section 4 to estimate the attack profits.

From Equation (13), the PDF of T_{DSA} requires the probabilities of two random events: one is the state progression time T_i in Equation (5); and the other is the event that a given state index i satisfies $\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}$. It has been well known that T_i follows Erlang distribution [22] given as

$$f_{T_i}(t) = \frac{\lambda_T(\lambda_T t)^{i-1} e^{-\lambda_T t}}{(i-1)!}. \tag{14}$$

We provide the probability for the latter event, i.e., $p_{DSA,i} = \Pr\left(\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}\right)$ in the following Lemma 1:

Lemma 1. *For a sample ω of random experiment $\mathrm{DS}(p_A, t_{cut}; N_{BC})$, the probability $p_{DSA,i} = \Pr\left(\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}\right)$ can be computed as*

$$p_{DSA,i} = \sum_{j=N_{BC}}^{j=2N_{BC}} \binom{j-1}{N_{BC}-1} C_{\frac{i-1}{2}-N_{BC},2N_{BC}-j} {p_A}^{\frac{i+1}{2}} {p_H}^{\frac{i-1}{2}} + \binom{i-1}{N_{BC}-1} {p_H}^{N_{BC}} {p_A}^{i-N_{BC}} \tag{15}$$

for odd $i > 2N_{BC}$, where $C_{n,m}$ is the ballot number [24] given by

$$C_{n,m} := \begin{cases} \frac{m+1}{n+m+1}\binom{2n+m}{n}, & n, m \in \mathbb{Z}^+ \cup \{0\}, \\ 0, & \text{otherwise}, \end{cases} \tag{16}$$

and for $i \leq 2N_{BC}$ and for all even-numbered i, $p_{DSA,i} = 0$.

Proof. See Appendix A. □

By taking infinite summations of $p_{DSA,i}$ in Lemma 1 for all indices $i \in \mathbb{N}$, we can compute the probability $\mathbb{P}_{DSA}$ that a DS attack will ever achieve the necessary conditions in Definition 1.

Corollary 1. *For a sample ω of random experiment $\mathrm{DS}(p_A, t_{cut}; N_{BC})$ with $t_{cut} = \infty$, the probability $\mathbb{P}_{DSA}$ has an algebraic expression*

$$\mathbb{P}_{DSA} = \begin{cases} 1, & p_H \leq p_A, \\ 1 - {p_A}^{N_{BC}+1} {p_H}^{N_{BC}} \sum_{j=N_{BC}}^{2N_{BC}} \binom{j-1}{N_{BC}-1} A_j, & p_H > p_A, \end{cases} \tag{17}$$

where

$$A_j := {p_A}^{j-2N_{BC}-1} - {p_H}^{j-2N_{BC}-1}. \tag{18}$$

Proof. See Appendix B. □

From Equation (13), the PDF of T_{DSA} follows the PDF of T_i at a given state index i, if at which it holds that $\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}$, with the probability of $p_{DSA,i}$. If there does not exist such an index i, with the probability of $1 - \mathbb{P}_{DSA}$, then $T_{DSA} = \infty$. Thus, we can write the PDF $f_{T_{DSA}}$ of T_{DSA} as follows,

$$\begin{aligned} f_{T_{DSA}}(t) = & \sum_{i=2N_{BC}+1}^{\infty} p_{DSA,i} f_{T_i}(t) \\ & + (1 - \mathbb{P}_{DSA})\delta(t-\infty), \end{aligned} \tag{19}$$

where $\delta(t)$ is the Dirac delta function.

Proposition 1. *The PDF $f_{T_{DSA}}$ has an analytic expression*:

$$f_{T_{DSA}}(t) = \frac{p_A\lambda_T e^{-\lambda_T t}\left(p_A p_H(\lambda_T t)^2\right)^{N_{BC}}}{(2N_{BC})!} \cdot \sum_{j=N_{BC}}^{j=2N_{BC}} \binom{j-1}{N_{BC}-1} {}_2F_3\left(\mathbf{a};\mathbf{b};p_A p_H(\lambda_T t)^2\right) + \frac{e^{-\lambda_T t}}{t}\frac{(p_H\lambda_T t)^{N_{BC}}}{(N_{BC}-1)!}\left(e^{p_A\lambda_T t} - \sum_{i=0}^{N_{BC}}\frac{(p_A\lambda_T t)^i}{i!}\right) + (1-\mathbb{P}_{DSA})\delta(t-\infty), \tag{20}$$

where ${}_pF_q(\mathbf{a};\mathbf{b};x)$ is the generalized hypergeometric function (See Appendix E for definition) with the parameter vectors

$$\mathbf{a} = \begin{bmatrix} N_{BC}+1-j/2 \\ N_{BC}+1/2-j/2 \end{bmatrix} \tag{21}$$

and

$$\mathbf{b} = \begin{bmatrix} 2N_{BC}+2-j \\ N_{BC}+1 \\ N_{BC}+1/2 \end{bmatrix}. \tag{22}$$

Proof. See Appendix C. □

By Definition 2, the probability $\mathbb{P}_{AS}$ that a DS attack $\mathrm{DS}(p_A, t_{cut}; N_{BC})$ succeeds equals

$$\mathbb{P}_{AS}(t_{cut}) = \Pr(T_{DSA} < t_{cut}) \tag{23}$$

Note that for a special case of $t_{cut} = \infty$, $\mathbb{P}_{AS}(t_{cut}) = \mathbb{P}_{DSA}$, which coincides with the result in Rosenfeld [6].

It will be shown to be convenient to define the attack success time T_{AS} of a DS attack as

$$T_{AS} := \begin{cases} T_{DSA}, & if\ T_{DSA} < t_{cut}, \\ \text{not defined}, & otherwise. \end{cases} \tag{24}$$

A random variable for $T_{DSA} > t_{cut}$ does not need to be defined since it is not useful. The PDF $f_{T_{AS}}$ of T_{AS} is just a scaled version of $f_{T_{DSA}}(t)$ for $0 < t < t_{cut}$, which is given in Equation (20), with a scaling factor of ${\mathbb{P}_{AS}}^{-1}$. Formally, the PDF $f_{T_{AS}}(t)$ equals

$$f_{T_{AS}}(t) = \begin{cases} \frac{f_{T_{DSA}}(t)}{\mathbb{P}_{AS}}, & for\ 0 \le t < t_{cut}, \\ 0, & for\ t \ge t_{cut}. \end{cases} \tag{25}$$

The expectation of attack success time is computed as

$$\mathbb{E}_{T_{AS}}(t_{cut}) = \frac{\int_0^{t_{cut}} t f_{T_{DSA}}(t)dt}{\mathbb{P}_{AS}(t_{cut})}. \tag{26}$$

The following Proposition 2 gives an explicit formula of $\mathbb{E}_{T_{AS}}$ for the special case when $t_{cut} = \infty$.

Proposition 2. *Let $p_M := \max(p_A, p_H)$, $p_m := \min(p_A, p_H)$. If $t_{cut} = \infty$, the expectation $\mathbb{E}_{T_{AS}}(t_{cut})$* has a closed-form expression:

$$\lim_{t_{cut}\to\infty} \mathbb{E}_{T_{AS}}(t_{cut}) = \frac{{\lambda_T}^{-1}\left(\sum_{j=N_{BC}}^{2N_{BC}} \binom{j-1}{N_{BC}-1} Z_j + \frac{N_{BC}}{p_H}\right)}{\mathbb{P}_{DSA}}, \tag{27}$$

where

$$Z_j := {p_A p_m}^{N_{BC}} {p_M}^{-(N_{BC}-j+1)} \left(\frac{2N_{BC} - 2jp_m + 1}{p_M - p_m} \right) - j{p_A}^{-(N_{BC}-j)} {p_H}^{N_{BC}}. \tag{28}$$

Proof. See Appendix B. □

4. Profitable DS Attacks

The previous probabilistic analyses in [1,6] have shown that the success of DS attacks is not guaranteed when $p_A < 0.5$. However, DS attacks with $p_A < 0.5$ can be vigorously pursued as long as they bring profit.

We analyze the profitability of DS attacks and to this end, we define a profit function P of a DS attack $\mathrm{DS}(C, p_A, t_{cut}; N_{BC})$, where C is the value of a fraudulent transaction, in terms of revenue and operating expense (OPEX) of the computing power.

The OPEX X (e.g., the rental fee for the computing power) and the block mining reward R tend to increase with respect to λ_A and the time t consumed during the attack. Thus, X and R are expressed as functions of λ_A and t, and they can be any increasing function; e.g., linear, exponential, or logarithm. We define X and R, respectively, as follows:

$$X(\lambda_A, t) := \gamma \lambda_A t \left(\log_{x_1} x_2 \right)^{\lambda_A} \left(\log_{x_3} x_4 \right)^t \tag{29}$$

for real constants $\gamma > 0$, $x_1, x_2 > 1$, and $x_3, x_4 > 1$, and

$$R(\lambda_A, t) := \beta \lambda_A t \left(\log_{r_1} r_2 \right)^{\lambda_A} \left(\log_{r_3} r_4 \right)^t \tag{30}$$

for real constants $\beta > 0$, $r_1, r_2 > 1$, and $r_3, r_4 > 1$. We denote the ratio of γ and β by

$$\mu := \beta \gamma^{-1}. \tag{31}$$

With regards to P, if an attack succeeds, the revenue comes from C, as it is double-spent, added to R for the number of blocks mined during the time duration T_{AS}, i.e., $R(\lambda_A, T_{AS})$. In this case, the cost is the OPEX for the time duration T_{AS}, i.e., $X(\lambda_A, T_{AS})$. If the attack fails, the cost is the OPEX $X(\lambda_A, t_{cut})$ for the time duration t_{cut}, and there is no revenue. Hence, for a DS attack $\mathrm{DS}(C, p_A, t_{cut}; N_{BC})$, we define P as follows,

$$P := \begin{cases} C + R(\lambda_A, T_{AS}) - X(\lambda_A, T_{AS}), & if \quad T_{DSA} < t_{cut}, \\ -X(\lambda_A, t_{cut}), & otherwise. \end{cases} \tag{32}$$

Subsequently, the expected profit function is

$$\begin{aligned} \mathbb{E}_P &= \mathbb{P}_{AS}(t_{cut}) \cdot (C + \mathbb{E}[R(\lambda_A, T_{AS})] - \mathbb{E}[X(\lambda_A, T_{AS})]) - (1 - \mathbb{P}_{AS}(t_{cut}))X(\lambda_A, t_{cut}) \\ &= \mathbb{P}_{AS}(t_{cut}) \cdot (C + \mathbb{E}[R(\lambda_A, T_{AS})]) - \mathbb{E}_X, \end{aligned} \tag{33}$$

where $\mathbb{E}_X$ is the expected OPEX defined as

$$\mathbb{E}_X := \mathbb{P}_{AS}(t_{cut}) \mathbb{E}[X(\lambda_A, T_{AS})] + (1 - \mathbb{P}_{AS}(t_{cut}))X(\lambda_A, t_{cut}). \tag{34}$$

Definition 5. *A DS attack* $\mathrm{DS}(C, p_A, t_{cut}; N_{BC})$ *is said to be profitable if and only if the expected profit* $\mathbb{E}_P > 0$ *, where* $\mathbb{E}_P$ *is defined in Equation (33).*

The key factor in determining the profitability of DS attacks is the value C of the fraudulent transaction. Thus, attackers would be interested in the minimum value required for profitable DS

attacks [25]. Definition 5 implies that a DS attack $\mathrm{DS}(C, p_A, t_{cut}; N_{BC})$ is profitable if and only if $C > C_{\mathrm{Req.}}$, where the required value of target transaction $C_{\mathrm{Req.}}$ is

$$C_{\mathrm{Req.}} = \frac{\mathbb{E}_X}{\mathbb{P}_{AS}} - \mathbb{E}[R(\lambda_A, T_{AS})]. \tag{35}$$

The following results in Theorem 1 and Theorem 2 focus on the case where both $X(\lambda_A, t)$ and $R(\lambda_A, t)$ are linearly increasing functions of λ_A and t.

Theorem 1. *Suppose $x_1 = x_2$ and $x_3 = x_4$ in Equation (29), and $r_1 = r_2$ and $r_3 = r_4$ in Equation (30). Then, a DS attack $\mathrm{DS}(C, p_A, t_{cut}; N_{BC})$ for any $p_A \in (0, 1)$ and for any $t_{cut} \in (0, \infty]$ is profitable if and only if $C > C_{\mathrm{Req.}}$, where*

$$C_{\mathrm{Req.}} = \frac{(1 - \mathbb{P}_{AS}(t_{cut}))}{\mathbb{P}_{AS}(t_{cut})} \gamma \lambda_A t_{cut} - (\mu - 1)\gamma \lambda_A \mathbb{E}_{T_{AS}}(t_{cut}). \tag{36}$$

Proof. Substituting $x_1 = x_2$, $x_3 = x_4$, $r_1 = r_2$, and $r_3 = r_4$ into Equation (35) results in Equation (36). □

Theorem 1 shows that not only superior attackers with $p_A \in (0.5, 1)$ but also inferior attackers with $p_A \in (0, 0.5)$ are able to expect profitable DS attacks once a high enough value C greater than $C_{\mathrm{Req.}}$ of the target transaction is designed. The condition $C_{\mathrm{Req.}}$ in Equation (36) can be pre-computed before carrying out an attack, as it stochastically estimates the future expected cost, for a given position $p_A \in (0, 1)$ and a cut-time t_{cut} of an attacker, and a given set of network environment parameters γ and β.

Tables 1 and 2 list the resources including $C_{\mathrm{Req.}}$, $\mathbb{E}_X$, and $\mathbb{E}_{T_{AS}}$ required for profitable DS attacks respectively using $p_A = 0.35$ and $p_A = 0.4$, when $t_{cut} = cN_{BC}\lambda_H^{-1}$ with $c = 4$. Note that the expectation of the time spent for the block confirmation equals $N_{BC}\lambda_H^{-1}$, and we let t_{cut} linear to it. In other words, as normal traders wait for $N_{BC}\lambda_H^{-1}$ seconds on the average, attackers shall be tolerable as well and wait for the same scale of time duration. Note that the $\mathbb{P}_{AS}$ for $N_{BC} = 1$ is smaller than that for $N_{BC} = 3$ due to not long enough t_{cut}. We scaled the results by parameters λ_H and γ, which we will explain how to obtain from the internet in the next subsection.

Table 1. Numerical computations of required resources for profitable double-spending (DS) attacks with $p_A = 0.35$ when $t_{cut} = cN_{BC}\lambda_H^{-1}$ with $c = 4$.

Block Confirmation Number (N_{BC})	**1**	**3**	**5**	**7**	**9**
Attack success probability ($\mathbb{P}_{AS}$)	0.315	0.279	0.218	0.170	0.132
Expected attack success time $\left(\mathbb{E}_{T_{AS}}\right)$(Scaled by ${\lambda_H}^{-1}$)	2.004	5.518	8.681	11.694	14.607
Expected OPEX ($\mathbb{E}_X$)(Scaled by γ)	1.815	5.487	9.440	13.588	17.859
Required value of target transaction ($C_{Suf.}$) (Scaled by γ)	1.079 $\cdot(1-\mu)$ $+$ 4.680	2.971 $\cdot(1-\mu)$ $+$ 16.68	4.675 $\cdot(1-\mu)$ $+$ 38.62	6.297 $\cdot(1-\mu)$ $+$ 73.84	7.866 $\cdot(1-\mu)$ $+$ 127.00

Table 2. Numerical computations of required resources for profitable DS attacks with $p_A = 0.4$ when $t_{cut} = cN_{BC}\lambda_H^{-1}$ with $c = 4$.

Block Confirmation Number (N_{BC})	**1**	**3**	**5**	**7**	**9**
Attack success probability ($\mathbb{P}_{AS}$)	0.411	0.419	0.376	0.334	0.297
Expected attack success time $\left(\mathbb{E}_{T_{AS}}\right)$(Scaled by ${\lambda_H}^{-1}$)	1.953	5.338	8.434	11.418	14.325
Expected OPEX ($\mathbb{E}_X$)(Scaled by γ)	2.106	6.139	10.436	14.977	19.716
Required value of target transaction ($C_{Suf.}$)(Scaled by γ)	1.302 $\cdot(1-\mu)$ $+ 3.819$	3.559 $\cdot(1-\mu)$ $+ 11.10$	5.622 $\cdot(1-\mu)$ $+ 22.15$	7.612 $\cdot(1-\mu)$ $+ 37.25$	9.550 $\cdot(1-\mu)$ $+ 56.96$

The following Theorem 2 is for the inferior attackers with $p_A \in (0, 0.5)$ and shows the importance of setting a finite t_{cut}.

Theorem 2. *Suppose $x_1 = x_2$ and $x_3 = x_4$ in Equation (29), and $r_1 = r_2$ and $r_3 = r_4$ in Equation (30). Then, a DS attack* $\mathrm{DS}(C, p_A, t_{cut}; N_{BC})$ *with $p_A \in (0, 0.5)$ is profitable only if $t_{cut} < \infty$.*

Proof. For any $p_A \in (0, 0.5)$, it always holds that $\mathbb{P}_{AS} < 1$. In this case, if $t_{cut} \to \infty$ then $C_{\text{Req.}} \to \infty$ from Equation (36); i.e., infinite value C of fraudulent transaction is required for a DS attack $\mathrm{DS}(C, p_A, t_{cut}; N_{BC})$ to be profitable. Thus, for a DS attack with $p_A \in (0, 0.5)$ to be profitable, a finite cut-time $t_{cut} < \infty$ must be set. □

Theorem 2 shows that for $p_A \in (0, 0.5)$, setting $t_{cut} = \infty$ is expected to incur infinite deficit. On the contrary, for $p_A \in (0.5, 1)$, what we have numerically checked out but omitted due to space limitation is the result that $\mathbb{E}_P$ is an increasing function of t_{cut}; i.e., setting $t_{cut} = \infty$ is the optimal choice in the superior attack regime. Applying $p_A \in (0.5, 1)$ and $t_{cut} = \infty$ into Equation (36) leads to $\mathbb{P}_{AS} = 1$, and thus $C_{\text{Req.}}$ turns into

$$C_{\text{Req.}} = -(\mu - 1)\gamma\lambda_A\mathbb{E}_{T_{AS}}, \tag{37}$$

where a closed-form expression of $\mathbb{E}_{T_{AS}}$ is given in Proposition 2. In this case, if $\beta > \gamma$; i.e., $\mu > 1$, DS attacks are always profitable regardless of C. According to nicehash.com, most networks maintain $\beta > \gamma$ by the economic equilibrium. As the result, in addition to the results in [1] and [6] that DS attacks with $p_A \in (0.5, 1)$ guarantee probabilistic success, we show that such attacks guarantee economic gain as well.

5. Practical Example of Profitable DS Attacks against BitcoinCash

We analyze resources required for profitable DS attacks against BitcoinCash network. The resources include the computing power proportion p_A, expected OPEX $\mathbb{E}_X$, expected attack success time $\mathbb{E}_{T_{AS}}$, and the required value of fraudulent transaction $C_{\text{Req.}}$.

To this end, we first recall the parameters involved in block mining reward R and the OPEX X. The parameters used in Equation (29) and Equation (30) are assumed to $x_1 = x_2$, $x_3 = x_4$, $r_1 = r_2$, and $r_3 = r_4$ which lead to linear functions $X(\lambda_A, t)$ and $R(\lambda_A, t)$ with respect to λ_A and t. There are three more parameters: γ, β, and ${\lambda_H}^{-1}$. From Equation (29) and Equation (30), the parameter γ is the expected cost spent per generating a block; and the parameter β is the reward per generating a block. Parameter ${\lambda_H}^{-1}$ is the average block generation time of the honest chain. All the parameters are different for each blockchain network.

In BitcoinCash, the reward β per block mining was 12.5 BCH (without transaction fees), which is around $\beta = 0.44$ BTC per block mining (as of 26 February 2020). The average block generation time was fixed at ${\lambda_H}^{-1} = 600$ seconds.

The parameter γ is obtainable from nicehash.com. BitcoinCash uses the SHA-256 cryptographic puzzle for which the unit of computation is hash. As of 26th Feb. 2020, the rental fee for 1-peta (P) hashes per second for a day was around 0.017 BTC, which was around 1.97×10^{-7} BTC per second. In other words, the rental fee was approximately 1.97×10^{-22} BTC per the computing of a hash. Referring to BTC.com, the network's computing speed is 3.57-exa (E) hashes per second; i.e., $3.57\text{E} \cdot 600 = 2142\text{E}$ hashes are needed to generate one block on the average. As the result, the parameter γ is obtained as

$$\begin{aligned} \gamma &= 1.97 \times 10^{-22}\ [\text{BTC/hash}] \times 2142\text{E}[\text{hashes/block mining}] \\ &\approx 0.422\ [\text{BTC/block mining}]. \end{aligned} \tag{38}$$

Note that it holds $\beta > \gamma$. From Equation (37), this relationship makes DS attack $\text{DS}(C, p_A, t_{cut}; N_{BC})$ with $p_A > 0.5$ and $t_{cut} = \infty$ always profitable regardless of the value C of target transaction.

In case of DS attacks with $p_A < 0.5$, the cut-time t_{cut} must be determined as a finite value for profitable DS attacks by Theorem 2. We set $t_{cut} = cN_{BC}\lambda_H^{-1} = 12000$ seconds with $c = 4$ and $p_A = 0.35$. We compute the resources required for profitable DS attacks against BitcoinCash when $N_{BC} = 5$. Results are obtainable from the values in Tables 1 and 2 by multiplying the scaling parameters $\gamma = 0.422$ and ${\lambda_H}^{-1} = 600$ and by substituting $\mu = \beta\gamma^{-1} = 1.04$ and $c = 4$.

As the results, we obtain $\mathbb{P}_{AS} \approx 0.218$, $\mathbb{E}_{T_{AS}} \approx 5200$ seconds, $\mathbb{E}_X \approx 3.98$ BTC, and $C_{\text{Req.}} \approx 16.22$ BTC. One can compute expected running time; i.e., the expected time spent for a single DS attack attempt as $\mathbb{P}_{AS}\mathbb{E}_{T_{AS}} + (1 - \mathbb{P}_{AS})t_{cut}$, which is around 2 h and 55 min. That is to say, attackers can repeatedly perform n number of attacks every 2 h and 55 min on the average. With the value C of target transaction, by the strong law of large numbers, the multiple attack attempts will return net profit $nP_{AS}(t_{cut}) \cdot \left(C - C_{\text{Req.}}\right)$ as $n \to \infty$ with probability 1.

6. Related Works

By Nakamoto [1] and Rosenfeld [6], the probabilities have been studied that a DS attack will ever succeed when there is no time limit, i.e., the cut-time is set to $t_{cut} = \infty$. Both of them applied PCPs to model the growth of chains $H(t)$ and $A(t)$. On one hand, the main difference between them was in probability calculations of the block confirmation process in Definition 1. Rosenfeld applied the PCPs to both $H(t)$ and $A(t)$, whereas Nakamoto assumed the time spent for $H(t) \geq N_{BC}$ deterministic to simplify the calculation. On the other hand, they both used the gambler's ruin approach to obtain the asymptotical behavior of S_i as $i \to \infty$ by manipulating the recurrence relationship between two adjacent states. Namely, their results are based on an assumption that an indefinite number of attack chances are given [12].

On the contrary, we introduce the cut-time t_{cut} which generalizes analytical framework to the more interesting finite attack time and inferior attacker regime. By setting t_{cut} infinite, the same result $\mathbb{P}_{DSA}$ was obtained in [6] as well. By setting a finite t_{cut}, our results shall be useful when attack chances are limited due to limited amount of resources such as time and cost. In addition, we show in Theorem 2 that DS attacks with $p_A < 0.5$ must set a finite t_{cut} in order to expect a non-negative profit. It shall be noted that there has been no intermediate result like $p_{DSA,i}$ in Lemma 1. We use Lemma 1 to derive the novel results.

Rosenfeld [6] and Bissias et al. [13] have analyzed the profitability of DS attacks. However, they put additional assumptions on the attack scenario to simplify the calculation of the attack time. Specifically, Rosenfeld assumed the attack time to be a constant. Bissias et al. assumed that the attack stops if either the normal peers or the attacker achieves the block confirmation first. On the contrary, in our model, an attack can be continued for a random attack time as long as it brings profit, even if the normal peers achieve the block confirmation before the attacker does.

In Zaghloul et al. [14], the profit of DS attack has been analyzed. Interestingly, they have discussed the need of cut-time for DS attacks with $p_A < 0.5$, which is theoretically proven in this paper in

Theorem 2. They also calculated the profit of DS attacks with a finite time-limit (see Section IV-C in [14]), but their calculation was not as precise as ours in three points:

First, the probability of attack success within a finite time-limit, i.e., $\mathbb{P}_{AS}(t_{cut})$ in Equation (23) was never considered, which requires the distribution of the DS achieving time, i.e., T_{DSA} given in Proposition 1. Instead, their calculation used $\mathbb{P}_{DSA}$ referring to the result in Rosenfeld [6]. This contradicts their time-limited attack scenario, since $\mathbb{P}_{DSA}$ in [6] was resulted from the assumption of infinite time-limit.

Second, they approximated costs and revenues of DS attack spent within a time-limit. Estimation of the costs and revenues requires estimations of the numbers of blocks respectively mined by honest nodes and attackers within a time-limit, but those were assumed to be constant. This was due to the absence of the time analysis we provide in Proposition 1.

Third, they assumed the average block generation rates λ_H, λ_A respectively by honest miners and by attackers are always the same. Since, the proportions p_H, p_A of computing power occupied by the two groups can be quite different in general, such a result is not very useful. We agree to their assumption that most blockchains control the difficulty of block mining puzzle to keep the average speed of block generation constant, and thus λ_H can be considered as a constant. However, λ_A should be left as a varying quantity by p_A. The fact is that the computing power invested by the attacker cannot be monitored by the honest network and thus it cannot be reflected in the difficulty control routine.

Budish [15] conducted simulations on the profitability of DS attacks only in the cases of $p_A > 0.5$. Under the cases, a condition on the value of the target transaction that makes DS attacks not profitable has been given based on the simulations. We give theoretical and numerically-calculable results for any $p_A \in (0, 1)$, which do not require massive simulations.

Gervais et al. [16] and Sompolinsky et al. [12] have used a Markov decision process (MDP) to analyze profits from DS attacks. These works differ from our contributions in the following regards:

First, they did not follow the DS attacks scenario considered by Nakamoto [1] and Rosenfeld [6]. Instead, the scenario in [12] was a special case of the pre-mining strategy which was introduced in [17,18]. We show that the success of DS attack under this scenario is even more difficult to occur than the success of the DS attack under the scenario of Nakamoto and Rosenfeld (see Appendix D for details). Also, the attack scenario in [16] went even further by modifying the condition for block confirmation in Definition 1. Specifically, under our definition, it is required for the honest chain to have added N_{BC} blocks, while under their condition it was fraudulent for the chain to do so (see Section 3 of [16]). Thus, it was not ensured that the potential victim has shipped the goods or service, and an attack success did not guarantee for the attacker to obtain the benefit of attacking.

Second, one important new advance in this paper is the derivation of the time analysis $f_{T_{AS}}$ given in Proposition 1. When one uses the MDP framework, one can obtain similar information such as the estimations for the attack success time $E_{T_{AS}}$, the future profit P that an attacker will earn in the end, and the minimum value of target transaction $C_{\text{Req.}}$. However, using MDP to make such estimations would have required extensive Monte Carlo simulations. Using our mathematical results, such estimations can be obtained without Monte Carlo simulations.

In addition, we believe that our mathematical results can be utilized in the MDP frameworks to improve the reliability of analyses. Conventionally, a rational user of an MDP will make a decision at every state whether to stop or to continue the process by comparing the rewards that will be incurred in the future by his/her decision. The rewards for stop actions are clear because such actions are either an attack success or a give-up. The reward for the continue action is complex because it needs to consider all the actions in all future possible states as well. In [12,16], the rewards for the continue action were over-simplified as they were evaluated only for the very next state and did not include the estimation of the profits in further future actions. To improve the reliability, the PDF $f_{T_{AS}}$ in Proposition 1 can be used at any intermediate Markov state to estimate the future profits. Specifically, the conditional expectation of the time left for an attack success T_{AS} given $T_{AS} > \tau$ can be calculated using $f_{T_{AS}}$, where τ is the observable time elapsed for reaching the current state. Once the time left is estimated, the estimation

of future profits can be updated by substituting it into Equation (33). That is to say, at each state, the estimation of profits can be updated and used as the rewards resulting from the continue action.

Goffard [19] and Karame et al. [20] have derived the PDFs of attack success time, but none of their DS attack scenarios matched with ours in Definition 1. In [19], Goffard derived the PDF of catch-up time spent for the fraudulent chain to catch up with the honest chain given that the length of honest chain is initially ahead by several blocks. The author used counting processes such as order statistic point process and renewal process which are more general than PCP, but there was no analytic result similar to what is given in Proposition 1. In [20], Karame et al. derived the PDF of the first attack success time under a fast-payment model which fixed $N_{BC} = 0$. To sum up, the attack success time in neither analysis included the time spent for achieving the first condition: the block confirmation should be realized.

7. Discussion and Conclusions

We showed that DS attacks using 50% or a lower proportion of computing power can be profitable and thus quite threatening. We provided how much quantitative resources are required to make a profitable DS attack. We derive the PDF of attack success time which enables us to figure out the operating time and the expense of mining rigs. We provided MATLAB codes on the website (https://codeocean.com/capsule/2308305/tree) for numerical evaluation of the expected profit function in Equation (33). We also listed an example of the minimum resources required for a profitable DS attack, which is applicable to any blockchain networks by substituting the network parameters γ, β, and λ_H. We also showed a more specific example of the required resources against BitcoinCash network.

Our results quantitatively guide how to set a block confirmation number for a safe transaction. The lower the block confirmation number is, the lower the minimum resource is required for a profitable attack. A solution can be utilized by the network developers to discourage such an attack. On the one hand, given a block confirmation number, we can have the value of any transaction to be set below the required value of making a profitable attack in a given network. On the other hand, given the value of transaction, the network can provide a service to inform the payee with the lowest block confirmation number that leads to negative DS attack profit.

Author Contributions: Conceptualization, J.J. and H.-N.L.; methodology, J.J.; software, J.J.; validation, J.J.; formal analysis, J.J.; investigation, J.J.; resources, J.J.; data curation, J.J.; writing—original draft preparation, J.J.; writing—review and editing, J.J. and H.-N.L.; visualization, J.J. and H.-N.L.; supervision, H.-N.L.; project administration, H.-N.L.; funding acquisition, H.-N.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by Institute of Information & Communications Technology Planning & Evaluation, grant number 2020-0-00958. This work was partially supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (NRF-2018R1A2A1A19018665).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Proof of Lemma 1. For a given sample ω and a given index i, the event $\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}$ is equivalent to the event that there exists an intermediate state index j such that $\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}$. By the mutual exclusiveness of $\mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}$ for integers j, such a state j is unique if it exists. Thus, we can write the probability $p_{DSA,i}$ as follows,

$$\begin{aligned} p_{DSA,i} &= \Pr\left(\exists j \in \mathbb{N}\colon \omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}\right) \\ &= \sum_{j=N_{BC}}^{\infty} \Pr\left(\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}\right). \end{aligned} \tag{A1}$$

Note that $\mathcal{D}_j^{(1)} \cap \left(\mathcal{D}_{i,j}^{(2)}\right) = \phi$ for $i \leq 2N_{BC}$, since the minimum number of states for an attack success is $2N_{BC}+1$: N_{BC} number of $+1$'s state transitions for the block confirmation; and $N_{BC}+1$ number of -1 's state transitions for the success of PoW competition. Thus, $p_{DSA,i} = 0$ for $i \leq 2N_{BC}$.

We further explore $\mathcal{D}_j^{(1)}$ and $\mathcal{D}_{i,j}^{(2)}$. We divide the domain of state index j in Equation (A1) into two exclusive domains; one is $j \leq 2N_{BC}$; and the other is $j > 2N_{BC}$. First, for $j \leq 2N_{BC}$, two sets $\mathcal{D}_j^{(1)}$ and $\mathcal{D}_{i,j}^{(2)}$ are independent, since their requirements on the state transitions are focusing on disjoint indices of state by their definitions. Formally, $\Pr\left(\omega \in \mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)}\right) = \Pr\left(\omega \in \mathcal{D}_j^{(1)}\right)\Pr\left(\omega \in \mathcal{D}_{i,j}^{(2)}\right)$. Second, we explore the domain $j > 2N_{BC}$. By the definition of $\mathcal{D}_j^{(1)}$, all $\omega \in \mathcal{D}_j^{(1)}$ satisfy $S_j = \sum_{k=1}^{j} \pi_{\Delta_k}(\omega) = 2N_{BC} - j$. Thus, for every $j > 2N_{BC}$, S_j is already negative, which implies all $\omega \in \mathcal{D}_j^{(1)}$ satisfy both and at state j. The set $\mathcal{D}_{i,j}^{(2)} = \phi$ for $j > 2N_{BC}$ and $j < i$, since the state $S_j = 2N_{BC} - j$ contradicts one requirement of $\mathcal{D}_{i,j}^{(2)}$: the interim transitions between the states j and i should be non-negative. For $j > 2N_{BC}$ and $j = i$, let us set $\mathcal{D}_{i,j}^{(2)} = \Omega$, since there is no interim state to apply the requirement to. To sum up, $\mathcal{D}_j^{(1)} \cap \mathcal{D}_{i,j}^{(2)} = \mathcal{D}_i^{(1)}$ for $j > 2N_{BC}$ and $i = j$, and $\mathcal{D}_j^{(1)} \cap \left(\mathcal{D}_{i,j}^{(2)}\right) = \phi$ for $j > 2N_{BC}$ and $i > j$. Subsequently, Equation (A1) is computed as

$$p_{DSA,i} = \sum_{j=N_{BC}}^{2N_{BC}} \Pr\left(\omega \in \mathcal{D}_j^{(1)}\right)\Pr\left(\omega \in \mathcal{D}_{i,j}^{(2)}\right) + \Pr\left(\omega \in \mathcal{D}_i^{(1)}\right). \tag{A2}$$

We now compute the ingredient probabilities $\Pr\left(\omega \in \mathcal{D}_j^{(1)}\right)$ and $\Pr\left(\omega \in \mathcal{D}_{i,j}^{(2)}\right)$ in Equation (A2). First, by the definition, all samples in $\mathcal{D}_j^{(1)}$ must have $N_{BC}-1$ number of $+1$'s state transitions among the first $j-1$ transitions. And the rest of the $j-1$ transitions must be valued by -1. In addition, the j-th transition must be valued by $+1$ so that the block confirmation is achieved exactly at the j-th state index. As the result, the probability $\Pr\left(\omega \in \mathcal{D}_j^{(1)}\right)$ equals the point mass function of a negative binomial distribution:

$$\Pr\left(\omega \in \mathcal{D}_j^{(1)}\right) = \binom{j-1}{N_{BC}-1} {p_H}^{N_{BC}} {p_A}^{j-N_{BC}}. \tag{A3}$$

Second, computing the probability $\Pr\left(\omega \in \mathcal{D}_{i,j}^{(2)}\right)$ starts from counting the number of combinations of state transitions satisfying the requirements of set $\mathcal{D}_{i,j}^{(2)}$. Recall the requirements on every element of $\mathcal{D}_{i,j}^{(2)}$, for $j = N_{BC}, \cdots, 2N_{BC}$, are that the state starts from the state $S_j = 2N_{BC} - j$ and ends at the state $S_i = -1$ while all the $i-j-1$ number of interim states remain nonnegative. The i-th transition should be $\Delta_i = -1$ so that the success of PoW competition is achieved exactly at the state index i. The number of combinations of such state transitions can be counted using the ballot number $C_{n,m}$ [24], which is the number of random walks that consist of $2n+m$ steps and never become negative, starting from the origin and ending at the point m. In our problem, the number of random walk steps is $2n+m = i-j-1$ with $m = 2N_{BC} - j$. As a result, by multiplying the probabilities p_A and p_H for state transitions, the probability $\Pr\left(\omega \in \mathcal{D}_{i,j}^{(2)}\right)$ is computed as

$$\Pr\left(\omega \in \mathcal{D}_{i,j}^{(2)}\right) = C_{n,m} {p_A}^{(n+m+1)} {p_H}^{n}, \tag{A4}$$

where $2n+m = i-j-1$ and $m = 2N_{BC} - j$.

Finally, substituting Equations (A3) and (A4) into Equation (A2) results in Equation (15). □

Appendix B

Proof of Corollary 1. Taking infinite summations of $p_{DSA,i}$ for all indices i results in $\mathbb{P}_{DSA}$:

$$\mathbb{P}_{DSA} = \sum_{i=2N_{BC}+1}^{\infty} p_{DSA,i} \tag{A5}$$

By substituting $p_{DSA,i}$ in Lemma 1 into Equation (A5), the probability $\mathbb{P}_{DSA}$ becomes

$$\mathbb{P}_{DSA} = \sum_{j=N_{BC}}^{2N_{BC}} \binom{j-1}{N_{BC}-1} p_A \sum_{i=2N_{BC}+1}^{\infty} C_{\frac{i-1}{2}-N_{BC},2N_{BC}-j} (p_A p_H)^{\frac{i-1}{2}} + \left(\frac{p_H}{p_A}\right)^{N_{BC}} \sum_{i=2N_{BC}+1}^{\infty} \binom{i-1}{N_{BC}-1} {p_A}^i. \tag{A6}$$

By rearranging the indices i in the summations, we can obtain

$$\begin{aligned} \mathbb{P}_{DSA} &= \sum_{j=N_{BC}}^{2N_{BC}} \binom{j-1}{N_{BC}-1} p_A \sum_{i=0}^{\infty} C_{i,2N_{BC}-j} (p_A p_H)^{i+N_{BC}} \\ &+ \left(\frac{p_H}{p_A}\right)^{N_{BC}} \left(\sum_{i=N_{BC}}^{\infty} \binom{i-1}{N_{BC}-1} {p_A}^i - \sum_{i=N_{BC}}^{2N_{BC}} \binom{i-1}{N_{BC}-1} {p_A}^i \right). \end{aligned} \tag{A7}$$

We define two generating functions as

$$M_k(x) := \sum_{i=0}^{\infty} C_{i,k} x^i, \tag{A8}$$

and

$$G_k(x) := \sum_{i=k}^{\infty} \binom{i}{k} x^i. \tag{A9}$$

By substituting M_k and G_k into Equation (A7), we can write

$$\begin{aligned} \mathbb{P}_{DSA} &= \sum_{j=N_{BC}}^{2N_{BC}} \binom{j-1}{N_{BC}-1} p_A (p_A p_H)^{N_{BC}} M_{2N_{BC}-j}(p_A p_H) \\ &+ \left(\frac{p_H}{p_A}\right)^{N_{BC}} \left(p_A G_{N_{BC}-1}(p_A) - \sum_{i=N_{BC}}^{2N_{BC}} \binom{i-1}{N_{BC}-1} {p_A}^i \right) \end{aligned} \tag{A10}$$

The function $M_k(x)$ is a generating function of the ballot numbers $C_{i,k}$, for which the algebraic expression given in [26] is

$$M_k(x) = \left(\frac{2}{1+\sqrt{1-4x}} \right)^{k+1}. \tag{A11}$$

Putting $x = p_A p_H$ into $M_k(x)$ results in

$$\begin{aligned} M_k(p_A p_H) &= \left(\frac{2}{1+\sqrt{1-4p_A p_H}} \right)^{k+1} \\ &= \begin{cases} \left(\frac{2}{1+\sqrt{1-4p_A(1-p_A)}} \right)^{k+1}, & if\ p_A < p_H, \\ \left(\frac{2}{1+\sqrt{1-4(1-p_H)p_H}} \right)^{k+1}, & if\ p_A \geq p_H \end{cases} \\ &= \left(\frac{1}{p_M} \right)^{k+1}, \end{aligned} \tag{A12}$$

where $p_M := \max(p_H, p_A)$. The function $G_k(x)$ is a generating function of binomial coefficients, and the algebraic expression for it is given in [27]:

$$G_k(x) = \frac{x^k}{(1-x)^{k+1}}. \tag{A13}$$

Putting $x = p_A$ into $G_k(x)$ results in

$$G_k(p_A) = {p_H}^{-1}\left(\frac{p_A}{p_H}\right)^k. \tag{A14}$$

Substituting Equation (A12) and Equation (A14) into Equation (A10) provides

$$\mathbb{P}_{DSA} = \sum_{j=N_{BC}}^{2N_{BC}} \binom{j-1}{N_{BC}-1} p_A (p_A p_H)^{N_{BC}} {p_M}^{-(2N_{BC}-j+1)} + 1 - \left(\frac{p_H}{p_A}\right)^{N_{BC}} \sum_{i=N_{BC}}^{2N_{BC}} \binom{i-1}{N_{BC}-1} {p_A}^i. \tag{A15}$$

We define $p_m := \min(p_A, p_H)$, then the relationship $p_A p_H = p_m p_M$ holds. By rearranging the order of operands, we can obtain

$$\mathbb{P}_{DSA} = 1 - \sum_{j=N_{BC}}^{2N_{BC}} \binom{j-1}{N_{BC}-1} \left(\left(\frac{p_H}{p_A}\right)^{N_{BC}} {p_A}^j - \frac{p_A}{p_M}\left(\frac{p_m}{p_M}\right)^{N_{BC}} {p_M}^j \right), \tag{A16}$$

which is equal to Equation (17). □

Proof of Proposition 2. From Equations (19) and (26), when $t_{cut} = \infty$, we obtain

$$\begin{aligned} \mathbb{E}_{T_{AS}} &= \frac{\lim_{t_{cut}\to\infty} \int_0^{t_{cut}} t f_{T_{DSA}}(t) dt}{\mathbb{P}_{AS}(t_{cut})} = \frac{\sum_{i=2N_{BC}+1}^{\infty} \mathbb{E}[T_i] p_{DSA,i}}{\mathbb{P}_{DSA}} \\ &= \frac{\sum_{i=2N_{BC}+1}^{\infty} \frac{i}{\lambda_T} p_{DSA,i}}{\mathbb{P}_{DSA}}, \end{aligned} \tag{A17}$$

where $E[T_i] = i{\lambda_T}^{-1}$ [22]. By substituting $P_{DsA,i}$ in Equation (15) into Equation (A17) and rearranging the order of operands, we obtain

$$\begin{aligned} \lambda_T \mathbb{P}_{DSA} \mathbb{E}_{T_{AS}} &= \sum_{j=N_{BC}}^{2N_{BC}} \binom{j-1}{N_{BC}-1} \sum_{i=2N_{BC}}^{\infty} (i+1) C_{\frac{i}{2}-N_{BC}, 2N_{BC}-j} {p_A}^{\frac{i+2}{2}} {p_H}^{\frac{i}{2}} \\ &+ \sum_{i=N_{BC}-1}^{\infty} (i+1) \binom{i}{N_{BC}-1} {p_A}^{i+1-N_{BC}} {p_H}^{N_{BC}} - \sum_{i=N_{BC}-1}^{2N_{BC}-1} (i+1) \binom{i}{N_{BC}-1} {p_A}^{i+1-N_{BC}} {p_H}^{N_{BC}}. \end{aligned} \tag{A18}$$

By rearranging the indices of summations, we arrive at

$$\begin{aligned} \lambda_T \mathbb{P}_{DSA} \mathbb{E}_{T_{AS}} &= \sum_{j=N_{BC}}^{2N_{BC}} \binom{j-1}{N_{BC}-1} {p_A}^{N_{BC}+1} {p_H}^{N_{BC}} \cdot \sum_{i=0}^{\infty} (2i + 2N_{BC} + 1) C_{i,2N_{BC}-j} (p_A p_H)^i \\ &+ p_A \left(\frac{p_H}{p_A}\right)^{N_{BC}} \sum_{i=N_{BC}-1}^{\infty} (i+1) \binom{i}{N_{BC}-1} {p_A}^i - \sum_{i=N_{BC}}^{2N_{BC}} i \binom{i-1}{N_{BC}-1} {p_A}^{i-N_{BC}} {p_H}^{N_{BC}}. \end{aligned} \tag{A19}$$

By substituting the generating functions $M_k(x)$ and $G_k(x)$ defined respectively in Equation (A8) and Equation (A9), Equation (A19) becomes

$$\begin{aligned}\lambda_T \mathbb{P}_{DSA} \mathbb{E}_{T_{AS}} &= \sum_{j=N_{BC}}^{2N_{BC}} \binom{j-1}{N_{BC}-1} {p_A}^{N_{BC}+1} {p_H}^{N_{BC}} \cdot \Big(2 \sum_{i=0}^{\infty} i C_{i,2N_{BC}-j} (p_A p_H)^i + (2N_{BC}+1) M_{2N_{BC}-j}(p_A p_H)\Big) \\ &+ p_A \Big(\frac{p_H}{p_A}\Big)^{N_{BC}} \Big(\sum_{i=N_{BC}-1}^{\infty} i \binom{i}{N_{BC}-1} {p_A}^i + G_{N_{BC}-1}(p_A)\Big) - \sum_{i=N_{BC}}^{2N_{BC}} i \binom{i-1}{N_{BC}-1} {p_A}^{i-N_{BC}} {p_H}^{N_{BC}}.\end{aligned} \tag{A20}$$

We use the following relationships,

$$\sum_{i=0}^{\infty} i C_{i,k} x^i = x {M'}_k(x) \tag{A21}$$

and

$$\sum_{i=k}^{\infty} i \binom{i}{k} x^i = x {G'}_k(x), \tag{A22}$$

and their derivatives are given by

$$\begin{aligned} {M'}_k(x) &:= \tfrac{d}{dx} M_k(x) = \sum_{i=0}^{\infty} i C_{i,k} x^{i-1} \\ &= \tfrac{(k+1)}{\sqrt{1-4x}} \Big(\tfrac{2}{1+\sqrt{1-4x}} \Big)^{k+2} \end{aligned} \tag{A23}$$

and

$$\begin{aligned} {G'}_k(x) : \quad &= \tfrac{d}{dx} G_k(x) \\ &= \sum_{i=k}^{\infty} i \binom{i}{k} x^{i-1} \\ &= \tfrac{\left(k x^{k-1} + x^k\right)}{(1-x)^{k+2}}. \end{aligned} \tag{A24}$$

By substituting Equation (A21) and Equation (A22) into Equation (A20), we obtain

$$\begin{aligned}\lambda_T \mathbb{P}_{DSA} \mathbb{E}_{T_{AS}} = &\sum_{j=N_{BC}}^{2N_{BC}} \binom{j-1}{N_{BC}-1} {p_A}^{N_{BC}+1} {p_H}^{N_{BC}} \cdot \Big(2 p_A p_H {M'}_{2N_{BC}-j}(p_A p_H) + (2N_{BC}+1) M_{2N_{BC}-j}(p_A p_H)\Big) \\ &+ p_A \Big(\frac{p_H}{p_A}\Big)^{N_{BC}} \Big(p_A {G'}_{N_{BC}-1}(p_A) + G_{N_{BC}-1}(p_A)\Big) - \sum_{i=N_{BC}}^{2N_{BC}} i \binom{i-1}{N_{BC}-1} {p_A}^{i-N_{BC}} {p_H}^{N_{BC}}\end{aligned} \tag{A25}$$

Putting $x = p_A p_H$ into ${M'}_k(x)$ in Equation (A23) results in

$${M'}_k(p_A p_H) = {M'}_k(p_m p_M) = \frac{(k+1)}{1-2p_m} \Big(\frac{1}{p_M} \Big)^{k+2}. \tag{A26}$$

Putting $x = p_A$ into ${G'}_k(x)$ in Equation (A24) gives

$${G'}_k(p_A) = \frac{\left(k {p_A}^{k-1} + {p_A}^k\right)}{{p_H}^{k+2}}. \tag{A27}$$

By substituting Equation (A12), Equation (A14), Equation (A26), and Equation (A27) into Equation (A25), we finally obtain Equation (27). □

Appendix C

Proof of Proposition 1. We use a generating function and generalized hypergeometric functions to compute the infinite summations in Equation (19).

By substituting $P_{DsA,i}$ in Equation (15) and $f_{T_i}(t)$ in Equation (14) into Equation (19), we arrive at

$$\begin{aligned} f_{T_{DSA}}(t) - (1 - \mathbb{P}_{DSA})\delta(t-\infty) &= \sum_{j=N_{BC}}^{j=2N_{BC}} \binom{j-1}{N_{BC}-1} \sum_{i=2N_{BC}+1}^{\infty} C_{\frac{i-1}{2}-N_{BC},2N_{BC}-j} {p_A}^{\frac{i+1}{2}} {p_H}^{\frac{i-1}{2}} \frac{\lambda_T{}^i t^{i-1} e^{-\lambda_T t}}{(i-1)!} \\ &\quad + \sum_{i=2N_{BC}+1}^{\infty} \binom{i-1}{N_{BC}-1} {p_H}^{N_{BC}} {p_A}^{i-N_{BC}} \frac{\lambda_T{}^i t^{i-1} e^{-\lambda_T t}}{(i-1)!}. \end{aligned} \tag{A28}$$

By rearranging the indices of summations and the order of operands, we obtain

$$\begin{aligned} f_{T_{DSA}}(t) - (1 - \mathbb{P}_{DSA})\delta(t-\infty) &= \sum_{j=N_{BC}}^{j=2N_{BC}} \binom{j-1}{N_{BC}-1} \sum_{i=0}^{\infty} \left(C_{i,2N_{BC}-j} {p_A}^{N_{BC}+i+1} {p_H}^{N_{BC}+i} \cdot \frac{\lambda_T{}^{2N_{BC}+2i+1} t^{2N_{BC}+2i} e^{-\lambda_T t}}{(2N_{BC}+2i)!} \right) \\ &\quad + \left(\frac{p_H}{p_A}\right)^{N_{BC}} e^{-\lambda_T t} \left(\sum_{i=N_{BC}}^{\infty} \binom{i-1}{N_{BC}-1} {p_A}^i \frac{\lambda_T{}^i t^{i-1}}{(i-1)!} - \sum_{i=N_{BC}}^{2N_{BC}} \binom{i-1}{N_{BC}-1} {p_A}^i \frac{\lambda_T{}^i t^{i-1}}{(i-1)!} \right). \end{aligned} \tag{A29}$$

We can define two generating functions as

$$B(x) := \sum_{i=0}^{\infty} C_{i,2N_{BC}-j} \frac{x^i}{(2N_{BC}+2i)!} = (2N_{BC}-j+1) \sum_{i=0}^{\infty} \frac{(2i+2N_{BC}-j)!}{i!(i+2N_{BC}-j+1)!} \frac{x^i}{(2N_{BC}+2i)!}, \tag{A30}$$

and

$$H(x) := \sum_{i=N_{BC}}^{\infty} \binom{i-1}{N_{BC}-1} \frac{x^{i-1}}{(i-1)!} = \sum_{i=N_{BC}-1}^{\infty} \binom{i}{N_{BC}-1} \frac{x^i}{i!}. \tag{A31}$$

By substituting $B(x)$ and $H(x)$ into Equation (A29), we obtain

$$\begin{aligned} f_{T_{DSA}}(t) - (1 - \mathbb{P}_{DSA})\delta(t-\infty) &= \sum_{j=N_{BC}}^{j=2N_{BC}} \binom{j-1}{N_{BC}-1} p_A \lambda_T e^{-\lambda_T t} \left(p_A p_H (\lambda_T t)^2\right)^{N_{BC}} B\left(p_A p_H (\lambda_T t)^2\right) \\ &\quad + \left(\frac{p_H}{p_A}\right)^{N_{BC}} e^{-\lambda_T t} \left(p_A \lambda_T H(p_A \lambda_T t) - \sum_{i=N_{BC}}^{2N_{BC}} \binom{i-1}{N_{BC}-1} {p_A}^i \frac{\lambda_T{}^i t^{i-1}}{(i-1)!} \right). \end{aligned} \tag{A32}$$

We replace function $B(x)$ in Equation (A30) with the generalized hypergeometric functions (See Appendix E for definition). For this purpose, we first denote the sequences in $B(x)$ by

$$\beta_i := \frac{(2i+2N_{BC}-j)!}{i!(i+2N_{BC}-j+1)!} \frac{1}{(2N_{BC}+2i)!}, \tag{A33}$$

and

$$\beta_0 := \frac{1}{(2N_{BC}-j+1)(2N_{BC})!}. \tag{A34}$$

Next, the function $B(x)$ can be rewritten as

$$B(x) = (2N_{BC}-j+1) \sum_{i=0}^{\infty} \beta_i x^i = (2N_{BC}-j+1)\beta_0 \left(x^0 + \frac{\beta_1}{\beta_0} x^1 + \frac{\beta_2}{\beta_1}\frac{\beta_1}{\beta_0} x^2 + \cdots \right). \tag{A35}$$

The reformulated sequence in Equation (A35) is computed by

$$\frac{\beta_{i+1}}{\beta_i} = \frac{(i+1+N_{BC}-j/2)(i+1/2+N_{BC}-j/2)}{(i+2+2N_{BC}-j)(i+1+N_{BC})(i+1/2+N_{BC})(i+1)}, \tag{A36}$$

which has 2 polynomials in i on the numerator and 3 polynomials in i except for $(i+1)$ on the denominator. $B(x)$ can be expressed in terms of a generalized hypergeometric function ${}_2F_3$ [28] as follows,

$$\begin{aligned} B(x) &= (2N_{BC}-j+1)\beta_0 \, {}_2F_3\left(\mathbf{a}_j; \mathbf{b}_j; x\right) \\ &= \tfrac{1}{(2N_{BC})!} \, {}_2F_3\left(\mathbf{a}_j; \mathbf{b}_j; x\right), \end{aligned} \tag{A37}$$

where vectors $\mathbf{a}_j$ and $\mathbf{b}_j$ respectively defined in Equations (21) and (22) are the constants in the polynomials in i of the numerator and denominator in Equation (A31), respectively.

We use a closed-form expression of generating function $H(x)$ in Equation (A31) given by

$$\begin{aligned} H(x) &= \sum_{i=N_{BC}-1}^{\infty} \binom{i}{N_{BC}-1} \frac{x^i}{i!} = \frac{1}{(N_{BC}-1)!} \sum_{i=N_{BC}-1}^{\infty} \frac{x^i}{(i-N_{BC}+1)!} \\ &= \frac{x^{N_{BC}-1}}{(N_{BC}-1)!} e^x, \end{aligned} \tag{A38}$$

where the following relationship is used [29]:

$$\sum_{i=0}^{\infty} \frac{x^i}{i!} = e^x. \tag{A39}$$

By substituting Equation (A37) and Equation (A38) into Equation (A32), we arrive at

$$\begin{aligned} & f_{T_{DSA}}(t) - (1 - \mathbb{P}_{DSA})\delta(t-\infty) \\ &= \frac{p_A \lambda_T e^{-\lambda_T t} \left(p_A p_H (\lambda_T t)^2\right)^{N_{BC}}}{(2N_{BC})!} \cdot \sum_{j=N_{BC}}^{j=2N_{BC}} \binom{j-1}{N_{BC}-1}_2 F_3\left(\mathbf{a}_j; \mathbf{b}_j; p_A p_H (\lambda_T t)^2\right) \\ &+ \left(\frac{p_H}{p_A}\right)^{N_{BC}} e^{-\lambda_T t} \left(p_A \lambda_T \frac{(p_A \lambda_T t)^{N_{BC}-1}}{(N_{BC}-1)!} e^{p_A \lambda_T t} - \sum_{i=N_{BC}}^{2N_{BC}} \binom{i-1}{N_{BC}-1} {p_A}^i \frac{{\lambda_T}^i t^{i-1}}{(i-1)!} \right) \\ &= \frac{p_A \lambda_T e^{-\lambda_T t} \left(p_A p_H (\lambda_T t)^2\right)^{N_{BC}}}{(2N_{BC})!} \cdot \sum_{j=N_{BC}}^{j=2N_{BC}} \binom{j-1}{N_{BC}-1}_2 F_3\left(\mathbf{a}_j; \mathbf{b}_j; p_A p_H (\lambda_T t)^2\right) \\ &+ \left(\frac{p_H}{p_A}\right)^{N_{BC}} e^{-\lambda_T t} \left(p_A \lambda_T \frac{(p_A \lambda_T t)^{N_{BC}-1}}{(N_{BC}-1)!} e^{p_A \lambda_T t} - \frac{1}{(N_{BC}-1)!} \sum_{i=N_{BC}}^{2N_{BC}} {p_A}^i \frac{{\lambda_T}^i t^{i-1}}{(i-N_{BC})!} \right). \end{aligned} \tag{A40}$$

We obtain Equation (20) by rearranging the indices of the summations and the order of operands. □

Appendix D

Comparison of Attack Success Probabilities of DS Attack and Pre-Mining Attack

In [12], a special case of pre-mining strategy has been considered, where the condition for a DS attack success was different from Definition 1. Specifically, the only condition was to have the fraudulent chain to grow longer than the honest chain by N_{BC}, i.e., $A(t) > H(t) + N_{BC}$ (see Section 7 of [12]). We refer to $\mathbb{P}_{\text{pre-mine}}$ as the probability of satisfying this condition. The literature has shown that satisfying this condition suffices a success of DS attack [12]. What they have not shown, however, is that this condition is not a necessary one. Thus, we here aim to show that their condition is indeed not a necessary condition, by showing that $\mathbb{P}_{DSA} > \mathbb{P}_{\text{pre-mine}}$ for all $p_A \in (0, 0.5)$. First, it has been given that $\mathbb{P}_{\text{pre-mine}} = (p_A/p_H)^{N_{BC}+1}$. Under the condition of [12], it is required that the fraudulent chain catches up with the honest chain with additional N_{BC} blocks. The catch-up probability has been derived by Nakamoto in [1] using the gambler's ruin approach as $(p_A/p_H)^k$, where k is the number of blocks that the honest chain leads by at the initial status. Next, we refer to an intermediate step in the derivation of $\mathbb{P}_{DSA}$ by Rosenfeld [6]:

$$\mathbb{P}_{DSA} = \sum_{k=0}^{N_{BC}+1} \binom{N_{BC}+k-1}{k} {p_H}^{N_{BC}} {p_A}^k \left(\frac{p_A}{p_H}\right)^{N_{BC}-k+1} + \sum_{k=N_{BC}+2}^{\infty} \binom{N_{BC}+k-1}{k} {p_H}^{N_{BC}} {p_A}^k. \tag{A41}$$

Finally, clear inequalities can be used to show $\mathbb{P}_{DSA} > \mathbb{P}_{\text{pre-mine}}$:

$$\begin{aligned}\mathbb{P}_{DSA} &> \sum_{k=0}^{N_{BC}+1}\binom{N_{BC}+k-1}{k}{p_H}^{N_{BC}}{p_A}^{k}\left(\frac{p_A}{p_H}\right)^{N_{BC}-k+1} + \sum_{k=N_{BC}+2}^{\infty}\binom{N_{BC}+k-1}{k}{p_H}^{N_{BC}}{p_A}^{k}\left(\frac{p_A}{p_H}\right)^{N_{BC}+1} \\ &> \left(\frac{p_A}{p_H}\right)^{N_{BC}+1}\sum_{k=0}^{\infty}\binom{N_{BC}+k-1}{k}{p_H}^{N_{BC}}{p_A}^{k} \\ &= \left(\frac{p_A}{p_H}\right)^{N_{BC}+1} = \mathbb{P}_{\text{pre-mine}}.\end{aligned} \tag{A42}$$

For numerical example, when $p_A = 0.35$ and $N_{BC} = 5$ the probabilities can be computed as $\mathbb{P}_{DSA} = 0.2287$ and $\mathbb{P}_{\text{pre-mine}} = 0.0244$. As the gap is significant, it is shown that the DS attack success condition defined in [12] was indeed only a sufficient condition, set to be too strict.

Appendix E

Generalized Hypergeometric Function

We define generalized hypergeometric series and generalized hypergeometric functions [28].

For a variable z and a given set of coefficients $\beta_0, \cdots, \beta_\infty$, if the ratio of coefficients b_n can be expressed in terms of two polynomials $A(n)$ and $B(n)$ in n as follows,

$$\frac{\beta_{n+1}}{\beta_n} = \frac{A(n)}{B(n)(n+1)} \tag{A43}$$

for all integer $n \geq 0$, a power series $\sum_{n\geq 0}\beta_n z^n$ is a generalized hypergeometric series, where the polynomials are in the forms of

$$A(n) = c(a_1 + n)\cdots\left(a_p + n\right) \tag{A44}$$

and

$$B(n) = d(b_1 + n)\cdots\left(b_q + n\right), \tag{A45}$$

for real numbers c and d and complex numbers $a_1, \cdots, a_p$ and $b_1, \cdots, b_q$. The generalized hypergeometric series is denoted by

$${}_pF_q(\mathbf{a};\mathbf{b};z) := \sum_{n\geq 0}\beta_n z^n, \tag{A46}$$

where $\mathbf{a}$ and $\mathbf{b}$ are the vectors of $a_1, \cdots, a_p$ and $b_1, \cdots, b_q$, respectively.

A generalized hypergeomteric series can be a generalized hypergeometric function, if it converges. If $p < q + 1$, the ratio Equation (A43) goes to zero as $n \to \infty$. This implies the series Equation (A46) converges for any finite value z and thus can be defined as a function.

References

1. Nakamoto, S. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008. Available online: https://bitcoin.org/bitcoin.pdf (accessed on 26 November 2020).
2. Ritzdorf, H.; Soriente, C.; Karame, G.O.; Marinovic, S.; Gruber, D.; Capkun, S. Toward Shared Ownership in the Cloud. *IEEE Trans. Inf. Forensics Secur.* **2018**, *13*, 3019–3034. [CrossRef]
3. Wu, S.; Chen, Y.; Wang, Q.; Li, M.; Wang, C.; Luo, X. CReam: A Smart Contract Enabled Collusion-Resistant e-Auction. *IEEE Trans. Inf. Forensics Secur.* **2018**, *14*, 1687–1701. [CrossRef]
4. Nguyen, G.-T.; Kim, K. A Survey about Consensus Algorithms Used in Blockchain. *J. Inf. Process. Syst.* **2018**, *14*, 101–128. [CrossRef]
5. Sompolinsky, Y.; Zohar, A. *Secure High-Rate Transaction Processing in Bitcoin*; Böhme, R., Okamoto, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2015; pp. 507–527.
6. Rosenfeld, M. Analysis of Hashrate-Based Double Spending. *arXiv* **2014**, arXiv:1402.2009 [cs].
7. Beikverdi, A.; Song, J. Trend of centralization in Bitcoin's distributed network. In Proceedings of the 2015 IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), Takamatsu, Japan, 1–3 June 2015; pp. 1–6.

8. Gervais, A.; Karame, G.O.; Capkun, V.; Capkun, S. Is Bitcoin a Decentralized Currency? *IEEE Secur. Priv.* **2014**, *12*, 54–60. [CrossRef]
9. Attah, E. Five most prolific 51% attacks in crypto: Verge, Ethereum Classic, Bitcoin Gold, Feathercoin, Vertcoin. CryptoSlate. Available online: https://cryptoslate.com/prolific-51-attacks-crypto-verge-ethereum-classic-bitcoin-gold-feathercoin-vertcoin/ (accessed on 26 November 2020).
10. Bonneau, J. *Why Buy When You Can Rent? Bribery Attacks on Bitcoin Consensus*; Springer: Berlin, Germany, 2016.
11. Sayeed, S.; Marco-Gisbert, H. Assessing Blockchain Consensus and Security Mechanisms against the 51% Attack. *Appl. Sci.* **2019**, *9*, 1788. [CrossRef]
12. Sompolinsky, Y.; Zohar, A. Bitcoin's Security Model Revisited. *arXiv* **2016**, arXiv:1605.09193 [cs].
13. Bissias, G.; Levine, B.N.; Ozisik, A.P.; Andresen, G. An Analysis of Attacks on Blockchain Consensus. *arXiv* **2016**, arXiv:1610.07985 [cs].
14. Zaghloul, E.; Li, T.; Mutka, M.W.; Ren, J. Bitcoin and Blockchain: Security and Privacy. *IEEE Internet Things J.* **2020**, *7*, 10288–10313. [CrossRef]
15. Budish, E.B. The Economic Limits of Bitcoin and the Blockchain. *SSRN J.* **2018**. [CrossRef]
16. Gervais, A.; Karame, G.O.; Wüst, K.; Glykantzis, V.; Ritzdorf, H.; Capkun, S. On the Security and Performance of Proof of Work Blockchains. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security—CCS'16, Vienna, Austria, 24–28 October 2016; pp. 3–16.
17. Ramezan, G.; Leung, C.; Jane Wang, Z. A Strong Adaptive, Strategic Double-Spending Attack on Blockchains. In Proceedings of the 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Halifax, NS, Canada, 30 July–3 August 2018; pp. 1219–1227.
18. Pinzón, C.; Rocha, C. Double-spend Attack Models with Time Advantange for Bitcoin. *Electron. Notes Theor. Comput. Sci.* **2016**, *329*, 79–103. [CrossRef]
19. Goffard, P.-O. Fraud risk assessment within blockchain transactions. *Adv. Appl. Probab.* **2019**, *51*, 443–467. [CrossRef]
20. Karame, G.O.; Androulaki, E.; Roeschlin, M.; Gervais, A.; Čapkun, S. Misbehavior in Bitcoin: A Study of Double-Spending and Accountability. *ACM Trans. Inf. Syst. Secur.* **2015**, *18*, 2:1–2:32. [CrossRef]
21. Karame, G.O.; Androulaki, E.; Capkun, S. Double-spending fast payments in bitcoin. In Proceedings of the 2012 ACM Conference on Computer and Communications Security—CCS '12, Raleigh, NC, USA, 16–18 October 2012; p. 906.
22. Papoulis, A.; Pillai, S.U. Random walks and other applications. In *Probability, Random Variables and Stochastic Processes*; McGraw-Hill Europe: Boston, MA, USA, 2002; ISBN 978-0-07-122661-5.
23. Bowden, R.; Keeler, H.P.; Krzesinski, A.E.; Taylor, P.G. Block arrivals in the Bitcoin blockchain. *arXiv* **2018**, arXiv:1801.07447 [cs].
24. Flajolet, P.; Sedgewick, R. Combinatorial structures and ordinary generating functions. In *Analytic Combinatorics*; Cambridge University Press: Cambridge, UK, 2009; ISBN 978-1-139-47716-1.
25. Conti, M.; Sandeep Kumar, E.; Lal, C.; Ruj, S. A Survey on Security and Privacy Issues of Bitcoin. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 3416–3452. [CrossRef]
26. Wilf, H.S. Analytic and asymptotic methods. In *Generatingfunctionology*, 3rd ed.; A K Peters/CRC Press: Wellesley, MA, USA, 2005; ISBN 978-1-56881-279-3.
27. Wilf, H.S. Introductory ideas and examples. In *Generatingfunctionology*, 3rd ed.; A K Peters/CRC Press: Wellesley, MA, USA, 2005; ISBN 978-1-56881-279-3.
28. Gasper, G.; Rahman, M. Basic Hypergeometric series. In *Basic Hypergeometric Series*; Encyclopedia of Mathematics and Its Applications; Cambridge University Press: Cambridge, UK, 2004; Volume 96, ISBN 978-0-521-83357-8.
29. Flajolet, P.; Sedgewick, R. Labelled structures and exponential generating functions. In *Analytic Combinatorics*; Cambridge University Press: Cambridge, UK, 2009; ISBN 978-1-139-47716-1.

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

symmetry

MDPI

Article

ECCPoW: Error-Correction Code based Proof-of-Work for ASIC Resistance

Hyunjun Jung [1] and Heung-No Lee [2,*]

[1] Blockchain Internet Economy Research Center, Gwangju Institute of Science and Technology, Gwangju 61005, Korea; junghj85@gist.ac.kr

[2] School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, Korea

* Correspondence: heungno@gist.ac.kr; Tel.: +82-62-715-2237

Received: 20 May 2020; Accepted: 5 June 2020; Published: 9 June 2020



Abstract: Bitcoin is the first cryptocurrency to participate in a network and receive compensation for online remittance and mining without any intervention from a third party, such as financial institutions. Bitcoin mining is done through proof of work (PoW). Given its characteristics, the higher hash rate results in a higher probability of mining, leading to the emergence of a mining pool, called a mining organization. Unlike central processing units or graphics processing units, high-cost application-specific integrated circuit miners have emerged with performance efficiency. The problem is that the obtained hash rate exposes Bitcoin's mining monopoly and causes the risk of a double-payment attack. To solve this problem, we propose the error-correction code PoW (ECCPoW), combining the low-density parity-check decoder and hash function. The ECCPoW contributes to the phenomenon of symmetry in the proof of work (PoW) blockchain. This paper proposes the implementation of ECCPoW, replacing the PoW in Bitcoin. Finally, we compare the mining centralization, security, and scalability of ECCPoW and Bitcoin.

Keywords: error-correction codes proof-of-work (ECCPoW); proof-of-work (PoW); ECCPoW implementation; ASIC resistance

1. Introduction

We use digital signatures from third-party trust agencies to promote trust in Internet commerce. We warrant proof of data forgery using middlemen. Satoshi Nakamoto proposed an electronic money system without a middleman in a peer-to-peer (P2P) network through a Bitcoin white paper [1]. Bitcoin applies blockchain technology to electronic monetary systems to guarantee transactions without intermediaries (e.g., banks). Blockchain is a ledger management technology based on a distributed computing technology that cannot be arbitrarily modified by storing the transaction content in a chain-based distributed data storage environment in the form of a block [2,3].

The blockchain stores the same ledger on a global network and is designed to pay certain rewards to maintain the block. This is called mining, and mining creates blocks and obtains cryptocurrency by executing a hash function. Miners belong to mining pools because of the probability and convenience of being rewarded for mining cryptocurrency [4,5].

The hash rate is the number of hash values calculated per second as a measure of computational processing power for mining cryptocurrency. The hash rate of cryptocurrency is determined by the total number of the participating nodes. Most miners belong to a hash pool and occupy a high percentage of the hash rate. We should be concerned about the risk of a double-payment attack if the mining pool in the

blockchain accounts for a high percentage of the total hash rate of cryptocurrency [6]. A double-payment attack occurs when the mining pool seizes at least 51% of the total hash rate to determine the branch of the blockchain to the desired side. Recent studies have shown that a double-payment attack can be made to benefit from a low share of the hash rate [7].

Bitcoin miners receive blockchain information (version, previous block hash, Merkle root, bits, and others) and execute hash functions (e.g., the secure hash algorithm (SHA 256). Miners create a block if the output value of the hash function is less than the target level of difficulty. Currently, Bitcoin requires an increasingly high operation to create blocks. Miners purchase application-specific integrated circuits (ASICs) with high per-second computing power to mine Bitcoin and join the mining pool. Low-power miners using central or graphics processing units (CPUs or GPUs) have difficulty mining.

Bitcoin uses the number of zeros in the front digits of the output from the SHA 256 function to generate blocks. The mining difficulty increases with the number of zeros. Miners buy ASICs to compute SHA functions quickly. The mining machine Antminer S9 (13,000 MH/s) is approximately 8800 times faster than GTX1060 (1478 MH/s). The recently released S19 has a speed of 95 TH/s, and S19Pro exhibits performance of 110 TH/s [8]. Miners using CPUs or GPUs in Bitcoin, and miners using ASIC chips do not have equal chances for success because ASIC miners have an 8800-fold greater chance of success in mining.

Blockchain was proposed to allow nodes to participate as miners freely and to share mining rewards fairly. Blockchain is not free to participate in as a miner. The participating miners compete against equity. Several methods have been proposed to curb the development of ASIC miners, but in the end, these methods have not prevented ASIC development. As a new mining function to prevent the development of ASIC miners, we proposed the error-correction code proof of work (ECCPoW) concept, which combines the low-density parity-check (LDPC) decoder and hash function [9]. In addition, we analyzed the hash cycle of ECCPoW and demonstrated that it could be used for blockchain mining.

This paper contributes to the phenomenon of symmetry in the proof of work (PoW) blockchain. The PoW blockchain tends to increase the hash rate along with the total size of the blockchain. The size of the hash rate is related to the computing power required by block generation. Some people take issue with the large amount of wasted power used to create the PoW blockchain. However, the PoW blockchain guarantees high stability with the power required to create the block. The transition to other PoWs due to problems in the PoW blockchain causes dangerous problems. Thus, ECCPoW mitigates the power problems in the PoW blockchain and helps reduce the symmetry of the hash rate, which increases in proportion to the size of the blockchain in the PoW blockchain.

The contribution of this paper is two-fold. It introduces the proposed ECCPoW and proposes an implementation method. The second contribution is to introduce the process of experimenting in Bitcoin by replacing the SHA 256 function with the ECCPoW function. This paper proposes the creation of a cryptographic puzzle that changes every block and shows how to apply the crypto puzzle decoder to the solution. We present the implementation of the proposed method by replacing ECCPoW in Bitcoin. We also measure the block generation time of the ECCPoW. Finally, we compare ECCPoW and Bitcoin by implementing them in the same environment.

This study contributes to the previous literature on ASIC resistance in PoW blockchain. Previous studies have used forced memory access, leading to the inefficient behavior of ASICs. Another method is to make ASICs challenging to produce using various hash functions. The last method is to change the existing hash function to another hash function when ASICs are released. Therefore, ECCPoW combines the LDPC decoder with a hash function to create the effect of releasing different hash functions in each block. Implementing ASICs is difficult for the LDPC decoder due to cost issues. Furthermore, the PoW blockchain without ASICs reduces the mining centralization. In addition, ECCPoW can reduce power consumption by maintaining the advantages of the PoW.

This paper is structured as follows. Section 2 introduces the relevant research on the development of the prevention of ASIC miners. Section 3 presents ECCPoW and development methods. Section 4 reveals the results of the experiment by loading the ECCPoW into Bitcoin. Section 5 evaluates the mining centralization, security, and scalability of the ECCPoW. Finally, Section 6 presents the conclusion of the paper.

2. Related Work

2.1. Ethereum

Ethereum is a distributed computing platform developed to implement smart contract functions based on blockchain technology. In July 2015, Vitalik Buterin developed Ethereum in the C++ and Go languages. Ethereum uses the Ethash algorithm. Ethereum plans to change from a PoW to a proof of stake in the future [10]. Ether Solarium is against the use of the nonlinear directed acyclic graph (DAG) in ASICs. The initial size of the DAG was about 1 GB, and it was designed to increase in size linearly over time. In October 2019, the size of the DAG was 3.99 GB, which was maintained until December 20, 2020 [11,12].

Ethereum approved the application of programmatic proof of work (ProgPoW) [13] to respond to the centralization of mining in ASIC in 2019. First, ProgPoW regularly changes mining problems to a problem in which GPUs can adapt quickly. Second, ProgPoW makes the most of all the components of the graphics card for mining. Moreover, ProgPoW uses randomly generated problems based on block numbers and is designed for the efficient operation of GPUs. This reduces performance differences compared to ASICs.

2.2. The X-11 Series

The X-11 series is a cryptocurrency mining algorithm that uses as many hash functions as the number indicated behind the X [14]. The X-11 series used hash functions to add depth and complexity to curb ASICs. The X-11 series connects several hash functions and uses the output value of the hash as the input value of the next hash. Typically, the algorithm is used in Dash. The concept of the X-11 series uses multiple hashes to increase security and prevent ASIC mining. Currently, however, the X-11 series has been upgraded to increase the number of hash functions, such as X-13, X-14, X-15, X-16R, and X-17 because ASIC mining is still possible.

2.3. CryptoNote

CryptoNote was designed to be more inefficiently executed with a GPU than a CPU [15] to prevent ASIC mining. The performance of CryptoNote is susceptible to memory latency because memory creation and subsequent read operations occur repeatedly. This is similar to Etherium's Ethash function. CryptoNote creates blocks by determining the hash function to be used after memory-intensive tasks.

Despite such attempts, Bitmain released ASICs optimized for CryptoNote algorithms in March 2018. Monero uses CryptoNote to change its mining algorithm twice a year to prevent ASICs. Monero's algorithm change has prevented the use of ASICs. However, the frequent hard fork execution (radical changes) caused participants to break away from mining. In the end, there was a risk of the centralization of mining. To prevent frequent hard forks, RandomX proposed a key block concept that periodically changes mining methods [16].

3. The Proposed Method

This chapter provides an overview of and the implementation method of the proposed ECCPoW. We proposed the concept of ECCPoW to increase the resistance to ASIC, using a hash function that makes ASIC development difficult. Moreover, ASIC resistivity research reviews methods of inducing the loading

of memory; for example, Ethereum uses several hash algorithms, such as X-11. However, ASICs for Ethereum and the X-11 series have been developed. Thus, ECCPoW is a method for miners to release different hash functions for each block to curb the emergence of ASICs.

The ECCPoW can help ease the centralization of mining. The conversion to ASICs in mining is made by ordinary people to avoid being excluded from mining and by a small group of people with capital power to monopolize mining. Mining centralization of a small group is likely to lead to mining blocks for malicious purposes and forging and tampering with the mined blocks. The implementation of ASICs in the LDPC decoder lacks flexibility due to structural cost issues [17]. Moreover, ECCPoW proposes a method of combining the LDPC decoder with a hash function. We help reduce the risk in the blockchain by mitigating mining centralization.

3.1. ECCPoW Overview

The ECCPoW is a POW that uses the ECC decoder that is used in communication, which can be implemented using ASICs. As a simple example, cell phones use ASICs to implement an ECC decoder quickly and at low power. The parity-check Matrix H determines the design of the ECC decoder based on ASICs. In other words, ASIC equipment can produce a decoder using parity-check matrices. For mobile phones, ASICs have standardized parity-check matrices that allow an ECC decoder design. Building ASICs to match the decoder supporting countless parity-check matrices is difficult due to cost problems and decoder size problems.

Randomly, ECCPoW changes each block parity-check matrix (i.e., ECCPoW uses an infinite number of parity-check matrices). As a result, ECCPoW inhibits the development of ASICs for ECC decoders. The ECC decoding algorithm only runs on a CPU or GPU. For example, even if the SHA function used in the PoW is executed quickly, a bottleneck occurs in the execution of ECC decoding algorithms.

3.2. Create a Cryptographic Puzzle That Changes Every Block

The ECCPoW aims to create a cryptographic puzzle that changes every block. We changed the composite function used to create the cryptographic puzzle using the generation method by Gallagher [18] and the previous hash value. In other words, ECCPoW randomly generates an LDPC matrix used by the decoder of the composite function. The Gallagher method requires variables. Table 1 displays the definition of the variables used in the LDPC.

Table 1. Variables in the low-density parity-check Matrix H.

Variables	Definition
n	Number of columns in H
m	Number of rows in H
w_c	Number of 1s in each column
w_r	Number of 1s in each row

The LDPC matrix satisfies the equation $nw_c = (n-k)w_r$, where k is $n-m$ and 2^k is the total number of symbols that can be generated. When the variables are given, the LDPC Matrix H of size A is generated by the following method:

Step 1: Create a partial matrix of size $\frac{m}{w_c} \times n$

$$A_1 := \begin{bmatrix} \underbrace{1\,1\,\cdots\,1}_{w_r} & & & \\ & \underbrace{1\,1\,\cdots\,1}_{w_r} & & \\ & & \ddots & \\ & & & \underbrace{1\,1\,\cdots\,1}_{w_r} \end{bmatrix} \in \{0,1\}^{\frac{m}{w_c} \times n}$$

Step 2: Generate $w_c - 1$ submatrices by randomly permutating the following matrices:

$$A_i := \prod i(A_i) \in \{0,1\}^{\frac{m}{w_c} \times n},$$

where $\prod i$ is the ith sequence, and $i = 2, 3, \cdots, w_c$.

Step 3: Construct the final LDPC matrix using all submatrices above:

$$\mathrm{H} := \begin{bmatrix} A_1^T & A_2^T & \cdots & A_{w_c}^T \end{bmatrix} \in \{0,1\}^{m \times n}.$$

Moreover, ECCPoW changes the sequence of permutations through the previous hash values and uses the previous hash value as the seed value to determine the sequence of permutations. The sequence of permutations is random because the hash values are random. The code is implemented in Reference [19]. Table 2 compares Matrix H produced using different hash values. The lower part of the generated Matrix H in Table 2 is different.

Table 2. Form of the resulting Matrix H using different hash values.

Generated Matrix H $n = 24$ $m = 16$ $w_c = 3$ $w_r = 4$	[1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1] [0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 1 0 0 0] [0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1] [0 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0] [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 1 0 0] [1 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 1 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0] [0 1 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0] [0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 1 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0 0 1] [0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0] [1 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0]
Previous hash value	0x00000000000000000000000000000001
Generated Matrix H $n = 24$ $m = 16$ $w_c = 3$ $w_r = 4$	[1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1] [1 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0] [0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0] [0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 1 0 1 0 0 0 0 0] [0 0 0 1 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0] [0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 1 0 0 0 0 0 0 0 1] [0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0] [0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 1 0 0 0 0 0] [0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1] [0 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0] [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 1 0 0] [1 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0] [0 0 1 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0]
Previous hash value	0x00000000000000000000000000000002

3.3. Crypto Puzzle Decoder That Changes Every Block

The LDPC decoder of ECCPoW was developed using a message-passing algorithm. The decoder receives an $m \times n$ LDPC matrix of hash values $r \in \{0,1\}^n$ of length n as input values. The decoder outputs a

value $c \in \{0,1\}^n$ of length n. The decoder can produce two types of answers depending on the input hash value r. The decoder outputs a sign $D_{MP} : \{r, H\} \mapsto c_i$ if the entered hash value r satisfies $\|r - c_i\|_h \leq t$ for any sign c_i, where t is the value determined by the LDPC matrix. If not satisfied, the decoder outputs a random vector $c \in \{0,1\}^n$. The code is implemented in Reference [19].

The two conditions for determining whether to solve a cryptographic puzzle are listed in Table 3. Condition 1 determines that the cryptographic puzzle has been solved if the decoder's output value c satisfies the conditions. In Condition 1, the output value is the code. Condition 1 is possible because the output value is less likely to code when the value is any input to the decoder. For example, there is a low probability of finding an answer to any input in SHA 256. In Condition 2, the Hamming weight of the output value is an element of the given set S. Set S is the range value of the output value. Condition 2 occurs because the Hamming weights of the possible codes may differ when Matrix H is given.

Table 3. Conditions for the determination of crypto puzzle resolution.

Condition 1	(Original method) If the result of the decoder is code and has a specific Hamming weight, the problem is solved.
Condition 2	(Existing proof of work) If the result of rehashing the result of the decoder is less than a specific value, the problem is solved.

The satisfactory probability of Condition 1 requires a minimum Hamming distance value of H. To calculate this value, all distinct codes in 2^k must be considered, which is possible when the number of codes is small, but it is impossible when the number is large. Litsyn [20] reported the upper and lower bounds of the minimum Hamming distance value of H at specific w_c and w_r values. Table 4 reveals the probability of finding the codes according to the variables in the LDPC matrix. In this way, the upper bound of the probability is small. This makes it less likely for the decoder to meet Condition 1 when any value is entered.

Table 4. Probability of finding a code according to the variables of the low-density parity-check matrix.

$w_c = 4, w_r = 5$	p_1 Upper Bounds	p_1 Lower Bounds
$n = 80, k = 12$	6.32×10^{-5}	2.12×10^{-8}
$n = 120, k = 24$	1.65×10^{-8}	1.49×10^{-13}
$n = 160, k = 32$	4.06×10^{-10}	1.34×10^{-17}

Condition 2 is used to increase the difficulty of cryptographic puzzles when variables n, m, w_c, and w_r are fixed. Table 5 displays the probability that Condition 2 is satisfied when given a set S and part of the distribution of Hamming weights of the codes that can be generated when $n = 256$, $m = 192$, and $w_c, w_r = 5$ are given.

Table 5. The probability that Condition 2 is satisfied.

Hamming Weight	Probability	Element of the Set S	Probability that Condition 2 is Satisfied
98	$\approx 5 \times 10^{-5}$	98	$\approx 5 \times 10^{-5}$
...	...	...	...
128	$\approx 9.7 \times 10^{-2}$	98, 100, ... , 126	$\approx 4 \times 10^{-1}$
128	$\approx 1 \times 10^{-1}$	98, 100, ... , 126, 128	$\approx 5 \times 10^{-1}$

The probability of satisfying both Conditions 1 and 2 is as follows:

$$p := Pr\{c|\mathrm{Hc} = 0\} \times Pr\{\|c\|_h \in S\}.$$

We produced the difficulty table, as shown in Table 6, by calculating the probability of meeting Conditions 1 and 2 at the same time when the variables n, w_c, and w_r and the set S are given. The probability value p in Table 6 is the difficulty level of the cryptographic puzzle. The closer the probability value is to zero, the higher the difficulty of the cryptographic puzzle.

Table 6. Difficulty table for ECCPoW.

Lv.	n	w_c	w_r	Set S	p
1	32	3	4	{10, 12, . . . , 20, 22}	$\approx 3.07 \times 10^{-5}$
2	32	3	4	{10, 12, . . . , 14, 16}	$\approx 2.02 \times 10^{-5}$
			...		
379	128	3	4	{34, 94}	$\approx 5.12 \times 10^{-23}$
380	128	3	4	{34}	$\approx 2.60 \times 10^{-23}$

Condition 2 determines the resolution of the cryptographic puzzle by comparing the output value of the decoder with the result value and comparing the nonce with the target. Figure 1 illustrates Condition 2 for determining whether ECCPoW cryptographic puzzles are resolved. If the composite function and the hash algorithm are recognized as one hash function, Figure 1 has the same structure as Bitcoin, so the difficulty control function of Bitcoin can be used.

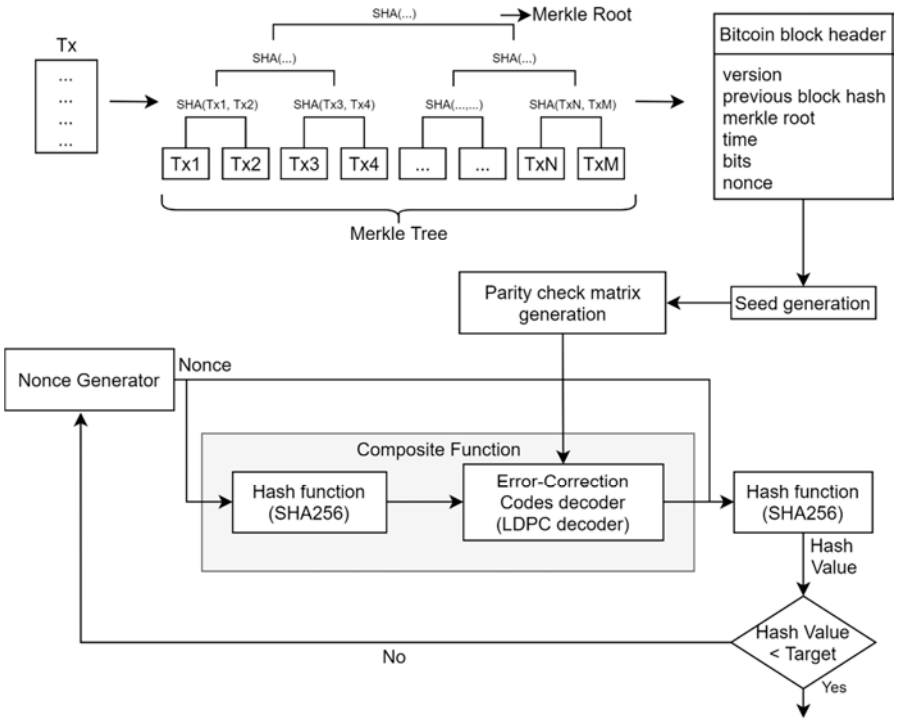


Figure 1. Condition 2 for determining the error-correction codes proof-of-work (ECCPoW) crypto puzzle resolution.

4. Experiment

This chapter presents experiments to verify ECCPoW. In the single-node experiment, the Bitcoin consensus algorithm was replaced by ECCPoW to verify the block generation function. In the multi-node experiment, we experimented with checking whether block generation, block synchronization, and transaction creation and transmission were performed correctly in a multi-node environment. In addition, the block generation time was tested in Bitcoin and Ethereum.

4.1. ECCPoW Operation Single/Multiple Node Experiment

We replaced Bitcoin's consensus algorithm with ECCPoW. The single-node experiment is a block generation experiment of the ECCPoW blockchain. Bitcoin uses the "generatetoaddress" command for block generation and receives the current address of the blockchain as a parameter. Then, a new block address is created using the "getnewaddress" command, and 10 blocks are created using "generatetoaddress." Figure 2 illustrates the change in 10 of the "blocks" after block generation.

Multiple node experiments use three random nodes (Nodes 1, 2, and 3) to experiment with block synchronization, to transmit transactions, and to verify them. This experiment demonstrates that each node mines different numbers of blocks, synchronizing to a long blockchain. Figure 3 indicates the results of mining and checking the blocks in Node 1.

Figure 4 reveals the results of mining and checking blocks on Node 2, and Figure 5 demonstrates the results of mining and checking blocks on Node 3.

Node 3 connects Node 1 (192.168.232.128) and Node 2 (192.168.232.129) using the "addnode" command. Then, Node 3 uses the "getadnednodeinfo" command to check the information on the connected nodes. Figure 6 illustrates the connection between Node 1 (192.168.232.128) and Node 2 (192.168.232.129) using the "addnode" command. Next, Figure 6 indicates the connected node information using the "getadnednodeinfo" command. Bitcoin's nodes synchronize to nodes that hold many blocks, and Bitcoin determines that nodes with many blocks are highly reliable. In the current experiment, Node 3 had the largest number of blocks. Nodes 1 and 2 synchronize to the blocks of Node 3, and these results are displayed in Figure 7.



Figure 2. Results of block generation.

```
getnewaddress
3FUFZ8ShBQpCGqqkZoFErG2aGo28QwaUDT
generatetoaddress 10 3FUFZ8ShBQpCGqqkZoFErG2aGo28QwaUDT
[
  "2b2d7485a399d93169ff8532de059db263e0ba42db73847489d4620a7b7510bf",
  "099640a372ec928e5af6aa1b2032014b24a04225724c5553c1a80bf53dd0d8c4",
  "1c44c7078afcdd49eb9ab8fd4ccbc64419cca96afda91c571a51de163b0ac9ce",
  "630b1b3127c74b8892ff7fcbae89f2733aeadf57f225354a05f60583200ac8a2",
  "74449f452fd35732dbfab5e3ada8f531d7df4ee1a7df23ef41ab91ca1b4c3ca3",
  "282768f094a38a6f7113e4ffc45f5cb93ae7fb01f9e164bd8482f59df49712ff",
  "4a16965fc7efab971b99cbd8cd807eb78b17852fe4be55a620cf00ca865bbafc",
  "5ea006ef89893ab591ba39d995c2b09bcba5ac21813a3c8ff9d7752765c820e7",
  "c43c6bb01c954f23b998f88ead512a7b7cbc1350e37225e6659f02f923ea8f17",
  "7a39ec86ae15a01d8d6930f2fcc5300d8830726fab16b4e733c59db1719f8e43"
]

getblockchaininfo
{
  "chain": "main",
  "blocks": 10,
  "headers": 10,
  "bestblockhash":
"7a39ec86ae15a01d8d6930f2fcc5300d8830726fab16b4e733c59db1719f8e43",
  "difficulty": 4.5230594683691 96e+73,
  "mediantime": 1569822743,
```

Figure 3. Block mining and verification in Node 1.

```
getnewaddress
34rqCn3Xf6PDtLoHLZ8wPp1wsRV6Gb9KFY
generatetoaddress 20 34rqCn3Xf6PDtLoHLZ8wPp1wsRV6Gb9KFY
[
  "404a6788ab0dcb5bd356b03b400c005fa8f49380ab978d4a7b3537dfc1f6ba23",
  "ca0404483d852a32abb8609073bbebab0d06c55c6fc5b9e141de211d570f5576",
  "129b0e34f22c2f745efcd708df8b653ac295ac7b08b9bf36de518d88e0cfaa9a",
  "807f58aa7d221d15c6acc9e3b0ea89bc96de2205c797547d72d10ea1d463a588",
  "d9265ae393201c60e988ccbeada6e82038e90997deaad64e5057d9c523d0abc3",
  "b518305ba918e8619c36c4d449e161dbe8ef9517bc2d2fd2cfaea92af4171008",
  "8d13e15542fda1e5950268f6d87b87a5ac737e4377231d06b79f683255d73869",
  "22c3b9e63d3b7f6fbc8959805fef052aea6fb136124fe5799097d61c6b29b4c4",
  "5658489c42ca5aecf600e324a8a07e4c1a38b5f0ec8c6367f4c6941f031fd496",
  "43cd9632b9879fa290c3f6c2e63818189d84e752791233ced6bc58e29977450e",
  "2c65177825619bbebc75caccc0afccb2e58fa69eef178cb472eff29c2eda6ca9",
  "67e39e0482caafc683b8e57c6961fd7a4a3f376117c4349acd5ba1a84574235a",

getblockchaininfo
{
  "chain": "main",
  "blocks": 20,
  "headers": 20,
  "bestblockhash":
"9e99fb700b43d5a59fa97a6d1550e536c3933c0899158d9300ea74ac50928279",
  "difficulty": 4.5230594683691 96e+73,
  "mediantime": 1569822820,
```

Figure 4. Block mining and verification in Node 2.

```
getnewaddress
3765hBQFFHMn8FProyKFuF4vHKayafka4d
generatetoaddress 30 3765hBQFFHMn8FProyKFuF4vHKayafka4d
[
  "28bbd5b26fd5b9530614279cf06a4e036056aa1b996d05d699f73d71b77311c9",
  "11902f08b5c19f9ce6551730311f076453ef620014f703643c9225d1279bf5ea",
  "0610136713ae049268c72dec6d32cd13346f12b9bcc20b1ae19d8d9492eeb955",
  "6848477bf878c2e38ff13d3aedcc047a8f68659da8454e1800ac963933c688c2",
  "1d20dd47c8400a5833fc84b4fe9d91d4f6dfc474fbaf939f45745121669d740c",
  "8f084bc0ea45ca54b81f0bc0d931591572209cf69210ccea9daf16edfc18ae56",
  "f8a52172a10c14283e9749c62aa9d4362d677f32d586b71e0ffced4ddf4f111f",
  "a0b47bde1bf2a11a35b0672e2a0be022eb94cd4dbb063bd35de44bc038cf0c43",
  "56221e5d9bc496a2cf2a61bc13efef967b87e74273e7dc6fbcd3da4f7486a532",
  "a93231e200e63364d56b5669144f2353ac22b018c1835220ada57ea471d68f58",
  "0f4913a21d6481cf7f476952dd6e5b7fc8c188c420569fd4365f5f12a766ab3c",
  "06d84fa3aa0b97b32194a5ff9533847250261d567e33c7b2b199cee0781f3cda",

getblockchaininfo
{
  "chain": "main",
  "blocks": 30,
  "headers": 30,
  "bestblockhash":
"0aaa7eb4a5bbd818e5a76383fd1f014df645ededd0ffb9568ee262b5080d7b5e",
  "difficulty": 4.5230594683691 96e+73,
  "mediantime": 1569822888,
```

Figure 5. Block mining and verification in Node 3.

```
addnode 192.168.232.128 add
null
addnode 192.168.232.129 add
null
```

```
getaddednodeinfo
[
  {
    "addednode": "192.168.232.128",
    "connected": true,
    "addresses": [
      {
        "address": "192.168.232.128:9777",
        "connected": "outbound"
      }
    ]
  },
  {
    "addednode": "192.168.232.129",
    "connected": true,
    "addresses": [
      {
        "address": "192.168.232.129:9777",
        "connected": "outbound"
      }
    ]
  }
]
```

Figure 6. Node connections in Node 3.

```
getblockchaininfo
{
  "chain": "main",
  "blocks": 30,
  "headers": 30,
  "bestblockhash":
"0aaa7eb4a5bbd818e5a76383fd1f014df645ededd0ffb9568ee262b5080d7b5e",
  "difficulty": 4.523059468369196e+73,
  "mediantime": 1569822888,
```

Figure 7. Blockchain synchronization of Nodes 1 and 2.

We checked the behavior of the blockchain by checking the synchronization. We sent the transactions between linked Nodes 2 and 3. Figure 8 indicates the balances held by Nodes 2 and 3. Currently, Node 2 has zero, and Node 3 has 1000.

```
getbalance
0.00000000

getbalance
1000.00000000
```

Figure 8. Balances for Node 2 (**top**) and Node 3 (**bottom**).

Figure 9 displays the address of Node 2 (in the Pay-To field) and the amount of coin to send (in the Amount field) for Node 3 to transmit 500 coins to Node 2. A transaction fee was set to the minimum cost of 0.00001. Figure 10 reveals the transaction transfer by identifying the amount of coin in Node 2 and displaying the recent transaction records. Figure 11 illustrates that the number of coins in Node 3 and the recent transaction record decreased.

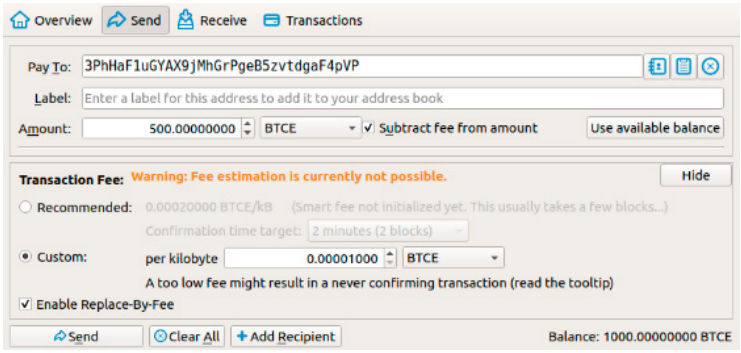


Figure 9. The input of the transaction transfer.

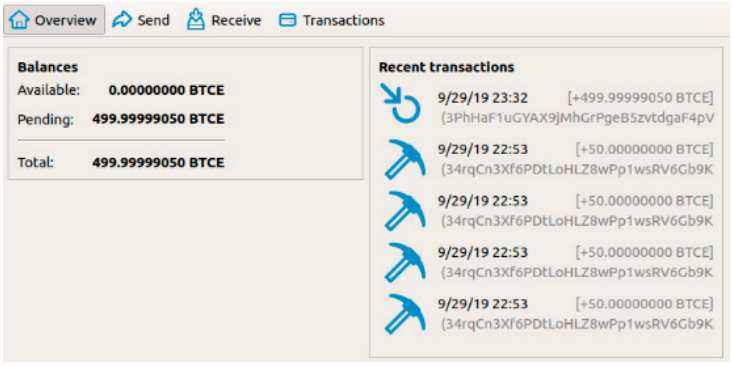


Figure 10. Balances and transaction logs for Node 2.

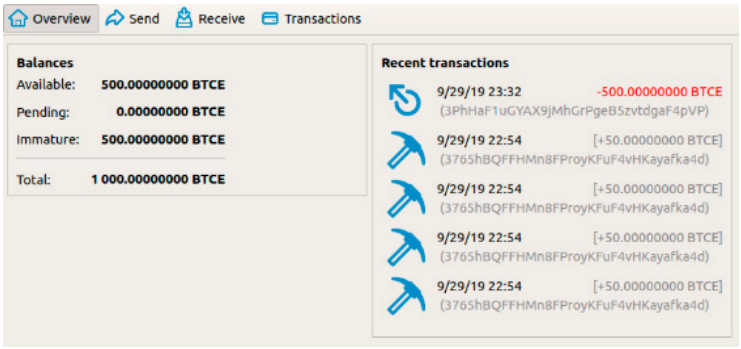


Figure 11. Balances and transaction logs for Node 3.

4.2. Block Generation Time

Figure 12 reveals the time-by-time block generation of Bitcoin with ECCPoW, using the difficulty table proposed in Section 3. Block generation time should be able to meet the target generation time for a stable blockchain. The blockchain adjusts the target block generation time by adjusting the difficulty level over time. The experimental environment used block generation with one node. As shown in Figure 12, the block generation time is unstable in the experiment. Sufficient nodes can confirm the stability of the block generation time. The difficulty table must be finely adjusted.

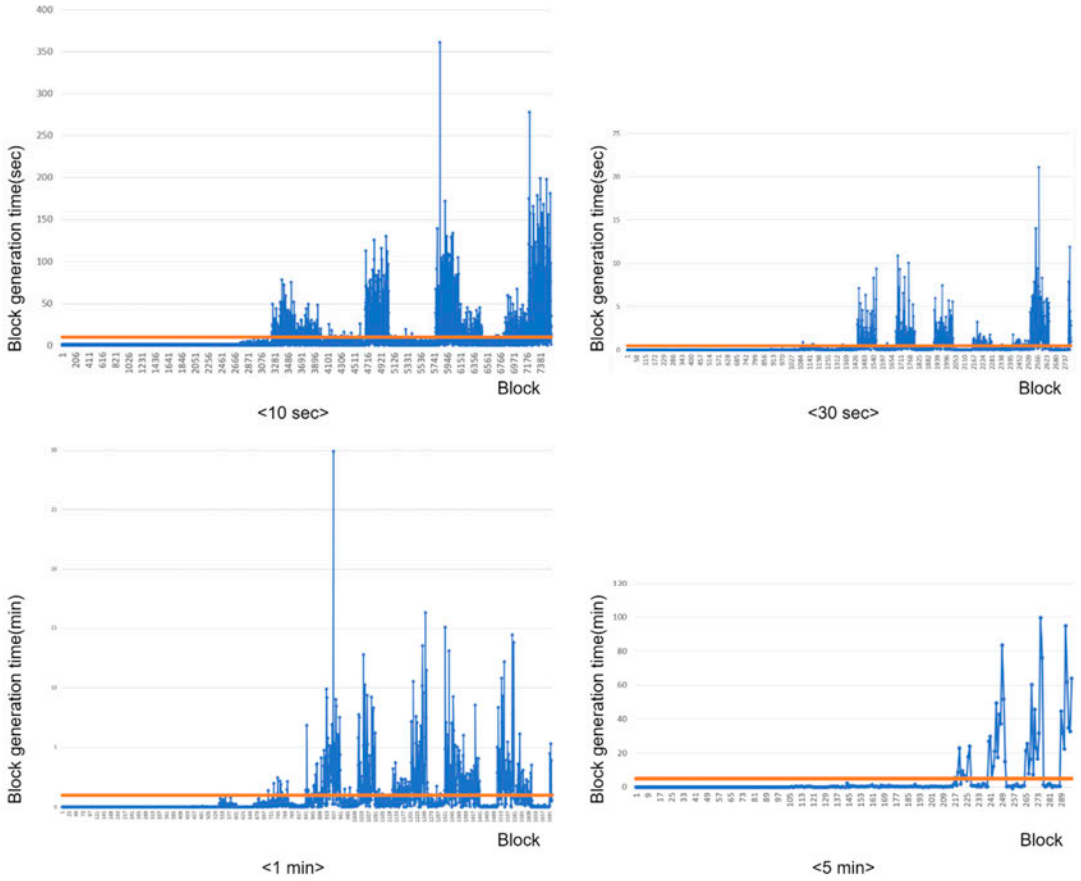


Figure 12. Block generation time of Bitcoin with ECCPoW.

5. Evaluation

We evaluated the mining centralization, security, and scalability of ECCPoW using Amazon Web Services. Table 7 presents the evaluation environment. A monitoring computer checked the network test results. The seed instance made the blockchain network connections between nodes. The mining instance was responsible for block mining. The implementation environment for each node was a Ubuntu Server 18.04 LTS, m5.large (vCPU processor 2 core with 8 GB of RAM, and a 20 GB solid-state drive).

Table 7. Implementation environment of the evaluation.

No	Role	CPU	Memory (GB)	HDD/SSD (GB)	Volume
1	Monitoring PC	Intel i7-8700 3.20 GHz	16	SSD 256	1
2	Seed Instance	AWS m5.xlarge vCPU 2	8	HDD 20	6
3	Mining Instance	AWS m5.xlarge vCPU 2	8	HDD 20	40

The blockchain trilemma problem is that trade-off relationships occur in the evaluation characteristics of the blockchain. The trilemma elements of a blockchain are decentralization, scalability, and security. Thus, we compared Bitcoin in terms of these elements (mining centralization instead of decentralization). Table 8 displays the evaluation standards and goals for evaluation.

Table 8. Description of evaluation features.

Evaluation Features	Unit of Measurement	Evaluation Standards	Evaluation Goal
Mining centralization	Distribution of mining success rate	40% (Estimate) (Bitcoin, Oct.–Dec. 2018)	40%
Security	Security of Bitcoin contrast	100% (Bitcoin)	100%
Scalability	Scalability of Bitcoin contrast	100% (Bitcoin)	100%

Bitcoin has a total hash rate of approximately 90 TH/s (March 2020). It is impossible to compare the current version of Bitcoin with ECCPoW. Moreover, ECCPoW (ver. 0.1.2) replaced the Bitcoin 0.17 version [21] of SHA 256 [22]. We compared two completely initialized blockchains (block count zero).

We compared the Bitcoin and ECCPoW initialized to set the evaluation environment (difficulty change cycle, target block generation time, 23 instances, among others). The blockchain typically sets a difficulty change cycle of 60 min and a target block creation time of 3 min for experimentation.

5.1. Mining the Centralization Evaluation

Mining centralization is when a network is no longer centralized and operates autonomously within a blockchain. We define mining centralization as when the nodes are fairly mined with the same performance. The indicators used in the mining centralization evaluation include the distribution of the mining success rates, which are defined by the formula below. According to the formula, the lower the dispersion of the probability of mining success, the higher the distribution of mining success. In other words, the distribution of mining success rates is higher for each participating node to have a uniform distribution of mining success rates, which means better dispersion. Bitcoin is estimated to have a 40% distribution of mining success rates (October to December 2018). Moreover, ECCPoW targets 40% dispersion:

$$A = \frac{C}{\sqrt{B + C^2}} \times 100 A = \text{The distribution of mining success rate } (\%),$$

$$B = \text{The dispersion of mining success rate,}$$

$$\text{C} = \text{Average of the mining success rate.}$$

We created three seed nodes for the ECCPoW blockchain and composed 20 mining nodes. Table 9 lists the number of mining successes of each mining node at the time of 100 blocks being mined. Table 10 indicates the distribution of mining success.

The dispersion of ECCPoW was measured by the distribution of the mining success rate and measured at 92.22%, which was 32% higher than the assessment target of 60% (Table 10). The results of the experiment revealed that the ECCPoW miner is more likely to be rewarded than a Bitcoin miner in an ideal environment where participating nodes exhibit the same performance. In other words, ECCPoW is stronger in mining centralization than Bitcoin.

Table 9. The number of mining success and number of nodes.

Number of Mining Nodes	Number of Mining Success	Number of Mining Nodes	Number of Mining Success
1	4	11	7
2	9	12	4
3	3	13	12
4	4	14	5
5	6	15	11
6	4	16	2
7	6	17	7
8	6	18	6
9	5	19	10
10	5	20	8

Table 10. Evaluation results of the mining centralization evaluation features.

Total Number of Mining Successes	Average Mining Successes	Square of the Average Number	Dispersion of the Average Number	Distribution of Mining Successes
124	6.2	38.44	6.76	92.21944

5.2. Security Evaluation

Security is associated with the difficulty of the blockchain being altered by a miner's attack. A typical attack on the blockchain is the double-payment issue. One of the causes of double-payment problems occurs in the branch of the blockchain. An orphan chain is a chain that has a branch other than the main chain, and an orphan block is a block belonging to an orphan chain. We assess the security as the ratio of orphaned blocks to the total number of blocks in the blockchain. The security assessment calculates the orphan block ratio of Bitcoin and ECCPoW using the expression below. We evaluated the security of ECCPoW based on the security of Bitcoin at 100%.

$$Orphan\ block\ ratio = \frac{number\ of\ orphan\ blocks}{Total\ Height\ of\ Blockchain} \times 100$$

We configured the blockchain test environment with three seed nodes and 20 mining nodes. In addition, we mined 100 blocks in the blockchain and checked the orphan blocks. Figure 13 depicts the orphan block identification for the security assessment. After mining 40 blocks, the evaluation identified blocks for stable synchronization of the blockchain. In the experiment, orphan blocks of a height of one frequently occurred in both environments. In our experiment, we classified the height of subchains of two or higher as orphan blocks. In Figure 13, orphan blocks correspond to Blocks 13, 14, 15, 16, and 18.

Figure 14 demonstrates the blockchain status based on the results of the security assessment in the Bitcoin environment.

Figure 15 reveals the blockchain status according to the results of the security assessment in the ECCPoW environment.

All blockchains were measured at 0% because of the assessment. The security of ECCPoW was the same on a Bitcoin and orphan block basis. Blockchain measurements inhibited the generation of orphan blocks.

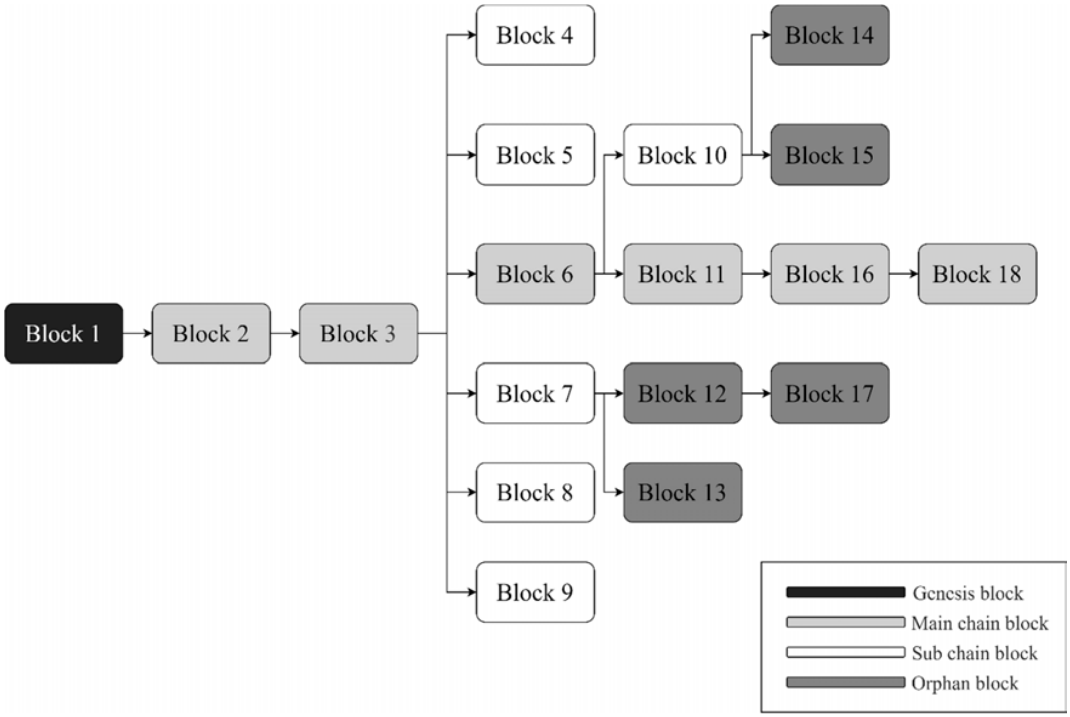


Figure 13. Diagram description of security evaluation.

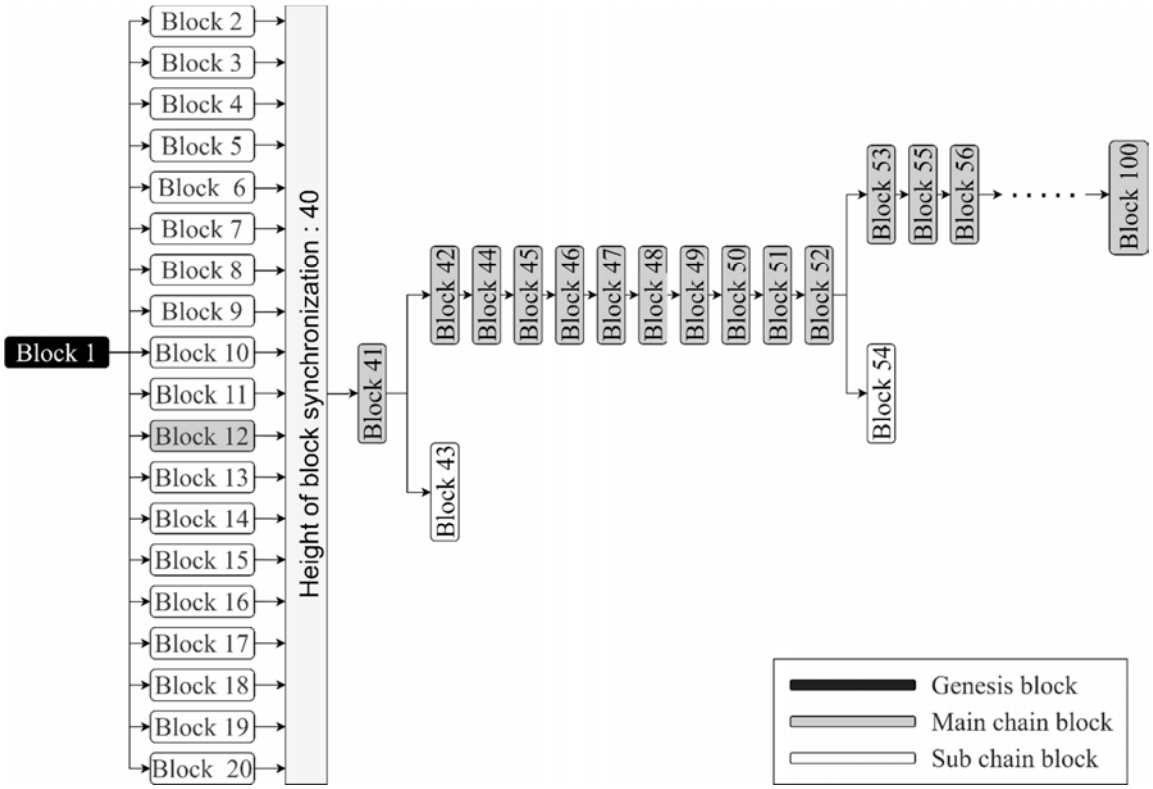


Figure 14. Evaluation diagram of Bitcoin security.

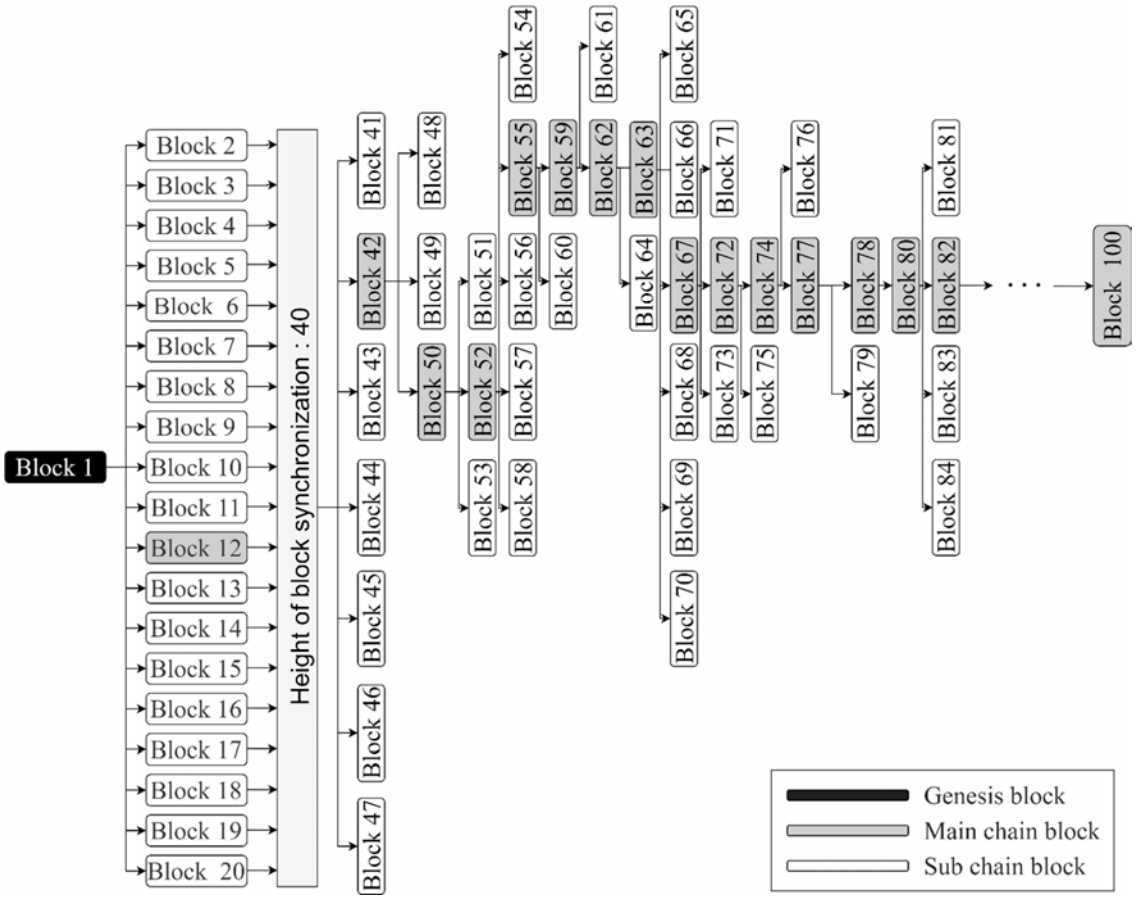


Figure 15. Evaluation diagram of ECCPoW security.

5.3. Scalability Evaluation

Scalability is the extent to which a system can flexibly respond to an increase in the number of users. Blockchain scalability is related to the transaction processing speed of blockchain. Transactions per second (TPS) is the transaction processing speed of the blockchain. Thus, we evaluated the scalability of the blockchain using TPS. The following equation defines how to obtain TPS using transactions in blocks.

$$\text{TPS} = \frac{Total\ number\ of\ transactions\ in\ block}{Generation\ time\ of\ total\ block}.$$

We configured the blockchain test environment with three seed nodes and 20 mining nodes. Transaction generation occurs continuously from height 91 to 100—the blockchain mines 100 blocks. We checked the number of transactions in the height blocks of the blockchain from 91 to 100. We calculated the TPS of the blockchain. Table 11 lists the generation time and number of transactions of blocks of height 91 to 100 in Bitcoin.

Table 12 displays the generation time and the number of transactions of blocks of height 91 to 100, in ECCPoW. The evaluation results revealed that Bitcoin's TPS is 0.95, and ECCPoW's average TPS is 0.94. Moreover, ECCPoW's scalability was measured at 98.95%, which was down 1.02% from Bitcoin's scalability. In addition, ECCPoW added a process to Bitcoin to resist ASICs. However, the scalability values of ECCPoW and Bitcoin were assessed at a similar level.

Table 11. The evaluation result of Bitcoin scalability.

Number	Height	Block Generation Time (Seconds)	Number of Transactions
1	91	719	672
2	92	19	22
3	93	94	88
4	94	144	141
5	95	275	263
6	96	40	40
7	97	313	296
8	98	6	11
9	99	574	534
10	100	146	137
Total		2330	2204
TPS		0.945922747 TPS	

Table 12. Evaluation results of ECCPoW scalability.

Number	Height	Block Generation Time (Seconds)	Number of Transactions
1	91	151	140
2	92	192	182
3	93	369	351
4	94	361	331
5	95	156	151
6	96	214	205
7	97	124	120
8	98	267	250
9	99	585	548
10	100	514	480
Total		2933	2758
TPS		0.940334129 TPS	

5.4. Comparison to Related Work

This section compares the features with the existing studies on ASIC resistance. Table 13 compares the proposed method and existing relevant research on ASIC resistance. The ASIC resistance study for the PoW blockchain has three approaches. The first method causes bottlenecks using memory access in a hash function—for example, Ethereum (Ethash) and Bytecoin. The second method prevents the generation of ASICs using a hash function overlay, for example, the X-11 series. Finally, the third method uses periodic hash function replacement for ASICs, for example, in Monero (2019.9, RandomX algorithm) and Ethereum (2020.7 applying scheduled, ProgPoW).

Table 13. Comparison results of the related work.

Comparison Features	Proposed Method	Memory Approach	Hash Function Overlay	Periodic Hash Function Replacement
Applied cryptocurrency	-	Ethereum, Bytecoin	X-11 series	Monero (2019.9, RandomX), Ethereum (2020.7 applying scheduled, ProgPoW)
Characteristic	Change hash function every block	Force memory access	Overlapping multiple hash functions	Hard fork manually or automatically with a hash function
Algorithm	ECCPoW	Ethash, ProgPoW, CryptoNight	Blake, Bmw, Groetl, etc.	RandomX, ProgPoW
ASIC appearance	-	Yes	Yes	Unknown
ASIC resistance induction method	Use of ASIC resistance by connecting the LDPC decoder and hash function	Induce cache miss when creating blocks (using a DAG, etc.)	Overlapping the difficulty of known hash functions	Hard fork every six months (formerly Monero). Convert the hash function for a period using the key-block concept (currently Monero)

The resistance induction of each ASIC in the existing studies is as follows. The memory approach induces a cache miss (using DAG, among others) when creating blocks. This method degrades ASIC performance as memory access increases. The hash function overlay method uses the overlap of the difficulty of known hash functions. This method relies on the security of the applied hash function. The periodic hash function replacement method periodically changes the hash function manually or automatically.

6. Conclusions

In this paper, we explained ECCPoW and applied the proposed method to Bitcoin. This paper addressed the problem of centralization of mining due to the emergence of ASICs. We proposed a PoW concept based on error-correction codes to solve this problem. The core of the ECCPoW is the connection of the hash function with the LDPC decoder. Blockchain applied with the ECCPoW determines the completion of the POW using the output value of the decoder. In addition, ASIC suppression is possible because ASICs use the LDPC decoder.

This paper contributes to the phenomenon of symmetry in the PoW blockchain. The total block size of the PoW blockchain symmetrically influences the hash rate. The PoW blockchain increases stability as the hash rate increases in size. However, a high hash rate causes the waste of computing power in block generation. Thus, we mitigate the causes of a high hash rate using ASIC resistance.

The ECCPoW offers a method of solving different puzzles in each block to avoid ASICs. This maximizes the benefits of how existing studies use a limited number of hash functions to solve each block of different hash functions. The proposed method offers more effective connections and the use of multiple hash functions. We presented the difficulty control, parity-check matrix generation method, hash vector generation, and code determination methods for implementing ECCPoW. Furthermore, ECCPoW was applied to Bitcoin to verify the proposed method. We assessed the mining centralization, security, and scalability of ECCPoW and Bitcoin. We found that ECCPoW maintained security and scalability, showing 32% higher mining centralization than Bitcoin. Using the proposed method, ECCPoW does not require high hash rates, and miners can compete more fairly.

This study contributes to the previous literature on ASIC resistance in the PoW blockchain. The ASIC resistance study of the PoW blockchain was conducted in response to hardware development by ASIC manufacturers. One of the studies on ASIC resistance induced forced access to memory in the hash function, thereby lowering the performance of ASICs. Other studies have suggested periodically altering the hash function to render ASICs useless. The ECCPoW provides a method of releasing different hash functions for each block, rather than changing the periodic hash function.

The ECCPoW revealed the limit of block generation time in the experiment. For the stable operation of the blockchain, the block generation time of the ECCPoW must be stable and controllable. The results of the block generation time test were unstable in the experiment using one node. Thus, ECCPoW requires further research regarding the difficulty of block generation for stable operation. We also need to increase the number of nodes in the ECCPoW to carry out mining.

Author Contributions: H.J. and H.-N.L. contributed equally to this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean government (MSIP) (NRF-2018R1A2A1A19018665).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nakamoto, S. *Bitcoin: A Peer-to-Peer Electronic Cash System*; 2008; Available online: https://bitcoin.org/bitcoin.pdf (accessed on 9 June 2020).
2. Nofer, M.; Gomber, P.; Hinz, O.; Schiereck, D. Blockchain. *Bus. Inf. Syst. Eng.* **2017**, *59*, 183–187. [CrossRef]
3. Saberi, S.; Kouhizadeh, M.; Sarkis, J.; Shen, L. Blockchain Technology and Its Relationships to Sustainable Supply Chain Management. *Int. J. Prod. Res.* **2019**, *57*, 2117–2135. [CrossRef]
4. Qin, R.; Yuan, Y.; Wang, F.-Y. Research on the Selection Strategies of Blockchain Mining Pools. *IEEE Trans. Comput. Soc. Syst.* **2018**, *5*, 748–757. [CrossRef]
5. Liu, X.; Wang, W.; Niyato, D.; Zhao, N.; Wang, P. Evolutionary Game for Mining Pool Selection in Blockchain Networks. *IEEE Wirel. Commun. Lett.* **2018**, *7*, 760–763. [CrossRef]
6. Karame, G.O.; Androulaki, E.; Capkun, S. Double-Spending Fast Payments in Bitcoin. In Proceedings of the 2012 ACM Conference on Computer and Communications Security, Raleigh, NC, USA, 16–18 October 2012; pp. 906–917. [CrossRef]
7. Jang, J.; Lee, H.-N. Profitable Double-Spending Attacks. *arXiv* **2019**. Available online: https://arxiv.org/abs/1903.01711 (accessed on 9 June 2020).
8. Antminer S19 Pro. Available online: https://support.bitmain.com/hc/en-us/articles/900000261726-S19-Pro-Specifications (accessed on 9 June 2020).
9. Park, S.; Kim, H.; Lee, H.N. Introduction to Error-Correction Codes Proof-of-WorK. *Mag. IEIE* **2019**, *5*, 26–32.
10. Buterin, V. A Next-Generation Smart Contract and Decentralized Application Platform. *White Pap.* Available online: https://ethereum.org/whitepaper/ (accessed on 9 June 2020).
11. Wood, G. Ethereum: A Secure Decentralised Generalised Sransaction Byzantium Version. *Yellow Pap.* **2014**. Available online: https://ethereum.github.io/yellowpaper/paper.pdf (accessed on 9 June 2020).
12. Zhou, Q.; Huang, H.; Zheng, Z.; Bian, J. Solutions to Scalability of Blockchain: A Survey. *IEEE Access* **2020**, *8*, 16440–16455. [CrossRef]
13. Greg, C.; Andrea, L.; Michael, C.; IfDefElse. ProgPoW, a Programmatic Proof-of-Work. Ethereum—EIPs, No. 1057. 2018. Available online: https://eips.ethereum.org/EIPS/eip-1057 (accessed on 9 June 2020).
14. Duffield, E.; Diaz, D. Dash: A PrivacyCentric Crypto-Currency. *White Pap.* **2014**. Available online: https://docs.dash.org/en/stable/introduction/about.html#whitepaper (accessed on 9 June 2020).
15. Saberhagen, N.V. Cryptonote v2.0. *White Pap.* **2013**. Available online: https://cryptonote.org/whitepaper.pdf (accessed on 9 June 2020).
16. RandomX. Available online: https://github.com/tevador/RandomX (accessed on 9 June 2020).
17. Shao, S.; Hailes, P.; Wang, T.-Y.; Wu, J.-Y.; Maunder, R.G.; Al-Hashimi, B.M.; Hanzo, L. Survey of Turbo, LDPC, and Polar Decoder ASIC Implementations. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 2309–2333. [CrossRef]
18. Gallager, R. Low-Density Parity-Check Codes. *IRE Trans. Inf. Theory* **1962**, *8*, 21–28. [CrossRef]
19. Bitcoin ECC LDPC. Available online: https://github.com/cryptoecc/bitcoin_ECC/blob/ecc-0.1/src/ldpc/LDPC.cpp (accessed on 9 June 2020).

20. Ben-Haim, Y.; Litsyn, S. Upper Bounds on the Rate of LDPC Codes as a Function of Minimum Distance. *IEEE Trans. Inf. Theory* **2006**, *52*, 2092–2100. [CrossRef]
21. Bitcoin Core 0.17 Version. Available online: https://github.com/bitcoin/bitcoin/tree/0.17 (accessed on 9 June 2020).
22. Bitcoin ECC 0.1.2 Version. Available online: https://github.com/cryptoecc/bitcoin_ECC/tree/ecc-0.1.2 (accessed on 9 June 2020).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN) | 979-8-3503-3538-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICUFN57995.2023.10201050

Smart Contract-Based Checkpoint for Initial PoW Network Security

Seungmin Kim, Gyeongdeok Maeng and Heung-No Lee*
Dept. of Electrical Engineering and Computer Sciences
Gwangju Institute of Science and Technology (GIST)
Gwangju, Korea
heungno@gist.ac.kr

***Abstract*—Proof of Work (PoW) is the most widely adopted consensus mechanism on public blockchains. The PoW blockchain network achieves consensus by solving computational problems. If a network participant owns more than 50% of the total computational power in the network, they can forge the blockchain. Hence, initial PoW blockchain networks with low computing power are vulnerable to block forgery attacks. We propose a smart contract-based checkpoint method to improve security vulnerabilities in the initial blockchain networks. In our method, participants periodically record block headers in an Ethereum smart contract. The recorded checkpoint block header validates the blockchain. Participants reject blocks with blocks that differ from the recorded checkpoints. Our method ensures the integrity of blocks until the height of the most recently created checkpoint, reducing the risk of double-spending. We optimize checkpoint costs by overlapping multiple checkpoint processes in a single transaction. The interval of our checkpoint method grows with the growth of the network, making the network less dependent on checkpoints. We analyze the performance of checkpoints in mitigating attacks and demonstrate that they significantly decrease the success probability of attacks in the network.**

***Index Terms*—Checkpoint, PoW blockchain, smart contract, double spending attack**

I. Introduction

Blockchain is a decentralized ledger that maintains the integrity of data through consensus. Blockchain uses a consensus mechanism to select decision-makers in each block. [1]. The decision-makers in the blockchain record the validated transactions in block and broadcast block to the network. The consensus mechanism is critical to blockchain decision-making and therefore has a major impact on the security and scalability of a blockchain [2]. Innumerable consensus mechanisms have been studied to increase the security and scalability of blockchains [3]–[6]. Proof of Work(PoW), introduced in Bitcoin, is the most popular consensus mechanism [7]. PoW is a consensus mechanism that involves solving mathematical problems to select decision-makers. The participant who solves the problem first gets the authority to create a block. The computational works involved in decision-making make it difficult for attackers to engage in malicious behavior. Proof of Work is still the most typical consensus mechanism, although Ethereum transitioned to Proof of Stake(PoS). [8].

In a PoW network, participants with more than 50% of the total computing power have majority decision-making power. This truth means they can monopolize the network's decision-making process and manipulate the data on the blockchain [9]. It is problematic that one participant possesses more than 50% of the network's computing power in a mature PoW network. One participant possessing more than 50% of the network's computing power in a blockchain network is complicated and becomes increasingly difficult as the number of participants increases. However, in early blockchain networks, where the network size is small, possessing more than 50% is relatively easy.

Checkpoints are state values recorded at regular intervals in a blockchain network. Blockchain networks regularly store state values as checkpoints. Network participants recognize blockchains that contain checkpoints as legitimate chains. This new regulation increases the network's security by making it harder for attackers to forge blockchains. Checkpoints also serve as reference points for new network participants. The new participant connects to the blockchain network and attempts to synchronize blocks. There are many forked chains in the network, which can cause new participants to synchronize the wrong chains. New participants can use checkpoints to validate forked chains and prevent incorrect synchronization.

Related Work. There have been various implementations that have applied checkpoints and numerous research studies proposing checkpoint mechanisms. Bitcoin Core, a client software for Bitcoin, utilizes hard-coded checkpoints internally to protect the initial network from potential attack vulnerabilities [10]. Hard-coded checkpoints are inflexible and centralized to the developers. Polygon [11] executes the Proof of Stake (PoS) consensus process in Ethereum smart contracts and stores the outcomes as checkpoints. Polygon does not have its inherent proof-of-work mechanism and relies on Ethereum for network security. In [12], Wang and Kim propose an additionally distributed ledger for external validators and a method for recording checkpoints in the ledger. Their recording checkpoints to a distributed ledger incur a high network cost. [13] proposes to create a minority committee for each checkpoint, which decides and records the checkpoint. Their work assumes a synchronous network, unlike asynchronous blockchain networks.

Contributions. We propose a smart contract-based checkpoint method to improve the security of initial proof-of-work networks. The core idea of our proposed method is

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on December 06,2023 at 13:19:58 UTC from IEEE Xplore. Restrictions apply.

to record checkpoints in the smart contracts of the Ethereum network. The Ethereum blockchain network's enormous size guarantees our checkpoints' security against tampering efforts. We design the checkpoint generation process to consider the asynchronous nature of blockchain networks. The checkpoint creation process consists of three timelines for each checkpoint generation interval. We overlap multiple checkpoint processes in a single transaction to reduce the cost of checkpoint generation. The checkpoint interval gradually increases with the total computing power of the network. As the interval between checkpoints recorded on Ethereum increases, the network becomes independent of Ethereum. We assume a selfish mining attack scenario and analyze the attack probabilities for checkpoints and typical networks. According to our analysis, our proposed method effectively reduces the attack probability compared to regular PoW networks.

Organization. The rest of this paper is structured as follows. In section II, we describe an overview of the research and background concepts related to checkpoints. Section III explains the checkpoint protocol and smart contracts. In section IV, we consider a scenario involving selfish mining attacks and compare the performance of checkpoints with a typical network. In section V, we conclude and discuss future work.

II. Background and Overview

A. PoW blockchain network attack

Attacks on proof-of-work (PoW) blockchain networks largely stem from the probabilistic nature of the PoW mechanism. PoW network follows the Longest Chain rule, which adopts the longest chain, i.e., the chain with the most accumulated work, as the legitimate chain of the network [14]. Whenever the PoW network finds a longer chain, it replaces the existing chain with the new one.

Double-spending attacks exploit the Longest Chain rule [15]. In a double-spend attack, the attacker pays through a legitimate transaction and receives a reward for the transaction. The attacker uses their computational power to create the longest blockchain that contains the forged transaction. The network recognizes the longest block, the forged blockchain, as legitimate and invalidates the legitimate transaction. The attacker obtains the forged transaction reward and reuses the currency used in the transaction. Selfish mining attack is another form of attack that exploits the Longest Chain rule [16]. In a selfish mining attack, the attacker intentionally withholds generated blocks instead of immediately broadcasting them. The attacker continues to increase the height of their private blockchain without revealing the blocks. The attacker announces the blocks after the attacker's blockchain surpasses the network's height. The transaction and block rewards on the network's chain become insignificant since the attacker's chain becomes the longest.

The attacker must catch up to the network's longest chain to execute such an attack. The attack has a probability of success of 1 if the attacker's computational power is over half of the network's total computing power since the attacker can catch up to the blockchain. However, even though the attacker's computational power does not exceed 50%, there is a possibility that they can succeed in executing the attack probabilistically [17].

B. Confirmation distance

Confirmation distance is a mechanism introduced in blockchain networks to mitigate double-spending attacks [18]. Confirmation distance refers to the number of blocks a user has to wait for a transaction to be sent and confirmed. A participant sends a transaction to the blockchain but does not immediately supply the transaction reward and waits for the confirmation distance. The participant sends the transaction reward after all the confirmation distance blocks have been generated. A double-spending attacker must forge blocks after the confirmation distance to invalidate transactions. The confirmation distance is effective against double-spending attacks employing less than 50% of the total computational power. The probability of the attacker catching up with the blocks decreases as block height increases, assuming the attacker has less than 50% computing power. To receive a transaction reward, an attacker must attempt a double-spend attack after the confirmation distance. Confirmation distance increases the difficulty of an attack by requiring the attacker to catch up to the block after the confirmation distance.

The confirmation distance in a blockchain network can vary depending on the network conditions. The confirmation distance increases with a higher rate of total computational power for an attacker and decreases with a lower rate. A longer confirmation distance means a longer transaction processing time, which can cause inconvenience for participants who require fast transaction confirmations. Furthermore, attacks with more than 50% computational power always generate longer blocks, ignoring confirmation distance. As a result, additional security protocols are necessary when the attacker can launch an attack with substantial computational power.

C. Smart contract

A smart contract is a self-executing contract where the terms and conditions are directly written into code, enabling automatic execution without the need for intermediaries [19]. These contracts operate based on predefined conditions and rules built on a blockchain platform. The smart contract code automatically executes the contract conditions, ensuring transparency and accuracy throughout the contractual process. Smart contracts are recorded and verified on the blockchain, providing high security and resistance to tampering. They reduce the need for intermediaries, streamline processes, and automate transactions, potentially saving time and cost. The self-executing nature of smart contracts is widely utilized and researched in many fields that rely on data-driven transactions [20]. Increasingly, developers and researchers are actively involved in the development and exploration of smart contracts [21]–[23].

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on December 06,2023 at 13:19:58 UTC from IEEE Xplore. Restrictions apply.

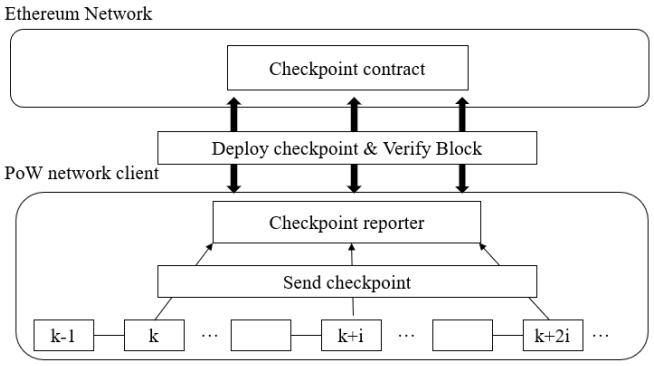


Fig. 1. The architecture of our method

III. Proposed method

In this section, we first present an architecture of our proposed checkpoint method, then introduce the details of the method.

A. Architecture

Fig. 1 depicts the architecture of our proposed method. The proposed method consists of two layers: the Ethereum network layer, which includes the checkpoint contract, and the PoW network layer, which the checkpoint aims to protect. Our method combines the Ethereum network layer with the PoW network layer to ensure the security of the PoW network. The checkpoint contract runs on the Ethereum network and serves as the administrator overseeing the management and execution of the checkpoint. The checkpoint contract defines the rules and conditions necessary to verify and record checkpoints on the Ethereum blockchain securely. The PoW network works similarly to a typical blockchain network but includes functionalities for generating, transmitting, and requesting checkpoints.

The PoW network validator generates a checkpoint when it reaches the checkpoint generation interval, which serves as a reference for the network's state. Checkpoint is nearly identical to the block header of the same block height, but they include additional data that marks them as a checkpoint. Checkpoints are generated every interval and sent to the Ethereum checkpoint contract. Participants in the network request the contract for a checkpoint when a block modification occurs such that the new block can be validated. Since checkpoint requests do not affect the state of the contract, they can be provided directly without submitting a transaction. The network client communicates with the checkpoint contract through the checkpoint reporter. The checkpoint reporter consists of an Ethereum light client that can communicate with Ethereum and is compatible with the same private key used by the PoW client.

Ethereum smart contract-based architecture provides the following advantages for the checkpoint. First, it mitigates the risk of checkpoint tampering. Ethereum is a network with many participants, and the likelihood of tampering is extremely low. Second, checkpoints operate transparently and according to predetermined rules, ensuring that checkpoint development remains decentralized and not centralized in the hands of checkpoint developers.

B. Checkpoint generation process

One of the things to consider when creating checkpoints is that Proof of Work networks operate as asynchronous distributed networks, which means that block propagation can be delayed or uncertain. We consider the asynchronous characteristics of the network during the checkpoint generation process. Despite a checkpoint created by one participant, other participants may need to be made aware of it and may still attempt to generate a checkpoint. We separate the checkpoint generation process into three timelines to reflect the network's asynchronous characteristics. Fig. 2 shows the checkpoint process timelines.

In Fig. 2, The first timeline occurs when the block height is at checkpoint time, defined as k. The decision maker creates block k but waits without creating a checkpoint. It is risky to trust a block that has just been created on an asynchronous network. Other blocks can easily replace blocks. Participants must wait until the network has transmitted block k to all participants.

The second timeline occurs at block height $k + i$, where i represents the checkpoint interval. We assume that i blocks have been generated from height k, so the propagation of block k is almost complete. The participant generates a checkpoint for block k and transmits it to the contract when the height $k + i$ has been achieved. The contract does not immediately complete the checkpoint operation for height k but instead stores the received checkpoints in a checkpoint pending pool and waits. The participant also performs the first timeline process at checkpoint $k + i$.

The third stage similarly occurs, following the completion of the second stage at height $k+i$. The second and third timelines i could differ according to the checkpoint interval adjustments described in the checkpoint interval section. However, we represent it as $k + 2i$ for ease of expression. The contract stores the checkpoints for height k in the checkpoint pending pool and waits. The first timeline for height $k + 2i$ and the second timeline for height $k + I$ execute when the network reaches height $k+2i$. During this process, participants transmit the checkpoints to contract for height $k+i$. The contract uses the checkpoint for height $k + i$ to recognize the block height, as it cannot identify the network's current block height. The contract finishes the vote for checkpoint k and decides the final checkpoint, k, after receiving the first checkpoint for height $k + i$. The final checkpoint is decided based on receiving the most votes. The checkpoint's recording guarantees the integrity of the blocks up to height k.

The actual network propagation time will probably be shorter than the checkpoint interval, but we await the propagation of the network until the next checkpoint interval. This wait causes multiple checkpoint processes to work simultaneously in an interval. Participants can merge transactions from

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on December 06,2023 at 13:19:58 UTC from IEEE Xplore. Restrictions apply.

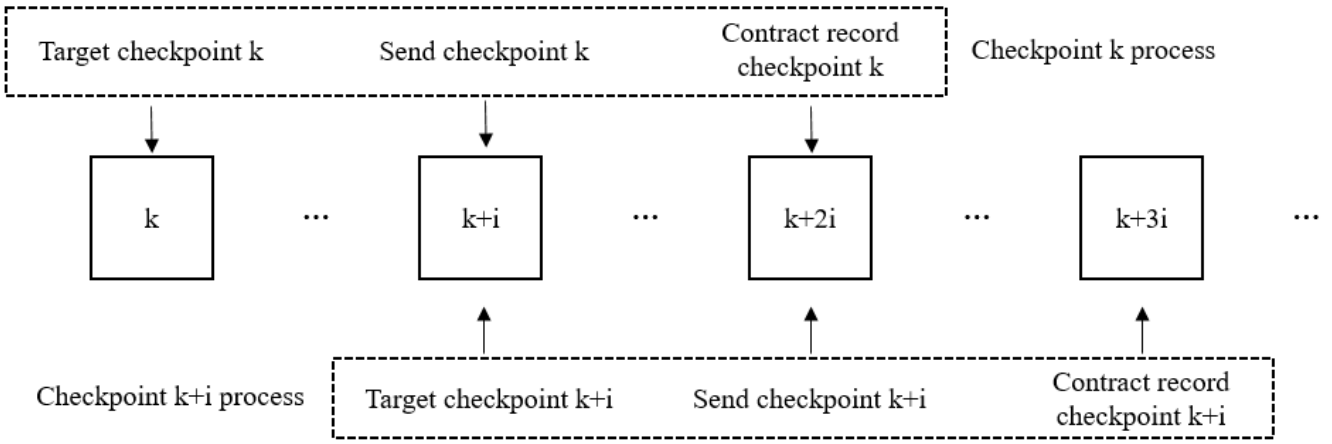


Fig. 2. Checkpoint process timeline

concurrently running checkpoints and send them in a single transaction. Checkpoint transaction merging reduces the cost of concurrently counting up to three processes by up to a third.

C. Checkpoint Smart contract

We establish a validator set that can generate checkpoints and manage the list in the contract. This list of validators forms a decentralized autonomous organization (DAO) that operates autonomously according to predefined rules within the contract. All validators have equal authority and participate in decision-making through voting. The contract records the checkpoint that receives the most votes from validators as the final checkpoint. Validator DAO can make decisions and execute processes without relying on centralized authority. The equal authority of the DAO ensures that decisions are made fairly and equitably.

The signatures of the validators determine the authenticity of a checkpoint transaction. We ensure the validators use the same private key for the PoW network and Ethereum. The validator signs the checkpoint with their private key and sends it to the contract. The contract verifies that the checkpoint's signature is included in the validators. Checkpoints with valid signatures are stored in the checkpoint pending pool. When a contract determines a checkpoint within a pending pool, the contract saves the checkpoint. The contract only stores the most recent checkpoint. The block generation process in PoW includes data from the previous block. The most recent checkpoint includes the previous one, which means the contract does not need to store the previous one. The contract calculates the attacker's computational power as the block difficulty for the spacing adjustment described in the following subsection. Algorithm 1 is a pseudo-code showing the process executed by the checkpoint contract when a checkpoint is transmitted.

Algorithm 1 Checkpoint contract process

function SUBMITCHECKPOINT(header , sig)
  $proposer, blocknumber, roothash \leftarrow header$
  if $proposer$ not in $validators$ **then**
    return $ValError$
  if $EcVerify(header, sig, proposer)$ is $false$ **then**
    return $SignError$
  if $CheckpointNumber < blocknumber$ **then**
    $checkpoint = PendingHeader.decideCheck()$
    $Interval \leftarrow CalNextInterval(checkpoint)$
  $PendingHeader.append(header)$
  return $(checkpoint, Interval)$

D. Checkpoint Interval

The smart contract calculates the checkpoint interval at the time of checkpoint creation. The checkpoint includes the block difficulty at the checkpoint's point in time, allowing us to assess the overall scale of the network through the block difficulty. Given the difficulty of block height k as $diff_k$ and block time as T_k, the total computational power of the network c_k is calculated as follows:

$$c_k = \frac{diff_k * 2^{32}}{t_k - t_{k-1}} \tag{1}$$

We find the attacker's computational power c_a to get the attacker's total computational power ratio. The c_a can be roughly gauged from the availability of computing resource rental platforms. For example, the Ethash computational power available for rent on nicehash.com in May 2023 is approximately 3 TH/s [24]. The attacker's total computational power ratio t can be expressed as follows:

$$t = \frac{c_a}{c_k} \tag{2}$$

Our checkpoints are confirmed through the three timelines and finalized two intervals later. A checkpoint interval of i will provide finality for a block up to $2i$ earlier. The confirmation distance z must be greater than $2i$ to be secure against double-spending attacks. The interval of checkpoints

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on December 06,2023 at 13:19:58 UTC from IEEE Xplore. Restrictions apply.

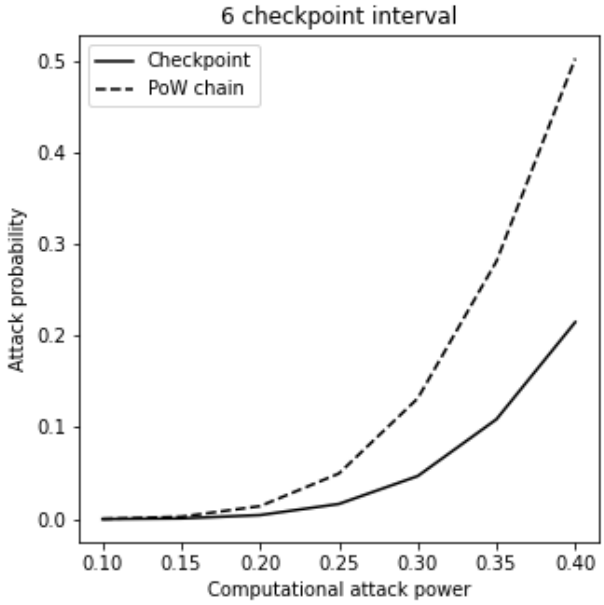


Fig. 3. probability of selfish mining attacks based on the attack power.

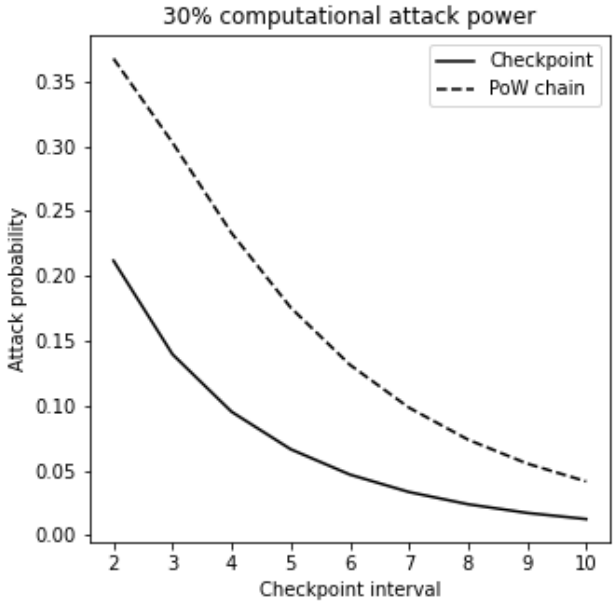


Fig. 4. probability of selfish mining attacks based on the checkpoint interval

significantly affects network performance. It provides significant completeness and security against short-term attacks but increases transaction costs.

Considering the cost of the attack, the network is sufficiently secure if the probability of a double-spend attack is less than 0.001. It means that the network is secure enough without checkpoints. We designed the interval of checkpoints to increase as the network expands. For every halving of the attack probability, the checkpoint interval doubles. This process repeats, gradually reducing the dependency on checkpoints. The attacker's probability p can be calculated from t obtained in Eq.(2) [15]. The network reaches a sufficiently large size to be secure against attacks. Given that the confirmation distance is z, and the attacker's attack probability is p, the interval i of the checkpoint can be obtained as follow:

$$i = \begin{cases} \frac{z}{2} & \text{if } p > 0.001 \\ 2^{\lceil \log_2(\frac{0.001}{p}) \rceil} \cdot z & \text{if } else \end{cases} \tag{3}$$

IV. Performance analysis

This section analyzes our proposed method's Selfish Mining Attack security performance. Our proposed method is resistant to double-spending attacks but only partially immune to selfish mining attacks. It is because Selfish Mining Attacks can occur even within confirmation distance. A typical PoW network has no checkpoints, so there is an infinite amount of time for a selfish mining attack to be attempted. In contrast to typical PoW networks, our method limits selfish mining attackers to only attacking until a checkpoint is generated.

The selfish mining attack scenario in our method is as follows. The attacker confirms that a checkpoint has been created and then prepares for a selfish mining attack before the next checkpoint. The attacker creates a block but does not broadcast it. The attacker creates and broadcasts the longest block before the checkpoint is created. Blocks and transactions created by other participants become invalid and monopolize the rewards of those blocks. The contract records the blocks generated by the attacker as checkpoints. This attack can harm network performance, but transactions are secure because the attacker's reward is limited to the mining revenue.

The probability of the selfish mining attack scenario is calculated similarly to the probability of a double-spend attack. We express the probability of a selfish attack scenario using a Poisson distribution. The attacker must create the longest chain and forge blocks before recording the checkpoint. We compute the probability distribution of blocks an attacker can generate during a checkpoint interval. We then calculate the total probability of the blocks generated by the attacker being more significant than the checkpoint interval. When the checkpoint interval is denoted as i, and the parameter λ is equal to $i * \frac{q}{1-q}$, the probability of the attacker's selfish mining attack can be computed as follows:

$$\text{attack}_i = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \cdot \begin{cases} 0 & \text{if } k \leq i \\ 1 & \text{if } k > i \end{cases} \tag{4}$$

We compare the performance of our method in the selfish mining attack scenario with a typical PoW network using Eq.(4). In our analysis, and we consider various attackers' computational power and the checkpoint interval. We conducted two experiments to evaluate the efficacy of checkpoints. In the first experiment, we set the checkpoint interval to 6 and analyzed the impact of varying the attacker's computational

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on December 06,2023 at 13:19:58 UTC from IEEE Xplore. Restrictions apply.

power on the success probability of attacks. The results of this experiment are presented in Fig. 3. It shows that our method exhibits a decrease in attack probability for all levels of attack power, with the reduction rate gradually increasing. It showcases a maximum reduction of approximately 60% in the attack probability. In the second experiment, we maintained the attacker's computational power at a fixed rate of 30% and observed the changes in the attack success probability by adjusting the checkpoint interval. Fig. 4 depicts the outcomes of this experiment. Similarly, Fig. 4 shows that our method reduces the attack success probability across all checkpoint intervals.

As a result, the two experiments show that checkpoint serves as an effective solution to prevent selfish mining attacks. The consistent decrease in attack probability for all situations is convincing evidence of the performance of our method.

V. Conclusion

In this paper, we present a smart contract-based checkpoint mechanism for Proof of Work (PoW) networks, addressing the challenges associated with network security and vulnerability to attacks. We record checkpoints in Ethereum smart contracts to reduce the probability of checkpoints being forged, as Ethereum is exceedingly difficult to forge. Since blockchain networks are asynchronous, we proposed a checkpoint generation process composed of three timelines that execute each checkpoint interval. To reduce checkpoint costs, We merge concurrently running processes and submit them as a single transaction. The checkpoint interval increases gradually as the PoW network expands. When the checkpoint interval is large enough, the network moves away from its dependence on checkpoints and becomes independent of Ethereum. We performed an attack probability analysis on a selfish mining attack to show the effectiveness of our method. The proposed checkpoint significantly lowers the probabilities of attacks, especially in the initial stages of PoW networks, compared to typical PoW networks. In our future work, we will apply our checkpoint method to existing blockchain networks and research methods to reduce the cost associated with checkpoints.

Acknowledgment

This work was supported by the MSIT, Korea, under the ITRC (Information Technology Research Center) support Program (IITP-2023-2021-0-01835) supervised by the IITP (Institute for Information & Communications Technology Planning Evaluation.)

References

[1] M. Platt, J. Sedlmeir, D. Platt, J. Xu, P. Tasca, N. Vadgama, and J. I. Ibañez, "The energy footprint of blockchain consensus mechanisms beyond proof-of-work," in *2021 IEEE 21st International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, 2021, pp. 1135–1144.

[2] S. Velliangiri and P. Karthikeyan, "Blockchain technology: Challenges and security issues in consensus algorithm," in *2020 International Conference on Computer Communication and Informatics (ICCCI)*, 2020, pp. 1–8.

[3] S. Chen, H. Mi, J. Ping, Z. Yan, Z. Shen, X. Liu, N. Zhang, Q. Xia, and C. Kang, "A blockchain consensus mechanism that uses proof of solution to optimize energy dispatch and trading," *Nature Energy*, vol. 7, no. 6, pp. 495–502, 2022.

[4] Y. Wang, H. Peng, Z. Su, T. H. Luan, A. Benslimane, and Y. Wu, "A platform-free proof of federated learning consensus mechanism for sustainable blockchains," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 12, pp. 3305–3324, 2022.

[5] "Proof-of-contribution consensus mechanism for blockchain and its application in intellectual property protection," *Information Processing Management*, vol. 58, no. 3, p. 102507, 2021.

[6] P. Fernando, K. Dadallage, T. Gamage, C. Seneviratne, A. Madanayake, and M. Liyanage, "Proof of sense: A novel consensus mechanism for spectrum misuse detection," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 12, pp. 9206–9216, 2022.

[7] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Decentralized business review*, p. 21260, 2008.

[8] E. Kapengut and B. Mizrach, "An event study of the ethereum transition to proof-of-stake," *Commodities*, vol. 2, no. 2, pp. 96–110, 2023. [Online]. Available: https://www.mdpi.com/2813-2432/2/2/6

[9] A. Gervais, G. O. Karame, K. Wüst, V. Glykantzis, H. Ritzdorf, and S. Capkun, "On the security and performance of proof of work blockchains," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 3–16. [Online]. Available: https://doi.org/10.1145/2976749.2978341

[10] "Bitcoin core 0.11.0," https://bitcoin.org/en/release/v0.11.0, 2015.

[11] "Matic whitepaper," https://github.com/maticnetwork/whitepaper/, 2020.

[12] K. Wang and H. S. Kim, "Reducing confirmation reversal probability of pow blockchains using checkpoints," in *2022 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, 2022, pp. 1–9.

[13] D. Karakostas and A. Kiayias, "Securing proof-of-work ledgers via checkpointing," in *2021 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, 2021, pp. 1–5.

[14] S. Sayeed and H. Marco-Gisbert, "Assessing blockchain consensus and security mechanisms against the 512019. [Online]. Available: https://www.mdpi.com/2076-3417/9/9/1788

[15] M. Rosenfeld, "Analysis of hashrate-based double spending," *arXiv preprint arXiv:1402.2009*, 2014.

[16] Q. Bai, X. Zhou, X. Wang, Y. Xu, X. Wang, and Q. Kong, "A deep dive into blockchain selfish mining," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, 2019, pp. 1–6.

[17] J. Jang and H.-N. Lee, "Profitable double-spending attacks," *Applied Sciences*, vol. 10, no. 23, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/23/8477

[18] Y. Kawase and S. Kasahara, "Transaction-confirmation time for bitcoin: A queueing analytical approach to blockchain mechanism," in *Queueing Theory and Network Applications*, W. Yue, Q.-L. Li, S. Jin, and Z. Ma, Eds. Cham: Springer International Publishing, 2017, pp. 75–88.

[19] N. Szabo, "Formalizing and securing relationships on public networks," *First monday*, 1997.

[20] W. Zou, D. Lo, P. S. Kochhar, X.-B. D. Le, X. Xia, Y. Feng, Z. Chen, and B. Xu, "Smart contract development: Challenges and opportunities," *IEEE Transactions on Software Engineering*, vol. 47, no. 10, pp. 2084–2106, 2021.

[21] S. Ahmadisheykhsarmast and R. Sonmez, "A smart contract system for security of payment of construction contracts," *Automation in Construction*, vol. 120, p. 103401, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092658052030981X

[22] A. Khatoon, "A blockchain-based smart contract system for healthcare management," *Electronics*, vol. 9, no. 1, 2020. [Online]. Available: https://www.mdpi.com/2079-9292/9/1/94

[23] H. Hasan, E. AlHadhrami, A. AlDhaheri, K. Salah, and R. Jayaraman, "Smart contract-based approach for efficient shipment management," *Computers Industrial Engineering*, vol. 136, pp. 149–159, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360835219304140

[24] "Nicehash marketplace," https://www.nicehash.com/marketplace, 2023.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on December 06,2023 at 13:19:58 UTC from IEEE Xplore. Restrictions apply.

지속 가능한 경제와 암호화폐

맹경덕○*, 김승민*, 이흥노*
광주과학기술원 전기전자컴퓨터공학부*
maengdeok@gm.gist.ac.kr, seungminkim@gm.gist.ac.kr, heungno@gist.ac.kr

초록

케인스주의, 신자유주의 그리고 실비오 게젤의 이론을 절충하여 새로운 통화 시스템을 제안한다. 현재의 통화 시스템은 중앙집권적이며 과도한 인플레이션과 이윤율 저하와 같은 구조적 한계를 가지고 있다. 통화정책의 중앙집권화와 불확실성을 제거하고 블록체인 기술을 통한 시스템의 보안과 신뢰성을 확보한다. 그리고 DAO와 SmartContract를 활용해 새로운 지식의 생산을 장려함으로써 전통적인 통화 시스템의 구조적 문제를 해소하고 지속 가능한 성장을 목표로 한다. 이 화폐 시스템은 또한 교환 기능을 중심으로 일관되고 투명한 통화팽창을 제공함으로써 안정성과 생산성 향상의 균형을 이룰 수 있음을 제시한다.

지속 가능한 경제와 암호화폐

맹경덕○*, 김승민*, 이흥노*
광주과학기술원 전기전자컴퓨터공학부*
maengdeok@gm.gist.ac.kr, seungmin@gm.gist.ac.kr, heungno@gist.ac.kr

1. 서론

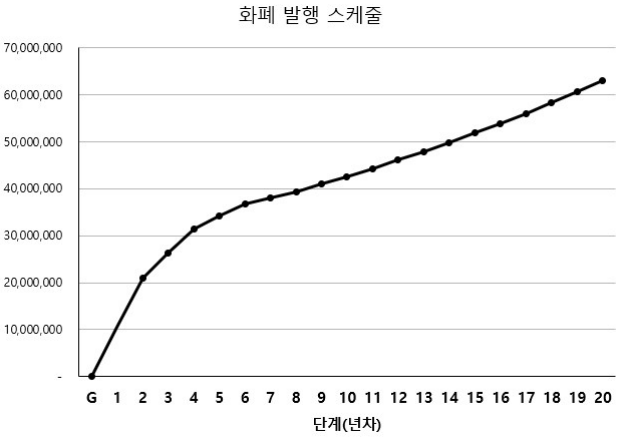
현대 통화 시스템은 케인스주의와 신자유주의 원칙의 혼합이다. 중앙은행들은 경제 성장을 견인하기 위한 케인스식 재정 자극과 함께 물가 안정을 유지하기 위한 감세, 규제 완화, 노동시장 유연화 등 하이에크식 통화 정책을 적절히 활용한다. 그러나 두 가지 접근 방식 모두 경제의 구조적 문제에 대한 일시적 완화만 제공했을 뿐 근본 원인을 해결하지 못하였다.

이러한 현대경제의 구조적 문제에 대한 대응으로 발생한 것이 Bitcoin이다. 2008년 경제 위기로 사람들은 중앙은행과 전통적인 금융기관에 대한 신뢰를 잃게 되었고, Satoshi Nakamoto는 어떠한 정부나 금융기관의 통제도 받지 않는 화폐인 Bitcoin를 탄생시켰다[1]. 이는 하이에크의 '화폐의 탈국가화' 와도 맥락을 같이하고 있다. 하이에크는 화폐가 정부 정책에 의해 결정되기보다, 금과 같은 Sound money로 존재해야 한다고 보았다. 이러한 Sound money는 정부나 중앙은행에서 가치를 조작할 수 없고, 수요와 공급에 따라 시장가치를 가지기 때문에 안정적으로 유지될 수 있다고 주장하였다[2].

현재 Bitcoin은 소수의 독점적인 권한 없이 탈중앙적으로 채굴, 유지되고 시장의 수급에 따라 그 가치가 결정되고 있다. 그러나, Bitcoin의 발행량은 제한되어 있고 반감기를 거치면서 생산량이 감소하는 구조로 설계되었기 때문에 장기적으로 그 가치가 계속해서 상승하게 된다. 이러한 구조적 특징으로 인해 Bitcoin을 소비하는 행동은 기회비용이 크고 Bitcoin을 소유한 사람들로 하여금 소비보다 저축에 경제적 유인을 제공한다. 이는 화폐의 유통 속도를 감소시키게 되고 경기순환에 경색을 야기함으로써 경기 침체를 유발할 수 있다. 결론적으로 Bitcoin은 장기적인 측면에서 교환의 목적으로 사용되기 어렵고 자산의 저축에 사용될 가능성이 크다. 따라서 본 연구에서 케인스와 게젤의 이론을 바탕으로 교환에 적합하며 건전한 소비와 생산을 장려할 수 있는 화폐 시스템을 제안한다.

2. 제안 모델

궁극적으로 본 연구의 목적은 케인스주의, 신자유주의 그리고 게젤 경제학의 핵심 원리와 블록체인 기술을 바탕으로 지속 가능한 화폐 시스템을 제안하는 것이다. 블록체인 기술의 활용은 보안성과 신뢰성의 확립을 가능하게 한다. 그리고 하이에크가 주창한 화폐의 탈 국가화를 실현시킴으로써 전통적인 통화 시스템의 중앙 집중적 성격으로 인한 구조적 한계를 해결할 수 있다.



[그림 1] 화폐 발행 스케줄

위의 [그림 1]은 본 연구에서 제안하는 지속 가능한 암호화폐의 발행 스케줄이다. 초기 8년간의 반감기 이후, 매년 4%의 고정된 통화팽창을 통해 예측 가능한 성장을 지향한다. 이로 인해 화폐의 액면 가치를 감소시키고 유통 속도의 증대를 불러옴으로써 케인스와 게젤의 관점을 종합한다. 이 통화 시스템은 저축의 기능이 강한 Bitcoin과 다르게 화폐의 교환 기능에 지향점을 두고 활발한 화폐 유통과 투명하게 이루어지는 점진적 통화팽창을 통해 생산성 향상과 안정성 간의 균형을 이루는 것을 목표로 한다. 이러한 제안에 대한 구체적인 근거들을 다음 목록에서 제시한다.

3. 화폐의 감가상각

케인스는 경제의 구조적 문제에 대해 유효수요 이론을 바탕으로 그의 해결책을 전개하였다. 그는 수요 부족을 근본적인 문제로 파악하고 지출 증가와 감세를 통한 정부 개입이 경제를 활성화할 수 있다고 주장하였다[3]. 반면 케인스의 이론이 낳게 될 결과를 앞서 예측한

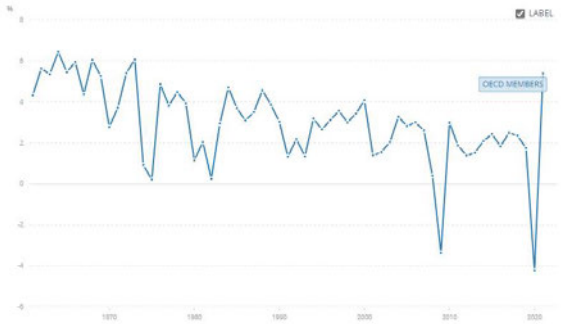
경제학자가 그 이전에 존재했다. 실비오 게젤은 이러한 접근법에 대해 화폐의 축적과 수요 부족은 원인이 아닌 하나의 결과이며, 정부의 화폐 공급정책으로 유효수요 부족 현상을 해소하는 것은 결국 더 많은 화폐의 발행, 과도한 축적 그리고 파괴적인 인플레이션을 야기할 수 있음을 경고하였다. 그는 화폐의 액면 가치가 불변하는 것이 모든 경제 문제의 근본적인 원인으로 보았고, 이러한 성질은 사람들로 하여금 화폐를 교환의 매개체로 사용하는 대신 화폐의 저축으로 유인한다고 주장하였다. 그리고 화폐를 저축의 대상으로 인식하는 경향은 수요 부족과 그에 따른 경기 침체로 이어진다고 지적하였다[4].

오늘날에 이르러서, 1970년대 스태그플레이션과 COVID-19 시기의 무제한 양적 완화로 인한 결과들은 게젤의 예측이 사실임을 증명하였다. 1970년대 동안 많은 국가들은 케인스의 이론을 바탕으로 경기 침체에 대응하기 위해 정부지출 증가와 통화팽창 정책을 시행하였다. 그러나 통화팽창과 수요 창출 정책은 중동발 오일쇼크와 이윤율 저하 등의 요인들에 의해 효과가 감소되었고, 인플레이션이 급증하는 동시에 성장 정체가 공존하는 스태그플레이션 현상이 발생하게 되었다. 이어서 2019년 COVID-19 범 유행 기간 동안 시행된 무제한 양적 완화는 단기적으로 경기를 회복시키고 소비시장을 활성화시킬 수 있었지만, 결국 소비자 물가 급등, 자산시장과 자원의 비생산적 투자 과열 그리고 정부의 부채 증가로 인한 재정 부실 등의 2차 문제를 초래함을 확인하였다.

이러한 사례들을 근거로 본 연구에서는 케인스의 접근법에서 중앙화적인 성질과 통화팽창률의 변동성을 제거하고 투명하고 점진적인 통화팽창정책으로 화폐의 액면 가치를 삭감한다. 이를 통해 화폐의 저축기능을 억제하고, 화폐가 교환의 목적으로 사용되는 것을 유도한다.

4. 예측 가능한 성장

투명한 고정 인플레이션 시스템 뒤에 있는 추론은 고용과 인플레이션의 관계에 뿌리를 두고 있다. 케인스 이론에 따르면 인플레이션은 실질 임금 감소로 인해 고용 증가와 실업 감소로 이어질 수 있다[2]. 그러나 노동자들이 실질 임금 감소에 대응하여 더 높은 임금을 요구함으로써 이러한 효과는 점차 감소하게 된다. 이 점은 로버트 루카스에 의해 증명되었는데, 노동자들은 더 높은 인플레이션을 기대하고 더 높은 임금을 요구하지만, 예상한 만큼 인플레이션율이 높지 않으면 실질 임금의 상승과 고용 감소를 초래할 수 있다[7]. 이러한 불확실성에 비추어 볼 때, 투명한 고정 인플레이션 정책이 변동 인플레이션 정책보다 더 지속 가능한 시스템이라는 것은 분명하다. 이에 본 연구에서는 연 4%의 고정된 통화팽창을 통해 예측 가능한 경제 성장을 설계하였다.



[그림 2] OECD 국가 연평균 성장률
(https://data.worldbank.org)

이러한 수치는 위의 [그림 2]를 바탕으로 산출하였다. [그림 2]는 OECD 국가의 연평균 GDP 성장률 통계치를 나타낸 것이다. OECD의 회원국들은 평균적으로 연간 4%의 성장을 달성하였다. 이러한 통계 데이터에 근거하여 일관되고 투명한 통화팽창을 제공하고 화폐의 액면가치를 감소시킨다. 이로 인해 화폐를 교환의 매개체로 사용함으로써 소비와 생산을 장려하고 자금경색과 금융시장의 버블을 해소할 수 있다.

5. 지속 가능한 성장

지속 가능한 성장은 '이윤율 저하를 동반하지 않는 생산성의 증대'에 부합하는 통화 시스템을 필요로 한다. 전통적인 경제 시스템에서 새롭게 발행된 화폐는 자산 저축의 용도로 사용되어 돈의 경색을 발생시키거나 생산에 적절히 투자될 경우에도 문제를 일으키게 된다. 마르크스의 이윤율 저하 법칙에 의하면 사회가 발전함에 따라 실물자본의 양의 증가하면서 잉여가치를 낮추게 되고 결국 실물자본의 이윤율이 감소하게 된다[5]. 이러한 현상에 의해 실물자본의 이윤율은 금융자본 이자율보다 낮아지게 되고 결과적으로 생산에 투자된 돈 또한 금융 시장으로 유입된다. 이러한 일련의 과정들을 거치면서 발생하는 금융 버블은 결국 금융 시장의 불안정뿐만 아니라 실물 자본시장의 침체를 동반하기 때문에 자본의 붕괴를 일으킨다. 본 연구에서는 새로운 지식의 생산에 경제적 유인을 발생시켜 생산성과 더불어 이윤율의 지수함수적 증대에 부합하는 통화시스템을 제안함으로써 그 구조적 한계점을 극복하고자 한다.

지식기반경제에서의 부의 창출은 새로운 지식의 생산을 통해 이루어지는데, 새로운 지식과 생산기술은 재화와 서비스의 비용을 감소시키는 동시에 품질과 이윤율을 지수함수적으로 증대시킴으로써 삶의 질을 향상시킨다. 이러한 성장을 육성하기 위해서 블록체인 상에서 탈중앙적으로 구현되는 DAO와 SmartContract 기술을 응용할 수 있다[6]. DAO의 탈중앙적 의사결정 기능과

SmartContract의 자동화 시스템을 통해 네트워크를 지속적으로 발전시키는 기여자에게 합당한 보상을 제공한다. 기여자는 dApp 개발, 네트워크 하드포크, 신규 유입에 대한 영향력 등의 일을 수행함으로써 이를 제공받을 수 있다.

연 4%의 고정된 통화팽창의 한도 내에서 이러한 지수함수적인 생산을 추구하는 합리적인 부의 분배는 돈이 축적, 경색되는 것을 제한하고 소비와 생산을 장려할 수 있다. 저축되어 소비와 생산에 기여하지 않는 돈은 통화팽창에 의해 점진적으로 그 가치를 잃어가지만, 소비되고 건전하게 투자되는 돈들은 새로운 가치를 창출한다. 그리고 이렇게 창출된 유의미한 가치와 부는 견고하고 지속 가능한 경제 발전을 가능하게 한다.

6. 결론

본 연구에서는 현재 통화시스템이 가지고 있는 한계점들을 명확히 하고 케인스, 하이에크 그리고 게젤의 경제학적 이론과 실제 데이터를 근거로 그 한계를 극복하고자 하였다. 이를 달성하기 위해 블록체인과 블록체인 상에서 구현할 수 있는 DAO, SmartContract 기술들을 기반으로 지속 가능한 화폐 체계를 제안하였다. 이 시스템은 블록체인 네트워크에 의해 보안과 탈중앙성 그리고 확장성을 보장받는다. 그리고 투명하고 예측가능한 통화 공급으로 화폐의 액면가치를 감소시킴으로써 유통 속도를 향상시키고 화폐의 교환 기능을 강화하여 지속적인 생산성 증대를 가능하게 할 것이다. 추후 본 연구의 결과를 적용시켜 발행한 암호화폐가 지속가능한 경제를 구축하기를 기대한다.

참고문헌

[1] Nakamoto, S., “Bitcoin white paper.” URL: https://bitcoin.org/bitcoin.pdf, 2008 Jan.

[2] Friedrich Hayek, The Denationalization of Money, Institute of Economic Affairs, 1976

[3] 존 메이너드 케인스, 고용, 이자, 화폐의 일반이론, 필맥, 2010

[4] 실비오 게젤, 자연스러운 경제질서, 클, 2021

[5] 카를 마르크스, 자본론, 비봉출판사, 2015

[6] Buterin, V., “Ethereum white paper”, GitHub repository. 2013 Jan.

[7] Lucas Jr, Robert E. "Expectations and the Neutrality of Money." Journal of economic theory, 1972

Gwangju Institute of
Science and Technology

School of Information and Communications



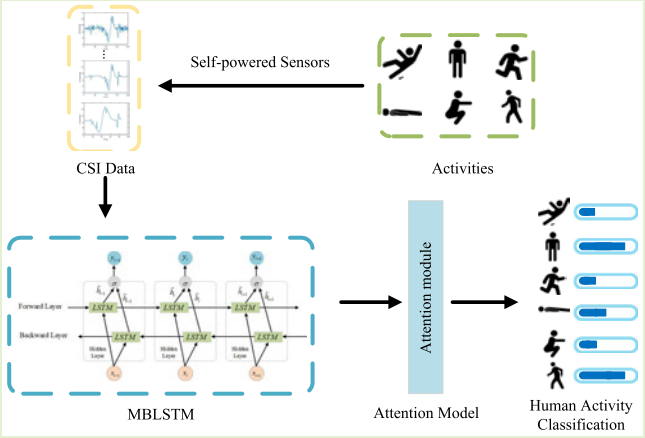
My AI Network™

- WorldLand, My AI Network 핵심 기술 연구 결과–

Human Activity Recognition Using Self-Powered Sensors Based on Multilayer Bidirectional Long Short-Term Memory Networks

Jian Su, *Member, IEEE*, Zhenlong Liao, Zhengguo Sheng, *Senior Member, IEEE*, Alex X. Liu, *Fellow, IEEE*, Dilbag Singh, and Heung-No Lee, *Senior Member, IEEE*

***Abstract*—Sensor-based human activity recognition (HAR) requires the acquisition of channel state information (CSI) data with time series based on sensors to predict human behavior. Many existing approaches are based on wearable sensors and cameras, which increases the burden and privacy issues for patients. Self-powered sensors are capable of noncontact collection of time series data generated by human activity while ensuring their own stable operation. In this article, we propose a deep-learning-based framework for contactless real-time activity detection of humans using self-powered sensors, which is called multilayer bidirectional long short-term memory (MBLSTM). The collected Wi-Fi CSI data are fed into our proposed network model, which is then used to learn representative features of both sides from the original continuous CSI measurements. The attention model is used to assign different weights to the learned features, and finally, activity recognition is performed. Experimental results show that our proposed method achieves an accuracy of more than 96% for the recognition of six activities in multiple rounds of testing, outperforming other benchmark methods used for comparison.**



***Index Terms*—Bidirectional long short-term memory (BLSTM), channel state information (CSI), deep learning, human activity recognition (HAR), self-powered sensors, Wi-Fi.**

Nomenclature

BLSTM	Bidirectional long short-term memory.
MBLSTM	Multilayer bidirectional long short-term memory.
CSI	Channel state information.
HAR	Human activity recognition.
RSS	Received signal strength.
APs	Access points.
KNN	k-nearest neighbor.
STFT	Short-time Fourier transform.
GAN	Generative adversarial network.
CNN	Convolutional neural network.
RF	Random forest.
SVM	Support vector machine.
LR	Logistic regression.
DT	Decision tree.
HMM	Hidden Markov model.
SAE	Sparse autoencoder.

Manuscript received 2 July 2022; revised 23 July 2022; accepted 23 July 2022. Date of publication 5 August 2022; date of current version 14 September 2023. This work was supported in part by the Natural Science Foundation of China under Grant 61802196, Grant 61872082, and Grant 61472184; in part by the Natural Science Foundation of Jiangsu Province under Grant BK20180791; in part by the Engineering Research Center of Digital Forensics, Ministry of Education; and in part by the National Research Foundation of Korea (NRF) funded by the Government of Korea (MSIP) under Grant NRF-2021R1A2B5B03002118. The associate editor coordinating the review of this article and approving it for publication was Dr. Shahid Mumtaz. *(Corresponding author: Jian Su.)*

Jian Su and Zhenlong Liao are with the School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044, China (e-mail: sj890718@gmail.com; liaozhenlong1105@163.com).

Zhengguo Sheng is with the Department of Engineering and Design, University of Sussex, Brighton BN1 9RH, U.K. (e-mail: z.sheng@sussex.ac.uk).

Alex X. Liu is with the Ant Financial Services Group, Hangzhou 310000, China (e-mail: alexliu360@gmail.com).

Dilbag Singh and Heung-No Lee are with the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea (e-mail: dilbagsingh@gist.ac.kr; heungno@gist.ac.kr).

Digital Object Identifier 10.1109/JSEN.2022.3195274

I. Introduction

In RECENT years, thanks to the rapid development of Internet-of-Things (IoT) technology, we can get a lot of useful information from different types of sensors in IoT. This information can help IoT technology to be applied in smart cities, smart farms, medical and health services, and so on. The application of IoT sensors in the livestock industry can help practitioners reduce costs and increase efficiency [1].

1558-1748 © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:49:05 UTC from IEEE Xplore. Restrictions apply.

Use fiber optic sensors for 3D sound source localization [2]. IoT technologies are also gradually entering our daily life and can identify human daily activities. HAR is also receiving increasing attention for research in the field of health detection. For example, we need to understand the health status of elderly people and need to monitor their daily activities [3] over time for fall detection [4] and identification of some diseases that the elderly are prone to, such as Parkinson's [5].

To identify various human activities, many methods have been used in previous work. Cameras [6]–[9], wearable sensors [10], [11], and radio frequency identification (RFID) [12]–[14] have been used for activity recognition. Camera-based systems have the advantage of being able to detect minor human movements. However, these systems face severe problems, such as object blocking and privacy issues. Because of the great recognition accuracy, wearable sensors are also useful in HAR [15]. Wearable sensor-based ones, on the other hand, need the use of additional devices for action recognition, which is both uncomfortable and ineffective. The mobile phone is another popular sensor for recognizing human activity. Smartphones may be considered electricity sensing platforms for HAR since different sensors, such as accelerometers, gyroscopes, and barometers, are incorporated in phones. If the user forgets to carry their smartphone, activity recognition will be turned off. Simultaneously, the operation of the sensors in the phone will be affected by its battery capacity. The usage of Wi-Fi devices for HAR has also been successful currently [16]–[19]. Wi-Fi provides new research directions for universal, nonvisual HAR due to its universality, low cost, and contactless operation. Use self-powered sensors to obtain stable and continuous Wi-Fi signal information.

The basic point of using Wi-Fi to recognize activities is that human motion influences the nearby Wi-Fi signal, and Wi-Fi signals reflected by different activities exhibit different characteristics. RSS, which is most widely practiced in the field of indoor positioning research [20], [21], is the most extensively utilized signal for Wi-Fi. Although it can be used to recognize human activity, it has disadvantages because of noise and unsteady RSS data. Distinguishing different human actions is mainly a matter of analyzing the pattern of the signal, CSI. The most advanced work showed pretty decent recognition accuracy while using a clean Wi-Fi channel in the experiment. However, in the real world, Wi-Fi channels are less than clean. Nowadays, wireless signals abound in indoor places, such as homes, offices, and supermarkets, and there are numerous private APs. Because most systems now utilize stationary Wi-Fi channels for action recognition and CSI acquisition, their performance is extremely vulnerable to cochannel interference, which can significantly decrease the quality of the receiver and distort the extracted recognition features. When classifying activities, traditional classification models utilized in present systems are highly influenced by such distortions. Recently, with the rapid development of deep learning techniques, the method of automatically learning activity features in CSI using deep learning has provided a completely new way of thinking for HAR [22]. There is also experience in combining machine learning and sensors in previous work, for example, fiber optic tactile sensors combined with machine learning algorithms for surface roughness recognition [23].

The advantage of long short-term memory (LSTM) networks to automatically learn meaningful features and encode data is widely used in deep learning. The traditional LSTM only handles the forward continuous CSI data, which means that the backward CSI data are not used effectively in training. Future CSI data, we believe, will be important for recognizing human activities. Furthermore, typical LSTM sequence properties may contribute differently to the HAR challenge. The learned characteristics, on the other hand, make an equivalent contribution to the final identification of human actions in the classic LSTM technique. We provide a MBLSTM based on Wi-Fi CSI data for HAR in this research paper. Stacking LSTM hidden layers gives more depth to the model and more accurate descriptions obtained as a deep learning technique while increasing the depth of the network, improving the efficiency of training, and obtaining higher accuracy. An MBLSTM network consisting of multiple forward and backward LSTM layers can handle both forward and backward continuous CSI measurements. Furthermore, the attention mechanism can give more weight to more essential characteristics and time steps, resulting in higher generalization for human activity detection. The effectiveness of the proposed personnel activity detection algorithm based on wireless CSI measurement is verified by real experiments. The results are compared to several published benchmark approaches.

In this article, the major contribution of our work is that we establish a framework called MBLSTM to recognize human activities. The following is a detailed description.

1) We designed an MBLSTM network to collect Wi-Fi CSI data for autonomous feature extraction and selection using self-powered sensors. Use self-powered sensors to continuously and steadily collect Wi-Fi time series information under different activities, and match this different information with different activities.
2) Continuous CSI data in both forward and reverse directions are processed by layering several BLSTM networks. The MBLSTM can simultaneously consider the information of different past and future actions in CSI data, thus bringing richer information reference for feature learning and using it can also speed up the convergence process of the training dataset.
3) The MBLSTM network uses an attention model to learn the relevance between activity features and time series. For the final personnel activity recognition, more main features and time series are assigned greater weights, resulting in improved recognition performance.

The rest of this article is organized as follows. Section II reviews some state-of-the-art work on using Wi-Fi signals to identify human activities, and Section III describes the channel sensing model and the MBLSTM network, as well as the proposed approach. Section IV describes the experimental setup and data. Then, this section shows and analyzes the experimental results. Finally, Section V summarizes this work.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:49:05 UTC from IEEE Xplore. Restrictions apply.

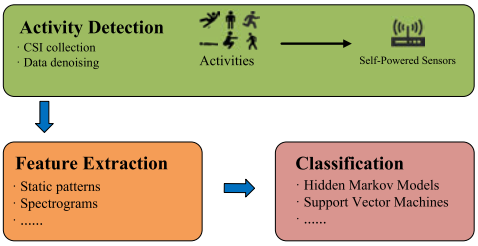


Fig. 1. Basic framework of the Wi-Fi-based activity recognition system.

II. Related Work

As illustrated in Fig. 1, a conventional Wi-Fi-based activity recognition system is composed of three parts.

1) *Filtering and Monitoring Channel Status:* Human activity is detected using a self-powered sensor. The human body activity affects the Wi-Fi signal, and this pair of signals can be observed. Therefore, the first step of the activity recognition system is to collect the original signal and denoise it to reveal the changes caused by human activity.
2) *Extraction of Features:* The CSI data from this denoising step are still not directly usable, and the next task is to discover and extract features from the existing behavioral data that are initially compatible with the technical requirements. At present, the signal feature extraction method contains the following three kinds: the time-domain analysis method, the frequency-domain analysis method, and the combined method of time–frequency analysis.
3) *Training and Recognition:* After getting the feature dataset, the first operation is to distinguish the dataset into the training set and the test set, and choosing the proper division ratio is a key part to ensure the effect of behavior recognition. The next step is to select the appropriate classification algorithm to train and test the data.

Due to the common presence of Wi-Fi in everyday life, many research teams have developed several activity recognition systems using Wi-Fi signals. Sigg *et al.* [24] proposed a wireless HAR system that analyzes the RSS information of the interfered Wi-Fi signal for activity recognition. They extracted several important features from RSS data and used a KNN classifier to recognize four daily activities. Abudulaziz Ali Haseeb and Parasuraman [25] designed an RSS-based gesture recognition system on cell phones. The system uses deep learning networks for gesture recognition and achieves high recognition accuracy. Due to multipath and fading effects, the collection of raw RSS data containing actions can be unstable and noisy, so the performance of using RSS to recognize activities with actions is very limited, even for simple actions. Wi-Fi's more steady and informative CSI has received a lot of attention recently. Zhang *et al.* [26] investigated the sensitivity of Wi-Fi signals theoretically and proposed a Fresnel zone method to recognize human activity using Wi-Fi CSI data.

Some special features may need to be carefully designed using domain knowledge in order to recognize certain actions using Wi-Fi CSI measurements. When used to recognize other activities, these features may not perform effectively. For example, the traditional KNN method, which has a simple idea, is applicable to multiclassification problems. However, when the sample distribution is unbalanced, the new sample will be classified as the dominant sample, so it cannot better approximate the actual classification result. Furthermore, handcrafted characteristics will gradually lose several of the implicit qualities that are important for recognizing human activity. Deep learning is a useful tool for automatically learning the differentiating features that are used to recognize the human activity.

Deep learning is a type of machine learning method that uses a deep neural network to classify data. In most cases, accurate features need to be identified for input to the training model, and the model classifies and outputs results based on these features. As a result, well-designed features are essential for accurate behavior recognition and have a significant effect on classification accuracy. Some feature extraction, on the other hand, may depend on empirical experience, lowering classification accuracy. Deep learning, unlike machine learning, generally does not require feature extraction stages since a deep neural network is able to automatically identify and extract features from training data. Deep learning allows us a new method to classify data and can deal with enormous amounts of data. In other words, the most significant advantage of deep learning is that it does not require preprocessing of data in order to obtain data features. Meanwhile, deep learning can automatically compute large-scale unknown parameters in neural networks through the training process. Usually, the process of neural network training consumes a lot of practice, but the results achieved are satisfactory. Deep learning algorithms are widely applied in various fields, including picture target identification, natural language processing, video classification, visual arts, and so on [27].

Damodaran *et al.* [28] used a device-free approach (CSI) to identify human activities. Wavelet analysis was used for preprocessing and feature extraction. As a result, they were able to recognize walking, sitting, standing, and running activities. High-bandwidth noise was removed using the principal component analysis by Moshiri *et al.* [29]. The signal was transformed to the frequency domain using STFT, and new data were generated using GANs. The LSTM algorithm was used for classification. The accuracy was 87.2% using 50% of the "real" data plus 50% of the synthetic data and 92.8% using a set of all "real" data.

CNN is a very popular deep learning method that automates feature extraction and can easily handle high-latitude data. However, when the network level is too deep, modifying the parameters using BP propagation will cause the parameters near the input layer to change more slowly, and the pooling layer will lose a lot of valuable information and ignore the local-to-whole correlation.

Since, for different activities, CSI measurements are continuous measurements with temporal information, BLSTM

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:49:05 UTC from IEEE Xplore. Restrictions apply.

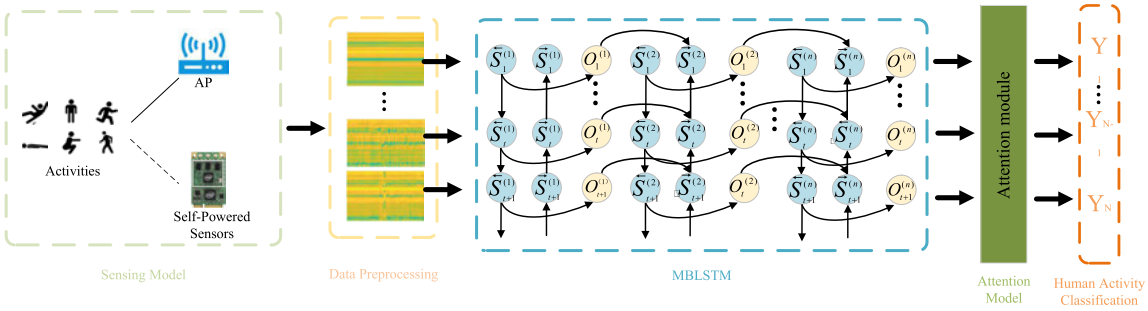


Fig. 2. Proposed MBLSTM framework for CSI-based HAR.

capable of encoding temporal information is a good candidate for automatic feature learning. BLSTM includes both forward and backward processes of feature learning. As a result, when evaluating the current hidden state of the LSTM, BLSTM can take into account both past and future information, resulting in richer information features. We propose stacked multilayer BLSTM networks for human action recognition. Each layer of the BLSTM neural network automatically learns the input action features and passes the learned features to the next layer. At the same time, the feature sequences learned in one temporal instance may contribute differently to the final HAR. Furthermore, the significance of CSI collected at different time stages may differ. Therefore, in order to assign different weights to different action features in the training for the purpose of reducing the training time and improving the accuracy of the model, we add an attention mechanism to the proposed network model.

III. Proposed Method

A. System Overview

The proposed MBLSTM framework is shown in Fig. 2. First, we use a router and a self-powered sensor to collect CSI signals from Wi-Fi of human actions. Second, we input the processed CSI signals into the MBLSTM framework to automatically learn the forward and backward features. There are 200 hidden nodes in the bidirectional LSTM used for feature learning in this experiment. Since the attention model has no available prior information, it can only use the features learned from BLSTM as input to derive an attention matrix representing the importance of features and time steps. Then, we use element multiplication to merge the learned features with the attention matrix to obtain the modified feature matrix with attention. After that, the feature matrix is flattened into feature vectors for final classification using the flattened layer. Finally, the softmax classification layer is used to identify different activities with the final feature vectors.

B. Channel Sensing Model

Wi-Fi signals are known to fluctuate significantly when objects move within the region of interest. The Fresnel zone model is introduced as a result of this to explore how the Wi-Fi signals on these receiving antennas change as a result of different activities. Furthermore, we infer potential behavioral information from such activity-induced signal fluctuations. Thus, we use the Intel 5300 NIC, a self-powered sensor, to collect the reflected Wi-Fi information.

In recent years, the Fresnel zone model has been applied to the research of human action recognition based on wireless sensing. It refers to the wireless electromagnetic wave in the transmission process, the formation of the transceiver at both ends of the transceiver device, and the location of the transceiver device as the focus of the ellipse-shaped area; the area is the wireless electromagnetic wave intensity concentration area. One of the most important zones is the first Fresnel zone, where most of the energy of the wireless signal is located. If there is an obstacle in this region, it will affect the wireless signal. The wireless signal will form multiple propagation paths from the receiver (Rx) to the transmitter (Tx), and the direct propagation path that passes through both the transmitter and receiver is called the line-of-sight (LoS) path. When the wireless signal transmission encounters an obstacle, the transmission path produces reflection, scattering and diffraction, called non-line-of-sight (NLoS).

Through the analysis and study of the received signals, the researchers found the characteristics of the changes brought by the human body movements on the signal propagation. Meanwhile, establish the relationship between these features and the mapping of different activities, which built the foundation for Wi-Fi-based HAR.

The phenomenon that different actions have different effects on Wi-Fi signals is a major discovery that the Fresnel zone model can be applied to the field of action recognition. Specifically, different activities cause significant differences in the speed of signal dynamic paths. Furthermore, CSI's amplitude attenuation and phase change can capture these specific patterns. It demonstrates the feasibility and application of using unique CSI variations to effectively and precisely identify and recognize different human activities.

C. MBLSTM Neural Network

In the case of multilayer stacking, each layer of the BLSTM neural network is composed of a forward recurrent network and a backward recurrent network. The combination of the output results of the forward LSTM and the backward LSTM of the previous layer is sent to the next layer of the network. Fig. 3 illustrates the MBLSTM

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:49:05 UTC from IEEE Xplore. Restrictions apply.

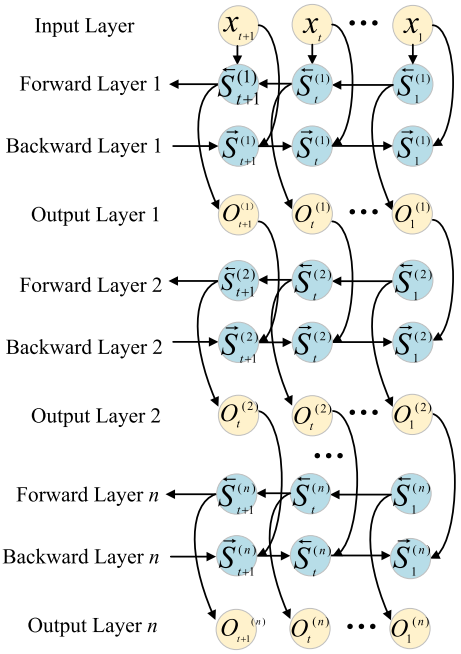


Fig. 3. Structure of the MBLSTM neural network.

network framework structure

$$\begin{aligned} S_t'^{(1)} &= f\left(U'^{(1)}x_t + W^{(1)}S'_{t-1}\right) \\ S_t^{(1)} &= f\left(U^{(1)}x_t + W^{(1)}S_{t-1}\right) \\ &\cdots \\ S_t'^{(i)} &= f\left(U'^{(i)}S_t^{(i-1)} + W'^{(i)}S'_{t+1}\right) \\ S_t^{(i)} &= f\left(U^{(i)}S_t^{(i-1)} + W^{(i)}S_{t-1}\right) \\ \mathrm{O}_t &= g\left(V^{(i)}S_t^{(i)} + V'^{(i)}S_t'^{(i)}\right). \end{aligned} \tag{1}$$

The output is determined by the sum of each layer's positive and negative computations. Here, $S_{t-1}^{(i)}$ and $S_t^{(i)}$ are the values of the ith hidden layer at times $t-1$ and t, respectively. Forward and backward computations do not share weights; $V^{(i)}$, $U^{(i)}$, and $W^{(i)}$ are the weight matrices of the ith hidden layer to the output layer, the input layer to the hidden layer, and the hidden layer. $V'^{(i)}$, $U'^{(i)}$, and $W'^{(i)}$ are the backward weight matrices used for the computations. i is the number of BLSTM layers, and $i = 0, 1, 2, \ldots, \infty$ is the output layer's value.

D. Attention Model

The attentional model is developed to be used for image recognition [30]. The concept was inspired by the human visual system, which says that, during picture recognition, humans always focus on a certain portion of the image and adjust their attention over time. During the recognition work, the attention model allows the computer to attend to the area of interest while blurring other areas. Recently, attention models have been used in language processing, proving that it is clearly effective [31]. For example, in the popular encoder-decoder method for natural language processing, the input sentence is encoded as a fixed vector that is translated throughout the translation process, meaning that, at every time step, all words in the input sentence contribute equally to the translation. This task of processing sentence translation is inefficient. When the encoder model is utilized with the attention model, translations will focus more on the words that are more relevant to the current translation process at different time steps. Since the MBLSTM network learns high-dimensional sequence features, individual features and time series may contribute differently to the final recognition results. We try to use an attention model to intelligently learn the effects of different actions of features and assign weights according to their importance.

In the recognition system, there is no usable *a priori* information for training. As a result, the attention model, also known as self-attention, will utilize the sequential features learned by MBLSTM as input. This attention model is shown in a simple example here. Given n feature vectors $\mathbf{h}_i, i = 1, 2, \ldots, n$ that can be obtained from the feature learning network, we build a score function $\Phi(\cdot)$ to evaluate the significance of each feature vector by computing the score s_i as follows:

$$s_i = \Phi\left(\mathbf{W}^T\mathbf{h}_i + b\right) \tag{2}$$

where $\mathbf{W}^T$ and b are the weight vector and bias, respectively. Any activation function in a neural network, such as tanh, relu, or linear, can be used to build the score function. We can normalize each feature vector's score utilizing the softmax function, which is written as

$$a_i = \mathrm{softmax}(s_i) = \frac{\mathrm{esp}(s_i)}{\sum_i(s_i)}. \tag{3}$$

The final output feature $\mathbf{O}$ of the attention model is the product of the vector and its normalization score as follows:

$$\mathbf{O} = \sum_{i=1}^{n} a_i * \mathbf{h}_i. \tag{4}$$

E. Training Proposed Method

To identify the model parameters, the proposed MBLSTM framework is trained using CSI data with real labels. First, all parameters are randomly given. The CSI data is then sent into MBLSTM, which uses it to predict the labels. The category cross-entropy errors are measured and backpropagated using a gradient-based optimization approach to adjust the model parameters utilizing the given true labels. We utilize ADAM [32] to calculate the adaptive learning rate for each parameter in the optimization process efficiently.

In learning-based systems, overfitting is a typical problem. To avoid overfitting, we utilize the ADAM optimizer. It provides adaptive learning rates for various parameters. Furthermore, the suggested attention method will only choose a few significant features and time series, decreasing the possibility of overfitting.

IV. Experiments and Analysis

In this section, we first introduce our experimental settings in detail and then present the extensive experimental results that validate the effectiveness of our model.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:49:05 UTC from IEEE Xplore. Restrictions apply.

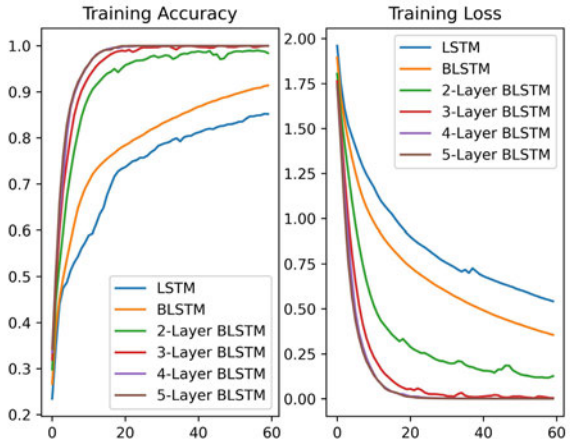


Fig. 4. Trend of accuracy rate of different network trainings.

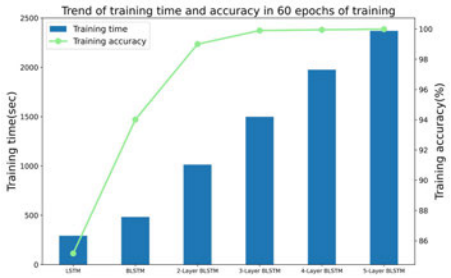


Fig. 5. Trend of training time and accuracy in 60 epochs of training.

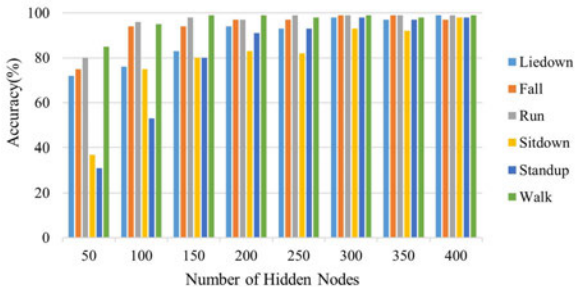


Fig. 6. Recognition accuracy of each activity with different numbers of hidden nodes.

A. Experiments' Settings

We compared the proposed method to several benchmark CSI-based human activity identification algorithms to evaluate how effective it is. According to Yousefi *et al.* [33], the RF model outperforms SVMs, LR, and DT in Wi-Fi-based HAR. In [34], HMMs have also been found to be useful for recognizing human activity. As a result, we compared our method to these two handmade methods. Manual feature extraction is described in detail in [33]. We also compare it to other deep-learning-based approaches that can learn features automatically, such as SAE [34], [35] and traditional LSTM [33]. The SAE algorithm is an unsupervised algorithm that automatically learns features from unlabeled data and can give a better feature description than the original data. Validation sets from the training examples were used to fine-tune the parameters of all methods. For evaluation, tenfold cross-validation has been used. We divided all of the data tenfold at random. Then, we select onefold of data for testing and the rest for training and, finally, get ten times. The average of all ten runs determines the final recognition accuracy. The dataset used for comparison was taken by Yousefi *et al.* [33] from an office. A router was used as a transmitter, and a laptop with an Intel 5300 NIC was used as a receiver. The sampling frequency was 1 kHz, with three antennas and 30 subcarriers, and the size of the original CSI data was 90. The window size used for data segmentation was a sliding window of 2 s. Transmitters and receivers were separated by three meters under LOS conditions. Each person performed each activity for 20 s during data collection. Note that the person remains stationary at the beginning and end of the activity. Six persons were involved in the data process of collecting, which included six normal daily activities: lie down, fall, run, sit down, stand up and walk. Every volunteer performed 20 rounds of each activity; the resulting dataset was approximately 17 GB in size. All experiments were performed on a workstation in our lab using python to run the code. The workstation is equipped with an eight-core, 16-thread Intel i9-9900 CPU and an NVIDIA GeForce RTX 2080 GPU.

We compare the trend of accuracy and loss of BLSTM networks with the different number of layers in the training dataset. Fig. 4 shows that the LSTM and BLSTM networks converge more slowly, with accuracy barely reaching 90% at the 60th round of training. The multilayer BLSTM network, on the other hand, converges quickly, with accuracy exceeding 90% at about ten rounds of training, approaching 100% at close to 20 rounds, and preserving stability in accuracy during subsequent training.

Although the training converges faster as the number of BLSTM layers increases, it is not better to have more layers. As the number of layers increases, the network structure becomes increasingly large, which means that more and more computational resources will be used, and more time will be consumed in training. As shown in Fig. 5, we run experiments using 200 hidden nodes. The results show that the more complex the network structure is, the time for training increases significantly. It is obvious that, with the same number of training rounds, the overall training accuracy does not improve much after increasing the BLSTM network to three layers, which are close to 100%, indicating that the limit has been approached. However, the training time spent by each network differs greatly. Considering all factors, we choose the three-layer BLSTM network as the network model for this experiment in order to minimize the computer resources consumed while ensuring high accuracy.

1) Impact of the Number of Hidden Nodes: We find that the number of LSTM hidden nodes has a large impact on the experimental results. As a result, we performed a second experiment to see how this parameter affected the accuracy of activity recognition. The results of the experiment are shown in Fig. 6. When using 50 hidden nodes, the recognition accuracy was low for actions, especially for the two activities "sit down" and "stand up," which, we guess, are too similar. When the number of hidden nodes is raised, the recognition performance

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:49:05 UTC from IEEE Xplore. Restrictions apply.

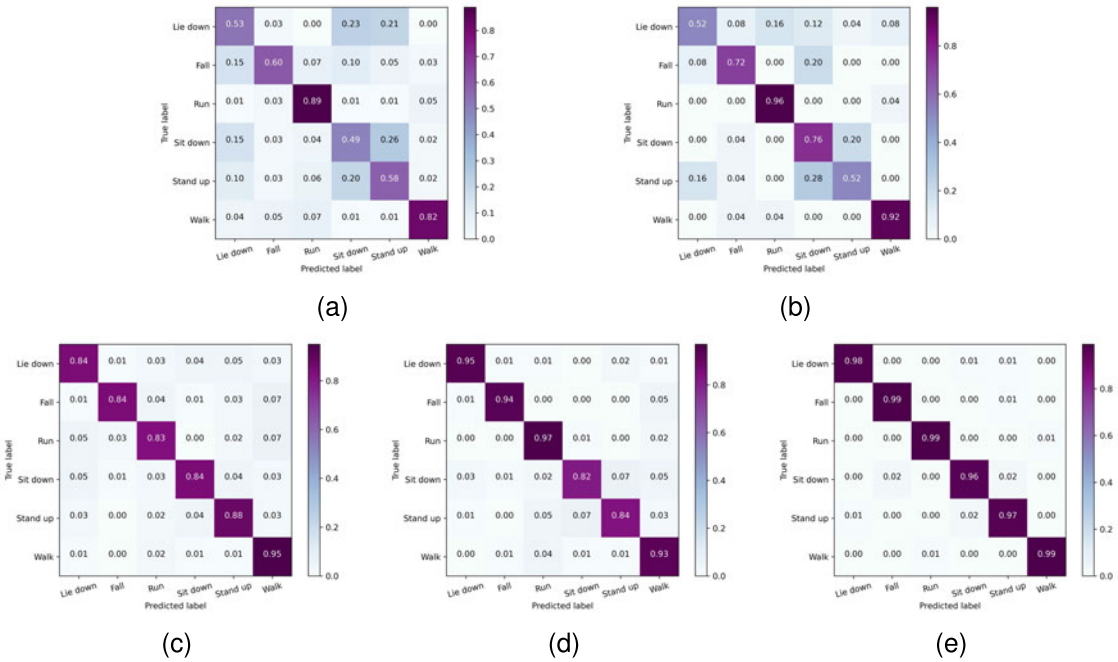


Fig. 7. Confusion matrix of all benchmarks and proposed MBLSTM methods on the dataset. (a) RF. (b) HMM. (c) SAE. (d) LSTM. (e) MBLSTM.

TABLE I
TRAINING AND VALIDATION TIMES FOR DIFFERENT NUMBERS OF HIDDEN NODES

Time \ Nodes	50	100	150	200	250	300	350	400
Training(s)	245.85	349.14	545.11	785.69	880.83	1043.76	1238.80	1444.24
Validation(s)	7.63	7.46	7.04	7.64	7.21	7.24	7.86	8.35

TABLE II
TRAINING AND TESTING TIMES FOR DIFFERENT METHODS

Time	RF	HMM	SAE	LSTM	MBLSTM
Training(s)	5.31	0.024	158.16	493.23	1551.14
Testing(s)	0.008	0.17	0.19	3.54	8.72

of each activity is improved, and after the number reaches 300, the accuracy tends to be stable. As shown in Table I, we use a three-layer BLSTM network, and in the same 30 rounds of training, the more the hidden nodes, the longer the training time, and we choose to use 200 hidden nodes in the MBLSTM.

2) Time Complexity: Deep-learning-based approaches' time complexity is a common issue. We compared the training time and testing time of some methods using the same dataset. Table II shows the training time and testing time for all methods. It can be clearly seen that algorithms using deep learning methods have much longer training times than a typical machine learning algorithm. The proposed MBLSTM consumes the longest training time of all methods using deep learning. All of the approaches have short testing times according to Table II. The proposed MBLSTM, for example, has a testing time of 8.72 s for 420 test samples. This signifies that each sample will be tested for 0.0208 s. The window size for data segmentation is 4 s for each case. We believe that our proposed MBLSTM approach, which is based on Wi-Fi CSI, may be utilized for real-time personnel activity recognition.

B. Experimental Results

Fig. 7 shows the confusion matrix of all benchmarks and proposed MBLSTM methods on the dataset. Activity recognition algorithms that need manual feature extraction, such as RF and HMM, perform the poorest. The HMM algorithm performs significantly better than the RF algorithm. Unlike RF and HMM manual feature extraction, SAE algorithms using deep learning methods have better performance. This demonstrates the effectiveness of using the SAE method for automatic feature learning. The LSTM network outperforms the SAE method because it incorporates the temporal factors in the CSI sequences into feature learning. Due to the inclusion of the attention model and the structure of the multilayer bidirectional LSTM in our proposed method, our MBLSTM method achieves excellent recognition results in recognizing six daily activities. For all six daily activity recognition, the accuracy is greater than or equal to 96%, which is sufficient for most recognition situations.

The accuracy of recognition varies greatly depending on the activity. Higher physical activities, such as "fall," "walk," and "run," show greater recognition performance. This is due to the fact that these activities have a large impact on the features of the collected CSI data. It is also evident that most methods have relatively low accuracy for recognizing the activity of "sit down." This might be because this activity has the same effect on CSI features as the "lie down" and "stand up" activities. It is worth noting that the RF method's recognition accuracy with handmade features is much lower than 50%. The "fall" activity is the most important of these six, especially for the elderly [36]. The proposed MBLSTM approach can

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:49:05 UTC from IEEE Xplore. Restrictions apply.

recognize "fall" activities with 99% accuracy, which will be useful in a wide variety of medical applications. The extended training period of the deep-learning-based approach is one of its drawbacks. However, this time-consuming procedure only has to be completed once. It is worth noting that deep-learning-based methods can be tested online quickly enough for most real-time applications.

V. Conclusion

In this article, we use self-powered sensors to collect Wi-Fi time series information and propose a multilayer BLSTM network for extracting Wi-Fi signal feature information used for HAR by improving the traditional LSTM model. In both directions, the BLSTM network can learn important sequential features from original Wi-Fi CSI data. The multilayer BLSTM network can enhance accuracy by accelerating convergence during training. We evaluated the method in real environments and compared it to a variety of benchmark methods, such as RF, HMMs, SAEs, and traditional LSTM. The proposed MBLSTM for Wi-Fi CSI-based personnel activity recognition has demonstrated higher performance in experiments. Although our method has a high recognition rate for single-person activities, there is still a big room for improvement in multiperson activities. For future work, we hope to improve the accuracy of multiperson activity recognition and the compatibility of the system with different environments. Glowinski *et al.* [37] used an inertial measurement device to calculate acceleration. This inspired our proposed method helps to recognize the type of body movement. In case of a car accident, it can help to determine the posture of the injured person.

References

[1] S. Neethirajan, "The role of sensors, big data and machine learning in modern animal farming," *Sens. Bio-Sens. Res.*, vol. 29, Aug. 2020, Art. no. 100367.

[2] S. E. Hayber and S. Keser, "3D sound source localization with fiber optic sensor array based on genetic algorithm," *Opt. Fiber Technol.*, vol. 57, Jul. 2020, Art. no. 102229. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1068520020302194

[3] T. Huynh-The, C.-H. Hua, N. A. Tu, and D.-S. Kim, "Physical activity recognition with statistical-deep fusion model using multiple sensory data for smart health," *IEEE Internet Things J.*, vol. 8, no. 3, pp. 1533–1543, Feb. 2021.

[4] Y. Wang, S. Yang, F. Li, Y. Wu, and Y. Wang, "FallViewer: A fine-grained indoor fall detection system with ubiquitous Wi-Fi devices," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 12455–12466, Aug. 2021.

[5] X. Cao *et al.*, "Video based shuffling step detection for parkinsonian patients using 3D convolution," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 641–649, 2021.

[6] J. K. Aggarwal and M. S. Ryoo, "Human activity analysis: A review," *ACM Comput. Surv.*, vol. 43, no. 3, pp. 1–43, Apr. 2011, doi: 10.1145/1922649.1922653.

[7] S. Jiang, Y. Qi, H. Zhang, Z. Bai, X. Lu, and P. Wang, "D3D: Dual 3-D convolutional network for real-time action recognition," *IEEE Trans. Ind. Informat.*, vol. 17, no. 7, pp. 4584–4593, Jul. 2021.

[8] D. C. Luvizon, D. Picard, and H. Tabia, "Multi-task deep learning for real-time 3D human pose estimation and action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 8, pp. 2752–2764, Aug. 2021.

[9] M. Moniruzzaman, Z. Yin, Z. He, R. Qin, and M. C. Leu, "Human action recognition by discriminative feature pooling and video segment attention model," *IEEE Trans. Multimedia*, vol. 24, pp. 689–701, 2022.

[10] L. Tong, H. Ma, Q. Lin, J. He, and L. Peng, "A novel deep learning Bi-GRU-I model for real-time human activity recognition using inertial sensors," *IEEE Sensors J.*, vol. 22, no. 6, pp. 6164–6174, Mar. 2022.

[11] D. Chen, S. Yongchareon, E. M.-K. Lai, J. Yu, and Q. Z. Sheng, "Hybrid fuzzy C-means CPD-based segmentation for improving sensor-based multiresident activity recognition," *IEEE Internet Things J.*, vol. 8, no. 14, pp. 11193–11207, Jul. 2021.

[12] X. Fan, W. Gong, and J. Liu, "TagFree activity identification with RFIDs," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 1, pp. 1–23, Mar. 2018, doi: 10.1145/3191739.

[13] G. A. Oguntala, Y. F. Hu, A. A. S. Alabdullah, R. A. Abd-Alhameed, M. Ali, and D. K. Luong, "Passive RFID module with LSTM recurrent neural network activity classification algorithm for ambient-assisted living," *IEEE Internet Things J.*, vol. 8, no. 13, pp. 10953–10962, Jul. 2021.

[14] F. Wang, J. Liu, and W. Gong, "Multi-adversarial in-car activity recognition using RFIDs," *IEEE Trans. Mobile Comput.*, vol. 20, no. 6, pp. 2224–2237, Jun. 2021.

[15] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1192–1209, 3rd Quart., 2013.

[16] J. Zhang *et al.*, "Data augmentation and dense-LSTM for human activity recognition using WiFi signal," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4628–4641, Mar. 2021.

[17] C. Xiao, Y. Lei, Y. Ma, F. Zhou, and Z. Qin, "DeepSeg: Deep-learning-based activity segmentation framework for activity recognition using WiFi," *IEEE Internet Things J.*, vol. 8, no. 7, pp. 5669–5681, Apr. 2021.

[18] R. H. Venkatnarayan, S. Mahmood, and M. Shahzad, "WiFi based multi-user gesture recognition," *IEEE Trans. Mobile Comput.*, vol. 20, no. 3, pp. 1242–1256, Mar. 2021.

[19] J. Zuo, X. Zhu, Y. Peng, Z. Zhao, X. Wei, and X. Wang, "A new method of posture recognition based on WiFi signal," *IEEE Commun. Lett.*, vol. 25, no. 8, pp. 2564–2568, Aug. 2021.

[20] C.-H. Lin *et al.*, "An indoor positioning algorithm based on fingerprint and mobility prediction in RSS fluctuation-prone WLANs," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 5, pp. 2926–2936, May 2021.

[21] D. Yu and C. Li, "An accurate WiFi indoor positioning algorithm for complex pedestrian environments," *IEEE Sensors J.*, vol. 21, no. 21, pp. 24440–24452, Nov. 2021.

[22] Z. Wang *et al.*, "A survey on human behavior recognition using channel state information," *IEEE Access*, vol. 7, pp. 155986–156024, 2019.

[23] S. Keser and Ş. E. Hayber, "Fiber optic tactile sensor for surface roughness recognition by machine learning algorithms," *Sens. Actuators A, Phys.*, vol. 332, Dec. 2021, Art. no. 113071.

[24] S. Sigg, S. Shi, F. Buesching, Y. Ji, and L. Wolf, "Leveraging RF-channel fluctuation for activity recognition: Active and passive systems, continuous and RSSI-based signal features," in *Proc. Int. Conf. Adv. Mobile Comput. Multimedia (MoMM)*. New York, NY, USA: Association for Computing Machinery, 2013, pp. 43–52, doi: 10.1145/2536853.2536873.

[25] M. Abudulaziz Ali Haseeb and R. Parasuraman, "Wisture: RNN-based learning of wireless signals for gesture recognition in unmodified smartphones," 2017, *arXiv:1707.08569*.

[26] D. Zhang, H. Wang, and D. Wu, "Toward centimeter-scale human activity sensing with Wi-Fi signals," *Computer*, vol. 50, no. 1, pp. 48–57, Jan. 2017.

[27] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, Feb. 2015, doi: 10.1038/nature14539.

[28] N. Damodaran, E. Haruni, M. Kokhkharova, and J. Schäfer, "Device free human activity and fall recognition using WiFi channel state information (CSI)," *CCF Trans. Pervasive Comput. Interact.*, vol. 2, no. 1, pp. 1–17, 2020, doi: 10.1007/s42486-020-00027-1.

[29] P. F. Moshiri, H. Navidan, R. Shahbazian, S. A. Ghorashi, and D. Windridge, "Using GAN to enhance the accuracy of indoor human activity recognition," Apr. 2020, *arXiv:2004.11228*.

[30] J. Ji, C. Xu, X. Zhang, B. Wang, and X. Song, "Spatio-temporal memory attention for image captioning," *IEEE Trans. Image Process.*, vol. 29, pp. 7615–7628, 2020.

[31] A. Galassi, M. Lippi, and P. Torroni, "Attention in natural language processing," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 10, pp. 4291–4308, Oct. 2021.

[32] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, 2014, pp. 1–15.

[33] S. Yousefi, H. Narui, S. Dayal, S. Ermon, and S. Valaee, "A survey on behavior recognition using WiFi channel state information," *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 98–104, Oct. 2017.

[34] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Device-free human activity recognition using commercial WiFi devices," *IEEE J. Sel. Area Commun.*, vol. 35, no. 5, pp. 1118–1131, Mar. 2017.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:49:05 UTC from IEEE Xplore. Restrictions apply.

[35] Q. Gao, J. Wang, X. Ma, X. Feng, and H. Wang, "CSI-based device-free wireless localization and activity recognition using radio image features," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10346–10356, Nov. 2017.
[36] C. Wang *et al.*, "Low-power fall detector using triaxial accelerometry and barometric pressure sensing," *IEEE Trans. Ind. Informat.*, vol. 12, no. 6, pp. 2302–2311, Dec. 2016.
[37] S. Głowinski, S. Majdanik, A. Głowinska, and E. Majdanik, "Trauma in a shaken infant? A case study," *Aggression Violent Behav.*, vol. 56, Jan. 2021, Art. no. 101515. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1359178920302196



Jian Su (Member, IEEE) received the B.S. degree in electronic and information engineering from Hankou University, Wuhan, China, in 2008, the M.S. degree in electronic circuit and system from Central China Normal University, Wuhan, in 2012, and the Ph.D. (Hons.) degree in communication and information systems from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2016.

He has been a Lecturer with the School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, China, since 2017. His current research interests cover the Internet of Things, radio frequency identification (RFID), and wireless sensors' networking.

Dr. Su is also a member of the Association for Computing Machinery (ACM).



Zhenlong Liao received the B.S. degree in Internet-of-Things Engineering from the Binjiang College, Nanjing University of Information Engineering, Nanjing, China, in 2020. He is currently pursuing the M.S. degree in computer science and technology with the School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing.

His research interests include Wi-Fi-based human activity sensing and the Internet of Things.



Zhengguo Sheng (Senior Member, IEEE) received the B.Sc. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2006, and the M.S. and Ph.D. (Hons.) degrees from Imperial College London, London, U.K., in 2007 and 2011, respectively.

He has been a Senior Lecturer with the Department of Engineering and Design, University of Sussex, Brighton, U.K., since 2015. His current research interests cover the Internet of Things (IoT), connected vehicles, and cloud/edge computing.



Alex X. Liu (Fellow, IEEE) received the Ph.D. degree in computer science from The University of Texas at Austin, Austin, TX, USA, in 2006.

He was a Professor with the Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, USA. He is currently an Adjunct Professor with the Shandong Provincial Key Laboratory of Computer Networks, Shandong Computer Science Center (National Supercomputer Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Jinan, China, and the Chief Scientist with the Ant Group, Hangzhou, China. His research interests focus on networking, security, and privacy.

Dr. Liu is also an IET Fellow and an Association for Computing Machinery (ACM) Distinguished Scientist. He received the IEEE & IFIP William C. Carter Award in 2004, the National Science Foundation CAREER Award in 2009, the Michigan State University Withrow Distinguished Scholar (Junior) Award in 2011, and the Michigan State University Withrow Distinguished Scholar (Senior) Award in 2019. He also received the Best Paper Awards from the IEEE International Conference on Sensing, Communication and Networking (SECON) 2018, the International Conference on Network Protocols (ICNP) 2012, the IEEE Symposium on Reliable Distributed Systems (SRDS) 2012, and the USENIX Large Installation System Administration Conference (LISA) 2010. He has served as the TPC Co-Chair of ICNP 2014 and IFIP Networking 2019. He has served as an Editor for IEEE/ACM TRANSACTIONS ON NETWORKING and an Area Editor for *Computer Communications*. He is also an Associate Editor for IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING and IEEE TRANSACTIONS ON MOBILE COMPUTING.



Dilbag Singh received the Ph.D. degree in computer science and engineering from the Thapar Institute of Engineering and Technology, Patiala, India, in 2019.

He has worked as an Assistant Professor at three well-known universities in India: Chandigarh University, Mohali, India, Manipal University Jaipur, Jaipur, India, and Bennett University, Greater Noida, India. In 2021, he moved to the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju, South Korea, where he is currently affiliated. He has published more than 100 research articles in Science Citation Index (SCI)-/Science Citation Index Expanded (SCIE)-indexed journals. He has also submitted five patents and has published three books and two book chapters. His H-index is 34. His research interests include image processing, computer vision, deep learning, metaheuristic techniques, and information security.

Dr. Singh has acted as a Lead Guest Editor/an Editorial Board Member of many SCI-/SCIE-indexed journals, such as *Journal of Healthcare Engineering*, *Mathematical Problems in Engineering*, and *Journal of Intelligent and Fuzzy Systems*. He was in the top 2% list issues by World Ranking of Top 2% Scientists in 2021. He was part of the 11 Web of Science-/Scopus-indexed conferences.



Heung-No Lee (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of California at Los Angeles, Los Angeles, CA, USA, in 1993, 1994, and 1999, respectively.

He was a Research Staff Member with HRL Laboratories, LLC, Malibu, CA, USA, from 1999 to 2002. From 2002 to 2008, he was an Assistant Professor with the University of Pittsburgh, Pittsburgh, PA, USA. In 2009, he moved to the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju, South Korea, where he is currently affiliated. His research interests include information theory, signal processing theory, blockchain, communications/networking theory, and their application to wireless communications and networking, compressive sensing, future Internet, and brain–computer interface.

Dr. Lee received several prestigious national awards, including the Top 100 National Research and Development Award in 2012, the Top 50 Achievements of Fundamental Researches Award in 2013, and the Science/Engineer of the Month in January 2014.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:49:05 UTC from IEEE Xplore. Restrictions apply.

Multimodal Sparse Representation-Based Classification Scheme for RF Fingerprinting

Kiwon Yang, Jusung Kang, Jehyuk Jang, and Heung-No Lee, *Senior Member, IEEE*

***Abstract*— In this letter, we propose a multimodal method for improving radio frequency (RF) fingerprinting performance that uses multiple features cultivated from RF signals. Combining multiple features, including a falling transient feature that has not previously been used in RF fingerprinting studies, we aim to demonstrate that the proposed method results in improved accuracy. We show that a sparse representation-based classification (SRC) scheme can be a good platform for combining multiple features. The experimental results on RF signals acquired from eight walkie-talkies show that the RF fingerprinting accuracy of the proposed method improves significantly as the number of features increases.**

***Index Terms*— Classification algorithm, feature extraction, multimodality, RF fingerprinting, radio frequency identification.**

I. Introduction

CLASSIFYING radio frequency (RF) signals is useful in electronic warfare to identify the radio transmission signals of adversaries [1]. For the classification to work well, the availability of good features and a simple but robust technique are essential. A feature is a sample vector cultivated from the transmitted RF signals and bears unique information about the pertinent device. Identification of RF transmitters using such features is called RF fingerprinting. Features are known to arise from many sources, including tiny differences in device fabrication process and electronic components [1].

RF fingerprinting has attracted significant attention [2]–[6]: In [2], RF-DNA features which contain information on variance, skewness, and kurtosis, within a preamble response were used with an ensemble method. In [3], four features—differential constellation trace figure, carrier frequency offset, modulation offset, and I/Q offset where classifications were done by calculating the minimum distance between test data and training data—were used. In [4], the mean of the instantaneous amplitude of the received signal and the modulation symbol were used with an optimal dimension-reduced matrix. In [5], an RF fingerprinting scheme based on low-rank representation of the original data with the robust classifier parameter was investigated. In [6], a convolutional neural network (CNN) for seven commercial Zigbee devices was used. They collected 1,000 data per class. In [2]–[6], RF fingerprinting schemes with multiple features, which exhibited a high accuracy rate, were proposed for Zigbee devices and satellite terminals.

The contributions and novelties of this letter are as follows:

- We propose a new RF fingerprinting algorithm and a set of three RF features—*rising transient*, *falling transient*, and *sync*—and show the possibility that each feature can provide unique information through a real-life experiment. The falling transient feature has never been used in RF fingerprinting studies. Our results indicate that the performance of RF fingerprinting improves as each feature is additionally employed.
- Even though SRC is a common algorithm in classification [7], no study on RF fingerprinting with a combination of SRC and multiple features has been reported. We show that SRC can be a good platform for RF fingerprinting.

The remainder of this letter is organized as follows. The experimental system is described in Section II. The proposed method is outlined in Section III. Results are presented and analyzed in Section IV. The conclusion is given in Section V.

II. Experimental System

A. Walkie-Talkie Signals

Our RF signals follow the digital mobile radio (DMR) standard. The DMR standard follows time-division multiple access (TDMA) and 4-level frequency-shift keying modulation [8]. A signal burst appears for 30ms and disappears for 30ms using the 2-slot TDMA method. This pattern is repeated in transmission.

The signal burst consists of rising transient, falling transient, and steady-state signals. The rising transient signal grows from zero to the designed level of the RF signal. Contrary to the rising transient signal, the falling transient signal decreases from the designed level to zero. The steady-state signal refers to the resting part between the rising and the falling transient signal and is composed of data and a sync signal. The data have 216 bits and the sync signal has 48 bits. The bit rate of the DMR standard is 9,600 bits/s. The sync signal is used to synchronize the transmitter and receiver.

From the pertinent signal part, a feature is obtained. Each feature is the main lobe of a spectrum of the pertinent signal part. We show how to extract each feature from the pertinent signal part in Section III. A. As we mentioned, the features arise from the inherent nonlinear properties of radio transmitters in the manufacturing process [1]. Owing to the presence of such features, RF fingerprinting can be accomplished.

Manuscript received December 24, 2018; revised January 31, 2019; accepted February 28, 2019. Date of publication March 20, 2019; date of current version May 8, 2019. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (NRF-2018R1A2A1A19018665). The associate editor coordinating the review of this letter and approving it for publication was F. Wang. *(Corresponding author: Heung-No Lee.)*

The authors are with the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea (e-mail: heungno@gist.ac.kr).

Digital Object Identifier 10.1109/LCOMM.2019.2905205

1558-2558 © 2019 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

B. Signal-Acquisition Setup

For our experiment, two walkie-talkie models were used: Motorola *Sl1M* and Hytera *BD-358*. Each model follows the DMR standard. Four units of each type, i.e., eight walkie-talkies in total, were used in the experiment.

Signal was transmitted from the transmitter and acquired from an SMA male mini car-mounted antenna, the receiving frequency band of which is 400–470 MHz. We then down-converted 423.1875 MHz to 10 MHz using an XL-11-411 RF mixer and an E4438C ESG vector signal generator. Then, we filtered the signal bandwidth and sampled the signal using an IF recording system with the PX14400 operator functioning as a low-pass filter and analog-to-digital converter. Signals sampled at 96 MHz were saved to a computer and loaded to MATLAB. As we captured 50 signals per walkie-talkie, 400 signals were saved to the computer.

III. Proposed Feature Extraction and Classification

A. Signal Burst to Features

To cultivate features, we extracted the three signal parts from a single signal burst. Then, each feature was selected from the pertinent extracted signal parts.

Each signal part is extracted from the first signal burst of the total received signal through time-windowing. To design the time-window for each signal, we used a thresholding method. For a rising transient signal $\mathbf{f}_R \in \mathbb{R}^{500,000\times 1}$, the starting threshold is the first time point at which the amplitude of the signal burst exceeds 10% of its maximum; the ending threshold is the time point it exceeds 90%. Similarly, for the falling transient signal $\mathbf{f}_F \in \mathbb{R}^{500,000\times 1}$, the starting threshold and the ending threshold are the latest time points at which the amplitude of the signal burst exceeds 90% and 10% of its maximum, respectively. Since the length of each transient signal fluctuates, we used zero padding method after the ending point of each transient signal to match the length. To design the time window for the sync signal, we referred to the DMR standard [8]. The sync signal $\mathbf{f}_S \in \mathbb{R}^{480,000\times 1}$ is located at the center of a signal burst. We set the center of the time-window for the sync signal to the central time point between the ending time of the rising transient signal and the starting time of the falling transient signal. The width of the time window was set to 0.005 s, as per the DMR standard [8]. Each extracted signal part was normalized to consider the case of signals with different power levels.

The extracted signal parts were transformed to the spectrum domain by fast Fourier transform (FFT) with the size of the time signal. Then, the operation of taking the absolute value of each element was executed to compare the energy and frequency information of the extracted signal part with those of the others. Since the main lobe occupies most of the energy of each signal part, the main lobe was taken from each spectrum. The main lobes, $F(\mathbf{f}_R)_{ML} \in \mathbb{R}^{2,000\times 1}$, $F(\mathbf{f}_F)_{ML} \in \mathbb{R}^{2,000\times 1}$, and $F(\mathbf{f}_S)_{ML} \in \mathbb{R}^{1,920\times 1}$, are the unique features used in our experiment, where $F(\cdot)$ is the FFT operation function and ML denotes the main lobe of the spectrum.

To extract the main lobe, we used a bandpass filter first. Then, the center frequency of the filtered spectrum was down-converted to zero. Finally, we decimated the signal to reduce the length of the sample sequence. The main lobe was set to occupy most of the energy of each signal part, considering the channel bandwidth.

B. Proposed SRC

SRC is a classification algorithm based on the compressed sensing theory [7] and is used to determine the class from the sparse solution of the representation equation

$$\mathbf{y} = \mathbf{D}\mathbf{s}, \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^{P\times 1}$ is a test data vector with length P, $\mathbf{D} \in \mathbb{R}^{P\times NL}$ is a training data matrix composed of N training data vectors for each class label of transmitters $l \in \{1, \cdots, L\}$, and $\mathbf{s} \in \mathbb{R}^{NL\times 1}$ is a vector of sparse representation coefficients. The sparse signal recovery algorithm in [7] was used to solve (1) with $P < NL$. In (1), $\mathbf{Ds}$ can be rewritten as

$$\mathbf{Ds} = \left[\, \mathbf{D}^{(1)} \quad \cdots \quad \mathbf{D}^{(L)} \,\right] \left[\, \left(\mathbf{s}^{(1)}\right)^T \quad \cdots \quad \left(\mathbf{s}^{(L)}\right)^T \,\right]^T, \tag{2}$$

where $\mathbf{D}^{(l)}$ is a submatrix of $\mathbf{D}$ corresponding to the lth class label of transmitters and $\mathbf{s}^{(l)}$ is a subvector of $\mathbf{s}$ corresponding to the lth class label of transmitters, and T denotes the transpose. To identify the class of test data, we solve

$$\text{class} = \underset{l\in\{1,\cdots,L\}}{\arg\min} \left\| \mathbf{y} - \mathbf{D}^{(l)}\mathbf{s}^{(l)} \right\|_2. \tag{3}$$

If the column vectors in $\mathbf{D}$ are less correlated, the solution $\mathbf{s}$ of (1) is approximated to be sparse since the condition of having a sparse solution depends on the mutual correlation between the columns of $\mathbf{D}$ [9]. Thus, the compressed sensing algorithms in [7] can be used to find a unique solution $\mathbf{s}$. However, when RF signals are taken directly to form the column vectors of $\mathbf{D}$, the solution $\mathbf{s}$ cannot be sparse because they may be highly correlated. Then, the performance of SRC may be poor. Thus, RF signals must be processed to remove correlation to obtain high performance in SRC [9].

To remove correlation among RF signals, the proposed method applies principal components analysis (PCA) to the column vectors, each of which combines multiple features. PCA is known to be good at geometrically separating the features in the Euclidean domain by removing the mutual correlation [10]. This section aims to show how three kinds of features are concatenated and how PCA is applied to the features. We first introduce the single modal method and then discuss the proposed multimodal method.

1) Single Modal RF Fingerprinting: Consider that one of the rising transient, falling transient, and sync features is used as the sole representative feature of a single transmitter. We first form a feature matrix in which the columns are the sample vectors of the feature of candidate RF transmitters. Mathematically, for L RF transmitters (classes) and N sample vectors of a feature of each RF transmitter, we construct the feature matrix $\mathbf{A} \in \mathbb{R}^{M\times NL}$ as follows:

$$\mathbf{A} = \left[\mathbf{a}_1^{(1)}, \cdots, \mathbf{a}_N^{(1)}, \mathbf{a}_1^{(2)}, \cdots, \mathbf{a}_N^{(L)}\right], \tag{4}$$

where column vector $\mathbf{a}_n^{(l)} \in \mathbb{R}^{M\times 1}$ is the nth sample vector of a feature of the lth RF transmitter for $n = 1, \cdots, N$ and $l = 1, \cdots, L$, and M is the length of a feature. We denote a feature of an unknown RF transmitter as $\mathbf{u} \in \mathbb{R}^{M\times 1}$. From the PCA operation, (7) and (9), $\mathbf{A}$ and $\mathbf{u}$ are changed to a training data matrix $\mathbf{D}$ and a test data vector $\mathbf{y}$, respectively.

2) Multimodal RF Fingerprinting: The proposed method is to concatenate the multiple features in the representation equation $\mathbf{u} = \mathbf{As}$, as shown in Fig. 1; the feature matrices are concatenated in a row-wise manner. Let us denote the nth sample vector of the kth feature of the lth RF transmitter for $n = 1, \cdots, N$, $l = 1, \cdots, L$, and $k = 1, \cdots, K$ as $\mathbf{a}_{k,n}^{(l)} \in \mathbb{R}^{M\times 1}$, where K is the number of features to be combined. The feature matrices are concatenated as follows:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1{}^T & \mathbf{A}_2{}^T & \cdots & \mathbf{A}_K{}^T \end{bmatrix}^T \in \mathbb{R}^{MK\times NL}, \tag{5}$$

where $\mathbf{a}_{k,n}^{(l)}$ forms the columns of feature matrix $\mathbf{A}_k$,

$$\mathbf{A}_k = \left[\mathbf{a}_{k,1}^{(1)}, \cdots, \mathbf{a}_{k,N}^{(1)}, \mathbf{a}_{k,1}^{(2)}, \cdots, \mathbf{a}_{k,N}^{(L)}\right] \in \mathbb{R}^{M\times NL}. \tag{6}$$

Then, PCA is applied to the feature matrix $\mathbf{A}$. We obtain the training data matrix $\mathbf{D}$ as follows:

$$\mathbf{D} = \mathbf{V}^T(\mathbf{A} - \mathbf{m1}) \in \mathbb{R}^{P\times NL}, \tag{7}$$

where $\mathbf{m} = \frac{1}{L\times N}\sum_{l=1}^{L}\sum_{n=1}^{N}\mathbf{a}_n^l \in \mathbb{R}^{MK\times 1}$ is the average vector of the columns of $\mathbf{A}$, $\mathbf{1} := [1\ 1\ \cdots\ 1]$ is the 1 by NL vector of 1 s, $\mathbf{a}_n^l = \left[(\mathbf{a}_{1,n}^l)^T(\mathbf{a}_{2,n}^l)^T\cdots(\mathbf{a}_{K,n}^l)^T\right]^T \in \mathbb{R}^{MK\times 1}$ is a column vector which combines K features, and $\mathbf{V} \in \mathbb{R}^{MK\times P}$ is a rearranged eigenvector matrix of the covariance matrix $(\mathbf{A} - \mathbf{m1})(\mathbf{A} - \mathbf{m1})^T \in \mathbb{R}^{MK\times MK}$. The eigenvectors of $(\mathbf{A}-\mathbf{m1})(\mathbf{A}-\mathbf{m1})^T$ are arranged according to the eigenvalues in descending order. Since the eigenvalue of the covariance matrix is proportional to the variance of the columns of $\mathbf{A}$ and the eigenvectors of the covariance matrices are orthonormal, the column vector of $\mathbf{V}$ becomes a basis of the new space on the variance of the columns of $\mathbf{A}$ [10]. The dimension of $\mathbf{V}$ can be selected by the user as $P \in \{1, \cdots, MK\}$. To obtain the test data vector $\mathbf{y}$ of SRC, we first concatenate different features of an unknown transmitter $\mathbf{u}_k \in \mathbb{R}^{M\times 1}$ as follows:

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_1^T & \mathbf{u}_2^T & \cdots & \mathbf{u}_K^T \end{bmatrix}^T \in \mathbb{R}^{MK\times 1}. \tag{8}$$

Finally, PCA is applied to a concatenated feature of unknown transmitter $\mathbf{u}$. The test data vector $\mathbf{y}$ is obtained by mapping the difference between concatenated feature $\mathbf{u}$ and $\mathbf{m}$, i.e., $\mathbf{u} - \mathbf{m}$, onto the space with the eigenvector matrix $\mathbf{V}$,

$$\mathbf{y} = \mathbf{V}^T(\mathbf{u} - \mathbf{m}) \in \mathbb{R}^{P\times 1}. \tag{9}$$

By using PCA, the equation in Fig. 1 is changed to (1), which has principal components as training and test data. The SRC solution in (1) was determined using the basis pursuit algorithm, which finds the unique sparse solution that has the minimum L1 norm [11].

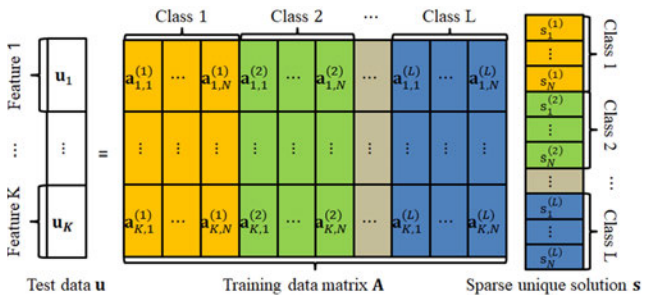


Fig. 1. Feature concatenation in the proposed method.

TABLE I
ACCURACY RATE OF THE PROPOSED METHOD

	4 BD-358	4 SL1M	4 BD-358 4 SL1M
	Accuracy rate (Minimum number of PC)		
TR(R)	88% (24)	82% (48)	90.5% (45)
TR(F)	87.5% (45)	90% (12)	92.25% (13)
TR(R + F)	93% (49)	92% (20)	95.5% (63)
Sync	99% (45)	83.5% (22)	93.75% (86)
TR(R + F) + Sync	99% (44)	98.5% (22)	98.75% (21)

R: Rising, F: Falling, PC: Principal components

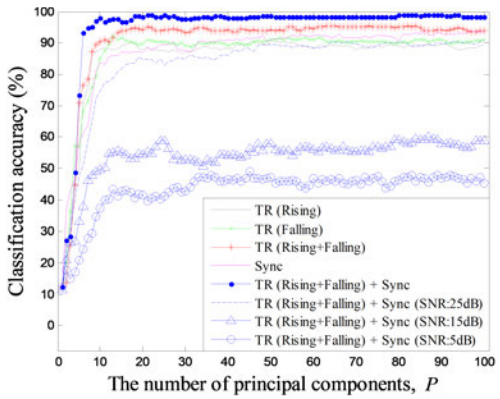
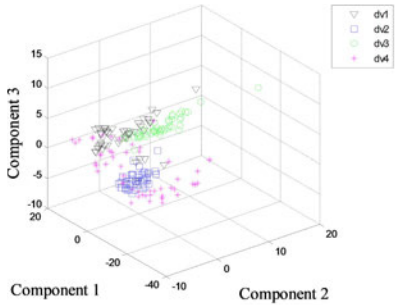


Fig. 2. Classification result of four BD-358 and four SL1M walkie-talkies.

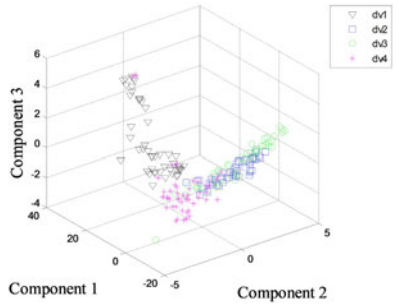
IV. EXPERIMENTAL RESULTS AND DISCUSSION

For our experiment, we set the decimation rate for all feature extractions to 250, considering the bandwidth of the RF signal following the DMR standard [8] and the sampling rate. To evaluate the performance of the proposed classifier, we used a five-fold cross validation technique. Fifty data were used per walkie-talkie, such that each test data was classified on the total of 320 training data. The experiment was performed in a line-of-sight environment and SNR was around 35-40 dB.

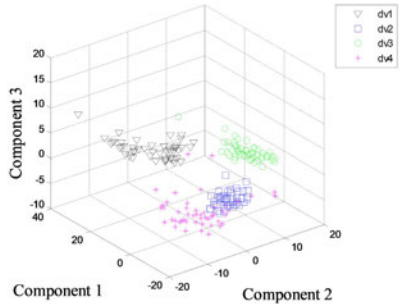
Table I shows the accuracy rate of the proposed method. The accuracy rate of the multimodal scheme is much better than that obtained from using only one feature. The minimum number of principal components is the minimum number of column vectors in the eigenvector matrix $\mathbf{V}$ that yields the highest accuracy rate. Fig. 2 shows eight results of classification using 1 to 100 principal components, 1) on the rising transient feature, 2) on the falling transient feature, 3) on both transient features combined, 4) on the sync feature, and



(a) SL1M transient (rising + falling) features



(b) SL1M sync features



(c) SL1M transient (rising + falling) + sync features

Fig. 3. Feature map of a 3-D principal components space.

5) on all features combined in SNR : 35-40dB, 6) 25dB, 7) 15dB, and 8) 5dB. Improved performance with an increased number of features indicates that each feature could contain unique information. Since the eigenvectors of the 21st and higher eigenvalues of $(\mathbf{A} - \mathbf{m1})(\mathbf{A} - \mathbf{m1})^T$ do not have enough information to represent the differences of data, the classification accuracy is not changed as $P > 21$. Fig. 3 shows a feature map in which the principal components of features of Motorola *Sl1M* are mapped onto a 3D plot. The label of each axis, such as Component 1 and Component 2, means the projection of $\mathbf{u} - \mathbf{m}$ to the Pth column vector of $\mathbf{V}$. The figure shows distinct cluster formation when a concatenated feature is used.

To compare the proposed method with convolutional neural network (CNN), we referred to the studies [6] and [12]. For additional tests, we used a five-fold cross validation technique. Five training neural networks were constructed using the same training dataset with the proposed method. We used the concatenated features as input data. The average classification accuracy rate of the CNN was 91.5%.

In our experiment, the size of the training data set is much smaller than that of [6]. Comparing CNN (91.5%) with the proposed method (98.75%), we show that SRC performs better than CNN for this small training data set. Because it is simple to add more training data and different kinds of features in SRC, the performance of the proposed method could be improved with additional training data and other kinds of features.

V. Conclusions

This letter proposed a multimodal RF fingerprinting scheme based on SRC. We showed that the proposed multimodal scheme, which concatenates multiple features in the row-wise manner and applies PCA to the concatenated dictionary matrix, improves accuracy significantly. We showed the possibility that the three signal features we have cultivated from RF signal samples could provide mutually independent information. The proposed scheme is efficient in the sense that improved RF fingerprinting accuracy is obtained. In addition, it is simple and easy in the proposed scheme to add more data and various kinds of features. The MATLAB source code for this study can be obtained at https://infonet.gist.ac.kr/?page_id=14.

References

[1] Y. Jia, S. Zhu, and L. Gan, "Specific emitter identification based on the natural measure," *Entropy*, vol. 19, no. 3, p. 117, 2017.

[2] H. J. Patel, M. A. Temple, and R. O. Baldwin, "Improving ZigBee device network authentication using ensemble decision tree classifiers with radio frequency distinct native attribute fingerprinting," *IEEE Trans. Rel.*, vol. 64, no. 1, pp. 221–233, Mar. 2015.

[3] L. Peng, A. Hu, J. Zhang, Y. Jiang, J. Yu, and Y. Yan, "Design of a hybrid RF fingerprint extraction and device classification scheme," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 349–360, Feb. 2019.

[4] Y. Jia, J. Ma, and L. Gan, "Combined optimization of feature reduction and classification for radiometric identification," *IEEE Signal Process. Lett.*, vol. 24, no. 5, pp. 584–588, May 2017.

[5] Y. Jia, J. Ma, and L. Gan, "Radiometric identification based on low-rank representation and minimum prediction error regularization," *IEEE Commun. Lett.*, vol. 21, no. 8, pp. 1847–1850, Aug. 2017.

[6] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, "Deep learning for RF device fingerprinting in cognitive communication network," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 160–167, Feb. 2018.

[7] J. Wright, A. Y. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[8] *Electromagnetic Compatibility and Radio Spectrum Matters (ERM); Digital Mobile Radio (DMR) Systems; Part 1: DMR Air Interface (AI) Protocol*, Standard ETSI TS 102 361-1, European Telecommun. Standards Inst., 2016.

[9] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, Feb. 2009.

[10] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdiscipl. Rev. Comput. Statist.*, vol. 2, no. 4, pp. 433–459, 2010.

[11] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis Pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001.

[12] L. Ding, S. Wang, F. Wang, and W. Zhang, "Specific emitter identification via convolutional neural networks," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2591–2594, Dec. 2018.

MLNet: Metaheuristics-Based Lightweight Deep Learning Network for Cervical Cancer Diagnosis

Manjit Kaur, *Senior Member, IEEE*, Dilbag Singh, *Senior Member, IEEE*, Vijay Kumar, and Heung-No Lee, *Senior Member, IEEE*

***Abstract*—One of the leading causes of cancer-related deaths among women is cervical cancer. Early diagnosis and treatment can minimize the complications of this cancer. Recently, researchers have designed and implemented many deep learning-based automated cervical cancer diagnosis models. However, the majority of these models suffer from over-fitting, parameter tuning, and gradient vanishing problems. To overcome these problems, in this paper a metaheuristics-based lightweight deep learning network (MLNet) is proposed. Initially, the hyper-parameters tuning problem of convolutional neural network (CNN) is defined as a multi-objective problem. Particle swarm optimization (PSO) is used to optimally define the CNN architecture. Thereafter, Dynamically hybrid niching differential evolution (DHDE) is utilized to optimize the hyper-parameters of CNN layers. Each particle of PSO and solution of DHDE together represent the possible CNN configuration. F-score is used as a fitness function. The proposed MLNet is trained and validated on three benchmark cervical cancer datasets. On the Herlev dataset, MLNet outperforms the existing models in terms of accuracy, f-measure, sensitivity, specificity, and precision by 1.6254%, 1.5178%, 1.5780%, 1.7145%, and 1.4890%, respectively. Also, on the SIPaKMeD dataset, MLNet achieves better performance than the existing models in terms of accuracy, f-measure, sensitivity, specificity, and precision by 2.1250%, 2.2455%, 1.9074%, 1.9258%, and 1.8975%, respectively. Finally, on the Mendeley LBC dataset, MLNet achieves better performance than the competitive models in terms of accuracy, f-measure, sensitivity, specificity, and precision by 1.4680%, 1.5845%, 1.3582%, 1.3926%, and 1.4125%, respectively.**

***Index Terms*—Cervical cancer, deep learning, diagnosis, dynamically hybrid niching differential evolution, lightweight deep learning network, metaheuristics, particle swarm optimization.**

Manuscript received 14 June 2022; revised 29 September 2022; accepted 31 October 2022. Date of publication 18 November 2022; date of current version 5 October 2023. This work was supported in part by the MSIT (Ministry of Science and ICT), Korea, through the ITRC (Information Technology Research Center) Support Program IITP-2021-2021-0-01835, supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation), and in part by the National Research Foundation of Korea (NRF), funded by the Korean government (MSIP) under Grant NRF-2021R1A2B5B03002118. *(Corresponding author: Heung-No Lee.)*

Manjit Kaur, Dilbag Singh, and Heung-No Lee are with the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea (e-mail: manjit.kr@yahoo.com; dggill2@gmail.com; heungno@gist.ac.kr).

Vijay Kumar is with the Computer Science and Engineering Department, NIT Hamirpur, Hamirpur, Himachal Pradesh 177005, India (e-mail: vijaykumarchahar@gmail.com).

Digital Object Identifier 10.1109/JBHI.2022.3223127

I. Introduction

CERVICAL cancer is found in the lower part of the uterus. The cell of cervix progressively grows in pre-cancerous cells and cultivates cancer. Cervical cancer is the fourth most life-threatening disease in females [1]. It is responsible for a 90% mortality rate in developing countries due to its unawareness of the disease and its impact on life [2]. The early detection of cervical cancer is done by using Pap testing [3]. In this test, the samples are taken from the cervix and examined by the pathologist for any abnormalities present in cervix [4]. Manual testing requires expert pathologists for abnormalities analysis. However, this process is time-consuming and error-prone. To handle these issues, automatic cell classification techniques are required.

The well-known automated analysis systems are BD Focal Point slide profiler [5] and ThinPrep [6]. These automated systems were approved by the Food and Drug Administration (FDA) of the US. These systems not only reduce the analysis time but also improve the classification performance. In recent years, artificial intelligence techniques are widely used for automatic cervical cell classification [7]. Most of the existing techniques utilize the concept of feature extraction to enhance the performance of the classifier and reduce the computational run time. Some of these techniques are support vector machine (SVM), K-nearest neighbor (KNN), and artificial neural networks (ANN). Machine learning techniques provide better results for small datasets. However, the medical dataset is large in size and complex in nature. Sometimes, the accurate classification through machine learning techniques is a little bit difficult due to the variations in shape and size of cells. The different feature extraction techniques provide different results. Hence, there is a need for an automatic feature extraction technique. For this, deep learning techniques are widely used.

Deep learning techniques are usually applied in several areas of medical imaging [8]. Among the existing deep learning techniques, convolutional neural networks (CNNs) are extensively used to classify cervical cells [9]. These techniques eliminate the need for pre-processing steps such as feature extraction. These techniques directly utilize the original image and reduce the

2168-2194 © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:50:11 UTC from IEEE Xplore. Restrictions apply.

control parameter of the training phase. The main aspects of deep learning techniques are pooling and weight sharing. Some of the well-known techniques are long-short term memory (LSTM), recurrent neural network (RNN), and generative adversarial networks (GAN). The performance of these techniques is still far from the optimal results. It is found that the existing models suffer from over-fitting, parameter tuning, and gradient vanishing problems. Model parameters are generally selected through trial-and-error, resulting in poorer performance. This motivates us to develop a novel technique for cervical cell classification. The main contributions of this paper are as follows:

1) A metaheuristics-based lightweight deep learning network (MLNet) is proposed.
2) The hyper-parameters tuning problem of convolutional neural network (CNN) is defined as a multi-objective problem.
3) Particle swarm optimization (PSO) is used to optimally define CNN architecture. Thereafter, Dynamically hybrid niching differential evolution (DHDE) is utilized to optimize the hyper-parameters of CNN layers. Each particle of PSO and solution of DHDE together represent the possible CNN configuration.

The remaining structure of this paper is as follows. Section II presents the related work done in the field of cervical cell classification. The utilized approaches are discussed in Section III. The proposed model is described in Section IV. Experimental results and discussion are mentioned in Section V followed by the conclusion in Section VI.

II. Related Work

Lin et al. [10] extended CNN-based image identification algorithm. Smooth L1 loss and regression methods were combined to enhance the classification accuracy. Cycle GAN was used due to its ease of implementation. They attained the classification accuracies of 98.1% and 97.6% for binary and seven class problems of the Herlev dataset, respectively. However, the classification accuracy of some samples is not good. Chen et al. [11] developed a lightweight CNN (LCNN) for the classification of cervical cells. The compression method was used to improve the performance of lightweight CNN. VGG, ResNet, and Inception-ResnetV2 were used to design the teacher model and determined labels from the designed model. Thereafter, Xception, MobileNet, MobileNetV2, and DenseNet were used to design the student model. The accuracy obtained from Inception-ResnetV2 was 73.58%. However, it suffers from a hyper-parameter tuning problem.

Sabeena et al. [12] utilized a pyramid scene parsing model (PSPM) for cell recognition from cervical images. The local and global priors were integrated with PSPM. The average classification accuracy obtained from their approach was 99.7%. However, the robustness of this approach has yet to be validated on different datasets. Allehaibi et al. [9] used a Mask regional CNN (Mask R-CNN) for the segmentation of cervical cells. Thereafter, a pretrained VGG model was used for classification. The accuracy obtained from Mask R-CNN was 96% for two-class classification. However, this approach requires high computational complexity. Zhao et al. [13] utilized dense U-Net (DU-Net) to design the deep learning framework for the segmentation of nuclei from cervical images. The ensemble strategy was used to eliminate the predictive bias during the training phase. Their method achieved precision and recall of 0.946 and 0.984, respectively.

Hussain et al. [14] proposed a fully CNN (FCNN) for the classification of cervical nuclei from the images. It attained an accuracy of 96%. However, it suffers from high computational time. Shi et al. [15] developed a cervical cell classification model using graph CNN (GCNN). They used intrinsic features of an image to capture the correlation among clusters. The average accuracy obtained from this method was 94.23%. This model suffers from annotation costs. Li et al. [16] integrated global information with an attention mechanism to classify the cervical nuclei. The features were extracted through ResNet50. These features were applied to the attention mechanism to find out the cell region. Long-short term memory (LSTM) was used to aggregate the cell features. The average classification accuracy obtained from this model was 98.89%.

Rahaman et al. [17] designed a novel deep learning framework named as DeepCervix for the classification of cervical cells. DeepCervix utilized VGG16, VGG19, XceptionNet, and ResNet50 for feature extraction. Dropout layer and batch normalization were incorporated in DeepCervix. DeepCervix was evaluated on the Herlev dataset and attained the accuracy of 98.91% for two-class and 90.32% for seven classes. However, the poison noise greatly degrades the performance of the model. Dong et al. [18] used InceptionV3 for cervical cell classification. The performance of InceptionV3 was evaluated on the Herlev dataset and attained an accuracy of 98.23%. The artificial features may be incorporated in InceptionV3 for better performance. Chankong et al. [19] designed an automatic cervical image classification. Fuzzy C-means (FCM) were utilized to extract the features. These features were applied on Bayesian, linear discriminant analysis (LDA), K-nearest neighbor (KNN), artificial neural network (ANN), and support vector machine (SVM). For the Herlev dataset, this model obtained the 93.78% and 99.27% accuracies for seven-class and binary-class, respectively. This method was not validated for multiple cell classification.

Wang et al. [20] developed an automatic cervical image classification. Mean-shift clustering was used to determine the regions of interest. The overlapping regions were spitted through morphological operations. The feature set was optimized through a chain-like agent genetic algorithm (CAGA). Their method outperformed the other methods in terms of specificity, sensitivity, and accuracy. The robustness of this method was not tested on different types of datasets. Martin et al. [21] used CNN for cervical cell classification. The average accuracy attained from their model was 60%. Yaman and Tuncer [22] utilized the exemplar pyramid deep feature extraction (EPDFE) technique to detect cervical cancer. Two pre-trained transfer learning models named DarkNet19 and DarkNet53 were used for extracting deep features from cervical images. Neighborhood component analysis was used to select the features obtained from pre-trained models. The selected deep features were applied to SVM for

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:50:11 UTC from IEEE Xplore. Restrictions apply.

classification. The average accuracy obtained from this approach was 96%. But this approach did not get better results for the SIPaKMeD dataset.

Manna et al. [23] designed an ensemble CNN model using fuzzy rank (CNN-F) for cervical image classification. The fuzzy rank-based fusion of three pre-trained models such as IncpetionV3, Xception, and DenseNet169 was incorporated in this model. It performed better than the other models in terms of accuracy. This model was unable to classify the overlapping and poor contrast images. Chen et al. [24] developed an automatic cervical cell classification. Compact VGG was used for developing a cell classifier. Their method performs better than the other models in terms of classification measures. Tripathi et al. [25] studied four deep learning models namely ResNet50, ResNet152, VGG16, and VGG19 for cell classification. The best accuracy obtained from ResNet152 was 94.89%. However, the parallel implementation of these models is required to improve the classification performance. Basak et al. [26] developed an automatic deep learning framework for cytology cell classification. CNN model was used to extract the deep features. These features were further reduced through principal component analysis (PCA). The reduced features were optimized by a grey wolf optimization algorithm. Thereafter, these optimized features were applied to SVM for classification. The classification accuracies obtained by this method were 98.32% and 97.87% for seven-class of Herlev and five-class of SIPaKMeD, respectively. Kumar and Krishna [27] classified cervical cancer by varying the channels of CNN (c-CNN). c-CNN with channels (32,64) achieved 96.89% accuracy, 94.15% F-score, 93.75% sensitivity, and 93.38% precision.

Martinez-Mas et al. [28] developed a CNN for PAP-smear cell classification. The cell merger approach was used in the developed model. The accuracy obtained from their model was 88.8%. This model can be further improved by merging networks. Pal et al. [29] developed a deep multiple instance learning (DMIL) for cell classification. Deep CNN was used to extract deep features from cervical cell images. The average classification accuracy obtained from this method was 84.55%. Ghoneim et al. [30] developed a cervical cell classification model using CNN and an extreme learning machine (ELM) referred to as CNN-ELM. The deep features were extracted using CNN. These features were applied to ELM for classification. The classification accuracy obtained by their method was 91.2% for the seven-class of Herlev dataset. The architectures of other models can be incorporated for further improvement. Jia et al. [31] developed a novel deep learning framework for cell classification. CNN was used to extract the deep features. Gabor and deep features were combined. These features were reduced through PCA. The reduced features were applied to SVM for classification. This method outperformed the other models in terms of accuracy, sensitivity, and specificity. Alyafeai and Ghouti [32] used YOLO algorithm for the extraction of a region of interest (ROI). Lightweight CNN was used to classify the cervical cells. The accuracy under the curve obtained from this approach was 0.82.

Zhang et al. [33] used the weight transfer learning approach for colposcopy image classification. The pre-trained DenseNet model was used for classification. The feature map concatenation was used to eliminate the gradient problem. Their approach obtained an accuracy of 73.08% over 600 test images. Xiang et al. [34] developed an automatic cervical cell classification (ACCC). YOLOv3 was used to extract the ROIs. CNN model was used to classify the cervical cell. ACCC performed better than the existing methods in terms of performance measures. Zorkafli et al. [35] developed a neural network based on hybrid multi-layered perceptron (HMLP) and GA for cervical cell identification. The average accuracy obtained from this model was 74.82%. Bhavani and Govardhan [36] designed a stacked ensemble technique for cervical cell classification. The synthetic minority over-sampling technique (SMOTE) was used for data balancing. Thereafter, KNN, SVM, Random forest, and logistic regression were used to build the model for classification. The developed model was evaluated on the UCI machine learning dataset and attained an accuracy of 91.2%. However, the computational time of the developed model is high.

It is observed that the performance of existing models is still far from the optimal results. The existing techniques suffer from over-fitting, parameter tuning, and gradient vanishing problems. Hence, there is a need to develop a metaheuristics-based lightweight deep learning network that can classify cervical cancer efficiently.

III. Preliminaries

A. Convolutional Neural Network

CNN is a deep learning architecture that learns the key features automatically from images. It can capture temporal and spatial dependencies using filters in an image. It also reduces the images without losing key features to minimize the computational complexity and improve performance. It also requires few parameters in comparison to traditional neural networks. CNN has mainly three layers: convolution, pooling, and fully connected.

In the convolution layer, key features are extracted from images using convolution operation and activation functions. The convolution operation facilitates the layers to identify similar objects in several images with different regions. It contains some parameters such as the size of filter, number of filter, padding, and stride that need to be set before the training process. In the pooling layer, the spatial size of convolved features is reduced through dimensionality reduction. Thus, it minimizes the computational power requirement. Its main benefit is that it extracts dominant features that are positional and rotational invariant. The different types of pooling can be utilized such as max, average, stochastic, spatial pyramid, and def-pooling [37]. But max pooling is mostly used as it provides better robustness and faster convergence. This layer does not contain any learnable parameters but has padding, stride, and filter size parameters. Both convolution and pooling layers together form the ith layer of CNN. Depending on the complexity of the images, the number of these layers can be increased. The output of the last pooling layer, i.e., a feature vector is converted into a one-dimensional vector and fed into a fully connected layer. Fully connected layers are also called densely connected layers because every

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:50:11 UTC from IEEE Xplore. Restrictions apply.

neuron in one layer is connected to every neuron of another layer. The final fully connected layer classifies the images.

B. Particle Swarm Optimization

PSO is a metaheuristic technique that is used to solve optimization problems.In PSO [38], each particle can represent as a solution. These particles contain position and velocity vectors and evolved through the problem space. The position of the particles is adjusted through personal- and global-best positions. In the personal best position, each particle keeps the track of its own best position that has been found so far. The global best position keeps the record of the best position which is founded among all particles. Particle i holds velocity vector $Vel_i = (vel_i^1, vel_i^2, \ldots, vel_i^D)$ and position vector $Pos_i = (pos_i^1, pos_i^1, \ldots, pos_i^D)$, where D is the number of dimensions. PSO finds the best solution by updating Vel_i and Pos_i through each iteration (it) as follows:

$$Vel_i(it+1) = \omega \cdot Vel_i(it) + c_1 r_1 (Best_{p_i}(it) - Pos_i(it)) + c_2 r_2 (Best_g(it) - Pos_i(it)) \quad (1)$$

$$Pos_i(it+1) = Pos_i(it) + Vel_i(it+1) \quad (2)$$

where $Best_{p_i}$ represents the personal best position of the particle i that has been searched so far. $Best_g$ denotes the best position that has been found by a whole swarm so far. r_1 and r_2 are the random numbers that are uniformly distributed in the range of [0,1]. The parameters such as cognitive coefficient (c_2), social coefficient (c_1), and inertia weight (ω) maintain the trade-off between exploitation and exploration. ω impacts the performance of PSO, so it should be calculated carefully. In this paper, sigmoid-like inertia weight, given by [39], is used to achieve a better trade-off and convergence. It is defined as:

$$\omega = \begin{cases} 0.9, & \text{if } it \leq \alpha M_{nt^1} \\ \frac{1}{ite^{(10it - 2M_{nt^1})/M_{nt^1}}} + 4, & \text{otherwise} \end{cases} \quad (3)$$

where M_{nt^1} is the maximum number of iterations. α is the predefined constant.

C. Dynamically Hybrid Niching Differential Evolution

Neural network ensembles are a multimodal optimization problem (MMOP) because it has multiple local and/or global optima, rough search space, and a large number of hyper-tunning parameters. To find the optimal solutions and improve the convergence accuracy, a dynamically hybrid niching-based differential evolution (DHDE) [40] is utilized. The reason is that it provides a better trade-off between convergence and diversity. DHDE utilizes crowding and speciation niching techniques dynamically during the execution to achieve the same. In DHDE, diversity is improved using the information of inferior offspring and crowding-based differential evolution (CDE-I_{OA}). Inferior offsprings are those who are weaker than their parents. Instead of discarding these inferior offsprings, DHDE saves them into an archive (i.e., inferior offspring archive (I_{OA})) to explore additional optimal solutions. The convergence is promoted by detecting and removing similar individuals from the population using improved neighborhood speciation-based differential evolution (INSDE). Optimal solution archive (O_{SA}) is used to save the obtained optimal solutions during the execution. The working of the DHDE algorithm is explained in Algorithm 2 (see Section IV). DHDE combines CDE-I_{OA} and INSDE to achieve better diversity and convergence. Firstly, CDE-I_{OA} is called and then INSDE is executed.

TABLE I
THE HYPER-PARAMETERS OF CNN ARCHITECTURE AND THEIR RANGES

Hyper-parameter	Range
Number of convolution layers (nCL)	[2,8]
Number of pooling layers (nPL)	[2,8]
Number of fully connected layers (nFL)	[2,8]

TABLE II
THE HYPER-PARAMETERS OF CNN'S LAYER AND THEIR RANGES

Layer	Hyper-parameter	Range
Fully connected	Number of neurons (F_{nn})	[1,1024]
Pooling	Padding pixels (pp_p)	[0,1]
	Stride size (ss_p)	[1,5]
	Filter size (fs_p)	[1,13]
Convolution	Stride size (ss_c)	[1,5]
	Padding pixels (pp_c)	[0,1]
	Filter size (fs_c)	[1,13]
	Number of filters (nf_c)	[64]

IV. PROPOSED MODEL

Although CNN extracts the features automatically from the images, it requires a number of hyper-parameters to be set before starting the training process. The value of these parameters plays a very important role to build an efficient CNN model. These parameters can be set by human expertise or using a trial-error basis. But it is a very lengthy and time-consuming process. Therefore, it is required to select the parameters optimally to build an optimized and efficient CNN model. Table I shows the hyper-parameters of CNN architecture with their ranges. The hyper-parameters of layers with their ranges are shown in Table II. Padding pixels of convolution and pooling layers are either set to 0 or 1. Here, 0 and 1 represent valid and the same paddings, respectively. Odd filter size is used in both convolution and pooling layers. The stride size is always taken smaller than the filter size.

Fig. 1 shows the proposed MLNet for cervical cancer diagnosis. The remaining section describes the optimization of CNN architecture and its hyper-parameters using PSO and DHDE. Inspired from [41], firstly, PSO is used to optimize the hyper-parameters of CNN architecture that are mentioned in Table I. Thereafter, DHDE is utilized to optimize the hyper-parameters of the layers which are mentioned in Table II. Each particle of PSO and solution of DHDE together represent the possible CNN configuration. Finally, the softmax layer that classifies cervical cancer. The fitness of particle and solution is tested using the obtained F-score. The hyper-parameters of CNN are optimized iteratively by PSO and DHDE and find a configuration with the best fitness value. Thereafter, the CNN model is configured using the optimal hyper-parameters and trained with many cervical

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:50:11 UTC from IEEE Xplore. Restrictions apply.

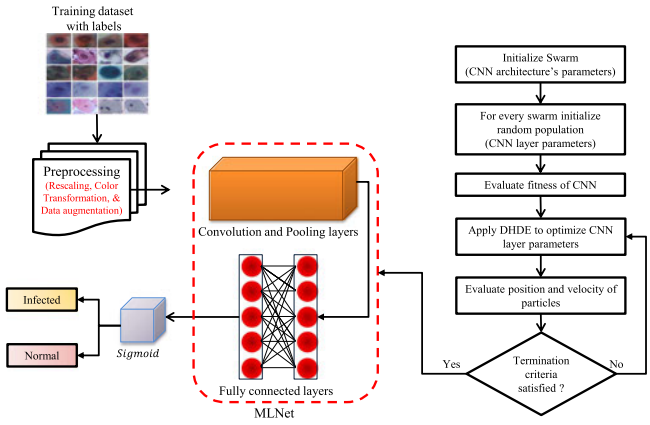


Fig. 1. MLNet: Metaheuristics-based lightweight deep learning network for cervical cancer diagnosis.

images. Finally, optimized CNN is applied for the classification of cervical cancer.

A. Optimization of CNN Architecture

Algorithm 1 describes the optimization of CNN architecture using PSO. A swarm $[SW_{p_1}, SW_{p_2}, \ldots, SW_{p_m}]$ of m particles is initialized randomly with specified ranges (line 3). Each particle SW_{p_i} has three-dimensions such as nCL, nPL, and nFL (refer Table I). Initially, $Best_{p_i}$ and $Best_g$ are set to null (line 2). From lines 4 to 6, similar particles are removed if exist in the swarm. From lines 7 to 25, PSO-CNN algorithm iteratively optimizes the hyper-parameters of CNN until M_{nt^1} is not reached. For every SW_{p_i}, optimal hyper-parameters such as nf_c, fs_c, pp_c, ss_c, fs_p, ss_p, pp_p, and F_{nn} (refer Table II) are returned by Algorithm 2 (DHDE algorithm) (line 11). These parameters save in O_{SA}. In line 13, CNN model is configured using current SW_{p_i} and R_{OS} (randomly selected optimal solution from O_{SA}). The fitness of CNN is evaluated in line 14. Based on the fitness value, $Best_{p_i}$ and $Best_g$ are updated (lines 15-20). In line 21, the position and velocity of the particle are updated.

Algorithm 1: PSO-CNN Algorithm.

Input : M_{nt^1} and search space for hyperparameters
Output: CNN hyperparameters: $nCL, nPL, nFL, nf_c, fs_c, pp_c, ss_c, fs_p, ss_p, pp_p$, and F_{nn}

1 **begin**
2 $Best_{p_i} = \phi$, $Best_g = \phi$;
3 Initialize the m particles (SW_{p_i}=$\{nCL, nPL, nFL\}$) of swarm randomly in specified range;
/* Remove similiar particles as: $\forall i, j \in m \ \& \ i \neq j$ */
4 **if** $SW_{p_i} == SW_{p_j}$ **then**
5 SW_{p_j} = NULL;
6 **end**
7 **while** $g \leq M_{nt^1}$ **do**
8 Evaluate ω using Eq. (3) ;
9 **for** $i = 1$ *to* m **do**
10 **if** $SW_{p_i} \neq NULL$ **then**
11 O_{SA} = Algorithm 2 (nCL, nPL, nFL);
12 Randomly select any optimal solution from O_{SA}, i.e., $R_{OS} = O_{SA}(rand)$;
13 Setup a CNN model: CNN(SW_{p_i}, R_{OS});
14 Evaluate fitness (f) using CNN;
15 **if** $f(SW_{p_i}) > f(Best_{p_i})$ *or* $Best_{p_i} == \phi$ **then**
16 $Best_{p_i} \leftarrow SW_{p_i}$;
17 **if** $f(Best_{p_i}) > f(Best_g)$ *or* $Best_g == \phi$ **then**
18 $Best_g \leftarrow Best_{p_i}$;
19 **end**
20 **end**
21 Update position (Pos_i) and velocity (Vel_i) of particle using Eq. (2) and Eq. (1), respectively;
22 **end**
23 **end**
24 **end**
25 **end**

B. Optimization of Hyper-Parameters of CNN's Layers

Table II shows the hyper-parameters of fully connected, pooling, and convolution layers of CNN. These hyper-parameters are optimized using DHDE algorithm. The process of optimization is illustrated in Algorithm 2. This algorithm is called by Algorithm 1 for every particle (SW_{p_i}). Algorithm 2 generates a archive of optimal solution (nf_c, fs_c, pp_c, ss_c, fs_p, ss_p, pp_p, and F_{nn}) for SW_{p_i}.

$SW_{p_i}\{nCL, nPL, nFL\}$, Population size S_P, crossover rate CR, scaling factor F, distance threshold d_t, minimal fitness threshold f_t, maximum counter M_C, maximum size of I_{OA} (M_{OA}), and maximum number of fitness evaluations ($M_{N_{fes}}$) are given as inputs to Algorithm 2. Initially, N_{fes} is set 0 and O_{SA} and I_{OA} are set to null. In line 3, a random population PO of S_P is generated. $PO = \{z_1, z_2, \ldots, z_{s_p}\}$, where $z_i = \{nf_c, fs_c, pp_c, ss_c, fs_p, ss_p, pp_p, F_{nn}\}$. In lines 4-5, CNN model is set and fitness is calculated. N_{fes} is set equal to S_P. Now, DHDE calls CDE-I_{OA} and INSDE one by one until N_{fes} is less than $M_{N_{fes}}$. In line 8, Algorithm 3 CDE-I_{OA} is called by passing PO, I_{OA}, N_{fes}, and SW_{p_i} parameters.

In Algorithm 3, three individuals are selected from PO randomly to produce an offspring t_i for every target parent z_i using mutation and crossover operations of DE [40] (see lines 2-3). For mutation, "DE/rand/1" [40] is operation is used. The fitness of t_i is evaluated using CNN model (lines 4-5). In line 9, Euclidean distance is used to find the nearest parent q from PO to t_i. In lines 10-13, if fitness of t_i comes better than q, then t_i replaces q. Otherwise, it is added into I_{OA}. In Algorithm 2, I_{OA} is added into PO to explore additional optimal solutions (line 12).

When the size of I_{OA} ($|I_{OA}|$) becomes greater or equal to M_{OA}, Algorithm 4 INSDE is called by using SW_{p_i}, PO, d_t,

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:50:11 UTC from IEEE Xplore. Restrictions apply.

Algorithm 2: DHDE Algorithm.

Input : $SW_{p_i}\{nCL, nPL, nFL\}$, S_P, CR, F, d_t, M_C, f_t, M_{OA}, M_{N_fes}
Output: PO and O_{SA}

begin
  $N_{fes} = 0$, $O_{SA} = \phi$, $I_{OA} = \phi$;
  Generate random population PO with S_P individuals;
  Setup a CNN model: CNN(SW_{p_i}, z_i);
  Evaluate $f(z_i)$ using CNN;
  $N_{fes} = N_{fes} + S_P$;
  while $N_fes < M_{N_fes}$ **do**
    Call Algorithm 3 CDE-I_{OA} (PO, I_{OA}, N_{fes}, SW_{p_i});
    $T = 1$;
    while $|I_{OA}| \geq M_{OA}$ & $N_{fes} < M_{N_{fes}}$ **do**
      if $T == 1$ **then**
        $PO = PO \cup I_{OA}$;
        $T = 0$;
      end
      Call Algorithm 4 INSDE(SW_{p_i}, PO, O_{SA}, d_t, f_t, M_C, N_{fes});
      if $|PO| < S_P$ **then**
        $I_{OA} = \phi$;
        generate $S_P - |PO|$ individuals randomly to fill PO;
      end
    end
  end
end

Algorithm 3: CDE-I_{OA} Algorithm.

Input : PO, I_{OA}, N_{fes}, SW_{p_i}
Output: PO and I_{OA}

for $i = 1$ *to* S_P **do**
  Select three individuals from PO randomly;
  Apply mutation and crossover of DE to generate an offspring t_i;
  Setup a CNN model: CNN(SW_{p_i}, t_i);
  Evaluate $f(t_i)$ using CNN;
  $N_{fes} = N_{fes} + 1$;
end
for $i = 1$ *to* S_P **do**
  Find the most similar individual q in PO to t_i;
  if $f(t_i) > f(q)$ **then**
    $q \leftarrow t_i$;
  else
    $I_{OA} = I_{OA} \cup t_i$;
  end
end

Algorithm 4: INSIDE Algorithm.

Input : SW_{p_i}, PO, O_{SA}, d_t, f_t, M_C, N_{fes}
Output: PO and O_{SA}

Individual PO's values are sorted from best to worst as PO_S;
for $i = 1$ *to* $\frac{S_P}{S_m}$ **do**
  Form sub_{pop_i} by finding the similar S_m individuals of the first individual from PO_S;
  Remove the used individuals from PO_S;
end
for $i = 1$ *to* $\frac{S_P}{S_m}$ **do**
  for $j = 1$ *to* $|sub_{pop_i}|$ **do**
    Use mutation and crossover of DE to obtain an offspring $t_{i,j}$ with in sub_{pop_i};
    Setup a CNN model: CNN($SW_{p_i}, t_{i,j}$);
    Evaluate $f(t_{i,j})$ using CNN;
    $N_{fes} = N_{fes} + 1$;
  end
end
$p = 1$;
for $i = 1$ *to* $\frac{S_P}{S_m}$ **do**
  for $j = 1$ *to* $|sub_{pop_i}|$ **do**
    if $f(t_{i,j}) > f(q_{i,j})$ **then**
      $C_p = 0$;
      $q_{i,j} \leftarrow t_{i,j}$;
    else
      $C_p = C_p + 1$;
    end
    $p = p + 1$;
  end
end
Merge every sub_{pop_i} into PO;
for $i = 1$ *to* S_p **do**
  Find Euclidean distances of ith individual to other individuals in PO, and save them in E_d;
  for $j = 1$ *to* S_p **do**
    if $i \neq j$ & $E_d(j) < d_t$ & $f(z_i) > f(z_j)$ & $f(z_i) - f(z_j) < f_t$ **then**
      Eliminate the jth individual from PO;
    end
  end
end
for $i = 1$ *to* $|PO|$ **do**
  if $C_i > M_C$ **then**
    Call Algorithm 4 as Archive_upadte(z_i, O_{SA}, 0.01);
    Eliminate the ith individual from PO;
  end
end

f_t, M_C, and N_{fes} parameters (line 15). In Algorithm 4, PO is sorted in descending order based on the fitness value and it becomes PO_S (line 1). PO_S is divided into S_m different sub-populations (sub_{pop}) (lines 2-4). For each sub_{pop_i}, an offspring $t_{i,j}$ is generated using mutation and crossover of DE (lines 6-8). The fitness of $t_{i,j}$ is evaluated using CNN (lines 9-10). The selection operation is performed from 15 to 19. A counter C_p ($p = 1, \ldots, S_P$) is assigned to every parent that records the

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:50:11 UTC from IEEE Xplore. Restrictions apply.

Algorithm 5: Archive_upadte Algorithm

```
Input  : z_i, O, e
Output: O
1  if |O| = 0 then
2      O = O ∪ z;
3      |O| = |O| + 1;
4  else
5      Find the nearest solution z_n ∈ O to z in decision
         space;
6      if ||z − z_n|| < e then
7          if f(z) > f(z_n) then
8              z_n ← z;
9          end
10     else
11         if |O| < M_O then
12             O = O ∪ z;
13             |O| = |O| + 1;
14         else
15             if f(z) > f(z_n) then
16                 z_n ← z;
17             end
18         end
19     end
20 end
```

times of parent not replaced with offspring. Initially, C_p is 0. In line 26, all sub_{pop_i} are merged into PO. In lines 27-32, Euclidean distance is calculated of every ith individual in PO and store into E_d. Thereafter, similar individuals are identified and eliminated from PO based on E_d, d_t, and f_t. In line 36, if the counter of the individual becomes greater than M_C, then it will be identified as an optimal solution and updated in O_{SA} using Algorithm 5.

In Algorithm 4, $|O|$ represents the size of O_{SA}. In lines 1-3, if $|O|$ is empty, then optimal solution z is added into the O and size is incremented. The nearest solution z_n to z is found in O (line 5). In line 6, crowding distance is calculated among z and z_n, and if it comes to less than $e = 0.01$, then the fitness of both solutions is compared. If the distance is greater than e, the optimal solution is added to the archive O. M_O represents the maximum size of O, i.e., $M_O = 5 \times S_P$. After execution of Algorithm 4, Algorithm 2 checks the size of PO ($|PO|$) (line 16). If $|PO|$ is less than S_p, then I_{OA} is set to empty. Again, Algorithm 3 (CDE-I_{OA}) will be used.

V. Performance Analysis

The proposed MLNet is implemented on 11th Gen Intel(R) Core(TM) $i5-11600$ @ $2.80\,GHz$ processor with $64\,GB$ RAM. MATLAB $2021a$ software is used. The proposed model is trained and tested on three benchmark datasets such as Herlev [42], SIPaKMeD [43], and Mendeley LBC [44] datasets. For all datasets, 65% cell images are used for training, 15% and 20% cell images are used for model validation and testing, respectively. To build more generalized models, Generative adversarial network (GAN)-based data augmentation (refer [45]) is also used on the training dataset. PSO and DHDE parameter settings are shown in Tables III and IV, respectively. Parameters for PSO and DHDE are set based on the information provided in [41] and [40], respectively.

TABLE III
Parameter Settings of PSO

Parameter	Value
Maximum number of iterations (M_{nt^1})	between 5 and 8
Cognitive coefficient (c_2)	2
Social coefficient (c_1)	2
Swarm size ($nPL \leq nCL$, $nFL \leq nCL$)	5

TABLE IV
Parameter Settings of DHDE

Parameter	Value
Population size (S_P)	80
Maximum size of I_{OA} (M_{OA})	$2S_P$
Minimal fitness threshold (f_t)	1.0E-5
Distance threshold (d_t)	0.1
Neighborhood size (S_m)	5
Crossover rate (CR)	0.3
Scaling Factor (F)	0.9
Maximum number of fitness evaluations ($M_{N_{fes}}$)	50,000

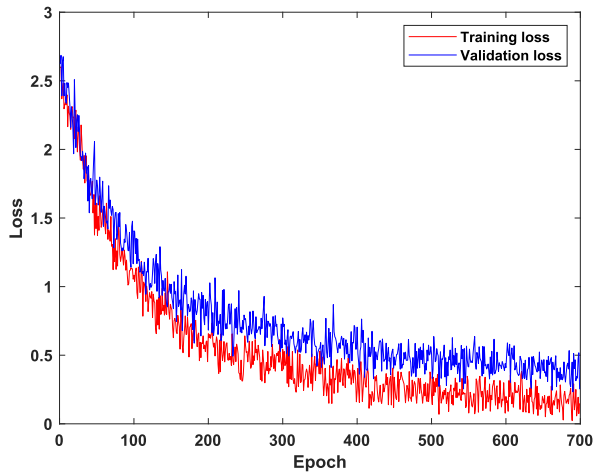


Fig. 2. Training and validation loss analyses of MLNet on Herlev dataset.

A. Comparative Analysis Based on Herlev Dataset

Fig. 2 shows the training and validation loss analyses of MLNet on the Herlev dataset. It is found that MLNet convergences at a good speed. Lesser difference between training and validation lines indicates that MLNet does not suffer from the overfitting problem. Overall the model achieves best training and validation accuracy as 99.92% and 99.67%, respectively.

A confusion matrix analysis of MLNet on the Herlev dataset is shown in Fig. 3. MLNet performs remarkably better on both normal and infected samples. Based on a better balancing of infected versus normal cases, MLNet yields better sensitivity

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:50:11 UTC from IEEE Xplore. Restrictions apply.

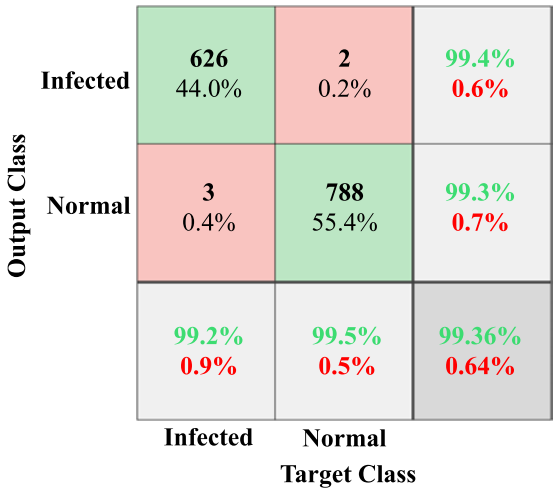


Fig. 3. Training and validation loss analyses of MLNet on Herlev dataset.

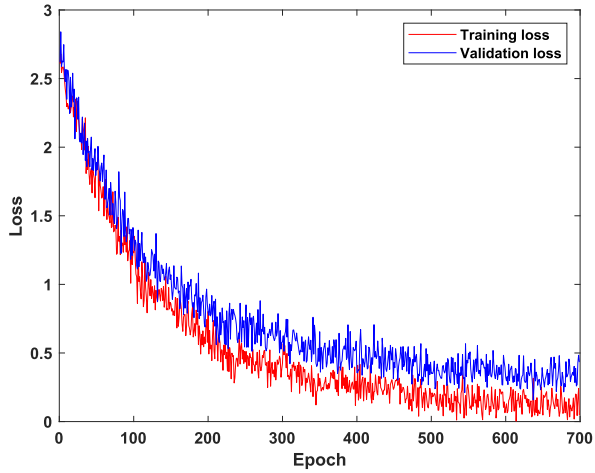


Fig. 4. Training and validation loss analyses of MLNet on SIPaKMeD dataset.

TABLE V
COMPARATIVE ANALYSIS BASED ON HERLEV DATASET

Models	Acc	Spec.	Sens./Rec	Pre	F1
R-CNN [9]	98.1	98.6	96.7	-	96.5
GAN [10]	98.1	98.6	96.7	-	-
LCNN [11]	73.58	-	-	-	-
PSPM [12]	**99.7**	99.3	99.8	-	-
DU-Net [13]	-	-	98.4	94.6	-
FCN [14]	98.8	-	95.8	96.1	-
DeepCervix [17]	98.91	-	98.0	99.25	98.5
InceptionV3 [18]	98.2	96.7	99.4	-	-
FCM [19]	99.27	96.53	**99.85**	-	-
VGG [24]	94.81	92.76	95.52	97.42	96.46
CNN [26]	98.32	-	97.65	98.66	98.12
CNN-ELM [30]	99.5	-	-	-	-
CNN-SVM [31]	99.3	99.4	98.9	-	-
MLNet	99.36	**99.53**	99.36	**99.35**	**99.28**

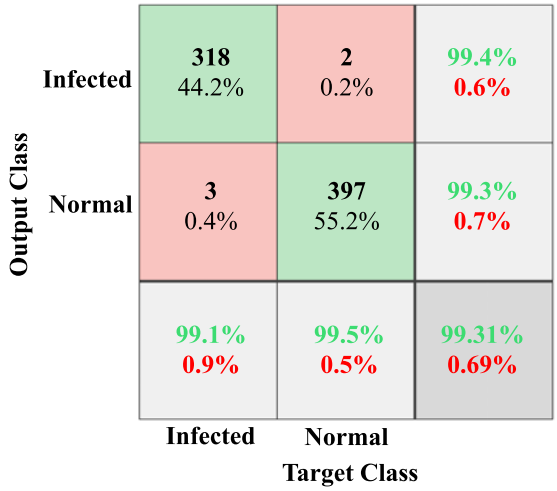


Fig. 5. Training and validation loss analyses of MLNet on SIPaKMeD dataset.

and specificity on the Herlev dataset. In total, MLNet achieves 99.36% accuracy.

Table V shows the comparative analysis among MLNet and the existing cervical cancer classification techniques on the Herlev dataset. It is found that each approach achieves remarkable results. PSPM [12] achieved remarkable accuracy (Acc) value as 99.7%. FCM [19] obtained better sensitivity (Sens.)/ recall (Rec) value as 99.85%. It is found that the proposed MLNet achieves remarkable performance in terms of accuracy, Specificity (Spec.), Sens./ Rec, precision (Pre) and F1-score (F1) as 99.36%, 99.36%, 99.53%, 99.32%, and 99.28%, respectively. Compared to PSPM [12] and FCM [19], MLNet achieves better specificity, precision and F1-score. Bold values indicate the higher performance.

B. Comparative Analysis Based on SIPaKMeD

Fig. 4 demonstrates the training and validation loss analyses of MLNet on the SIPaKMeD dataset. MLNet converges at a better speed. MLNet does not suffer from the overfitting problem since there are lesser differences between training and validation lines. The model achieves an overall accuracy of 99.98% and 99.76% on the training and validation dataset, respectively.

In Fig. 5, the confusion matrix of MLNet on the SIPaKMeD dataset is shown. MLNet obtains remarkable performance on normal and infected samples. As a result of better balancing of infected versus normal cases, MLNet shows higher sensitivity and specificity on the SIPaKMeD dataset. MLNet achieves an accuracy of 99.31% in total.

Table VI shows the comparative analysis among MLNet and the existing cervical cancer classification techniques on SIPaKMeD dataset. It is found that each technique achieved remarkable results. VGG [24] achieved remarkable specificity and F1-score values as 99.50% and 99.26%, respectively. GCNN [15] obtained better F1-score value as 99.3%. In terms of accuracy, specificity, sensitivity/recall, precision, and F1-score, the proposed MLNet performs remarkably better with values of 99.31%, 99.37%, 99.26%, 99.29%, and 99.24%. MLNet

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:50:11 UTC from IEEE Xplore. Restrictions apply.

TABLE VI
COMPARATIVE ANALYSIS BASED ON SIPAKMED DATASET

Models	Acc	Spec.	Sens./Rec	Pre	F1
GCNN [15]	98.3	98.6	98.8	-	**99.3**
ResNet50 [16]	98.8	98.7	99.1	-	98.7
EPDFE [22]	98.26	-	98.28	98.27	98.28
CNN-F [23]	98.55	-	98.52	98.57	98.54
VGG [24]	97.80	**99.50**	98.10	99.26	98.43
ResNet152 [25]	94.89	-	96.0	96.0	96.0
CNN [26]	97.87	-	99.12	98.56	98.89
MLNet	**99.31**	99.37	**99.26**	**99.29**	**99.24**

TABLE VII
COMPARATIVE ANALYSIS BASED ON MENDELEY LBC

Models	Acc	Spec.	Sens./ Rec	Pre	F1
EPDFE [22]	99.27	-	98.21	99.26	98.73
CNN-F [23]	99.23	-	99.23	99.13	99.18
CNN [26]	99.27	-	99.27	99.14	99.20
c-CNN [27]	96.89	-	93.75	93.38	94.15
MLNet	**99.36**	**99.47**	**99.32**	**99.35**	**99.32**

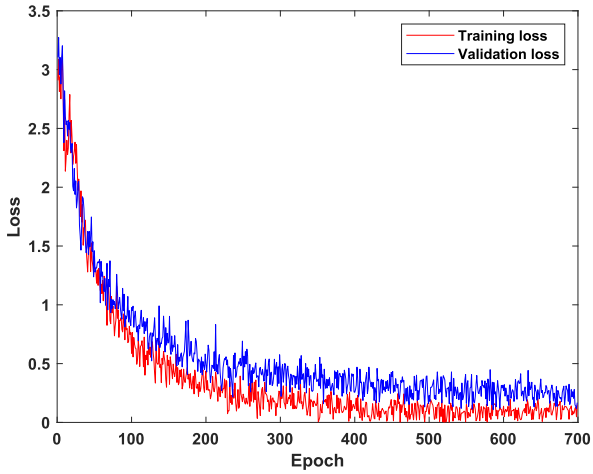


Fig. 6. Training and validation loss analyses of MLNet on Mendeley LBC dataset.

achieves higher accuracy, specificity, and precision values than GCNN [15] and ResNet50 [16].

C. Comparative Analysis Based on Mendeley LBC

Fig. 6 illustrates the training and validation losses of MLNet on the Mendeley LBC dataset. The convergence rate of MLNet is faster. Since there are fewer differences between training and validation lines, MLNet does not suffer from the overfitting problem. It achieves an accuracy of 100.00% and 99.76%, respectively on training and validation datasets.

Fig. 7 represents the confusion matrix of MLNet on Mendeley's LBC dataset. MLNet performs well on both infected and normal samples. MLNet demonstrates a high level of sensitivity and specificity on the Mendeley LBC dataset since it has a better balance between infected and normal cases. A total of 99.36% accuracy is achieved by MLNet.

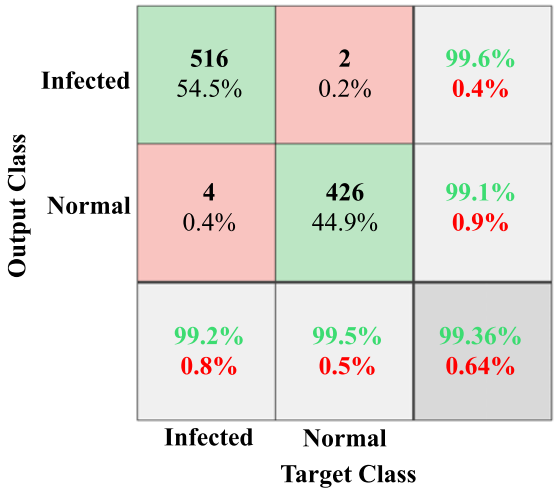


Fig. 7. Training and validation loss analyses of MLNet on Mendeley LBC dataset.

Comparative analysis of MLNet and existing cervical models on Mendeley LBC dataset is shown in Table V. It demonstrates that MLNet provides remarkable performances in terms of accuracy, specificity, sensitivity/recall, precision, and F1-score, at 99.76%, 99.87%, 99.57%, 99.68%, and 99.62%, respectively.

D. Discussion

MLNet successfully overcomes the over-fitting, parameter tuning, and gradient vanishing problems. Since it solves the hyper-parameter tuning problem of CNN by defining it as a multi-objective problem. PSO is used to optimally define CNN architecture and DHDE is utilized to optimize the hyper-parameters of CNN layers. Thus, MLNet achieves better performance than the competitive models. Since the majority of competitive models used a trial and error approach to define the architecture and respective hyper-parameters. The proposed model is compared with some well-known existing cervical cancer classification models by considering three different datasets.

On the Herlev dataset, PSPM [12] achieved better results among the competitive models in terms of accuracy, specificity, and sensitivity by 99.7%, 99.3%, and 99.8%, respectively. These competitive models includes R-CNN [9], GAN [10], LCNN [11], DU-Net [13], FCN [14], DeepCervix [17], InceptionV3 [18], FCM [19], VGG [24], CNN [26], CNN-ELM [30], and CNN-SVM [31]. Also, FCM [19] achieved significantly better sensitivity than the competitive models including PSPM [12]. But PSPM [12] and FCM [19] only limited to specific classes only. Whereas the proposed MLNet achieves better overall performance in terms of accuracy, f-measure, sensitivity, specificity, and precision by 1.6254%, 1.5178%, 1.5780%, 1.7145%, and 1.4890%, respectively.

For comparative analysis on SIPaKMeD dataset, various competitive models are considered such as GCNN [15], ResNet50 [16], EPDFE [22], CNN-F [23], VGG [24], ResNet152 [25], and CNN [26]. Among these models, VGG [24] achieved better performance in terms of accuracy, f-measure,

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:50:11 UTC from IEEE Xplore. Restrictions apply.

TABLE VIII
COMPUTATIONAL SPEED ANALYSIS OF MLNET BY VARYING THE IMAGE SIZE

Image size	Time (in milliseconds)
256×256	0.58 ± 0.17
512×512	0.81 ± 0.21
1024×1024	1.28 ± 0.32
2048×2048	1.51 ± 0.48
4096×4096	3.25 ± 0.59
8192×8192	5.16 ± 1.29

sensitivity, specificity, and precision by 97.80%, 99.50%, 98.10%, 99.26%, and 98.43, respectively. Whereas MLNet outperforms every model in terms of performance metrics except VGG [24] in terms of specificity. But overall in terms of other performance metrics, the proposed model is significantly superior to VGG [24]. Overall, MLNet obtains better performance on the SIPaKMeD dataset than the existing models in terms of accuracy, f-measure, sensitivity, specificity, and precision by 2.1250%, 2.2455%, 1.9074%, 1.9258%, and 1.8975%, respectively.

On Mendeley LBC, various models have been considered such as EPDFE [22], CNN-F [23], CNN [26], and c-CNN [27]. Overall, MLNet achieves better performance on the Mendeley LBC dataset in terms of accuracy, f-measure, sensitivity, specificity, and precision by 1.4680%, 1.5845%, 1.3582%, 1.3926%, and 1.4125%, respectively.

By scaling the image sizes, the computational speed analysis of MLNet is achieved (see Table VIII). The images have been scaled to different sizes, including 256×256, 512×512, 1024×1024, 2048×2048, 4096×4096, and 8192×8192. It is observed that MLNet can provide diagnosis results at a better computational speed.

VI. CONCLUSION

We proposed a MLNet to overcome overfitting, parameter tuning, and gradient vanishing problems with existing deep learning models. A multi-objective problem was defined to optimize the hyper-parameters of CNN. The architecture of CNN was optimized by PSO. DHDE was then applied to optimize the hyper-parameters of CNN layers. By concatenating the PSO's particle with the DHDE's solution, optimal CNN configuration was achieved. F-score was used to measure the performance of CNN. MLNet was trained and validated on three benchmark datasets. Performance analyses revealed that MLNet achieved better performance on Herlev, SIPaKMeD, and Mendeley LBC datasets in terms of accuracy, f-measure, sensitivity, specificity, and precision. The proposed MLNet is not limited only to cervical cancer diagnosis. In near future, the proposed MLNet can be utilized for the classification of various computer vision and biomedical problems. Besides one can use MLNet for other kind of datasets such as audio, text, etc. by changing the CNN layers including activation function.

REFERENCES

[1] W. Small Jr. et al., "Cervical cancer: A global health crisis," *Cancer*, vol. 123, no. 13, pp. 2404–2412, 2017.

[2] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *Cancer J. Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[3] T. A. Kessler, "Cervical cancer: Prevention and early detection," *Seminars Oncol. Nurs.*, vol. 33, pp. 172–183, 2017.

[4] M. E., "Pap-smear classification," M.S. thesis, Dept. Automat., 2003.

[5] D. C. Wilbur et al., "The becton Dickinson FocalPoint GS imaging system: Clinical trials demonstrate significantly improved sensitivity for the detection of important cervical lesions," *Amer. J. Clin. Pathol.*, vol. 132, no. 5, pp. 767–775, 2009.

[6] C. V. Biscotti et al., "Assisted primary screening using the automated ThinPrep imaging system," *Amer. J. Clin. Pathol.*, vol. 123, no. 2, pp. 281–287, 2005.

[7] K. Bora, M. Chowdhury, L. B. Mahanta, M. K. Kundu, and A. K. Das, "Automated classification of Pap smear images to detect cervical dysplasia," *Comput. Methods Programs Biomed.*, vol. 138, pp. 31–47, 2017.

[8] L. Zhang, L. Lu, I. Nogues, R. M. Summers, S. Liu, and J. Yao, "DeepPap: Deep convolutional networks for cervical cell classification," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 6, pp. 1633–1643, Nov. 2017.

[9] K. H. S. Allehaibi et al., "Segmentation and classification of cervical cells using deep learning," *IEEE Access*, vol. 7, pp. 116925–116941, 2019.

[10] Z. Lin, Z. Gao, H. Ji, R. Zhai, X. Shen, and T. Mei, "Classification of cervical cells leveraging simultaneous super-resolution and ordinal regression," *Appl. Soft Comput.*, vol. 115, 2022, Art. no. 108208.

[11] W. Chen, L. Gao, X. Li, and W. Shen, "Lightweight convolutional neural network with knowledge distillation for cervical cells classification," *Biomed. Signal Process. Control*, vol. 71, 2022, Art. no. 103177.

[12] K. Sabeena and C. Gopakumar, "A hybrid model for efficient cervical cell classification," *Biomed. Signal Process. Control*, vol. 72, 2022, Art. no. 103288.

[13] J. Zhao, Q. Li, X. Li, H. Li, and L. Zhang, "Automated segmentation of cervical nuclei in Pap smear images using deformable multi-path ensemble model," in *Proc. IEEE 16th Int. Symp. Biomed. Imag.*, 2019, pp. 1514–1518.

[14] E. Hussain, L. B. Mahanta, C. R. Das, M. Choudhury, and M. Chowdhury, "A shape context fully convolutional neural network for segmentation and classification of cervical nuclei in Pap smear images," *Artif. Intell. Med.*, vol. 107, 2020, Art. no. 101897.

[15] J. Shi, R. Wang, Y. Zheng, Z. Jiang, H. Zhang, and L. Yu, "Cervical cell classification with graph convolutional network," *Comput. Methods Programs Biomed.*, vol. 198, 2021, Art. no. 105807.

[16] J. Li et al., "Cervical cell multi-classification algorithm using global context information and attention mechanism," *Tissue Cell*, vol. 74, 2022, Art. no. 101677.

[17] M. M. Rahaman et al., "DeepCervix: A deep learning-based framework for the classification of cervical cells using hybrid deep feature fusion techniques," *Comput. Biol. Med.*, vol. 136, 2021, Art. no. 104649.

[18] N. Dong, L. Zhao, C.-H. Wu, and J.-F. Chang, "Inception V3 based cervical cell classification combined with artificially extracted features," *Appl. Soft Comput.*, vol. 93, 2020, Art. no. 106311.

[19] T. Chankong, N. Theera-Umpon, and S. Auephanwiriyakul, "Automatic cervical cell segmentation and classification in Pap smears," *Comput. Methods Programs Biomed.*, vol. 113, no. 2, pp. 539–556, 2014.

[20] P. Wang, L. Wang, Y. Li, Q. Song, S. Lv, and X. Hu, "Automatic cell nuclei segmentation and classification of cervical Pap smear images," *Biomed. Signal Process. Control*, vol. 48, pp. 93–103, 2019.

[21] V. Martin, T. H. Kim, M. Kwon, M. Kuko, M. Pourhomayoun, and S. Martin, "A more comprehensive cervical cell classification using convolutional neural network," *J. Amer. Soc. Cytopathol.*, vol. 7, no. 5, 2018, Art. no. S66.

[22] O. Yaman and T. Tuncer, "Exemplar pyramid deep feature extraction based cervical cancer image classification model using Pap-smear images," *Biomed. Signal Process. Control*, vol. 73, 2022, Art. no. 103428.

[23] A. Manna, R. Kundu, D. Kaplun, A. Sinitca, and R. Sarkar, "A fuzzy rank-based ensemble of CNN models for classification of cervical cytology," *Sci. Rep.*, vol. 11, no. 1, pp. 1–18, 2021.

[24] H. Chen et al., "CytoBrain: Cervical cancer screening system based on deep learning technology," *J. Comput. Sci. Technol.*, vol. 36, no. 2, pp. 347–360, 2021.

[25] A. Tripathi, A. Arora, and A. Bhan, "Classification of cervical cancer using deep learning algorithm," in *Proc. 5th IEEE Int. Conf. Intell. Comput. Control Syst.*, 2021, pp. 1210–1218.

[26] H. Basak, R. Kundu, S. Chakraborty, and N. Das, "Cervical cytology classification using PCA and GWO enhanced deep features selection," *SN Comput. Sci.*, vol. 2, 2021, Art. no. 369.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:50:11 UTC from IEEE Xplore. Restrictions apply.

[27] N. K. Chauhan and K. Singh, "Impact of variation in number of channels in CNN classification model for cervical cancer detection," in *Proc. 9th Int. Conf. Rel. Infocom Technol. Optim.*, 2021, pp. 1–6.
[28] J. Martínez-Más et al., "Classifying papanicolaou cervical smears through a cell merger approach by deep learning technique," *Expert Syst. Appl.*, vol. 160, 2020, Art. no. 113707.
[29] A. Pal et al., "Deep multiple-instance learning for abnormal cell detection in cervical histopathology images," *Comput. Biol. Med.*, vol. 138, 2021, Art. no. 104890.
[30] A. Ghoneim, G. Muhammad, and M. S. Hossain, "Cervical cancer classification using convolutional neural networks and extreme learning machines," *Future Gener. Comput. Syst.*, vol. 102, pp. 643–649, 2020.
[31] A. D. Jia, B. Z. Li, and C. C. Zhang, "Detection of cervical cancer cells based on strong feature CNN-SVM network," *Neurocomputing*, vol. 411, pp. 112–127, 2020.
[32] Z. Alyafeai and L. Ghouti, "A fully-automated deep learning pipeline for cervical cancer classification," *Expert Syst. Appl.*, vol. 141, 2020, Art. no. 112951.
[33] T. Zhang et al., "Cervical precancerous lesions classification using pre-trained densely connected convolutional networks with colposcopy images," *Biomed. Signal Process. Control*, vol. 55, 2020, Art. no. 101566.
[34] Y. Xiang, W. Sun, C. Pan, M. Yan, Z. Yin, and Y. Liang, "A novel automation-assisted cervical cancer reading method based on convolutional neural network," *Biocybernetics Biomed. Eng.*, vol. 40, no. 2, pp. 611–623, 2020.
[35] M. F. Zorkafli, M. K. Osman, I. S. Isa, F. Ahmad, and S. N. Sulaiman, "Classification of cervical cancer using hybrid multi-layered perceptron network trained by genetic algorithm," *Procedia Comput. Sci.*, vol. 163, pp. 494–501, 2019.
[36] C. Bhavani and A. Govardhan, "Cervical cancer prediction using stacked ensemble algorithm with SMOTE and RFERF," *Mater. Today Proc.*, 2021.
[37] N. Akhtar and U. Ragavendran, "Interpretation of intelligence in CNN-pooling processes: A methodological survey," *Neural Comput. Appl.*, vol. 32, no. 3, pp. 879–898, 2020.
[38] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. Int. Conf. Neural Netw.*, 1995, pp. 1942–1948.
[39] D. Tian and Z. Shi, "MPSO: Modified particle swarm optimization and its applications," *Swarm Evol. Computation*, vol. 41, pp. 49–68, 2018.
[40] K. Wang, W. Gong, L. Deng, and L. Wang, "Multimodal optimization via dynamically hybrid niching differential evolution," *Knowl.-Based Syst.*, vol. 238, 2022, Art. no. 107972.
[41] P. Singh, S. Chaudhury, and B. K. Panigrahi, "Hybrid MPSO-CNN: Multi-level particle swarm optimized hyperparameters of convolutional neural network," *Swarm Evol. Computation*, vol. 63, 2021, Art. no. 100863.
[42] J. Jantzen, "Pap-smear (DTU/HERLEV) databases." 2008. Accessed: Dec. 17, 2021. [Online]. Available: mde-lab.aegean.gr/downloads
[43] M. E. Plissiti, P. Dimitrakopoulos, G. Sfikas, C. Nikou, O. Krikoni, and A. Charchanti, "Sipakmed: A new dataset for feature and image based classification of normal and pathological cervical cells in Pap smear images," in *Proc. 25th IEEE Int. Conf. Image Process.*, 2018, pp. 3144–3148.
[44] E. Hussain, "Liquid based cytology Pap smear images for multi-class diagnosis of cervical cancer," Mendeley Data, V4. 2019. Accessed: Dec. 28, 2021, doi: 10.17632/zddtpgzv63.4. [Online]. Available: data.mendeley.com/datasets/zddtpgzv63/4
[45] S. Yu et al., "Generative adversarial network based data augmentation to improve cervical cell classification model," *Math. Biosci. Eng.*, vol. 18, pp. 1740–1752, 2021.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:50:11 UTC from IEEE Xplore. Restrictions apply.

Efficient Evolving Deep Ensemble Medical Image Captioning Network

Dilbag Singh, *Senior Member, IEEE*, Manjit Kaur, *Senior Member, IEEE*, Jazem Mutared Alanazi, Ahmad Ali AlZubi, and Heung-No Lee, *Senior Member, IEEE*

***Abstract*—With the advancement in artificial intelligence (AI) based E-healthcare applications, the role of automated diagnosis of various diseases has increased at a rapid rate. However, most of the existing diagnosis models provide results in a binary fashion such as whether the patient is infected with a specific disease or not. But there are many cases where it is required to provide suitable explanatory information such as the patient being infected from a particular disease along with the infection rate. Therefore, in this paper, to provide explanatory information to the doctors and patients, an efficient deep ensemble medical image captioning network (DCNet) is proposed. DCNet ensembles three well-known pre-trained models such as VGG16, ResNet152V2, and DenseNet201. Ensembling of these models achieves better results by preventing an overfitting problem. However, DCNet is sensitive to its control parameters. Thus, to tune the control parameters, an evolving DCNet (EDC-Net) was proposed. Evolution process is achieved using the self-adaptive parameter control-based differential evolution (SAPCDE). Experimental results show that EDC-Net can efficiently extract the potential features of biomedical images. Comparative analysis shows that on the Open-i dataset, EDC-Net outperforms the existing models in terms of BLUE-1, BLUE-2, BLUE-3, BLUE-4, and kappa statistics (KS) by 1.258%, 1.185%, 1.289%, 1.098%, and 1.548%, respectively.**

***Index Terms*—Medical, diagnosis, pre-trained models, explanatory information.**

I. Introduction

With the advancement in artificial intelligence (AI) based E-healthcare applications, the role of automated diagnosis of various diseases has increased at a rapid rate. Deep learning models have recently been used by many researchers to classify patients suffering from a particular disease. But these models generally provide results in a binary fashion such as whether the patient is infected from a specific disease or not. However, there are many cases where it is required to provide suitable information to the patients such as the patient being infected with some disease with this much infection rate. Therefore, these days many researchers have started utilizing image captioning techniques to provide explanatory information to doctors and patients.

Recently, semantic concepts were used to detect the captions from the image. Image-caption pairs were used to train the concept detector. However, it suffers from vocabulary discrepancy and deficiency of required information [1]. To increase the accuracy of pathological information in diagnostic reports, Semantic fusion networks (SFNet) were utilized [2]. Attention-based models were also used in image captioning. The attention masks were queried using hidden states of LSTM from image features. It provided better image information for training deep sequential networks. However, these models did not ensure that layers significantly focused on regions of interest due to indirectly supervised learning [3]. These problems were overcome by utilizing differentiable neural networks to obtain the captions from the images [4]. To describe the disease information in ultrasound images, LSTM was used to decode the encoding vectors [5].

Most image captioning techniques are based on visual information and rarely rely on semantic content. Therefore, it is also needed to express the emotions of text descriptions for better captioning. Yang et al. [6] made sentences using both emotional and visual information by combining latent codes. The quality of AI-generated diagnostic reports is equally important as the development of models. However, [6] performs poorly for images with poor visibility.

Babar et al. [7] proposed a new measure to evaluate the quality of generated diagnostic reports. Convolutional neural network (CNN) -based image captioning models can retain only a few features of the original image. The image captioning DCNet based on a recurrent neural network (RNN) suffers from a gradient vanishing problem. The multimodal fusion method solved these issues by using CNN and RNN into the same DCNet for image captioning [8]. A semi-supervised deep generative DCNet generated the description more naturally from images [9]. An adaptive multimodal attention network was also designed to produce better quality captioned images [10].

Manuscript received 4 February 2022; revised 22 August 2022; accepted 30 September 2022. Date of publication 18 November 2022; date of current version 6 February 2023. This work was supported in part by the Ministry of Science and ICT, Korea, through the Information Technology Research Center Support Program under Grant IITP-2021-0-01835 supervised by the Institute for Information and Communications Technology Planning and Evaluation, in part by the National Research Foundation of Korea funded by the Korean Government, MSIP, under Grant NRF-2021R1A2B5B03002118, and in part by the Researchers Supporting Project, King Saud University, Riyadh, Saudi Arabia, under Grant RSP-2021/395. *(Corresponding author: Heung-No Lee.)*

Dilbag Singh, Manjit Kaur, and Heung-No Lee are with the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea (e-mail: dggill2@gmail.com; manjitkaur@gist.ac.kr; heungno@gist.ac.kr).

Jazem Mutared Alanazi and Ahmad Ali AlZubi are with the Computer Science Department, Community College, King Saud University, Riyadh 11451, Saudi Arabia (e-mail: ajazem@ksu.edu.sa; aalzubi@ksu.edu.sa).

Digital Object Identifier 10.1109/JBHI.2022.3223181

2168-2194 © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:55:33 UTC from IEEE Xplore. Restrictions apply.

Kavur et al. [11] utilized various ensemble models such as majority voting, average com-biner, product combiner, and min/max combiner. To prevent overfitting problems with deep learning models, U-Net, deepmedic, V-Net, and dense V-networks were ensembled to enhance the liver segmentation from CT images.

However, it is found that the hyper-parameters selection of the existing models was achieved using the trial-and-error basis. Additionally, the designed model suffers from the over-fitting and gradient vanishing problem. To overcome these problems, an efficient EDC-Net model is designed.

The main contributions of this paper are as follows:

1) To provide explanatory information to the doctors and patients, an efficient deep ensemble medical image captioning network (DCNet) is proposed.
2) The proposed DCNet ensembles three well-known pre-trained models such as VGG16, ResNet152V2, and DenseNet201.
3) To tune the control parameters, evolving DCNet (EDC-Net) was proposed. The evolution process is achieved using the self-adaptive parameter control-based differential evolution (SAPCDE)
4) By considering benchmark datasets, extensive experiments are conducted.

The remainder of the paper is organized as follows: Section II describes related work. Section III provides a mathematical formulation of EDC-Net. Section IV presents the comparative analysis. Section V concludes the paper.

II. Related Work

Yu et al. [12] used order embedding to caption the topic-oriented image. A convolutional neural network (CNN)-based classifier is used to select the topics for images from candidates. Wang et al. [13] captioned the images using a recurrent memory network (RMN). In training, topic words were recorded from a topic repository. After that, the image was tested using the retrieval method for the generation of a topic word. Finally, the sentence was generated through a recurrent memory network by incorporating the retrieved topic words. Zhang et al. [1] proposed an image captioning DCNet using missing concepts mining and online positive recall. The suitable captions were automatically generated using the element-wise selection method. Zhang et al. [3] captioned the images using the visual aligning attention DCNet (VAA). The visual aligning loss was designed to optimize the attention layer in the training stage. The non-visual words were filtered out using visual vocab to minimize their effect on the attention layer. Zhou et al. [14] enhanced single-phase image captioning using a saliency-enhanced re-captioning model. In this, two-phase learning was applied to get better results. In the first phase, semantic and visual saliency cues were extracted from the model. In the second phase, cues were fused for the self-boosting of the model. Zhao et al. [15] captioned the images using DCNet adaption and cross-modal retrieval in cross-domain. Firstly, the source domain was used to pre-train the cross-modal and then utilized in the target domain to extract the initial image-sentence pairs. These pairs were further reefed using a retrieval model. Yang et al. [16] proposed cross-domain image captioning using Multitask learning (ML). Textual explanations of the images were generated using CNN-LSTM. Images were synthesized using a Conditional generative adversarial network (C-GAN) based on generated text descriptions. Hoxha et al. [17] captioned the images using CNN and Recurrent neural network (RNN). Firstly, visual features were extracted and translated into a textual explanation. Secondly, embedding techniques were used to convert the textual explanation into feature vectors. Finally, similar images were retrieved by calculating the similarity between vectors of textual explanation and archive images. Huang et al. [18] utilized a Multimodal attribute detector (MAD) and Subsequent attribute predictor (SAP) to improve the performance of image captioning. MAD used image features as well as word embedding to improve attribute detection accuracy. A concise attribute was predicted with SAP every time to reduce the diversity of image attributes. Yang et al. [19] proposed a Dual generator generative adversarial network (DGGAN) based image captioning model. This technique ensembled the generation-based and retrieval-based image captioning models. Yuan et al. [20] captioned the images using multi-label attribute graph convolution and multi-level attention. The attention module focused on both specific spatial and scale features. For image captioning, attribute features were learned using the attribute graph convolution module. Huang et al. [21] utilized the Denoising-based multi-scale feature fusion (DMSFF) technique to caption the images. Monay et al. [22] captioned the images using the probabilistic latent semantic analysis (PLSA) model. The textual and visual modalities are assumed equally by an expectation-maximization algorithm. Yu et al. [23] proposed an image captioning DCNet using Multimodal Transformer (MT) model. Multi-view visual features were also introduced to improve the performance. Xian et al. [24] used multimodal LSTM to caption the images. Amirian et al. [25] proposed a deep learning-based image captioning model. Li et al. [26] designed an efficient image captioning model that can be used to extract potential information from digital images. It has shown effective performance over competitive models. But this model suffers from the over-fitting issue.

Park et al. [27] designed a multi-difference average pooling LSTM (mDiAP-LSTM)-based medical image captioning model. The most appropriate method for capturing the differences was determined through a study of feature representation methods. Hou et al. proposed an efficient Full-adversarial reinforcement learning (Full-ARL) model for medical image captioning. Li et al. [28] designed an efficient knowledge-driven encode, retrieve, paraphrase (KERP) model to obtain better medical image captions. It decomposes captions into explicit disease-related abnormality graphs, which were then analyzed using natural language modeling. Hou et al. [29]. An overall score based on accuracy and fluency was provided by two additional discriminators as the evaluator. A report generator was implemented, which produces discrete sequences of words based on decision probabilities, as opposed to generative adversarial networks (GANs) used in image generation. Furthermore, it prevented

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:55:33 UTC from IEEE Xplore. Restrictions apply.

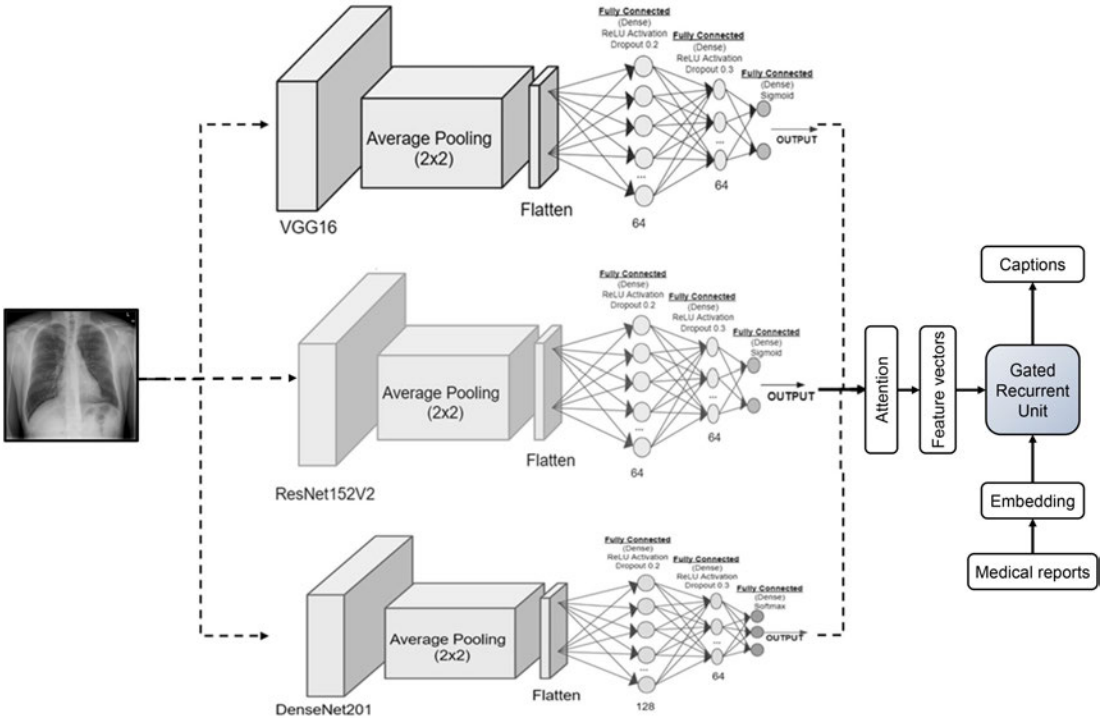


Fig. 1. Proposed ensemble deep transfer network with gated recurrent unit.

the gradients from being transmitted between discriminators and generators.

Wang et al. [30] designed an efficient relational-topic retrieval and generation framework (R-paraNet). Reports incorporated semantically consistent medical terms and encourage the generation of sentences for rare abnormal descriptions. Li et al. [31] designed a hybrid retrieval-generation reinforced agent (HRGR-Agent) that integrated retrieval-based strategies with sophisticated learning techniques to provide robust image captions. Hierarchical decision-making was employed by HRGR-Agent. To obtain a caption, a low-level generation module invoked either a high-level retrieval policy or a high-level retrieval policy module. Reinforcement learning guided the HRGR-Agent's updates via rewards at the word and sentence levels. Jing et al. [32] designed a multitask learning model that predicted tags and paragraphs simultaneously. It utilized co-attention to locate abnormal regions and generate narrations for them. The hierarchical LSTM (HLSTM) was proposed for generating long paragraphs.

It is observed that the existing models suffer from hyperparameters tuning problems. Moreover, the designed model exhibits over-fitting and gradient vanishing problems.

III. Proposed Model

This section discusses the proposed medical image captioning model.

A. Ensemble Deep Transfer Model

An ensemble deep transfer network (DCNet) is proposed. VGG16 [33], ResNet152V2 [34], and DenseNet201 [35] models have been used to develop the proposed model. Fig. 1 shows the EDC-Net with a gated recurrent unit. Ensembling of models provides better results by preventing over-fitting [36], [37]. It also enhances the extraction of features and performance of supervised models [38], [39]. Fig. 1 demonstrates the proposed ensemble deep transfer model. 128 neurons have been utilized for input dense layer [40]. VGG16 [33], ResNet152V2 [34], and DenseNet201 [35] models are used to obtain potential features. These models have been trained using 20 epochs with a batch size of 8. Fully connected layers [41] having size 128 neurons along with dropouts of 0.2 and 0.25, respectively have been utilized to prevent memorization problems with the competitive models. $l_r = 0.001$ has been utilized as learning rate.

B. Gated Recurrent Unit

Cho et al. [42] designed gated recurrent unit (GRU). GRU allows every recurrent unit to dynamically obtain dependencies of different time scales. Similar to LSTM, GRU has gating units that modulate the knowledge flow within the unit, but, without utilizing any additional memory cells.

Activation (α_n^k) of GRU at time t is a linear interpolation among candidate activation ($\tilde{\alpha}_n^k$) and succeeding activation (α_{t-1}^k). α_n^k can be computed as:

$$\alpha_n^k = (1 - \beta_n^k)\alpha_{t-1}^k + \beta_n^k \tilde{\alpha}_n^k, \tag{1}$$

Here, update gate (β_n^k) monitors and controls the activation. An update gate can be evaluated as:

$$\beta_n^k = \sigma \left(W_z \gamma_n + U_z \alpha_{t-1}\right)^k. \tag{2}$$

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:55:33 UTC from IEEE Xplore. Restrictions apply.

Here, $W_r\gamma_n$ represents the weighted matrix. GRU is unable to limit the degree to which its state is exposed. However, it can expose the whole state during every iteration.

A candidate activation ($\tilde{\alpha}_n^k$) can be computed [43] as:

$$\tilde{\alpha}_n^k = \tanh\left(W_r\gamma_n + U\left(\mathbf{r}_n \odot \alpha_{t-1}\right)\right)^k \tag{3}$$

Here, $\odot$ shows an element-wise multiplication. $\mathbf{r}_n$ contains a group of reset gates.

When r_n^k approaches 0, the reset gate can efficiently develop a unit act as if it is utilizing the first symbol of the input sequence by forgetting the earlier evaluated state.

Reset gate (r_n^k) can be evaluated according to the update gate as:

$$r_n^k = \sigma\left(W_r\gamma_n + U_r\alpha_{t-1}\right)^k. \tag{4}$$

C. Adaptive Differential Evolution

The proposed DCNet suffers from the hyper-parameters tuning issue. Therefore, in this paper, a differential evolution variant is used to evolve the proposed model. Bilingual Evaluation Understudy (BLUE) [44] is used as an objective function. Differential evolution (DE) is used in various fields due to its advantages such as strong robustness, good performance, and simple structure to solve optimization problems. The performance of DE mainly depends on the selection of control parameters (crossover rate CR, scale factor F, and population size NZ) and trial vector generation strategy (crossover and mutation). According to the problem nature, these parameters should be selected for better optimization results. The setting of these parameters is a challenging task for any problem. The solution to this problem is given by Cui et al. [45] by proposing self-adaptive parameter control-based DE (SAPCDE). It is based on the idea that good parameters should be propagated from generation to generation and the bad parameters should learn from the good parameters. In SAPCDE, two populations are utilized i.e., the solution population and the parameter population. Every solution has its control parameters. Parameter population is also evolved with each generation. SAPCDE is a combination of basic DE and parameter self-adaptation control methods.

Lets suppose that initial parameter population for DCNet is represented as $C^0 = \{C_1^0, C_2^0, \ldots, C_{NZ}^0\}$ with $C_i^0 = \{F_{i,1}^0, CR_{i,2}^0\}$, where NZ denotes the population size. The initial solution population is represented as $S^0 = \{S_1^0, S_2^0, \ldots, S_{NZ}^0\}$ with $S_i^0 = \{s_{i,1}^0, s_{i,2}^0, \ldots, s_{i,D}^0\}$, where D represents the number of variables. G_n represents the number of generations. The parameter population evolved in the same way as the solution population in DE. Initially, parameter population is generated uniformly and randomly between [0, 0] and [1, 1]. Thereafter, for each individual $C_i^{G_n}$, mutation parameter ($VC_i^{G_n}$) is generated using mutation operator such as

$$VC_i^{G_n} = C_{r1}^{G_n} + CF \cdot (C_{r2}^{G_n} - C_{r3}^{G_n}) \tag{5}$$

$$VC_i^{G_n} = C_{r1}^{G_n} + CF \cdot (SC_k^{G_n} - C_{r2}^{G_n}) \tag{6}$$

Here, C_{r1}, C_{r2}, and C_{r3} are selected from parameter population randomly. $SC_k^{G_n}$ denotes a good parameter that is selected randomly. Next, the trailing parameter ($TC_i^{G_n}$) is generated using crossover operator such as

$$TC_i^{G_n} = \begin{cases} VC_{i,j}^{G_n}, & \text{if } (rand_{i,j} \le CCR \quad \text{or} \quad j == j_{rand}) \\ C_{i,j}^{G_n}, & Otherwise \end{cases} \tag{7}$$

where $j = \{1, 2\}$ and $i = \{1, 2, \ldots, NZ\}$. $rand_{i,j} \in [0, 1]$ represents the uniform random number. $NCR \in [0, 1]$ denotes the new CR. Lastly, the selection operator is applied to select the good parameter for the next generation. In SAPCDE, good parameter individual $C_i^{G_n}$ is one which helps the $S_i^{G_n}$ to produce better offspring $TS_i^{G_n}$. Otherwise, $C_i^{G_n}$ is considered as a bad control parameter. The selection operator for a good parameter is defined as

if $C_i^{G_n}$ is a good parameter

$$C_i^{G_{n+1}} = \begin{cases} C_i^{G_n}, & \text{if } \ rand(0,1) < \lambda_1 \\ TC_i^{G_{n+1}}, & \text{Otherwise} \end{cases} \tag{8}$$

else

$$C_i^{G_{n+1}} = \begin{cases} TC_i^{G_{n+1}}, & \text{if } \ rand(0,1) < \lambda_2 \\ C_i^{G_n}, & \text{Otherwise} \end{cases} \tag{9}$$

where λ_1 and λ_2 control the values of parameters to explore the new values and keep the previous values. The basic working of SAPCDE is illustrated in Algorithm 1. Initially, solution and parameter populations are generated in line 1 and line 2, respectively. In line 3, number of generation (G_n) is initialized to 1. F_{eval} represents the number of function evaluations. Line 5 represents the termination condition of the algorithm when F_{eval} reaches $F_{\max}$. GP stores the good parameter individuals (line 6). It is initialized with 0. The solution population is evolved through lines 7 to 16. In line 7, the mutation operator is applied to generate a mutant vector ($SV_i^{G_n}$) using parameter individual $C_i^{G_n}$ such as

$$SV_i^{G_n} = S_{r1}^{G_n} + F_{i,1}^{G_n} \cdot (S_{r2}^{G_n} - S_{r3}^{G_n}) \tag{10}$$

where S_{r1}, S_{r2}, and S_{r3} are randomly selected from solution population. $F_{i,1}^{G_n} \in C_i^{G_n}$ represents the scale factor. In line 8, crossover operator is applied to obtain a trial vector ($TS_i^{G_n}$) using C_iG_n such as

$$TS_i^{G_n} = \begin{cases} SV_{i,j}^{G_n}, & \text{if } (rand_{i,j} \le CR_{i,2} \quad \text{or} \quad j == j_{rand}) \\ S_{i,j}^{G_n}, & Otherwise \end{cases} \tag{11}$$

where $CR_{i,2} \in C_i^{G_n}$ and $i = \{1, 2, \ldots, NZ\}$. A selection operator is applied to select the best solution (lines 10–15). If fitness of $TS_i^{G_n}$ is better than $S_i^{G_n}$, the corresponding $C_i^{G_n}$ is considered as a good parameter individual and flagged as 1 (i.e., $val = 1$) (line 11). It is also added in the GP in line 12. If $C_i^{G_n}$ is a bad parameter, then it is flagged as 0 (line 14). Next, the parameter population is evolved using mutation, crossover, and selection operators (lines 17–37). If $GP = 0$ at any G_n, then bad parameters are initialized randomly (line 31) to explore the new values.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:55:33 UTC from IEEE Xplore. Restrictions apply.

Algorithm 1: Self-Adaptive Parameter Control-Based Differential Evolution for Evolving Ensemble Model.

Generate initial solution (S^0) and parameter population (C^0)
Set $G_n = 1$ and $F_{eval} = 0$
while $F_{eval} < F_{max}$ **do**
  Set $GP = 0$
  for $i = 0$ *to* NZ **do**
    Mutant vector $SV_i^{G_n}$ is obtained using Eq. 10 and $C_i^{G_n}$
    Trial vector $TS_i^{G_n}$ is obtained using Eq. 11 and $C_i^{G_n}$
    if $f(TS_i^{G_n}) \geq f(S_i^{G_n})$ **then**
      $S_i^{G_{n+1}} = TS_i^{G_n}, val(i) = 1$
      Put $C_i^{G_n}$ into GP
    else
      $S_i^{G_{n+1}} = S_i^{G_n}, val(i) = 0$
    end
  end
  for $i = 1$ *to* NZ **do**
    if $val(i) == 1$ **then**
      if $rand(0,1) < \lambda_1$ **then**
        $C_i^{G_{n+1}} = C_i^{G_n}$
      else
        Generate $TC_i^{G_n}$ using Eq. (5) and Eq. (7)
        $C_i^{G_{n+1}} = TC_i^{G_n}$
      end
    else
      if $rand(0,1) < \lambda_2$ **then**
        if $GP \neq 0$ **then**
          Generate $TC_i^{G_n}$ using Eqs. (6) and (7)
          $C_i^{G_{n+1}} = TC_i^{G_n}$
        else
          initialize $C_i^{G_{n+1}}$ randomly
        end
      else
        $C_i^{G_{n+1}} = C_i^{G_n}$
      end
    end
  end
  $G_n = G_{n+1}$
end

IV. Performance Analysis

In this paper, a dataset is obtained from Open-i [46]. It contains chest X-ray images and radiology text reports. Each image is paired with respective captions. 7,471 chest X-ray images are available with frontal and lateral views of given patients. Additionally, VGG16 [33], ResNet152V2 [34], and DenseNet201 [35] models are also ensembled using majority voting (MV), min, and max combiner. These models are renamed as MVE_1, Min_4, and Max_5, respectively. Also, inspired from [47], during the evolution phase, validation data is used to evaluate the performance of EDC-Net. The final trained EDC-Net is then tested on the training dataset.

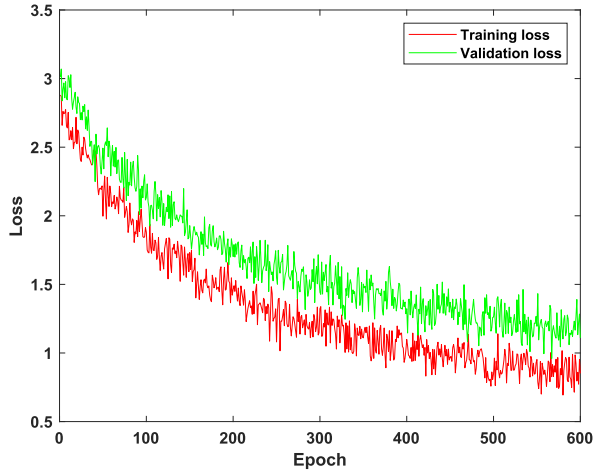


Fig. 2. Training and validation loss analysis of VGG16.

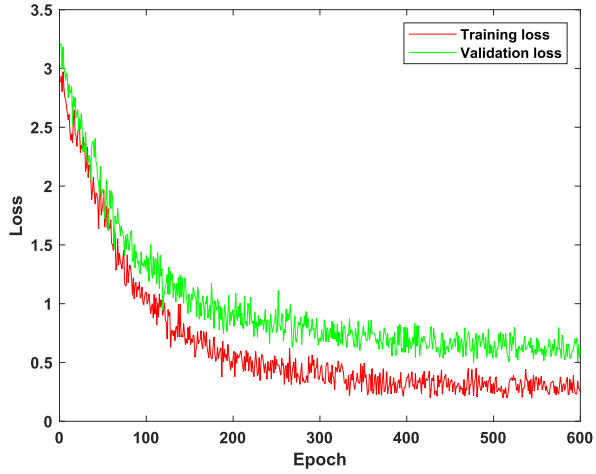


Fig. 3. Training and validation loss analysis of ResNet152V2.

A. Training and Validation Loss Analysis

This section discusses the training and validation analysis using the loss curves. Fig. 2 shows the training and validation loss analysis of VGG16 with respect to a number of epochs. It is found that the VGG16 achieves the best training and validation loss values as 0.7387 and 1.3481, respectively. It shows that there is an impact of over-fitting too.

Fig. 3 demonstrates the training and validation loss analysis of ResNet152V2. It is found that RESNET152 V achieves the best training and validation loss values of 0.6297 and 1.1726, respectively. It shows that there is an impact of over-fitting. But it shows better performance than VGG-16. Fig. 4 shows the training and validation loss analysis of DenseNet201. It is found that the best training and validation loss values are 0.3207 and 0.5218, respectively), thus DenseNet201 is least affected by the

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:55:33 UTC from IEEE Xplore. Restrictions apply.

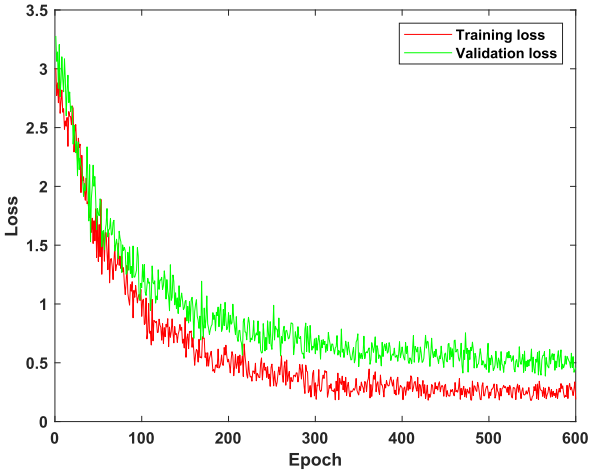


Fig. 4. Training and validation loss analysis of DenseNet201.

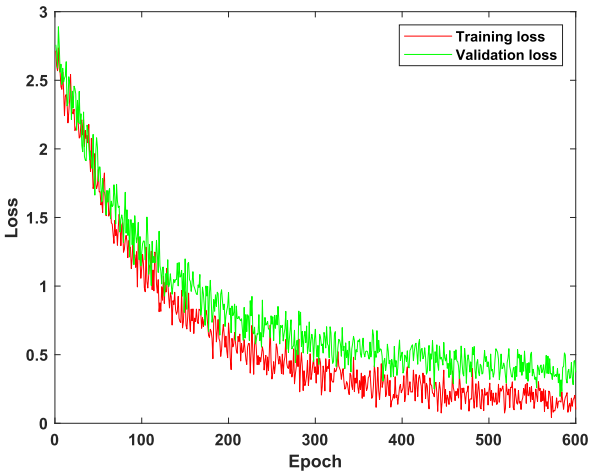


Fig. 5. Training and validation loss analysis of the proposed EDC-Net.

over-fitting problem. But still, there is room for improvement in the convergence curve. Fig. 5 demonstrates the training and validation loss analysis of EDC-Net. It is found that EDC-Net approaches toward minimum possible training loss. Also, there is a lesser difference between best training and validation loss values, (i.e., 0.2189 and 0.4368, respectively), thus EDC-Net is least affected by the over-fitting problem.

B. Visual Analysis

Fig. 6 shows the correctly obtained captions. It is found that the predicted captions are identical to the actual captions. So, for such captions, we received a maximum BLUE score. Since we have considered it as a classification problem, therefore, it represents the truly predicted class too. It clearly shows that EDC-Net can effectively provide captions for medical images.

Fig. 7 demonstrates the incorrectly obtained captions. It is found that the predicted captions are far away from the actual captions. So, for such captions, we received a minimum BLUE score. Since we have considered it as a classification problem, therefore, it represents the falsely predicted class too. It clearly shows that in certain cases when the visibility of the images is poor then EDC-Net fails to provide correct captions.

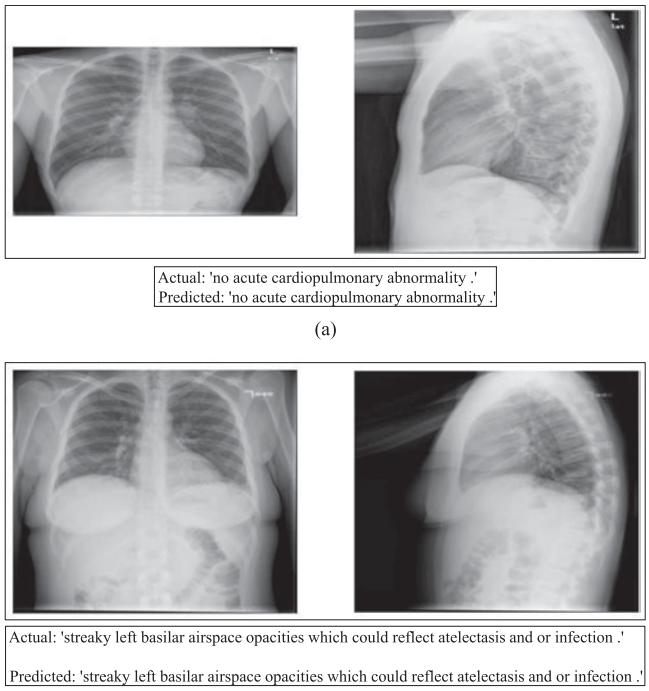


Fig. 6. Correctly classified captions.

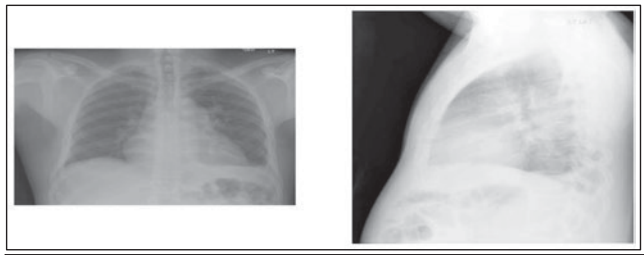
TABLE I
BLUE SCORE AND KAPPA STATISTICS (KS) ANALYSIS OF THE EDC-NET

Model	BLUE-1	BLUE-2	BLUE-3	BLUE-4	KS
RMN	0.563	0.342	0.229	0.125	0.919
VGG16	0.568	0.364	0.249	0.126	0.923
CNN-RNN	0.511	0.351	0.253	0.123	0.912
DMSFF	0.539	0.353	0.244	0.126	0.924
DGGAN	0.563	0.365	0.238	0.127	0.932
ResNet152V2	0.571	0.309	0.256	0.119	0.896
DenseNet201	0.574	0.327	0.227	0.128	0.938
MVE_1	0.573	0.363	0.252	0.127	0.934
Min_1	0.571	0.361	0.250	0.124	0.915
Max_1	0.569	0.359	0.249	0.121	0.908
EDC-Net	**0.577**	**0.367**	**0.258**	**0.129**	**0.952**

C. Quantitative Analysis

This section discusses the performance analyses using the confusion matrix. Fig. 8 shows the confusion matrix obtained from the VGG16 on the medical image captioning dataset. It is found that the VGG16 achieves 93.5% accuracy. Fig. 9 shows the confusion matrix obtained from the ResNet152V2 on the medical image captioning dataset. It is found that the ResNet152V2 achieves 94.8% accuracy. Fig. 10 shows the confusion matrix obtained from the DenseNet201 on the medical image captioning dataset. It is found that the DenseNet201 achieves 95.4% accuracy. Fig. 11 shows the confusion matrix obtained from EDC-Net on the medical image captioning dataset. It is found

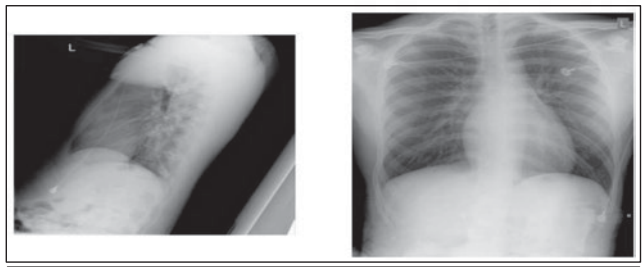
Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:55:33 UTC from IEEE Xplore. Restrictions apply.



Actual: 'suspected left hilar mass with obstructive atelectasis . thorax for further characterization.'

Predicted: 'streaky left basilar airspace opacities which could reflect atelectasis and or infection .'

(a)



Actual: 'negative for acute abnormality . discrete projectile not seen .'

Predicted: 'streaky left basilar airspace opacities which could reflect atelectasis and or infection .'

(b)

Fig. 7. Incorrectly classified captions.

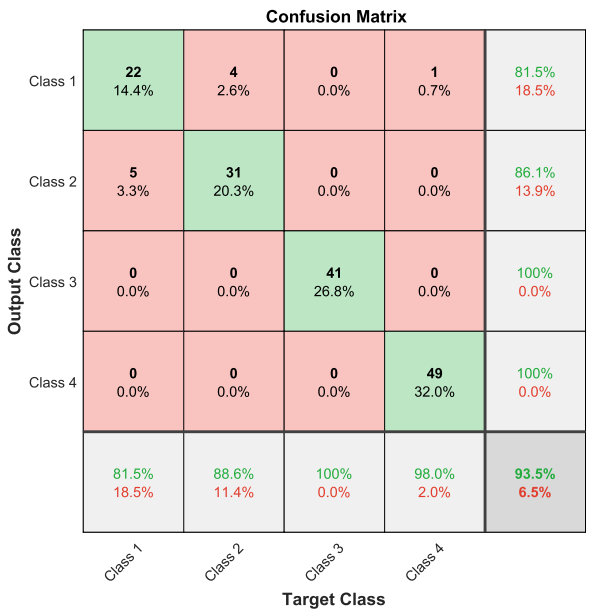


Fig. 8. Confusion matrix analysis of VGG16 on medical image captioning dataset.

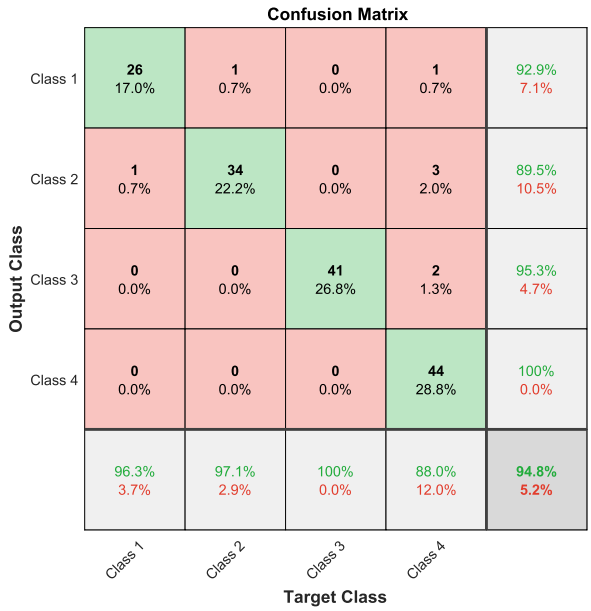


Fig. 9. Confusion matrix analysis of ResNet152V2 on medical image captioning dataset.

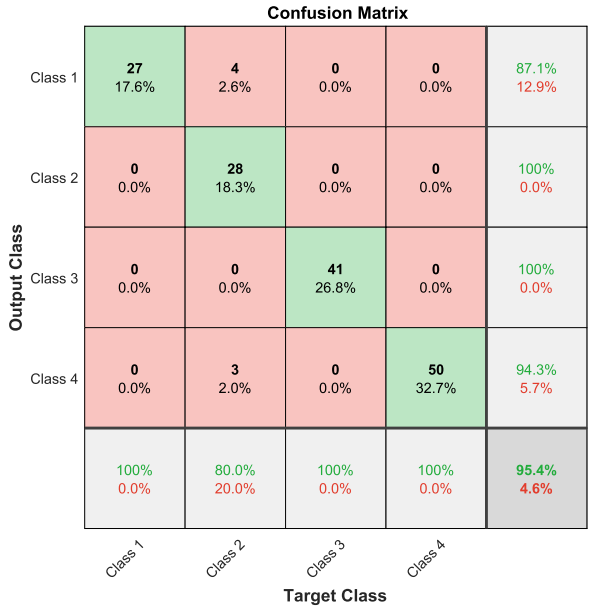


Fig. 10. Confusion matrix analysis of DenseNet201 on medical image captioning dataset.

that EDC-Net achieves 97.4% accuracy. Thus, EDC-Net outperforms the competitive models with an average improvement of 2.0%.

D. Discussion

Table I shows the BLUE score [44] analysis of EDC-Net. Comparisons are drawn between the proposed and the competitive models such as RMN [13], VGG16 [33], CNN-RNN [7], DMSFF 9057472, DGGAN [19], ResNet152V2 [34], and DenseNet201 [35] using BLUE scores [44] and kappa statistics (KS) [48], [49]. It is found that DenseNet201 [35] achieved the highest BLUE-1 as 0.574 as compared to the other competitive models. DGGAN [19] and VGG16 [33] models have achieved highest BLUE-2 values as 0.365 and 0.364, respectively. ResNet152V2 [34] achieved highest BLUE-3 value as 0.56 than the existing models. DenseNet201 [35]. Ensemble-based models, i.e., MVE_1, Min_4, and Max_5 achieved better results than most of the existing models. Out of these models,

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:55:33 UTC from IEEE Xplore. Restrictions apply.

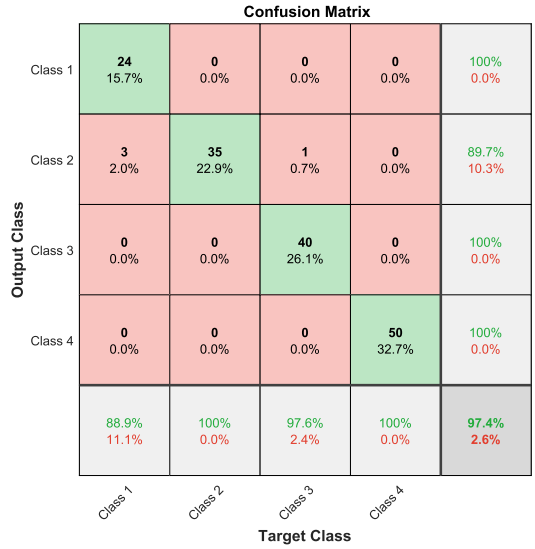


Fig. 11. Confusion matrix analysis of EDC-Net on medical image captioning dataset.

ANOVA analysis of BLUE-1					
Source	*SS*	*df*	*MS*	*F*	*Prob>F*
Columns	2.267	10	0.2267	37.7833	2.016e-04
Error	0.294	49	0.0060		
Total	2.561	59			

Fig. 12. ANOVA table of BLUE-1.

ANOVA analysis of BLUE-2					
Source	*SS*	*df*	*MS*	*F*	*Prob>F*
Columns	2.836	10	0.2836	32.9767	0.0156
Error	0.423	49	0.0086		
Total	3.259	59			

Fig. 13. ANOVA table of BLUE-2.

ANOVA analysis of BLUE-3					
Source	*SS*	*df*	*MS*	*F*	*Prob>F*
Columns	0.53724	10	0.3301	47.8405	0.0023
Error	0.44408	20	0.0069		
Total	0.98132	24			

Fig. 14. ANOVA table of BLUE-3.

MVE_1 outperformed the Min_4 and Max_5. In terms of KS, DGGAN achieved remarkable performance over the existing models. From Table I, it is found that EDC-Net outperforms the existing models by achieving higher values of BLUE scores. Bold values indicate higher performance. It is found that the proposed EDC-Net outperforms the existing models in terms

ANOVA analysis of BLUE-4					
Source	*SS*	*df*	*MS*	*F*	*Prob>F*
Columns	3.198	10	0.3198	53.3000	0.0003
Error	0.297	49	0.0060		
Total	3.495	59			

Fig. 15. ANOVA table of BLUE-4.

ANOVA analysis of Kappa Statistics					
Source	*SS*	*df*	*MS*	*F*	*Prob>F*
Columns	1.748	10	0.1748	22.4102	1.435e-03
Error	0.385	49	0.0078		
Total	2.133	59			

Fig. 16. ANOVA table of kappa statistics (KS).

TABLE II
COMPARATIVE ANALYSIS OF EDC-NET WITH THE STATE-OF-THE-ART MODELS

Model	BLEU-1	BLEU-2	BLEU-3	BLEU-4
mDiAP-LSTM [27]	0.373	0.226	0.147	0.101
Full-ARL [29]	-	-	-	0.125
R-paraNet (VGG-19)[30]	0.505	0.329	0.230	0.168
R-paraNet (DneseNet-19)[30]	0.503	0.333	0.236	0.175
HRGR-Agent [31]	0.438	0.298	0.208	0.151
HLSTM [32]	0.455	0.288	0.205	0.154
KERP [28]	0.216	0.214	0.087	0.066
EDC-Net	**0.516**	**0.348**	**0.238**	**0.178**

of BLUE-1, BLUE-2, BLUE-3, BLUE-4, and KS by 1.258%, 1.185%, 1.289%, 1.098%, and 1.548%, respectively. Also, additional ANOVA analyses are performed on each performance metric which has shown that the EDC-Net significantly outperforms the competitive models.

For statistical analysis, ANOVA is used. Each performance metric has the following hypotheses:

$$\begin{cases} H_0 & \mu M_1 = \mu M_2 = \cdots = \mu M_{11}, \\ H_A & \text{Means are not equal.} \end{cases} \tag{12}$$

Here, Null and alternate hypotheses are defined by H_0 and H_A, respectively. μM_i represents EDC-Net and the existing image captioning models. M_{11} depicts EDC-Net. The ANOVA table contains degrees of freedom, a sum of squares, the mean sum of squares, F-statistics (F_V), and P-value (p). If p value of F_V falls below the significance level, we will reject H_0 and state that there is a significant difference between the models. For all the metrics analysis presented in Table I, H_A is accepted (see Figs. 12–16). It shows that there is a significant difference in performance between different models.

Table II shows the comparative analysis between the EDC-Net and state-of-the-art models on IU X-RAY [50]. It is found that EDC-Net outperforms the state-of-the-art models in terms of BLUE-1, BLUE-2, BLUE-3, and BLUE-4 by 1.357%, 1.249%, 1.115%, and 1.031%, respectively.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:55:33 UTC from IEEE Xplore. Restrictions apply.

V. Conclusion

To provide explanatory information to the doctors and patients, an efficient deep ensemble medical image captioning network (DCNet) was proposed. DCNet has ensembled three well-known pre-trained models such as VGG16, ResNet152V2, and DenseNet201. Ensembling of these models has shown better results by preventing over-fitting. DCNet can efficiently extract the potential features of biomedical images, but it is sensitive to its control parameters. Therefore, to tune the control parameters, evolving DCNet (EDC-Net) was proposed. Evolution was achieved using the self-adaptive parameter control-based differential evolution. Comparative analysis has shown that EDC-Net achieves higher performance than the existing models to provide explanatory information for biomedical images. Comparative analysis has shown that EDC-Net achieves 97.4% accuracy. According to the results obtained using Open-i dataset, the proposed EDC-Net outperformed the existing models in terms of BLUE-1, BLUE-2, BLUE-3, BLUE-4, and KS by 1.258%, 1.185%, 1.289%, 1.098%, and 1.548%, respectively. Additionally, the EDC-Net model outperformed the state-of-the-art models on IU X-RAY dataset in terms of BLUE-1, BLUE-2, BLUE-3, and BLUE-4 by 1.357%, 1.249%, 1.115%, and 1.031%, respectively.

In the near future, to handle the poor visibility, noise, poor registration, etc. kind of problems with medical images, we will integrate various preprocessing techniques such as image filtering, image registration, visibility restoration, etc. with the proposed EDC-Net. Since preprocessing techniques may reduce the performance of EDC-Net, therefore, a novel selective preprocessing model will be designed which will decide whether preprocessing is required or not. If preprocessing is required then which operation(s) should be applied on the given image(s).

Data Availability Statement: The data collected during the data collection phase are available from the corresponding authors upon request.

Conflicts of Interest: The authors would like to confirm there are no conflicts of interest regarding the study.

References

[1] M. Zhang, Y. Yang, H. Zhang, Y. Ji, H. T. Shen, and T.-S. Chua, "More is better: Precise and detailed image captioning using online positive recall and missing concepts mining," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 32–44, Jan. 2019.

[2] X. Zeng, L. Wen, Y. Xu, and C. Ji, "Generating diagnostic report for medical image by high-middle-level visual information incorporation on double deep learning models," *Comput. Methods Programs Biomed.*, vol. 197, 2020, Art. no. 105700.

[3] Z. Zhang, W. Zhang, W. Diao, M. Yan, X. Gao, and X. Sun, "VAA: Visual aligning attention model for remote sensing image captioning," *IEEE Access*, vol. 7, pp. 137355–137364, 2019.

[4] R. Sharma, A. Kumar, D. Meena, and S. Pushp, "Employing differentiable neural computers for image captioning and neural machine translation," in *Proc. Int. Conf. Smart Sustain. Intell. Comput. Appl. Procedia Comput. Sci.*, 2020, vol. 173, pp. 234–244.

[5] X. Zeng, L. Wen, B. Liu, and X. Qi, "Deep learning for ultrasound image caption generation based on object detection," *Neurocomputing*, vol. 392, pp. 132–141, 2020.

[6] J. Yang, Y. Sun, J. Liang, B. Ren, and S.-H. Lai, "Image captioning by incorporating affective concepts learned from both visual and textual components," *Neurocomputing*, vol. 328, pp. 56–68, 2019.

[7] Z. Babar, T. van Laarhoven, F. M. Zanzotto, and E. Marchiori, "Evaluating diagnostic content of AI-generated radiology reports of chest x-rays," *Artif. Intell. Med.*, vol. 116, 2021, Art. no. 102075.

[8] D. Zhao, Z. Chang, and S. Guo, "A multimodal fusion approach for image captioning," *Neurocomputing*, vol. 329, pp. 476–485, 2019.

[9] N. Zakharov, H. Su, J. Zhu, and J. Gläscher, "Towards controllable image descriptions with semi-supervised vae," *J. Vis. Commun. Image Representation*, vol. 63, 2019, Art. no. 102574.

[10] S. Yang, J. Niu, J. Wu, Y. Wang, X. Liu, and Q. Li, "Automatic ultrasound image report generation with adaptive multimodal attention mechanism," *Neurocomputing*, vol. 427, pp. 40–49, 2021.

[11] A. E. Kavur, L. I. Kuncheva, and M. A. Selver, "Basic ensembles of vanilla-style deep learning models improve liver segmentation from ct images," 2020, *arXiv:2001.09647*.

[12] N. Yu, X. Hu, B. Song, J. Yang, and J. Zhang, "Topic-oriented image captioning based on order-embedding," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 2743–2754, Jun. 2019.

[13] B. Wang, X. Zheng, B. Qu, and X. Lu, "Retrieval topic recurrent memory network for remote sensing image captioning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 256–270, 2020.

[14] L. Zhou, Y. Zhang, Y.-G. Jiang, T. Zhang, and W. Fan, "Re-caption: Saliency-enhanced image captioning through two-phase learning," *IEEE Trans. Image Process.*, vol. 29, pp. 694–709, 2020.

[15] W. Zhao, X. Wu, and J. Luo, "Cross-domain image captioning via cross-modal retrieval and model adaptation," *IEEE Trans. Image Process.*, vol. 30, pp. 1180–1192, 2021.

[16] M. Yang et al., "Multitask learning for cross-domain image captioning," *IEEE Trans. Multimedia*, vol. 21, no. 4, pp. 1047–1061, Apr. 2019.

[17] G. Hoxha, F. Melgani, and B. Demir, "Toward remote sensing image retrieval under a deep image captioning perspective," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 4462–4475, 2020.

[18] Y. Huang, J. Chen, W. Ouyang, W. Wan, and Y. Xue, "Image captioning with end-to-end attribute detection and subsequent attributes prediction," *IEEE Trans. Image Process.*, vol. 29, pp. 4013–4026, 2020.

[19] M. Yang et al., "An ensemble of generation- and retrieval-based image captioning with dual generator generative adversarial network," *IEEE Trans. Image Process.*, vol. 29, pp. 9627–9640, 2020.

[20] Z. Yuan, X. Li, and Q. Wang, "Exploring multi-level attention and semantic relationship for remote sensing image captioning," *IEEE Access*, vol. 8, pp. 2608–2620, 2020.

[21] W. Huang, Q. Wang, and X. Li, "Denoising-based multiscale feature fusion for remote sensing image captioning," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 3, pp. 436–440, Mar. 2021.

[22] F. Monay and D. Gatica-Perez, "Modeling semantic aspects for cross-media image indexing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 10, pp. 1802–1817, Oct. 2007.

[23] J. Yu, J. Li, Z. Yu, and Q. Huang, "Multimodal transformer with multi-view visual representation for image captioning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 12, pp. 4467–4480, Dec. 2020.

[24] Y. Xian and Y. Tian, "Self-guiding multimodal LSTM—when we do not have a perfect training dataset for image captioning," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5241–5252, Nov. 2019.

[25] S. Amirian, K. Rasheed, T. R. Taha, and H. R. Arabnia, "Automatic image and video caption generation with deep learning: A concise review and algorithmic overlap," *IEEE Access*, vol. 8, pp. 218386–218400, 2020.

[26] X. Li et al., "COCO-CN for cross-lingual image tagging, captioning, and retrieval," *IEEE Trans. Multimedia*, vol. 21, no. 9, pp. 2347–2360, Sep. 2019.

[27] H. Park, K. Kim, S. Park, and J.-W. Choi, "Medical image captioning model to convey more details: Methodological comparison of feature difference generation," *IEEE Access*, vol. 9, pp. 150560–150568, 2021.

[28] C. Y. Li, X. Liang, Z. Hu, and E. P. Xing, "Knowledge-driven encode, retrieve, paraphrase for medical image report generation," in *Proc. AAAI Conf. Artif. Intell.*, 2019, vol. 33, pp. 6666–6673.

[29] D. Hou, Z. Zhao, Y. Liu, F. Chang, and S. Hu, "Automatic report generation for chest x-ray images via adversarial reinforcement learning," *IEEE Access*, vol. 9, pp. 21236–21250, 2021.

[30] F. Wang, X. Liang, L. Xu, and L. Lin, "Unifying relational sentence generation and retrieval for medical image report composi-tion," *IEEE Trans. Cybern.*, vol. 52, no. 6, pp. 5015–5025, Jun. 2022.

[31] Y. Li, X. Liang, Z. Hu, and E. P. Xing, "Hybrid retrieval-generation reinforced agent for medical image report generation," *Adv. Neural Inf. Process. Syst.*, vol. 31, 2018.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:55:33 UTC from IEEE Xplore. Restrictions apply.

[32] B. Jing, P. Xie, and E. Xing, "On the automatic generation of medical imaging reports," in *Proc. Annu. Meeting Assoc. Comput. Linguistics*, 2018, pp. 2577–2586.
[33] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.
[35] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4700–4708.
[36] K. Siwek, S. Osowski, and R. Szupiluk, "Ensemble neural network approach for accurate load forecasting in a power system," *Int. J. Appl. Math. Comput. Sci.*, vol. 19, no. 2, pp. 303–315, 2009.
[37] J. Islam and Y. Zhang, "Brain MRI analysis for alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks," *Brain Informat.*, vol. 5, no. 2, 2018, Art. no. 2.
[38] H. Alshazly, C. Linse, E. Barth, and T. Martinetz, "Ensembles of deep learning models and transfer learning for ear recognition," *Sensors*, vol. 19, no. 19, 2019, Art. no. 4139.
[39] D. Singh, V. Kumar, M. Kaur, M. Y. Jabarulla, and H.-N. Lee, "Screening of covid-19 suspected subjects using multi-crossover genetic algorithm based dense convolutional neural network," *IEEE Access*, vol. 9, pp. 142566–142580, 2021.
[40] D. Singh, V. Kumar, V. Yadav, and M. Kaur, "Deep neural network-based screening model for covid-19-infected patients using chest x-ray images," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 35, no. 03, 2021, Art. no. 2151004.
[41] D. Singh, V. Kumar, and M. Kaur, "Densely connected convolutional networks-based COVID-19 screening model," *Appl. Intell.*, vol. 51, no. 5, pp. 3044–3051, 2021.
[42] K. Cho, B. Van Merriënboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," 2014, *arXiv:1409.1259*.
[43] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, *arXiv:1409.0473*.
[44] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: A method for automatic evaluation of machine translation," in *Proc. 40th Annu. Meeting Assoc. Comput. Linguistics*, 2002, pp. 311–318.
[45] L. Cui, G. Li, Z. Zhu, Z. Wen, N. Lu, and J. Lu, "A novel differential evolution algorithm with a self-adaptation parameter control method by differential evolution," *Soft Comput.*, vol. 22, no. 18, pp. 6171–6190, 2018.
[46] U. L. of Medicine, "Open-i service of the national library of medicine," 2022. [Online]. Available: https://openi.nlm.nih.gov
[47] A. E. Kavur et al., "Chaos challenge - combined (CT-MR) healthy abdominal organ segmentation," *Med. Image Anal.*, vol. 69, 2021, Art. no. 101950.
[48] T. Toprak, B. Belenlioglu, B. Aydın, C. Guzelis, and M. A. Selver, "Conditional weighted ensemble of transferred models for camera based onboard pedestrian detection in railway driver support systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5041–5054, May 2020.
[49] M. A. Selver, "A robotic system for warped stitching based compressive strength prediction of marbles," *IEEE Trans. Ind. Informat.*, vol. 16, no. 11, pp. 6796–6805, Nov. 2020.
[50] H. Park, K. Kim, J. Yoon, S. Park, and J. Choi, "Feature difference makes sense: A medical image captioning model exploiting feature difference and tag information," in *Proc. 58th Annu. Meeting Assoc. Comput. Linguistics: Student Res. Workshop*, 2020, pp. 95–102.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 08,2024 at 08:55:33 UTC from IEEE Xplore. Restrictions apply.

IEEE Journal of Translational Engineering in Health and Medicine

Received 3 March 2022; revised 4 April 2022; accepted 22 April 2022.
Date of publication 28 April 2022; date of current version 12 May 2022.

Digital Object Identifier 10.1109/JTEHM.2022.3171078

Artificial Intelligence-Based Cyber-Physical System for Severity Classification of Chikungunya Disease

DILBAG SINGH[1], (Senior Member, IEEE), MANJIT KAUR[1], (Member, IEEE), VIJAY KUMAR[2], MOHAMED YASEEN JABARULLA[1], (Member, IEEE), AND HEUNG-NO LEE[1], (Senior Member, IEEE)

[1]School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea
[2]Department of Computer Science and Engineering, NIT Hamirpur, Hamirpur 177005, India

CORRESPONDING AUTHOR: HEUNG-NO LEE (heungno@gist.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MSIP) under Grant NRF-2021R1A2B5B03002118; and in part by the Ministry of Science and ICT (MSIT), South Korea, through the Information Technology Research Center (ITRC) support program supervised by the Institute of Information and Communications Technology Planning and Evaluation (IITP) under Grant IITP-2021-0-01835.

ABSTRACT Background: Artificial intelligence techniques are widely used in solving medical problems. Recently, researchers have used various deep learning techniques for the severity classification of Chikungunya disease. But these techniques suffer from overfitting and hyper-parameters tuning problems. Methods: In this paper, an artificial intelligence-based cyber-physical system (CPS) is proposed for the severity classification of Chikungunya disease. In CPS system, the physical components are integrated with computational algorithms to provide better results. Random forest (RF) is used to design the severity classification model for Chikungunya disease. However, RF suffers from overfitting and poor computational speed problems due to complex architectures and large amounts of connection weights. Therefore, an evolving RF model is proposed using the adaptive crossover-based genetic algorithm (ACGA). Results: ACGA can efficiently optimize the architecture of RF to achieve better results with better computational speed. Extensive experiments are performed by utilizing the Chikungunya disease dataset. Conclusion: Performance analysis demonstrates that ACGA-RF achieves higher performance as compared to the competitive models in terms of F-measure, accuracy, sensitivity, and specificity. The proposed CPS system can prevent users from visiting hospitals and can render services to patients living far away from hospitals.

INDEX TERMS Artificial intelligence, cyber-physical system, automated diagnosis, Chikungunya disease, random forest, adaptive crossover, genetic algorithm, severity classification.

Clinical translation statement: The proposed model can be utilized for the severity classification of Chikungunya disease. The research findings are impactful as the proposed model can prevent users from visiting hospitals and can render services to patients living far away from hospitals.

I. INTRODUCTION

The Healthcare system plays a vital role in the development of any nation. Government is responsible to design proper healthcare policies to protect their citizens from any outbreak of diseases [1]. Hence, the outbreak of any new disease such as coronavirus is a major challenge for the healthcare system. Several viruses exist that can affect both animals and human beings. Chikungunya is one of the viruses that can be spread very rapidly and may create a big problem for the health system. Two basic types of infected mosquitoes i.e., Aedes albopictis and Aedes agypti transmit this virus in the human body [2]. The symptoms of Chikungunya are joint pain, sudden high fever, and rash. Some infected persons have headaches, fatigue, digestive complaints, and conjunctivitis [3]. The symptoms of Chikungunya are very similar to dengue because the same mosquito carries both viruses [4]. But in Chikungunya, joint pain is more severe as well as redness of the eyes. The symptoms of sore throat are different from dengue infection. Chikungunya may not cause death. As per literature, it is found that the patient

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

recovers within a week of this disease [5]. But, the joint pain may last for a few months. The doctors start the treatment by perceiving the symptoms of patients. However, the exact measurement of these symptoms is not possible. Therefore, the treatment of patients may not be effective.

Instead of symptoms found in the patients, reverse transcription-polymerase chain reaction (RT-PCR) and serological tests are used to diagnose Chikungunya. Both tests require blood samples of patients [6]. However, these tests are unable to provide reliable performance for this disease. Supervised learning techniques such as machine learning and deep learning can be used to evaluate the severity of this disease by considering the symptoms of patients and laboratory tests [7]. The severity classification of Chikungunya infected persons is still an ill-posed problem.

In [8], a fog-based framework for Chikungunya disease diagnosis was designed. J48 was utilized to classify Chikungunya infected patients. In [9], wearable internet of things (IoT) and fog-based framework for classification and controlling the Chikungunya disease was proposed. Fuzzy-C means (FCM) classifier was utilized for Chikungunya classification. But J48 [8] and FCM [9] suffer from over-fitting and hyper-parameters tuning problems. In [10], a particle swarm optimization-based ANFIS (PANFIS) model was implemented for the diagnosis of Chikungunya disease. Initially, an adaptive neuro-fuzzy inference system (ANFIS) classifier was used to classify the infected patients. Thereafter, particle swarm optimization (PSO) was utilized to overcome the parameter tuning problem with ANFIS. It achieved remarkable results compared to artificial neural networks (ANN). But, PANFIS [10], [11] suffers from the over-fitting problem. Also, sometimes PSO may be stuck in local optima and suffers from premature convergence problems [12], [13].

Therefore, to overcome the over-fitting and hyper-parameters tuning problems, an efficient evolving Random forest (RF) model is proposed for the severity classification of Chikungunya disease. The main contributions of this paper are as follows:

1) A cyber-physical system (CPS) based severity classification model is proposed for Chikungunya disease. In CPS system, the physical components are integrated with computational algorithms to attain better results.
2) Evolving RF model is proposed for severity classification of Chikungunya disease. An adaptive crossover-based genetic algorithm (ACGA) is utilized to evolve RF model.
3) Deep learning model is also implemented and compared with ACGA-RF for severity classification of Chikungunya disease.

The remainder of this paper is organized as follows. Section II presents the related work. The proposed ACGA-RF model is described in Section III. Performance analyses are presented in Section IV. The concluding remarks are discussed in Section V.

II. RELATED WORK

Artificial intelligence techniques are widely used in solving medical problems. Recently, researchers used machine learning techniques for the classification of Chikungunya disease. Hossain *et al.* [5] utilized the different symptoms of patients for the accurate assessment of Chikungunya disease. Their proposed framework collected the data from the interviews of patients. They used a belief-based rule system for predicting the level of Chikungunya. Their model attained an accuracy of 92%. Yang [14] developed a decision system for the diagnosis of Chikungunya disease. The neural network was used for classification by considering the uncertainty of the disease's symptoms. However, the uncertainty of some symptoms is not considered in this approach. Ganesan *et al.* [15] presented three different models to diagnose the Chikungunya disease. However, these models require human intervention for the assessment of this disease. Caicedo-Torres *et al.* [16] proposed a machine learning-based classifier for differentiating the dengue and Chikungunya patients. Their classifier was tested on 447 patients. The logistic regression model outperformed the other models. The accuracy obtained from logistic regression was 87%.

Ibrahim *et al.* [17] presented the backpropagation method for predicting the epidemic disease. They used epidemic disease factors for prediction. Thereafter, these factors were applied to the clustering technique. Their method is capable to identify the epidemic disease using feature classification. Coelho *et al.* [18] used a transfer learning model for predicting mosquito-borne diseases. They used time-series data from two Brazilian cities. Both the long short-term memory neural network model and random quantile forest model provided the same prediction performance. Caicedo-Torres *et al.* [19] utilized the machine learning techniques for envisaging the morbidity of Chikungunya in Colombia. Kernel ridge regression was used for forecasting the Chikungunya cases. Cross-validation and mean absolute error were used.

Sippy *et al.* [20] developed two prediction models on the Machala dataset. The first model namely the severity index for suspected arbovirus model (SISA) that utilized demographic data. Another model namely the severity index for suspected arbovirus (SISA1) with laboratory utilized the laboratory data. The accuracies obtained from SISA and SISAl were 91% and 95%, respectively. Both models are capable to envisage arbovirus hospitalization.

Shimpi *et al.* [21] used a backpropagation algorithm to predict the Chikungunya disease. Five gradient-based optimization techniques were used. The pre-processed features were applied to the backpropagation algorithm for classification. The accuracy obtained from this model was 95%. However, a small dataset was used for validation purposes. Eng *et al.* [22] used machine learning techniques for predicting the binding affinity of T-cell epitopes of Chikungunya. They built prediction models for identifying binders and non-binder. This model will be helpful for vaccine development.

IEEE Journal of Translational Engineering in Health and Medicine

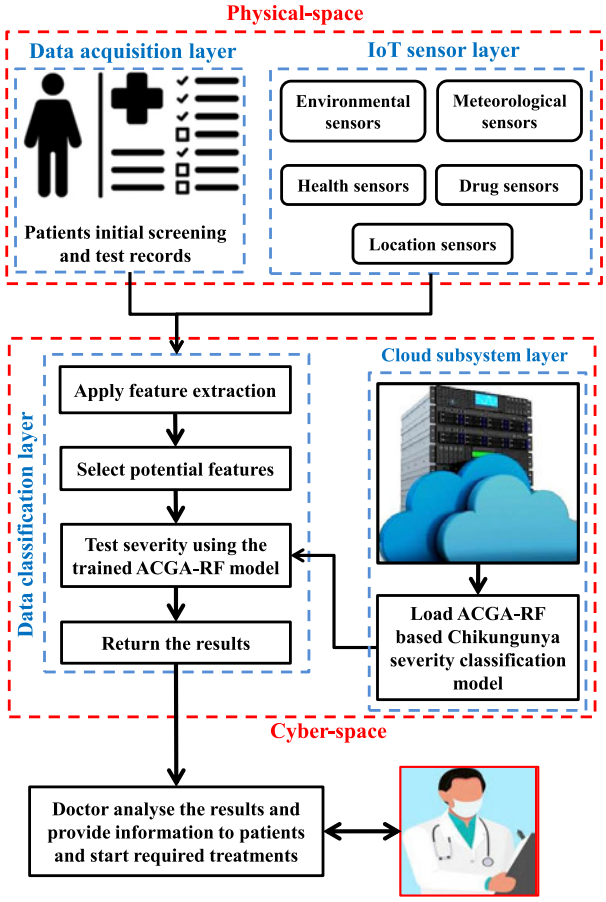


FIGURE 1. Proposed ACGA-RF model based Chikungunya severity classification model.

From the existing literature, it is found that the majority of the existing models suffer from hyper-parameters tuning, (i.e., optimization of initial control parameters), over-fitting and data insensitivity problems [23], [24]. Hence, there is a need to develop an efficient model for severity classification for Chikungunya disease.

III. PROPOSED FRAMEWORK

Motivated from [8], [10], an efficient model is proposed in this paper for diagnosis of Chikungunya disease. Figure 1 shows the proposed artificial intelligence-based cyber-physical system (CPS) for the diagnosis of Chikungunya disease. There are two main components in the proposed model, i.e., physical space and cyberspace. In physical space. the data related to users' health is collected and forwarded to cyberspace for predicting the severity of the Chikungunya virus [25], [26]. In cyberspace, there are two sub-layers, i.e., the cloud subsystem layer and data classification layer [27], [28]. In the cloud subsystem layer, data and the proposed trained model are stored. At the data classification layer, the severity prediction of the Chikungunya virus is done in real-time by applying the proposed ACGA-RF model after extracting the potential features. Finally, doctors are involved for further assessment of the results.

A. PHYSICAL SPACE

In physical space, there is a data acquisition layer that collects users' personal data, symptoms related to Chikungunya, and records of initial screening with the help of sensors.

The Chikungunya dataset is obtained from [8] and [9]. Sood and Mahajan [8] and [9] built the dataset by taking the symptoms-based dataset from [29] that contains eleven health features, i.e., abdominal pain, muscle pain, bleeding disorder, fatigue, eyes pain, itching, nausea, sore throat, joint pain, skin rash, and fever, of 2367 patients. Dataset with 5032 cases comprising environmental variables was taken from [30] and [31]. Monthly climate variables, i.e., rainfall, temperature, and humidity, were taken from [32]–[34]. The symptoms-based data was combined with climate and environmental features to validate ACGA-RF (for further details refer to [8] and [9]).

TABLE 1. Data attributes related to Chikungunya virus.

Attribute type	Attribute Name	Sensors used
User Personal data	UID (unique identification), Name, Age, Gender, Location of user workplace, Location of user residence, and Mobile number	GPS sensors
Health-related data	Joint pain, Muscle Pain, Rashes on the skin, Appetite loss, Fever, Redness in the eyes, Headache, nausea, sore throat, fatigue, vomiting	Biosensors
Environmental and location data	Water quality, air temperature, humidity, mosquito density, rainfall, mosquito breeding sites, level of carbon dioxide	Mosquito and climatic sensors

Table 1 shows the common features of Chikungunya virus. It contains several sensing devices to observe the health of patients. The observed information is then transferred to the data classification layer to classify the health of a particular client. The description of the dataset is mentioned in Table 1 (refer [8], [10]).

1) User Personal Data: This comprises the user's personal data. GPS sensors are used to collect data such as UID, name, age, address, and mobile number.
2) Health-related Data: This set consists of vital signs. It provides information related to rashes on the skin, muscle pain, appetite loss, joint pain, fever, redness in the eyes, headache, nausea, sore throat, fatigue, and vomiting. Biosensors are used to collect this information.
3) Environmental and location Information: It provides the position of patients, susceptible users, uninfected users, and positions of mosquito dense areas, water quality, humidity, and mosquito breeding sites. These positions are evaluated by Global Positioning System (GPS) sensors to obtain the travel history of every patient. Radio Frequency Identification (RFID) tags and mosquito sensors are also utilized to store the proximity interactions between infected/uninfected/ susceptible users, mosquito densities, and breeding sites.

Table 2 depicts the personal-parameters of registered users. It presents a brief description of the user's personal parameters such as name, age, gender, address, mobile number, and contact details of guardians (refer [8], [10]). Table 3 shows

TABLE 2. Chikungunya's user personal-parameters.

Sr. No.	Parameters	Explanation
1.	Name	Patient name
2.	Age	Age of user in years
3.	Gender	Male or Female
4.	GPS	Geographic location of user
5.	Address	Address of the registered patient
6.	Telephone No.	Mobile number
7.	Patient Guardian's	Contact details of guardians

TABLE 3. Chikungunya symptoms related-dataset.

Sr. No.	Attributes	Decision	Description
1	Headache	Binary	Whether a user has a headache
2	Exposure to risky area	Binary	Whether a user works or lives in risk area
3	Nausea	Binary	User is feeling tired and sleepy
4	Fatigue	Binary	User feels weak and loss weight
5	Rash and Vomiting	Binary	Itching problem and has body marks
6	Muscular Pain	Binary	User suffers from muscular pain
7	Joints Swelling	Binary	User suffers from swelling
8	Joint pain	Binary	User suffers from joint pain symptom
9	Fever	Integer	Temperature in °C
10	Initiation post infection	Yes	3 to 7 days

Chikungunya symptoms such as headache, exposure to the risky area, nausea, fever, rash, vomiting, etc.

B. CYBER SPACE

Cyberspace is comprised of two sub-layers, i.e., the cloud subsystem layer and the data classification layer. The detail of registered users and the proposed trained model is stored in cloud subsystem layers. Whereas, at the data classification layer, diagnosis of the Chikungunya virus is done by applying ACGA-RF to the potential features. Also, RF-based model is used as it has the abilities of generalization, learning, fault tolerance, and adoption. However, RF model is sensitive to its control parameters. An efficient tuning of these control parameters can improve the performance of RF. Therefore, to automatically optimize control parameters of RF, an ACGA-RF model is proposed. Finally, the predicted severity report of a patient is transmitted to doctors for further treatment.

IV. PROPOSED EVOLVING RANDOM FOREST MODEL

This section discusses the proposed evolving random forest model. Initially, RF model is briefly discussed to understand its basic notations and hyper-parameters. Random forest (RF) is an ensemble of classification/regression trees [35], where every tree shows a mapping from feature space to the response. Trees can be obtained either using a subsampled data set of actual data or bootstrapped. Every tree is conditionally independent of one another. However, RF model is sensitive to its architecture and hyper-parameters. Thus, an adaptive crossover genetic algorithm (ACGA) is utilized to obtain the optimized architecture and hyper-parameters of RF.

In literature, genetic algorithm is widely accepted to optimize various classifiers such as deep learning models [23], [36]. Genetic algorithms utilize crossover and mutation operators during evolution phases to obtain the final solution. It has been found that the selection of efficient crossover and mutation operators is a challenging problem. To efficiently select crossover operator(s), Xue *et al.* [37] proposed an adaptive crossover genetic algorithm (ACGA). A group of crossover operators was utilized during the evolution process. Based on the performance of crossover operators, roulette wheel selection was utilized to select a specific crossover. In this paper, ACGA [37] is utilized to form evolving RF model. Three different crossover operators are used. The working of the proposed ACGA-RF is depicted in Algorithm 1.

init_P(): Population P is created using normal distribution by creating M vectors. Each vector represents the architecture and hyper-parameters-related values of RF.

init_A(): Adaptive crossover selector (A_S) is utilized by assigning probabilities to each crossover as $1/C$. C shows a number of utilized crossovers.

init_R(): Given crossover is selected according to roulette wheel selection and probabilities obtained from A_S.

init_C(): Apply selected crossover on parents' offspring to form child offsprings.

init_M(): Mutation operator is utilized to obtain child offspring. Compute child offsprings are saved in the offspring population(P_δ).

init_CA(): Dominated offsprings are then evaluated by using the actual and child offsprings. The respective outcomes are saved in rW and pL. P_δ is computed during $M/2^{th}$ step.

init_S(): Crowded distance [38] and Non-dominated sorting [39] are used to obtain M solutions from R ($P \cup P_\delta$).

init_D(): To allocate reward/ penalty to selected offsprings, dominance comparison is utilized.

Penalty and reward of offsprings is saved in $nP_{IT\times C}$ and $nR_{IT\times C}$, respectively. After IT number of phases, A_S is updated by considering $nP_{IT\times C}$ and $nR_{IT\times C}$. All steps are repeated until the termination criterion (i.e., FE) is satisfied.

The succeeding subsections discuss the steps of ACGA.

A. FITNESS FUNCTION

Fitness function is designed to optimize RF by using sensitivity and specificity. It is defined as:

$$\max F(X) = \{f_1(X) \text{ and } f_2(X)\} \tag{1}$$

Here, X is an offspring. f_1 and f_2 represent the sensitivity and specificity parameters, respectively.

B. CROSSOVER OPERATORS

Three different crossover operators, i.e., single-point [38], chaotic crossover [40], and reduced surrogate [41], [42] are utilized. Single-point [38] has significant results to solve many computationally hard problems. It has shown better

IEEE Journal of Translational Engineering in Health and Medicine

Algorithm 1 ACGA Based RF

Input: Max population (M), A_S, iterative threshold (IT), No. of crossovers (C), and No. of fitness evaluations (FE)
Output: Optimized values for RF

P ← init_P(M)
Initialize rW, pL, $nR_{IT\times C}$, and $nP_{IT\times C}$.
$\widehat{P} = p_1, p_2, \ldots, p_C \leftarrow init\_A(C)$
$nFE \leftarrow 0$ and $k \leftarrow 0$
$P_\delta \leftarrow \phi$
while $nFE < FE$ **do**
  for $i = 1$ to $M/2$ **do**
    $a_v \leftarrow$ init_R $(\widehat{P})$
    Two offsprings are selected as parents: O_p
    $O_c \leftarrow$ init_C(O_p, a_v)
    $O_c \leftarrow$ init_M(O_c)
    $nFE \leftarrow nFE + 2$
    $[rW, pL] \leftarrow$ init_CA(O_p, O_c)
    Add O_c to P_δ
  end for
  k ← k + 1
  Update rW to k^{th} row of $nR_{IT\times C}$
  Update pL to k^{th} row of $nP_{IT\times C}$
  if k = IT **then**
    $\widehat{P} \leftarrow$ init_A($nR_{IT\times C}$, $nP_{IT\times C}$)
    $k = 0$
  end if
  $R \leftarrow P \cup P_\delta$
  $P \leftarrow$ init_S(R)
  Assign non-dominated offsprings in P to T_P
  Optimized values ← T_P
end while
return Optimized values

computational speed compared to the existing crossover operators [37].

Reduced surrogate [41], [42] can avoid unnecessary crossover operations when parents have similar offspring. Initially, it evaluates parents and forms a group of crossover points where both parents have different genes. In the absence of such a crossover point, no crossover operator is implemented. Chaotic crossover [40] can obtain a better converged and distributed group of Pareto-optimal offspring.

C. PENALTY AND REWARD

Penalties and rewards are allocated to offsprings by utilizing two matrices namely pL and rW as:

$$rW = [0 \ldots 0]_{1\times C} \tag{2}$$

$$pL = [0 \ldots 0]_{1\times C} \tag{3}$$

Pareto optima among the offsprings is used to modify rW and pL.

Algorithm 2 Credit Allocation (init_CA ())

Input: Parents (ρ), Children (δ), crossover selected using init_A (q)
Output: rW, pL $[n_d, d_s]$ ← init_D(ρ) // d_s and n_d define dominated and non-dominated offsprings, respectively.

Dominated parent (assume $\rho_1 \prec \rho_2$).
if $d_s \neq \phi$ **then**
  for i = 1 to 2 **do**
    if $\rho_1 \prec \delta_i$ **then**
      $pL_q \leftarrow pL_q + 1$
    else
      $rW_q \leftarrow rWrW_q + 1$
    end if
  end for
else
  // If parent is non-dominated.
  for i = 1 to 2 **do**
    if $\rho_1 \not\prec \delta_i$ $\rho_2 \not\prec \delta_i$ **then**
      $rW_q \leftarrow rWrW_q + 1$
    else
      $pL_q \leftarrow pL_q + 1$
    end if
  end for
end if
P ← init_P (M)
return Optimized values for RF

1) PARENT IS NON-DOMINATED

Pareto optima is evaluated between child and respective parent offsprings. If child offsprings are dominated by parents, then append pL_q by 1, otherwise append rW_q by 1. The pseudocode of updation of penalty and reward is depicted in Algorithm 2.

2) DOMINATED PARENT

If parent 1 (ρ_1) is dominated by parent 2 (ρ_2), then pareto optima of each child offspring is compared with ρ_2. If child offspring is dominated by ρ_2, then append pL_q by 1. Otherwise, update rW_q by 1.

D. UPDATION OF ADAPTIVE CROSSOVER SELECTOR

A_S is utilized to update the crossover selection probabilities. During evolution process, it is implemented after every IT steps (refer [43]). Two matrices, i.e., $nP_{IT\times C}$ and $nR_{IT\times C}$ are used to hold the values of pL and rW, respectively. Recently updated IT's pL and n_R values are utilized to modify A_S. To evaluate the probability for q^{th} $(q = 1, 2, \ldots, C)$ crossover, addition of qth column is utilized (refer [43]).

V. PERFORMANCE ANALYSIS

To analyze the efficiency of proposed model, health-related attributes are collected from [8], [34], [44], [45]. It mainly consists of attributes such as age, sex, location, fever, skin

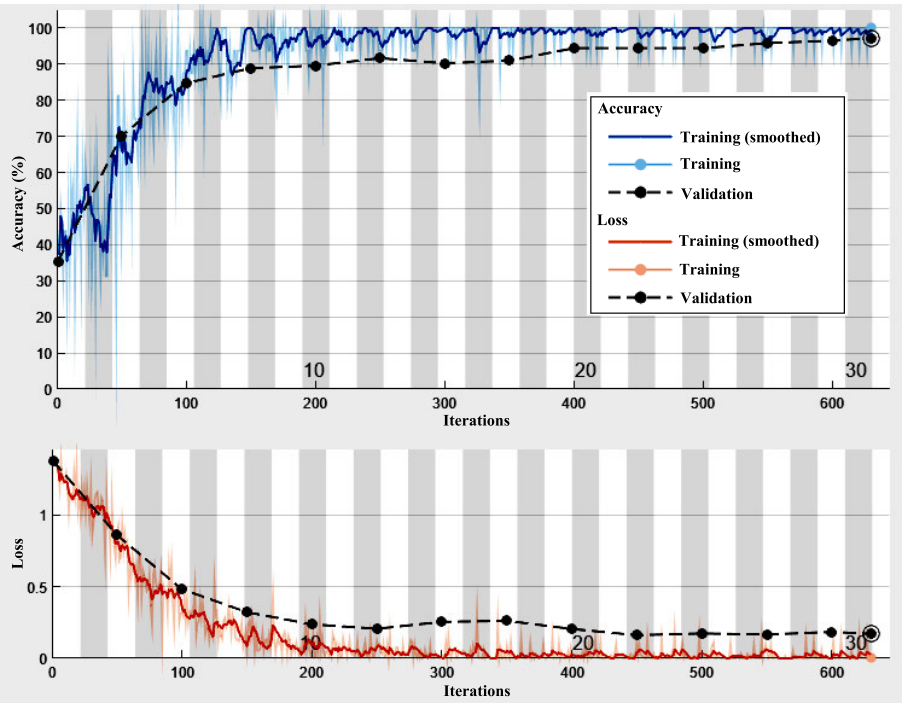


FIGURE 2. Accuracy and loss analysis of the deep learning-based severity classification of Chikungunya disease.

rashes, and joint pain. It is collected for approximately 10000 users. Around 1805 users are at no risk while 3890 are at normal risk and 4305 are at high risk. The acquired data is stored in cloud storage and is used for recognition by an optimized RF model. ACGA-RF is compared with J48, SVM, ANN, RF, adaptive neuro-fuzzy inference system (ANFIS), PANFIS [10], and deep learning (DL).

To implement DL model, various layers (i.e., feature input layer, fully connected layer, batch normalization layer [46], [47], ReLU layer [46], [47], softmax layer [46], [47], and classification layer) are utilized. For normalizing the input data, Z-score normalization [48] is used. Minibatch size is set to be 8. Adam [49] a stochastic gradient descent optimizer is used to achieve the better convergence of DL model.

A. TRAINING ANALYSIS

Figure 2 shows the accuracy and loss analysis of the deep learning-based severity classification model for Chikungunya disease. It clearly shows that the deep learning achieves significantly better convergence speed. It achieves 98.3% validation accuracy, therefore, least affected by the impact of over-fitting as training accuracy is 100%.

The performance of RF achieves the best training and validation accuracy values are 99.5% and 97.6%, respectively. Although it shows remarkable results, the performance of the proposed approach is still far from the optimal results. It shows an over-fitting problem as there is a high difference between training and validation accuracy values. The performance of ACGA-RF model achieves the best training and validation accuracy values are 100.0% and 99.6%, respectively. Thus, ACGA-RF model achieves remarkable results than both deep learning and RF models. Also, ACGA-RF is least affected by the over-fitting issue as there is only a 0.4% difference between the training and validation accuracy which are 1.7 and 1.9 for deep learning and RF, respectively. Also, ACGA-RF shows an enhancement in validation accuracy over deep learning and RF as 1.3% and 2.0%, respectively.

TABLE 4. Confusion matrix analysis among the proposed ACGA-RF, DL, GA-RF, PSO-RF, and RF models.

Model	RF	DL	GA-RF	PSO-RF	ACGA-RF
TP	203	206	207	207	209
TN	314	315	315	317	318
FP	7	4	3	3	1
FN	6	5	5	3	2
Accuracy	0.9754	0.9831	0.9837	0.9855	0.9943
Sensitivity	0.9712	0.9763	0.9732	0.9855	0.9905
Specificity	0.9781	0.9874	0.9690	0.9902	0.9968
F-measure	0.9747	0.9818	0.9829	0.9902	0.9936

B. CONFUSION MATRIX ANALYSIS

To evaluate the performance of ACGA-RF, confusion matrix analyses are also achieved. Table 4 depicts the confusion matrix analysis proposed ACGA-RF, RF, and deep learning (DL) models. It is found that RF achieves the testing accuracy of 97.54%. It is found that the deep learning (DL) model achieves the testing accuracy of 98.31%. DL shows an average improvement of 0.77% over RF model. It is found

IEEE Journal of Translational Engineering in Health and Medicine

that ACGA-RF achieves an accuracy of 99.43%. Thus, the proposed ACGA-RF achieved an average improvement of 1.89% and 1.12% over RF and deep learning-based severity classification models, respectively.

C. COMPARATIVE ANALYSIS

Boxplot and ANOVA are used for statistical analysis. The hypotheses for every performance measure can be defined as:

$$\begin{cases} H_0 & \mu M_1 = \mu M_2 = \ldots = \mu M_7, \\ H_A & \text{Means are not equal.} \end{cases} \quad (4)$$

where μM_i shows various severity classification models for Chikungunya disease. M_7 shows the proposed ACGA-RF. H_0 and H_A define the null and alternate hypotheses, respectively. ANOVA table consists of various attributes such as a sum of squares (SS), degrees of freedom (df), the mean sum of squares (MS), F-statistics (F), and P-value (P). If P value of F is lesser than the level of significance, then we can reject H_0) and conclude that the models are significantly different from each other.

ANOVA analysis of Testing Accuracy

Source	*SS*	*Df*	*MS*	*F*	*Prob>F*
Columns	0.00177	6	0.0003	43.25	$1.639e^{-20}$
Error	0.00043	63	0.00001		
Total	0.00212	69			

FIGURE 3. ANOVA analysis of severity classification models for Chikungunya disease in terms of testing accuracy.

ANOVA analysis of Testing Sensitivity

Source	*SS*	*df*	*MS*	*F*	*Prob>F*
Columns	0.00207	6	0.00034	55.29	$3.057e^{-23}$
Error	0.00039	63	0.00001		
Total	0.00246	69			

FIGURE 4. ANOVA analysis of severity classification models for Chikungunya disease in terms of testing sensitivity.

ANOVA analysis of Testing Specificity

Source	*SS*	*Df*	*MS*	*F*	*Prob>F*
Columns	0.00247	6	0.00041	43.55	$1.376e^{-20}$
Error	0.00059	63	0.00001		
Total	0.00306	69			

FIGURE 5. ANOVA analysis of severity classification models for Chikungunya disease in terms of testing specificity.

Figures 3 - 6 show that H_A is accepted for all the considered performance metrics as the evaluated $p - values$ are lower than 0.01. Thus, the performance of different models is significantly different from each other. But, it is not possible

ANOVA analysis of Testing F-measure

Source	*SS*	*df*	*MS*	*F*	*Prob>F*
Columns	0.00221	6	0.00037	49.16	$2.057e^{-21}$
Error	0.00043	63	0.00001		
Total	0.00272	69			

FIGURE 6. ANOVA analysis of severity classification models for Chikungunya disease in terms of testing F-measure.

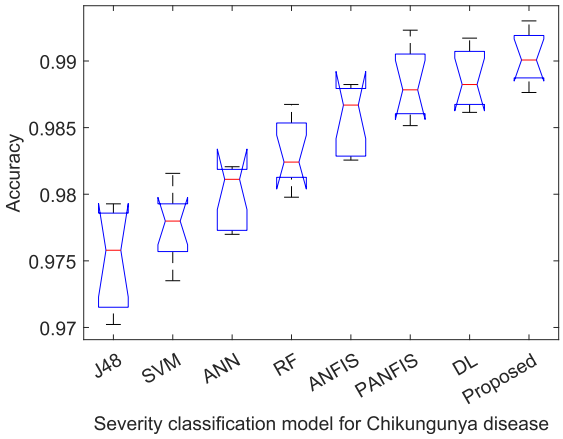


FIGURE 7. Accuracy analysis of ACGA-RF model.

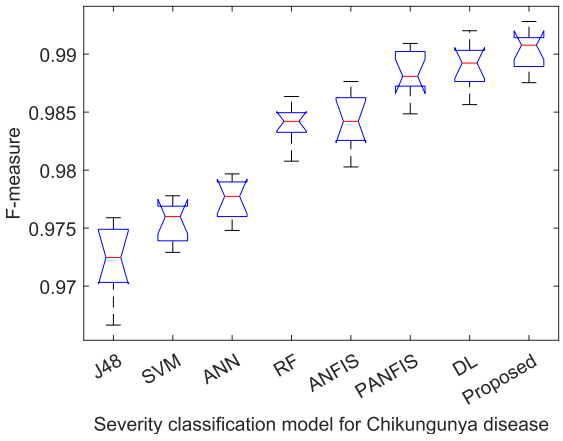


FIGURE 8. F-measure analysis of ACGA-RF model.

to find which technique outperforms the others. Thus, the boxplots are obtained to evaluate which technique performs significantly better than the others (see Figures 7 to 10).

Figure 7 demonstrates the accuracy analysis among ACGA-RF and the existing recognition Chikungunya disease recognition models. It is clearly shown that ACGA-RF achieves remarkably significant and consistent accuracy values. Compared to the existing Chikungunya disease recognition models, ACGA-RF achieves 1.3822% improvement in terms of accuracy.

Figure 8 shows F-measure analysis among ACGA-RF and the existing Chikungunya disease recognition models. ACGA-RF outperforms the existing Chikungunya disease recognition models by an average improvement of 1.4145%.

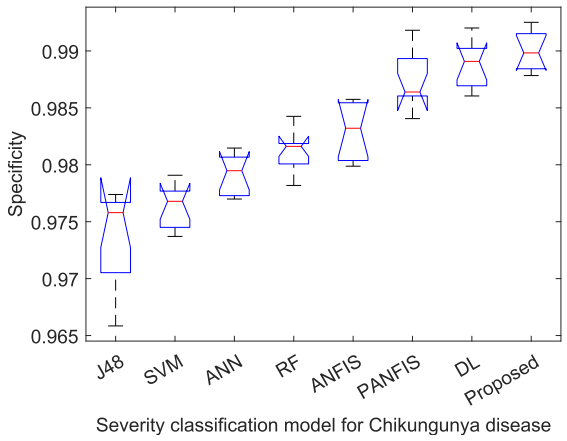


FIGURE 9. Specificity analysis of ACGA-RF model.

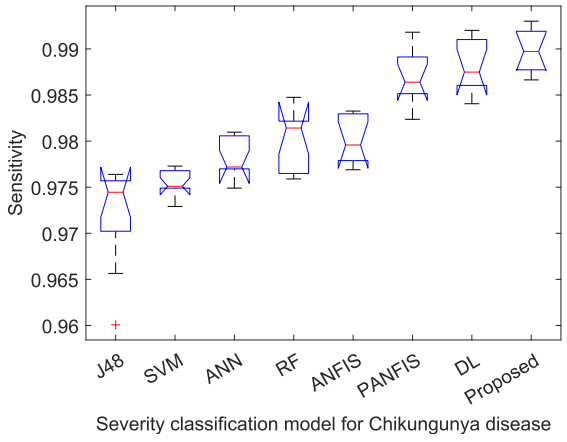


FIGURE 10. Sensitivity analysis of ACGA-RF model.

Figure 9 shows the specificity analysis among the existing models and the proposed Chikungunya disease recognition model. ACGA-RF achieves a 99.21% (median) value, which is significantly better than the competitive models by an average improvement of 1.3972%. Therefore, ACGA-RF can efficiently diagnose normal risk Chikungunya disease-infected patients. Figure 10 shows the sensitivity analysis among the existing models and proposed Chikungunya disease recognition model. ACGA-RF achieves 99.24% (median) value, which is significantly better than the competitive models by 1.4172%. Therefore, ACGA-RF can efficiently diagnose the high risk Chikungunya disease patients.

It is found that ACGA-RF takes on an average of 1 hour 36 minutes during the training process. During the testing process, it takes only 2.746 minutes to achieve the results. Additionally, ACGA-RF provides the testing results on an average of 1.674 seconds for a single patient. Therefore, ACGA-RF can be used for real-time applications.

D. DISCUSSION

Table 5 shows the comparison among the proposed model with state-of-the-art machine learning and deep learning models. It is found that in [8], J48 was utilized for severity classification of Chikungunya disease. It achieved the sensitivity, specificity, F-measure, and accuracy values as 0.935, 0.965, 0.827, and 92.7865, respectively. Fuzzy-C means (FCM) [9] achieved the sensitivity, specificity, and accuracy values as 0.867, 0.888, and 0.934, respectively. PANFIS [10] achieved sensitivity, specificity, F-measure, and accuracy values as 0.9578, 0.9787, 0.9431, and 0.9871, respectively. DL based Chikungunya diagnosis model achieved sensitivity, specificity, F-measure, and accuracy values as 0.9831, 0.9763, 0.9874, and 0.9818, respectively. Compared to these models, ACGA-RF achieved a sensitivity, specificity, F-measure, and accuracy values of 0.9943, 0.9905, 0.9968, and 0.9936, respectively. Therefore, the proposed ACGA-RF achieved significantly better results than the existing models in terms of accuracy, specificity, sensitivity, and F-measure by 1.3822%, 1.3972%, 1.4172%, and 1.4145%, respectively.

TABLE 5. Comparative analysis among the proposed ACGA-RF, DL, and state-of-the art models.

Model	Accuracy	Sensitivity	Specificity	F-measure
J48 [8]	0.9278	0.935	0.883	0.827
FCM [9]	0.934	0.904	0.912	–
ANN [10]	0.9674	0.9438	0.9617	0.9328
PANFIS [10]	0.9871	0.9578	0.9787	0.9431
DL	0.9831	0.9763	0.9874	0.9818
Proposed ACGA-RF	0.9943	0.9905	0.9968	0.9936

VI. CONCLUSION

In this paper, a cloud-based CPS is designed and implemented for the recognition of Chikungunya disease. The proposed system is divided into two main categories, i.e., physical space and cyberspace. Once, the data related to user-health are collected, it is stored in the cloud sub-system layer. An evolving RF model was proposed for the severity classification of Chikungunya disease by using ACGA. ACGA can efficiently optimize RF architecture to achieve better results with better computational speed. The comparative analysis demonstrates that ACGA-RF achieves significantly better testing performance than the existing models in terms of accuracy, specificity, sensitivity, and F-measure by 1.3822%, 1.3972%, 1.4172%, and 1.4145%, respectively. Thus, the proposed Chikungunya disease recognition model is beneficial for real-time medical applications.

In near future, the deep transfer learning models can be used to obtain more efficient results. Further novel meta-heuristic techniques can be designed to efficiently tune the deep learning architectures. Also, the proposed model can be applied to other kinds of datasets.

CONFLICT OF INTEREST

The authors declare that no conflict of interest.

REFERENCES

[1] S. C. Weaver and M. Lecuit, "Chikungunya virus and the global spread of a mosquito-borne disease," *New England J. Med.*, vol. 372, no. 13, pp. 1231–1239, Mar. 2015.

[2] Y. Lin, X. Lu, F. Fang, and J. Fan, "Personal health care monitoring and emergency response mechanisms," in *Proc. 1st Int. Symp. Future Inf. Commun. Technol. Ubiquitous HealthCare (Ubi-HealthTech)*, Jul. 2013, pp. 1–5.

[3] B. Wahid, A. Ali, S. Rafique, and M. Idrees, "Global expansion of chikungunya virus: Mapping the 64-year history," *Int. J. Infectious Diseases*, vol. 58, pp. 69–76, May 2017.

[4] M. Kaur *et al.*, "Coinfection of chikungunya and dengue viruses: A serological study from north western region of Punjab, India," *J. Lab. Physicians*, vol. 10, no. 4, pp. 443–447, Oct. 2018.

[5] M. S. Hossain, Z. Sultana, L. Nahar, and K. Andersson, "An intelligent system to diagnose chikungunya under uncertainty," *J. Wireless Mobile Netw., Ubiquitous Comput., Dependable Appl.*, vol. 10, no. 2, pp. 37–54, 2019.

[6] V. Lakshmi *et al.*, "Clinical features and molecular diagnosis of chikungunya fever from south India," *Clin. Infectious Diseases*, vol. 46, no. 9, pp. 1436–1442, May 2008.

[7] S. K. Mardekian and A. L. Roberts, "Diagnostic options and challenges for dengue and chikungunya viruses," *BioMed Res. Int.*, vol. 2015, pp. 1–8, Oct. 2015.

[8] S. K. Sood and I. Mahajan, "A fog-based healthcare framework for chikungunya," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 794–801, Oct. 2018.

[9] S. K. Sood and I. Mahajan, "Wearable IoT sensor based healthcare system for identifying and controlling chikungunya virus," *Comput. Ind.*, vol. 91, pp. 33–44, Oct. 2017.

[10] S. Kaur and K. K. Chahal, "Prediction of chikungunya disease using PSO-based adaptive neuro-fuzzy inference system model," *Int. J. Comput. Appl.*, pp. 1–9, Jan. 2021.

[11] A. Rezaeipanah and M. Mojarad, "Modeling the scheduling problem in cellular manufacturing systems using genetic algorithm as an efficient meta-heuristic approach," *J. Artif. Intell. Technol.*, vol. 1, pp. 228–234, Oct. 2021.

[12] G. D. Singh, M. Prateek, S. Kumar, M. Verma, D. Singh, and H.-N. Lee, "Hybrid genetic firefly algorithm-based routing protocol for VANETs," *IEEE Access*, vol. 10, pp. 9142–9151, 2022.

[13] S. Ndichu, S. Kim, and S. Ozawa, "Deobfuscation, unpacking, and decoding of obfuscated malicious Javascript for machine learning models detection performance improvement," *CAAI Trans. Intell. Technol.*, vol. 5, no. 3, pp. 184–192, Sep. 2020.

[14] J. B. Yang, "Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties," *Eur. J. Oper. Res.*, vol. 131, no. 1, pp. 31–61, 2001.

[15] V. Ganesan, B. Duan, and S. Reid, "Chikungunya virus: Pathophysiology, mechanism, and modeling," *Viruses*, vol. 9, no. 12, p. 368, Dec. 2017.

[16] W. Caicedo-Torres, A. Paternina-Caicedo, H. Pinzón-Redondo, and J. Gutiérrez, "Differential diagnosis of dengue and chikungunya in Colombian children using machine learning," in *Proc. Ibero-Amer. Conf. Artif. Intell.*, Cham, Switzerland: Springer, 2018, pp. 181–192.

[17] N. Ibrahim, N. S. M. Akhir, and F. H. Hassan, "Predictive analysis effectiveness in determining the epidemic disease infected area," in *Proc. AIP Conf.*, 2017, Art. no. 020064.

[18] F. C. Coelho, N. L. D. Holanda, and B. M. Coimbra, "Transfer learning applied to the forecast of mosquito-borne diseases," *MedRxiv*, to be published, doi: 10.1101/2020.02.03.20020164.

[19] W. Caicedo-Torres, D. Montes-Grajales, W. Miranda-Castro, M. Fennix-Agudelo, and N. Agudelo-Herrera, "Kernel-based machine learning models for the prediction of dengue and chikungunya morbidity in Colombia," in *Proc. Colombian Conf. Comput.*, Cham, Switzerland: Springer, 2017, pp. 472–484.

[20] R. Sippy *et al.*, "Severity index for suspected arbovirus (SISA): Machine learning for accurate prediction of hospitalization in subjects suspected of arboviral infection," *PLoS Neglected Tropical Diseases*, vol. 14, no. 2, 2020, Art. no. e0007969.

[21] P. Shimpi, S. Shah, M. Shroff, and A. Godbole, "An artificial neural network approach for classification of vector-borne diseases," in *Proc. Int. Conf. Comput. Methodolog. Commun. (ICCMC)*, Jul. 2017, pp. 412–415.

[22] C. L. P. Eng, T. W. Tan, and J. C. Tong, "T-cell epitope prediction of chikungunya virus," in *Chikungunya Virus*. Cham, Switzerland: Springer, 2016, pp. 201–207.

[23] D. Singh, V. Kumar, M. Kaur, M. Y. Jabarulla, and H.-N. Lee, "Screening of COVID-19 suspected subjects using multi-crossover genetic algorithm based dense convolutional neural network," *IEEE Access*, vol. 9, pp. 142566–142580, 2021.

[24] X. Zhang and G. Wang, "Stud pose detection based on photometric stereo and lightweight yolov4," *J. Artif. Intell. Technol.*, vol. 2, no. 1, pp. 32–37, 2022.

[25] Y. Zhang, M. Qiu, C.-W. Tsai, M. M. Hassan, and A. Alamri, "Health-CPS: Healthcare cyber-physical system assisted by cloud and big data," *IEEE Syst. J.*, vol. 11, no. 1, pp. 88–95, Mar. 2017.

[26] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE ICNN*, vol. 4, Nov./Dec. 1995, pp. 1942–1948.

[27] M. Janalipour and A. Mohammadzadeh, "Building damage detection using object-based image analysis and ANFIS from high-resolution image (case study: BAM earthquake, Iran)," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 5, pp. 1937–1945, May 2016.

[28] S. A. Hosseini and I. Esmaili Paeen Afrakoti, "Evaluation of a new neutron energy spectrum unfolding code based on an adaptive neuro-fuzzy inference system (anfis)," *J. Radiat. Res.*, vol. 59, no. 4, pp. 436–441, 2017.

[29] *Data Catalog. Now Cast Predictions for Local Transmission of Chikungunya Virus*. Accessed: Oct. 8, 2017. [Online]. Available: https://catalog.data.gov/dataset?tags=chikungunya

[30] *DHS Program Demographic and Health Surveys*. Survey Dataset Files. Accessed: Oct. 8, 2017. [Online]. Available: https://dhsprogram.com/data/dataset/Ethiopia_StandardDHS_2016.cfm

[31] *DHS Program Demographic and Health Surveys*. Survey Dataset Files. Accessed: Oct. 8, 2017. [Online]. Available: https://dhsprogram.com/data/dataset/Angola_StandardDHS_2015.cfm

[32] *Data Catalog*. Surface Air Temperature. Accessed: Oct. 8, 2017. [Online]. Available: https://data.gov.sg/dataset/surface-air-temperaturemonthly-mean

[33] *Un Data: A World of Information*. Relative Humidity. Accessed: Oct. 8, 2017. [Online]. Available: http://data.un.org/Data.aspx?d=CLINO&f=ElementCode%3a11

[34] *Data Catalog*. Monthly Total Rainfall. Accessed: Oct. 8, 2017. [Online]. Available: https://data.gov.sg/dataset/rainfall-monthly-total

[35] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[36] I. M. El-Hasnony, S. I. Barakat, and R. R. Mostafa, "Optimized ANFIS model using hybrid metaheuristic algorithms for Parkinson's disease prediction in IoT environment," *IEEE Access*, vol. 8, pp. 119252–119270, 2020.

[37] Y. Xue, H. Zhu, J. Liang, and A. Słowik, "Adaptive crossover operator based multi-objective binary genetic algorithm for feature selection in classification," *Knowl.-Based Syst.*, vol. 227, Sep. 2021, Art. no. 107218.

[38] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Aug. 2002.

[39] M. Premkumar, P. Jangir, R. Sowmya, H. H. Alhelou, A. A. Heidari, and H. Chen, "MOSMA: Multi-objective slime mould algorithm based on elitist non-dominated sorting," *IEEE Access*, vol. 9, pp. 3229–3248, 2020.

[40] L. dos Santos Coelho and P. Alotto, "Multiobjective electromagnetic optimization based on a nondominated sorting genetic approach with a chaotic crossover operator," *IEEE Trans. Magn.*, vol. 44, no. 6, pp. 1078–1081, Jun. 2008.

[41] D. Whitley, "An executable model of a simple genetic algorithm," in *Foundations of Genetic Algorithms*, vol. 2. Amsterdam, The Netherlands: Elsevier, 1993, pp. 45–62.

[42] H. Elmaraghy, V. Patel, and I. B. Abdallah, "Scheduling of manufacturing systems under dual-resource constraints using genetic algorithms," *J. Manuf. Syst.*, vol. 19, no. 3, pp. 186–201, 2000.

[43] Y. Xue, B. Xue, and M. Zhang, "Self-adaptive particle swarm optimization for large-scale feature selection in classification," *ACM Trans. Knowl. Discovery From Data*, vol. 13, no. 5, pp. 1–27, Oct. 2019.

[44] (Jun. 8, 2018). *Centers for Disease Control and Prevention*. Chikungunya Diseases. [Online]. Available: https://www.cdc.gov/chikungunya/geo/default.html

[45] (May 16, 2018). *Pan American Health Organization. Chikungunya Disease*. [Online]. Available: https://www.paho.org/data/index.php/en/mnu-topics/chikv-en.html

[46] R. Vecchi, S. Scardapane, D. Comminiello, and A. Uncini, "Compressing deep-quaternion neural networks with targeted regularisation," *CAAI Trans. Intell. Technol.*, vol. 5, no. 3, pp. 172–176, Sep. 2020.

[47] Y. Yang, L. Juntao, and P. Lingling, "Multi-robot path planning based on a deep reinforcement learning DQN algorithm," *CAAI Trans. Intell. Technol.*, vol. 5, no. 3, pp. 177–183, 2020.

[48] S. G. K. Patro and K. K. Sahu, "Normalization: A preprocessing stage," 2015, *arXiv:1503.06462*.

[49] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 20 June 2022, accepted 4 July 2022, date of publication 15 July 2022, date of current version 25 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3191429

RESEARCH ARTICLE

DKNet: Deep Kuzushiji Characters Recognition Network

DILBAG SINGH[1], (Senior Member, IEEE), C. V. ARAVINDA[2], MANJIT KAUR[1], (Senior Member, IEEE), MENG LIN[3], JYOTHI SHETTY[2], VIKRAM RAJU REDDICHERLA[2], AND HEUNG-NO LEE[1], (Senior Member, IEEE)

[1]School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea
[2]N.M.A.M. Institute of Technology, Nitte, Karkala 574110, India
[3]Department of Electronic and Computer Engineering, The Graduate School of Science and Engineering, Ritsumeikan University, Kusatsu Shiga 525-8577, Japan

Corresponding author: Heung-No Lee (heungno@gist.ac.kr)

This work was supported in part by the Institute of Information and Communications Technology Planning and Evaluation (IITP) Grant by the Korean Government through MSIT (Artificial Intelligence Graduate School Program—GIST), under Grant 2019-0-01842; and in part by the Ministry of Science and ICT (MSIT), South Korea, under the Information Technology Research Center (ITRC) Support Program supervised by the Institute of Information and Communications Technology Planning and Evaluation (IITP) under Grant IITP-2021-0-01835.

ABSTRACT Kuzushiji, a cursive writing style, had been extensively utilized in Japan for over a thousand years starting from the 8th century. In 1900, Kuzushiji was not included in regular school curricula due to the change in the Japanese writing system. Nowadays Japanese natives are unable to read historical books that were written using Kuzushiji language. Therefore, libraries and museums have decided to build digital copies of the documents and books that were written in Kuzushiji language. Due to a limited number of trained experts, researchers have utilized machine and deep learning models to convert historical documents and books into a modern script that can be easily read by human beings. However, the existing deep learning techniques suffer from over-fitting and gradient vanishing problems. To overcome these problems, an efficient deep Kuzushiji characters recognition network (DKNet) is proposed. Initially, to remove noise from the training images, a trilateral joint filter is applied. Contrast limited adaptive histogram equalization (CLAHE) is then applied to enhance the visibility of filtered images. Thereafter, a pre-trained MobileNet is utilized to extract the features of Kuzushiji characters. MobileNet's final layers are removed, including the fully connected layer and softmax. The flatten layer is then applied to the input. A fully connected classification layer is then used with Rectified linear units (ReLUs) and dropouts. Dropouts are used to generalize the model, thus preventing the over-fitting problem. Finally, the softmax activation function is employed to provide the recognition results. To test the proposed model, actual documents are first segmented by using the proposed Maximally stable extremal regions (MSERs) and convexhull-based segmentation approach. Segmented characters are then recognized using the trained DKNet. Extensive comparative analyses reveal that DKNet achieves better performance than the competitive models in terms of various performance metrics. An efficient Application Programming Interface (API) is also designed for Japanese Kuzushiji ancient heritage character recognition to help the end-users.

INDEX TERMS Maximally stable extremal regions, deep learning models, Kuzushiji character recognition, MobileNet, InceptionNet, ResNet.

I. INTRODUCTION

Kuzushiji, a cursive writing style, had been extensively utilized in Japan for over a thousand years starting from the 8th century. But, in 1900, Kuzushiji was not included in regular school curricula due to the change in Japanese writing system. Nowadays Japanese natives are unable to read historical books that were written using Kuzushiji language. Therefore, libraries and museums have decided to build digital copies of the documents and books that were written

The associate editor coordinating the review of this manuscript and approving it for publication was Donato Impedovo.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

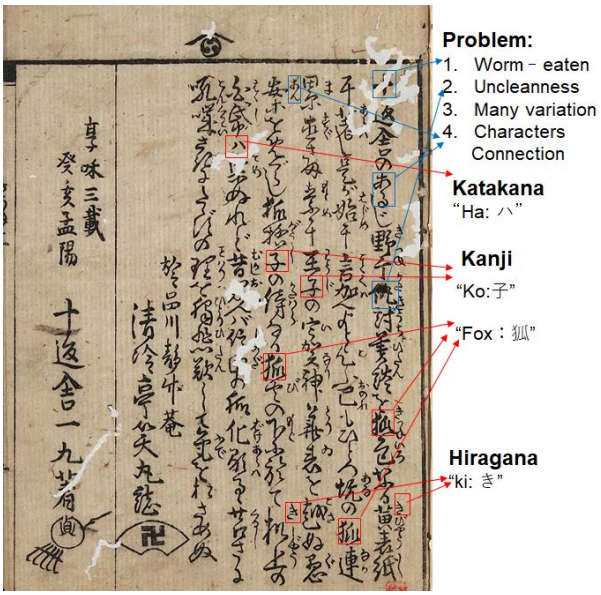


FIGURE 1. Typical page of a classical Japanese book.

in Kuzushiji language. But, due to a limited number of trained experts, researchers have utilized machine and deep learning models to convert historical documents and books into a modern script that can be easily read by human beings [1], [2].

To convert a single document or book page at a time, it is required to efficiently segment each character at a time [3]–[5]. In some cases, text detection becomes even more important than just text extraction until it can detect the shapes of regular or irregular shapes in the given images. Adding two Chinese syllabaries and a set of Chinese logograms into a complex logosyllabary system is the Japanese writing pattern, which is a fascinating study of tradition and innovation. The Japanese language uses incurved verbs and postpositions. It utilizes suffixes and particles to create words and clauses in sentences. Chinese characters (kanji) have been extensively utilized in a long history of use dating back centuries. The Japanese also used some Chinese characters with sound values for grammatical purposes. These characters were not instantly interpreted as phonetic or logographic representations of meaning [6], [7].

There were many variants of Kuzushiji characters and in some documents and books, these characters were connected with each other. Aging causes the documents to become unclean and some have been infected with worms [8]. Figure 1 shows a page of an early Japanese book that contains Kuzushiji characters.

- **Hiragana:** Manyoshu describes a system of writing native words in ancient Japan that uses manyogana. In Figure 1, Ki illustrates how these signs were reduced and simplified into sogana and finally to hiragana.
- **Katakana:** According to the Chinese Buddhist scriptures, the word kanji is an essential pronunciation aid. This character accumulated as grammatical suffixes, particles, and postpositions. But the literal meaning of kanji did not change. In Figure 1, Ha indicates that katakana is widely used to write non-Chinese words.
- **Allophones:** In Figure 1, Ko and Fox illustrate what may appear to be allophones in the Japanese syllabograms. People speaking the same language perceive these sounds as similar sounds.

A. CHALLENGES

Figure 1 demonstrates challenges and problems associated with Kuzushiji character recognition.

1) **Modern and Kuzushiji characters:** Some Kuzushiji characters are written in such a way that only context, particularly the identity of preceding characters can be used to recognize them. A simple line, for example, might signify a variety of characters based on the characters used before it. KuroNet can disambiguate by using the surrounding characters since it employs both local and global contexts. However, for efficient classification of characters, it is required to segment each character prior to the recognition process.
2) **Variation:** Although Kuzushiji has a huge number of characters (e.g., used dataset comprises 4328), their distribution is long-tailed. A significant portion of the characters only appears once or twice in the sample. This is owing to the Japanese language's peculiar structure that has two sorts of character sets, i.e., a phonetic alphabet (with simple and extremely frequent characters) and non-phonetic Kanji characters. Kanji is made up of both common and uncommon characters, some are very intricate and others are merely one or two straight lines.
3) **Hentaigana characters:** One feature of classical Hiragana or Hentaigana (the word literally translates to ''Character Variations'') that has a significant impact on recognition. Many characters that can be written exclusively in modern Japanese can be written in a variety of ways in pre-modern Japanese language [35].
4) **Aging processing:** It is hard to recognize the unclean characters due to infestation by worms.

Recently, researchers have utilized various deep learning models to improve the performance of Kuzushiji character recognition. Some of the popular models are as LeNet [16], AlexNet [17], GoogLeNet [18], VGG16, VGG19 [19], ResNet152V2 [21], InceptionNet [24], InceptionResNetV2 [20], XCeptionNet [25], and MobileNet [26]. But most of these models suffer from over-fitting and gradient vanishing problems. Additionally, an efficient approach is required to segment the characters prior to the recognition process.

B. CONTRIBUTIONS

To overcome the aforementioned problems, an efficient deep Kuzushiji characters recognition network (DKNet) is proposed. The main contributions of this paper are as follows:

- To overcome the over-fitting and gradient vanishing problems, an efficient deep Kuzushiji characters recognition network (DKNet) is proposed.
- To remove noise from the training images, a trilateral joint filter is employed. Contrast limited adaptive histogram equalization (CLAHE) is also applied to enhance the visibility of filtered images.
- In DKNet, a MobileNet's final layers are removed. The flatten layer is then applied to the input. Thereafter, a fully connected classification layer is implemented with Rectified linear units (ReLUs) and dropouts. Finally, the softmax activation function is employed to provide the recognition results.
- To test the proposed model, actual documents are first segmented by using the proposed Maximally stable extremal regions MSERs) and convexhull-based segmentation approach. Segmented characters are then recognized using the trained DKNet.
- An efficient Application Programming Interface (API) is also designed for Japanese Kuzushiji ancient heritage characters recognition to help the end-users.

The remaining paper is organized as follows. Section II discusses the literature review. Section III discusses the proposed DKNet. Section IV presents the proposed MSER and convexhull-based segmentation approach. Experimental results are presented in Section V. Section VI concludes the paper.

II. LITERATURE SURVEY

In [9], an adaptive neural network was utilized for character recognition by considering features of directional elements as feature vectors. In [11], line-by-line recognition was used to classify characters according to the attention mechanism. Lyu *et al.* [12] detected the text lines prior to the recognization process. Image processing operations and ARU-Net were used to identify text lines. It has achieved precision, recall, and F-score values as 95.2%,, −98.3%, and 96.6%, respectively. Le *et al.* [13] designed a Recognition system using an attention-based encoder-decoder (RAED). It reprocess data by detecting the line's start and recognizing each character until the line has been completed.

Li *et al.* [10] utilized three and five-cross point non-directed graphs of Oracle-bone inscriptions (OBIs) such as blocks and holes. Another recognition method was also used (refer to [14]). Li *et al.* [15] developed a DNA-based model for detection of OBIs. However, Meng *et al.* [5] and Liu *et al.* [28] proved that all OBIs needed to be cut in advance which is not convenient for practical applications. Meng *et al.* [5] extended the Single shot multibox detector (SSD) for OBI detection and achieved precision, recall, and F values as 98%, 83%, and 88%, respectively. Clanuwat *et al.* [40] recognized the Kuzushiji characters using a deep learning model. Residual U-Net architecture (RU-Net) was applied to recognize the entire page with identification of all characters given on it. In [41], Japanese historical text was recognized by attention based encoder-decoder (AED) model. Multiscale features were extracted using a dense convolutional neural network. The target text was generated using Long short-term memory (LSTM) decoder.

In [45], handwritten Japanese text was recognized using Attention augmented convolutional recurrent network (AACRN). The performance of the model was evaluated on Kuzushiji and TUAT Kondate datasets. In [46], the Deep learning model(DLM) was applied to recognize the Kuzushiji characters. Ueki *et al.* [47] proposed a recognition system for the Kuzushiji characters by utilizing Multiple softmax outputs (MSOs). In [48], a 2-dimensional Context box proposal network (2DCbpn) was proposed to identify the Kuzushiji Characters. Features were extracted using VGG16 and then features were processed through a 2-D context. Thereafter, horizontal and vertical contexts were explored using Bidirectional LSTM (BLSTM).

Bing *et al.* [42] proposed an method to understand the ancient Japanese books using conventional image processing and ARU-Net. These methods were used for the detection of frames and segmentation od lines. Character recognition was done using by applying AlexNet. Nguyen *et al.* [43] used Gaussian mixture mode, LSTM, and CNN to recognize the Japanese kanji characters. Septiana *et al.* [44] proposed template matching algorithm to recognize the handwritten Japanese characters. It achieved the accuracy of 89.8%.

Horie *et al.* [6] recognized OBIs using deep learning models. While they achieved good accuracy, the experimental data were drawn from recent research, not the original rubbings. OBIs are a kind of hieroglyphs that were the original version of the modern Chinese characters that are widely used in China, Japan, and other Asian countries. One OBI character is one hieroglyph graph, with a meaning corresponding to one thing or one object. Lyu *et al.* [12] developed a Japanese historical character recognition system for the reading support system for historical documents by using the directional elements as feature vectors and designing a modular neural network as a pattern classification model.

Nguyen *et al.* [36] segmented the image into patches for specific characters and then classified each patch separately. This strategy is computationally better, but due to the contextual nature of many characters, it is unsuited for the overall Kuzushiji recognition job. Le *et al.* created datasets with separately segmented Kazushi characters but did not take into account the Kuzushiji recognition [13].

Unger [23] considered several characters that can be exclusively rewritten for current Japanese in different ways in pre-modern Japanese scripts. It seems to be one element of classical Hiragana or Hentaigana (meaning ''Character Variations'') that has a considerable impact on recognition. Japanese MNIST dataset [13] is significantly more difficult than the original MNIST dataset [16]. Because many pre-modern Japanese characters can be written in numerous ways. Therefore, there is a need to develop good models that

must be able to reflect the multi-modal distribution of each class.

III. DEEP KUZUSHIJI CHARACTERS RECOGNITION NETWORK (DKNet)

Deep neural networks can be made lightweight using pre-trained models. Pre-trained models can reduce the number of parameters significantly when compared to networks with regular convolutions. To train the proposed model, MobileNet is used as a transfer learning model. MobileNet was designed by Google and it is an open-sourced model. This section initially briefly discusses various pre-trained models. Finally, the proposed deep Kuzushiji characters recognition network (DKNet) is discussed.

A. PRE-TRAINED LEARNING MODELS

This section discusses various pre-trained learning models which were trained on the ImageNet dataset.

1) RESIDUAL NEURAL NETWORK (ResNet)

It was presented at ILSVRC 2015. This architecture utilized the concept of skip connections. It comes up with 152 layers that allow us to train very deep neural networks without suffering from the gradient vanishing problem. Skipping connections are also called gated-recurrent units and are similar to recurrent neural networks based on recent successes [11], [24].

2) INCEPTION

The main aim of this architecture is to reduce the computational workload of deep neural networks. Originally proposed by GoogleNet, this model architecture was later applied to both Inception-V2 and Inception-V3. It can manipulate similar inputs in parallel to achieve better convergence [21], [25].

3) MobileNet

MobileNet is a well-known pre-trained model. It is an open-sourced model developed by Google as shown in the Table 1. To develop a lightweight model, it utilizes depth-wise separable convolutions. It minimizes the number of attributes than the network with regular convolutions with similar depth. Thus, it is a lightweight deep learning model [23], [26]. Two different hyper-parameters are used that can efficiently trade-off between accuracy and latency.

4) XCeptionNet

It is an extension of the inception model which can be viewed as an intermediate step between regular convolutions and depth-wise separable convolutions [22].

5) LeNet

It has five convolution layers. 32 images having 32×32 size for improving the classification performance. However, it has relatively few layers and simple architecture. LeNet cannot work properly for higher-resolution images.

TABLE 1. Mobile-net model architecture.

Type	Filter	Input
Conv/S2	3×3×3×32	224×224×3
Conv/dw/s1	3×3×32dw	112×112×32
Conv/s1	1×1×32×64dw	112×112×32
Conv/dw/s2	3×3×64dw	112×112×64
Conv/s1	1×1×128	56×56×64
Conv/s1	3×3×128dw	56×56×128
Conv/s1	1×1×128dw	56×56×128
Conv/dw/s2	3×3×128dw	56×56×128
Conv/s1	1×1×128×256	28×28×128
Conv/dw/s1	3×3×256dw	28×28×256
Conv/s1	1×1×256×512	14×14×256
5×Conv/s1/dw	3×3×512	14×14×512
Avgpool/s1	Pool×7×7	7×7×1024
FC/s1	1024×1000	1×1×1024
Softmax/s1	Classifier	1×1×1000

6) VGG

This was proposed by ILSVRC in 2014. It consists of 16 convolutional layers for uniform architecture. It is similar to AlexNet and has many filters and 138 million parameters. Hence, it has more computation power. Later, VGG-16 is extended to VGG-19 layers.

7) GoogLeNet

It achieved an error rate close to the human level in 2014. It consists of 22 layers with 224×224 sized images. The parameters are reduced from 60 million to 4 million [15].

8) InceptionNet

InceptionNet is created to reduce the computational burden of deep neural networks while maintaining performance. The idea is firstly proposed in GoogLeNet and then extended to Inception-V2, V3, and Inception-Res (Inception-V4). The key idea of InceptionNet is to compute multiple transformations over the same input map in parallel and provide the results into a single output. It has a compressed version of the spatial information that reduces the computation cost [16].

B. JOINT TRILATERAL FILTER

A joint trilateral filter is a well-known edge-preserving filter used to filer the Kuzushiji images. It considers a guided image G_{im} to filter the input image. Generally, the actual image I_m acts as a guided image. It can be implemented as:

$$J_{tf}(I_m) = \frac{\sum_{q\in\Omega} \rho^{pq}(G_{im}) \times I_q \times \sigma^2(I_q, G_{im})}{\sum_{q\in\Omega} \rho^{pq}(G_{im})} \tag{1}$$

Here, Ω is a kernel window at coordinate k dependent on the bilateral filter [37], [38].

A kernel weight, i.e., $\rho^{pq}(G_{im})$ can be computed as:

$$\rho^{pq}(G_{im}) = \frac{1}{|n|^2} \sum_{n:(p,q)\in\Omega} \left(1 + \frac{\left(G_{imp} - \mu_n\right)(G_{imq} - \mu_n)}{\sigma_n^2 + \epsilon}\right) \tag{2}$$

where $|n|$ shows pixels in window. μ_n and σ_n^2 are average and variance of G_{im} in local Ω. The weight associated to pixel q is maximum, if G_{imp} and G_{imq} are on similar sides of an edge, otherwise, a minimum weight is assigned to q.

C. CONTRAST LIMITED ADAPTIVE HISTOGRAM EQUALIZATION

The precision of Kuzushiji recognition is generally low if the images have poor visibility. The most of the enhancement methods are not effective for uneven lightening circumstances due to over-enhancement of images [39]. Therefore, CLAHE is used that can prevent excessive amplification of noise by limiting the contrast. It can be evaluated as:

$$f(D) = (1-\omega y)((1-\omega x)f_{ul}(D) + \omega x f_{bl}(D)) + \omega y((1-\omega x)f_{ur}(D) + \omega x f_{br}(D)) \quad (3)$$

The mapping of four adjacent values of the histogram cumulative distribution function to pixel are desirable for every pixel. ωx and ωy indicate the distance among center of the left upper mask and pixels. $f()$ is cumulative distribution function. f_{br}, f_{bl}, f_{ul}, and f_{ur} represent below right, below left, upper left and upper right values in current patch, respectively. D represents the pixel coordinates.

D. PROPOSED DKNet

Deep neural networks can be made lightweight using MobileNet's depth-separable convolutions. It reduces the number of parameters significantly when compared with networks with regular convolutions. Figure 2 shows the proposed DKNet model. Initially, joint trilateral filter is used to remove the noise from training images. Thereafter, contrast limited adaptive histogram equalization (CLAHE) is used to enhance the visibility of filtered images. Thereafter, MobileNet is used to extract the potential filters. Final layers of MobileNet, i.e., fully connected layer and softmax are removed. Flatten layer is then used to flatten the input. Thereafter, a fully connected classification layer is used. It utilizes two fully connected layers with a Rectified linear unit (ReLU) and dropouts are used. Dropouts are used to generalized the model, thus can prevent the over-fitting problem. Finally, the softmax activation function is used to provide the recognition results.

E. VALIDATION ON SELECTED DOCUMENTS

The trained DKNet model is tested on Kuzushiji documents using the following steps.

1) Initially, MSER and convexhull are used to extract the characters.
2) Apply joint trilateral filter on each extracted character image.
3) Apply CLAHE on the filtered character image.
4) Use trained DKNet model to recognize the character.

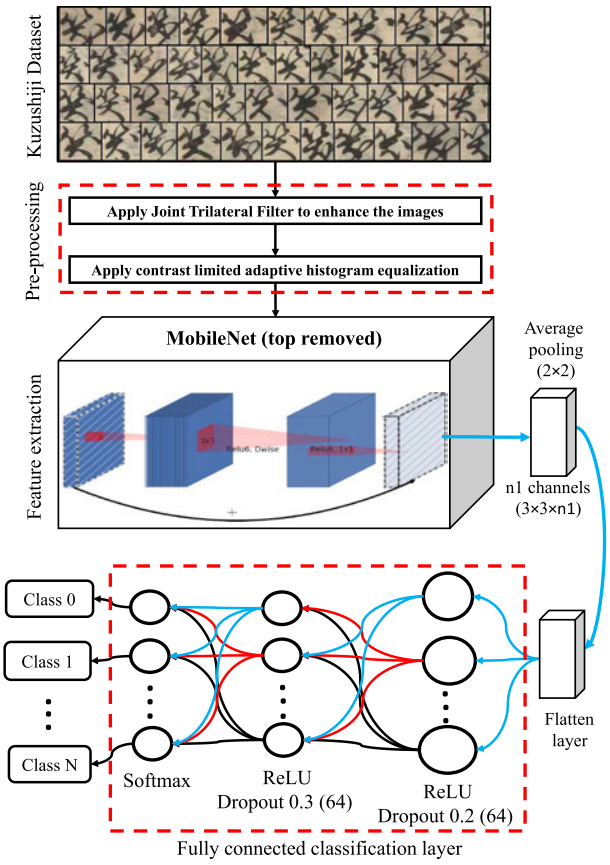


FIGURE 2. Diagrammatic flow of the proposed DKNet model.

IV. PROPOSED SEGMENTATION APPROACH FOR KUZUSHIJI CHARACTERS

This section discusses the proposed segmentation approach for Kuzushiji characters. Maximally stable extremal regions (MSER) and convexhull approaches are used to segment the characters from a given document.

Like a Scale-invariant feature detector (SIFT), MSER is a feature detector. It decomposes a given document to several co-variant regions, i.e., stable connected components or so-called MSERs. Figure 3 shows the diagrammatic flow of MSER technique.

To identify individual MSERs by characteristics of the Kuzushiji sample, MSER is applied to each document. To develop this pattern, binarization is achieved using OTSU threshold. The Pixel region of an area control is obtained by calculating the pixels region and constructing an ellipse using MinDiversity, MaxVariation, and Delta techniques. The contour is drawn around the arbitrary shape of the image using an elliptical frame that transposes the row index into a column index and vice versa. Based on a region seed, it returns a list of pixels within the region.

A. REGION OF GEOMETRIC PROPERTIES

The process of character detection and recognition can be achieved using following steps.

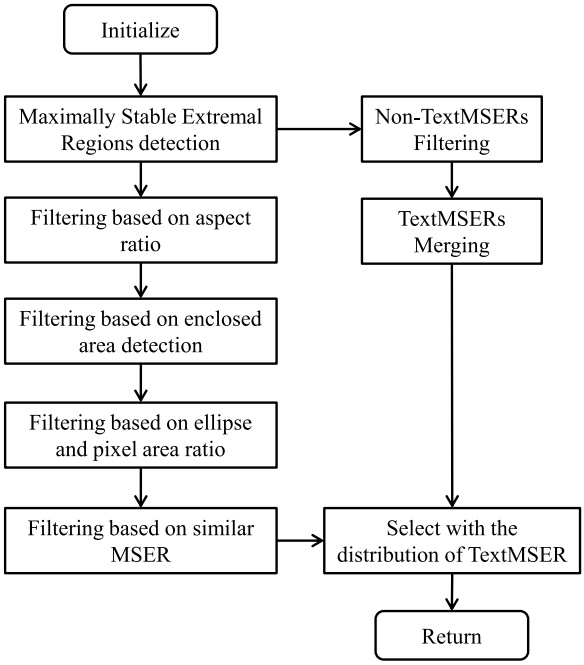


FIGURE 3. Diagrammatic flow of maximally stable extremal regions technique.

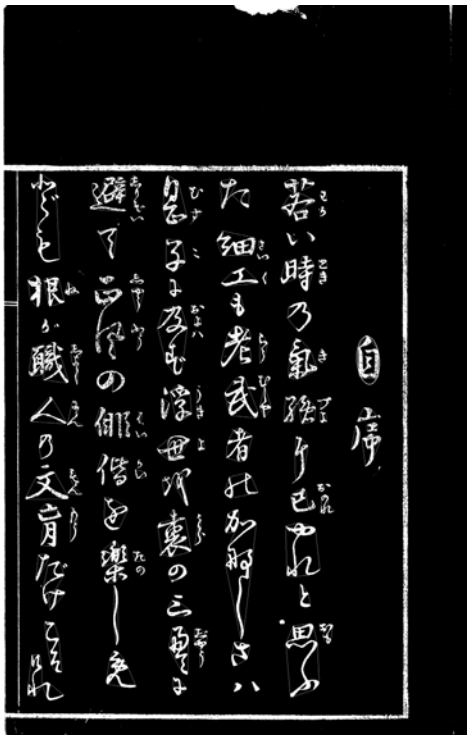


FIGURE 4. Region of image on which geometric properties are calculated.

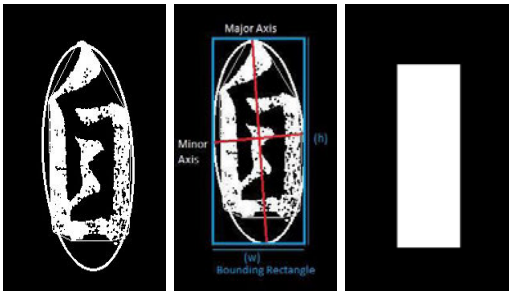


FIGURE 5. Geometric analysis using MSERs: (a) visualization of geometric properties, (b) rectangle image mask, and (c) sample image to calculate geometric properties using MSERs.

- **Pre-processing:** Various conversion processes are used to convert RGB images to binary as shown in Figure 4.
- **Process:** An image was processed with each MSER algorithm as feature detector.
- **Properties:** Geometric attributes are employed to discriminate between non-text regions and text regions in the first pass for deleting the non-text regions from the obtained features (see Figure 5(a)).
- **Text regions:** To compute the remaining text area, the overlapping bounding boxes of these regions are combined to generate a bounded text region (see Figures 5(b) and 5(c)).
- **Detected Image:** The identified text bounding boxes in an image are examined using deep learning models to evaluate the actual and the predicted text.
- **Aspect ratio:** The bounding box ratio between width and height of the character is calculated using aspect ratio.
- **Eccentricity:** A given character's circularity could be found using Eccentricity.
- **Extent:** To identify a region for character encoding, the rectangle enclosing text is determined to be a given value using Extent.
- **Solidity:** The convex hull area ratio is calculated for a given region of a character based on the pixels in the convex hull area.

The aspect ratio, eccentricity, extent, and solidity values are set to be 0.48130841121495327, 0.8839539479029316, 0.8026494873423464, and 0.07339475469138594, respectively.

MSER calculates the mean square error of the pixels of an image over its background that have a constant intensity over its background. Text characters are assumed to have a comparable intensity and color contrast to their background. Some characters can be ignored if these characters have a small font size or represent the font's pronunciation.

B. PRE-PROCESSING OF MSER

- Initially, white to black intensity thresholds are swept by using a simple luminance thresholding algorithm.
- External regions can be calculated as

$$S_r = p|I(p) > I(q)\forall_p \in S_r, \forall_q \in B(S_r) \tag{4}$$

The gray-scale image I is indexed by p and q. $B(S_r)$ contains the boundaries of S_r as mentioned in Eq. (4).

An external region is maximally stable if it satisfies the following conditions, reject otherwise.

- The threshold δ, is the number of intensity threshold levels that must be reached for a region to be considered maximally stable during MSER computation.
- MSERs demonstrate a minimum level of variation which implies a high level of stability than the regions from previous threshold levels. It can be computed as follows:

$$V = \frac{area(R+\delta) - area(R)}{area(R)} \tag{5}$$

Here, R represents a threshold value for excising the region.

- A region beyond the area range or area size of the region in pixels should be excluded. For instance, an object whose area size is beyond 90% range of the whole image or less than 5% range would be rejected since it has no relevance for the actual test scenario as mentioned in Eq. (5).
- At different intensity thresholds, the maximum area variation exists between extremal regions. Therefore, the stability of results is limited. The variation in intensity thresholds leads to providing stable regions with tiny areas.

C. MSER REGIONS

The output of MSER algorithm is a set of pixels representing MSER regions. Ellipses can also be calculated to represent MSER regions. To represent a pixel list, it is intuitive to draw a minimum rectangle that encloses all of its coordinates. To represent ellipses, some mathematical operators are used to obtain the bounding box of each ellipse.

In the first method, a minimum enclosing rectangle for pixels is obtained within the region starting from different representations of MSER region. In the second method, a sampling set is used to find ROI coordinates. A bounding box of (x_{min}, y_{min}) and (x_max, y_{max}) can be easily derived for the entire list of pixels. These parameters identify the minimum and maximum coordinate values for the horizontal and vertical axes at any point. By geometrical transformation of MSER ellipses, ROI can be obtained in the form of bounding boxes as follows:

$$width = max(2a\cos\theta,\ 2b\sin\theta) \tag{6}$$

$$height = max(2a\sin\theta,\ 2b\cos\theta) \tag{7}$$

Here, a and b define the center values of extracted region. θ shows the rotation transformation.

To generate bounding boxes for MSER regions, ellipses are used. MSER regions often contain extra pixels near the text especially when text blends with its surroundings.

D. CHARACTER DETECTION APPROACHES

- **Approach 1: Morphological Image Operations** The characters in this approach are enlarged to a specified threshold limit. It means that all the single characters are grouped and the character's pronunciations are ignored.

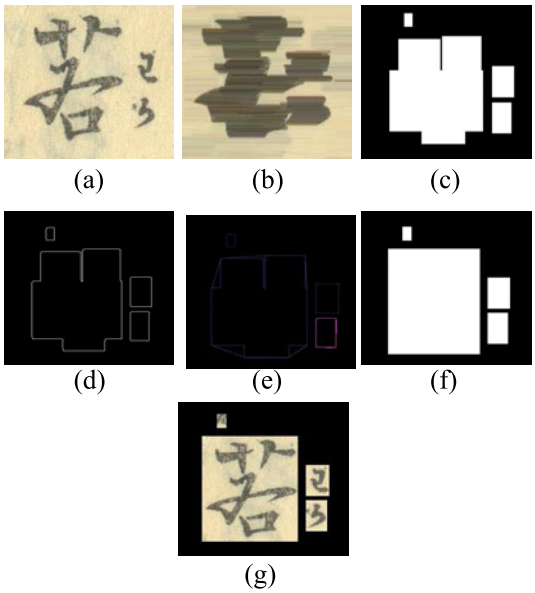


FIGURE 6. Analysis of morphological operation: (a) Input image, (b) dilatation, (c) bounding box extension, (d) canny edge detection, (e) contour technique, (f) mask applied on new contours, and (g) character extracted after operations.

Figure 6(a) and 6(b) illustrate the failed attempt to find the dilated characters for the given image and a sample of such a failure is shown below.

- **Approach 2: Contour technique** It is found that the gap between a single disconnected character is approximately in the range of 7 to 10 pixels as opposed to the gap between the character and its pronunciation which is approximately 14 pixels. An extension of the bounding box is made for this image about $7px$ to match the other disconnected component of the same character. The boundary mask area is drawn outward by $7px$ and the sample is shown in Figures 6(c), 6(d), and 6(e), respectively. Approximately 90% of all disconnected characters within the boundary mask meet (see Figures 6(f) and 6(g)).

V. EXPERIMENTAL RESULTS

Experiments are performed on Xeon Ef 16200 v4 CPU and GPU GeForce GTX1080Ti ×4 with memory as 64 GB and operating system is Ubuntu14.4. 3212 classes of Kuzushiji characters are used for building the DKNet model. The total number of training images is more than 600, 000. Kuzushiji dataset ID is obtained from the Center for Open Data in the Humanities (CODH) (refer to [29]). The parameter setting of DKNet and competitive models is presented in Table 2.

A. DATA AUGMENTATION

To augment the training data, three different augmentation techniques are also used. These techniques are random rotate, random blur, and random crop. The main objective is to augment the amount of training data to prevent the

TABLE 2. Parameters setting of DKNet and competitive models.

Parameter	Range
Padding pixels	0.1
Stride size	2
Filter size	7
Number of filters	6
Learning rate	0.01
epochs	100

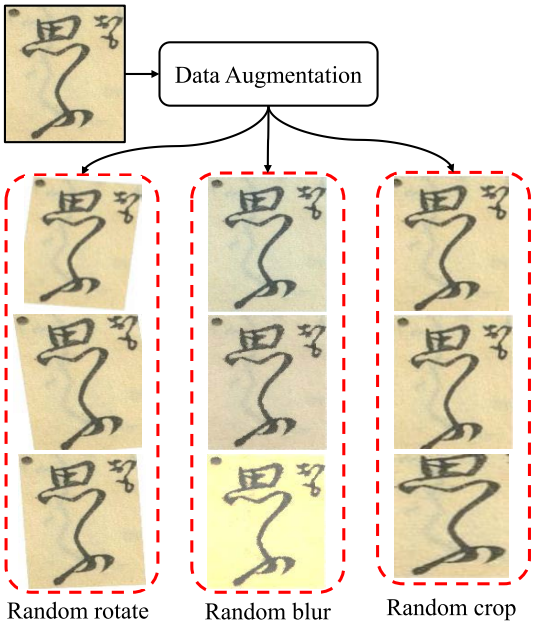


FIGURE 7. Impact of data augmentation techniques.

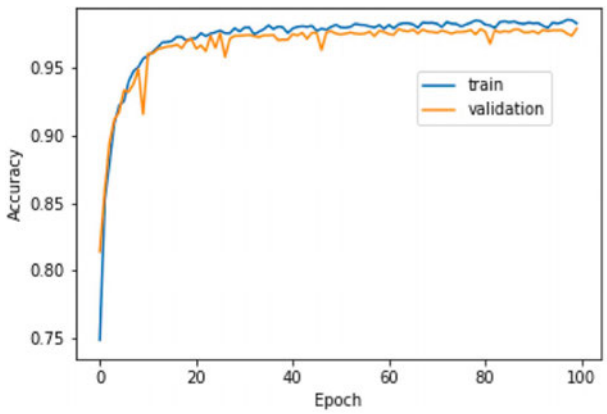


FIGURE 8. Training and validation accuracy of MobileNet.

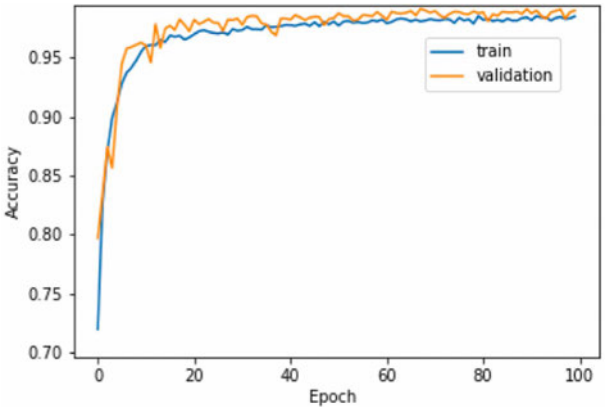


FIGURE 9. Training and validation accuracy of DKNet.

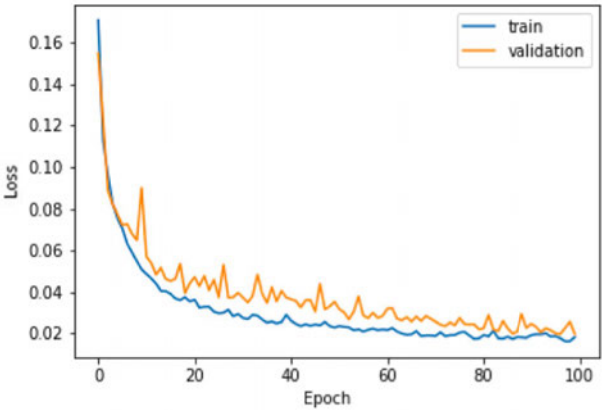


FIGURE 10. Training and validation loss of MobileNet.

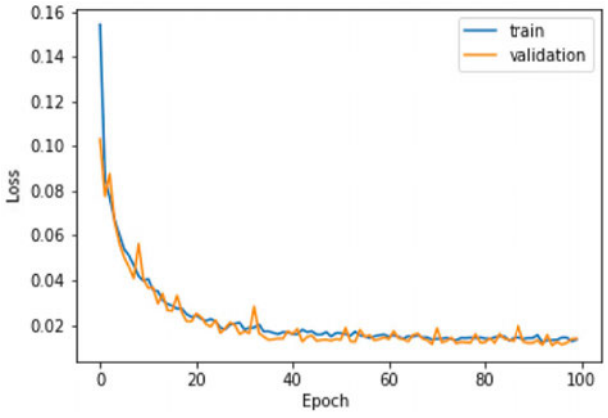


FIGURE 11. Training and validation loss of DKNet.

over-fitting and gradient vanishing problems. Figure 7 shows the impact of data augmentation techniques on a Kuzushiji character image.

B. TRAINING AND TESTING ANALYSIS

The training and validation accuracy analysis of MobileNet and DKNet are shown in Figures 8 and 9, respectively. The training and validation loss analysis of MobileNet and DKnet are shown in Figures 10 and 11, respectively. It is found that DKNet has achieved the better training and validation accuracy as compared to MobileNet. Hence, DKNet is selected and implemented in our online API.

C. PERFORMANCE ANALYSIS

The performance of DKNet is evaluated using confusion matrix-based metrics such as accuracy, F-measure, and area

TABLE 3. Training analysis of DKNet.

Model	AUC	CER	F-measure	SER	Accuracy
LeNet	94.15 ± 1.81	20.48 ± 1.58	94.87 ± 1.14	58.59 ± 1.43	94.91 ± 1.45
GoogLeNet	97.45 ± 1.23	20.14 ± 1.18	97.34 ± 1.52	57.18 ± 1.54	97.32 ± 1.34
VGG16	97.12 ± 1.34	19.29 ± 1.23	94.95 ± 1.11	56.08 ± 0.98	98.03 ± 0.99
VGG19	97.72 ± 0.44	18.04 ± 0.64	95.49 ± 0.44	55.72 ± 0.34	97.17 ± 0.32
ResNet152V2	97.59 ± 0.94	17.52 ± 0.93	97.21 ± 0.79	54.25 ± 0.74	97.24 ± 0.72
InceptionNet	97.41 ± 1.13	16.19 ± 1.53	97.25 ± 1.07	53.51 ± 0.95	97.53 ± 0.93
InceptionResNetV2	97.81 ± 1.03	15.84 ± 0.93	97.82 ± 0.74	52.94 ± 0.78	97.94 ± 0.84
XCeptionNet	97.81 ± 1.33	14.48 ± 1.01	97.83 ± 0.78	51.85 ± 1.04	97.88 ± 1.08
MobileNet	98.42 ± 0.68	13.41 ± 0.52	98.47 ± 0.63	50.74 ± 0.75	98.21 ± 0.85
Proposed DKNet model	**99.59 ± 0.39**	**11.36 ± 0.21**	**99.64 ± 0.35**	**48.61 ± 0.29**	**99.67 ± 0.31**

TABLE 4. Testing analysis of DKNet.

Model	AUC	CER	F-measure	SER	Accuracy
LeNet	94.75 ± 1.01	21.62 ± 1.72	94.74 ± 1.93	59.58 ± 1.92	94.71 ± 1.78
GoogLeNet	93.61 ± 1.31	21.54 ± 1.33	93.57 ± 1.64	58.51 ± 1.21	93.64 ± 1.23
VGG16	93.91 ± 1.24	20.77 ± 1.49	93.72 ± 1.21	57.74 ± 1.44	93.62 ± 1.44
VGG19	95.49 ± 1.41	19.19 ± 1.41	93.94 ± 1.41	56.21 ± 1.27	93.54 ± 1.22
ResNet152V2	94.31 ± 1.19	18.13 ± 1.14	94.15 ± 1.44	55.53 ± 1.43	93.51 ± 1.52
InceptionNet	95.21 ± 1.09	17.05 ± 1.14	93.44 ± 1.13	54.25 ± 1.09	94.83 ± 1.13
InceptionResNetV2	95.41 ± 1.17	16.81 ± 0.47	95.21 ± 0.89	53.29 ± 0.81	95.31 ± 0.22
XceptionNet	95.51 ± 1.25	15.12 ± 1.22	93.94 ± 1.22	52.02 ± 1.22	95.12 ± 1.46
MobileNet	96.63 ± 1.15	14.76 ± 1.51	95.32 ± 1.52	51.54 ± 0.42	96.34 ± 0.76
Proposed DKNet model	**98.42 ± 0.81**	**12.78 ± 0.17**	**98.75 ± 0.47**	**49.53 ± 0.83**	**98.57 ± 0.95**

under the curve (AUC). For evaluating handwriting recognition systems, we also use the Sequence Error Rate (SER) and Character Error Rate (CER) metrics [13]. DKNet's performance is evaluated through medians and uncertainty values (i.e., median $\pm$ $IQR \times 1.5$).

The nine models such as LeNet, GoogLeNet, VGG16, VGG19, ResNet152V2, Inception, InceptionResNetV2, XCeptionNet, MobileNet, and DKNet are evaluated and the results are summarized in Tables 3 and 4.

Tables 3 and 4 illustrate the performance analysis of DKNet. DKNet achieves 99.67 and 98.57 accuracy values during training and testing. This means that the overfitting issue is not an issue. Further, DKNet was able to achieve an AUC value of 99.59 during training and 98.42 during testing. The proposed model is therefore least affected by the false positives and false negatives.

D. DISCUSSION

From the existing literature, it has been found that machine learning and deep learning techniques have been successfully used for Kuzushiji character recognition. Many models have been designed and implemented to recognize Kuzushiji characters from Japanese historical documents. Some well-known recognition models are RU-Net [40], AED [41], AACRN [45], 2DCbpn [48], MSOs [47], and RAED [13]. Although these models have shown significant results, the majority of these models have shown poor performance analysis. Therefore, an efficient deep Japanese recognition model using MSER is proposed. Table 5 demonstrates the performance analyses of DKNet and existing models.

RU-Net provides 86.83% accuracy and 85.53% F-measure. 2DCbpn provides better accuracy i.e., 93.32%, and F-measure i.e., 92.42 as compared to RU-Net. The accuracy of MSOs is 94.67% which is significantly better than RU-Net and 2DCbpn. DKNet provides 98.57% accuracy and 98.57% F-measure. Hence, it can be seen that DKNet outperforms the existing techniques in terms of accuracy and F-measure. The CER of AED is 32.34%. CER and SER of AACRN are 21.49% and 94.97%. CER of AACRN is better than AED. RAED provides 13.07% CER and 53.81% SER which is significantly better than AED and AACRN. DKNet provides 12.78% CER and 49.53% SER. It can be seen that DKNet outperforms AED, AACRN, and ADE in terms of CER and SER. Therefore, the proposed DKNet model is an efficient recognition system. In future, performnce of DKNet can be improved by tuning its hyper-parameters using metaheuristic techniques such as bat algorithm [50], particle swarm optimization [49], [51], etc.

TABLE 5. Performance comparison of DKNet with existing models.

Model	Database	Accuracy	F-measure	CER	SER
RU-Net [40]	CODH	86.83	85.48	-	-
AED [41]	-	-	-	32.34	-
AACRN [45]	CODH	-	-	21.48	94.97
2DCbpn [48]	CODH	93.32	92.42	-	-
MSOs [47]	CODH	94.67	-	-	-
RAED [13]	CODH	-	-	13.07	53.81
DKNet	CODH	**98.57**	**98.75**	**12.78**	**49.53**

E. EXPERIMENTAL ANALYSIS OF API

The recognition system is equipped on sever for online access. The CPU is Intel(R) Xeon(R) CPU E5-1410 v2 @ 2.80 GHz, RAM is 8G, and OS is Ubuntu 18.04.3 LTS. Apache HTTP Server was used for designing the API website.

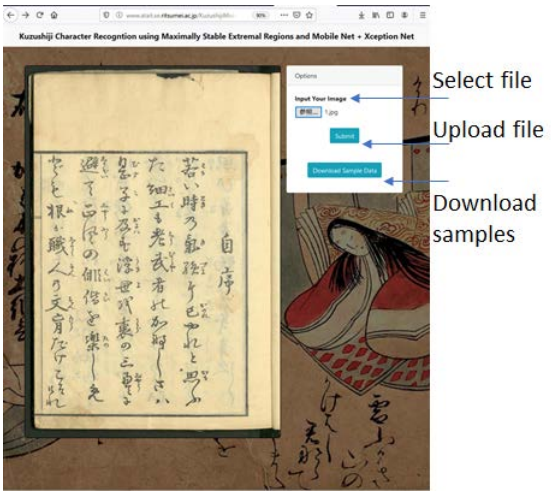


FIGURE 12. Interface of API.

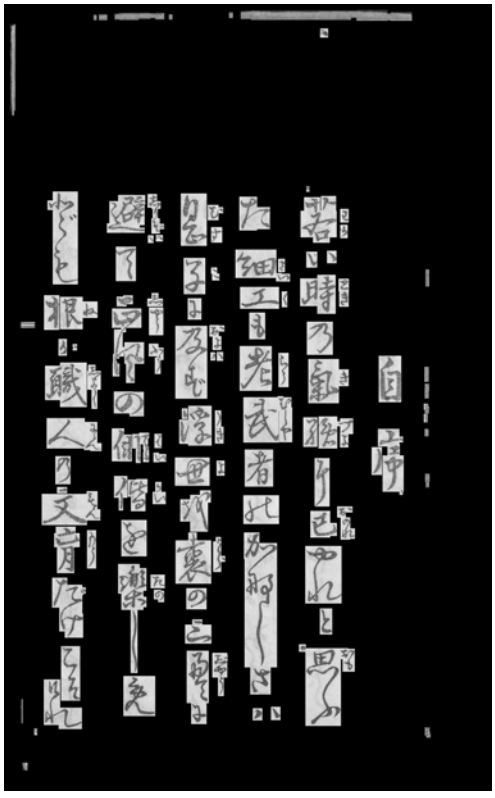


FIGURE 13. Masking MSER Technique.

Figure 12 shows the designed API, the target file can be selected from the local computer and submitted to our server. The user may download the sample files for testing the API.

Figures 13 and 14 show the obtained results of the proposed API. It is found that lots of characters are segmented correctly. The segmented characters are shown in the bottom of interface (see Figures 15 and 16). It is found that MSER with a bounding mask can achieve better results when there are significant gaps between characters. The proposed MSER and convex hull-based segmentation approach can segment

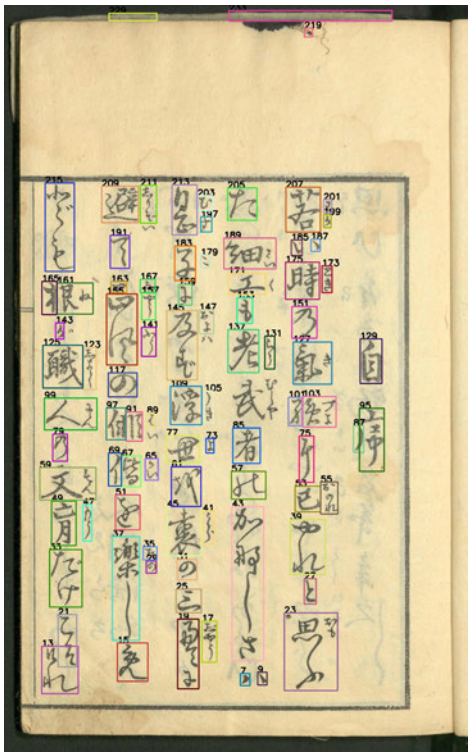


FIGURE 14. Segmentation of characters.

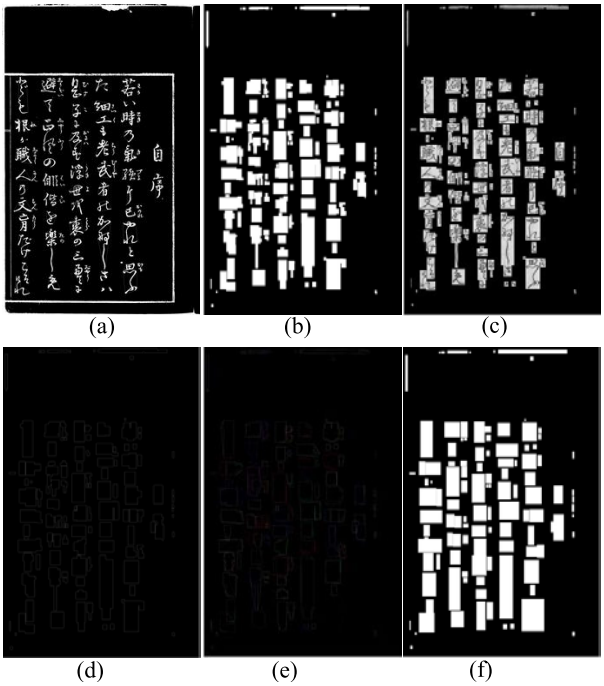


FIGURE 15. Extraction of characters: (a) MSER resultant image, (b) bounding mask drawn around possible characters, (c) text extracted with mask on the input image, (d) canny edge extracted from image mask, (e) convex hulls drawn around the canny edge extracted image, and (f) extraction of bounding mask from the convex hulls that contains characters.

characters when there is no or lesser gap between the characters (refer Fig. 16(a)). Whenever the user clicks the



(a) (b)

FIGURE 16. Extraction of single character: (a) extraction of text using Fig. 15(b) and (b) extraction of text using Fig.15(f).

characters, they will be recognized by the proposed model (refer Fig. 16(b)).

VI. CONCLUSION

A deep Kuzushiji characters recognition network (DKNet) was proposed to overcome over-fitting and gradient vanishing problems. First, the training images were filtered with a joint trilateral filter to remove noise. The filtered images were then enhanced with a contrast limited adaptive histogram equalization (CLAHE). A pre-trained MobileNet was then used to extract features from Kuzushiji characters. MobileNet's final layers, such as the fully connected layer and softmax, were removed. The flatten layer was then applied to the input. It was followed by a fully connected classification layer with rectified linear units (ReLU) and dropouts. Dropouts were used to generalize the model, thus preventing the over-fitting problem. Finally, the recognition results were obtained by applying the softmax activation function. To test the proposed model, actual documents were segmented using the proposed MSERs and a convexhull-based segmentation approach. Segmented characters were then recognized by the trained DKNet. Analyses of the competitive models revealed that DKNet surpassed the competitive models in terms of several performance metrics such as accuracy, F-measure, AUC, CER, and SER by 2.1675%, 1.9842%, 2.1075%, 1.6914%, and 1.7265%, respectively. An efficient Application Programming Interface (API) was also developed for Japanese Kuzushiji ancient heritage character recognition to help the end-users.

REFERENCES

[1] W. H. Bangyal and J. Q. A. Abbas, "Recognition of off-line isolated handwritten character using counter propagation network," *Int. J. Eng. Technol.*, vol. 5, no. 2, p. 227, 2013.
[2] Q. Abbas, W. H. Bangyal, and J. Ahmad, "Analysis of learning rate using BP algorithm for hand written digit recognition application," in *Proc. Int. Conf. Inf. Emerg. Technol.*, Jun. 2010, pp. 1–5.
[3] N. Fujisaki, Y. Ishikawa, M. Takata, and K. Joe, "Crawling low appearance frequency character images for early-modern Japanese printed character recognition," in *Advances in Parallel & Distributed Processing, and Applications*. Cham, Switzerland: Springer, 2021, pp. 683–695.
[4] C. V. Aravinda, M. Lin, A. Masahiko, and P. G. Amar, "A complete methodology for Kuzushiji historical character recognition using multiple features approach and deep learning Model," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 8, pp. 1–7, 2020.
[5] L. Meng, B. Lyu, Z. Zhang, C. V. Aravinda, N. Kamitoku, and K. Yamazaki, "Oracle bone inscription detector based on SSD," in *Proc. Int. Conf. Image Anal. Process.* Cham, Switzerland: Springer, 2019, pp. 126–136.
[6] F. Horie, H. Goto, and T. Suganuma, "Synthetic scene character generator and ensemble scheme with the random image feature method for Japanese and Chinese scene character recognition," *IEICE Trans. Inf. Syst.*, vol. 104, no. 11, pp. 2002–2010, Nov. 2021.
[7] B. Lyu, Z. Wang, H. Li, A. Tanaka, K. Funumoto, and L. Meng, "Deep leaning based medicine packaging information recognition for medication use in the elderly," *Proc. Comput. Sci.*, vol. 187, pp. 194–199, Jan. 2021.
[8] A. Sayeed, J. Shin, M. A. M. Hasan, A. Y. Srizon, and M. M. Hasan, "BengaliNet: A low-cost novel convolutional neural network for Bengali handwritten characters recognition," *Appl. Sci.*, vol. 11, no. 15, p. 6845, Jul. 2021.
[9] T. Horiuchi and S. Kato, "A study on Japanese historical character recognition using modular neural networks," *Int. J. Innov. Comput., Inf. Control*, vol. 7, no. 8, pp. 5003–5014, 2005.
[10] F. Li and P.-Y. Woo, "The coding principle and method for automatic recognition of Jia Gu Wen characters," *Int. J. Hum. Comput. Stud.*, vol. 53, no. 2, pp. 289–299, Aug. 2000.
[11] X. Hu, M. Inamoto, and A. Konagaya, "Recognition of Kuzushi-Ji with deep learning method a case study of Kiritsubo chapter in the Tale of Genji," in *Proc. 33rd Annu. Conf. Jpn. Soc. Artif. Intell.*, 2019, pp. 1–2.
[12] B. Lyu, R. Akama, H. Tomiyama, and L. Meng, "The early Japanese books text line segmentation base on image processing and deep learning," in *Proc. Int. Conf. Adv. Mech. Syst. (ICAMechS)*, Aug. 2019, pp. 299–304.
[13] A. D. Le, T. Clanuwat, and A. Kitamoto, "A human-inspired recognition system for pre-modern Japanese historical documents," *IEEE Access*, vol. 7, pp. 84163–84169, 2019.
[14] Q. Li, Y. Yang, and A. Wang, "Recognition of inscriptions on bones or tortoise shells based on graph isomorphism," *Comput. Eng. Appl.*, vol. 47, no. 8, pp. 112–114, 2011.
[15] Q. S. Li and Y. X. Yang, "Sticker DNA algorithm of Oracle-bone inscriptions retrieving," *Comput. Eng. Appl.*, vol. 44, no. 28, pp. 140–142, 2008.
[16] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
[18] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.
[19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[20] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-V4, inception-ResNet and the impact of residual connections on learning," 2016, *arXiv:1602.07261*.
[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015, *arXiv:1512.03385*.
[22] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2016, pp. 4700–4708.
[23] K. Brown, *Encyclopedia of Language & Linguistics*, J. M. Unger, Ed., 2nd ed. Amsterdam, The Netherlands: Elsevier, 2006, pp. 95–102, doi: 10.1016/B0-08-044854-2/04566-1.
[24] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.
[25] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017.

[26] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[27] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottleneck," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 4510–4520.

[28] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot MultiBox detector," 2015, *arXiv:1512.02325*.

[29] Center for Open Data in the Humanities (CODH). *Japanese Classic Cursive Script Data Set Book Title List [Data Set]*. Accessed: Feb. 26, 2020. [Online]. Available: http://codh.rois.ac.jp/char-shape/book/

[30] C. V. Aravinda, M. Lin, and P. G. Amar. *Kuzashi Recognition API*. Accessed: Feb. 26, 2020. [Online]. Available: http://www.atait.se.ritsumei.ac.jp/KuzushijiMser/

[31] L. Meng, C. V. Aravinda, K. R. U. K. Reddy, T. Izumi, and K. Yamazaki, "Ancient Asian character recognition for literature preservation and understanding," in *Proc. Euro-Mediterranean Conf.* Cham, Switzerland: Springer, 2018, pp. 741–751.

[32] C. V. Aravinda and H. N. Prakash, "A review on various handwritten cursive characters segmentation techniques," in *Proc. 2nd Int. Conf. Appl. Theor. Comput. Commun. Technol. (iCATccT)*, 2016, pp. 647–651.

[33] C. V. Aravinda and H. N. Prakash, "Template matching method for Kannada handwritten recognition based on correlation analysis," in *Proc. Int. Conf. Contemp. Comput. Informat. (ICI)*, Nov. 2014, pp. 857–861.

[34] L. Meng, "Recognition of Oracle bone inscriptions by extracting line features on image processing," in *Proc. 6th Int. Conf. Pattern Recognit. Appl. Methods (ICPRAM)*, 2017, pp. 606–611, doi: 10.5220/0006225706060611.

[35] L. Meng, "Two-stage recognition for Oracle bone inscriptions," in *Proc. ICIAP*, 2017, pp. 672–682, doi: 10.1007/978-3-319-68548-9.

[36] H. T. Nguyen, N. T. Ly, K. C. Nguyen, C. T. Nguyen, and M. Nakagawa, "Attempts to recognize anomalously deformed Kana in Japanese historical documents," in *Proc. 4th Int. Workshop Historical Document Imag. Process.*, 2017, pp. 31–36.

[37] D. Singh and V. Kumar, "Dehazing of remote sensing images using improved restoration model based dark channel prior," *Imag. Sci. J.*, vol. 65, no. 5, pp. 282–292, 2017.

[38] M. Kaur, D. Singh, V. Kumar, and K. Sun, "Color image dehazing using gradient channel prior and guided L_0 filter," *Inf. Sci.*, vol. 521, pp. 326–342, Jun. 2020.

[39] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," *J. VLSI Signal Process. Syst. Signal, Image Video Technol.*, vol. 38, no. 1, pp. 35–44, 2004.

[40] T. Clanuwat, A. Lamb, and A. Kitamoto, "KuroNet: Pre-modern Japanese kuzushiji character recognition with deep learning," in *Proc. Int. Conf. Document Anal. Recognit. (ICDAR)*, Sep. 2019, pp. 607–614.

[41] A. D. Le, D. Mochihashi, K. Masuda, and H. Mima. (2018). *An Attention-Based Encoder–Decoder for Recognizing Japanese Historical Documents*. Pattern Recognition and Machine Understanding (PRMU). [Online]. Available: https://www.ieice.org/ken/paper/20181213H1IM/eng

[42] L. Bing, H. Tomiyama, and L. Meng, "Frame detection and text line segmentation for early Japanese books understanding," in *Proc. 9th Int. Conf. Pattern Recognit. Appl. Methods*, 2020, pp. 600–606, doi: 10.5220/0009179306000606.

[43] H. T. Nguyen, T. Nakamura, C. T. Nguyen, and M. Nakawaga, "Online trajectory recovery from offline handwritten Japanese kanji characters of multiple strokes," in *Proc. 25th Int. Conf. Pattern Recognit. (ICPR)*, Jan. 2021, pp. 8320–8327.

[44] Y. Septiana, A. Mulyani, D. Kurniadi, and H. Hasanudin, "Handwritten recognition of Hiragana and Katakana characters based on template matching algorithm," in *Proc. IOP Conf. Mater. Sci. Eng.*, vol. 1098, no. 3, Mar. 2021, Art. no. 032093.

[45] N. T. Ly, C. T. Nguyen, and M. Nakagawa, "Attention augmented convolutional recurrent network for handwritten Japanese text recognition," in *Proc. 17th Int. Conf. Frontiers Handwriting Recognit. (ICFHR)*, Sep. 2020, pp. 163–168.

[46] L. Sichao and H. Miwa, "Algorithm using deep learning for recognition of Japanese historical characters in photo image of historical book," in *Proc. Int. Conf. Intell. Netw. Collaborative Syst.* Cham, Switzerland: Springer, 2019, pp. 181–189.

[47] K. Ueki, T. Kojima, R. Mutou, R. S. Nezhad, and Y. Hagiwara, "Recognition of Japanese connected cursive characters using multiple softmax outputs," in *Proc. IEEE Conf. Multimedia Inf. Process. Retr. (MIPR)*, Aug. 2020, pp. 127–130.

[48] A. D. Le, "Detecting Kuzushiji characters from historical documents by two-dimensional context box proposal network," in *Proc. Int. Conf. Future Data Secur. Eng.* Cham, Switzerland: Springer, 2019, pp. 731–738.

[49] W. H. Bangyal, A. Hameed, J. Ahmad, K. Nisar, M. R. Haque, A. A. A. Ibrahim, J. J. P. C. Rodrigues, M. A. Khan, D. B. Rawat, and R. Etengu, "New modified controlled bat algorithm for numerical optimization problem," *Comput., Mater. Continua*, vol. 70, no. 2, pp. 2241–2259, 2022.

[50] S. Pervaiz, Z. Ul-Qayyum, W. H. Bangyal, L. Gao, and J. Ahmad, "A systematic literature review on particle swarm optimization techniques for medical diseases detection," *Comput. Math. Methods Med.*, vol. 2021, pp. 1–10, Sep. 2021.

[51] W. H. Bangyal, A. Hameed, W. Alosaimi, and H. Alyami, "A new initialization approach in particle swarm optimization for global optimization problems," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1–17, May 2021.

Predicting the Performance of Cooperative Wireless Networking Schemes With Random Network Coding

Jin-Taek Seong and Heung-No Lee, *Senior Member, IEEE*

***Abstract*—In this paper, we consider a cooperative wireless network in which there are multiple sources and multiple relays. Owing to unreliable wireless channels, the quality of network links between nodes can vary. This results in the failure of intermediate nodes that generate linear combinations of incoming messages in network coding schemes. We propose an analytical framework to evaluate the recovery performance of source messages at the base station. To this end, we consider a random transmission matrix in which each element of the transmission matrix is processed a random variable, where its distribution is a function of the outage probability. We derive an upper bound for the reconstruction performance, i.e., decoding failure probability and nullity. The proposed framework provides an evaluation tool that enables us to investigate the impact of a large number of sources and relays, as well as the field size of the network codes on system performance.**

***Index Terms*—Cooperative network, network coding, upper bound, rank.**

I. Introduction

CHANNEL fading is one of the underlying causes of the performance degradation in wireless networks. One naive approach to combat channel fading is to increase the transmit power. A more advanced approach is to utilize modern diversity techniques, which can be performed without increasing the transmit power. To date, numerous diversity techniques have been proposed and employed in the time, frequency, and space domains. Cooperative networking is one of the current approaches that aim to utilize spatial diversity via user cooperation. Each user participates collaboratively, and shares the benefit of a virtual antenna array in transceiver messages that are available through another user's antenna [1].

Network coding [2] first proposed by Ahlswede *et al.* is shown to achieve maximum information flow in a single source multicast network. Numerous efforts have subsequently been attempted; these efforts focused on elucidating if network coding can provide additional advantages compared to other cooperative networking schemes. For example, network coding over the binary field [3], [4] has shown to improve diversity gain and provide higher spectral efficiency in wireless networks, whereas network coding with a nonbinary field further increase those benefits [5]–[9]. In particular, numerous studies have investigated the extent to which network coding can improve the performance of media access control and routing protocols, in terms of energy efficiency [22], transmission delay [24], and throughput [23], [25], compared to traditional forward-and-relay only based designs [19]–[21]. The performance of cooperative wireless networking schemes with network coding has been analyzed, and compared with erasure channel models [5]–[8], and error propagation models [30]–[33]. We will further address recent cooperative communication schemes in Section II by categorizing them with respect to their decoding techniques, spectral efficiencies, and cooperative strategies.

Xiao and Skoglund [5], [6] recently proposed a network coding scheme called Dynamic Network Codes (DNC) to handle a dynamic network topology. The inherent nature of wireless channels implies that links are unreliable and that link failures will occur randomly in the inter-user channels. DNC is perform successfully over such a dynamic network channel topology, in conjunction with techniques such as enhanced diversity order. In the DNC schemes, multiple network code matrices are used; each one is designed to handle a particular occurrence of link outages. In particular, an intermediate node in a network may fail to decode some of the messages received from the other nodes. The intermediate node creates, and later on forwards it to the base station, a linear combination of the messages which it could only successfully decode, and then forwards it to the base station. That is, a certain occurrence of link outages results in a particular restriction to the elements of the network code matrix. Thus, each network code matrix in the DNC scheme, referred to as the *transmission matrix* in this paper, is designed carefully so as to work effectively for the occurrence of a specific set of link failures. In addition, Rebelatto *et al.* in [7] and [8] extended the two-phase transmission framework of the DNC to multiple phases, in the Generalized Dynamic Network Codes (GDNC), to further enhance the transmission rate and the diversity order.

In this paper, our goal is to focus on the system models of the DNC and GDNC schemes, and provide a novel analysis framework for them. As noted earlier, there are other recently studied cooperative communications schemes, each with an advantage in a different perspective such as spectral efficiency and higher decoding performance. These will be discussed further in Section II. The DNC and GDNC schemes are found to be interesting due at least to the following two aspects: i) DNC is the first network coding schemes designed for dynamic network topology. To the best of the authors' knowledge, DNC is the first attempt to adaptively use different network code

Manuscript received October 10, 2013; revised March 11, 2014 and May 18, 2014; accepted June 4, 2014. Date of publication June 13, 2014; date of current version August 20, 2014. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST) through the Do-Yak Research Program under Grant 2010-0017944. The associate editor coordinating the review of this paper and approving it for publication was T. Tsiftsis.

The authors are with the School of Information and Communications, Gwangju Institute of Science and Technology, Gwangju 500-712, Korea (e-mail: jtseong@gist.ac.kr; heungno@gist.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2014.2330825

0090-6778 © 2014 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

matrix as the link failure varies, and is designed to achieve the so-called the *min-cut capacity* of randomly changing links [5], ii) GDNC is shown to achieve full diversity order and increases the transmission rate [7]. While a series of performance analyses for DNC and GDNC.

Although a series of performance analyses for DNC and GDNC are provided in [5]–[8], the authors rely on the exhaustive investigation of all individual network code matrices to determine if the resultant transmission matrix at the base station is sufficiently able to decode the source messages. This is an exceedingly time-consuming and tedious process; thus, it cannot be extended to larger and more general networks where the link outage probabilities throughout the networks are, in general, different from each other.

In particular, the performance analyses in [5]–[8] to determine the probability of successful decoding at the base station are performed only for small and non-general networks. A successful decoding is assumed to be achieved when the network code matrix at the base station has a sufficient number of linearly independent vectors that at least equals the number of unknown source messages. The successful event begins by determining whether the rank of the network code matrix at the base station is full. Then, the success probability is obtained by adding all individual probabilities of such events over all possible link failures. To achieve this outcome, the authors, using Theorem 1 in [5] and Section V in [7], followed the approach of tracking down each network code matrix individually, and determining if each was full in rank. This is an exhaustive process. When the number of nodes in a network increases, it is evident that this approach becomes intractable, because of the exponential increase in possible combinations. For example, the total number of distinct $N \times N$ random matrices with full rank is $\prod_{i=1}^{N}(q^N - q^{i-1})$ for the finite field of size q [10].

As a result, the analyses performed in [5]–[8] are limited to small and homogeneous networks, i.e., a network of fewer than 10 nodes, with link failure probabilities set to be equal throughout the network. These methods are not suitable for analyzing networks where nodes are randomly deployed in an area of interests, and wireless networks present heterogeneous link outage probabilities. The lack of a general and systematic performance analysis framework to deal with such networks has motivated this study.

In this paper, our main goal is to propose a novel evaluation framework for cooperative network coding schemes. The contributions of this work are summarized as follows.

- (*Design of random transmission matrix*) We model a random transmission matrix with uniform and maximum distance separable (MDS) distributions in (6)–(8) of Section IV. The elements of this matrix are represented with random variables as a function of the outage probability of each wireless link. This new system model, enables us to avoid the exhaustive counting of each network code metric occurrence.
- (*Tight upper bounds to probability of decoding failure*) We derive a series of tight upper bounds on the probability of failure. In particular, the dimension of the nullspace of the random transmission matrix is used to derive an upper bound, as discussed in Proposition 3. It is then linked to the decoding failure probability where the rank of the network code matrix is not full, which is shown in Theorem 4. The upper bounds have proven to be considerably tight in comparison to simulation results.
- (*Generality and scalability*) The developed analysis framework is general and scalable, offering the capability of analyzing large wireless networks with random deployments where all outage probabilities of wireless links are different. For example, consider the network scenario of randomly deployed nodes shown in Fig. 6 in Section VI. In addition, the developed framework can handle a large cooperative network that has more than 100 nodes. To the best of our knowledge, this scale of wireless networks is unprecedented in DNC and GDNC performance evaluations. The proposed framework enables us to investigate its impact on the successful reconstruction of source messages based on varying outage probabilities, and on various key parameters such as the number of relays and the field sizes in DNC and GDNC schemes.

TABLE I
SUMMARY OF RECENT TECHNIQUES FOR COOPERATIVE COMMUNICATIONS

	Recent technique	References
Decoding techniques	Maximum ratio combining	[4], [13]–[15]
	Rank-based decoding	[5]–[8], [10], [11]
High spectral efficiency	Multiuser detection	[17]
	Interference cancellation	[18]
	Non-orthogonal channel	[16]
Cooperative strategies	Amplify-and-Forward	[1], [35]
	Decode-and-Forward	[1], [36]
	Compute-and-Forward	[34]

The remainder of this paper is organized as follows. We review other recent cooperative communications techniques in Section II. We explain our cooperative model for wireless networks in Section III. In Section IV, we model a transmission matrix in terms of an outage probability. In Section V, we calculate upper bounds for the reconstruction performance of cooperative frameworks with various types of link connectivity. Finally, we evaluate the proposed framework in Section VI, and conclude the paper in Section VII.

II. OTHER RECENT WORKS AND RELATION TO OUR WORK

In this section, and in Table I, we provide an overview of the prevalent cooperative communications schemes believed to be closely related to the network coding schemes considered in this paper.

Two decoding approaches in cooperative wireless communications have been recently considered. The first is the Maximum Ratio Combining (MRC) decoding scheme [4], [13]–[15]. The main idea is to configure the network to coordinate the transmissions, and repeat a signal with a weak signal-to-noise ratio (SNR) multiple times over independent fading channels. In this manner, MRC allows the destination to maximize SNR. To implement an MRC scheme, all decoding information, as well as the success or failure of each source message at the base station, should be identified and forwarded to the relays. This is required to determine the source of the SNR, for which

retransmission was required. To enable its deployment in large-scale networks, the scheduling issue must be resolved.

Recently, two research groups have proposed advanced cooperative network schemes to achieve high spectral efficiency. In [16], Youssef and Amat have proposed the use of non-orthogonal channel allocation to improve spectral efficiency. For wireless networks where a multiuser detection receiver is utilized at the base station, cooperative transmission protocols with high spectral efficiency have been developed [17]. Furthermore, the work describe in [18] has aimed at improving the spectral efficiency of cooperative systems using superposition coding and iterative detection methods.

We believe our analysis framework could be utilized in [16] with a necessary but simple modification. The first aspect to consider is that all wireless channels should not be modeled independently from each other anymore. The outage probabilities are not independent from each other. This can be achieved by designing a joint probability distribution for the random matrix. The probability that the random transmission matrix is not full rank can be obtained by considering such an event over the joint probability distribution. For more details, we will show Example 3 for channel correlation cases in Section V-B of this paper.

Using network coding on lattice codes, Nazer and Gastpar recently proposed the compute-and-forward (CF) relaying scheme [34]. In CF, a relay is configured with a linear combination of multiple codeword signals, which are simultaneously transmitted and superposed in the air. The key idea of using CF relaying in network coding is to utilize the property of the lattice code property stating that the *integer combination* of lattice codewords remains a lattice codeword. Thus, the relay receives a codeword with additive noise. After a denoising step, the relay retransmits the decoded lattice codeword to the base station. Therefore, the benefit of using CF relaying in network coding is evident. Because the transmissions from all sources to the relay are performed simultaneously, the spectral efficiency is significantly enhanced.

There are two widely recognized cooperative relaying strategies, referred to as amplify-and-forward (AF) [1], [35] and decode-and-forward (DF) [1], [36]. AF and DF cooperative relaying strategies perform effectively in either low or high SNR regimes, while CF approach offers advantages in moderate SNR regimes where both interference and noise are significant factors [34]. The DNC and GDNC schemes considered in this paper are categorized as DF based network coding strategies.

Cooperative wireless communications with multiple sources and multiple relays closely related to our work have attracted significant attention because of their higher achievability rate [26], better error performance [28], [29] and diversity-multiplexing tradeoffs [9], [27]. There are two types of error propagation models worth that should be considered here. The authors in [30]–[33] assume a network channel model where erroneous messages are permitted to propagate throughout the network. For the erasure channel model, [5]–[8], erroneous messages at the relays are discarded, to avoid unnecessary error propagation caused by encoding and forwarding operations.

Recent studies [5]–[12] closely related to this paper have focused on the performance analysis and the design of network code matrices for cooperative networks with erasure channel models. In particular, in order to maximize diversity order in a multiple-access network, the problem of designing network code matrices subject to link failures was studied [5]–[8], to maximized diversity order in a multiple-access network. In [7], it is shown that the design of network code matrices is equivalent to the design of linear block codes for erasure correction coding. It was shown that maximum diversity order is guaranteed if an MDS code generator matrix of MDS codes is utilized as the network code matrix. In addition, Nguyen *et al.* [11] have defined upper and lower bounds on GDNC scheme recovery performance. For random linear network coding schemes, Trullols-Cruces *et al.* [10] have derived the exact decoding probability of obtaining network codes of full rank.

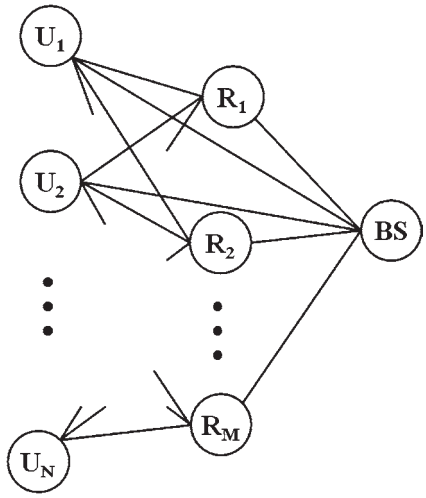


Fig. 1. An (N, M) cooperative network with N sources and M relays.

III. Cooperative Network

We consider an (N, M) cooperative scheme for wireless networks as shown in Fig. 1, in which there are N sources, $\{U_1, U_2, \ldots, U_N\}$, and M relays, $\{R_1, R_2, \ldots, R_M\}$. There are two cooperating transmission phases: *broadcasting* and *relaying*. In the broadcasting phase, each source transmits its message to the base station (BS). Owing to the nature of wireless channels, the relays in this phase can, in general, receive and successfully decode the messages from the sources. In the relay phase, each relay can generate a parity message constructed from a linear combination of these messages, and forward it to the BS. In this work, we assume that the received message for a single channel is considered either completely corrupted—an outage and therefore not available at the receiver—or error-free, i.e., no outage. For more complicated cooperative communications error models that have been studied in [30]–[33], we have included a discussion in Section II.

For both transmission types, we assume that in both transmissions all transmitters send their signals through orthogonal channels using either time- or frequency-division multiple access, and that all channels are spatially independent, flat-faded, and perturbed by additive white Gaussian noise (AWGN). We further assume that the channel gains are independent in both the broadcasting and relay phases. A discussion is included in

Section II in which no orthogonal transmissions are utilized and inter-user channels are not independent.

In the broadcasting phase, the signal y_{u,d_1} received at node d_1 for $d_1 \in \{R_1, R_2, \ldots, R_M, BS\}$ is given by $y_{u,d_1} = \sqrt{P_u} h_{u,d_1} x_{u,d_1} + n_{u,d_1}$, where u denotes the transmitter node, i.e., $u \in \{U_1, U_2, \ldots, U_N\}$; P_u denotes the transmit power at node u; h_{u,d_1} denotes the channel gain between the two nodes u and d_1, which is a circular symmetric complex-valued Gaussian random variable with zero mean and variance $\sigma^2_{u,d_1}/2$ per dimension; x_{u,d_1} is the signal transmitted from node u; and noise n_{u,d_1} denotes the complex-valued AWGN with zero mean and variance $N_0/2$ per dimension. In the relay phase, the signal y_{r,d_2} received at the BS is $y_{r,d_2} = \sqrt{P_r} h_{r,d_2} x_{r,d_2} + n_{r,d_2}$, where r denotes the relay node, i.e., $r \in \{R_1, R_2, \ldots, R_M\}$; d_2 denotes the BS; P_r denotes the transmit power at relay node r; h_{r,d_2} denotes the channel gain between the relay node r and the BS, which is a circular symmetric complex-valued Gaussian random variable with zero mean and variance $\sigma^2_{r,d_2}/2$ per dimension; x_{r,d_2} is the signal transmitted from relay node r; and noise n_{r,d_2} denotes the same AWGN as in the broadcasting phase. For Rayleigh fading channels, the variances of channel gains are defined as $\sigma^2_{u,d_1} := \rho^{-\eta}_{u,d_1}$ and $\sigma^2_{r,d_2} := \rho^{-\eta}_{r,d_2}$, letting ρ_{u,d_1} and ρ_{r,d_2} be the distances for u-to-d_1 and r-to-d_2, respectively, and η be the path-loss exponent, i.e., $2 \le \eta \le 6$ [37]. Throughout this paper, we use $\eta = 3$. The instantaneous SNRs of the two channels are denoted as $\gamma_{u,d_1} := |h_{u,d_1}|^2 P_u / N_0$ and $\gamma_{r,d_2} := |h_{r,d_2}|^2 P_r / N_0$.

Let R_{th} be the predefined threshold of the spectral efficiency in bits/s/Hz. For both phases, we utilize the following outage probabilities based on [1] and [38],

$$\delta_{u,d_1} = \Pr\left\{\log\left(1 + \gamma_{u,d_1}\right) < R_{th}\right\}, \tag{1}$$

and

$$\delta_{r,d_2} = \Pr\left\{\log\left(1 + \gamma_{r,d_2}\right) < R_{th}\right\}. \tag{2}$$

Throughout this paper, we use $R_{th} = 1$ bit/s/Hz. Each of the outage probabilities is a function of the instantaneous SNR and the distance between two nodes. We can use these outage probabilities to model the elements of a transmission matrix.

IV. Modeling of Transmission Matrices

A. Transmission Matrix

We utilize the outage probabilities defined in Section III to model the elements of the transmission matrix. A random transmission matrix can be used to represent a family of network coding matrices for an (N, M) cooperative scheme shown in Fig. 1 in which two transmissions occur over a multiple access network with N sources and M relays. Let $\mathbb{F}_q$ be a finite field of size q. Let $\mathbf{x} \in \mathbb{F}_q^{N\times 1}$ denote the $N \times 1$ vector of transmitted messages, $\mathbf{y} \in \mathbb{F}_q^{(N+M)\times 1}$ denote the $(N+M) \times 1$ vector of messages received at the BS, and $\mathbf{A} \in \mathbb{F}_q^{(N+M)\times N}$ denote the $(N+M) \times N$ transmission matrix. The vector $\mathbf{y}$ received at the BS is then given by

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{3}$$

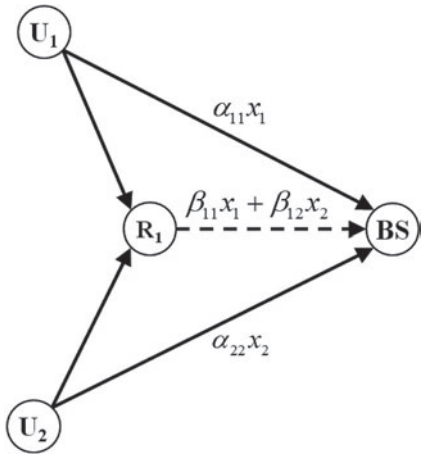


Fig. 2. The (2,1) cooperative wireless network with $N = 2$ and $M = 1$. The solid lines indicate the broadcasting phase, and the dashed line indicates the relay phase.

where the transmission matrix $\mathbf{A}$ consists of the $N \times N$ direct matrix $\mathbf{D}$ and the $M \times N$ combination matrix $\mathbf{P}$; i.e., $\mathbf{A} := \begin{bmatrix}\mathbf{D}\\ \mathbf{P}\end{bmatrix}$. Note that all of the arithmetic operations are performed over finite fields.

The direct matrix $\mathbf{D}$ can be modeled as a diagonal matrix, i.e., one with zeroes for all off-diagonal elements. If there are no outage events for the channel links between the sources and the BS, the diagonal elements of this direct matrix are all set to one; otherwise, the corresponding elements are set to zero. Let α_{ii} denote the ith diagonal element of $\mathbf{D}$, i.e., $\alpha_{ii} \in \{0, 1\}$ for $i \in \{1, 2, \ldots, N\}$, then the ith element $y_{i,1}$ of $\mathbf{y}$ is represented as $y_{i,1} = \alpha_{ii} x_i$, where $x_i \in \mathbb{F}_q$ denotes the ith element of $\mathbf{x}$, and 1 in the subscript of $y_{j,1}$ indicates the broadcasting phase for $j \in \{1, 2, \ldots, M\}$. Let $\beta_{ji} \in \mathbb{F}_q$ denote an element of $\mathbf{P}$, then the $(N+j)$th element $y_{j,2}$ of $\mathbf{y}$ is represented by $y_{j,2} = \sum_{i=1}^{N} \beta_{ji} x_i$, where 2 in the subscript of $y_{j,2}$ indicates the relay phase. As the BS receives $N + M$ messages from N sources and M relays, we can represent the transmission matrix as

$$\begin{bmatrix} y_{1,1} \\ \vdots \\ y_{N,1} \\ y_{1,2} \\ \vdots \\ y_{M,2} \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_{NN} \\ \beta_{11} & \cdots & \beta_{1N} \\ \vdots & \ddots & \vdots \\ \beta_{M1} & \cdots & \beta_{MN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}. \tag{4}$$

Note that all elements of the transmission matrix $\mathbf{A}$ are random variables except for the off—diagonal terms of $\mathbf{D}$. The following simple example illustrates the method for determining the elements of the transmission matrix.

Example 1: Consider two sources (U_1 and U_2) and one relay (R_1) in a (2, 1) cooperative wireless network shown in Fig. 2. Let the size of the finite field for the network coding be 2, $q = 2$. In the broadcasting phase, source U_1 transmits message x_1 to the BS and relay R_1, while U_2 transmits message x_2 to the BS and R_1. The relay overhears, decodes, and then linearly combines the decoded messages to generate a parity message that is forwarded to the BS. Thus, the BS receives three messages: x_1 and x_2 from the respective sources and a

TABLE II
DETERMINATION OF THE TRANSMISSION MATRIX FOR ALL CASES OF FAILURES FOR $q = 2$, WHERE "O" INDICATES NO OUTAGE, "×" INDICATES AN OUTAGE, AND "–" INDICATES DON'T CARE

U_1-BS	U_2-BS	**D**	U_1-R_1	U_2-R_1	R_1-BS	**P**
O	O	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	O	O	O	(1 1)
O	×	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	O	×	O	(1 0)
×	O	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	×	O	O	(0 1)
×	×	$\begin{pmatrix}0&0\\0&0\end{pmatrix}$	×/–	×/–	O/×	(0 0)

parity message from the relay. This transmission mechanism is depicted in Fig. 2.

The transmission matrix in this example is given by

$$\begin{bmatrix} y_{1,1} \\ y_{2,1} \\ y_{1,2} \end{bmatrix} = \begin{bmatrix} \alpha_{11} & 0 \\ 0 & \alpha_{22} \\ \beta_{11} & \beta_{12} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{5}$$

Here, the connectivities of the channel links (U_1-BS and U_2-BS) are represented by α_{11} and α_{22} in the transmission matrix. If a channel link (U_1-BS) or (U_2-BS) incurs an outage, then its associated connectivity, α_{11} or α_{22}, will be set to zero; otherwise, this element is set to one. Similarly, two elements β_{11} and β_{12} represent the joint factors of the qualities of the three channel links: $(U_1 - R_1)$, $(U_2 - R_1)$, and (R_1-BS). If both links, $(U_2 - R_1)$ and (R_1-BS), do not simultaneously incurs outages, β_{12} is set one; conversely, if either of the two links undergoes an outage, β_{11} will be set to zero. Thus, the transmission matrix will be determined by the condition of all five channel links in the wireless network, and if the transmission matrix at a given condition has full rank, the BS can successfully decode the two source messages x_1 and x_2. Table II summarizes all outage events of the transmission matrix for the (2,1) cooperative wireless network in which there are two sources and one relay. ■

B. Modeling of Random Elements

In this subsection, we provide techniques for defining the elements of the transmission matrix as random variables. We assume that all outage events are mutually independent from each other, which is reasonable for typical wireless networks. Then, the probability distribution of the elements can be determined based on the outage probabilities of the wireless channels as follows. First, the probability distribution for each diagonal element of $\mathbf{D}$ is modeled using the outage probability of the source-to-BS channels. Second, by simultaneously considering the outage events in both channels (i.e., source-relay and relay-BS), we determine the probability distribution for each element of $\mathbf{P}$.

To model each diagonal element of $\mathbf{D}$, the probability of the ith diagonal element α_{ii}, $i \in \{1, 2, \ldots, N\}$, can be defined from the set of possible outage events between the sources and the BS in the broadcasting phase as:

$$\Pr\{\alpha_{ii} = \theta\} = \begin{cases} \delta_{U_i,BS} & \text{if } \theta = 0, \\ 1 - \delta_{U_i,BS} & \text{if } \theta = 1, \end{cases} \tag{6}$$

where $\delta_{U_i,BS}$ the outage probability defined in (1) where an outage occurs in the single link between the ith source U_i and the BS.

Next, to model each element of the combination matrix $\mathbf{P}$, we consider two types of probability distributions. The first is to permit the nonzero values of each element in $\mathbf{P}$ to be uniformly distributed. This distribution is reasonable, considering the recent result in [39] where it is acknowledged that a uniform distribution for linear network coding provides various benefits, inlcuding decentralized operation and robustness to network changes or link failures in multisource, multicast networks. The second is to allow the nonzero value to be predetermined. This specific value can be set using MDS codes [40] in coding theory. It is well known that MDS codes achieve the Singleton bounds. This supports the consideration of MDS codes for optimum reconstruction performance. In the latest literature, Rebelatto *et al.* [7] have proved that a systematic MDS code generator matrix, operating over sufficiently large finite fields such as the transmission matrix, is sufficient for obtaining full diversity in cooperative networks. However, MDS codes use a large field size so that may result in excessive complexity, especially in the cases where the dimension of the code is large.

1) Uniform Distribution: When modeling the elements of the combination matrix $\mathbf{P}$, we have to consider the outage events in both the source-to-relay and relay-to-BS links, as the occurrence of either or both of these events will prevent the relay from delivering the source message to the BS. Let $\bar{\mathcal{E}}_j$ and $\mathcal{E}_j$ denote the nonoccurrence and occurrence of an outage from the jth relay R_j, $j \in \{1, 2, \ldots, M\}$, to the BS, respectively. Thus, both probabilities are: $\Pr\{\bar{\mathcal{E}}_j\} = 1 - \delta_{R_j,BS}$ and $\Pr\{\mathcal{E}_j\} = \delta_{R_j,BS}$. Because the outage event from a source to a relay is independent of any other outage events, the conditional probability $\Pr\{\beta_{ji} = \theta|\bar{\mathcal{E}}_j\}$ of this element of the combination matrix $\mathbf{P}$ can be modeled as

$$\Pr\{\beta_{ji} = \theta|\bar{\mathcal{E}}_j\} = \begin{cases} \delta_{U_i,R_j} & \text{if } \theta = 0, \\ \left(1 - \delta_{U_i,R_j}\right)/(q-1) & \text{if } \theta \neq 0, \end{cases} \tag{7}$$

where δ_{U_i,R_j} denotes the probability that the outage occurs from the ith source U_i to the jth relay R_j. Each outage probability δ_{U_i,R_j} can be determined independently in (1).

In (7), the elements of $\mathbf{P}$ are nonzero when both outage events do not occur simultaneously; however, when an outage event from the jth relay to the BS occurs, i.e., when $\mathcal{E}_j$ is true, the conditional probability is set as $\Pr\{\beta_{ji} = 0|\mathcal{E}_j\} = 1$, regardless of the condition of the outage event (source-relay).

2) MDS Distribution: Next, we consider modeling the elements of $\mathbf{P}$ based on the systematic generator matrix of MDS codes. The difference from the aforementioned uniform distribution is that the nonzero value of each element should be taken from the pertinent value of a predefined MDS code. In this subsection, we refer to this as the MDS distribution. By considering the MDS distribution, we can compare its reconstruction performance to that of the uniform distribution given in (7). For the MDS distribution, the conditional probability $\Pr\{\beta_{ji} = \theta|\bar{\mathcal{E}}_j\}$ defined similarly to that in (7) is given by

$$\Pr\{\beta_{ji} = \theta|\bar{\mathcal{E}}_j\} = \begin{cases} \delta_{U_i,R_j} & \text{if } \theta = 0, \\ 1 - \delta_{U_i,R_j} & \text{if } \theta = \chi, \\ 0 & \text{otherwise}, \end{cases} \tag{8}$$

where χ denotes the coefficient that is predefined from the systematic generator matrix of MDS codes. To generate this code, we used the software application SAGE [41]. For $N = 8$ and $M = 4$, for example, the 4×8 submatrix of the systematic MDS code is:

$$\begin{bmatrix} 9 & 13 & 14 & 7 & 2 & 15 & 13 & 12 \\ 15 & 3 & 9 & 12 & 12 & 10 & 12 & 2 \\ 14 & 9 & 12 & 7 & 8 & 1 & 3 & 7 \\ 4 & 5 & 5 & 10 & 9 & 3 & 4 & 1 \end{bmatrix}. \tag{9}$$

In example (9), the conditional probability $\Pr\{\beta_{11} = 9|\bar{\mathcal{E}}_1\}$ is $1 - \delta_{U_1,R_1}$, and for any $\theta \in \mathbb{F}_{16} \setminus \{0, 9\}$, it is set to zero, i.e., $\Pr\{\beta_{11} = \theta|\bar{\mathcal{E}}_1\} = 0$. Based on this, we can investigate the improvement in the reconstruction performance that is achieved when using MDS codes in a cooperative wireless network.

Remark 1: In this work, we assume that all the inter-node channels are independent from each other. Thereby, probability distributions of random elements are defined independently. If channel correlations are considered, the distributions, i.e., (6)–(8), should be modeled as a joint distribution corresponding to the channel correlation. Using joint distributions, we can evaluate the performance of correlated wireless network coding schemes. Example 3 in Section V-B of this paper shows that the proposed framework can be extended to correlation cases. A generalized version of the proposed framework for channel correlations will be another direction of future research.

V. Upper Bound on Reconstruction of Messages

If a transmission matrix for a dynamic network topology randomly generated using the probability distributions given in (6)–(8) has full rank, the BS can uniquely decode all messages from all sources. In this section, we aim to derive an upper bound on the decoding failure probability, and the dimension of the nullspace of the random transmission matrix of an (N, M) cooperative wireless network. We then connect them to investigate the manner in which network coding performance varies based on wireless channel conditions, the number of relays, field sizes, and the positions of nodes deployed in a 2D space. Throughout this paper, we use the random transmission matrix as a bold face, i.e., $\mathbf{A}$, while the realized transmission matrix in sans-serif style, i.e., A.

We define the dimension of the nullspace within the column space of a transmission matrix as follows. Let A be an $(N+M)\times N$ matrix over the finite field with size q as $\mathbb{F}_q$. Based on linear algebra theory, the columns $A_1, \ldots, A_N$ of A are linearly dependent if and only if a vector $\mathbf{c} = (c_1, \ldots, c_N) \in \mathbb{F}_q^N$ such that exists, with at least one nonzero c_i, such that

$$\sum_{i=1}^{N} c_i A_i = 0. \tag{10}$$

Definition 1. (Number of Nonzero Coefficient Vectors): Let $L(\mathsf{A})$ be the number of all such nonzero vectors $\mathbf{c}$ belonging to the nullspace of the given matrix A. Let the column rank of a realized transmission matrix be $\mathsf{rank}(\mathsf{A})$. Thus, $L(\mathsf{A})$ can be represented as

$$L(\mathsf{A}) = q^{N-\mathsf{rank}(\mathsf{A})} - 1. \tag{11}$$

Definition 2. (Nullity): Let $\mathsf{nullity}(\mathsf{A})$ be the dimension of the nullspace in the column space of A.

Proposition 3: For a random matrix $\mathbf{A}$, the expectation of the nullity of $\mathbf{A}$ is upper bounded by $\mathbb{E}[\mathsf{nullity}(\mathbf{A})] \leq \log_q(\mathbb{E}[L(\mathbf{A})] + 1)$, where $\mathbb{E}[\cdot]$ denotes the expectation.

Proof: For any $(N + M) \times N$ matrix A, we follow $\mathsf{nullity}(\mathsf{A}) = N - \mathsf{rank}(\mathsf{A})$, known as the rank-nullity theorem of linear algebra [42]. Considering the expectation for a random transmission matrix in both sides of (11), we obtain the following upper bound using Jensen's inequality:

$$\begin{aligned} \mathbb{E}\left[\mathsf{nullity}(\mathbf{A})\right] &:= N - \mathbb{E}\left[\mathsf{rank}(\mathbf{A})\right] \\ &= \mathbb{E}\left[\log_q\left(L(\mathbf{A}) + 1\right)\right] \\ &\leq \log_q\left(\mathbb{E}\left[L(\mathbf{A})\right] + 1\right). \end{aligned} \tag{12}$$

The proof of Proposition 3 is complete. ■

Theorem 4: Let P_{fail} be the decoding failure probability for the reconstruction of source messages. Then, $P_{fail} \leq \min\{1, 1/(q-1)\mathbb{E}[L(\mathbf{A})]\}$.

Proof: The probability P_{fail} is defined and upper bounded by

$$\begin{aligned} P_{fail} &= \Pr\left\{\mathsf{rank}(\mathbf{A}) < N\right\} \\ &= \Pr\left\{\exists \mathbf{c} : \sum_{i=1}^{N} c_i A_i = 0\right\} \\ &\overset{(a)}{\leq} \sum_{\mathbf{c} \in \mathbb{F}_q^N \setminus \{0^T\}} \Pr\{\mathbf{A}\mathbf{c} = 0^T\} \\ &= \mathbb{E}\left[L(\mathbf{A})\right]. \end{aligned} \tag{13}$$

where inequality (a) is due to the union bound; note that $\mathbb{E}[L(\mathbf{A})] = \sum_{\mathbf{c} \in \mathbb{F}_q^N \setminus \{0^T\}} \Pr\{\mathbf{A}\mathbf{c} = 0^T\}$. Then, the upper bound on the probability P_{fail} can be tightened as

$$P_{fail} \leq \min\left\{1, \frac{1}{q-1}\mathbb{E}\left[L(\mathbf{A})\right]\right\}, \tag{14}$$

where the $1/(q-1)$ factor is due to the following reason. Suppose a nonzero vector $\mathbf{c}$ exists such that $\mathbf{A}\mathbf{c} = 0^T$. Then, other $q-2$ nonzero vectors $\theta\mathbf{c}$, $\theta^2\mathbf{c}, \ldots, \theta^{q-2}\mathbf{c}$ certainly exist for a primitive element $\theta \in \mathbb{F}_q \setminus \{0\}$, where each satisfies $\mathbf{A}\theta^i\mathbf{c} = 0^T$ for $i \in \{1, \ldots, q-2\}$. Then, we note $\bigcup_{\mathbf{c}_1 \in \{\theta\mathbf{c}, \ldots, \theta^{q-2}\mathbf{c}\}} \{\mathbf{A} : \mathbf{A}\mathbf{c}_1 = 0^T\} = \{\mathbf{A} : \mathbf{A}\mathbf{c} = 0^T\}$. ■

Remark 2: Proposition 3 and Theorem 4 provide a groundwork novel performance evaluation framework of cooperative wireless network coding schemes. These results enable us to calculate the decoding failure probability without an exhaustive search of all possible individual cases. They are new key steps that enable the evaluation framework to be computationally efficient, are not available in the literature, for example [5]–[8].

Next, we will derive $\mathbb{E}[L(\mathbf{A})]$ for three types of cooperative wireless networks.

A. Homogeneous and Heterogeneous Connectivity

In this subsection, we aim to find $\mathbb{E}[L(\mathbf{A})]$ for two cases: i) *homogeneous connectivity* in which all outage probabilities in the wireless network are equal; i.e., $\delta = \delta_{U_i,BS} = \delta_{U_i,R_j} = \delta_{R_j,BS}$ given in (1) and (2) for $i \in \{1,2,\ldots,N\}$ and $j \in \{1,2,\ldots,M\}$, assuming that all the channel qualities in networks are equal, and ii) *heterogeneous connectivity* in which two types of outage probabilities exist, i.e., $\delta_1 = \delta_{U_i,BS} = \delta_{U_i,R_j}$ and $\delta_2 = \delta_{R_j,BS}$, with each outage probability assumed to be merely a function of transmit power. Note that each element of a random matrix follows the probability distributions defined in (6) and (7).

Let S_k denote the probability, $S_k := \Pr\{\sum_{i=1}^{k} \beta_{ji} = 0\}$, for the sum of the first k random elements, $k \in \{1,2,\ldots,N\}$, in the jth row of a combination matrix $\mathbf{P}$. For the homogeneous case, Lemma 5 provides this probability S_k.

Lemma 5: For the homogeneous connectivity with the distributions (6) and (7), the probability S_k is given by

$$S_k = \delta + (1-\delta)\left(q^{-1} + (1-q^{-1})\left(1 - \frac{1-\delta}{1-q^{-1}}\right)^k\right). \quad (15)$$

Proof: See Appendix A. ■

Before attempting to derive $\mathbb{E}[L(\mathbf{A})]$ of a random matrix $\mathbf{A}$ from Lemma 5, recall that $L(\mathsf{A})$ is the number of all nonzero vectors $\mathbf{c}$ satisfying the linear dependency in (10). The following Proposition 6 gives $\mathbb{E}[L(\mathbf{A})]$ for the homogeneous (N,M) wireless cooperative network.

Proposition 6: Given an (N,M) cooperative network with the homogeneous connectivity based on some outage probability δ, $\mathbb{E}[L(\mathbf{A})]$ of a $(N+M)\times N$ random transmission matrix $\mathbf{A}$ over the finite field $\mathbb{F}_q$ is

$$\mathbb{E}\left[L(\mathbf{A})\right] = \sum_{k=1}^{N} \binom{N}{k} (q-1)^k \delta^k \left[\delta + (1-\delta) \times \left(q^{-1} + (1-q^{-1})\left(1 - \frac{1-\delta}{1-q^{-1}}\right)^k\right)\right]^M. \quad (16)$$

Proof: See Appendix B. ■

Consider $\mathbb{E}[L(\mathbf{A})]$ under the heterogeneous case; i.e., $\delta_1 = \delta_{U_i,BS} = \delta_{U_i,R_j}$ and $\delta_2 = \delta_{R_j,BS}$. In the following section, we aim to study the manner in which the outage probabilities δ_1 and δ_2 affect the recovery performance, assuming that these outage probabilities rely on the transmit power at the sources and relays.

Proposition 7: Given the heterogeneous (N,M) cooperative network defined by the two outage probabilities δ_1 and δ_2, $\mathbb{E}[L(\mathbf{A})]$ of a $(N+M)\times N$ random transmission matrix $\mathbf{A}$ over finite fields $\mathbb{F}_q$ is given by

$$\mathbb{E}\left[L(\mathbf{A})\right] = \sum_{k=1}^{N} \binom{N}{k} (q-1)^k \delta_1^k \left[\delta_2 + (1-\delta_2) \times \left(q^{-1} + (1-q^{-1})\left(1 - \frac{1-\delta_1}{1-q^{-1}}\right)^k\right)\right]^M. \quad (17)$$

The proof is omitted. However, it can be proved by following the formalism given in Proposition 6, using two outage probabilities, δ_1 and δ_2, instead of single outage probability as in Proposition 6.

B. General Connectivity

Thus far, we have obtained $\mathbb{E}[L(\mathbf{A})]$ of a random transmission matrix $\mathbf{A}$ for homogeneous and heterogeneous cases. In this subsection, we extend it to a more general case where δ_{U_i,R_j}, $\delta_{R_j,BS}$, and $\delta_{U_i,BS}$ are used as defined in Section III. Outage probabilities for the wireless links are obtained using (1) and (2), which are functions of transmit power and variance of the channel gain. After obtaining the outage probability for the each link, we determine the probability distributions for all elements in $\mathbf{A}$. We call this the *general connectivity* case. Since it involves an exhaustive search of combinations of column vectors in a random matrix, the general connectivity requires a more complicated computation than that of the homogeneous and heterogeneous connectivity to derive $\mathbb{E}[L(\mathbf{A})]$ where the proposed approach using the upper bound on the dimension of nullspace will be most effective.

Proposition 8: Given an (N,M) cooperative network with the general connectivity based on the outage probabilities defined in (1) and (2), $\mathbb{E}[L(\mathbf{A})]$ of a $(N+M)\times N$ random transmission matrix $\mathbf{A}$ over finite fields $\mathbb{F}_q$ is

$$\mathbb{E}\left[L(\mathbf{A})\right] = \sum_{k=1}^{N} (q-1)^k Q_k, \quad (18)$$

where $Q_k := \sum_{l=1}^{|\mathcal{L}_k|} Q_{k,l}$, $l \in \{1,2,\ldots,|\mathcal{L}_k|\}$, $|\mathcal{L}_k| := \binom{N}{k}$, and $\mathcal{L}_{k,l}$ is the lth entry of a set $\mathcal{L}_k$. Let $\mathcal{L}_k$ denote the collection of the sets of k distinct indices among $[N] := \{1,2,\ldots,N\}$, i.e., $\mathcal{L}_k := \{\{\lambda_1,\lambda_2,\ldots,\lambda_k\} : \lambda_i \in \{1,2,\ldots,N\}, \lambda_i \neq \lambda_j, i \neq j\}$. $Q_{k,l} := \Pr\{\sum_{i\in\mathcal{L}_{k,l}} c_i A_i = 0\}$.

Proof: See Appendix C. ■

We use Proposition 8 to obtain $\mathbb{E}[L(\mathbf{A})]$ for $q=2$ in a (2, 1) cooperative wireless network as follows.

Example 2: Let us consider a (2, 1) cooperative wireless network for $q=2$, $N=2$, and $M=1$. There are three nonzero vectors $\mathbf{c}$ in $\mathbb{F}_2^2$: (10), (01), and (11). For each nonzero vector, we obtain the probability $Q_{k,l}$ as follows. First, the probability $Q_{1,1}$ is

$$\begin{aligned} Q_{1,1} &= \Pr\{c_1 A_1 = 0\} \\ &= \Pr\{\alpha_{11} = 0\}\Pr\{\beta_{11} = 0\} \\ &= \delta_{U_1,BS}\left(\delta_{R_1,BS} + (1-\delta_{R_1,BS})\,\delta_{U_1,R_1}\right). \end{aligned} \quad (19)$$

The probability $Q_{1,2}$ is

$$\begin{aligned} Q_{1,2} &= \Pr\{c_2 A_2 = 0\} \\ &= \Pr\{\alpha_{22} = 0\}\Pr\{\beta_{12} = 0\} \\ &= \delta_{U_2,BS}\left(\delta_{R_1,BS} + (1-\delta_{R_1,BS})\,\delta_{U_2,R_1}\right). \end{aligned} \quad (20)$$

The probability $Q_{2,1}$ is

$$\begin{aligned} Q_{2,1} &= \Pr\{c_1 A_1 + c_2 A_2 = 0\} \\ &= \Pr\{\alpha_{11} = 0\}\Pr\{\alpha_{22} = 0\}\Pr\{\beta_{11} + \beta_{12} = 0\} \\ &= \delta_{U_1,BS}\delta_{U_2,BS}\left(\delta_{R_1,BS}\right. \\ &\quad \left. + (1-\delta_{R_1,BS})\Pr\{\beta_{11} + \beta_{12} = 0 | \bar{\mathcal{E}}_1\}\right). \end{aligned} \quad (21)$$

In this example, $\mathbb{E}[L(\mathbf{A})]$ is then given by

$$\mathbb{E}\left[L(\mathbf{A})\right] = Q_{1,1} + Q_{1,2} + Q_{2,1}. \tag{22}$$

■

In addition, it would be intriguing to determine if the proposed evaluation framework developed thus far can be extended to cases where the outages between different links are not independent, but correlated. Such cases may occur when the channels between two nodes are not perfectly orthogonal. Then, such a case becomes an interesting problem to show how the proposed evaluation framework can be utilized to compute the decoding failure probability in the correlated link outages cases. In general, this is a difficult task and would require an additional research paper to effectively provide all details. These details will be provided in a future work. In this paper, we aim to show that the framework is extendible to correlated link outage cases.

For this purpose, we again utilize Proposition 8 to compute $\mathbb{E}[L(\mathbf{A})]$ and extend Example 2 for correlated cases. The outage probabilities are not independent from each other. This can be addressed by considering a joint probability distribution for the random matrix. Using the joint probability distribution, we can again compute the last line of (44) in Appendix, instead of the product of probabilities. This is the main change that allows correlated cases to extend the proposed evaluation framework. In Example 3, given the joint probability distributions, we compute $Q_{1,1}$, $Q_{1,2}$, and $Q_{2,1}$ as shown in Example 2.

Example 3: Consider a (2, 1) cooperative wireless network for $q = 2$, $N = 2$, and $M = 1$. There are two sets of channel correlations that are assumed in this example. The first set of correlated channels is between $U_1 - R_1$ and $U_2 - R_1$; the second set of correlated channels is between U_1-BS and U_2-BS. We assume that all other combinations of channels are mutually independent. Note that both sets of correlations occur in the broadcasting phase. A pair of two outage events, U_1-BS and U_2-BS, create a joint probability as $\Pr\{\alpha_{11} = \theta_1, \alpha_{22} = \theta_2\} = \Theta_{\theta_1,\theta_2}$ for each $(\theta_1, \theta_2) \in \mathbb{F}_2^2$, where $\sum_{\theta_1,\theta_2} \Theta_{\theta_1,\theta_2} = 1$. For example, when both channels are simultaneously successful during the broadcasting phase, we can set the particular probability as $\Pr\{\alpha_{11} = 1, \alpha_{22} = 1\} = \Theta_{1,1}$. Similarly, other probabilities can be defined according to the conditions of the two outage events, U_1-BS and U_2-BS. Note that the conditional joint probability is set as $\Pr\{\beta_{11} = 0, \beta_{12} = 0|\mathcal{E}_1\} = 1$ since the two elements, β_{11} and β_{12}, are zero when the channel outage between R_1 and the BS occurs. In addition, a set of the two outage events, $U_1 - R_1$ and $U_2 - R_1$, can determine the values of the two random elements β_{11} and β_{12} once the channel outage between R_1 and BS does not occur, i.e., when $\bar{\mathcal{E}}_1$ is true. In this case, let the conditional joint probability distribution be known acknowledged and given as $\Pr\{\beta_{11} = \gamma_1, \beta_{12} = \gamma_2|\bar{\mathcal{E}}_1\} = \Gamma_{\gamma_1,\gamma_2}$ for each $(\gamma_1, \gamma_2) \in \mathbb{F}_2^2$, where $\sum_{\gamma_1,\gamma_2} \Gamma_{\gamma_1,\gamma_2} = 1$. For $q = 2$, Table III summarizes this conditional joint probability distribution according to the conditions of the two outage events $U_1 - R_1$ and $U_2 - R_1$.

For three nonzero vectors $\mathbf{c}$ in $\mathbb{F}_2^2$, we can again compute $Q_{1,1}$, $Q_{1,2}$, and $Q_{2,1}$. The computation of $Q_{1,1}$ and $Q_{1,2}$ is straightforward because of our assumption that the two sets of channel correlations are independent. The results are

$$Q_{1,1} = (\Theta_{0,0} + \Theta_{0,1})\left(\delta_{R_1,BS} + (1 - \delta_{R_1,BS})\left(\Gamma_{0,0} + \Gamma_{0,1}\right)\right) \tag{23}$$

and

$$Q_{1,1} = (\Theta_{0,0} + \Theta_{1,0})\left(\delta_{R_1,BS} + (1 - \delta_{R_1,BS})\left(\Gamma_{0,0} + \Gamma_{1,0}\right)\right) \tag{24}$$

The computation of $Q_{2,1}$ is given as follows:

$$\begin{aligned}
Q_{2,1} &= \Pr\{c_1 A_1 + c_2 A_2 = 0\} \\
&= \Pr\{\alpha_{11} = 0, \alpha_{22} = 0, \beta_{11} + \beta_{12} = 0\} \\
&= \Pr\{\alpha_{11} = 0, \alpha_{22} = 0, \beta_{11} + \beta_{12} = 0|\mathcal{E}_1\}\Pr\{\mathcal{E}_1\} \\
&\quad + \Pr\{\alpha_{11} = 0, \alpha_{22} = 0, \beta_{11} + \beta_{12} = 0|\bar{\mathcal{E}}_1\}\Pr\{\bar{\mathcal{E}}_1\} \\
&\overset{(a)}{=} \Pr\{\alpha_{11} = 0, \alpha_{22} = 0\} \\
&\quad \times \left(\Pr\{\beta_{11} + \beta_{12} = 0|\mathcal{E}_1\}\Pr\{\mathcal{E}_1\}\right. \\
&\quad\quad \left. + \Pr\{\beta_{11} + \beta_{12} = 0|\bar{\mathcal{E}}_1\}\Pr\{\bar{\mathcal{E}}_1\}\right) \\
&= \Theta_{0,0}\left(\delta_{R_1,BS} + (\Gamma_{0,0} + \Gamma_{1,1})(1 - \delta_{R_1,BS})\right),
\end{aligned} \tag{25}$$

where equality (a) is based on the fact that the relation between the two sets, $(\alpha_{11}, \alpha_{22})$ and (β_{11}, β_{12}), is independent. We finally obtain $\mathbb{E}[L(\mathbf{A})] = Q_{1,1} + Q_{1,2} + Q_{2,1}$ for correlated cases by using the proposed evaluation framework. ■

TABLE III
DETERMINATION OF THE CONDITIONAL JOINT PROBABILITY FOR THE TWO ELEMENTS β_{11} AND β_{12}, WHERE "O" INDICATES NO OUTAGE AND "×" INDICATES AN OUTAGE

(U_1-R_1, U_2-R_1)	(β_{11},β_{12})	$\Pr\{\beta_{11} = \gamma_1, \beta_{12} = \gamma_2 \,\vert\, \bar{\mathcal{E}}_1\}$
(O, O)	(1, 1)	$\Gamma_{1,1}$
(O, ×)	(1, 0)	$\Gamma_{1,0}$
(×, O)	(0, 1)	$\Gamma_{0,1}$
(×, ×)	(0, 0)	$\Gamma_{0,0}$

C. Asymptotic Nullity

In practice, the computation of (18) requires a significant amount of time, because all combinations of column vectors must be collected as the number of sources and relays increases. In larger networks, this process is complicated; therefore, it will be beneficial to reduce the resources required for this computation. In this subsection, we aim to obtain an asymptotic form of (18) for utilization in large-scale networks.

As previously mentioned, the homogeneous connectivity scheme is a specific case among general connectivity schemes. We can exhibit a simple form of $\mathbb{E}[L(\mathbf{A})]$ in terms of Q_k for the homogeneous topology of cooperative networks. Based on this approach, we can obtain an asymptotic result of (18) in general connectivity schemes. Let us consider $\mathbb{E}[L(\mathbf{A})]$ for $q = 2$ in the homogeneous connectivity. Thus, $\mathbb{E}[L(\mathbf{A})] = \sum_{k=1}^{N} Q_k$ in (18). Using (44), Q_1 is given by

$$Q_1 = N\delta S_1. \tag{26}$$

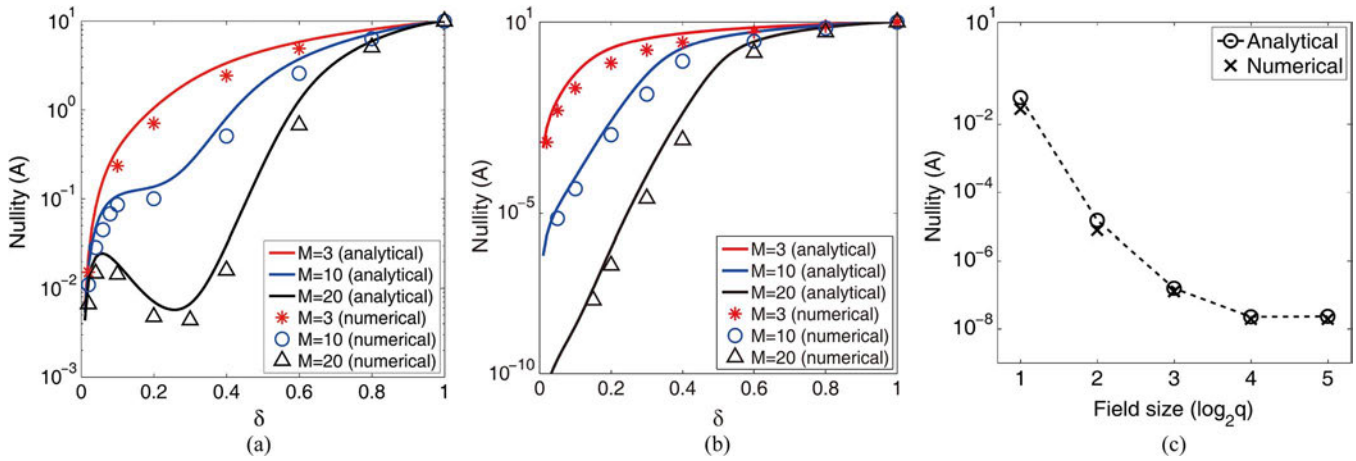


Fig. 3. The nullity of random matrix $\mathbf{A}$ for a homogeneous $(10, M)$ cooperative wireless network with $N=10$ and $M=3, 10,$ and 20. Solid lines indicate the upper bounds on $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ using Proposition 6, and markers indicate numerically simulated results of $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$, respectively: (a) $q=2$ and (b) $q=4$. Fig. 3(c) shows $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ with different field sizes from $q=2$ to 32 under fixed $N=10$, $M=10$, and $\delta=0.05$.

For $k=2$, we have $Q_2=\binom{N}{2}\delta^2 S_2$. We further obtain $Q_3=\binom{N}{3}\delta^3 S_3$ for $k=3$. The general expression of Q_k is given by

$$Q_k=\binom{N}{k}\delta^k S_k. \tag{27}$$

In this case, $\mathbb{E}[L(\mathbf{A})]$ in (18) is

$$\mathbb{E}\left[L(\mathbf{A})\right]=\sum_{k=1}^{N}\binom{N}{k}\delta^k S_k. \tag{28}$$

In high SNR regions, assuming δ is small, an approximation form of (28) is obtained as

$$\begin{aligned}\mathbb{E}\left[L(\mathbf{A})\right]&=\sum_{k=1}^{N}Q_k\\&\overset{(a)}{\approx}\binom{N}{1}\delta S_1+\binom{N}{2}\delta^2 S_2,\end{aligned} \tag{29}$$

where (a) is based on the fact that the order of Q_k for $k\geq 3$ is greater than two with respect to δ. This approximation indicates that for the computation of $\mathbb{E}[L(\mathbf{A})]$, two terms Q_1 and Q_2 are sufficient in high SNR regions. Therefore, in the high SNR regions, $\mathbb{E}[L(\mathbf{A})]$ converges to the second order of the transmit SNR. For any finite field and the general connectivity, this approximation is satisfied.

Corollary 9: Given an (N, M) cooperative network with general connectivity having the distributions (6) and (7), $\mathbb{E}[L(\mathbf{A})]$ is simplified in the high SNR regime

$$\mathbb{E}\left[L(\mathbf{A})\right]\approx(q-1)Q_1+(q-1)^2Q_2. \tag{30}$$

Remark 3: Proposition 8 provides a closed-form solution to the expectation of the number of nonzero vectors in the nullspace of the random transmission matrix $\mathbf{A}$. This enables us to evaluate the performance of a general network with randomly deployed nodes, without separately processing each count of transmission matrix count. Corollary 9 is a closed-form approximation of (18) that is useful for performance evaluations of large size networks.

VI. Numerical and Simulation Results

In this section, the performance of source message reconstruction at the BS is investigated by utilizing the proposed evaluation framework, i.e., $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ and P_{fail}. For the homogeneous connectivity scheme, we employ Proposition 6 to analytically derive the upper bound on $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ defined in Section V as a function of the outage probability of the single channel link. We compare the upper bounds with the numerically simulated results of $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ as well as P_{fail}. Subsequently, we employ Proposition 8 to investigate the upper bounds of a general cooperative network in which sources and relays are deployed in a 2D space. Furthermore, we show the results of the upper bound on $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ and P_{fail} for a given transmission matrix to investigate the impact of the number of relays and the field size of network coding.

Fig. 3 shows analytically obtained upper bounds and numerically averaged results of $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ for a random transmission matrix in a $(10, M)$ cooperative wireless network given the homogeneous connectivity scheme, where $N=10$ and $M=3$, 10, and 20 for $q=2$ in Fig. 3(a), and 4 in Fig. 3(b). We observe that $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ increases as the outage probability slightly increases. Based on Fig. 3(b), it is evident that a nonbinary network coding scheme provides superior reconstruction performance for source messages at the BS when compared to binary coding; moreover, increasing the field size of network coding also improves recovery performance. As the outage probability is reduced to zero, $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ approaches zero for all field sizes.

Fig. 4 shows the analytically derived upper bounds using Proposition 7 for a heterogeneously connected (20, 20) cooperative wireless network. Regardless of the value of δ_2, when the outage probability δ_1 approaches one, $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ barely reaches 20 for both field sizes, i.e., $q=2$ and 4. This indicates that all channel links undergo outage events, causing all elements of the transmission matrix to become zero. In Fig. 4(a), for $q=2$, there is an oscillation in the proximity of $\delta_1=0.3$ such that $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ decreases as δ_1 increases to 0.3, and beyond this point $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ increases. This oscillation also appears in Fig. 3(a). This behavior results from the fact that the rows of $\mathbf{P}$ tend to be identical as the outage probabilities

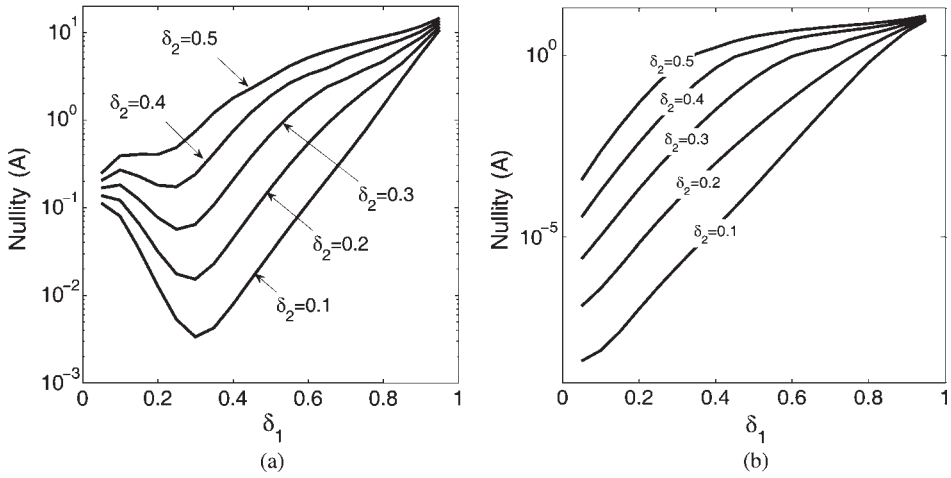


Fig. 4. Upper bounds on $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ using Proposition 3 and 7 in a heterogeneous (20, 20) cooperative wireless network with two outage probabilities δ_1 and δ_2 for (a) $q = 2$, and (b) $q = 4$.

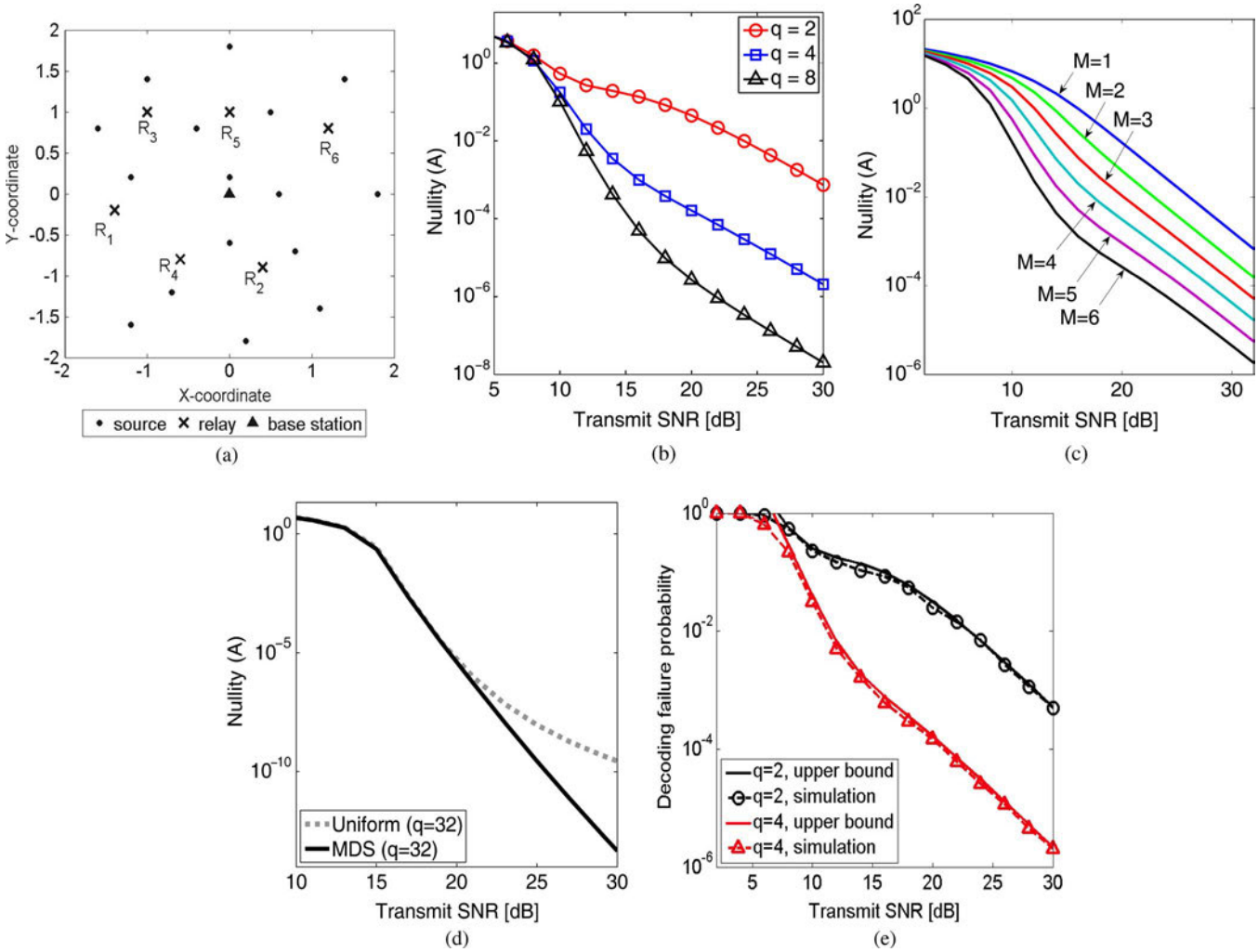


Fig. 5. (a) Location of 16 sources and 6 relays in 2D space for an (16, 6) cooperative wireless network. (b) Results of upper bounds on $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ with differing network coding field sizes $q=2$, 4, and 8, (c) varying the number of relays at $q=4$. (d) Comparison of upper bounds on $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ for the uniform and MDS distributions. (e) Comparison of the decoding failure probabilities with the numerical simulation and the upper bound using Proposition 8 for $q=2$ and 4.

δ_1 and δ_2 approach zero. For $q=4$, however, this behavior disappears owing to the extension of the field size from binary to quaternary.

Now, let us consider the (16, 6) cooperative wireless network shown in Fig. 5(a), in which there are 16 sources and 6 relays: R_1 through R_6. We randomly deploy these relays in a 2D space. We assume that all the transmit powers of sources and relays in both transmission phases are equal. Fig. 5(b) shows the upper bound on $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ of the transmission matrix for $q = 2, 4,$ and 8; the benefit of increasing the field size of

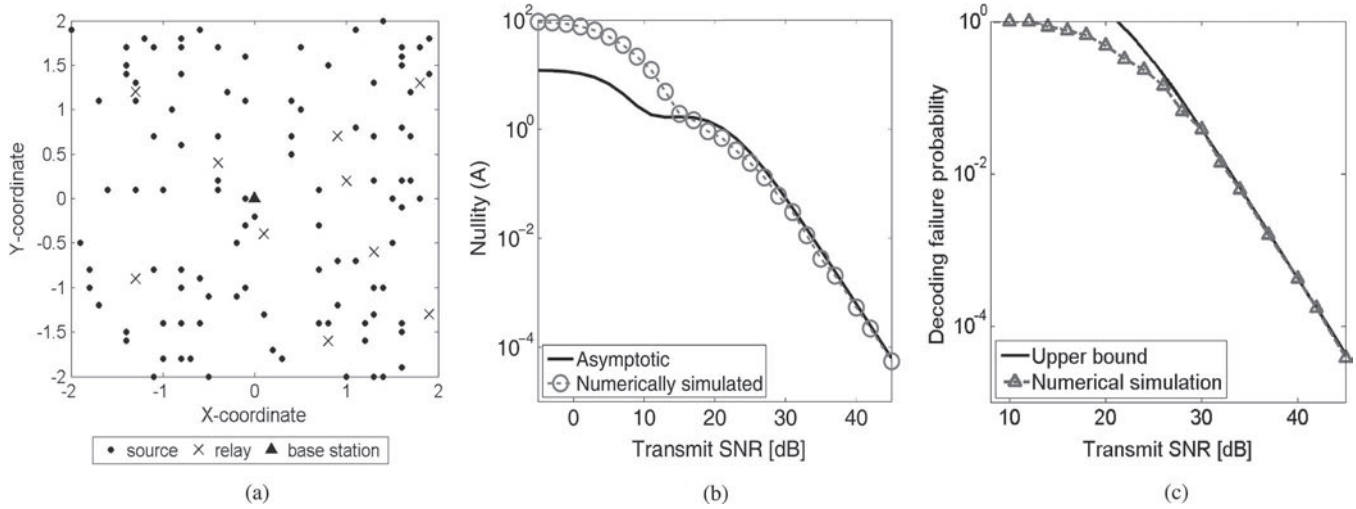


Fig. 6. (a) Locations of 100 sources and 10 relays in a 2D space for an (100, 10) cooperative wireless network, (b) Comparison of $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ with numerically simulated result and the upper bound using Corollary 9 with $q = 2$ and the uniform distribution. (c) Comparison of the decoding failure probability with the numerical simulation and the upper bound using Corollary 9.

network coding appears in this scheme. Fig. 5(c) shows the upper bound on $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ with respect to the number of relays. When $M = 1$, R_1 is used, while R_1 and R_2 are used as relays for $M = 2$, and relays R_1, R_2, and R_3 are used for $M = 3$. For $M = 4, 5,$ and 6, we deploy one relay in order. We investigate the impact of the number of relays, as shown in Fig. 5(c), where increasing the number of relays contributes to the increasingly high likelihood of deriving random transmission matrices of full rank.

Other intriguing results indicate that the value of $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ differs slightly for the uniform and MDS distributions of the combination matrix defined in Section IV-B; furthermore, the recovery performance obtained using the MDS distribution is superior to that of the uniform distribution in high SNR regions. Comparative results of $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ for the two cooperative networks are shown in Fig. 5(d). We observe that there are minimal differences between the uniform and MDS distributions for the recovery performance in the low SNR regions. In particular, the benefit of using the systematic generator of MDS codes appears only in the high SNR regions.

To validate the usefulness of our asymptotic nullity, we consider an (100, 10) cooperative wireless network as a large-scale network in which 100 sources and 10 relays are deployed in the 2D space shown in Fig. 6(a). For Corollary 9, we show that in the high SNR regions, the asymptotic nullity of (30) is similar to the numerical results that were obtained from randomly generated transmission matrices. A comparison of those results is shown in Fig. 6(b). Using the asymptotic nullity, the complexity of (18) can be significantly reduced. In addition, the nullity of the random transmission matrix can be obtained efficiently. Our proposed framework provides the ability to evaluate reconstruction performance in large-scale networks.

Fig. 5(e) shows the comparison of numerically simulated decoding failure probabilities and upper bounds using (14) for an (16, 6) cooperative wireless network with $q = 2$ and 4. The gap between the results is evident in small SNR regions. However, the upper bound on the decoding failure probability is tight in high SNR regions. This behavior is shown in Fig. 6(c) in which the upper bound is obtained from the approximation form of $\mathbb{E}[\mathsf{nullity}(\mathbf{A})]$ in (30). Based on those results, we show that predicting the performance of source message reconstruction for an (N, M) cooperative wireless network is straightforwardly possible in large-scale networks.

VII. Conclusion

In this paper, we considered a cooperative wireless network where N sources are assisted with M relays in two phase transmissions. Our main goal was to propose a new performance analysis framework for evaluating the reconstruction performance of source messages at the BS. To handle dynamic network topologies, we modeled the elements of the transmission matrix as random variables. This enabled us to develop a systematic approach, to avoid the exhaustive evaluations used in DNC and GDNC schemes [5]–[8]. To complete the performance evaluation, we derived two tight upper bounds on the expected dimension of the nullspace of the random transmission matrix, as well as the decoding failure probability. The result is a framework that is more effective than the rank-based method proposed in the previous literature.

Three types of connectivity schemes are considered in this paper, as they make the framework to be general and scalable. They enabled us to show the reconstruction performance of our proposed framework using multiple sources and multiple relays randomly deployed in a 2D space. In addition, it enables us to investigate the impact of the number of relays and the field size of network coding on the system performance; an example is shown in Figs. 5 and 6 for example. In particular, the ability to generate a precise prediction of network coding performance for a network with a large number of sources and relays is a significant benefit. We can formulate challenging scenarios and generate accurate response in an efficient manner, without resorting to extensive computer simulations. For example, we can determine the advantage that using an MDS code, rather than random code, provides when designing the transmission matrices, as relays are added and field sizes are increased; we

can also determine how the position of relays and sources, with respect to the base station locations, affect the performance of cooperative communications. These questions are important engineering inquires in terms of wireless networks design. These questions, which could not be readily answered in the past, but can now be answered precisely using the proposed framework describe in this paper. In addition, we demonstrate that the proposed framework can be extended to channel correlation cases; however, it is necessary to generalize for any categorization of cooperative wireless network coding schemes.

APPENDIX A
PROOF OF LEMMA 5

For the homogeneous connectivity, we use $\delta = \delta_{U_i,BS} = \delta_{U_i,R_j} = \delta_{R_j,BS}$. The conditional probability $\Pr\{\beta_{ji} = \theta|\bar{\mathcal{E}}_j\}$ of each element β_{ji} is defined using (7), and the other conditional probability can be set as $\Pr\{\beta_{ji} = 0|\mathcal{E}_j\} = 1$. Using the total probability theorem, the probability $\Pr\{\sum_{i=1}^{k}\beta_{ji} = 0\}$ can be decomposed by the condition of the outage event $\mathcal{E}_j$, and then given as follows:

$$\begin{aligned} S_k &= \Pr\{\mathcal{E}_j\}\Pr\left\{\sum_{i=1}^{k}\beta_{ji} = 0\middle|\mathcal{E}_j\right\} \\ &\quad + \Pr\{\bar{\mathcal{E}}_j\}\Pr\left\{\sum_{i=1}^{k}\beta_{ji} = 0\middle|\bar{\mathcal{E}}_j\right\} \\ &\overset{(a)}{=} \delta + (1-\delta)\Pr\left\{\sum_{i=1}^{k}\beta_{ji} = 0\middle|\bar{\mathcal{E}}_j\right\}, \end{aligned} \tag{31}$$

where (a) follows from the fact that $\Pr\{\sum_{i=1}^{k}\beta_{ji} = 0|\mathcal{E}_j\} = 1$, as $\Pr\{\beta_{ji} = 0|\mathcal{E}_j\} = 1$ (note that the conditional probability $\Pr\{\beta_{ji} = \theta|\bar{\mathcal{E}}_j\}$ is independent of this). Let f_k be the probability, i.e., $f_k := \Pr\{\sum_{i=1}^{k}\beta_{ji} = 0|\bar{\mathcal{E}}_j\}$. Given the conditional probability defined in (7) and $f_0 = 1$, the probability f_k can be rewritten by [43]

$$\begin{aligned} f_k &= \Pr\left\{\sum_{i=1}^{k-1}\beta_{ji} = 0\middle|\bar{\mathcal{E}}_j\right\}\Pr\{\beta_{jk} = 0|\bar{\mathcal{E}}_j\} \\ &\quad + \sum_{\theta\in\mathbb{F}_q\setminus\{0\}}\Pr\left\{\sum_{i=1}^{k-1}\beta_{ji} = \theta\middle|\bar{\mathcal{E}}_j\right\}\Pr\{\beta_{jk} = -\theta|\bar{\mathcal{E}}_j\} \\ &= f_{k-1}\delta + (1-f_{k-1})\frac{1-\delta}{q-1}. \end{aligned} \tag{32}$$

Let $g_k := f_k - q^{-1}$. By rewriting (32) as a function of g_k, we have a simple closed form:

$$g_k = g_{k-1}\left(1 - \frac{1-\delta}{1-q^{-1}}\right). \tag{33}$$

Applying a geometric series to (33), we obtain f_k as

$$f_k = q^{-1} + (1-q^{-1})\left(1 - \frac{1-\delta}{1-q^{-1}}\right)^k. \tag{34}$$

Finally, the probability S_k can be obtained by substituting (34) into (31) as:

$$S_k = \delta + (1-\delta)\left(q^{-1} + (1-q^{-1})\left(1 - \frac{1-\delta}{1-q^{-1}}\right)^k\right). \tag{35}$$

APPENDIX B
PROOF OF PROPOSITION 6

Let us consider a vector $\mathbf{c} = (c_1, \ldots, c_N) \in \mathbb{F}_q^N$ in which the first k entries (and only the first k entries) are nonzero, i.e., $\mathbf{c} = (c_1, \ldots, c_k, 0, \ldots, 0)$. Let P_k be the probability that the sum of the first k column vectors is zero, i.e., $P_k := \Pr\{\sum_{i=1}^{k} c_i A_i = 0\}$. As (13), $\mathbb{E}[L(\mathbf{A})]$ is given by

$$\begin{aligned} \mathbb{E}\left[L(\mathbf{A})\right] &= \sum_{\mathbf{c}\in\mathbb{F}_q^N\setminus\{0^T\}}\Pr\{\mathbf{A}\mathbf{c} = 0^T\} \\ &= \sum_{k=1}^{N}\binom{N}{k}(q-1)^k P_k. \end{aligned} \tag{36}$$

Since all links in the wireless network are assumed to be spatially and temporally independent, the rows of the transmission matrix are also independent. Thus, P_k is given by

$$\begin{aligned} P_k &= \Pr\left\{\sum_{i=1}^{k} c_i A_i = 0\right\} \\ &= \prod_{i=1}^{k}\Pr\{c_i\alpha_{ii} = 0\}\prod_{j=1}^{M}\Pr\left\{\sum_{i=1}^{k} c_i\beta_{ji} = 0\right\}. \end{aligned} \tag{37}$$

Let H_k be the probability as $H_k := \Pr\{\sum_{i=1}^{k} c_i\beta_{ji} = 0\}$. For $k = 1$, it is easy to show $\Pr\{c_1\beta_{j1} = 0\} = \Pr\{\beta_{j1} = 0\}$ for $c_1 \in \mathbb{F}_q \setminus \{0\}$ because of the multiplication property in finite fields. Next, we prove that $H_k = S_k$ for $k \geq 2$ where $c_1, c_2, \ldots, c_k \in \mathbb{F}_q \setminus \{0\}$ denote the k nonzero elements. The probability H_k is represented by

$$\begin{aligned} H_k &= \sum_{\theta\in\mathbb{F}_q}\Pr\left\{\sum_{i=1}^{k-1} c_i\beta_{ji} = \theta, c_k\beta_{jk} = -\theta\right\} \\ &= \Pr\left\{\sum_{i=1}^{k-1} c_i\beta_{ji} = 0, c_k\beta_{jk} = 0\right\} \\ &\quad + \sum_{\theta\in\mathbb{F}_q\setminus\{0\}}\Pr\left\{\sum_{i=1}^{k-1} c_i\beta_{ji} = \theta, c_k\beta_{jk} = -\theta\right\}. \end{aligned} \tag{38}$$

Decomposing the outage event $\mathcal{E}_j$, (38) can be rewritten by

$$\begin{aligned} H_k &= \Pr\left\{\sum_{i=1}^{k-1} c_i\beta_{ji} = 0, c_k\beta_{jk} = 0\middle|\bar{\mathcal{E}}_j\right\}\Pr\{\bar{\mathcal{E}}_j\} \\ &\quad + \Pr\left\{\sum_{i=1}^{k-1} c_i\beta_{ji} = 0, c_k\beta_{jk} = 0\middle|\mathcal{E}_j\right\}\Pr\{\mathcal{E}_j\} \\ &\quad + \sum_{\theta\in\mathbb{F}_q\setminus\{0\}}\left(\Pr\left\{\sum_{i=1}^{k-1} c_i\beta_{ji} = \theta, c_k\beta_{jk} = -\theta\middle|\bar{\mathcal{E}}_j\right\}\Pr\{\bar{\mathcal{E}}_j\}\right. \\ &\quad \left. + \Pr\left\{\sum_{i=1}^{k-1} c_i\beta_{ji} = \theta, c_k\beta_{jk} = -\theta\middle|\mathcal{E}_j\right\}\Pr\{\mathcal{E}_j\}\right). \end{aligned} \tag{39}$$

Since $\Pr\{\beta_{ji} = 0|\mathcal{E}_j\} = 1$, (39) can be represented by

$$\begin{aligned} H_k &= \Pr\left\{\sum_{i=1}^{k-1} c_i\beta_{ji} = 0, c_k\beta_{jk} = 0\middle|\bar{\mathcal{E}}_j\right\}\Pr\{\bar{\mathcal{E}}_j\} + \Pr\{\mathcal{E}_j\} \\ &\quad + \sum_{\theta\in\mathbb{F}_q\setminus\{0\}}\Pr\left\{\sum_{i=1}^{k-1} c_i\beta_{ji} = \theta, c_k\beta_{jk} = -\theta\middle|\bar{\mathcal{E}}_j\right\}\Pr\{\bar{\mathcal{E}}_j\}. \end{aligned} \tag{40}$$

Noting that wireless channels are independent from each other under the condition of $\bar{\mathcal{E}}_j$, (40) can be decomposed by

$$\begin{aligned} H_k &= \Pr\left\{\sum_{i=1}^{k-1} c_i\beta_{ji}=0 \middle| \bar{\mathcal{E}}_j\right\}\Pr\left\{c_k\beta_{jk}=0|\bar{\mathcal{E}}_j\right\}\Pr\{\bar{\mathcal{E}}_j\} \\ &\quad + \Pr\{\mathcal{E}_j\} + \sum_{\theta\in\mathbb{F}_q\setminus\{0\}} \Pr\left\{\sum_{i=1}^{k-1} c_i\beta_{ji}=\theta \middle| \bar{\mathcal{E}}_j\right\} \\ &\quad \times \Pr\left\{c_k\beta_{jk}=-\theta|\bar{\mathcal{E}}_j\right\}\Pr\{\bar{\mathcal{E}}_j\}. \end{aligned} \tag{41}$$

Using recursion, the probability H_k is given by

$$\begin{aligned} H_k &= \Pr\left\{\sum_{i=1}^{k-1} \beta_{ji}=0 \middle| \bar{\mathcal{E}}_j\right\}\Pr\{\beta_{jk}=0|\bar{\mathcal{E}}_j\}\Pr\{\bar{\mathcal{E}}_j\} \\ &\quad + \Pr\{\mathcal{E}_j\} + \sum_{\theta\in\mathbb{F}_q\setminus\{0\}} \Pr\left\{\sum_{i=1}^{k-1} \beta_{ji}=\theta \middle| \bar{\mathcal{E}}_j\right\} \\ &\quad \times \Pr\{\beta_{jk}=-\theta|\bar{\mathcal{E}}_j\}\Pr\{\bar{\mathcal{E}}_j\} \\ &= \sum_{\theta\in\mathbb{F}_q} \Pr\left\{\sum_{i=1}^{k-1}\beta_{ji}=\theta, \beta_{jk}=-\theta\right\} \\ &= \Pr\left\{\sum_{i=1}^{k}\beta_{ji}=0\right\}. \end{aligned} \tag{42}$$

Thus, we simply rewrite P_k as follows

$$P_k = \delta^k S_k^M. \tag{43}$$

The proof of Proposition 6 is complete.

Appendix C
Proof of Proposition 8

For general connectivity, each element of $\mathbf{A}$ have a different the probability distribution. This result in different probabilities $Q_{k,l}$ for that any k column vectors of $\mathbf{A}$ that are linearly dependent. The total number of $Q_{k,l}$ is $|\mathcal{L}_k| := \binom{N}{k}$. We have to consider all different probabilities $Q_{k,l}$ with respect to all sets $\mathcal{L}_{k,l}$. The probability Q_k should be summed over all different probabilities $Q_{k,l}$, i.e., $Q_k := \sum_{l=1}^{|\mathcal{L}_k|} Q_{k,l}$ for $l \in \{1,2,\ldots,|\mathcal{L}_k|\}$ and $k \in \{1,2,\ldots,N\}$. Thus, all $Q_{k,l}$ are enumerated and collected to obtain the probability Q_k, which is derived as follows:

$$\begin{aligned} Q_k &= \sum_{l=1}^{|\mathcal{L}_k|} Q_{k,l} = \sum_{l=1}^{|\mathcal{L}_k|} \Pr\left\{\sum_{i\in\mathcal{L}_{k,l}} c_i A_i = 0\right\} \\ &= \sum_{l=1}^{|\mathcal{L}_k|}\prod_{m=1}^{k}\Pr\{\alpha_{l_m l_m}=0\}\prod_{j=1}^{M}\Pr\left\{\sum_{m=1}^{k}\beta_{jl_m}=0\right\}, \end{aligned} \tag{44}$$

where l_m is the mth entry of the set $\mathcal{L}_{k,l}$, $m \in \{1,2,\ldots,k\}$. As similarly obtained in (31), (44) can be rewritten as

$$\begin{aligned} Q_k = \sum_{l=1}^{|\mathcal{L}_k|}\prod_{m=1}^{k}\Pr\{\alpha_{l_m l_m}=0\}\prod_{j=1}^{M}\Bigg(&\delta_{R_j,BS} \\ &+ \left(1-\delta_{R_j,BS}\right)\Pr\left\{\sum_{m=1}^{k}\beta_{jl_m}=0 \middle| \bar{\mathcal{E}}_j\right\}\Bigg). \end{aligned} \tag{45}$$

In order to determine $\mathbb{E}[L(\mathbf{A})]$, we count the number of vectors $\mathbf{c}$ having the first k nonzero elements, i.e., $(q-1)^k$. Finally, we perform the summation over all k, and obtain $\mathbb{E}[L(\mathbf{A})]$ as follows,

$$\mathbb{E}\left[L(\mathbf{A})\right] = \sum_{k=1}^{N}(q-1)^k Q_k. \tag{46}$$

The proof of Proposition 8 is complete.

References

[1] J. L. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[2] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[3] Y. Chen, S. Kishore, and J. Li, "Wireless diversity through network coding," in *Proc. IEEE WCNC*, Las Vegas, NV, USA, Apr. 2006, pp. 1681–1686.

[4] D. H. Woldegebreal and H. Karl, "Network-coding based adaptive decode and forward cooperative transmission in a wireless network: Outage analysis," in *Proc. 13th Eur. Wireless Conf.*, Paris, France, Apr. 2007, pp. 1–6.

[5] M. Xiao and M. Skoglund, "Multiple-user cooperative communications based on linear network coding," *IEEE Trans. Commun.*, vol. 58, no. 12, pp. 3345–3351, Dec. 2010.

[6] M. Xiao and M. Skoglund, "Design of network codes for multiple-user multiple-relay wireless networks," in *Proc. IEEE ISIT*, Seoul, Korea, Jun. 2009, pp. 2562–2566.

[7] J. L. Rebelatto, B. F. Uchoa-Filho, Y. Li, and B. Vucetic, "Multiuser cooperative diversity through network coding based on classical coding theory," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 916–926, Feb. 2012.

[8] J. L. Rebelatto, B. F. Uchoa-Filho, Y. Li, and B. Vucetic, "Adaptive distributed network-channel coding," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 2818–2822, Sep. 2011.

[9] H. Topakkaya and Z. Wang, "Wireless network code design and performance analysis using diversity-multiplexing tradeoff," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 488–496, Feb. 2011.

[10] O. Trullols-Cruces, J. M. Barcelo-Ordinas, and M. Fiore, "Exact decoding probability under random linear network coding," *IEEE Commun. Lett.*, vol. 15, no. 1, pp. 67–69, Jan. 2011.

[11] H. V. Nguyen, S. X. Ng, and L. Hanzo, "Performance bounds of network coding aided cooperative multiuser systems," *IEEE Signal Process. Lett.*, vol. 18, no. 7, pp. 435–438, Jul. 2011.

[12] J.-T. Seong and H.-N. Lee, "4-ary network coding for two nodes in cooperative wireless networks: Exact outage probability and coverage expansion," *EURASIP J. Wireless Commun. Netw.*, vol. 2012, p. 366, Dec. 2012.

[13] T. Wang and G. B. Giannakis, "Complex field network coding for multiuser cooperative communications," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 3, pp. 561–571, Apr. 2008.

[14] T. Islam, A. Nasri, R. Schober, R. K. Mallik, and V. K. Bhargava, "Network coded multi-source cooperative communication in BICM-OFDM networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3180–3193, Sep. 2012.

[15] X.-T. Vu, M. D. Renzo, and P. Duhamel, "BER analysis of joint network/channel decoding in block Rayleigh fading channels," in *Proc. IEEE Int. Symp. PIMRC*, London, U.K., Sep. 2013, pp. 698–702.

[16] R. Youssef and A. Graell i Amat, "Distributed serially concatenated codes for multi-source cooperative relay networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 253–263, Jan. 2011.

[17] Z. Han, X. Zhang, and H. V. Poor, "High performance cooperative transmission protocols based on multiuser detection and network coding," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2352–2361, May 2009.

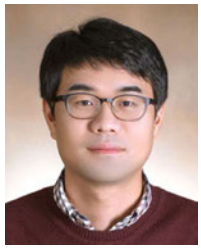
[18] T.-W. Yune, D. Kim, and G.-H. Im, "Iterative detection for spectral efficient user cooperative transmissions over multipath fading channels," *IEEE Trans. Commun.*, vol. 58, no. 4, pp. 1121–1128, Apr. 2010.

[19] S. Katti *et al.*, "XORs in the air: Practical wireless network coding," *IEEE/ACM Trans. Netw.*, vol. 16, no. 3, pp. 497–510, Jun. 2008.

[20] A. Argyriou, "Wireless network coding with improved opportunistic listening," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 2014–2023, Apr. 2009.

[21] J. Zhang, Y. P. Chen, and I. Marsic, "MAC-layer proactive mixing for network coding in multi-hop wireless networks," *Comput. Netw.*, vol. 54, no. 2, pp. 196–207, Feb. 2010.

[22] A. Antonopoulos, C. Verikoukis, C. Skianis, and O. B. Akan, "Energy efficient network coding-based MAC for cooperative ARQ wireless networks," *Ad Hoc Netw.*, vol. 11, no. 1, pp. 190–200, Jan. 2013.
[23] X. Wang, J. Li, and F. Tang, "Network coding aware cooperative MAC protocol for wireless ad hoc networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 1, pp. 167–179, Jan. 2014.
[24] M. H. Firooz, Z. Chen, S. Roy, and H. Liu, "Wireless network coding via modified 802.11 MAC/PHY: Design and implementation on SDR," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1618–1628, Aug. 2013.
[25] S. Wang, Q. Song, X. Wang, and A. Jamalipour, "Distributed MAC protocol supporting physical-layer network coding," *IEEE Trans. Mobile Comput.*, vol. 12, no. 5, pp. 1023–1036, May 2013.
[26] J. Li, J. Yuan, R. Malaney, M. Azmi, and M. Xiao, "Network coding based LDPC code design for a multi-source relaying system," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1538–1551, May 2011.
[27] C. Wang, M. Xiao, and M. Skoglund, "Diversity-multiplexing tradeoff analysis of coded multi-user relay networks," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 1995–2005, Jul. 2011.
[28] M. Xiao and T. Aulin, "On the bit error probability of noisy channel networks with intermediate node encoding," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5188–5198, Nov. 2008.
[29] M. Xiao and T. Aulin, "Optimal decoding and performance analysis of a noisy channel network with network coding," *IEEE Trans. Commun.*, vol. 57, no. 5, pp. 1402–1412, May 2009.
[30] H.-T. Lin, Y.-Y. Lin, and H.-J. Kang, "Adaptive network coding for broadband wireless access networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 1, pp. 4–18, 2013.
[31] M. D. Renzo, M. Iezzi, and F. Graziosi, "Error performance and diversity analysis of multi-source multi-relay wireless networks with binary network coding and cooperative MRC," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2883–2903, 2013.
[32] M. D. Renzo, M. Iezzi, and F. Graziosi, "On diversity order and coding gain of multisource multirelay cooperative wireless networks with binary network coding," *IEEE Trans. Veh. Technol.*, vol. 62, no. 3, pp. 1138–1157, Mar. 2013.
[33] J. Li, J. Yuan, R. Malaney, M. Xiao, and W. Chen, "Full-diversity binary frame-wise network coding for multiple-source multiple-relay networks over slow-fading channels," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1346–1360, Mar. 2012.
[34] B. Nazer and M. Gastp, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.
[35] S. Borade, L. Zheng, and R. Gallager, "Amplify-and-forward in wireless relay networks: Rate, diversity, and network size," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3302–3318, Oct. 2007.
[36] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.
[37] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.
[38] J. N. Laneman and G. W. Wornell, "Distributed space–time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.
[39] T. Ho *et al.*, "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.
[40] F. J. Macwilliams and N. J. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North Holland, 1977.
[41] SAGE, Open Source Mathematics Software. [Online]. Available: http://www.sagemath.org
[42] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA, USA: SIAM, 2000.
[43] J.-T. Seong and H.-N. Lee, "Necessary and sufficient conditions for recovery of sparse signals over finite fields," *IEEE Commun. Lett.*, vol. 17, no. 10, pp. 1976–1979, Oct. 2013.



Jin-Taek Seong received the B.S. degree from the University of Seoul, Seoul, Korea, in 2006 and the M.S. degree from Gwangju Institute of Science and Technology (GIST), Gwangju, Korea, in 2008. He is currently working toward the Ph.D. degree at the School of Information and Communications, GIST. He worked for LG Electronics as a Junior Engineer from 2008 until he started his Ph.D. program in 2010. His research interests include cooperative communications, compressed sensing, and network coding.



Heung-No Lee (SM'13) received the B.S., M.S., and Ph.D. degrees from the University of California, Los Angeles, CA, USA, in 1993, 1994, and 1999, respectively, all in electrical engineering. He then moved to HRL Laboratories, LLC, Malibu, CA, USA, and worked there as a Research Staff Member from 1999 to 2002. In 2002, he was appointed an Assistant Professor with the University of Pittsburgh, Pittsburgh, PA, USA, where he stayed until 2008. In 2009, he then moved to the School of Information and Communications, Gwangju Institute of Science and Technology, Gwangju, Korea, where he is currently affiliated. His general areas of research include information theory, signal processing theory, communications/networking theory, and their application to wireless communications and networking, compressive sensing, future internet, and brain–computer interface. He has served as a member of technical program committees for several IEEE conferences, including the IEEE International Conference on Communications and the IEEE Global Communications Conference (Globecom). He served as the Lead Guest Editor for the European Association for Signal Processing *Journal on Wireless Communications and Networking* in 2010–2011. He has served as an Area Editor for the AEU International *Journal of Electronics and Communications* since January 2013. His research efforts have been recognized with prestigious national awards, including the Top 100 National Research and Development Award from the Korean Ministry of Science, ICT, and Future Planning in 2012, the Top 50 Achievements of Fundamental Researches Award from the National Research Foundation of Korea in 2013, and the Science/Engineer of the Month (January 2014) from the National Research Foundation of Korea.

Computation of an Equilibrium in Spectrum Markets for Cognitive Radio Networks

Sang-Seon Byun, *Member*, *IEEE*, Ilangko Balashingham, *Senior Member*, *IEEE*, Athanasios V. Vasilakos, *Senior Member*, *IEEE*, and Heung-No Lee, *Member*, *IEEE*

Abstract—In this paper, we investigate a market equilibrium in multichannel sharing cognitive radio networks (CRNs): it is assumed that every subchannel is orthogonally licensed to a single primary user (PU), and can be shared with multiple secondary users (SUs). We model this sharing as a spectrum market where PUs offer SUs their subchannels with limiting the interference from SUs; the SUs purchase the right to transmit over the subchannels while observing the inference limits set by the PUs and their budget constraints. Moreover, we consider each SU limits the total interference that can be invoked from all other SUs, and assume that every transmitting SU marks the interference charges to other transmitting SUs. The utility function of SU is defined as least achievable transmission rate, and that of PU is given by the net profit. We define a market equilibrium in the context of extended Fisher model, and show that the equilibrium is yielded by solving an optimization problem, Eisenberg-Gale convex program. To make the solutions of the convex program meet the market equilibrium, we apply monotone-transformation to the utility function of each SU. Furthermore, we develop a distributed algorithm that yields the stationary solutions asymptotically equivalent to the solutions given by the convex program.

Index Terms—Cognitive radio networks, Fisher model, market equilibrium, Eisenberg-Gale convex program, distributed algorithm

1 Introduction

In general markets, when the demand and supply depend upon current price only, there exists an equilibrium under certain conditions, called market equilibrium, where 1) all the traders (i.e., suppliers and consumers) can achieve maximum utilities at least in the Pareto sense; 2) the total demand for each commodity is equal to the total supply of that commodity; and 3) all the budgets possessed by the consumers are spent completely [1].

Recently, market-based approaches have started being deployed to various cognitive radio network (CRN) scenarios since the behaviors of the wireless users in CRN can be cast easily into those of the traders in the general market. Furthermore, the market equilibria comply with the key requirement of the CRN, i.e., spectral efficiency; therefore, lately the US Federal Communications Commission (FCC) has employed policies and procedures to bring spectrum trading into CRN, and analogous regulatory efforts are commenced by EU [2], [3], [4].

In the market-based approaches for CRNs, spectra and interference are regarded as marketable products. In words, primary users (PUs) offer the interference on their licensed spectra to transmitting secondary users (SUs)[1] with collecting certain monetary rewards from SUs; SUs purchase the offered interference on each spectrum by adjusting their transmission powers. Furthermore, every PU has a limitation in her production (i.e., spectra and interference); meanwhile, every SU has limited budget.

In this work, we consider a multichannel sharing CRN where the frequency range is divided into multiple subchannels, and each subchannel is orthogonally licensed to a single PU. Each PU offers the interference on her subchannels to SUs with limiting the interference from the SUs; the SUs purchase the offered interference observing their budget constraints and the interference limits set by the PUs. Moreover, we reflect the interference among SUs to the spectrum market: every SU limits the total interference from other SUs, and every transmitting SU is required to pay charges for interfering with other SUs. The utility function of SU is defined as the least achievable transmission rate, and the utility of PU is given by the net profit the PU makes.

During the history of market theory, there have been lots of market models developed in accordance with various market scenarios: for instance, the *Cournot model* for oligopoly, the *Stackelberg model* considering leadership in the production, the *Bertrand model* for describing the fierce competition among the sellers, and the *Edgeworth model* considering a capacity-limited market.[2] In our work, we apply the market model developed in [6] where the authors have extended *Fisher model* for a market consisting of multisellers as well as multibuyers, and divisible goods. Considering concave utility functions for buyers and linear utility functions for sellers, the authors have proved that we can achieve an equilibrium that clears the market by solving

- *S.-S. Byun is with Daegu-Gyeongbuk Medical Innovation Foundation, 701-310, Dong-Nae Ro 88, Dong-Gu, Daegu, Korea. E-mail: ssbyun@dgmif.re.kr.*
- *I. Balasingham is with the Intervention Center, Rikshospitalet, Oslo University Hospital, Oslo N0027, Norway. E-mail: ilangkob@medisin.uio.no.*
- *A.V. Vasilakos is with the Department of Computer and Telecommunications Engineering, University of Western Macedonia, Parko Agiou Dimitriou, Kozani GR50100, Greece. E-mail: vasilako@ath.forthnet.gr.*
- *H.-N. Lee is with the School of Information and Communications, Gwangju Institute of Science and Technology, 500-712, Oh-Ryong-Dong, Buk-Gu, Gwangju, Korea. E-mail: heungno@gist.ac.kr.*

Manuscript received 14 Mar. 2012; revised 9 Aug. 2012; accepted 12 Aug. 2012; published online 28 Aug. 2012.
Recommended for acceptance by V. Eramo.
For information on obtaining reprints of this article, please send e-mail to: tc@computer.org, and reference IEEECS Log Number TC-2012-03-0193.
Digital Object Identifier no. 10.1109/TC.2012.211.

1. Hereafter, SU implies transmitting SU.
2. Refer to [5] for the details.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 11,2024 at 05:40:35 UTC from IEEE Xplore. Restrictions apply.
0018-9340/14/$31.00 © 2014 IEEE Published by the IEEE Computer Society

a convex optimization problem called *Eisenberg-Gale convex program*. However, in our spectrum market, the purchasing amount is restricted by other buyers as well as sellers unlike the model in [6]: the transmission rate of SU is restricted by the interference limits defined by other SUs as well as PUs. To derive an equilibrium with this restriction, the buyers (i.e., SUs) should pay additional charges to other buyers.

Including the interference limit among SUs into the market model in [6], we define the market equilibrium such that the following conditions hold:

1. The transmission power vector yielded by the market equilibrium maximizes the joint utility of SUs in the Pareto sense satisfying the constraints of budget, interference to PUs, and interference to other SUs.
2. The transmission power vector yielded by the market equilibrium maximizes the utility of every PU, and completely consumes the interference exactly up to the limits set by the PUs.
3. With the prices and charges yielded by the market equilibrium, the sum of the initial budget possessed by all SUs equals the sum of the profits made by all PUs plus the sum of the interference charges gathered by all the SUs.

We model our spectrum market using the Eisenberg-Gale convex program whose objective is to maximize the joint utility function of all SUs, which is given by log-sum-utility, over a convex region defined via a set of linear constraints, and prove that the convex program yields the equilibrium holding the above-mentioned conditions. Meanwhile, this convex program becomes *Nash bargaining problem* whose solution satisfies the weak Pareto optimality, and its Lagrangian dual variables turn out to be the prices and charges given by the equilibrium. Furthermore, we show that the equilibrium satisfies the *core stability* of PUs. However, the utility functions in the convex program should be concave and homogeneous of degree one. Since we formulate the joint utility function of all SUs as a concave function, we apply a monotone-transformation that transforms a concave function into an equivalent function that is homogeneous of degree one with maintaining the concavity.

To find the Lagrangian dual variables, we should solve the system of the linear equations that consist of the Karush-Kuhn-Tucker (KKT) optimality conditions of the convex program. However, the system is generally inconsistent, and which implies that it is impossible to find the exact Lagrangian variables. For this reason, we solve the system with a certain small precision bound. In numerical experiments, we show that approximate Lagrangian variables are found with quite small precision error, and yield an approximate equilibrium quite close to the exact one.

We also consider a distributed implementation for solving the convex program, which enables PUs and SUs to make their decisions autonomously as follows: given the price of each channel, 1) each SU computes her optimal transmission power vector using *best response dynamics*; 2) using linear dynamics, each PU updates the price of her each subchannel in proportion to the interference invoked from the SUs, and, meanwhile, each SU updates the interference charge on each subchannel in proportion to the interference invoked from all the other SUs; and 3) repeat "1" and "2" until the linear dynamics become stationary. We show that the linear dynamics are *asymptotically stable* with any initial points, and the solutions yielded at the stable state is equivalent to the solutions given by the convex program (i.e., *market equilibrium*). By numerical evaluations, we illustrate the market equilibrium given by the distributed algorithm is quite close to the equilibrium yielded by the convex program.

Recently, a distributed algorithm for finding the market equilibrium in the Fisher model has been studied in [7]; the authors have applied proportional response dynamics and proved its convergence. To our best knowledge, there is no previous work that has developed a distributed implementation for finding the equilibrium in the extended model given in [6]. Furthermore, in this paper, we consider the situation where the purchasing amount is restricted by buyers (SUs) as well sellers (PUs).

The remainder of this paper is organized as follows: In Section 2, we summarize some important characteristics of the market model studied in [6], and survey some recent market-based approaches for CRNs. In Section 3, we give our spectrum market model. In Section 4, we define the market equilibrium in our spectrum market. In Section 5, we give the Eisenberg-Gale convex program for our spectrum market, and investigate its properties. In Section 6, we present the distributed algorithm for solving the convex program, and analyze its stability and asymptotic optimality. In Section 7, we give several results of the numerical evaluations. Finally, we conclude this paper in Section 8.

2 Related Work

2.1 Market-Based Approaches for CRNs

Recently, the market-based approach has been deployed to CRNs by several research efforts.

Xing et al. [8] have considered a spectrum market where different consumers evaluate the same supplier differently according to their applications and locations. Considering limited information, they have developed price dynamics with a stochastic learning algorithm to find the optimal price yielding maximum benefit of the suppliers. However, they have not addressed the utility of the consumers.

Hong and Garcia [9] have proposed a fully distributed algorithm—no collaboration among SUs and PUs—that achieves a market equilibrium in multichannel sharing CRN such that the supply of the spectrum equals to its demand, and the network of SU is stable. They have investigated and presented the convergence condition of the algorithm in terms of the channel gain. Unlike other approaches, they have considered the utilities of PUs as well as SUs. However, they have not investigated the social optimality of the market equilibrium.

Niyato and Hossain [10], [11] have modeled a multilevel bandwidth sharing in CRN into an interrelated market. They have proposed a price-demand decision algorithm that guarantees the convergence to a market equilibrium with which all the primary and secondary services are satisfied. However, they have assumed that the price-demand decision algorithm is directed by a central authority in each level, and have not considered the way of allocating the spectra to users in each level.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 11,2024 at 05:40:35 UTC from IEEE Xplore. Restrictions apply.

Xu et al. [12] have proposed a secondary network where SUs trade among themselves their channels purchased from PUs in the direction of asymptotic optimal spectrum utilization. To this end, they have devised dynamic double auction mechanism that is conducted by a centralized spectrum broker, and proved the truthfulness and asymptotic efficiency in the total social welfare.

Li et al. [13] have addressed a spectrum auction mechanism between SUs and spectrum owners without any central auctioneer. They have deployed an iterative matching algorithm that achieves the price set in core where no SUs and spectrum owners can negotiate to do better for both. However, they have assumed fixed transmission power, and there is no consideration of the market clearance.

In [14], Xu et al. have handled a two-tier market: spectrum contracts from a PU to SUs in Tier-1, and spectrum redistribution among SUs to satisfy SUs dynamic traffic demands in Tier-2. They have applied Nash bargaining solution in Tier-1 market to achieve the fairness between the utility of the PU and the aggregate utility of all the SUs. For Tier-2 market, they have deployed random matching and bilateral bargaining. However, they have considered a single PU, and have not addressed the market clearance.

Xie et al. [15] have addressed the spectrum trading between wireless users—that can be regarded as SUs in a CRN—and a single price manager—that can be regarded as a PU or spectrum broker in a CRN, and investigated a market equilibrium where the market clears, and the budgets of the wireless users are completely consumed. Unlike our work, they have addressed the actual interference among the wireless users in the utility function. Then they have shown that the market equilibrium is given by the solution of a linear complementarity problem, and under the symmetric channel gain and low-rank conditions, they have proved that this problem becomes equivalent to the problem of finding KKT points of a quadratic program. Furthermore, they have developed a decentralized tatonnement process that converges to the equilibrium. However, they have not included the manager's utility in the market equilibrium. Moreover, the KKT points of the quadratic program do not guarantee the optimality, and due to this reason, it is not verified whether the distributed algorithm (tatonnement process) converges to the optimal solution even asymptotically.

Koutsopoulos and Iosifidis [16] have surveyed various action mechanisms for spectrum allocation in CRN. They have indicated that auction mechanisms (including double auction mechanisms) involve a single seller and multiple buyers, and have no interaction among buyers.

2.2 Brief of the Extended Market Model

The Fisher model considers a market consisting of multi-buyers and divisible goods. In the Fisher model, the budget for each buyer and the amount of each commodity need to be specified, and the utility functions of buyers are assumed to be concave. Then the market equilibrium in the Fisher model is given by the price of each commodity that yields optimal utilities of buyers at least in Pareto sense and clears the market: there should be neither surplus nor deficiency in any of the commodities and the budgets [17]. In 1959, Eisenberg and Gale gave a convex program for computing market equilibrium for the Fisher model of linear utility functions [18], and in 1961, Eisenberg generalized this to concave homogeneous functions of degree one [19]. In 1954, Nobel laureates Arrow and Debrue generalized the Fisher model considering agents who come to the market with initial endowments of goods, and at any set of prices, want to sell all their goods and buy optimal bundles at these prices. The problem again is to find market clearing prices [20].

Jain et al. [6] have extended the Fisher model considering the utilities that are homothetic, quasiconcave, and homogeneous functions of arbitrary degree; they also have included sellers' utilities into their model. They have applied the Eisenberg-Gale convex program with the buyers' utilities monotone-transformed. They show that the equilibrium price, given by the Lagrangian dual variables of the convex program, maximizes sellers' utilities as well as buyers' utilities, and clears the market.

In our work, we employ this extended model to our spectrum market since it deals with the market where the buyers are clearly distinguished from sellers, and considers the utilities of sellers as well as those of buyers. However, in our spectrum market, the purchasing amount is restricted by other buyers as well as sellers: that is, the amount of interference that can be purchased by each SU is restricted by the interference limits imposed by other SUs as well as PUs. To derive an equilibrium with this restriction, we consider that each buyer (i.e., SU) pays additional charges to all the other buyers. Furthermore, we envisage a distributed implementation of our market model.

We emphasize that the extended Fisher model is a unique market model that enables the following considerations altogether:

1. strictly distinguished multiple sellers (PUs) from multiple buyers (SUs);
2. guarantee of Pareto optimality for SUs' utility and the core-existence for PUs' utilities;
3. guarantee of the market clearance;
4. distributed implementation with no central authority; and
5. providing the way of allocating the spectra to SUs.

3 Spectrum Market Model

We consider a CRN where all the PUs and SUs are located within a limited geographical region. Then we address the spectral resource on the frequency domain taking the multichannel sharing into account—that is, the whole frequency range is split into multiple subchannels, and each subchannel can be shared by multiple users. In this work, we premise that each subchannel is exclusively licensed to a single PU; it, however, can be shared with multiple SUs concurrently unless the SUs invoke interference larger than certain limits (see Fig. 1).

Applying a market concept to our CRN scenario, the subchannels and interference are interpreted into the types of commodity and the quantity of each commodity, respectively. Upon the current prices, each SU decides subchannels and the amount of interference she would like to purchase; each PU updates the price on every her subchannel according to the interference invoked from the SUs.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 11,2024 at 05:40:35 UTC from IEEE Xplore. Restrictions apply.

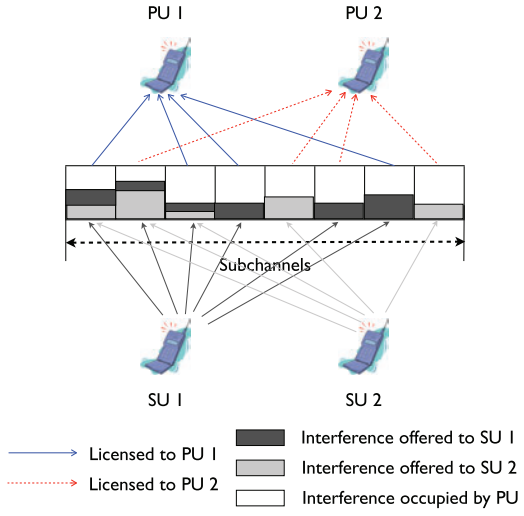


Fig. 1. Multichannel sharing model with two PUs, two SUs, and eight subchannels. Each subchannel is orthogonally allocated to a single PU, and can be shared with all the SUs unless the SUs invoke less interference than the PUs can tolerate.

Ideally, SUs' utility function should reflect actual interference from all the other SUs as well as interference from PUs, which, however, makes the problem nonconvex if the utility function is given as Shannon capacity [21]. Therefore, we consider the maximum allowable interference and reflect it to the utility function; the utility function is given by least achievable transmission rate. Then we let SUs make and charge for every subchannel—on which she is transmitting—in proportion to the amount of interference from all the other SUs. Moreover, we assume that the maximum allowable interference is given by a central authority to let SUs share the interference fairly.

The price as well as the interference charge are marked on every subchannel, and given as a *price per unit interference*.

As a supplier in general market has a limitation on the net supply of its commodity, we envisage that every PU has a limitation on the interference she can offer to SUs over every her subchannel. In addition, like consumers in general markets, every SU cannot spend more budget than she possesses on purchasing the interference from the PUs and paying the interference charges to all the other SUs.

Prior to giving the formal definition of the market equilibrium, we define the following denotations:

- $\mathcal{I}$: Set of transmitting SUs.
- $\mathcal{L}$: Set of PUs, and we let $m := |\mathcal{L}|$.
- $\mathcal{J}$: Set of subchannels, and we let $n := |\mathcal{J}|$
- $u_i : \mathbf{R}_+^n \to \mathbf{R}_+$: Utility function of SU $i \in \mathcal{I}$.
- $\mathbf{p}_i = [p_{i1}, \ldots, p_{in}]^T$: Transmission power vector of SU i.
- p_{ij}: SU i's transmission power on subchannel $j \in \mathcal{J}$.
- $\mathcal{J}_l$: Set of subchannels licensed to PU $l \in \mathcal{L}$.
- $v_l : \mathbf{R}^n \to \mathbf{R}$: Utility function of PU l.
- π_{lj}: Price marked by PU l on subchannel j.
- $\pi_l = [\pi_{l1}, \ldots, \pi_{ln}]^T$: Price vector of PU l.
- η_{ij}: Interference charge decided by SU i on subchannel j.
- $\eta_i = [\eta_{i1}, \ldots, \eta_{in}]^T$: Vector of interference charges decided by SU i.
- y_{lj}: Limit of allowable interference from all the SUs to PU l on subchannel j.
- γ_{ij}: Maximum allowable interference from all the other SUs to SU i on subchannel j.

Now, we develop the spectrum market model with the following key considerations:

1. Each SU $i \in \mathcal{I}$ has a concave, scalable utility function u_i with respect to its transmission power vector p_i. In this paper, the utility of SU is defined as the summation of least achievable rate on every subchannel, that is,

$$u_i(\mathbf{p}_i) = \sum_{j \in \mathcal{J}} B_j \log_2 \left(1 + \frac{p_{ij} G_{ij}}{N_0 + \gamma_{ij} + \Gamma_{ij}}\right), \quad (1)$$

   where B_j is the bandwidth of subchannel $j \in \mathcal{J}$, G_{ij} is the channel gain on subchannel j between SU i and its target, Γ_{ij} is the interference invoked from the PU who owns subchannel j to SU i, and N_0 is the thermal noise.
2. SUs cannot purchase the interference larger than the limits set by PUs. That is, $\forall l \in \mathcal{L}$ and $\forall j \in \mathcal{J}_l$,

$$\sum_{i \in \mathcal{I}} p_{ij} G_{ij}^l \le y_{lj}, \quad (2)$$

   where G_{ij}^l is the channel gain on subchannel j between SU i and PU l. We assume that each PU has a minimum QoS requirement and sets the interference limit (i.e., y_{lj}) in order that the QoS may be guaranteed.
3. Each SU i has an initial endowment of budget $e_i > 0$, and the total budget spent on the purchase of subchannels and the interference charges cannot exceed e_i.
4. Each PU l offers her subchannels, i.e., subchannels in $\mathcal{J}_l$, to SUs with price π_{lj}, and its utility function v_l is defined as the net profit it makes, i.e.,

$$v_l = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}_l} \pi_{lj} p_{ij} G_{ij}. \quad (3)$$

5. In every SU, the total interference on subchannel j from all the other SUs cannot exceed the maximum allowable interference: $\forall i \in \mathcal{I}$ and $\forall j \in \mathcal{J}$

$$\sum_{k \in \mathcal{I}, \neq i} p_{kj} G_{kj}^i \le \gamma_{ij}, \quad (4)$$

   where G_{kj}^i is the channel gain on subchannel j between SU k and the target of SU i. In this paper, we assume a symmetric channel gain among SUs: that is, $G_{kj}^i = G_{ij}^k$ where $i \neq k$.

The channel gain reflects the free space path loss defined by Friis transmission equation [22]:

$$G_{ij} = \frac{g_i g_k \lambda_j^2}{(4\pi)^2 d^2 L}, \quad (5)$$

where g_i is the transmitter antenna gain, g_k is the receiver antenna gain, d is the transmitter-receiver separation distance in meters, L is the system loss factor not related to propagation, and λ_j is the wavelength in meters on subchannel j. Since the channel gain reflects the wave length of each subchannel, the higher frequency a subchannel has,

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 11,2024 at 05:40:35 UTC from IEEE Xplore. Restrictions apply.

the less amount of information a user can transmit on it given a fixed transmission power and subchannels with equal bandwidth [8]. To let every subchannel have equal opportunity of being purchased, we assume that the frequency range is divided into subchannels in the way that every subchannel yields similar transmission rate given a constant transmission power and separate distance.

Additional assumptions are the following:

1. All PUs and SUs can access noncontiguous subchannels in parallel.
2. All SUs have the transmission power enough to fully utilize the interference offered by PUs.
3. The spectrum trade occurs on every predefined epoch, and all SUs perform their transmissions ceaselessly during the epoch.
4. All PUs perform their transmission ceaselessly without any changes in their transmission powers during the epoch.

4 Market Equilibrium

Henceforth, we let y_{lj} have a very large number for all $l \in \mathcal{L}$ and $j \in \mathcal{J}$ but $j \notin \mathcal{J}_l$.[3] Based on the key considerations mentioned in Section 3, we define the equilibrium in the spectrum market following [6]: the equilibrium is defined as a pair of a nonnegative price vector $\pi = [\pi_1, \ldots, \pi_m]^T$ and a vector of interference charges $\eta = [\eta_1, \ldots, \eta_{|\mathcal{I}|}]^T$ at which there exists a transmission power vector p_i for each SU i such that the following conditions hold:

1. The vector $\mathbf{p}_i$ maximizes the utility of SU i given her initial endowment of budget e_i and the equilibrium π and η, that is, $\mathbf{p_i}$ maximizes u_i over all $\mathbf{p}_i \in \mathbf{R}_+^n$ subject to

$$\sum_{l \in \mathcal{L}} \sum_{j \in \mathcal{J}} \pi_{lj} p_{ij} G_{ij}^l + \sum_{j \in \mathcal{J}} p_{ij} \sum_{k \in \mathcal{I}, k \neq i} \eta_{kj} G_{ij}^k \leq e_i, \qquad (6)$$

   and constraint (4).
2. For each PU l, the vector $\mathbf{p}_i$ maximizes the profit v_l subject to constraint (2).
3. The total interference offered by all PUs equals the total interference consumed by all SUs, that is, $\sum_{l \in \mathcal{L}} \sum_{j \in \mathcal{J}} y_{lj} = \sum_{i \in \mathcal{I}} \sum_{l \in \mathcal{L}} \sum_{j \in \mathcal{J}} p_{ij} G_{ij}^l$.
4. The sum of the initial budget possessed by SU i equals the sum of the prices paid to all PUs by SU i plus the sum of the interference charges paid to all the other SUs by SU i, that is,

$$e_i = \sum_{l \in \mathcal{L}} \sum_{j \in \mathcal{J}} \pi_{lj} p_{ij} G_{ij}^l + \sum_{j \in \mathcal{J}} p_{ij} \sum_{k \in \mathcal{I}, k \neq i} \eta_{kj} G_{ij}^k.$$

Therefore, the market equilibrium in this model is also known as market clearing equilibrium since it clears not only all the commodities offered by suppliers but also all the initial budget possessed by consumers; that is, it clears all the budget possessed by SUs as well as all the interference offered by PUs.

3. Over subchannels that a PU does not own, we set the interference limit very large numbers. Then we can drop the subscript l from $\mathcal{J}_l$. Surely, $\forall j \notin \mathcal{J}_l$, y_{lj} does not belong to the commodities of PU l.

However, if SU's maximum allowable interference, i.e., γ_{ij} for any i and j, is set significantly small or null in the worst case, the interference offered by PUs may not be consumed entirely. Then the market clearance cannot be guaranteed. We present this problem in Section 5 in detail.

5 Eisenberg-Gale Convex Program

In this section, we develop a convex program, called Eisenberg-Gale convex program [6], [23], that yields the market equilibrium defined in the previous section. The convex program is to maximize the joint utility function of all SUs given by log-sum-utility over a convex region defined via a set of linear constraints.

The Eisenberg-Gale convex program can compute the market equilibrium only when the utility functions are homogeneous of degree one as well as concave. As given in (1), the utility function is concave but not homogeneous of degree one. Therefore, prior to presenting the convex program, we give a formal method of transforming a nonhomogeneous function into an equivalent function homogeneous of degree one.

5.1 Monotone-Transformation of the Utility Function

Most of all, we need the following definition [24]:

Definition 1. *Given a function $u : \mathbf{R}_+^n \to \mathbf{R}_+$:*

1. *u is* strictly monotonic *if for any $\mathbf{p}$, $\bar{\mathbf{p}} \in \mathbf{R}_+^n$, $\mathbf{p} > \bar{\mathbf{p}}$ implies that $u(\mathbf{p}) > u(\bar{\mathbf{p}})$;*
2. *let u be a strictly monotonic function. Then u is* homothetic *if for any $\mathbf{p}$, $\bar{\mathbf{p}} \in \mathbf{R}_+^n$ and any $\alpha > 0$, $u(\mathbf{p}) \geq u(\bar{\mathbf{p}})$ iff $u(\alpha\mathbf{p}) \geq u(\alpha\bar{\mathbf{p}})$; and*
3. *u is* homogeneous of degree *one if for any $\mathbf{p} \in \mathbf{R}_+^n$ and any $\alpha > 0$, $u(\alpha\mathbf{p}) = \alpha u(\mathbf{p})$.*

It is not difficult to check whether u_i given by (1) is continuous, strictly monotonic and homothetic, but not homogeneous of degree one. Therefore, we apply a *monotone-transformation* [6] that preserves strict monotonicity, concavity, and homotheticity. The monotone-transformation is given by the following theorem:

Theorem 1. *Let $u : \mathbf{R}_+^n \to \mathbf{R}_+$ be a continuous, strict monotonic, concave, homothetic function. Then there is a monotone-transformation yielding a function $f : \mathbf{R}_+^n \to \mathbf{R}_+$ that is homogeneous of degree one, and preserves continuity, strict monotonicity, concavity, and homotheticity, and satisfies:*

1. *If $u(\mathbf{p}) = 0$, then $f(\mathbf{p}) = 0$.*
2. *If $u(\mathbf{p}) \neq 0$, then there exists a unique $\alpha \in \mathbf{R}_+$ such that $u(\mathbf{p}/\alpha) = 1$, and $f(\mathbf{p}) = \alpha$.*

Consequently, the monotone-transformation can be done by finding α that satisfies the following nonlinear equation:

$$u(\mathbf{p}/\alpha) = 1. \qquad (7)$$

5.2 Convex Program Yielding Market Equilibrium

By the monotone-transformation, we can transform u_i to f_i that is homogeneous of degree one. Then we develop the Eisenberg-Gale convex program that yields the market equilibrium as follows:

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 11,2024 at 05:40:35 UTC from IEEE Xplore. Restrictions apply.

$$\text{maximize} \quad \sum_{i\in\mathcal{I}} e_i \ln(f_i) \tag{8}$$

subject to

$$\sum_{k\in\mathcal{I},\neq i} p_{kj} G^i_{kj} \leq \gamma_{ij}, \quad \forall i \in \mathcal{I} \text{ and } \forall j \in \mathcal{J}, \tag{9}$$

$$\sum_{i\in\mathcal{I}} p_{ij} G^l_{ij} \leq y_{lj}, \quad \forall l \in \mathcal{L} \text{ and } \forall j \in \mathcal{J}. \tag{10}$$

As we have addressed in Section 4, if γ_{ij} is set too small, the market clearance cannot be guaranteed. Considering nonzero channel gains, we present the problem in detail:

- In case $\gamma_{ij} = 0$ for all i and j, all p_{ij} should be null. Then, the conditions for the market equilibrium (i.e., conditions 3 and 4) can never be met.
- In case γ_{ij} for all i and j is nonzero, but it has significantly small value, p_{ij} should have also small value to meet the constraint given in (9). Then the constraint (10) may not be tight always. The tightness of the constraint (10) is one of the condition for the market equilibrium, i.e., 3). If the tightness of the constraint (10) is not met, the price of the subchannels that are not entirely purchased by SUs is given as zero according to the KKT condition given in (15), and in turn the condition 4) may not hold either.

To prevent the above problem, we need to add the following necessity condition for the existence of the market equilibrium: every $\gamma_{ij}\ \forall i \in \mathcal{I}$ and $\forall j \in \mathcal{J}$ is nonzero and has a value that makes a solution vector $\mathbf{p}$ satisfy $\sum_{i\in\mathcal{I}} p_{ij} G^l_{ij} = y_{lj}\ \forall l \in \mathcal{L}$ and $\forall j \in \mathcal{J}_l$, and $\sum_{k\in\mathcal{I},\neq i} p_{kj} G^i_{kj} \leq \gamma_{ij}\ \forall i \in \mathcal{I}$ and $\forall j \in \mathcal{J}$ always. Intuitively, SUs will always utilize the available interference entirely bounded by constraint (10) to maximize their collaborative utility given in (8) if constraint (10) is set tighter than constraint (9), and therefore the market clearance can be guaranteed. As an instance, we consider a spectrum market where the channel gains over all SUs, PUs, and channels have an identical value, i.e., identical G^l_{ij} and G^i_{kj} for all $i, k \in \mathcal{I}$, $j \in \mathcal{J}$, and $l \in \mathcal{L}$. Additionally, we let $\gamma_{ij} = c$ and $y_{lj} = f$ for all $i \in \mathcal{I}$, $j \in \mathcal{J}$, and $l \in \mathcal{L}$. Then, if $c \geq f$, constraint (10) over all l and j becomes always tight to achieve the maximum utility.

For all l and $j \notin \mathcal{J}_l$, we give a large value to y_{lj}. In this case, π_{lj} should be zero to satisfy the optimality condition given in (15), and this setting is reasonable. Then we draw the following remark from the convex program:

Remark 1. The solution of the above Eisenberg-Gale convex program (henceforth, convex program) is often regarded as the *Nash bargaining solution* [25] with zero disagreement point. Therefore, it has a unique solution vector that satisfies *Pareto optimality* due to the strict concavity of the objective function and linearity of the constraints [26].

Let $\widetilde{p}_{ij}$ denote the optimal solutions to the convex program. Notice that $f_i(\widetilde{\mathbf{p}}_i) > 0$ for all i. Now, we have the following KKT optimality conditions with the corresponding Lagrangian multipliers η_{ij} and π_{lj}:

$$\widetilde{p}_{ij}\left(\frac{e_i}{f_i(\widetilde{\mathbf{p}}_i)}\frac{\partial f_i(\widetilde{\mathbf{p}}_i)}{\partial p_{ij}} - \sum_{l\in\mathcal{L}} \pi_{lj} G^l_{ij} - \sum_{k\in\mathcal{I},k\neq i} \eta_{kj} G^i_{kj}\right) = 0, \quad \forall i \in \mathcal{I} \text{ and } \forall j \in \mathcal{J}, \tag{11}$$

$$\frac{e_i}{f_i(\widetilde{\mathbf{p}}_i)}\frac{\partial f_i(\widetilde{\mathbf{p}}_i)}{\partial p_{ij}} - \sum_{l\in\mathcal{L}} \pi_{lj} G^l_{ij} - \sum_{k\in\mathcal{I},k\neq i} \eta_{kj} G^i_{kj} \leq 0, \quad \forall i \in \mathcal{I} \text{ and } \forall j \in \mathcal{J}, \tag{12}$$

$$\sum_{k\in\mathcal{I},\neq i} \widetilde{p}_{kj} G^i_{kj} \leq \gamma_{ij}, \quad \forall i \in \mathcal{I} \text{ and } \forall j \in \mathcal{J}, \tag{13}$$

$$\sum_{i\in\mathcal{I}} \widetilde{p}_{ij} G^l_{ij} \leq y_{lj}, \quad \forall l \in \mathcal{L} \text{ and } \forall j \in \mathcal{J}, \tag{14}$$

$$\pi_{lj}\left(\sum_{i\in\mathcal{I}} \widetilde{p}_{ij} G^l_{ij} - y_{lj}\right) = 0, \quad \forall l \in \mathcal{L} \text{ and } \forall j \in \mathcal{J}, \tag{15}$$

$$\eta_{ij}\left(\sum_{k\in\mathcal{I},\neq i} \widetilde{p}_{kj} G^i_{kj} - \gamma_{ij}\right) = 0, \quad \forall i \in \mathcal{I} \text{ and } \forall j \in \mathcal{J}, \tag{16}$$

and (11) and (12).

Subsequently, we establish the following linear program for each PU l:

LP1:

$$\text{maximize} \quad \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \pi_{lj} p_{ij} G^l_{ij} \tag{17}$$

subject to

$$\sum_{i\in\mathcal{I}} p_{ij} G^l_{ij} \leq y_{lj}, \quad \forall j \in \mathcal{J}. \tag{18}$$

We prove that π_l and η_i is the equilibrium price, and $\widetilde{\mathbf{p}}$ is the optimal solution of LP1 as well. This proof begins with Euler's theorem [24]:

Theorem 2 (Euler's Theorem). *Let $f(\mathbf{p})$ be a homogeneous function of degree 1 on $\mathbf{R}^n_+$. Then, for all $\mathbf{p}$,*

$$p_1\frac{\partial f(\mathbf{p})}{\partial p_1} + p_2\frac{\partial f(\mathbf{p})}{\partial p_2} + \cdots + p_n\frac{\partial f(\mathbf{p})}{\partial p_n} = f(\mathbf{p}). \tag{19}$$

Then the following core theorem holds:

Theorem 3. *The optimal solution to the convex program optimizes the utility of each SU i and the utility of each PU l, and the Lagrangian multipliers, π and η, are the equilibrium. In addition, the interference offered by all PUs is entirely exhausted by SUs, and the initial budget possessed by all SUs is fully spent and precisely equals to the total profit earned by the PUs plus the total interference charges. Namely, the market clears.*

Proof. Summing (11) over $j \in \mathcal{J}$, we get

$$\sum_{j\in\mathcal{J}} \frac{e_i}{f_i(\widetilde{\mathbf{p}}_i)}\frac{\partial f_i(\widetilde{\mathbf{p}}_i)}{\partial p_{ij}}\widetilde{p}_{ij} - \sum_{j\in\mathcal{J}} \widetilde{p}_{ij}\sum_{l\in\mathcal{L}} \pi_{lj} G^l_{ij} - \sum_{j\in\mathcal{J}} \widetilde{p}_{ij}\sum_{k\in\mathcal{I},k\neq i} \eta_{kj} G^i_{kj} = 0, \quad \forall i \in \mathcal{I}. \tag{20}$$

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 11,2024 at 05:40:35 UTC from IEEE Xplore. Restrictions apply.

By Euler's theorem, the first term in the left side of (20) is reduced to e_i simply. Then, since we assume $G^i_{kj} = G^k_{ij}$, $\forall i, k \in \mathcal{I}$ and $\forall j \in \mathcal{J}$ where $i \neq k$,

$$e_i = \sum_{j\in\mathcal{J}} \widetilde{p}_{ij} \sum_{l\in\mathcal{L}} \pi_{lj} G^l_{ij} + \sum_{j\in\mathcal{J}} \widetilde{p}_{ij} \sum_{k\in\mathcal{I},k\neq i} \eta_{kj} G^k_{ij} = 0, \quad \forall i \in \mathcal{I}, \tag{21}$$

which implies that each SU spends her initial budget completely under the equilibrium π and η. The first term in the right side of (21) indicates the total price to be paid to all PUs by SU i, and the second term indicates the total interference charge to be paid to all other SUs by SU i.

We next consider the dual program of LP1 for each PU l with dual variables w:

LP2:

$$\text{minimize} \quad \sum_{j\in\mathcal{J}} y_{lj} w_{lj} \tag{22}$$

subject to

$$w_{lj} G^l_{ij} = \pi_{lj} G^l_{ij}, \quad \forall i \in \mathcal{I} \text{ and } \forall j \in \mathcal{J}, \tag{23}$$

$$w_{lj} \geq 0, \quad \forall j \in \mathcal{J}. \tag{24}$$

Let $\bar{\mathbf{p}}$ be the optimal solution of LP2. By the complementary slackness condition [27], the following equation holds as well:

$$w_{lj} \left(\sum_{i\in\mathcal{I}} \bar{p}_{ij} G^l_{ij} - y_{lj} \right) = 0, \quad \forall l \in \mathcal{L} \text{ and } \forall j \in \mathcal{J}. \tag{25}$$

Assuming that $G^l_{ij} \neq 0$ for all i, j, and l, (23) becomes $w_{lj} = \pi_{lj}$. Thus,

$$\pi_{lj} \left(\sum_{i\in\mathcal{I}} \bar{p}_{ij} G^l_{ij} - y_{lj} \right) = 0, \quad \forall l \in \mathcal{L} \text{ and } \forall j \in \mathcal{J}. \tag{26}$$

We see that $\widetilde{\mathbf{p}}$ also satisfies (26), and thus it can be the optimal solution of the LP2 as well. Hence, together with Remark 1, we draw a conclusion that all the PUs and SUs can achieve maximum utilities with respect to the equilibrium π and η.

Finally, we show that the market clears. Summing (21) over all $i \in \mathcal{I}$, we get

$$\sum_{i\in\mathcal{I}} e_i = \sum_{i\in\mathcal{I}} \sum_{j\in J} \widetilde{p}_{ij} \sum_{l\in\mathcal{L}} \pi_{lj} G^l_{ij} + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \widetilde{p}_{ij} \sum_{k\in\mathcal{I},k\neq i} \eta_{kj} G^k_{ij}, \tag{27}$$

and by (15),

$$\sum_{i\in\mathcal{I}} e_i = \sum_{j\in\mathcal{J}} \sum_{l\in\mathcal{L}} \pi_{lj} y_{lj} + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \widetilde{p}_{ij} \sum_{k\in\mathcal{I},k\neq i} \eta_{kj} G^k_{ij}. \tag{28}$$

Eventually, we conclude that the equilibrium given by the Lagrangian multipliers clears the market. □

The equilibrium can be computed by solving the system of the linear equations (11), (12), (15), and (16) using the optimal solutions of the convex program. It is known that the system of the linear equations has a unique solution if it is consistent [17]. However, the system is normally inconsistent since there are more equations than the unknown variables. Thus, we need to provide a certain precision bound on each linear equation to achieve an approximate equilibrium at least. The precision bound is dependent on the instance of the problem such as the sizes of PUs, SUs, subchannels, and nonzero solutions. In Section 7, we evaluate numerically the quality of the solution varying the precision bound given a CRN instance.

Besides, to solve the system of the linear equations, we need to compute the partial derivatives of the monotone-transformed function (required in (11) and (12)), and it is given by the following lemma [6]:

Lemma 1. *If we let $f(\mathbf{p}) = \alpha$, then the partial derivatives to p_j of f is given by*

$$\frac{\partial f(\mathbf{p})}{\partial p_j} = \alpha \frac{\partial u(\mathbf{p}/\alpha)/\partial p_j}{\nabla u(\mathbf{p}/\alpha)^T \mathbf{p}}. \tag{29}$$

5.3 Core Stability of the Production Vector

In this section, we show that the solution $\widetilde{\mathbf{p}}$ yielded by the convex program is in the *core* of nontransferable utility (NTU) game in the context of cooperative game theory [28].

In a cooperative game, the existence of nonempty core guarantees that no player will break away from the grand coaltion (i.e., the cooperation of all the players) since the payoffs achieved by the cooperation within any subcoalition are not larger than the payoff yielded by the cooperation of the grand coalition. Therefore, if the solution $\widetilde{\mathbf{p}}$ is in the core of the spectrum market, we guarantee that no PU will leave this market.

Denoting the core of the spectrum market as $C(V)$, the core of an NTU game is defined as:

Definition 2. *For a $S \subseteq \mathcal{L}$, let $V(S) = \{v_l(\mathbf{p}) : l \in S\}$. Then the core of the spectrum market, $C(V)$ is defined as the set of all undominated imputations, i.e., $\widetilde{\mathbf{p}} \in C(V)$, if and only if there is neither $S \subset \mathcal{L}$, $S \neq \emptyset$, nor $\mathbf{p}$ such that $v_l(\mathbf{p}) > v_l(\widetilde{\mathbf{p}})$ for all $l \in S$.*

Theorem 4. *The solution vector $\widetilde{\mathbf{p}}$ of the convex program belongs to $C(V)$.*

Proof. In this proof, we apply the proof by contradiction. Thus, this proof begins with supposing $\widetilde{\mathbf{p}} \notin C(V)$. Then there exists $S \subset L$, $S \neq \emptyset$, and some allocation $\widehat{\mathbf{p}}$ such that

$$v_l(\widehat{\mathbf{p}}) > v_l(\widetilde{\mathbf{p}}), \tag{30}$$

for all $l \in S$. That is,

$$\sum_{j\in\mathcal{J}} \pi_{lj} \sum_{i\in\mathcal{I}} \widehat{p}_{ij} G^l_{ij} > \sum_{j\in\mathcal{J}} \pi_{lj} \sum_{i\in\mathcal{I}} \widetilde{p}_{ij} G^l_{ij}, \quad \forall l \in S. \tag{31}$$

Then, by (15),

$$\sum_{j\in\mathcal{J}} \pi_{lj} \sum_{i\in\mathcal{I}} \widehat{p}_{ij} G^l_{ij} > \sum_{j\in\mathcal{J}} \pi_{lj} y_{lj}, \quad \forall l \in S. \tag{32}$$

To satisfy (32), the following relation should be satisfied in any l and j:

$$\sum_{i\in\mathcal{I}} \widehat{p}_{ij} G^l_{ij} > \sum_{j\in\mathcal{J}} y_{lj}, \quad \forall l \in S. \tag{33}$$

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 11,2024 at 05:40:35 UTC from IEEE Xplore. Restrictions apply.

However, the inequality (33) violates the constraint of the total interference. This contradiction proves that $\widetilde{\mathbf{p}}$ is undominated, and therefore belongs to $C(V)$. □

Consequently, it is guaranteed that the interference solutions given by the convex program let no PU break away from the spectrum market.

6 Distributed Algorithm

In this section, we develop a distributed approach whose stationary point is asymptotically equivalent to the optimal solution given by the convex program.

6.1 The Algorithm

A natural class of dynamics in multiplayer noncooperative system is the *best-response dynamics* where each player updates her strategy to maximize her utility given the strategies of other players [29]. In this algorithm, each SU deploys the best-response dynamics to maximize her individual utility given the vectors of price and interference charges. That is, the best response of SU i is given by

$$\beta_i(\pi,\eta) = \underset{\mathbf{p}_i:\text{s.t.}(6)}{\arg\max}\, e_i \ln f_i(\mathbf{p}_i). \tag{34}$$

Accordingly, the algorithm with the best-response dynamics is given as follows:

1: Initialize $\pi(0)$ and $\eta(0)$;
2: $t \leftarrow 0$;
3: **loop**
4: **for** each $i \in \mathcal{I}$ **do**
5: Find the best response $\beta_i(t)$ given $\pi(t)$ and $\eta(t)$;
6: **end for**
7: **for** each $l \in \mathcal{L}$ and $j \in \mathcal{J}$ **do**
8: **if** $j \notin \mathcal{J}$ **then**
9: $\pi_{lj} = 0$;
10: **else**
11: Update the price such that

$$\dot{\pi}_{lj} = \alpha\left(\sum_{i\in\mathcal{I}} p_{ij}(t)G^l_{ij} - y_{lj}\right)^+_{\pi_{lj}}; \tag{35}$$

12: **end if**
13: **end for**
14: **for** each $i \in \mathcal{I}$ and $j \in \mathcal{J}$ **do**
15: **if** $p_{ij}(t) = 0$ **then**
16: $\eta_{ij} = 0$;
17: **else**
18: Adjust the interference charge by

$$\dot{\eta}_{ij} = \alpha\left(\sum_{k\in I, k\neq i} p_{kj}(t)G^l_{kj} - \gamma_{ij}\right)^+_{\eta_{ij}} \tag{36}$$

19: **end if**
20: **end for**
21: **if** $\dot{\pi}_{lj} < \epsilon$ and $\dot{\eta}_{lj} < \epsilon$ for all $l \in \mathcal{L}$ and $j \in \mathcal{J}$ **then**
22: Terminate the loop;
23: **else**
24: $t \leftarrow t + 1$;
25: Continue the loop;
26: **end if**
27: **end loop**

In this algorithm, α and ϵ indicate the adjustment speed of the linear dynamics and the termination condition of the algorithm, respectively. If we set the adjustment speed too small, the algorithm requires more iterations until it converges. On the other hand, setting it too large can make the algorithm oscillate. There is no formal way of finding an adequate value for the adjustment speed except trial-and-error. In this paper, an adequate adjustment speed is found by testing the algorithm with various adjustment speed. $(a)^+_b$ implies $\max(a,0)$ if $b = 0$, and equal to a if $b > 0$. In every time t, by (34), each SU finds the transmission power vector that is her best response to the current price and interference charges. When each SU computes its best response dynamics, she should know the amount of interference from PUs as well as the channel gain to her target. By (35), each PU updates the price on each her subchannel according to the interference from all SUs. By (36), each SU updates her interference charge on each subchannel according to the interference from other SUs. We note that PUs and SUs update their interference charges with only the amount of interference they observe. Then the price updated by each PU (i.e., π) is delivered to all SUs, and the price updated by each SU (i.e., η) is delivered to other SUs. These price lists are the only feedbacks that should be delivered.

6.2 Stability Analysis

In this section, we prove the asymptotic stability of the linear dynamics given by (35) and (36). To this end, we develop the following theorem.

Theorem 5. *The linear dynamics given by (35) and (36) are globally asymptotically stable.*

Proof. See Appendix A, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TC.2012.211. □

We see that, as the distributed algorithm proceeds, SUs' utilities approach to those yielded by the convex program.

6.3 Investigation of Asymptotic Equivalence to the Convex Program

We have shown the linear dynamics are asymptotically stable. Then, assuming the system of the linear equations (11), (12), (15), and (16) is consistent, we can prove that the distributed algorithm yields the solutions *asymptotically equivalent* to the solutions of the convex program. That is, the solutions yielded at $t = \infty$ is equivalent to the solutions of the convex program. However, as discussed in Section 5.2, it is not easy to validate the equivalence because of the inconsistency in the system of the linear equations. Nonetheless, by numerical experiments, we illustrate that the utilities of SUs determined by the distributed algorithm are almost identical to those yielded by the convex program in spite of the inconsistency, and which will be explained in Section 7. We also show that the equilibrium yielded by the distributed algorithm also meets the KKT conditions within a small tolerance.

We give the following theorem.

Theorem 6. *Postulating that the system of the linear equations (11), (12), (15), and (16) are consistent, the solutions and equilibrium obtained by the distributed algorithm are asymptotically equivalent to those yielded by the convex program and its KKT conditions.*

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 11,2024 at 05:40:35 UTC from IEEE Xplore. Restrictions apply.

Proof. We denote the best response of SU i at $t=\infty$ as $\widetilde{\beta}_i$, the price made by PU l on subchannel j at $t=\infty$ as $\widetilde{\pi}_{lj}$, and the interference charge made by SU i on subchannel j at $t=\infty$ as $\widetilde{\eta}_{ij}$, then the KKT conditions (37)-(39) hold with Lagrangian multiplier $\kappa_i \geq 0$:

$$\widetilde{\beta}_{ij}\left(\frac{e_i}{f_i(\widetilde{\beta}_i)}\frac{\partial f_i(\widetilde{\beta}_i)}{\partial p_{ij}} - \kappa_i\left(\sum_{l\in\mathcal{L}}\widetilde{\pi}_{lj}G^l_{ij} + \sum_{k\in\mathcal{I},k\neq i}\widetilde{\eta}_{ij}G^k_{ij}\right)\right) = 0, \quad \forall j \in \mathcal{J}, \tag{37}$$

$$\frac{e_i}{f_i(\widetilde{\beta}_i)}\frac{\partial f_i(\widetilde{\beta}_i)}{\partial p_{ij}} - \kappa_i\left(\sum_{l\in\mathcal{L}}\widetilde{\pi}_{lj}G^l_{ij} + \sum_{k\in\mathcal{I},k\neq i}\widetilde{\eta}_{ij}G^k_{ij}\right) \leq 0, \quad \forall j \in \mathcal{J}, \tag{38}$$

$$\kappa_i\left(\sum_{j\in\mathcal{J}}\widetilde{\beta}_{ij}\sum_{l\in\mathcal{L}}\widetilde{\pi}_{lj}G^l_{ij} + \sum_{j\in\mathcal{J}}\widetilde{\beta}_{ij}\sum_{k\in\mathcal{I},k\neq i}\widetilde{\eta}_{lj}G^k_{ij} - e_i\right) = 0, \tag{39}$$

$$\sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{L}}\widetilde{\pi}_{lj}\widetilde{\beta}_{ij}G^l_{ij} + \sum_{j\in\mathcal{J}}\widetilde{\beta}_{ij}\sum_{k\in\mathcal{I},k\neq i}\widetilde{\eta}_{lj}G^k_{ij} \leq e_i. \tag{40}$$

Summing (37) over all $j \in \mathcal{J}$, we get (41), and, by Euler's theorem,

$$\sum_{j\in\mathcal{J}}\frac{e_i}{f_i(\widetilde{\beta}_i)}\frac{\partial f_i(\widetilde{\beta}_i)}{\partial p_{ij}}\widetilde{\beta}_{ij} = \kappa_i\left(\sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{L}}\widetilde{\pi}_{lj}\widetilde{\beta}_{ij}G^l_{ij} + \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{I},k\neq i}\widetilde{\eta}_{lj}\widetilde{\beta}_{ij}G^k_{ij}\right), \tag{41}$$

$$e_i = \kappa_i\left(\sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{L}}\widetilde{\pi}_{lj}\widetilde{\beta}_{ij}G^l_{ij} + \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{I},k\neq i}\widetilde{\eta}_{lj}\widetilde{\beta}_{ij}G^k_{ij}\right). \tag{42}$$

Substituting e_i in (39) with (42), we get the following equation:

$$\kappa_i(1-\kappa_i)\left(\sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{L}}\widetilde{\pi}_{lj}\widetilde{\beta}_{ij}G^l_{ij} + \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{I},k\neq i}\widetilde{\eta}_{lj}\widetilde{\beta}_{ij}G^k_{ij}\right) = 0. \tag{43}$$

For all $i \in \mathcal{I}$, at least one of $\widetilde{\mathbf{p}}_i$ should be nonzero to make $f_i(\widetilde{\beta}_i) > 0$. Moreover, due to the strict monotonicity of f_i, the partial derivatives of f_i should be always positive. Therefore, at least one element of $\widetilde{\pi}$ and $\widetilde{\eta}$ should be nonzero to make (37) hold. We also set all the channel gains nonzeros. Subsequently, to make both (37) and (43) hold for all i, κ_i should be 1. Then, the KKT condition (37) and (38) turn out to be identical with (11) and (12), respectively, if we let $\widetilde{\beta}_i = \widetilde{\mathbf{p}}_i$ for all i, $\widetilde{\pi} = \pi$, and $\widetilde{\eta} = \eta$.

Since the two linear dynamics (i.e., (35) and (36)) are stable, there exist some $l \in \mathcal{L}$ and $j \in \mathcal{J}$ such that

$$\sum_{i\in\mathcal{I}}\widetilde{\beta}_{ij}(t)G^l_{ij} = y_{lj}, \tag{44}$$

if $\widetilde{\pi}_{lj} > 0$, or

$$\sum_{i\in\mathcal{I}}\widetilde{\beta}_{ij}(t)G^l_{ij} < y_{lj}, \tag{45}$$

if $\widetilde{\pi}_{lj} = 0$; in a similar way, there exist some $i \in \mathcal{I}$ and $j \in \mathcal{J}$ such that

$$\sum_{k\in\mathcal{I},k\neq i}\widetilde{\beta}_{ij}(t)G^i_{kj} = \gamma_{ij}, \tag{46}$$

if $\widetilde{\eta}_{ij} > 0$, and

$$\sum_{k\in\mathcal{I},k\neq i}\widetilde{\beta}_{ij}(t)G^i_{kj} < \gamma_{ij}, \tag{47}$$

if $\widetilde{\eta}_{ij} = 0$. We notice that (44) and (45) are equivalent to the KKT condition (13), and (46) and (47) are equivalent to the KKT condition (16) if we let $\widetilde{\beta}_i = \widetilde{\mathbf{p}}_i$ for all i, $\widetilde{\pi} = \pi$, and $\widetilde{\eta} = \eta$. Finally, (44)-(47) satisfy both (13) and (14) as well if we let $\widetilde{\beta}_i = \widetilde{\mathbf{p}}_i$ for all i.

For the reasons stated above, the stable solutions and the equilibrium yielded by the distributed algorithm satisfies the KKT condition of the convex program. Since the equilibrium as well as the solutions of the convex program are unique, we confirm $\widetilde{\beta}_i = \widetilde{\mathbf{p}}_i$ for all i, $\widetilde{\pi} = \pi$, and $\widetilde{\eta} = \eta$. Consequently, we conclude that the solutions and equilibrium yielded by the distributed algorithm are asymptotically equivalent to those given by the convex program and its KKT conditions. □

7 Numerical Evaluations

7.1 Experimental Setup

We generate a CRN within a 500×500 square, and consider the frequency range of 54-862-MHz TV band following the IEEE 802.22 standard [30]. As mentioned in Section 3, we divide the frequency range into subchannels in the way that every subchannel yields equal transmission rate as possible given a constant transmission power. We vary the sizes of SUs, PUs, and subchannels according to the type of the experiment we perform. Besides, we use interior point optimizer (IPOPT) [31] for solving the convex program and the best-response dynamics, and GNU linear programming kit (GLPK) [32] for solving the system of the linear equations. Additional experimental parameters are the following:

- e_i: randomly chosen from (0, 1.0].
- y_{lj}: $\forall l \in \mathcal{L}$ and $\forall j \in \mathcal{J}_l$, 1e-08.
- γ_{ij}: $\forall i \in \mathcal{I}$ and $\forall j \in \mathcal{J}$, 1e-08.
- Initial π_{lj} for the distributed algorithm: $\forall l \in \mathcal{L}$ and $\forall j \in \mathcal{J}_l$, 6e06.
- Initial η_{ij} for the distributed algorithm: $\forall i \in \mathcal{I}$ and $\forall j \in \mathcal{J}$, 6e06.
- Antenna gain: 1.0 for both transmitter side and receiver side.
- System loss factor: 1.0.
- Speed of light: 3e08 m/s.
- Thermal noise: 1e-10.
- PU's transmission power: 0.1W for all PUs.

7.2 Illustration of the Strict Monotonicity

First, we illustrate the strict monotonicity of the monotone-transformed utility function, i.e., f_i. To this end, we consider two subchannels and compute the function values of f_i varying the transmission power on each subchannel. Fig. 2b shows the results. We plot also the function values of the original utility function, i.e., u_i in Fig. 2a. As shown in

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 11,2024 at 05:40:35 UTC from IEEE Xplore. Restrictions apply.

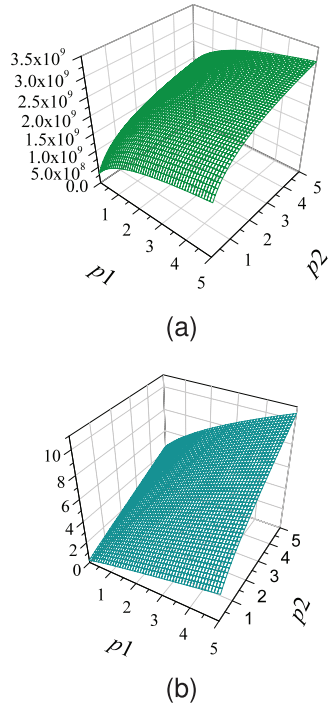


Fig. 2. Illustrations of the strict monotonicity of u_i and f_i.

Fig. 2b, the function value increases strictly monotonically as the transmission power on each subchannel increases, and which illustrates the strict monotonicity.

7.3 Illustration and Evaluation of Equilibrium Price

Next, we measure the transition of the total interference demand according to the change of the prices given by PUs when the function (8) is maximized subject to the budget constraint (i.e., (6)). For these experiments, we locate three SUs and two PUs accommodating two subchannels,[4] and ignore the limit of the interference from other SUs, that is ignore (9). Thus, all the budget possessed by the SUs will be paid to the PUs. Setting $y_{lj} = 8\text{e-}08$ for all l, j, the measured results are shown in Fig. 3, and it is observed that the demand decreases as the prices increase. By the definition, the equilibrium price is obtained when the total interference demand equals $2 \times \sum_{l,j} y_{lj}$, that is, 1.6e-07.

The next set of experiments is done to evaluate the precision of the equilibrium price obtained by solving the system of the linear equations (11)-(16) under various precision bound. For this set of experiments, we locate eight SUs and eight PUs, and arrange 32 subchannels. Then we measure the absolute gap between the initial budget and the payment of each SU under three different precision bounds. Fig. 4 shows the results. If the precision bound is made lower than 1.4508e-5, then the system becomes inconsistent. As shown in the graph, the absolute gaps are measured at most around 4.0e-4 with the smallest feasible precision bound.

7.4 Evaluation of the Distributed Algorithm

In this section, we illustrate the convergence process of the distributed algorithm, and evaluate it in terms of convergence speed and solution quality.

4. We let each PU have one subchannel.

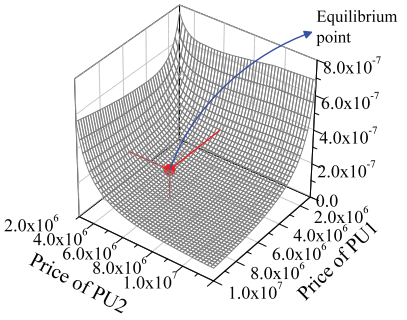


Fig. 3. The transition of the total interference demand according to the prices when we locate three SUs, two PUs, and eight subchannels. This figure also plots the equilibrium point; the price pair given by the equilibrium point is the equilibrium price.

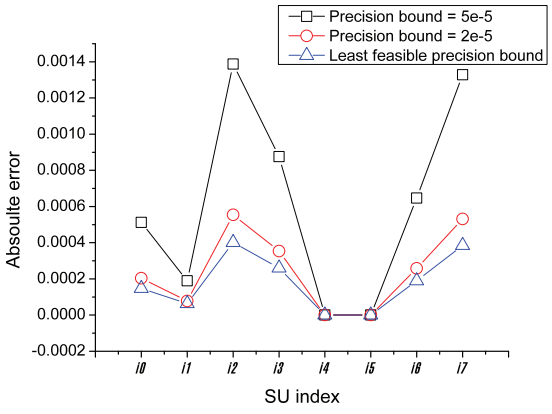


Fig. 4. The absolute gap between the initial budget of each SU and her payment decided by the KKT conditions of the convex program.

7.4.1 Illustration of the Convergence Process

First, we illustrate the convergence process to the equilibrium point with the distributed algorithm. In this set of experiments, we consider three SUs, three PUs, and eight subchannels, and apply the constant adaptive size of 1e13. In Fig. 5, we plot the convergence trajectories as the iteration proceeds. Each axis on each graph indicates the utility of each SU (see Fig. 5a), the utility of each PU (see Fig. 5b), the total price gathered by each PU (see Fig. 5c), and the total interference charge gathered by each SU (see Fig. 5d). It is observed that, as the algorithm approaches to the equilibrium point, the amount of update in each iteration decreases, and which illustrates the asymptotic stability of the algorithm.

7.4.2 Illustration of the Convergence Speed and the Solution Quality

Next, we evaluate the distributed algorithm in terms of the convergence speed and the solution quality. In these experiments, eight SUs and eight PUs are located accommodating 32 subchannels. First, we measure the number of iterations required to reach the termination condition. Here, we define the termination condition as the errors in the KKT optimality conditions. The results are shown in Fig. 6 where we plot the objective function value of (8) on every iteration. In addition, the speed of adjustment is set 2e13. As shown in the graph, as the KKT error is larger, the distributed algorithm converges faster. The function value yielded by the convex program is 79.53536. With KTT error of 1e-2, the algorithm terminates after 33 iterations, and the value of

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 11,2024 at 05:40:35 UTC from IEEE Xplore. Restrictions apply.

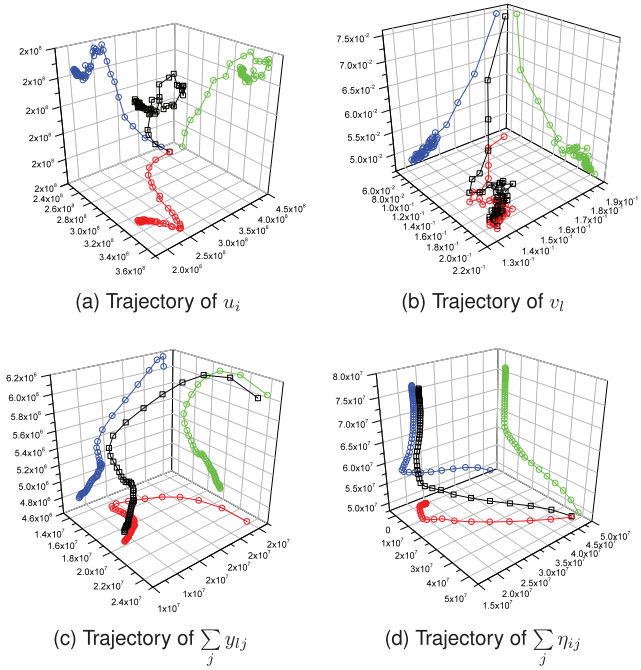


Fig. 5. The trajectories of the transitions of the utilities of SUs and PUs, prices, and interference charges as the distributed algorithm proceeds. We also plot the projection of each trajectory to the surfaces.

the objective function measured 79.43998. With KKT error of 4e-3, the algorithm stops at 188th iteration, and the value is measured 79.59514. With KKT error of 2e-3, the algorithm terminates at 303rd iteration with the function value of 79.53107. Therefore, we notice the tradeoff between the convergence speed and the solution quality.

Furthermore, we measure the utilities of the SUs obtained by the distributed algorithm and those yielded by the convex program. Fig. 7 plots the results. The largest difference is measured around 3e6 at i6, but in percentage, it is 0.67 percent.

We also measure the absolute gaps between the initial budget of each SU and her payment yielded by the distributed algorithm as we have done with the KKT conditions of the convex program in Section 7.3. The measured results are plotted in Fig. 8 together with the results with the KKT conditions. We see that the distributed algorithm yields larger gaps than the KKT conditions, and the largest absolute gap is measured 3.69e-4 at most.

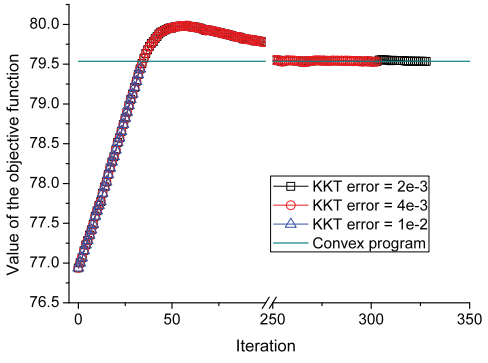


Fig. 6. Graphical presentation of the number of iterations with various termination conditions (i.e., KKT error). The solid green line indicates the optimal value of the objective function yielded by the convex program.

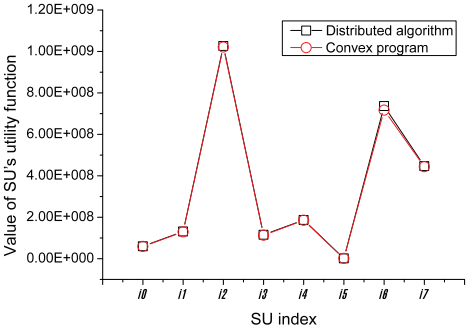


Fig. 7. The comparison of SUs utility.

8 Concluding Remark

In this paper, we consider a market equilibrium in multi-channel sharing CRN. PUs and SUs act as suppliers and purchasers, respectively: PUs offer their subchannels to SUs with bounding the total amount of interference invoked from SUs transmissions, and SUs purchase the offered subchannels observing their budget constraints and the interference bounds given by the PUs. Moreover, we consider that each SU sets the constraint of interference from other SUs. Accordingly, SUs pay the interference charges to other SUs she interferes. The utility functions of SUs and PUs are given as the least achievable transmission rates and the net profits, respectively.

The market equilibrium not only optimizes all traders' (PUs and SUs) utilities but also achieves the market clearance. We show that the market equilibrium is yielded by solving the optimization problem called the Eisenberg-Gale convex program, and the equilibrium price is given by the Lagrangian dual variables of the convex program. The convex program yields the equilibrium only when the utility functions of SUs are homogeneous of degree one. Therefore, we apply a monotone-transformation to SUs utility functions with maintaining the strict monotonicity and concavity.

We also develop a distributed algorithm with which the traders can reach the market equilibrium without any central authority. However, it is impossible to yield the exact equilibrium price since the system of the linear equations—that are composed of the KKT conditions of the convex program—are normally inconsistent, and the convergence behavior of the distributed algorithm is asymptotic. For these reasons, we provide the system of the linear

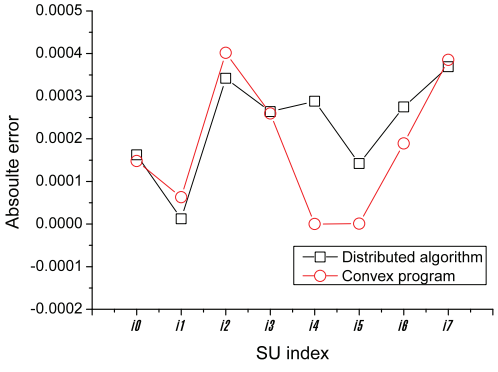


Fig. 8. The comparison of the absolute gaps between the initial budget of each SU and her total payment.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 11,2024 at 05:40:35 UTC from IEEE Xplore. Restrictions apply.

equations with a certain precision bound that makes the system consistent, and give a termination threshold to the distributed algorithm.

By the numerical experiments, we illustrate the strict monotonicity of the monotone-transformed function, and present graphically the existence of the equilibrium and the convergence process of the distributed algorithm. We also measure the absolute errors in the solutions and KKT conditions, which is yielded by the precision bound in the system of the linear equations and the asymptotic optimality of the distributed algorithm. The measured results show that the solutions achieved by the distributed algorithm are quite close to those of the convex program.

REFERENCES

[1] A. O'Sullivan and S.M. Sheffrin, *Economics: Principles in Action.* Prentice Hall, 2003.

[2] A. Al Daoud, M. Alanyali, and D. Starobinski, "Pricing Strategies for Spectrum Lease in Secondary Markets," *IEEE/ACM Trans. Networking,* vol. 18, no. 2, pp. 462-475, Apr. 2010.

[3] "Secondary Markets Initiative. Federal Communications Commision," http://wireless.fcc.gov/licensing/secondarymarkets, 2013.

[4] "Radio Spectrum Policy Group," http://rspg.groups.eu.int, 2013.

[5] K. Binmore, *Playing for Real: A Text on Game Theory.* Oxford Univ. Press, 2007.

[6] K. Jain, V.V. Vazirani, and Y. Ye, "Market Equilibria for Homothetic, Quasi-Concave Utilities and Economies of Scale in Production," *Proc. ACM-SIAM Symp. Discrete Algorithms,* pp. 63-71, Jan. 2005.

[7] B. Birnbaum, N.R. Devanur, and L. Xiao, "Distributed Algorithms via Gradient Descent for Fisher Markets," *Proc. ACM Conf. Electronic Commerce,* pp. 127-136, June 2011.

[8] Y. Xing, R. Chandramouli, and C. Cordeiro, "Price Dynamics in Competitive Agile Spectrum Access Markets," *IEEE J. Selected Areas Comm.,* vol. 25, no. 3, pp. 613-621, Apr. 2007.

[9] M. Hong and A. Garcia, "Competitive Sharing of the Spectrum in Cognitive Radio Network: A Market Equilibrium Framework," *Proc. Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt),* pp. 40-49, June 2010.

[10] D. Niyato and E. Hossain, "Hierarchical Spectrum Sharing in Cognitive Radio: A Microeconomic Approach," *Proc. IEEE Wireless Comm. and Networking Conf. (WCNS),* pp. 3822-3826, Mar. 2007.

[11] D. Niyato and E. Hossain, "A Microeconomic Model for Hierarchical Bandwidth Sharing in Dynamic Spectrum Access Networks," *IEEE Trans. Computers,* vol. 59, no. 7, pp. 865-877, July 2010.

[12] H. Xu, J. Jin, and B. Li, "A Secondary Market for Spectrum," *Proc. IEEE INFOCOM,* Mar. 2010.

[13] D. Li, Y. Xu, J. Liu, X. Wang, and Z. Han, "A Market Game for Dynamic multiBand Sharing in Cognitive Radio Networks," *Proc. IEEE Int'l Conf. Comm. (ICC),* pp. 1-5, May. 2010.

[14] D. Xu, X. Liu, and Z. Han, "A Two-Tier Market for Decentralized Dynamic Spectrum Access in Cognitive Radio Networks," *Proc. IEEE Sensor Mesh and Ad Hoc Comm. and Networks (SECON),* pp. 1-9, June 2010,

[15] Y. Xie, B. Armbruster, and Y. Ye, "Dynamic Spectrum Management with the Competitive Market Model," *IEEE Trans. Signal Processing,* vol. 58, no. 4, pp. 2442-2446, Apr. 2010.

[16] I. Koutsopoulos and G. Iosifidis, "Auction Mechanisms for Network Resource Allocation," *Proc. Int'l Symp. Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt),* pp. 554-563, 2010.

[17] N.R. Devanur, C.H. Papadimitriou, A. Saberi, and V.V. Vazirani, "Market Equilibrium via a Primal-Dual Algorithm for a Convex Program," *J. ACM,* vol. 55, no. 5, pp. 22:1-22:18, Nov. 2008.

[18] E. Eisenberg and D. Gale, "Consensus of Subjective Probablities: The Pari-Mutuel Method," *Annals of Math. Statistics,* vol. 30, pp. 165-168, 1959.

[19] E. Eisenberg, "Aggregation of Utility Functions," *Management Sciences,* vol. 7, no. 4, pp. 337-350, 1961.

[20] K. Arrow and G. Debreu, "Existence of an Equilibrium for a Competitive Economy," *Econometrica,* vol. 22, no. 3, pp. 265-290, 1954.

[21] M. Chiang, C.W. Tan, D.P. Palomar, D. O'Neill, and D. Julian, "Power Control by Geometric Programming," *IEEE Trans. Wireless Comm.,* vol. 6, no. 7, pp. 2640-2651, July 2007.

[22] M. Rice, *RF and Microwave Wireless Systems.* Wiley, 2009.

[23] B. Codenotti and K. Varadarajan, "Computation of Market Equilibria by Convex Programming," *Algorithmic Game Theory,* chapter 6, N. Nisan T. Roughgarden E. Tardos, and V.V. Vazirani, eds., pp. 135-158, Cambridge Univ. Press, 2007.

[24] C.P. Simon and L. Blume, *Mathematics for Economists.* Norton, 1994.

[25] J.F. Nash, "The Bargaining Problem," *Econometrica,* vol. 28, pp. 155-162, 1950.

[26] R.B. Myerson, *Game Theory: Analysis of Conflict.* Harvard Univ. Press, 1991.

[27] A. Ravindran, D.T. Philips, and J.J. Solberg, *Operations Reasearch: Principles and Practice,* second, ed. Wiley, 2007.

[28] G. Owen, *Game Theory,* third, ed. Emerald, 2008.

[29] U. Candogan, I. Menache, A. Ozdaglar, and P. Parrilo, "Near-Optimal Power Control in Wireless Networks: A Potential Game Approach," *Proc. IEEE INFOCOM,* pp. 1-9, Mar. 2010.

[30] C.-K. Yu, K.-C. Chen, and S.-M. Cheng, "Cognitive Radio Network Tomography," *IEEE Trans. Vehicular Technology,* vol. 59, no. 4, pp. 1980-1997, May 2010.

[31] A. Wachter and L.T. Biegler, "On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming," *Math. Programming,* vol. 106, no. 1, pp. 25-57, Jan. 2006.

[32] "GNU Linear Programming Kit (GLPK)," http://www.gnu.org/software/glpk/, 2013.



Sang-Seon Byun received the BS, MS, and PhD degrees in computer science from Korea University, Seoul, in 1996, 2002, and 2007, respectively. He was an assistant research professor of the Graduate School of Embedded Software, Korea University, in 2007. From 2007 to 2011, he has served as a postdoctoral researcher and research scientist of the Department Electronics and Telecommunications, Norwegian University of Science and Technology, Trondheim, Norway. He has also worked as a research professor in the School of Information and Communications, Gwanju Institute of Science and Technology, Korea. Currently, he is a senior researcher of Daegu-Gyeongbuk Medical Innovation Foundation, Korea. His research interests include the fields of cognitive radio networks and software-defined radio architecture. He is a member of the IEEE.



Ilangko Balasingham received the MSc and PhD degrees in signal processing from the Department of Telecommunications, Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 1993 and 1998, respectively. He performed the master's degree thesis at the Department of Electrical and Computer Engineering, University of California, Santa Barbara. From 1998 to 2002, he worked as a research scientist at Fast Search and Transfer ASA, Oslo, Norway. Since 2002, he has been with the Intervention Centre, Oslo University Hospital, Norway, as a senior research scientist, where he heads the Wireless Sensor Network Research Group. He was appointed as a professor in signal processing in medical applications at NTNU in 2006. His research interests include super robust short-range communications for both in-body and on-body sensors, body area sensor network, microwave short-range sensing of vital signs, and short-range localization and tracking sensors, catheters, and micro robots. He is the general chair of the 2012 Bodynets conference in Norway and serves an area editor of the Elsevier Nano Communication Networks. He is a senior member of the IEEE.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 11,2024 at 05:40:35 UTC from IEEE Xplore. Restrictions apply.



Athanasios V. Vasilakos is currently a professor at the Department of Computer and Telecommunications Engineering, University of Western Macedonia, Greece, and a visiting professor at the Graduate Programme of the Department of Electrical and Computer Engineering, National Technical University of Athens. He has authored or coauthored more than 200 technical papers in major international journals and conferences. He is author/coauthor of five books, 20 book chapters in the areas of communications. He served as a general chair, TPC chair, and symposium chair for many international conferences. He served or is serving as an editor or/and guest editor for many technical journals, i.e., *IEEE TNSM*, *IEEE TSMC-Part B*, *IEEE TITB*, and *JSAC*. He is founding editor-in-chief of the journals: *International Journal of Adaptive and Autonomous Communications Systems* (IJAACS, http://www.inderscience.com/ijaacs), *International Journal of Arts and Technology* (IJART, http://www.inderscience.com/ijart). He is the chairman of the European Alliance for Innovation. He is a senior member of the IEEE.



Heung-No Lee received the BS, MS, and PhD degrees in electrical engineering from the University of California at Los Angeles, in 1993, 1994, and 1999, respectively. He then moved to HRL Laboratory, Malibu, California, and worked there as a research staff member from 1999 to 2002. He was then appointed as an assistant professor at the University of Pittsburgh, Pennsylvania, in 2002 where he stayed until 2008. He then moved to Gwangju Institute of Science and Technology in 2009. His research interests include information theory, signal processing theory, and communications/networking theory, and their application to wireless communications and networking, compressive sensing, future Internet, and brain computer interface. He has served as the lead guest editor for *EURASIP Journal on Wireless Communications and Networking* from 2010 to 2011. He has also served as a member of technical program committees for several renowned IEEE conferences, including the IEEE International Conference on Communications, and Globecom. He is a member of the IEEE.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.

Authorized licensed use limited to: Kwangju Institute of Science and Technology. Downloaded on January 11,2024 at 05:40:35 UTC from IEEE Xplore. Restrictions apply.

Signal Processing 90 (2010) 373–377

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Fast communication

MB iterative decoding algorithm on systematic LDGM codes: Performance evaluation

Cheng-Chun Chang [a], Zhi-Hong Mao [a], Heung-No Lee [b,*]

[a] Department of Electrical and Computer Engineering, University of Pittsburgh, Pittsburgh, PA, USA
[b] Gwangju Institute of Science and Technology (GIST), Republic of Korea

ARTICLE INFO

Article history:
Received 9 May 2008
Received in revised form
22 May 2009
Accepted 25 May 2009
Available online 31 May 2009

Keywords:
Systematic LDGM codes
Majority-rule decoding
Performance bounds

ABSTRACT

We investigate the performance of regular systematic low-density generator matrix (LDGM) codes under the majority rule based (MB) iterative decoding algorithm. We derive a recursive form which can be used to extract the error performance of the code. Based on the recursive expression, we derive a tight non-recursive lower bound. These results can serve as efficient tools to evaluate the performance of the code for different degrees.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Systematic low-density generator matrix (LDGM) codes with moderate code length are of interest not only because they can provide satisfying performance at moderate block length while maintaining low encoding and decoding complexities [1–4], but also because the systematic form of LDGM codes makes the code useful in new applications such as cooperative wireless multiple access relay network [5] and joint source-channel encoding systems [6].

In this paper, we are interested in the performance of the majority rule based (MB) iterative decoding algorithm for systematic LDGM codes. Although the MB algorithm is based on hard-decision and thus its performance cannot match those based on soft-decision, it has drawn significant interest in the past owing to its simplicity and low computation complexity, which allow fast decoding [4,7–9].

We investigate the asymptotic performance of the code consisting of regular systematic LDGM codes and the MB iterative decoding algorithm. By assuming infinite block length, we derive a recursive expression which predicts both the threshold and error floor behaviors of the code. Gallager has analyzed an MB iterative decoding for low-density parity-check (LDPC) codes; we build our analysis on systematic LDGM codes by extending his results reported in [9]. Based on the recursive expression, we further derive a non-recursive lower bound expression which is simply a function of the *degree of variable nodes*. We show that the bound is tight in simulation, and thus it can be useful to quickly assess the performance of the code for given degrees.

The rest of the paper is organized as follows. In Section 2, we briefly introduce the systematic LDGM codes and the majority rule based iterative decoding algorithm. The recursive expression and the lower bound expression are derived in Section 3. In Section 4, computer simulation results and analysis results are compared and discussion is provided. Finally, we make a conclusion in Section 5.

2. Systematic LDGM codes and MB algorithm

Similar to the well known LDPC codes, systematic LDGM codes can be represented by sparse matrices,

* Corresponding author. Tel.: +82 62 970 2237; fax: +82 62 970 2204.
E-mail addresses: chc55@pitt.edu (C.-C. Chang), maozh@engr.pitt.edu (Z.-H. Mao), heungno@gist.ac.kr (H.-N. Lee).

0165-1684/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2009.05.019

see [10]. We briefly give the definition of systematic LDGM codes, and then introduce the majority rule based iterative decoding algorithm.

2.1. Systematic LDGM codes

Systematic LDGM codes are linear block codes with parity check matrix $\mathbf{H}=[\mathbf{P};\mathbf{I}]$, where $\mathbf{P}$ is an $(n-k)$ by k sparse matrix and $\mathbf{I}$ is the $(n-k)$ by $(n-k)$ identity matrix. The positive integer k denotes the number of input bits and n denotes the number of output bits of a systematic LDGM encoder. The matrix $\mathbf{P}$ of ones and zeros can be generated at random. A systematic LDGM code will be called *regular* if both the number of 1's in column in the $\mathbf{P}$ matrix and that in row stay fixed for all columns and rows. Though irregularity can provide performance improvement, regularity could lead to simplified modular implementation in hardware realization. We will study the regular version only in this paper. We denote the *degree* of a variable node as d_v, which is the number of ones in each column in the $\mathbf{P}$ matrix. Similarly, the degree of a check node, d_c, represents the number of ones in each row in the $\mathbf{H}$ matrix. The code can be completely specified by a bipartite graph [11] consisting of check nodes and variable nodes. Since a systematic codeword is composed of message bits and parity-check bits, the variable nodes can be further separated into *message-bit variable* (MV) *nodes* and *parity-check-bit variable* (PV) *nodes*. Based on the structure of the $\mathbf{H}$ matrix, the code rate R of (d_v, d_c)-regular systematic LDGM codes is given as $R=1/(d_v/(d_c-1)+1)$.

It should be noticed that, when a code has a generator matrix with rows of constant weight, the code contains code words of the specified constant weight, and hence the minimum distance of the code could have been specified. In the case of regular systematic LDGM codes, the minimum distance of the code is not larger than d_v+1, i.e., the weight of the rows of the generator matrix.

2.2. The majority-rule based iterative decoding algorithms

There are two steps in each iteration for the MB iterative decoding algorithm. The first step is done in a check node. The output binary message from the ith check node, toward the jth of its d_c variable nodes, is the result of the XOR operation on the rest of d_c-1 incoming binary messages. That is, $c_{i,j}=\oplus_{k=1,(k\neq j)}^{d_c-1}(v_{k,j})$, where the summation is done in modular-2 addition and $v_{k,j}$ is the binary message from the kth variable to the ith check node. The second step is done in a variable node at which the majority rule is applied. Let $f_j, f_j\in\{0,1\}$, denote the hard-decision binary value of the received signal for the jth bit transmission. The output binary message from the jth variable node, toward the ith of its d_v check nodes, is obtained from the rest of d_v-1 incoming messages, and is given by

$$v_{j,i}=\begin{cases}\tilde{f}_j & \text{if}\left(\sum_{k=1,(k\neq i)}^{d_v-1} XOR(f_j,c_{k,j})\right)\geq m,\\ f_j & \text{o.w.}\end{cases} \tag{1}$$

That is, if m or more incoming messages are violated, then the message $v_{j,i}$ is the complement of f_j; otherwise, it holds the value of f_j. At the last iteration, the jth bit is decoded to be $\tilde{f}_j$ if $(\sum_{k=1}^{d_v} XOR(f_j,c_{k,j}))\geq m$; otherwise, the jth bit is decoded to be f_j. In the algorithm, the weight m is an integer between 0 and d_v. The weight m needs to be carefully chosen in each iteration as it affects the performance of the MB iterative decoding algorithm. From the above description, we note that the MB iterative decoding algorithm is extremely simple.

3. Error performance analysis

In this section, we derive the recursive expression (2) and the tight lower bound expression (11). These expressions serve as efficient tools to extract the performance of the codec for given degrees of systematic LDGM codes.

Due to the hard decision characteristic of the MB decoding algorithm, we may assume all the coded bits are transmitted through a binary symmetric channel with error probability P_0. Consider the error performance on an MV node. Assume infinite code length and unfold the MB iterative decoding onto a cycle free decoding tree. Then, the error probability for the message on the MV node after the ith iteration can be expressed by the recursive form

$$P_{i+1}=P_0(1-f(m,P_i))+(1-P_0)(g(m,P_i)). \tag{2}$$

where

$$f(m,x)=\sum_{l=m}^{d_v-1}\binom{d_v-1}{l}\left(\frac{1+(1-2P_0)(1-2x)^{d_c-2}}{2}\right)^{l}\times\left(\frac{1-(1-2P_0)(1-2x)^{d_c-2}}{2}\right)^{d_v-1-l} \tag{3}$$

and

$$g(m,x)=\sum_{l=m}^{d_v-1}\binom{d_v-1}{l}\left(\frac{1-(1-2P_0)(1-2x)^{d_c-2}}{2}\right)^{l}\times\left(\frac{1+(1-2P_0)(1-2x)^{d_c-2}}{2}\right)^{d_v-1-l}. \tag{4}$$

The first term in (2) represents the probability of an event that the MV node was in the error state originally and the error correction mechanism of MB algorithm is not triggered because less than m extrinsic messages, out of d_v-1 total, are in violation. Thus the error in the variable node remains unchanged. The second term represents the probability of an event that the MV node was in the correct state but the error correction mechanism of MB algorithm is falsely triggered—because of m or more extrinsic messages in violation—and forces an error.

It is interesting to compare this recursion result (2) to Gallager's result on regular LDPC codes [9, p. 46]. The difference is that we have $(1-2P_0)(1-2x)^{d_c-2}$ in the recursion equation, instead of $(1-2x)^{d_c-1}$. This belongs to one of the characteristic results of the systematic LDGM codes. It is caused by the one and only one connection made by each PV node to the corresponding check node in the bipartite graph. This causes the error floor effect in systematic LDGM codes.

For a given channel error probability P_0, the weight m and the degrees d_v and d_c determine the behavior of the recursive process (2) and hence the error performance. The optimal weight m which minimizes P_{i+1} in (2) for the ith iteration can be found by exhaustively searching for the integer between 0 and d_v, or by solving the smallest integer m which satisfies the following inequality [9]:

$$\frac{1-P_0}{P_0} \le \left(\frac{1+(1-2P_0)(1-2P_i)^{d_c-2}}{1-(1-2P_0)(1-2P_i)^{d_c-2}}\right)^{2m-d_v+1}. \quad (5)$$

In the following, we derive the lower bound expression based on the recursive expression (2). Taking the partial derivative of (2) with respect to P_i, we obtain

$$\frac{\partial P_{i+1}}{\partial P_i} = -P_0\frac{\partial f}{\partial P_i} + (1-P_0)\frac{\partial g}{\partial P_i}, \quad (6)$$

where

$$\frac{\partial f}{\partial P_i} = \binom{d_v-1}{m}(m\xi^{+^{m-1}}\xi^{-^{d_v-1-m}}\eta^-) \quad (7)$$

and

$$\frac{\partial g}{\partial P_i} = \binom{d_v-1}{m}(m\xi^{-^{m-1}}\xi^{+^{d_v-1-m}}\eta^+). \quad (8)$$

In (7) and (8), the notations ξ^+, ξ^-, η^+, and η^- are defined as $\xi^+ = (1+(1-2P_0)(1-2P_i)^{d_c-2})/2$, $\xi^- = (1-(1-2P_0)(1-2P_i)^{d_c-2})/2$, $\eta^+ = (d_c-2)(1-2P_o)(1-2P_i)^{d_c-3}$, and $\eta^- = -(d_c-2)(1-2P_o)(1-2P_i)^{d_c-3}$.

Without loss of generality, we may assume that P_0 and P_i are restricted in the interval [0, 0.5]. Then, we observe that (6) is always non-negative, i.e., $\partial P_{i+1}/\partial P_i \ge 0$. This shows that (2) is a monotone increasing function of P_i. Therefore, by substituting $P_i = 0$ into (2), we can obtain a lower bound expression of P_{i+1}.

The weight m used in the lower bound expression can be determined via (5). By substituting $P_i = 0$ into (5), we have

$$\frac{1-P_0}{P_0} \le \left(\frac{1+(1-2P_0)}{1-(1-2P_0)}\right)^{2m-d_v+1}. \quad (9)$$

If P_0 is restricted within the interval [0, 0.5], then $(1-P_0)/P_0$ is not less than 1. The inequality is satisfied if and only if the exponent of the right hand side is greater than 1, i.e., $2m-d_v+1 \ge 1$. The smallest integer that satisfies this inequality is $m = \lceil d_v/2 \rceil$, where $\lceil\cdot\rceil$ is the ceiling operation. Notice that, at the last iteration, the number of available extrinsic messages for an MV node is d_v, instead of d_v-1. Hence, we choose the weight m^* for a given (d_v, d_c) regular systematic LDGM as

$$m^* = \left\lceil \frac{d_v+1}{2} \right\rceil. \quad (10)$$

Therefore, the lower bound expression, which is only a function of the degree of variable nodes d_v, is given by

$$P_{LB} = P_0\left(1-\sum_{l=m*}^{d_v}\binom{d_v}{l}^l(1-P_0)^l(P_0)^{d_v-l}\right) + (1-P_0)\left(\sum_{l=m^*}^{d_v}\binom{d_v}{l}^l(P_0)^l(1-P_0)^{d_v-l}\right). \quad (11)$$

We note that the results (2) and (11) are based on the cycle-free assumption, i.e., infinite code length. Thus, they render the best performance for given degrees of systematic LDGM codes. The tightness of the lower bound is illustrated through simulation in Section 4.

4. Simulation results and discussion

Assuming BPSK modulation over AWGN channels, the error probability of the equivalent binary symmetric channel (BSC) for the MB algorithm is obtained by $P_0 = 0.5erfc(\sqrt{RE_b/N_0})$, where R is the code rate, E_b is the energy per bit, and N_0 is the one sided power spectral density of the noise. We assess the best possible performance of the code by using the recursion form (2) in the following manner. While numerically evaluating the recursion expression (2), we test out all the possible choices of m in each iteration and then select the best value of m that results in the lowest error probability at the end of each iteration. In addition, we let a large number of iterations (more than 50) to ensure the convergence of the recursive form (2).

Fig. 1(a) shows the BER curves for rate half systematic LDGM codes with degree (8, 9), (9, 10), (10, 11), (11, 12), (12, 13), and (13, 14). Fig. 1(b) shows the BER curves for rate around 1/3 systematic LDGM codes with degree (9,6), (10,6), (11,6), and (11,7). The dashed curves in the figure are obtained from the recursive method (2), whereas the solid curves are obtained from the non-recursive lower bound (11). We note that the lower bound is asymptotically tight with respect to channel signal to noise ratio (SNR). This is expected because, at high SNR, P_i in (2) can evolve to a value very close to zero, and hence the assumption $P_i = 0$ we made to derive the lower bound becomes more valid. We also note that the lower bound expression predicts the performance well in the entire error floor region. Defining threshold as the SNR the waterfall starts, moreover, we note that a code with small degrees exhibits a high error floor but a small threshold; whereas a code with large degrees shows a low error floor but a larger threshold. Considering the trade-off relation between the error floor and the threshold behavior, the best degree d_v for systematic LDGM codes under MB iterative decoding algorithm can be selected. For example, we may select it to be 10 based on our results. Systematic LDGM codes with other degrees are not good, since they exhibit either a high error floor or a large threshold. In addition, we observe that, for rate half systematic LDGM codes, the curves of (d_v, d_v+1) and (d_v+1, d_v+2) converge asymptotically for even d_v. This is one of the characteristic behaviors of systematic LDGM codes which was also reported in [2].

Fig. 2 shows the BER curves obtained from the Monte Carlo computer simulation. Ten iterations are used in the MB algorithm. Two randomly constructed (8, 9) and (9, 10) systematic LDGM codes with length 6000 are used in simulation.

To draw best threshold behavior while maintaining a low error floor, we use the following strategy for selecting

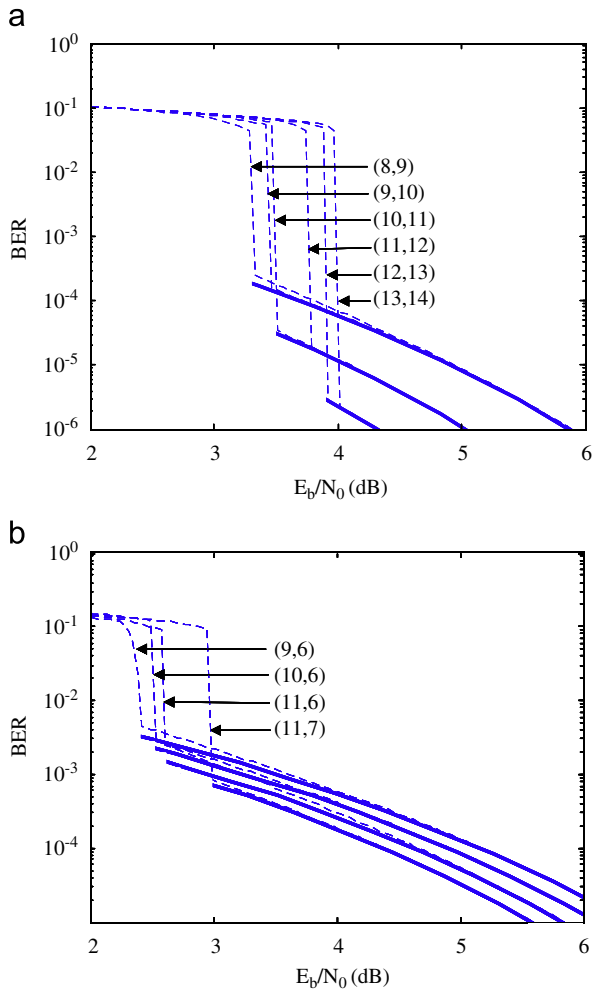


Fig. 1. BER performance of (a) rate half systematic LDGM codes and (b) rate around $\frac{1}{3}$ systematic LDGM codes. In the figure, the dashed curves represent the BER obtained from the recursive expression (2) with dynamic optimized weight m; the solid curves represent the BER obtained by the non-recursive lower-bound expression (11).

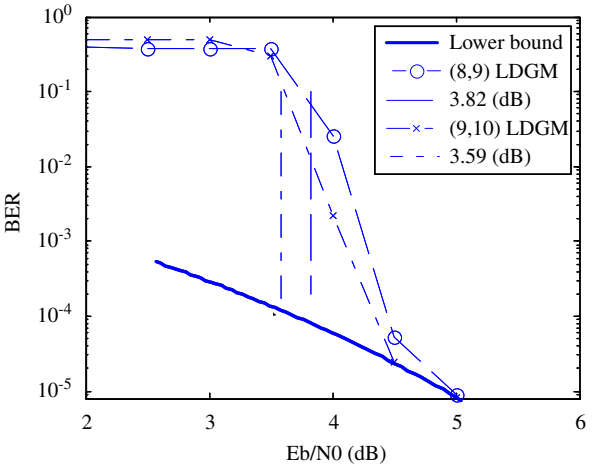


Fig. 2. Comparison between the Monte Carlo computer simulation results and the analytic results.

the weight m. Initially, we choose the weight m which has the smallest threshold. This initial weight is used all the way through the last iteration, and at the last iteration the weight calculated from (10) is used to push down the error floor. For the (8, 9) systematic LDGM code, the weight of the smallest threshold is $m = 5$ and the weight calculated from (10) is also $m = 5$. Hence, we select the weight to be 5 throughout the iterations. For the (9, 10) systematic LDGM codes, the weight of the smallest threshold is $m = 6$, whereas the weight calculated from (10) is $m = 5$. We choose $m = 6$ for the iterations all the way until the last one, and then choose $m = 5$ for the last iteration.

The simulation results show that both the (8, 9) systematic LDGM code and the (9, 10) systematic LDGM code not only can achieve the lower bound, but also can achieve the thresholds. In other words, the derived recursive expression and non-recursive lower bound are tight and can successfully serve as efficient tools to access the error performance of the codec.

5. Conclusion

Systematic LDGM codes and the majority-rule based iterative decoding algorithm may be of interest for communications system engineers because they render simple encoding and decoding complexities. The codec exhibits two eminent error performance behaviors, the threshold and the error floor. We have provided the analytic expression (2) for efficiently assessing the best possible performance of the codec for given degrees of systematic LDGM codes. Furthermore, a simple tight lower bound expression (11) has been derived, which can be readily evaluated once the degree of variable nodes is given.

References

[1] J.F. Cheng, R.J. McEliece, Some high-rate near capacity codes for the Gaussian channel, in: Proceedings of 34th Allerton Conference on Communication, Control and Computing, Monticello, IL, October 1996.
[2] J. Garcia-Frias, W. Zhong, Approaching Shannon performance by iterative decoding of linear codes with low-density generator matrix, IEEE Communication Letters 7 (6) (2003) 266–268.
[3] L. Ping, S. Chan, K.L. Yeung, Iterative decoding of multi-dimensional concatenated single parity check codes, in: Proceedings of IEEE International Conference on Communication, Atlanta, GA, June 1998, pp. 131–135.
[4] M.G. Luby, M. Mitzenmacher, M.A. Shokrollahi, D.A. Spielman, V. Stemann, Practical loss-resilient codes, in: Proceedings of 29th ACM Symposium on Theory of Computing, El Paso, TX, May 1997, pp. 150–159.
[5] X. Bao, J. Li, Matching code-on-graph with network-on-graph: adaptive network coding for wireless relay networks, in: Proceedings of 43th Allerton Conference on Communication, Control, and Computing. Champaign, IL, September 2005.
[6] W. Zhong, J. Garcia-Frias, LDGM codes for channel coding and joint source-channel coding of correlated sources, EURASIP Journal on Applied Signal Processing 2005 (1) (2005) 942–955.
[7] P. Zarrinkhat, A.H. Banihashemi, Threshold values and convergence properties of majority-based algorithms for decoding regular low-density parity-check codes, IEEE Transactions on Communication 52 (12) (2004) 2087–2097.
[8] L. Bazzi, T. Richardson, R. Urbanke, Exact thresholds and optimal codes for the binary symmetric channel and Gallager's decoding

algorithm A, IEEE Transactions on Information Theory 50 (9) (2004) 2010–2021.
[9] R.G. Gallager, Low-Density Parity-Check Codes, MIT Press, Cambridge, MA, 1963.
[10] T. Richardson, R. Urbanke, Modern Coding Theory, Cambridge University Press, Cambridge, 2008.
[11] R.M. Tanner, A recursive approach to low complexity codes, IEEE Transactions on Information Theory 27 (5) (1981) 533–547.

Date of submission xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2023.0322000

Enhancing Zero-Shot Crypto Sentiment with Fine-tuned Language Model and Prompt Engineering

RAHMAN S M WAHIDUR[1], ISHMAM TASHDEED[2], MANJIT KAUR[3], (Senior Member, IEEE), and HEUNG-NO-LEE [1], (Senior Member, IEEE)

[1]School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea
[2]Department of Computer Science and Engineering, Islamic University of Technology, Dhaka, Bangladesh.
[3]School of Computer Science and Artificial Intelligence, SR University, Warangal, Telangana, India-506371.

Corresponding author: Heung-No Lee (heungno@gist.ac.kr).

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2023-2021-0-00118, Development of decentralized consensus composition technology for large-scale nodes) and This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2023-2021-0-01835) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation)

arXiv:2310.13226v1 [cs.CL] 20 Oct 2023

ABSTRACT Blockchain technology has revolutionized the financial landscape, with cryptocurrencies gaining widespread adoption for their decentralized and transparent nature. As the sentiment expressed on social media platforms can significantly influence cryptocurrency discussions and market movements, sentiment analysis has emerged as a crucial tool for understanding public opinion and predicting market trends. Motivated by the aim to enhance sentiment analysis accuracy in the cryptocurrency domain, this paper investigates fine-tuning techniques on large language models. This paper also investigates the efficacy of supervised fine-tuning and instruction-based fine-tuning on large language models for unseen tasks. Experimental results demonstrate a significant average zero-shot performance gain of 40% after fine-tuning, highlighting the potential of this technique in optimizing pre-trained language model efficiency. Additionally, the impact of instruction tuning on models of varying scales is examined, revealing that larger models benefit from instruction tuning, achieving the highest average accuracy score of 75.16%. In contrast, smaller-scale models may experience reduced generalization due to the complete utilization of model capacity. To gain deeper insight about how instruction works with these language models, this paper presents an experimental investigation into the response of an instruction-based model under different instruction tuning setups. The investigation demonstrates that the model achieves an average accuracy score of 72.38% for short and simple instructions. This performance significantly outperforms its accuracy under long and complex instructions by over 12%, thereby effectively highlighting the profound significance of instruction characteristics in maximizing model performance. Finally, this paper explores the relationship between fine-tuning corpus size and model performance, finding an optimal corpus size of 6,000 data points for achieving the highest performance across different language models. Significantly, a distillation model of BERT called MiniLM, published by the Microsoft research team, stands out for its exceptional data efficiency, effectively optimizing its performance while making efficient use of data. Conversely, Fine-tuned LAnguage Net, FLAN-T5 developed by the Google research team, impressively maintains consistent and reliable performance across diverse corpora, further affirming its robustness and versatility.

INDEX TERMS Zero-Shot Learning, In-context Learning, Supervised fine-tuning, Instruction Tuned, Prompt Engineering.

I. INTRODUCTION

IN the recent decade, cryptocurrency has gained much traction in finance and business for its decentralization, permissionless, and open nature built on the public blockchain [1]. This architecture can create an immutable and highly interoperable financial system with unprecedented trans-

parency, equal access rights, and little need for central authorities that smart contracts can control [2]. Cryptocurrencies employ cryptographic ciphers to facilitate financial transactions, distinguishing them from traditional forms of currency [3]. Cryptocurrencies are the first pure digital assets to be included by asset managers [4]. They mitigate double-spending by leveraging multiple verifications from neighboring nodes within the blockchain network. As the number of confirmations grows, the transaction gains enhanced reliability and become highly irreversible. Because of these favorable attributes and the widespread accessibility of cryptocurrencies, they can be a means of transaction and a store of wealth [4], [5].

In cryptocurrency discussions, social media networks have gained significant importance as information-sharing platforms. Communicating crypto events and market conditions through social media is widely recognized. Moreover, there is a prevalent belief in the correlation between altcoin prices and the sentiment expressed within the Twitter community [6]. The demand for cryptocurrency is intricately tied to people's trust in Bitcoin and its underlying technology. As people's trust plays a significant role in the growth of the cryptocurrency market, the sentiment of the general population has a substantial impact on the future market capitalization of cryptocurrencies [7]. The effect of social media on crypto discourse is growing [8]. The sentiment expressed on social media platforms can greatly influence cryptocurrency discussions, shaping public opinion, driving market movements, disseminating information, and generating buzz. In 2019, over 300 million active monthly users shared their emotions in multiple languages on various social media platforms, including Twitter, Facebook, and YouTube. Among these platforms, Twitter has emerged as one of the most influential social media networks [9], [10]. Twitter stands out as a unique platform due to its straightforward way of gauging individuals' sentiments toward a text. Performing real-time data analysis is also possible on the Twitter platform [11].

Sentiment analysis is a sub-research area of computational Natural Language Processing (NLP) studies and a widely used contextual mining technique for extracting valuable and subjective information from text-based data [6], [12]. There are several approaches to sentiment analysis, which is the process of determining the sentiment or emotion expressed in the text based on the type and size of the corpus. The rule-based approaches use a set of predefined rules, such as regular expressions and dictionaries, to classify text as positive, negative, or neutral [12]. Machine learning approaches try to find patterns from the provided data. This approach involves training a machine learning model on a labeled dataset to classify new text. These machine-learning techniques are further expanded over supervised and unsupervised. Wherein supervised approaches are trained on an annotated dataset, unsupervised does not require work on experience to improve accuracy [13]. However, pre-trained models can be Fine-tuned on specific datasets, reducing the need for labeled data and computational resources [14].

In text mining, sentence-level sentiment analysis has emerged as a burgeoning area of research. This analytical process encompasses six fundamental stages: data collection, data preprocessing, feature extraction, model training, model evaluation, and model deployment. Feature extraction, in particular, holds significant importance in optimizing the model's efficacy. Conventionally, Bag-of-Words (BoW) [15] and N-gram [16] techniques are widely employed for this purpose. However, their one-hot word representation approach creates high-dimensional feature spaces and scalability challenges, failing to capture word sequences and their syntactic and semantic nuances. Word embedding models like word2vec [17] and Glove [18] have gained popularity for their ability to capture word semantics in high-dimensional spaces, although they require substantial training data and are susceptible to data sparsity issues for rare or out-of-vocabulary (OVV) [12] words, which can impact performance.

To address the abovementioned limitations, a transformer-based NLP model was proposed [19] in the paper "Attention is All You Need." The transformer architecture is based on self-attention, which allows the model to weigh the importance of different words in a sentence when making predictions. In short, it creates contextual awareness between the words of a sentence. BERT [20], T5 [21], GPT-3 [22], and LLaMA [23] are a few examples among all pre-trained transformer-based models for NLP tasks. All these models have shown remarkable performance on various NLP tasks, especially in few-shot learning. However, they are less successful in zero-short learning [24]. This paper explores a simple but powerful method called instruction tuned to improve the performance of zero-shot learning of Large Language Models (LLMs) for cryptocurrency sentiment classification tasks. Moreover, this paper utilizes an In-context learning (ICL) and prompt engineering method to generate effective instructions using LLMs that allow the creation of effective instructional datasets for fine-tuning the pre-trained language models to extract the cryptocurrency sentiment from social media data.

The key contributions of this paper are summarized as follows:

- **Improved Model Efficiency through fine-tuning:** The study demonstrates that supervised fine-tuning and instruction tuning significantly enhance pre-trained language models' performance on unseen tasks related to cryptocurrency sentiment analysis. This research experimentally shows that the average accuracy scores significantly increased after fine-tuning. The findings provide compelling evidence that fine-tuning enhances model efficiency, with an average performance gain of 40%. This contributes to the practical application of fine-tuning as a powerful tool for optimizing pre-trained language model performance.
- **Benefits of Instruction Tuning and Model Scale:** Building upon previous research, the study investigates the impact of instruction tuning on different-sized lan-

guage models. By analyzing FLAN-T5 models of varying scales, it was discovered that larger models benefit from instruction tuning by improving generalization to new tasks. However, for small-scale models, instruction tuning had a detrimental effect on generalization, potentially due to the complete utilization of model capacity for learning the mixture of instruction tuning tasks. This observation enhances our understanding of the interplay between instruction tuning and model scale, providing valuable insights for future model development and deployment.

- **Impact of model performance under different instruction tuning setups:** The conducted experimentation aimed to compare the quality of models under different instruction tuning setups, specifically focusing on the response of the instruction-based model. By introducing diverse instructions of varying lengths and complexities, the study provided insights into the model's handling of different instruction types, revealing its effectiveness in understanding and executing short and simple instructions compared to long and complex ones, thereby emphasizing the importance of considering instruction characteristics in instruction tuning setups.
- **Impact of fine-tuning Corpus Size on Model Performance:** The research explores the relationship between the size of the fine-tuning corpus and the performance of language models. By varying the sample size, the study investigates the influence of data availability on model performance. The findings highlight the optimal corpus size of 6000 data points for achieving the highest performance for each model and the data efficiency of MiniLM in leveraging limited data.

The remaining paper is organized as follows: Section II provides background information and reviews of related works. Section III presents the comprehensive architecture of the proposed model. Section IV showcases the analysis and metrics of the experimental results used for evaluation. Finally, in Section V, the paper concludes and summarizes the findings while also discussing the potential areas for future research and research limitations.

II. RELATED WORK

Recently, researchers have perceived a growing interest in leveraging sentiment analysis techniques for cryptocurrency. Several models are available for sentiment acquisition, each with varying precision and applicability [25]. Hasan et al. [26] investigated public sentiment analysis by employing a machine learning algorithm, namely the Support Vector Machine. The researchers also utilized Chi-square for feature selection to mitigate sentence noise. A similar approach was used by Satrya et al. [27] to determine the polarity of the sentiment. Additionally, they used TF-IDF weighting to transform the data from text to numeric values. The methodology employed by Padmalatha et al. [28] is based on Naive Bayes model to analyze social media opinions. Prasad et al. [29] designed an ensemble classifier to classify YouTube comments based on cryptocurrency. They used Decision Tree, K Nearest Neighbors, Random Forest Classifier, XGBoost, and a Logistic Regression base classifier to create a stacked ensemble model. Sasmaz et al. [6] studied the feasibility of automated sentiment analysis for cryptocurrencies using the Random Forest Classifier.

The Valence Aware Dictionary for Sentiment Reasoning (VADER), a widely utilized and straightforward sentiment calculation model, is commonly employed to predict cryptocurrency prices by analyzing cryptocurrency-related news and Tweets. Suardi et al. [30] used VADER to investigate the predictive power of information contained in social media tweets on bitcoin market dynamics. The extent to which Twitter sentiment analysis can predict price fluctuations for cryptocurrencies was examined by Oikonomopoulos et al. [31]. They employed VADER for sentiment analysis in their study. Jagini et al. [32] intend to analyze the effect of tweets on the stock price of Bitcoin. To calculate the associated sentiment, they also used VADER. Parekh et al. [33] proposed a hybrid and robust DL-Gues framework for cryptocurrency price prediction. They also utilized a similar VADER technique in their framework to extract the polarity from the Twitter sentiment.

Regarding the recent superior performance transformers-based model, Dwivedi et al. [34] utilized the BERT (Bidirectional Encoder Representation) to predict the sentiments of cryptocurrency news articles. Kim et al. [35] introduced CBITS, a Fine-tuned version of BERT designed explicitly for cryptocurrency sentiment analysis, mainly focusing on the Korean crypto market. Widianto et al. [36] have created a BERT model to assess sentiment analysis on cryptocurrency and NFT by utilizing data crawling and pre-processing using Rapiminer. The findings of Ortu et al. [37] indicate that incorporating features derived from BERT-based emotion classification of comments on GitHub and Reddit results in a notable improvement in the predictability of Bitcoin and Ethereum's hourly and daily return direction. Despite extensive exploration of sentiment extraction methods, including rule-based and machine learning-based approaches, they often prove inadequate in cryptocurrencies due to domain-specific jargon, slang, and sentiment ambiguity not adequately covered by general-purpose texts [35]. Rule-based systems, although less accurate, are limited by the need for more powerful linguistic resources [12]. On the other hand, machine learning-based methods tend to achieve higher accuracy but require large amounts of labeled training data and significant computational resources. Existing literature reveals a notable gap in utilizing large language models for sentiment extraction. Additionally, there is a lack of research using fine-tuning techniques, despite their potential to enhance performance and adaptability.

This study proposes integrating large language models and utilizing fine-tuning techniques in sentiment extraction to address these research gaps. By leveraging the capabilities of large language models and optimizing their performance

through fine-tuning, this research aims to overcome the limitations of current approaches and advance the accuracy and applicability of sentiment extraction, especially in cryptocurrency. This contribution expects to facilitate the development of more effective and comprehensive sentiment analysis methodologies, ultimately enhancing decision-making processes in related industries and domains.

III. PROPOSED SYSTEM

This section illustrates and describes the entire procedure adopted in this paper. It is subdivided into five sub-sections. Subsection III-A summarizes the proposed model along with its visualization. The overview of the core concept related to the proposed model is discussed in the following subsections.

A. PROBLEM FORMULATION

Let X be the set of tweets related to cryptocurrency where each tweet $x \in X$ is represented as a feature matrix $x \in \mathbb{R}^{1\times n}$ thus, X can be represented as

$$X \in \mathbb{R}^{m\times n} \tag{1}$$

where m is the number of tweets and n is the number of features used to represent each tweet. Let Y be a set of sentiments represented as a target matrix

$$Y \in \mathbb{R}^{m\times 1} \tag{2}$$

where m is the number of tweet labels. For this experiment, only *positive, negative* sentiment tweets are considered. Thus dataset D can be represented as

$$D = (X^{(i)}, Y^{(i)} | X^{(i)} \in \mathbb{R}^{(1\times n)}, Y^{(i)} \in \mathbb{R}^{(1\times 1)}) \tag{3}$$

which consisting of pairs $(X^{(i)}, Y^{(i)})$. We can then define the model

$$f : \mathbb{R}^{(m\times n)} \rightarrow \mathbb{R}^{(m\times 1)} \tag{4}$$

that maps each tweet matrix $X \in \mathbb{R}^{(m\times n)}$ to its sentiment matrix $Y \in \mathbb{R}^{(m\times 1)}$. The goal is to train the model f on the data set D such that

$$\forall x \in X : f(x, \Theta) \in Y \tag{5}$$

where Θ is the model parameters. We can use various mathematical techniques, such as gradient descent, backpropagation, and cross-entropy loss, to train/fine-tune the model f on D. Moreover, We can use various optimization algorithms, such as stochastic gradient descent or Adam, to find the optimal set of parameters Θ^* that minimizes the cross-entropy loss. The loss function can be defined as:

$$\begin{aligned} L(\Theta) = -\sum_{i=1}^{m} y^{(i)} log(f(x^{(i)}, \Theta)) + \\ (1 - y^{(i)}) log(1 - f(x^{(i)}, \Theta)) \end{aligned} \tag{6}$$

where m is the number of training examples, $x^{(i)}$ is the ith tweet vector in the training set, $y^{(i)}$ is the corresponding sentiment label (either 0 or 1 for negative or positive sentiment, respectively), and $f(x^{(i)}, \Theta)$ is the predicted sentiment score for $x^{(i)}$ given the current set of parameters Θ. We aim to minimize the cross-entropy loss $L(\Theta)$ by finding the optimal set of parameters Θ^* that maximizes the likelihood of the training data D given the model.

Once the model is trained/fine-tuned, we can use it to predict the sentiment of new tweets related to cryptocurrency by feeding their feature matrix representations into the model and obtaining the predicted sentiment matrix as output. Mathematically the final model can be represented as

$$M_{\text{eval}} = (f_{\text{tuned}}, D_{\text{test}}) \tag{7}$$

where f_{tuned} is the trained/Fine-tuned model and D_{test} is the evaluation dataset.

B. PROPOSED SYSTEM ARCHITECTURE

The design overview of our proposed model is shown in Figure 1. The proposed architecture begins with the initial user interaction step and prompt generation. The user engages with the OpenAI "text-davinci-003" model, initiating the prompt generation process. The users provide contextual information using prompt templates, which input the subsequent stages. The provided text is transmitted to the language model's backend through an API call, facilitating communication between the users and the large language mode. Next, a content moderator component is employed to evaluate the content. If the content is determined to be unsafe, the model responds with a default message. In the case of safe content, the system proceeds through a safety gateway to engage with the core parts of the large language model. This model encompasses several key components, like a response generator, context processor, knowledge domain, and system response generator. The system leverages its knowledge domain through interaction with these components to generate an effective response based on the provided input. The resulting response, an AI-generated prompt, is then delivered to the user. Upon receiving the AI-generated prompt, a filtering process takes place. Human feedback is vital in determining whether the prompt should progress to the subsequent stages.

The primary objective of this filtering is to generate adequate instructions by harnessing the advantages of in-context learning inherent in large language models and to prune low-quality and repeated instructions before adding them to the task pool. Finally, the prompt generation is refined by incorporating user feedback and capitalizing on the model's contextual understanding to enhance its efficacy. The process involving user interaction, prompt generation, filtering, and human feedback collectively constitutes the in-context learning and prompt engineering phase. The overall process of generating instructions using in-context learning can be seen in Algorithm 1. This phase aims to iteratively improve the prompt generation procedure by leveraging user feedback and maximizing the model's contextual comprehension capabilities.

Following the generation of effective instructions, the subsequent objective is to create an augmented dataset. This dataset is constructed by concatenating the introductions with

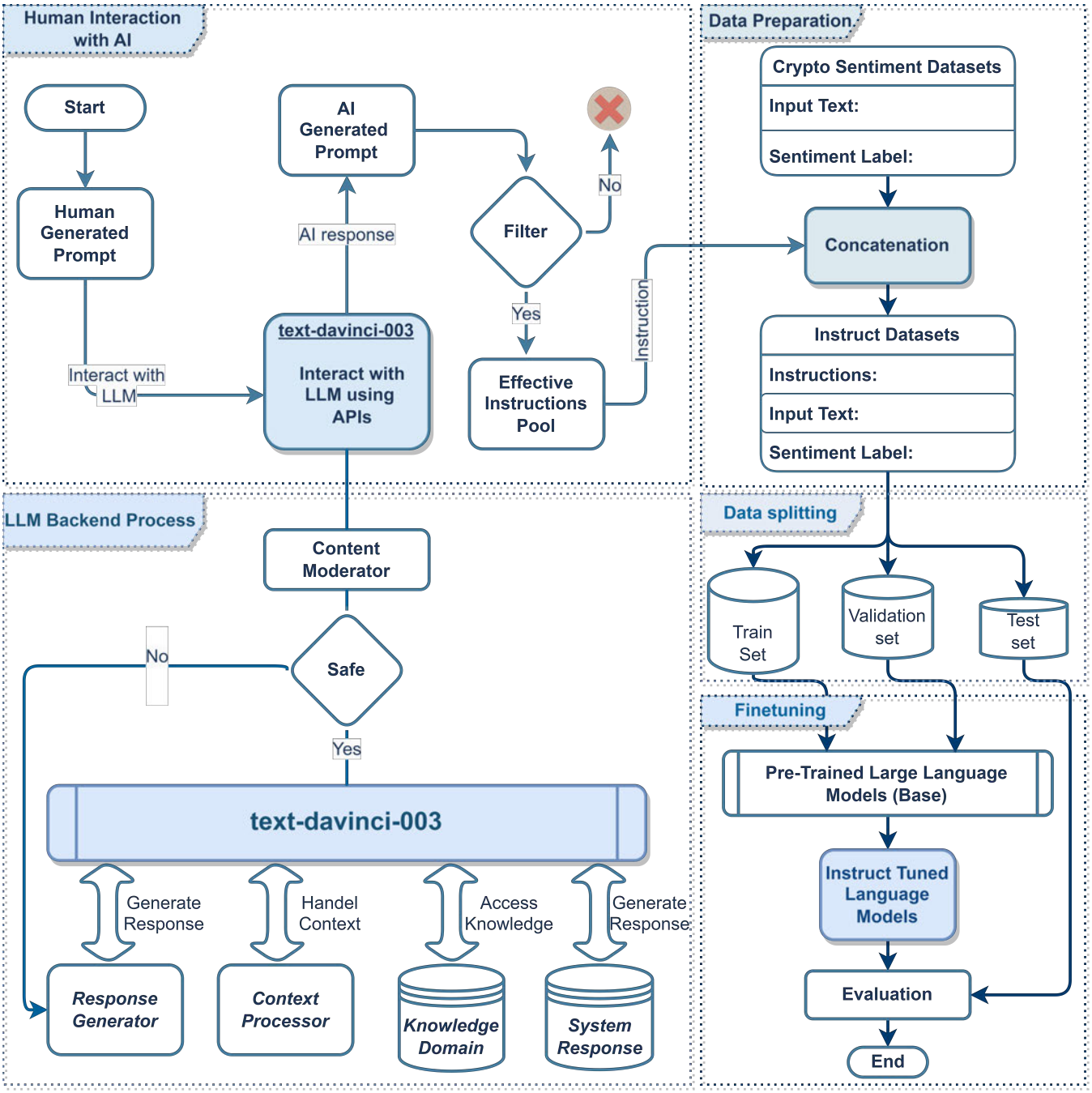


FIGURE 1. The Proposed System Model.

the original dataset, ensuring a comprehensive collection of relevant information for training purposes. Algorithm 2 shows the process of generating an augmented crypto sentiment dataset. The instructional dataset is then divided into separate training and validation sets. Three additional datasets are withheld to measure zero-shot performance, remaining untouched for evaluation. The large language models undergo fine-tuning at this stage, utilizing the instructional dataset. The initial model weights are modified, resulting in a newly instructed, Fine-tuned model version. The overall fine-tuning process of a large language model can bee observed in Algorithm 3. After the fine-tuning process, the performance of various models is evaluated. This evaluation provides valuable insights into the efficacy and effectiveness of the in-context learning and prompt engineering techniques employed in instructing fine-tuning for Zero-shot learning.

C. PRE-TRAINED LANGUAGE MODELS (PLMS)

Recently, there has been a significant emphasis on pre-trained language models (PLMs) that utilize self-supervised learning

Algorithm 1 Generating instructions using in-context learning.

Constant model, model, temp, max_len, top_p, penalty
Input A human-generated prompt $p \in P$.
Output An effective instructions pool Y.

Define the Language Model (LLM) as a function $G : P \rightarrow R$ that maps a given prompt $p \in P$ to a model-generated response $r \in R$.
$r \leftarrow G(p; mode, model, temp, max\_len, top\_p)$
Pass the response through a filter $F : R \rightarrow (True, False)$
for response $r \in R$ **do**
  if $F(R) = True$ **then**
    Add r to the effective instructions pool Y.
  else
    Discard the response r.
  end if
end for
Return Effective instructions pool Y

Algorithm 2 Generating augmented crypto sentiment dataset.

Input Effective instructions pool Y, and Crypto sentiment dataset X.
Output Augmented crypto sentiment dataset $X_{augmented}$.

Let function $C : x \rightarrow x$ which cleans a dataset entry.
$X_{filtered} \subseteq X$ which contains only non-neutral sentiments.
for $i \leftarrow (1 \ldots len(X_{filtered}))$ **do**
  $X^{(i)}_{cleaned} \leftarrow C(X^{(i)}_{filtered})$.
end for
Select an instruction $y \in Y$.
Let function $A : (x, y) \rightarrow x$ which augments each entry $x \in X$ with an instruction y.
for $i \leftarrow (1 \ldots len(X_{cleaned}))$ **do**
  $X^{(i)}_{augmented} \leftarrow A(X^{(i)}_{cleaned}, y)$.
end for
Return Augmented dataset $X_{augmented}$

Algorithm 3 Fine-tuning of a pre-trained language model.

Constant random_seed, input sequence size, number of layers, number of hidden layer nodes, number of classifier outputs.
Input Training set $(X^{(1)}_{train}, Y^{(1)}_{train}), \ldots, (X^{(M)}_{train}, Y^{(M)}_{train})$.
Output Trained language model network parameters..

for *epoch* $\in$ *epochs* **do**
  for *batch* $\in$ *batches* **do**
    for $i \leftarrow (1 \ldots len(batch))$ **do**
      $M^{(i)}_{train} \leftarrow tokenizer(X^{(i)}_{train})$
      $\hat{M}^{(i)}_{train} \leftarrow model(M^{(i)}_{train})$
    end for
    $loss \leftarrow E(Y^{(batch)}_{train}, Y^{(batch)}_{train})$
    Calculate $\nabla\Theta$ for backpropagation.
    Adjust parameters using an optimizer to minimize the loss.
  end for
end for
Return Task-specific Fine-tuned model.

on extensive raw text data [38]. Notable examples of such models include GPT-3 [22], PaLM [39], Chinchilla [40], LLaMA [23], and Falcon 40B [41]. By training on large-scale texts using self-learning tasks like masked word prediction, sentence sequence recognition, text completion, and text generation [19], [42] PLMs acquire a comprehensive understanding of language. In addition, these models enhance the semantic representation of words by considering contextual dynamics and provide a unified framework for various NLP tasks. Currently, there are three standard models [43] structures in PLMs: autoregressive language models, autoencoding language models, and hybrid language models. Representative models for each design are GPT [22], BERT [20], and T5 [21], respectively. Autoregressive language models follow a standard approach where language modeling is done decoder-only, predicting words one by one through one-way language encoding-decoding and token-by-token prediction of words. Autoencoding language models randomly mask words in a sentence, use bidirectional encoding to capture context, and then predict the masked words based on the encoded information. Finally, hybrid language models combine the approaches of the previous two models. They mask words randomly in a sentence, apply bidirectional encoding, and predict subsequent words step by step by inputting the earlier text in one direction [43].

The advancement of artificial intelligence technology has demonstrated that LLMs can acquire a deeper understanding of language and exhibit stronger capabilities in understanding and generating data. These models learn abstract knowledge from raw data, resulting in better generality and generalization. The autoregressive language model adopted by GPT-3 and its subsequent models, such as GPT3.5 and GPT4, has proven advantageous in utilizing natural language for various tasks in different fields [43]. Initially, it was believed that increasing the number of parameters in models would lead to better performance. However, recent research by Hoffmann et al. [23] has shown that smaller models trained on more data can achieve the best performances given a specific computing budget.

According to Figure 2, the data analysis shows a clear trend of increasing pre-trained token usage over the years. In the initial years (2020-2021), models exhibited relatively low token counts during the pre-training phase. However, in recent years (2022-2023), there has been a notable surge in the number of pre-trained tokens utilized by language models. Such expansion signifies model developers' recognition of the benefits of leveraging more extensive and diverse pre-training corpora, enabling improved contextual understanding and enhanced performance in downstream tasks.

Contrary to the trend observed in pre-trained token usage, model parameter sizes exhibit a different pattern. In 2021 and 2022, models with considerably large parameter sizes emerged. However, in 2023, a noticeable decrease in model parameter sizes is observed. This shift suggests a growing focus on optimizing computational efficiency and addressing the resource-intensive nature of large models. As a result, model developers are actively exploring methods to achieve comparable performance with fewer parameters, which may reduce computational costs and carbon footprint.

According to Figure 3, language models can categorize into three distinct clusters: Models with low token usage and small parameter sizes, representing the majority, are suitable for resource-constrained environments. Models with substantial token counts but relatively small parameter sizes exhibit an exciting trade-off, leveraging massive amounts of pre-training data while keeping the parameter sizes reasonably small. Models with low token usage and large parameter sizes prioritize performance and strike a balance between computational resources and model capacity. The observed

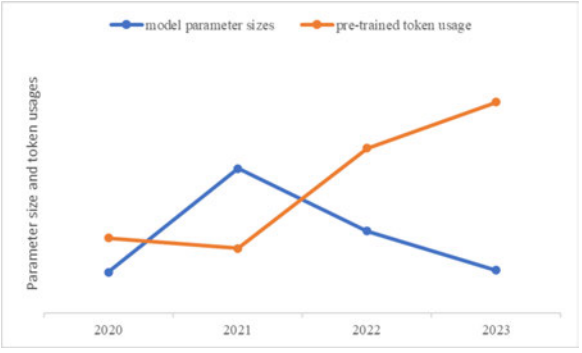


FIGURE 2. Yearly trend analysis of model parameter size and token usage.

trends in language model development hold significant implications for researchers and practitioners. The increasing usage of pre-trained tokens emphasizes the importance of diverse and extensive training data for capturing the degree of language understanding. Conversely, the fluctuation in model parameter sizes highlights ongoing efforts to balance model capacity and computational efficiency. Further research is required to explore novel techniques and architectures that optimize the trade-offs in token usage and parameter sizes.

D. FINE-TUNING OF LARGE LANGUAGE MODELS (LLMS)

LLMs have exhibited exceptional abilities in various NLP tasks [39], [44], [45]. Nevertheless, these models can sometimes display unintended behaviors, such as generating false information, pursuing inaccurate objectives, and producing harmful, misleading, and biased expressions [46], [47]. In the pre-training stage, pre-trained language models acquire non-task-specific language knowledge. The subsequent fine-tuning stage facilitates task-specific adjustments of the model, enabling it to be utilized for various downstream tasks [48]. There are two primary approaches [49] to adapt the pre-trained language models for target tasks: feature-based approach and fine-tuning. The feature-based approach involves loading a pre-trained LLM and utilizing it on a target dataset. The primary focus is creating output embeddings for the training set, which can be used as input features for a classification model. While this approach is often used for embedding-centric models like BERT, embeddings can also be extracted from generative GPT-style models like "text-embedding-ada-002". The classification model can be any desired model, such as a logistic regression model, random forest, or XGBoost. However, linear classifiers, specifically logistic regression, have demonstrated superior performance [49].

fine-tuning is essential for adapting pre-trained language models to perform specific tasks using labeled training data. Initially, a pre-trained language model is used as a starting point and then Fine-tuned on a task-specific dataset with labeled examples. This process is called supervised fine-tuning (SFT). This SFT process is necessary to apply PLMs to tasks like sentence classification, named entity recognition, and question-answering. Unlike pre-training, this fine-tuning

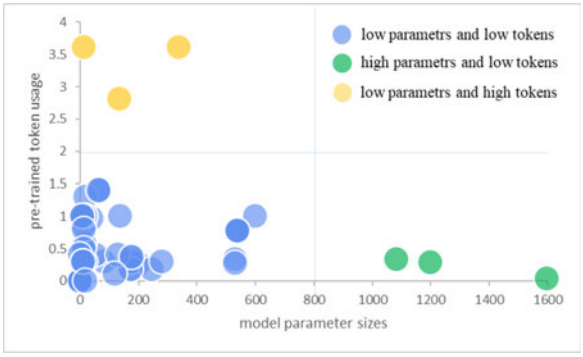


FIGURE 3. Cluster distribution analysis considering model parameters and token utilization.

requires less data, typically around 100k words [50]. During SFT, a task-specific layer is added to the PLMs, and the model parameters, including those of the task-specific layer, are updated through gradient descent using an appropriate loss function [49]. One advantage of the PLMs is the ability to freeze specific layers while fine-tuning the remaining ones, potentially improving performance. However, it has been observed that freezing too many layers may lead to poor performance [50]. The power and popularity of PLMs come from the fact that the pre-training process only needs to be done once in a task-agnostic manner. Subsequently, a simple and more cost-effective fine-tuning process is sufficient for each specific task. This is possible because the dataset size required for fine-tuning is considerably smaller, reducing time and resource requirements [51].

Another form of fine-tuning is instruction tuning (IT) – fine-tuning language models on a collection of datasets described via instructions. Fundamentally, instruction tuning involves fine-tuning pre-trained LLMs using a set of formatted instances in natural language form [24]. This approach is

closely related to SFT [46]. The initial step involves gathering or constructing instances formatted as instructions to initiate instruction tuning. Subsequently, these formatted instances are employed to finetune LLMs using supervised learning techniques, such as training with the sequence-to-sequence loss. Following instruction tuning, LLMs exhibit enhanced generalization capabilities towards unseen tasks [24], [52], [53], even within a multilingual context [54]. IT is a more efficient alternative to pre-training, as it relies on a moderate number of instances for training. This supervised training process introduces distinct optimization considerations compared to pre-training, including utilizing a sequence-to-sequence loss as the training objective and special attention to factors like smaller batch size and learning rate in the optimization configuration [53]. IT significantly impacts LLMs by improving performance across various models and enhancing task generalization by enabling LLMs to understand and follow natural language instructions [47], [55].

E. IN-CONTEXT LEARNING (ICL)

In-Context Learning (ICL) pertains to comprehending the context of the input information and leveraging it accurately to produce the intended output. It exhibits similarities to the human decision-making process, wherein individuals learn from analogy [56]. In LLM utilization, ICL was initially introduced as a distinctive prompting method, specifically alongside GPT-3 [44], and has since emerged as a prominent approach [43]. In contrast to supervised learning, which involves a training phase that utilizes backward gradients to modify model parameters, ICL does not engage in parameter updates but instead performs direct predictions on pre-trained language models. The effectiveness of ICL lies in its dependence on the existing knowledge of the model and its ability to decipher the concealed pattern present in the demonstrations, thus facilitating precise predictions. The approach holds significant potential in situations that demand swift adaptation to new tasks, as it prevents the need for extensive training periods [57]. The vanilla GPT-3 model shows considerable potential for ICL, as pre-training adaptation has been shown to enhance its capabilities [58]. Moreover, the efficacy of ICL is contingent upon various parameters such as the choice of prompting template, in-context examples, and their sequence [59]. Despite being plausible, the underlying mechanism of ICL remains obscure, and only a few studies have provided initial insights into its workings [45], [60].

The context window size in LLMs [43], such as GPT-4, determines the maximum amount of input text the model can consider when generating its output. With the release of GPT-4, the context window's size doubled. GPT-3 was limited to 2048 tokens. The context window for the GPT-4 API is 8195 tokens, and a 32K context window exists in the most significant model. However, its performance typically falls short of fine-tuning, as it needs to update the model's parameters for a specific task, which may limit its adaptability to task-specific nuances [49]. Moreover, there are potential risks associated with ICL. The risk of prejudice and misinformation is a significant concern, as the LLMs cannot fact-check the input provided as part of the prompt, resulting in the possibility of incorporating erroneous and biased information into any generated output, including fabricated news or blog posts [61].

F. PROMPT ENGINEERING

As technological progress continues to unfold, the significance of prompt engineering grows in parallel. The ascendancy of large language models, such as ChatGPT [43], necessitates a skillset that can proficiently engage with them. By furnishing appropriate prompts, adherence to prescribed norms and the automation of processes can be ensured, resulting in outputs that align with desired quality and quantity benchmarks. Like programming, prompts facilitate the customization of interactions with large language models, optimizing their utilization to meet specific requirements [62].

A prompt comprises a specific set of directives furnished to a LLM, enabling customization and refinement of the model's capabilities through programming [63]. In short, it is a text that provides context and instructions for LLMs to generate a response. Prompts are pivotal in facilitating LLMs to undertake extensive linguistic tasks, including language translation, text classification, text summarization, question answering, and even producing human-like and coherent text [61]. The use of prompting techniques in NLP tasks has been extensively studied, covering both zero-short and few-shot settings, as indicated by various research papers [64]–[67]. Current prompt-based models predominantly rely on Transformers [19]. Prompting is particularly prevalent in online demonstrations, where generative models are interacted with using Transformers as assistive agents. By leveraging prompt engineering, the challenging task of language understanding can be addressed with improved efficiency [68]. The 'show-and-tell' technique [69], where examples and instructions are provided within the prompt, is the most effective approach for obtaining the desired output from the LLMs. We provide an example prompt in our LLM to demonstrate how prompts for LLMs can be engineered to elicit desired outcomes. The structure of the prompt is presented in Figure 4. Prompt engineering possesses several notable advantages: Firstly, it enables the pre-training of the language model on extensive volumes of raw text. Secondly, by defining a new prompting function, the model demonstrates the capability for few-shot or even zero-shot learning, effectively adapting to new scenarios with limited or no labeled data [63].

IV. RESULT ANALYSIS

This result analysis section is subdivided into three subsections. Together, they describe the dataset, narrate the evaluation metrics, illustrate the detailed experimental result, and compare the performance of the models. The experiments were repeated multiple times to check for anomalies and to get rid of any bias.

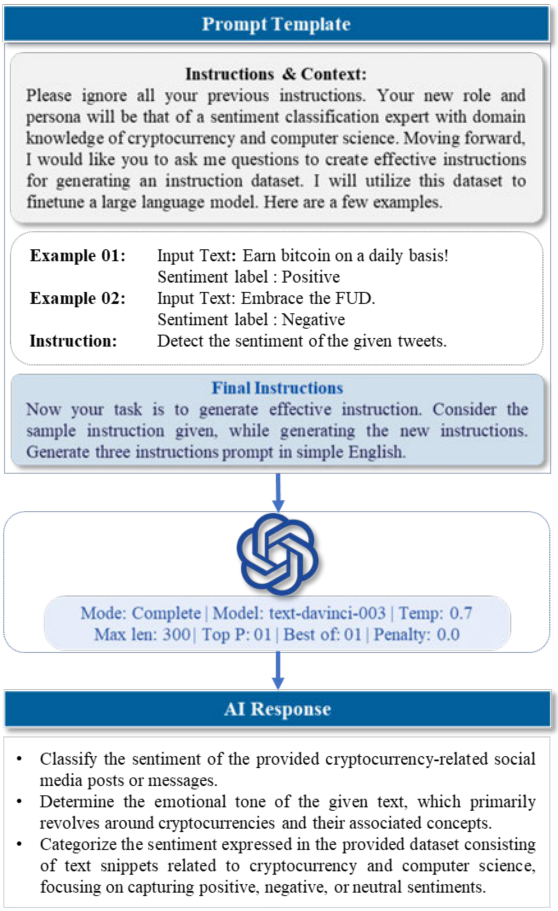


FIGURE 4. Prompt engineering template.

A. DATASET DESCRIPTION AND PRE-PROCESSING

Four datasets, namely the Neo, Reddit, Bitcoin sentiment, and Cryptocurrency sentiment datasets are utilized in the experimental analysis. These datasets are the foundation for conducting comprehensive investigations and analyses related to sentiment analysis in cryptocurrency. Table 1 shows the statistics.

The Neo dataset consists of tweets related to cryptocurrencies, specifically focused on the sentiment towards the Neo cryptocurrency. The dataset contains 12,000 tweets, evenly distributed between positive and negative emotions, with 6,000 tweets in each category. The Reddit dataset contains posts and comments extracted from the popular social media platform Reddit. The dataset focuses on the sentiment expressed towards various cryptocurrencies. It includes a total of 562 posts. Among these posts, 302 are classified as positive sentiment, while 260 are classified as negative sentiment. The Bitcoin sentiment dataset consists of tweets specifically related to the Bitcoin cryptocurrency. The dataset comprises 1,029 tweets, with 779 tweets classified as positive sentiment and 250 tweets classified as negative sentiment. It is important to note that this dataset exhibits an imbalance in sentiment distribution, with more positive than negative tweets taken intentionally for the experiments. The Cryptocurrency sentiment dataset is a collection of tweets that cover a wide range of cryptocurrencies. The dataset includes 500 tweets, evenly split between positive and negative sentiments, with 250 tweets in each category.

TABLE 1. Volume of all tweets and volume of tweets for each sentiment label and dataset.

Description	Volume	Percentage
All tweets	14,091	100.00%
Positive label	7,331	52.03%
Negative label	6,760	47.97%
Neo dataset	12,000	85.16%
Bitcoin sentiment dataset	1,029	7.30%
Reddit dataset	562	3.99%
Cryptocurrency sentiment dataset	500	3.55%

This study uses the Neo dataset to finetune the pre-trained language models. The other datasets, Reddit, Bitcoin sentiment, and Cryptocurrency sentiment datasets, are employed to evaluate the performance of the Fine-tuned models on unseen tasks. The concept of unseen tasks is defined based on prior work, which disallows the same dataset to appear during training. By separating the training and evaluation datasets in this manner, the study aims to provide a robust assessment of the model's ability to generalize and perform effectively on new, unseen data.

B. EVALUATION METRICS

To evaluate the effectiveness of the Fine-tuned models, accuracy (binary classification accuracy) and F1-score were selected as the evaluation metrics. The calculation of binary classification accuracy can be represented using the following equation:

$$\mathcal{T} = \frac{\theta + \emptyset}{\theta + \Lambda + \emptyset + \Psi} \tag{8}$$

where $\mathcal{T}$ represents the Accuracy (the percentage of the correct number predicted by the model to the total number of samples). θ corresponds to the true positive (the predicted result is positive, and the true result is also positive). $\emptyset$ is the true negative (the predicted result is negative, and the true result is also negative). Λ corresponds to the false positive (the predicted result is positive, but the true result is negative). Ψ represents the true false negative (the predicted result is negative, but the true result is positive).

For evaluating the value of F1 score, which is the harmonic mean of precision and recall, the equation would be represented as

$$F1 = \frac{2}{\text{Precision}^{-1} + \text{Recall}^{-1}} \tag{9}$$

$$F1 = \frac{\theta}{\theta + 0.5(\emptyset + \Psi)} \tag{10}$$

where the precision is defined as the number of true positives over the number of true positives plus the number of false positives, and recall is the number of true positives over the number of true positives plus the number of false negatives.

C. COMPARISON EXPERIMENT AND DISCUSSION

This section will describe the research questions, the experiment detail, the result, and the discussion

1) How does supervised fine-tuning and instruction tuning impact the efficiency of pre-trained language models in terms of performance on unseen tasks?

The research question of whether fine-tuning improves the model efficiency was investigated through a series of experiments. Three pre-trained language models (DistilBert, MiniLM, and FLAN-T5-Base) and four datasets were used to train and evaluate the performance of the models. This work was mainly focused on the zero-shot setup under the default parameter setup except for the learning rate, batch size, and no of epochs. The models were trained for a maximum of 3 epochs, and the Adam optimizer with a learning rate of 0.00002, and batch size of 8 was utilized.

Based on the experimental setup the pre-trained language models were divided into three groups: untuned model, SFT model, and IT model. Untuned model is the vanilla LM having the original checkpoint (only pre-training, no additional fine-tuning). The SFT model is considered the fine-tuning setups in a standard supervised way without instruction. It follows a no-template setup, only inputs and outputs were given to the model (e.g., for text classification, the input would be "Earn bitcoin on a daily basis!" and the output would be "Positive"). On the other hand, the IT model is considered fine-tuning setups in a standard supervised way with natural instruction. For IT model, first the instruction and instance input were concatenated to make a single prompt (e.g., for text classification. the input would be "Detect the sentiment of the given text, Text: Earn bitcoin on a daily basis!" and the output would be "Positive") and then trained the model to generate the instance output.

Tables 2, 3, and 4 present the results of experiments that were conducted to classify cryptocurrency-related tweets into two categories: positive and negative. The performances were compared between three language models with their vanilla, supervised Fine-tuned, and instruction-based Fine-tuned versions in terms of accuracy, F1 score, precision, and recall.

The overall analysis of the results based on the accuracy score is shown in Figure 5, which indicates that fine-tuning improves the model performance dramatically. The baseline performance for this study was established with the vanilla LM models. The DistiBERT-vanilla model achieved an accuracy score of 35.66%, 58.95%, and 60.33% across three datasets Bitcoin sentiment dataset, Reddit dataset, and Cryptocurrency sentiment dataset respectively. The MiniLM-vanilla model achieved an accuracy score of 41.67%, 46.67%, and 51.25%. And the Flan-T5-Base-vanilla model achieved an accuracy score of 25.10%, 46.07%, and 45.89%. The average accuracy score of all three vanilla LM models across all three different datasets is 45.73%. However, after fine-tuning, the average performance improved significantly to 59.52% for SFT model and 68.80% for IT model. The experimental results provide compelling evidence that fine-tuning enhances model efficiency. The substantial improvements in accuracy scores demonstrate the effectiveness of fine-tuning across different datasets and models. The average performance gain of 40% highlights the significant impact of fine-tuning on model performance, indicating its potential for practical applications.

Additionally, a comparison was made between the SFT and IT models to identify the role of instructions. The results revealed that the Fine-tuned models with instructions outperformed their counterparts without instructions. On average, the performance of the IT model was 16% higher than the SFT model. The consistent improvements observed across multiple models and datasets reinforce the generalizability of the findings. The results indicate that fine-tuning can be a valuable technique for optimizing model efficiency in various domains and tasks, providing researchers and practitioners with a powerful tool to enhance model performance.

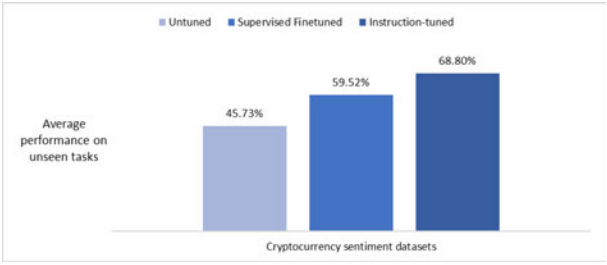


FIGURE 5. Zero-shot performance analysis between untuned, supervised, and instruction-base Fine-tuned models on unseen tasks.

TABLE 2. Classification results on Bitcoin sentiment dataset.

Model	Accuracy	F1 score	Precision	Recall
DistilBERT-vanilla	35.66%	24.42%	62.60%	15.97%
MiniLM-vanilla	41.67%	27.99%	25.00%	33.33%
Flan-T5-Base-vanilla	25.10%	0.26%	0.78%	0.16%
Avg_vanilla	34.14%	17.56%	29.46%	16.49%
DistilBERT-SFT	42.83%	41.28%	70.68%	31.68%
MiniLM-SFT	66.07%	63.03%	70.91%	61.21%
Flan-T5-Base-SFT	65.89%	62.13%	72.57%	60.13%
Avg_SFT	58.26%	55.48%	71.39%	51.00%
DistilBERT-IT	68.12%	70.93%	**91.87%**	61.53%
MiniLM-IT	70.38%	78.01%	78.59%	81.40%
Flan-T5-Base-IT	**75.00%**	**83.98%**	75.00%	**100.00%**
Avg_IT	71.17%	77.64%	81.82%	80.98%
Best_Score	**75.00%**	**83.98%**	**91.87%**	**100.00%**

2) How the benefits of instruction tuning are affected by model scale?

Based on the study conducted by Brown et al. [44], which revealed that zero and few-shot capabilities of larger language models substantially improve for larger models, the present research delves into investigating how the scale of the model influences the benefits of instruction tuning. The impact of instruction tuning was evaluated across FLAN-T5 models of different sizes: small (80M), base (250M), and large (780M), based on their parameters. The model's architecture and their comparative analysis have been summarized in Table 5.

TABLE 3. Classification results on Reddit dataset.

Model	Accuracy	F1 score	Precision	Recall
DistilBERT-vanilla	58.95%	30.79%	49.60%	23.45%
MiniLM-vanilla	46.67%	0.96%	1.90%	0.77%
Flan-T5-Base-vanilla	46.07%	1.05%	2.86%	0.64%
Avg_vanilla	50.56%	10.93%	18.12%	8.29%
DistilBERT-SFT	66.07%	63.03%	70.91%	61.21%
MiniLM-SFT	60.71%	57.00%	63.48%	57.08%
Flan-T5-Base-SFT	53.75%	67.42%	53.70%	**99.71%**
Avg_SFT	60.18%	62.48%	62.70%	72.67%
DistilBERT-IT	64.82%	55.60%	**76.86%**	48.04%
MiniLM-IT	63.75%	62.03%	65.06%	65.94%
Flan-T5-Base-IT	**74.29%**	**75.00%**	70.19%	**86.60%**
Avg_IT	67.62%	64.21%	70.70%	66.86%
Best_Score	**74.29%**	**75.00%**	**76.86%**	**99.71%**

TABLE 4. Classification results on Cryptocurrency sentiment dataset.

Model	Accuracy	F1 score	Precision	Recall
DistilBERT-vanilla	60.33%	30.76%	49.80%	23.15%
MiniLM-vanilla	51.25%	45.64%	35.83%	66.67%
Flan-T5-Base-vanilla	45.89%	0.95%	2.86%	0.57%
Avg_vanilla	52.49%	25.78%	29.50%	30.13%
DistilBERT-SFT	65.89%	62.13%	72.57%	60.13%
MiniLM-SFT	60.54%	58.19%	67.19%	57.16%
Flan-T5-Base-SFT	53.93%	68.49%	53.88%	**99.71%**
Avg_SFT	60.12%	62.94%	64.54%	72.34%
DistilBERT-IT	64.82%	55.60%	**76.86%**	48.04%
MiniLM-IT	63.75%	62.03%	65.06%	65.94%
Flan-T5-Base-IT	**74.29%**	**75.00%**	70.19%	**86.60%**
Avg_IT	67.62%	64.21%	70.70%	66.86%
Best_Score	**74.29%**	**75.00%**	**76.86%**	**99.71%**

Table 6 shows the experimental results of classifying the crypto-related tweets for three models on three datasets. The models are further subdivided into two groups. One is the vanilla LM, and the other one is the instruction-based Fine-tuned model. As depicted in Figure 6, the results shed light on the effectiveness of instruction tuning with a larger model size enhancing the performance on unseen tasks. The untuned models achieved an average accuracy of 54.28% for the small model, 39.02% for the base model, and 39.28% for the large model. The achieved accuracy result was inconsistent compared to the model's size. However, after applying instruction tuning, the accuracy improved to 57.98% for the small model, 73.10% for the base model, and 75.17% for the large model.

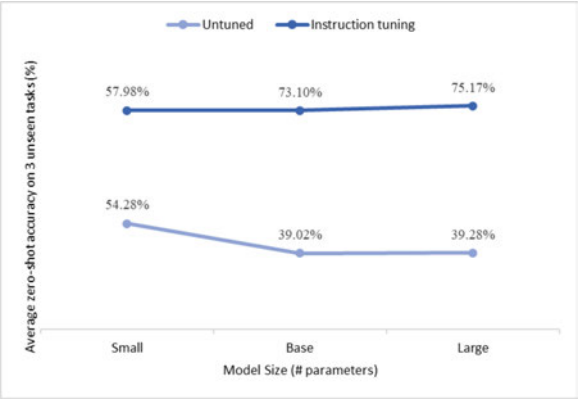


FIGURE 6. Evaluating instruction tuning efficacy for sentiment detection in FLAN-T5 models of varying sizes.

This potential result could be explained based on the study conducted by Wei et al. [19] that instruction tuning serves two purposes for larger-scale models. Firstly, it occupies some of the model's capacity. Secondly, it instructs these models on following instructions, enabling them to apply this skill to new tasks using the remaining capacity, which helps large models generalize to new tasks. But for small models, it actually hurts generalization to unseen tasks, potentially because all model capacity is used to learn the mixture of instruction tuning tasks.

3) How does the instruction-based model respond over different instruction tuning setups?

The conducted experimentation aimed to measure and compare the quality of models under different instruction tuning setups, focusing on the response of the instruction-based model. By introducing diversity in the styles and formats of tasks through instructions of varying lengths and complexities, the study provided insights into how the model handles diverse instructions and facilitates prompt tuning. Generally, they can be formulated as the following equations:

$$\text{argmax}_{M_{\text{tuned}}} \sum_{i \in I} Q(M_{\text{tuned}}, L_i, C_i) \tag{11}$$

where $I = \{i_1, i_2, i_3, \ldots, i_n\}$. $Q(M_{\text{tuned}}, L_i, C_i)$ represents the quality of the instruction-tuned model. M_{tuned} be the instruction-tuned model. i_j represents the instruction. L_i denotes the length of instruction i. C_i denotes the complexity of instruction i. Six instructions were created for this experiment, representing different lengths and complexities as shown in Table 7.

The results of prompt tuning experiments were analyzed for both short and simple instructions, as well as long and complex instructions, with the baseline being the FLAN-T5-BASE vanilla LM model as shown in Figure 7. Preliminary experiments revealed that prompt tuning had no significant impact on the performance of the vanilla LM model, which achieved an average accuracy score of 46.43% for both setups. However, prompt tuning demonstrated a significant im-

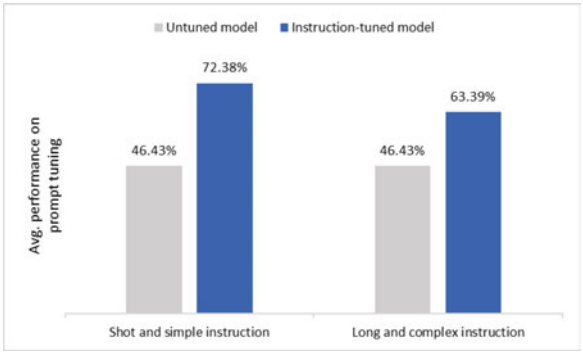


FIGURE 7. Performance analysis of untuned and instruction-based Fine-tuned FLAN-T5 model across various prompts.

TABLE 5. Architectural details of FLAN-T5 models.

Model	Architecture	#Heads	#Layers	Estimated model params size	#Parameter	#Trainable parameter	Model Checkpoint
FLAN-T5-SMALL	encoder-decoder	6	8	308MB	80M	77M	google/flan-t5-small
FLAN-T5-BASE	encoder-decoder	12	12	990MB	250M	247M	google/flan-t5-base
FLAN-T5-LARGE	encoder-decoder	16	24	3133MB	780MB	783M	google/flan-t5-large

TABLE 6. Performance analysis of untuned and instruction-based Fine-tuned FLAN-T5 model across different sizes and datasets.

Model	Untuned			Instruction-based tuned		
	Bitcoin sentiment	Reddit	Cryptocurrency sentiment	Bitcoin sentiment	Reddit	Cryptocurrency sentiment
FLAN-T5-SMALL	59.98%	51.43%	51.43%	75.00%	49.46%	49.46%
FLAN-T5-BASE	25.10%	46.07%	45.89%	75.00%	72.14%	72.14%
FLAN-T5-LARGE	25.87%	46.07%	45.89%	76.94%	74.29%	74.29%

provement in the performance of the instruction-based model. Under short and simple instructions, the model achieved an average accuracy score of 72.38%, showcasing its effectiveness in understanding and executing such instructions. On the other hand, the model exhibited slightly inferior performance for long and complex instructions, with an average accuracy score of 63.39%.

The findings suggest that the instruction-based model excels in responding to and executing short and simple instructions, outperforming its performance under long and complex instructions by over 12%.

TABLE 7. Prompt type variations for sentiment detection and instruction-based Fine-tuned model evaluation across diverse text types.

Type	Prompt
Shot and simple	*Please detect the sentiment.*
	Detect the sentiment of the text.
	Please detect the sentiment of the given text.
Long and complex	*Classify the sentiment of the provided cryptocurrency related social media posts or messages.*
	Determine the emotional tone of the given text, which primarily revolves around cryptocurrencies and their associated concepts.
	Categorize the sentiment expressed in the provided dataset consisting of the text snippets related to cryptocurrency and computer science, focusing on capturing positive or negative sentiments.

TABLE 8. Average zero-shot accuracy at different sample sizes on supervised Fine-tuned models.

Sample size	DistilBert	MiniLM	FLAN-T5	Average	Best Score
2K	54.58%	63.03%	58.51%	58.71%	63.03%
4K	56.54%	69.08%	60.27%	61.93%	69.08%
6K	**66.64%**	**69.67%**	**61.16%**	**65.82%**	**69.67%**
8K	57.83%	66.82%	60.89%	61.85%	66.82%
10K	55.93%	54.58%	61.04%	57.19%	61.04%
12K	58.26%	59.72%	60.93%	59.64%	60.93%

4) How does the size of the fine-tuning dataset impact the performance of different language models, and what is the optimal sample size for achieving the highest performance of any model?

The experimental evaluation aimed to explore the relationship between the size of the fine-tuning dataset and the SFT models' performance. Three models, namely DistilBert, MiniLM, and FLAN-T5-Base, were examined, and the sample size was varied from 2,000 to 12,000 data points to assess the influence of data availability. Table 8 presents the average zero-shot accuracy achieved by each model at different sample sizes on unseen tasks. The experimental findings provide valuable insights into the impact of fine-tuning dataset size on model performance, highlighting data efficiency and consistency considerations across the examined models. Among the three models, the best accuracy result was consistently achieved with a sample size of 6,000 data points. At this sample size, DistilBert reached an accuracy of 66.64%, MiniLM achieved 69.67%, and FLAN-T5-Base demonstrated an accuracy of 61.16%. This indicates that 6,000 data points serve as an optimal balance for achieving the highest performance for each model. MiniLM exhibited significant data efficiency by achieving an average zero-shot accuracy of 63.03% with a sample size of 2,000 data points. In comparison, FLAN-T5-Base and DistilBert required larger sample sizes, approximately double and triple the data, respectively, to attain comparable accuracy levels. This showcases MiniLM's ability to leverage limited data and deliver competitive performance effectively.

Furthermore, FLAN-T5-Base displayed consistent performance across varying sample sizes, maintaining relatively stable average zero-shot accuracy values. This suggests the model's robustness and resilience to variations in the dataset size. Conversely, DistilBert exhibited performance fluctuations and a declining trend as the sample size increased beyond 6,000 data points. This implies that DistilBert may reach a saturation point where additional data does not contribute significantly to its performance improvement.

The findings have practical implications for model selection in the fine-tuning process. The data efficiency of MiniLM, combined with the consistent performance of FLAN-T5-Base, offers advantages in scenarios with limited data availability or where performance stability is crucial. Meanwhile, considering the appropriate sample size, such as 6,000 data points, is essential to optimize the performance of DistilBert.

V. CONCLUSION

In conclusion, the above research establishes the value of large language model fine-tuning strategies for sentiment analysis in the context of cryptocurrencies. Experimental results show a significant 40% average gain in zero-shot performance after fine-tuning, highlighting the potential of this strategy in maximizing the effectiveness of pre-trained language models. Additionally, the result shows that instruction adjustment improves model performance, with larger models reaching a remarkable average accuracy score of 75.17%. Notably, it has been found that 6,000 data points were the ideal corpus size, with MiniLM showing outstanding data efficiency and FLAN-T5 performing consistently across a range of corpus sizes. The research also shows that the model performs well when given short, clear instructions, surpassing its performance when given longer, more complicated instructions by almost 12%. These observations help sentiment analysis in the cryptocurrency space improve while also offering helpful advice for large language model optimization. They should stimulate additional studies in supervised and instruction-based NLP, zero-shot learning, instruction tuning, and the use of labeled data to improve the effectiveness of big language models in cryptocurrency applications.

REFERENCES

[1] Muhammad Saad and Aziz Mohaisen. Towards characterizing blockchain-based cryptocurrencies for highly-accurate predictions. *INFOCOM 2018 - IEEE Conference on Computer Communications Workshops*, pages 704–709, 7 2018.

[2] Vitalik Buterin. Ethereum: A next-generation smart contract and decentralized application platform. 2014.

[3] Zeinab Shahbazi and Yung Cheol Byun. Improving the cryptocurrency price prediction performance based on reinforcement learning. *IEEE Access*, 9:162651–162659, 2021.

[4] Fan Fang, Carmine Ventre, Michail Basios, Leslie Kanthan, David Martinez-Rego, Fan Wu, and Lingbo Li. Cryptocurrency trading: a comprehensive survey. *Financial Innovation*, 8:1–59, 12 2022.

[5] Patel Jay, Vasu Kalariya, Pushpendra Parmar, Sudeep Tanwar, Neeraj Kumar, and Mamoun Alazab. Stochastic neural networks for cryptocurrency price prediction. *IEEE Access*, 8:82804–82818, 2020.

[6] Emre Şaşmaz and F. Boray Tek. Tweet sentiment analysis for cryptocurrencies. *Proceedings - 6th International Conference on Computer Science and Engineering, UBMK 2021*, pages 613–618, 2021.

[7] Pratikkumar Prajapati. Predictive analysis of bitcoin price considering social sentiments. 1 2020.

[8] Nikolaos Passalis, Loukia Avramelou, Solon Seficha, Avraam Tsantekidis, Stavros Doropoulos, Giorgos Makris, and Anastasios Tefas. Multisource financial sentiment analysis for detecting bitcoin price change indications using deep learning. *Neural Computing and Applications*, 34:19441–19452, 11 2022.

[9] Yogev Matalon, Ofir Magdaci, Adam Almozlino, and Dan Yamin. Using sentiment analysis to predict opinion inversion in tweets of political communication. *Scientific Reports 2021 11:1*, 11:1–9, 3 2021.

[10] Kogilavani Shanmugavadivel, V. E. Sathishkumar, Sandhiya Raja, T. Bheema Lingaiah, S. Neelakandan, and Malliga Subramanian. Deep learning based sentiment analysis and offensive language identification on multilingual code-mixed data. *Scientific Reports 2022 12:1*, 12:1–12, 12 2022.

[11] Binder Tweet. Twitter real time monitoring and twitter live analytics. https://www.tweetbinder.com/blog/real-time-twitter/, 2019.

[12] Sayyida Tabinda Kokab, Sohail Asghar, and Shehneela Naz. Transformer-based deep learning models for the sentiment analysis of social media data. *Array*, 14:100157, 7 2022.

[13] María Hernández-Rubio, Iván Cantador, and Alejandro Bellogín. A comparative analysis of recommender systems based on item aspect opinions extracted from user reviews. *User Modeling and User-Adapted Interaction*, 29:381–441, 4 2019.

[14] Sonali Sharma, Manoj Diwakar, Kapil Joshi, Prabhishek Singh, Shaik Vaseem Akram, and Anita Gehlot. A critical review on sentiment analysis techniques. *Proceedings of 3rd International Conference on Intelligent Engineering and Management, ICIEM 2022*, pages 741–746, 2022.

[15] Zhao Jianqiang, Gui Xiaolin, and Zhang Xuejun. Deep convolution neural networks for twitter sentiment analysis. *IEEE Access*, 6:23253–23260, 1 2018.

[16] Ummara Ahmed Chauhan, Muhammad Tanvir Afzal, Abdul Shahid, Moloud Abdar, Mohammad Ehsan Basiri, and Xujuan Zhou. A comprehensive analysis of adverb types for mining user sentiments on amazon product reviews. *World Wide Web*, 23:1811–1829, 5 2020.

[17] Sani Kamiş and Dionysis Goularas. Evaluation of deep learning techniques in sentiment analysis from twitter data. *Proceedings - 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications, Deep-ML 2019*, pages 12–17, 8 2019.

[18] Jeffrey Pennington, Richard Socher, and Christopher Manning. Glove: Global vectors for word representation. pages 1532–1543. Association for Computational Linguistics, 2014.

[19] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017-December:5999–6009, 6 2017.

[20] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. 10 2018.

[21] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21:1–67, 10 2019.

[22] Shaden Smith, Mostofa Patwary, Brandon Norick, Patrick LeGresley, Samyam Rajbhandari, Jared Casper, Zhun Liu, Shrimai Prabhumoye, George Zerveas, Vijay Korthikanti, Elton Zhang, Rewon Child, Reza Yazdani Aminabadi, Julie Bernauer, Xia Song, Mohammad Shoeybi, Yuxiong He, Michael Houston, Saurabh Tiwary, and Bryan Catanzaro. Using deepspeed and megatron to train megatron-turing nlg 530b, a large-scale generative language model. 1 2022.

[23] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. Llama: Open and efficient foundation language models. 2 2023.

[24] Jason Wei, Maarten Bosma, Vincent Y Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M Dai, and Quoc V Le Google Research. Finetuned language models are zero-shot learners. 9 2021.

[25] Lin Yue, Weitong Chen, Xue Li, Wanli Zuo, and Minghao Yin. A survey of sentiment analysis in social media. *Knowledge and Information Systems*, 60:617–663, 8 2019.

[26] Isabella Donita Hasan, Raymond Sunardi Oetama, and Aldo Lionel Saonard. Sentiment analysis on cryptocurrency based on tweets and retweets using support vector machines and chi-square. *2022 7th International Conference on Informatics and Computing, ICIC 2022*, 2022.

[27] Raja Nanda Satrya, Oktariani Nurul Pratiwi, Riska Yanu Farifah, and Jemal Abawajy. Cryptocurrency sentiment analysis on the twitter platform using support vector machine (svm) algorithm. *Proceedings - International Conference Advancement in Data Science, E-Learning and Information Systems, ICADEIS 2022*, 2022.

[28] E. Padmalatha, Sailekya Sheral, K. Dhanush, S. Samveeth, B.Ruchi Datta, and K.Tarun Krishna. Sentiment analysis of bitcoin data by tweets through naive bayes. pages 1–4. IEEE, 11 2022.

[29] Divesh, Gaurav Prasad, Gaurav Sharma, and Dinesh Kumar Vishwakarma. Sentiment analysis on cryptocurrency using youtube comments. pages 730–733. IEEE, 3 2022.

[30] Sandy Suardi, Atiqur Rahman Rasel, and Bin Liu. On the predictive power of tweet sentiments and attention on bitcoin. *International Review of Economics & Finance*, 79:289–301, 5 2022.

[31] Sotirios Oikonomopoulos, Katerina Tzafilkou, Dimitrios Karapiperis, and Vassilios Verykios. Cryptocurrency price prediction using social media sentiment analysis. *13th International Conference on Information, Intelligence, Systems and Applications, IISA 2022*, 2022.

[32] Achyut Jagini, Kaushal Mahajan, Namita Aluvathingal, Vedanth Mohan, and Prajwala TR. Twitter sentiment analysis for bitcoin price prediction.

2023 3rd International Conference on Smart Data Intelligence (ICSMDI), pages 32–37, 3 2023.

[33] Raj Parekh, Nisarg P. Patel, Nihar Thakkar, Rajesh Gupta, Sudeep Tanwar, Gulshan Sharma, Innocent E. Davidson, and Ravi Sharma. dl-guess : Deep learning and sentiment analysis-based cryptocurrency price prediction. *IEEE Access*, 10:35398–35409, 2022.

[34] Himanshu Dwivedi. Cryptocurrency sentiment analysis using bidirectional transformation. pages 140–142. IEEE, 3 2023.

[35] Gyeongmin Kim, Minsuk Kim, Byungchul Kim, and Heuiseok Lim. Cbits: Crypto bert incorporated trading system. *IEEE Access*, 11:6912–6921, 2023.

[36] Mochammad Haldi Widianto and Yhudi Cornelius. Sentiment analysis towards cryptocurrency and nft in bahasa indonesia for twitter large amount data using bert. *International Journal of Intelligent Systems and Applications in Engineering*, 11:303 – 309, 2 2023.

[37] Marco Ortu, Nicola Uras, Claudio Conversano, Giuseppe Destefanis, and Silvia Bartolucci. On technical trading and social media indicators in cryptocurrencies' price classification through deep learning. 2 2021.

[38] Xipeng Qiu, Tianxiang Sun, Yige Xu, Yunfan Shao, Ning Dai, and Xuanjing Huang. Pre-trained models for natural language processing: A survey. 3 2020.

[39] Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, and et al. Maarten Bosma. Palm: Scaling language modeling with pathways. 4 2022.

[40] Jordan Hoffmann, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego De, Las Casas, Lisa Anne Hendricks, Johannes Welbl, Aidan Clark, Tom Hennigan, Eric Noland, Katie Millican, George Van Den Driessche, Bogdan Damoc, Aurelia Guy, Simon Osindero, Karen Simonyan, Erich Elsen, Jack W Rae, Oriol Vinyals, Laurent Sifre, and Equal. Training compute-optimal large language models. 3 2022.

[41] Guilherme Penedo, Quentin Malartic, Daniel Hesslow, Ruxandra Cojocaru, Alessandro Cappelli, Hamza Alobeidli, Baptiste Pannier, Ebtesam Almazrouei, and Julien Launay. The refinedweb dataset for falcon llm: Outperforming curated corpora with web data, and web data only. 6 2023.

[42] Hongyang Du, Zonghang Li, Dusit Niyato, Jiawen Kang, Zehui Xiong, Xuemin, Shen, and Dong In Kim. Enabling ai-generated content (aigc) services in wireless edge networks. 1 2023.

[43] Tianyu Wu, Shizhu He, Jingping Liu, Siqi Sun, Kang Liu, Qing Long Han, and Yang Tang. A brief overview of chatgpt: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica*, 10:1122–1136, 5 2023.

[44] Tom B Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam Mccandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33:1877–1901, 2020.

[45] Susan Zhang, Stephen Roller, Naman Goyal, Mikel Artetxe, Moya Chen, Shuohui Chen, Christopher Dewan, Mona Diab, Xian Li, Victoria Lin, Todor Mihaylov, Myle Ott, Sam Shleifer, Kurt Shuster, Daniel Simig, Singh Koura, Anjali Sridhar, Tianlu Wang, Luke Zettlemoyer, and Meta Ai. Opt: Open pre-trained transformer language models. 5 2022.

[46] Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. Training language models to follow instructions with human feedback. 3 2022.

[47] Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, Yifan Du, Chen Yang, Yushuo Chen, Zhipeng Chen, Jinhao Jiang, Ruiyang Ren, Yifan Li, Xinyu Tang, Zikang Liu, Peiyu Liu, Jian-Yun Nie, and Ji-Rong Wen. A survey of large language models. 3 2023.

[48] Chengwei Wei, Yun-Cheng Wang, Bin Wang, and C. C. Jay Kuo. An overview on language models: Recent developments and outlook. 3 2023.

[49] Raschka Sebastian. Finetuning large language models. https://magazine.sebastianraschka.com/p/finetuning-large-language-models?subscribe_prompt=free, 4 2023.

[50] Shah Kushal. Pre-training, fine-tuning and in-context learning in large language models (llms) | by kushal shah | medium. https://medium.com/@atmabodha/pre-training-fine-tuning-and-in-context-learning-in-large-language-models-llms-dd4 8 2022.

[51] Vithursan Thangarasa, Abhay Gupta, William Marshall, Tianda Li, Kevin Leong, Dennis DeCoste, Sean Lie, and Shreyas Saxena. Spdf: Sparse pre-training and dense fine-tuning for large language models. 3 2023.

[52] Victor Sanh, Albert Webson, Colin Raffel, Stephen H. Bach, Lintang Sutawika, Zaid Alyafeai, Antoine Chaffin, Arnaud Stiegler, Teven Le Scao, Arun Raja, Manan Dey, M Saiful Bari, Canwen Xu, Urmish Thakker, Shanya Sharma Sharma, Eliza Szczechla, Taewoon Kim, Gunjan Chhablani, Nihal Nayak, Debajyoti Datta, Jonathan Chang, Mike Tian-Jian Jiang, Han Wang, Matteo Manica, Sheng Shen, Zheng Xin Yong, Harshit Pandey, Rachel Bawden, Thomas Wang, Trishala Neeraj, Jos Rozen, Abheesht Sharma, Andrea Santilli, Thibault Fevry, Jason Alan Fries, Ryan Teehan, Tali Bers, Stella Biderman, Leo Gao, Thomas Wolf, and Alexander M. Rush. Multitask prompted training enables zero-shot task generalization. 10 2021.

[53] Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Alex Castro-Ros, Marie Pellat, Kevin Robinson, Dasha Valter, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Zhao, Yanping Huang, Andrew Dai, Hongkun Yu, Slav Petrov, Ed H. Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V. Le, and Jason Wei. Scaling instruction-finetuned language models. 10 2022.

[54] Niklas Muennighoff, Thomas Wang, Lintang Sutawika, Adam Roberts, Stella Biderman, Teven Le Scao, M Saiful Bari, Sheng Shen, Zheng-Xin Yong, Hailey Schoelkopf, Xiangru Tang, Dragomir Radev, Alham Fikri Aji, Khalid Almubarak, Samuel Albanie, Zaid Alyafeai, Albert Webson, Edward Raff, Colin Raffel, Hugging Face, Google Research, Brain Team, and Booz Allen Hamilton. Crosslingual generalization through multitask finetuning. 11 2022.

[55] BigScience Workshop, Teven Le Scao, Angela Fan, and Christopher Akiki et al. Bloom: A 176b-parameter open-access multilingual language model. 11 2022.

[56] Qingxiu Dong, Lei Li, Damai Dai, Ce Zheng, Zhiyong Wu, Baobao Chang, Xu Sun, Jingjing Xu, Lei Li, and Zhifang Sui. A survey on in-context learning. 12 2022.

[57] Layton Dennis. Prompt engineering - leveraging in context learning. https://www.linkedin.com/pulse/prompt-engineering-leveraging-context-learning-dennis-layton/, 4 2023.

[58] C. Montgomery and Patrick H. Winston. Learning and reasoning by analogy. *Communications of the ACM*, 23:689–703, 12 1980.

[59] Sewon Min, Mike Lewis, Luke Zettlemoyer, and Hannaneh Hajishirzi. Metaicl: Learning to learn in context. *NAACL 2022 - 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*, pages 2791–2809, 10 2021.

[60] Damai Dai, Yutao Sun, Li Dong, Yaru Hao, Shuming Ma, Zhifang Sui, Furu Wei, and Moe Key. Why can gpt learn in-context? language models implicitly perform gradient descent as meta-optimizers. 12 2022.

[61] Alejandro Lopez-Lira and Yuehua Tang. Can chatgpt forecast stock price movements? return predictability and large language models. 4 2023.

[62] Jules White, Quchen Fu, Sam Hays, Michael Sandborn, Carlos Olea, Henry Gilbert, Ashraf Elnashar, Jesse Spencer-Smith, and Douglas C. Schmidt. A prompt pattern catalog to enhance prompt engineering with chatgpt. 2 2023.

[63] Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55, 7 2021.

[64] Tianyu Gao, Adam Fisch, and Danqi Chen. Making pre-trained language models better few-shot learners. pages 3816–3830. Association for Computational Linguistics, 2021.

[65] Teven Le Scao and Alexander M. Rush. How many data points is a prompt worth? *NAACL-HLT 2021 - 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*, pages 2627–2636, 3 2021.

[66] Timo Schick and Hinrich Schütze. It's not just size that matters: Small language models are also few-shot learners. *NAACL-HLT 2021 - 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*, pages 2339–2352, 9 2020.

[67] Hendrik Strobelt, Albert Webson, Victor Sanh, Benjamin Hoover, Johanna Beyer, Hanspeter Pfister, and Alexander M. Rush. Interactive and visual prompt engineering for ad-hoc task adaptation with large language models. *IEEE Transactions on Visualization and Computer Graphics*, 29:1146–1156, 1 2023.
[68] Paula Maddigan and Teo Susnjak. Chat2vis: Generating data visualisations via natural language using chatgpt, codex and gpt-3 large language models. 2 2023.
[69] OpenAI Platform. Gpt best practices - openai api. https://platform.openai.com/docs/guides/gpt-best-practices.

Gwangju Institute of
Science and Technology

School of Information and Communications



Hyper Spectrometer Imaging system

- WorldLand, My AI Network 핵심 기술 연구 결과–

scientific reports

Check for updates

OPEN

Mass production-enabled computational spectrometers based on multilayer thin films

Cheolsun Kim[1], Pavel Ni[1], Kang Ryeol Lee[2] & Heung-No Lee[1]✉

Multilayer thin film (MTF) filter arrays for computational spectroscopy are fabricated using stencil lithography. The MTF filter array is a 6 × 6 square grid, and 169 identical arrays are fabricated on a single wafer. A computational spectrometer is formed by attaching the MTF filter array on a complementary metal–oxide–semiconductor (CMOS) image sensor. With a single exposure, 36 unique intensities of incident light are collected. The spectrum of the incident light is recovered using collected intensities and numerical optimization techniques. Varied light sources in the wavelength range of 500 to 849 nm are recovered with a spacing of 1 nm. The reconstructed spectra are a good match with the reference spectra, measured by a grating-based spectrometer. We also demonstrate computational pinhole spectral imaging using the MTF filter array. Adapting a spectral scanning method, we collect 36 monochromatic filtered images and reconstructed 350 monochromatic images in the wavelength range of 500 to 849 nm, with a spacing of 1 nm. These computational spectrometers could be useful for various applications that require compact size, high resolution, and wide working range.

Spectrometers are powerful tools for remote sensing and medical applications[1–3]. However, these uses are restricted to research and development due to limitations based on the spectrometers' bulky size, high cost, and long measuring time. There have been tremendous efforts to overcome spectrometer limitations and go beyond restricted applications[4–21]. One promising candidate to achieve this is optical filter array-based spectrometers: complementary metal–oxide–semiconductor (CMOS) image sensors with a filter array used as a spectrometer. These spectrometers are smaller and have faster measuring times, so they are useful in portable applications, such as on-site detection and small unmanned aerial vehicle (UAV)-based remote sensing. However, the number of filters that can be attached to a CMOS image sensor is limited due to its small sensing area. Thus, these spectrometers offer a low spectral resolution.

Over the past decade, computational approaches[22,23] have been adapted for filter-based spectrometers. The spectral resolution in conventional filter array-based spectrometers has been improved using computational approaches. New optical filter types have been proposed that work well in computational approaches and achieve further improvements[6–8,10,11,13,14,16,21]. Unlike conventional optical filters, which selectively transmit incident light in specific wavelengths and reflect the remaining wavelengths, these filters, called random spectral filters, modulate and transmit incident light with unique spectral features in the entire wavelength ranges of interest. Various types of random spectral filters have been proposed, such as etalon filters[10,11,20], quantum dot filters[8,21], photonic crystal slabs[7,9,14,16], and multilayer thin films (MTF) filters[6,13]. The spectral resolvability of computational spectroscopy has been successfully demonstrated using random spectral filters with low correlation among filters.

In contrast to transmission functions of an etalon filter, which consists of repetitive narrow peaks, and a quantum dot filter, which consists of a broadband peak, the fabricated random spectral filter has a transmission function of multiple peaks with various full widths at half maximums (FWHMs) and has a large difference between maximal and minimal transmission in the transmission function. By utilizing the computational approaches, a wide wavelength range can be covered with a small number of MTF filters. In this work, a small number of MTF-based random spectral filters were fabricated in the form of an array. 169 identical filter arrays, consisting of 36 MTF filters, were fabricated on a single wafer. We realized MTF filters by stacking multiple layers of two alternating materials with high and low refractive indices. Using stencil lithography based on shadow masks, we could fabricate MTF filters with different spectral features simultaneously as a filter array form. This idea of random spectral filters can be applied to various wavelength ranges by changing the MTFs' design properties.

[1]School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea. [2]SNI Co., Ltd., Seongnam, Gyeonggi 13509, South Korea. ✉email: heungno@gist.ac.kr

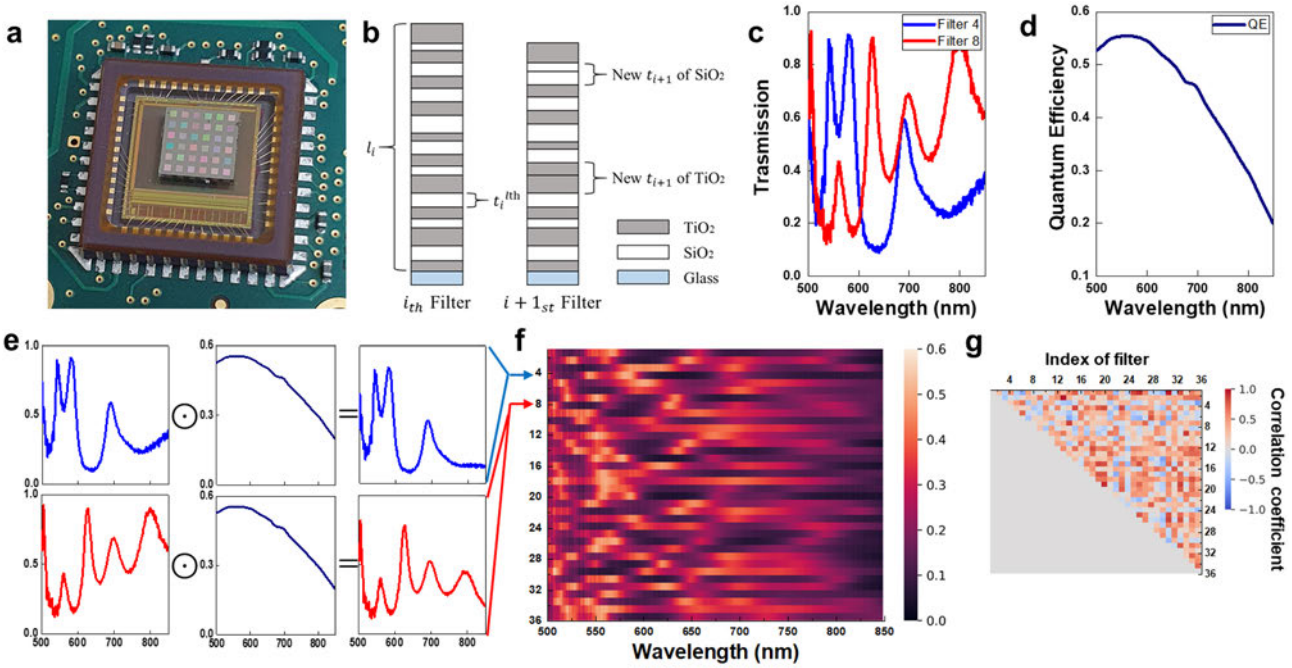


Figure 1. Multilayer thin films (MTF) based computational spectrometer. (**a**) Photograph of the MTF filter array, which is directly attached to the CMOS image sensor. (**b**) Schematic of MTF filters: TiO_2 and SiO_2 are deposited on a glass substrate with design properties such as the number of layers of the i-th MTF filter (l_i) and thickness of the l-th layer of the i-th MTF filter $\left(t_i^{\text{lth}}\right)$. (**c**) Two transmission functions of MTF filters. (**d**) Spectral response of the CMOS image sensor. (**e**) Spectral sensitivity of an MTF filter with the CMOS image sensor, which can be calibrated by element-wise multiplication of the transmission function and the spectral response of the CMOS image sensor. (**f**) Heatmap of the sensing matrix. Each row represents the spectral sensitivity with respect to wavelength. (**g**) Upper triangular matrix of correlation coefficients which are pairwise compared among rows of the sensing matrix.

According to the usage of MTF filters, they can also be mass-produced in various shapes using stencil lithography techniques.

Here, we demonstrate the competence of spectral reconstructions over the wavelength range of 500 to 849 nm using a fabricated filter array. The fabricated filter array consists of 36 MTF filters in the shape of a square. Varied spectra of incident light such as monochromatic, broadband, and continuous light is used to test the filter array's reconstruction performances. Additionally, we perform pinhole spectral imaging using the filter array, showing that computational spectral imaging is possible.

Results

Spectrometers based on MTF filters. The computational spectrometer consists of MTF filters and the CMOS image camera. As shown in Fig. 1a, MTF filters are in the form of an array and are directly attached to the CMOS image sensor. Each filter has unique spectral features that can be realized by stacking multiple layers of thin films. A schematic of the MTF filter is shown in Fig. 1b. The transmission function of the i-th MTF filter is determined by design properties, such as the number of layers (l_i) and the thickness of the l-th layer $\left(t_i^{\text{lth}}\right)$. Using the transfer-matrix method, the transmission function of the MTF filter can be calculated[24,25]. We choose a set of different design properties to produce a set of filters with a unique transmission function. Let us denote the transmission function of the i-th MTF filter in the wavelength range $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_N]$ as $\mathbf{T}_i = [T_i(\lambda_1), T_i(\lambda_2), \ldots, T_i(\lambda_N)]$. Figure 1c shows two measured transmission functions of the MTF filters (see Sect. 4 for measuring transmission functions of MTF filters). The intensity, y_i, measured by CMOS image sensor for an unknown incident spectrum $\mathbf{x} = [x(\lambda_1), x(\lambda_2), \cdots, x(\lambda_N)]^T$, can be expressed as:

$$y_i = \sum_{k=1}^{N} T_i(\lambda_k) Q(\lambda_k) x(\lambda_k), \tag{1}$$

where $\mathbf{Q} = [Q(\lambda_1), Q(\lambda_2), \ldots, Q(\lambda_N)]$ is the spectral response of CMOS image sensor in the wavelength range $\boldsymbol{\lambda}$. The spectral response is represented in Fig. 1d. Let us set $R_i(\lambda_k) = T_i(\lambda_k)Q(\lambda_k)$, where $R_i(\lambda_k)$ represents spectral sensitivity of the i-th filter of the CMOS image sensor at the wavelength λ_k, the Eq. (1) becomes $y_i = \sum_{k=1}^{N} R_i(\lambda_k)Q(\lambda_k)$. Considering an M number of filters, there is a set of M equations for $i = 1, 2, \ldots, M$. The set of M equations can be represented in matrix formation:

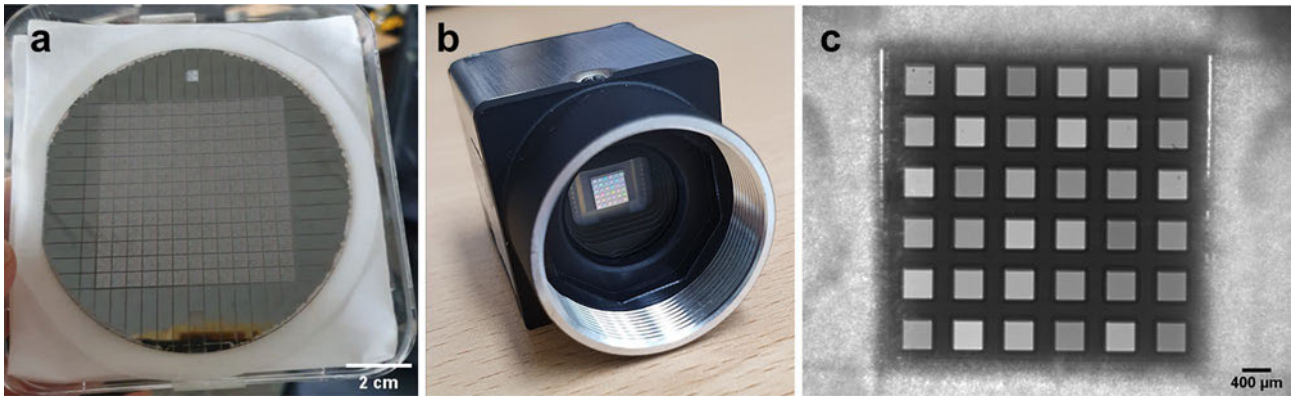


Figure 2. Fabricated MTF filter array. (**a**) 169 identical MTF filter arrays fabricated in a single wafer. (**b**) Photograph of the CMOS image camera with the fabricated MTF filter array. (**c**) Monochrome image of the fabricated MTF filter array illuminated by a halogen light source.

$$\begin{bmatrix} y_1 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} R_1(\lambda_1) & \cdots & R_1(\lambda_N) \\ \vdots & \ddots & \vdots \\ R_M(\lambda_1) & \cdots & R_M(\lambda_N) \end{bmatrix} \begin{bmatrix} x(\lambda_1) \\ \vdots \\ x(\lambda_N) \end{bmatrix}, \tag{2}$$

where $\mathbf{y} \in \mathbb{R}^{M\times 1}$ is a column vector with measured intensities from M filters and $\mathbf{R} \in \mathbb{R}^{M\times N}$ is the sensing matrix where each row represents the spectral sensitivity with respect to the wavelength. The spectral sensitivity can be calibrated by element-wise multiplication of the transmission functions of MTF filters and the spectral response of the CMOS image sensor, as depicted in Fig. 1e,f.

Conventional spectrometers read out **y** as the incident spectrum **x**. In order to make the measured intensities **y** as close as possible to the incident spectrum **x**, the sensing matrix **R** should be an identity matrix with the dimension of $N \times N(M = N)$. This means that N number of filters are needed in conventional manners. In practice, it may be difficult to fabricate a narrow FWHM filter and, since the number of filters required increases as the wavelength of interest increases, it is more challenging to make a compact spectrometer operating in a wide wavelength range. Unlike conventional filter-based spectrometers, computational spectrometers modulate and measure a wide wavelength range of the incident spectrum using a small number of MTF filters. We consider the sensing matrix **R** with dimensions $M \times N(M < N)$. The set of M equations becomes an underdetermined problem. Reconstruction algorithms[26–28] can be applied to restore the incident spectrum in high resolution by solving the underdetermined problem.

Figure 1f shows the heatmap of the sensing matrix of the fabricated MTF filter array-based computational spectrometer. Each row represents the spectral sensitivity with respect to wavelength. The correlation coefficients for each pair of two rows of the sensing matrix are shown as the upper triangular matrix in Fig. 1g. The average value of the correlation coefficients is 0.231, which can be described as a weak or moderate correlation among sensitivities. With the weakly correlated spectral sensitivities, the incident spectrum was measured as unique intensities, which allow the reconstruction algorithms to work effectively.

Fabrication of MTF filter arrays. We fabricated 169 identical filter arrays on a single wafer, as shown in Fig. 2a. The filter array is the shape of a 6×6 square grid. The size of the square is 400×400 μm^2, and the space between the squares is 300 μm. Accordingly, the size of filter array is 4.5×4.5 mm^2. To fabricate filter arrays, we use TiO_2 and SiO_2 as a high and low refractive index materials, respectively. The refractive indices for TiO_2 and SiO_2 are approximately 2.6 and 1.45 at 600 nm, respectively.

Unlike etalon filters that were fabricated by changing the thickness of interspacing dielectric layers[20,29], we fabricated the MTF filters by changing the number of layers and thicknesses of layers. An MTF filter with a unique transmission function can be obtained by repeatedly alternating the two materials and depositing them with different thicknesses. 36 MTF filters with unique transmission functions were fabricated by selectively omitting certain layers of different MTF filters using shadow masks during the deposition of the MTF filter array. As shown in Fig. 1b, the upper and lower layers sum up to form one layer with a different thickness by omitting an intermediate layer. The designed thicknesses of layers for MTF filters are shown in Supplementary Information Table S1. The deposition process for creating filter arrays follows.

TiO_2 and SiO_2 films were deposited onto a borosilicate glass wafer whose refractive index is approximately 1.472 at 588 nm. In order to distinguish where the material should be deposited, shadow masks were used. The desired thickness of TiO_2 is deposited on the desired locations using direct current (DC) magnetron sputter. For TiO_2 deposition, a Ti target was sputtered in a mixture of argon (Ar) and oxygen (O_2). The mixture gas flow of 188 sccm of Ar and 12 sccm of O_2 was utilized and the DC power was 700 W. The TiO_2 deposition is performed only on the desired region designated by the shadow mask. Then, the shadow mask is changed, with different patterns on the other mask, and we deposit SiO_2 at the intended thickness. Radio frequency (RF) magnetron sputter was used for the SiO_2 deposition. A Si target was sputtered in a mixture of Ar and O_2. The mixture gas

flow of 185 sccm of Ar and 15 sccm of O_2 was utilized. The RF power was 300 W. The deposition is repeated 17 additional times by changing the shadow mask and alternating between TiO_2 and SiO_2. Hence, we conducted ten individual depositions of TiO_2 and nine individual depositions of SiO_2. The number of shadow masks used in depositions was 19. After completing thin film deposition, we coated the surface of thin films with a photoresist. Germanium (Ge) was deposited over the entire wafer area using an e-beam evaporator. Lift-off of the photoresist was performed by soaking the deposited wafer in acetone. When the photoresist was washed away, Ge deposited on the top of the photoresist was lifted off and washed. After lift-off, a square grid of Ge with the size of 400 μm and spacing of 300 μm was formed. The Ge grid was formed to separate MTF filters and prevent incident lights from entering among MTF filters. The wafer cleaning process was then performed, and, finally, the wafer was diced to produce MTF filter arrays.

Unlike the previous work in that SiNx was used as the high refractive index material to fabricate an MTF filter array[13], we used TiO_2 as the high refractive index material and could reduce the number of layers for realizing the unique transmission functions. In addition, using stencil lithography, MTF filter arrays could be fabricated in a simplified process that does not involve an etching process.

The MTF filter array-based spectrometer was built by attaching the fabricated MTF filter array to the front of a CMOS monochrome camera, as shown in Fig. 2b. Figure 2c is a monochrome image of the fabricated MTF filter array illuminated by a halogen light source. The image was taken by the CMOS monochrome camera, whose number of pixels is 1280×1024. As shown in Fig. 2c, we measure uniform intensity using pixels under a single MTF filter. Also, the pixels have unique intensity according to the MTF filter. Using these unique intensities from MTF filters, we can reconstruct the spectrum of unknown incident light.

Spectral reconstruction experiments. Here, we address the spectral resolvability of the MTF filter array-based computational spectrometer. An unknown spectrum, $\mathbf{x}$, consisting of 350 ($N = 350$) spectral components from wavelengths ranging 500 to 849 nm, is retrieved using measured intensities $\mathbf{y}$ with size 36 ($M = 36$). For retrieving the unknown spectrum, we use a sparse representation-based l_1-norm minimization problem. $\mathbf{x}$ can be represented as the multiplication of a sparsifying basis $\mathbf{G} \in \mathbb{R}^{N \times N}$ and a sparse signal $\mathbf{s} \in \mathbb{R}^{N \times 1}$, i.e., $\mathbf{x} = \mathbf{Gs}$. Then, Eq. (1) becomes $\mathbf{y} = \mathbf{RGs}$. The solution of the sparse signal, $\hat{\mathbf{s}}$, can be retrieved by solving the following minimization problem with nonnegativity constraints:

$$\min_{\mathbf{s}} \left\|\mathbf{y} - \mathbf{RGs}\right\|_2^2 + \gamma \|\mathbf{s}\|_1 \text{ subject to } s_k \geq 0 \text{ for } k = 1, 2, \ldots, N, \tag{3}$$

where γ is the non-negative regularization parameter and $\|\mathbf{s}\|_p$ is defined as $\left(\sum_{k=1}^{N} |s_k|^p\right)^{1/p}$. We use the collection of Gaussian distribution functions for the sparsifying basis $\mathbf{G}$[4]. The linear combination of Gaussian distribution functions represents the line shape of the spectrum. The retrieved spectrum $\hat{\mathbf{x}}$ is $\mathbf{G}\hat{\mathbf{s}}$. There exist open-source program that can be easily accessed to solve the numerical optimization problem[30,31]. In this work, we use the l1_ls package[31] to solve the problem. All the spectral reconstructions were done in MATLAB R2017b with an Intel Core i7-5820 K CPU computer. The reconstruction of a single spectrum was done within ~ 0.1 s.

Before conducting spectral reconstructions, we first measured the transmission functions of MTF filters. A beam from the halogen light source (KLS-150H-LS-150D, Kwangwoo) was fed into a monochromator (MMAC-200, Mi Optics). From the monochromator, a monochromatic light with an FWHM of 4 nm was generated. After passing through a collimator, the collimated monochromatic light was fed into the CMOS monochrome camera (EO-1312M, Edmund optics). Using the CMOS camera, we measured the light intensities with and without the MTF filter array. Then, the transmission T of i-th filter at wavelength λ_k is calculated by:

$$T_i(\lambda_k) = \frac{IWF_i(\lambda_k) - BI_i(\lambda_k)}{IWOF_i(\lambda_k) - BI_i(\lambda_k)}, \tag{4}$$

where IWF_i, $IWOF_i$, and BI_i are intensity with i-th filter, intensity without i-th filter, and background intensity, respectively. Using the monochromator, we could generate series of monochromatic light at the peak locations from 500 to 849 nm with the step of 1 nm. We captured 350 pairs of monochrome images with and without filters in the wavelength range of 500 to 849 nm. Using Eq. (4), we could obtain transmission functions of 36 MTF filters. The transmission functions were calibrated by element-wise multiplication of the spectral response of the CMOS image sensor as shown in Fig. 1e.

To analyze the dual-peak resolution of the fabricated MTF filter array spectrometer, we conducted simulations of dual-peak spectra reconstructions. Figure 3a shows an example of a dual-peak spectrum. The recovery performances were investigated for noisy environments ranging from 10 to 35 dB of signal-to-noise ratios (SNRs), more details are in the Methods section. The root mean squared error (RMSE), which is defined as $\sqrt{\left\|\mathbf{x}_{refer} - \mathbf{x}_{recon}\right\|_2^2 / N}$, was used to evaluate the performances. The result of dual-peak spectra reconstructions with respect to SNRs is shown in Fig. 3b. We considered four kinds of dual-peak spectra. The FWHMs of a peak were 1 and 2 nm, respectively and the gaps between peaks were 2 and 3 nm, respectively. For each kind of dual-peak, spectra were created by changing the location of dual-peak in the wavelength range of 500 to 849 nm. The reconstructions were performed on all these spectra, and RMSEs were calculated. We averaged the RMSEs and regarded the average RMSE value as the performance of the fabricated MTF filter array to reconstruct dual-peak in noisy environments. As shown in Fig. 3b, the average RMSEs of dual-peak with the FWHM of 1 nm and the gap of 2 nm were 0.0268 for 30 dB, 0.0481 for 20 dB, and 0.0643 for 10 dB, respectively. Similar performances were obtained from the other three kinds of dual-peak spectra. We could find that the fabricated MTF filter array performs well to reconstruct dual-peak spectra in noisy environments.

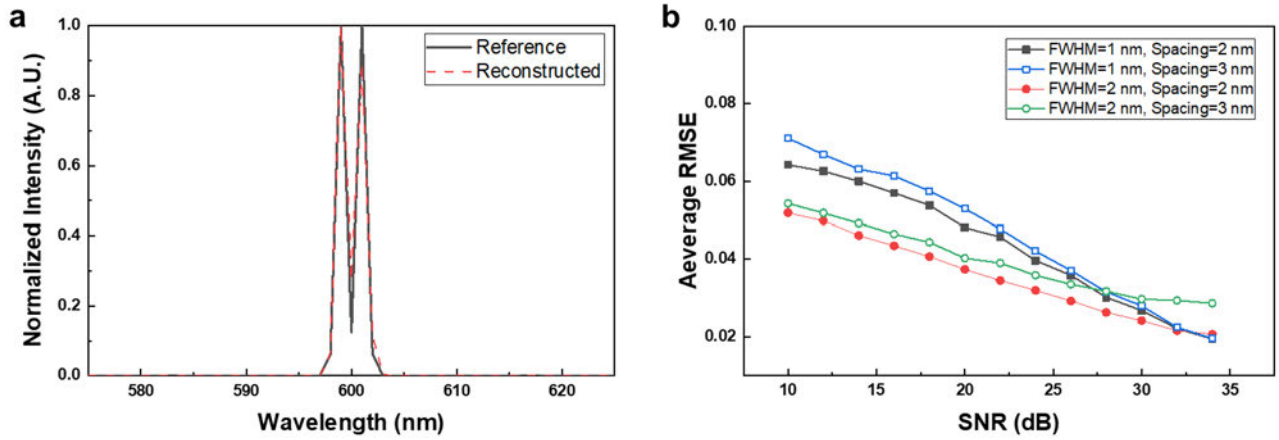


Figure 3. Simulation result of dual-peak spectra reconstructions using MTF filter array. (**a**) An example of dual-peak spectra with an FWHM of 1 nm with 2 nm apart. (**b**) Reconstruction performances of dual-peak spectra with respect to signal-to-noise ratios.

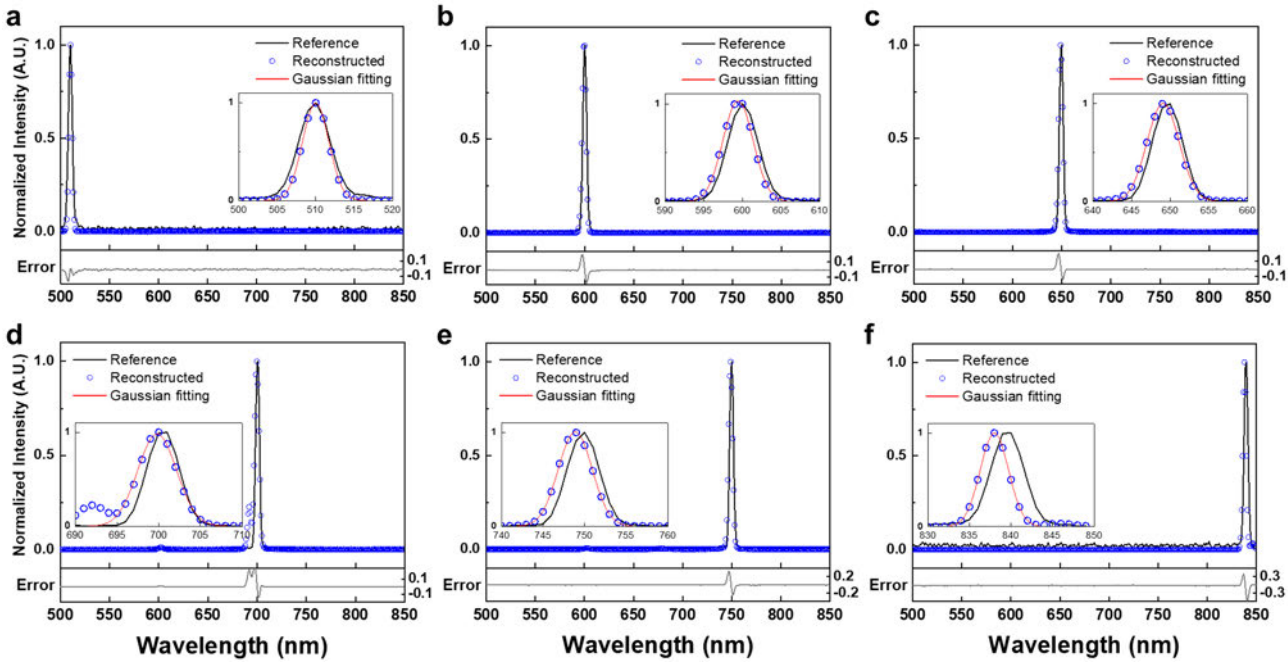


Figure 4. Spectral reconstruction of monochromatic light sources. Monochromatic light with an FWHM of 4 nm at peak wavelengths of (**a**) 510 nm, (**b**) 600 nm, (**c**) 650 nm, (**d**) 700 nm, (**e**) 750 nm, and (**f**) 840 nm. Solid black lines represent reference spectra which are measured by the grating-based spectrometer. Blue circles represent reconstructed spectra using the MTF filter array spectrometer. Solid red lines represent the results of Gaussian fitting. Solid light gray lines represent error between reconstructed and reference spectra.

After analyzing the dual-peak resolution of the MTF filter array in simulations, we tested monochromatic lights by varying the peak wavelength. The CMOS monochrome camera was used for measuring the intensities of test lights. The pixel size of the CMOS image sensor is 5.2×5.2 μm. Underneath each filter, there are approximately 60×60 pixels. However, considering a case where the layer's location mismatch may occur during the fabrication process of the MTF filter, we excluded the boundary pixels. The averaged intensities from 30×30 pixels at the center of each filter were used for the spectral reconstruction experiments. Using a grating-based spectrometer (Black-Comet, StellarNet), monochromatic lights were measured for use as a reference.

Figure 4 shows the reconstruction results for the monochromatic light. For ease of comparison, reference spectra and reconstructed spectra are normalized. Solid black lines and blue circles in Fig. 4 represent reference spectra and reconstructed spectra, respectively. Reference spectra have peak wavelengths at 510, 600, 650, 700, 750 and 840 nm with FWHMs of 4 nm. As depicted in the inset enlarged graph, the reconstructed spectra using the MTF filter array spectrometer matched the reference spectra. More specifically, differences of peak wavelengths between reference and reconstructed spectra were within 2 nm. The RMSEs were 0.023, 0.023,

Monochromatic light	510 nm	600 nm	650 nm	700 nm	750 nm	840 nm
Peak Center (nm)	509.995	599.438	649.092	699.808	748.878	838.001
Peak shift (nm)	0.005	0.562	0.908	0.192	1.122	1.999
FWHM (nm)	4.008	4.597	4.887	5.486	4.79	4.003

Table 1. Evaluation of monochromatic lights reconstructions using the Gaussian curve fitting.

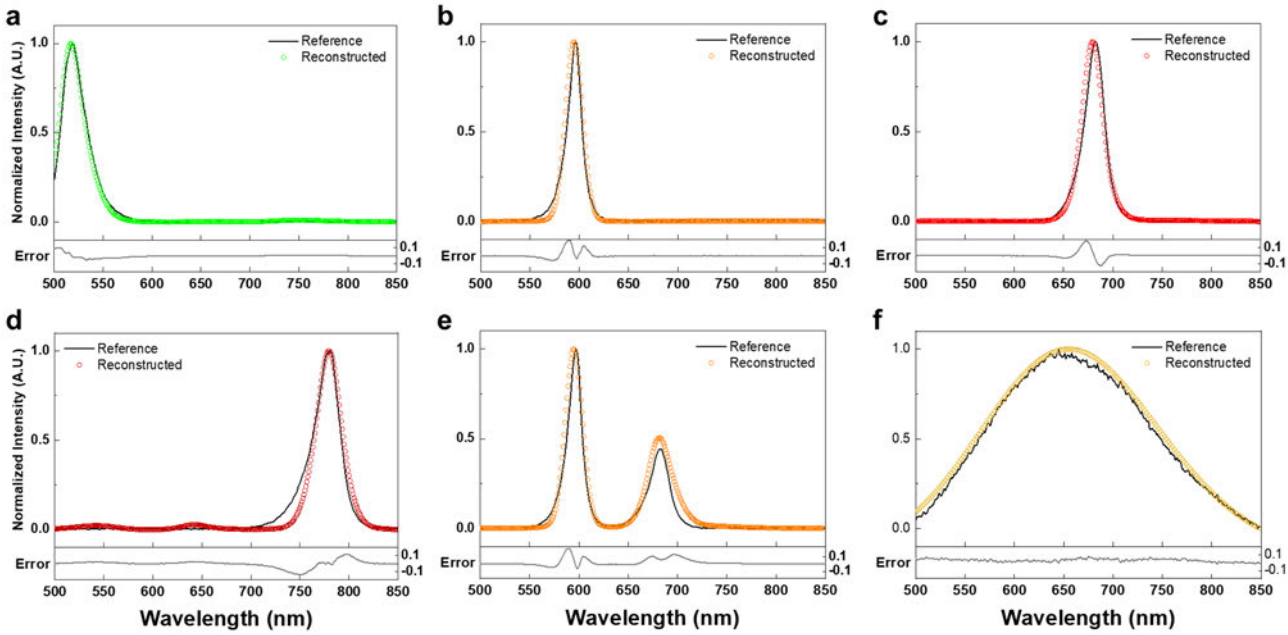


Figure 5. Spectral reconstructions of LEDs and a halogen light source. (**a**) a green LED, (**b**) an orange LED, (**c**) a red LED, (**d**) an infrared LED, (**e**) Combined two LEDs (orange and red), and (**f**) a halogen light source. Solid black lines represent reference spectra which are measured by the grating-based spectrometer. Colored circles represent reconstructed spectra using the MTF filter array spectrometer. Solid light gray lines represent error between reconstructed and reference spectra.

0.021, 0.035, 0.035, and 0.061, for wavelengths 510, 600, 650, 700, 750, and 840 nm, respectively, as shown in Fig. 4a–f. Table 1 presents the evaluation of monochromatic light reconstructions using Gaussian fittings. Over monochromatic lights, peak shifts and FWHMs were within 2 nm and 5.5 nm, respectively. Spectral reconstruction performance seems to degrade in the long-wavelength range due to the low spectral response of the CMOS image sensor and the monotonous spectral features of MTF filters.

We further explored the performance of the MTF filter array spectrometer using broadband light sources, such as LEDs and a halogen light source. Figure 5 shows the spectral reconstruction results. Solid black lines represent reference spectra, which are measured using the grating-based spectrometer. Colored circles represent reconstructed spectra using the MTF filter array spectrometer. Three single-color visible LEDs and one single-color infrared LED were used for spectral reconstruction experiments, as shown in Fig. 5a–d. A green LED (LED 525E, Thorlabs) with an FWHM of 32 nm was reconstructed with an RMSE of 0.021. An orange LED (LED 600L, Thorlabs) with an FWHM of 12 nm was reconstructed with an RMSE of 0.034. A red LED (LED 680L, Thorlabs) with an FWHM of 16 nm was reconstructed with an RMSE of 0.035. An infrared LED (LED 780E, Thorlabs) with an FWHM of 25 nm was reconstructed with an RMSE of 0.044. Similar to experimental results of monochromatic light, the reconstruction performance is relatively poor for spectrum in the long-wavelength range. In addition, we conducted the spectral reconstruction for combined LEDs (an orange LED and a red LED), as shown in Fig. 5e. A beam splitter is used to measure the light of the combined LED. The combined LED was reconstructed with an RMSE of 0.044. Finally, the spectral reconstruction of the halogen light source was conducted. The halogen light source, with an FWHM of 180 nm, was reconstructed with an RMSE of 0.034, as shown in Fig. 5f. As evidenced by low RMSE values, reconstructed spectra agree well with reference spectra measured by the grating-based spectrometer.

Computational pinhole spectral imaging. Furthermore, we demonstrated spectral imaging using the MTF filter array. As shown in Fig. 6a, the pinhole imaging system was constructed by combining a pinhole (Edmund optics), whose aperture diameter is 150 μm, with the monochrome CMOS image camera. The MTF filter array was placed in front of the pinhole. A single filter was adjusted to the pinhole to allow an incident image to pass through the filter and pinhole, and the filtered image was measured by the CMOS image sensor.

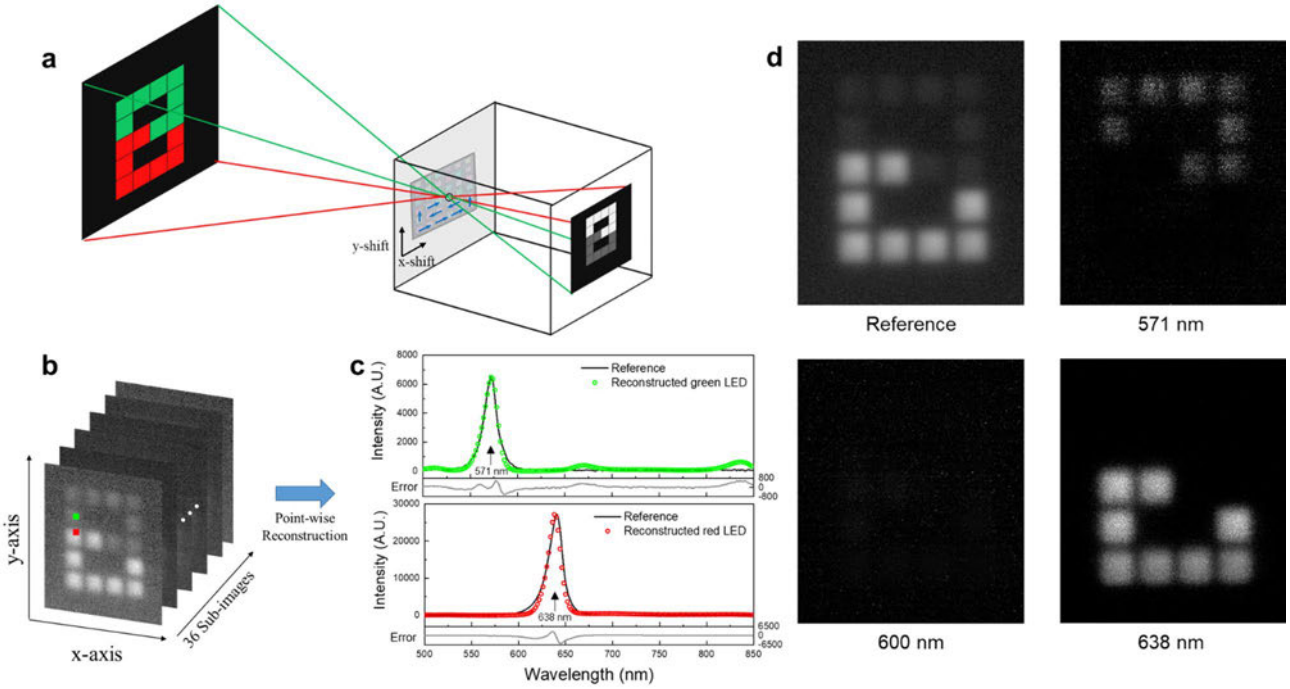


Figure 6. Computational pinhole spectral imaging. (**a**) Schematic of pinhole imaging; The MTF filter array is placed in front of the pinhole camera. A single filter is adjusted to the pinhole and the filtered image is acquired. By changing filters, 36 filtered images are obtained. (**b**) 36 filtered images of 8 × 8 LED matrix showing the number "8". The upper part consists of green LEDs and the lower part consists of red LEDs. (**c**) Point-wise spectral reconstruction: a pixel of a green LED block and a red LED block which are denoted in (**b**). Solid light gray lines represent error between reconstructed and reference. (**d**) Monochrome image of reference and reconstructed monochrome images at 571, 600 and 638 nm, respectively.

By changing filters using a linear translation stage (Newport), 36 filtered images are obtained. Bi-Color 8 × 8 LED matrix (Adafruit) was used to generate a target. We made a small display by connecting the LED matrix to an Arduino Uno (Arduino) and by controlling the color of the 64 blocks. The number "8" was represented by the LED cube. The upper blocks consist of green LEDs, and the lower blocks consist of red LEDs. Figure 6b shows a stack of the filtered 36 sub-images. A 1280 × 1024 size image was reduced to a sub-image size of 350 × 300 by discarding unnecessary pixels. Thus, a data cube with 350 × 300 × 36 in size was obtained. Spectral reconstruction was performed for each pixel, and the data cube was restored with a size of 350 × 300 × 350. It took ~ 1.8 h to reconstruct the data cube. As shown in Fig. 6c, the reference spectra measured by the grating-based spectrometer are shown in solid black lines. As denoted pixels in Fig. 6b, the reconstructed spectra of a pixel in the green LED block and a pixel in the red LED block are represented in Fig. 6c as green circles and red circles, respectively. The RMSE was calculated after normalizing reference spectra and reconstructed spectra. The green LED with an FWHM of 15 nm was reconstructed with an RMSE of 0.0315. The red LED with an FWHM of 20 nm was reconstructed with an RMSE of 0.0370. Figure 6d shows the monochrome image of reference and reconstructed monochrome images at 571, 600, and 638 nm. The pinhole imaging system also measured the monochrome image of reference without the MTF filter array. Since the spectral component of the red LED does not exist at 571 nm, only the upper blocks of the number "8" are shown in the reconstructed monochrome image at 571 nm. On the other hand, only the lower blocks of the number "8" are shown in the reconstructed monochrome image at 638 nm, where the spectral component of the green LED does not exist. Finally, nothing is displayed in the reconstructed monochrome image at 600 nm, where spectral components of the green and red LED do not exist.

As proof-of-principle of the spectral imaging, we implemented the spectral scanning method on the pinhole imaging system. While reconstructed spectra of the pinhole spectral imaging match well with reference spectra, there are improvements to consider. In the spectral scanning method, the data cube acquisition time is long so that spectral smearing may occur in the case of a moving target. The non-scanning method, such as in snapshot spectral imaging systems, can solve these problems by acquiring the data cube in a single exposure. We assume that it is possible to construct a snapshot spectral imaging system combining the MTF filter array and a thin observation module by bound optics (TOMBO)[32–34] structure. This spectral imaging system requires a microlens array and a single separator but does not need the MTF filter to be as small as pixel size. Rather, the MTF filter should be made large so that many pixels are underlying the filter. The MTF filter array can be fabricated in scalable using stencil lithography techniques according to the spatial resolution of the spectral imaging system.

Discussion

In conclusion, we mass-produced MTF filter arrays using stencil lithography and experimentally demonstrated the spectral resolvability of an MTF filter array-based computational spectrometer. 169 identical filter arrays with 36 MTF filters were fabricated on a single wafer. Although the MTF filter size was larger than that of the photonic crystal slabs[14,16], it can be improved to a smaller size by using advanced lithography techniques and facilities. In addition, by using a higher refractive index material, the number of layers of the MTF filter can be reduced so that manufacturing efficiency can be improved.

Using the random spectral features of MTF filters and numerical optimization techniques, we recover varied spectra from the visible range to the near-infrared range (500 to 849 nm) with 1 nm spacing. The spectral reconstruction performance in the near-infrared range is relatively inferior to the visible range, but it can be further improved by using a CMOS image sensor with a high spectral response in the near-infrared region. Also, computational spectral imaging with the MTF filter array was demonstrated using the spectral scanning method. The reconstructed data cube was found to match well with spatial and spectral references. However, to use the spectral imaging system in mobile applications, a shorter data cube acquisition time is required. By utilizing the TOMBO structure with the MTF filter array, it is possible to construct a snapshot spectral imaging system that has a short acquisition time.

Finally, the production of the MTF filter arrays is an important step towards the industrialization and practical uses of computational spectrometers. This study will be helpful for computational spectroscopy to be used in various applications where compact size, high resolution, and wide working range are required.

Methods

Simulation details. Using Gaussian distribution functions, we generated a dual-peak spectrum $\mathbf{x}$ as shown in Fig. 3a. The spectrum $\mathbf{x}$ is computationally measured as $\mathbf{y}$ by multiplying the sensing matrix $\mathbf{R}$, i.e., $\mathbf{y} = \mathbf{Rx}$. In addition, we made noisy measurement $\tilde{\mathbf{y}}$ by adding additive noise $\mathbf{n}$ as $\tilde{\mathbf{y}} = \mathbf{y} + \mathbf{n} = \mathbf{Rx} + \mathbf{n}$. SNR in decibels is defined as $10 \log \left(\|\mathbf{x}\|_2^2 / N\sigma^2 \right)$ where σ is the standard deviation of the noise.

Data availability

Raw data is available from the corresponding author upon reasonable request.

Received: 6 September 2021; Accepted: 8 February 2022
Published online: 08 March 2022

References

1. Clark, R. N. & Roush, T. L. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *J. Geophys. Res. Solid Earth* **89**, 6329–6340 (1984).
2. Bacon, C. P., Mattley, Y. & DeFrece, R. Miniature spectroscopic instrumentation: Applications to biology and chemistry. *Rev. Sci. Instrum.* **75**, 1–16 (2004).
3. Kim, S. *et al.* Smartphone-based multispectral imaging: System development and potential for mobile skin diagnosis. *Biomed. Opt. Express* **7**, 5294–5307 (2016).
4. Kurokawa, U., Choi, B. I. & Chang, C.-C. Filter-based miniature spectrometers: Spectrum reconstruction using adaptive regularization. *IEEE Sens. J.* **11**, 1556–1563 (2011).
5. Oliver, J., Lee, W., Park, S. & Lee, H.-N. Improving resolution of miniature spectrometers by exploiting sparse nature of signals. *Opt. Express* **20**, 2613–2625 (2012).
6. Oliver, J., Lee, W.-B. & Lee, H.-N. Filters with random transmittance for improving resolution in filter-array-based spectrometers. *Opt. Express* **21**, 3969–3989 (2013).
7. Wang, Z. & Yu, Z. Spectral analysis based on compressive sensing in nanophotonic structures. *Opt. Express* **22**, 25608–25614 (2014).
8. Bao, J. & Bawendi, M. G. A colloidal quantum dot spectrometer. *Nature* **523**, 67–70 (2015).
9. Yang, T. *et al.* Miniature spectrometer based on diffraction in a dispersive hole array. *Opt. Lett.* **40**, 3217–3220 (2015).
10. Huang, E., Ma, Q. & Liu, Z. Etalon array reconstructive spectrometry. *Sci. Rep.* https://doi.org/10.1038/srep40693 (2017).
11. Oiknine, Y., August, I., Blumberg, D. G. & Stern, A. Compressive sensing resonator spectroscopy. *Opt. Lett.* **42**, 25–28 (2017).
12. Cerjan, B. & Halas, N. J. Toward a nanophotonic nose: A compressive sensing-enhanced, optoelectronic mid-infrared Spectrometer. *ACS Photonics* **6**, 79–86 (2018).
13. Kim, C., Lee, W.-B., Lee, S. K., Lee, Y. T. & Lee, H.-N. Fabrication of 2D thin-film filter-array for compressive sensing spectroscopy. *Opt. Lasers Eng.* **115**, 53–58 (2019).
14. Wang, Z. *et al.* Single-shot on-chip spectral sensors based on photonic crystal slabs. *Nat. Commun.* **10**, 1020 (2019).
15. Yang, Z. *et al.* Single-nanowire spectrometers. *Science* **365**, 1017–1020 (2019).
16. Zhu, Y., Lei, X., Wang, K. X. & Yu, Z. Compact CMOS spectral sensor for the visible spectrum. *Photonics Res.* **7**, 961–966 (2019).
17. Yang, Z., Albrow-Owen, T., Cai, W. & Hasan, T. Miniaturization of optical spectrometers. *Science* **371**, 0722 (2021).
18. August, Y. & Stern, A. Compressive sensing spectrometry based on liquid crystal devices. *Opt. Lett.* **38**, 4996–4999 (2013).
19. Kwak, Y., Park, S. M., Ku, Z., Urbas, A. & Kim, Y. L. A pearl spectrometer. *Nano Lett.* **21**, 921 (2020).
20. Liu, C. & Sun, Z. Design and fabrication of a metallic irregular F-P filter array for a miniature spectrometer. *Appl. Opt.* **60**, 4948–4953 (2021).
21. Li, H. *et al.* A near-infrared miniature quantum dot spectrometer. *Adv. Opt. Mater.* **9**, 2100376 (2021).
22. Baraniuk, R. G. Compressive sensing [lecture notes]. *IEEE Signal Process. Mag.* **24**, 118–121 (2007).
23. Donoho, D. L. Compressed sensing. *IEEE Trans. Inf. Theory* **52**, 1289–1306 (2006).
24. Barry, J. R. & Kahn, J. M. Link design for nondirected wireless infrared communications. *Appl. Opt.* **34**, 3764–3776 (1995).
25. Macleod, H. A. *Thin-Film Optical Filters* (CRC Press, 2001).
26. Koh, K., Kim, S.-J. & Boyd, S. An interior-point method for large-scale l1-regularized logistic regression. *J. Mach. Learn. Res.* **8**, 1519–1555 (2007).
27. Bioucas-Dias, J. M. & Figueiredo, M. A. A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Trans. Image Process.* **16**, 2992–3004 (2007).
28. Wright, S. J., Nowak, R. D. & Figueiredo, M. A. Sparse reconstruction by separable approximation. *IEEE Trans. Signal Process.* **57**, 2479–2493 (2009).

29. Wang, S.-W. *et al.* 128 channels of integrated filter array rapidly fabricated by using the combinatorial deposition technique. *Appl. Phys. B* **88**, 281–284 (2007).
30. Candes, E. & Romberg, J. *l1-magic: Recovery of Sparse Signals Via Convex Programming*, Vol. 4, 14 (2005) www.acm.caltech.edu/l1magic/downloads/l1magic.pdf.
31. Koh, K., Kim, S. & Boyd, S. *l1 ls: A Matlab Solver for Large-Scale ℓ1-Regularized Least Squares Problems* (2008).
32. Tanida, J. *et al.* Thin observation module by bound optics (TOMBO): Concept and experimental verification. *Appl. Opt.* **40**, 1806–1813 (2001).
33. Gupta, N., Ashe, P. R. & Tan, S. Miniature snapshot multispectral imager. *Opt. Eng.* **50**, 033203 (2011).
34. Geelen, B., Tack, N. & Lambrechts, A. A snapshot multispectral imager with integrated tiled filters and optical duplication. In *Advanced Fabrication Technologies for Micro/Nano Optics and Photonics VI*, Vol. 8613, 861314 (International Society for Optics and Photonics, 2013).

Acknowledgements

This work was supported by a National Research Foundation of Korea (NRF) Grant funded by the Korean government (MSIP) (NRF-2021R1A2B5B03002118).

Author contributions

C.K. and H.-N.L. conceptualized the idea. C.K., P.N. and K.R.L. performed the design and fabrication of the MTF filter arrays. C.K. and P.N. conducted optical experiments and reconstructions. C.K. visualized the experiments results. All authors wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-08037-y.

Correspondence and requests for materials should be addressed to H.-N.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022

sensors

MDPI

Article

Compressive Sensing Spectroscopy Using a Residual Convolutional Neural Network

Cheolsun Kim, Dongju Park and Heung-No Lee *

School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, Korea; csk0315@gist.ac.kr (C.K.); toriving@gist.ac.kr (D.P.)

* Correspondence: heungno@gist.ac.kr; Tel.: +82-62-715-2237

Received: 20 December 2019; Accepted: 20 January 2020; Published: 21 January 2020



Abstract: Compressive sensing (CS) spectroscopy is well known for developing a compact spectrometer which consists of two parts: compressively measuring an input spectrum and recovering the spectrum using reconstruction techniques. Our goal here is to propose a novel residual convolutional neural network (ResCNN) for reconstructing the spectrum from the compressed measurements. The proposed ResCNN comprises learnable layers and a residual connection between the input and the output of these learnable layers. The ResCNN is trained using both synthetic and measured spectral datasets. The results demonstrate that ResCNN shows better spectral recovery performance in terms of average root mean squared errors (RMSEs) and peak signal to noise ratios (PSNRs) than existing approaches such as the sparse recovery methods and the spectral recovery using CNN. Unlike sparse recovery methods, ResCNN does not require *a priori* knowledge of a sparsifying basis nor prior information on the spectral features of the dataset. Moreover, ResCNN produces stable reconstructions under noisy conditions. Finally, ResCNN is converged faster than CNN.

Keywords: spectroscopy; compressed sensing; deep learning; inverse problems; sparse recovery; dictionary learning

1. Introduction

There has been considerable interest in producing compact spectrometers having a high spectral resolution, wide working range, and short measuring time. Such a spectrometer can be used in a broad range of fields such as remote sensing [1], forensics [2], and medical applications [3]. Spectrometers that exploit advanced signal-processing methods are promising candidates. The compressive sensing (CS) [4,5] framework makes it possible for a spectrometer to improve its spectral resolution while retaining its compact size. CS spectroscopy comprises two parts: Capturing a spectrum with a small number of compressed measurements and reconstructing the spectrum from the compressed measurements using reconstruction techniques.

To date, for effective signal recovery in CS spectroscopy, three requirements should be satisfied. First, the spectrum should be a sparse signal or capable of sparse representation on a certain basis. Second, the sensing patterns of optical structures should be designed to have a small mutual coherence [6]. Third, appropriate reconstruction algorithms are required. Note that several sparsifying bases have been used in CS spectroscopy such as a family of orthogonal Daubechies wavelets [7], a Gaussian line shape matrix [8,9], and a learned dictionary [10]. Furthermore, numerous optical structures have been proposed to attain the necessary small mutual coherence for sensing patterns such as thin-film filters [11,12], a liquid crystal phase retarder [13], Fabry–Perot filters [7,14], and photonic crystal slabs [15,16]. As algorithms for reconstructing the original signal, two types of basic reconstruction techniques have been developed: greedy iterative algorithms [17,18] and convex relaxation [19,20]. In CS spectroscopy, the reconstruction algorithms have been used with a sparsity

constraint. Additionally, a non-negativity constraint is used in Reference [16,21]. Combining these three considerations, CS spectrometers have shown stable performance for light-emitting diodes (LEDs) and monochromatic lights.

Since not all signals can be represented as sparse on a fixed basis, prior information on structural features of the spectral dataset is therefore required to generate a best-fit sparsifying basis. Furthermore, a high computational cost is required for reconstruction techniques. Recently, deep learning [22] has been emerging as a promising alternative framework for reconstructing the original signal from the compressed measurements.

Mousavi et al. [23] was the first study on image recovery from structured measurements using deep learning. Moreover, a deep-learning framework for inverse problems has been applied in biomedical imaging for imaging through scattering media [24], magnetic resonance imaging [25,26], and X-ray computed tomography [27]. Kim et al. [28] reported the first attempt to use deep learning in CS spectroscopy. They trained a convolutional neural network (CNN) to output the reconstructed signal from the network. From here on the network reported by Kim et al. will be referred to as CNN.

Unlike CNN [28] in which learnable layers were simply stacked and trained to directly reconstruct the original spectrum, we make a residual connection [29] between the input and output of CNN and train the network to reconstruct the original spectrum by referring the input of the network. As a result, the network learns residuals between the input of the network and the original spectrum. It has been reported that it is easier to train a network when using residual connections than to train a plain network that was simply stacked with learnable layers [25,29]. Lee et al. [25] analyzed the topological structure of magnetic resonance (MR) images and the residuals of MR images. They showed that the residuals possessed a simpler topological structure, thus making learning residuals easier than learning the original MR images. In addition, He et al. [29] demonstrated with empirical results that the residual networks are easy to optimize and they achieved improvements in image-recognition tasks. From these works, we gain insights such that adding residual connections to CNN would improve the spectral reconstruction performance in CS spectroscopy.

In this paper, we aim to propose a novel residual convolutional neural network (ResCNN) for recovering an input spectrum from the compressed sensing measurements in CS spectroscopy. The novelty lies in the proposed ResCNN structure, with a moderate depth of learnable layers and a single residual connection, which provides the desired spectral reconstruction performance. The desired performance here means that the proposed ResCNN offers a performance which is better than that of CNNs as well as that of CS reconstruction with its sparsifying base known. In CS reconstruction, the prior knowledge of a fixed sparsifying basis is useful and offers good sparse representation results. However, in general it is a difficult problem to identify a sparsifying basis for various kinds of spectra and apply the identified basis to have the recovery performance improved. In this regard, it is an important advance to find a simple ResCNN which offers good enough performance. It is also worth to note that the proposed ResCNN is tested with the array type CS spectroscopy, discussed in Section 2, which we have designed with an array of multilayer thin-film filters.

The previous works on CS spectroscopy [7,11,13,14,16] have shown decent reconstruction performance but on limited simple sources such as LEDs and monochromatic lights. Using ResCNNs, we are now able to reconstruct more complex spectra, such as spectra with multiplicity of peaks mixed with a gradual rise-and-fall.

The remainder of this paper is organized as follows. In Section 2, we model the optical structure which is used for CS spectroscopy. In Section 3, we describe the system of CS spectroscopy and the proposed ResCNN. In Section 4, simulated experiments are described. Section 5 presents the results of experiments. In Section 6, we discuss the results. Finally, we conclude this paper in Section 7.

2. Optical Structure

Numerous optical structures have been proposed for CS spectroscopy. It has been reported that CS spectrometers, which have various spectral features in the transmission spectrum, show high

spectral-resolving performance [16]. In this work, we used thin-film filters to model CS spectrometers. Thin-film filters demonstrate a variety of spectral features depending on the materials used, the number of layers, and the thicknesses of the layers. Once the structure of thin-film is determined, a transmission value at a given wavelength λ is defined as follows [30]:

$$T(\lambda) = 1 - \frac{1}{2}\left(\left|\rho_{TE}(\lambda)\right|^2 + \left|\rho_{TM}(\lambda)\right|^2\right), \quad (1)$$

where $\rho_{TE}(\lambda)$ and $\rho_{TM}(\lambda)$ are amplitude reflection coefficients. The coefficients represent the fraction of the power reflected by a multilayer thin-film in the transverse electric (TE) and transverse magnetic (TM) modes of an incident light, respectively. We summarized recursive processes for calculating amplitude reflection coefficients in Algorithm 1 [11,12,31].

Algorithm 1: Recursive processes for amplitude reflection coefficients.

Input:	λ
	Structure parameters: θ_1,$\mathbf{n} = \{n_1, n_2, \cdots, n_l\}$,$\mathbf{d} = \{d_2, d_3, \cdots, d_l\}$.
Step 1:	Calculate θ_k,β_k, and N_k using structure parameters.
	$\theta_k = \sin^{-}\left(\frac{n_{k-1}}{n_k}\sin\theta_{k-1}\right)$, for k = 2, 3, ..., l.
	$\beta_k = 2\pi\cos(\theta_k)n_k d_k/\lambda$, for k = 2, 3, ..., l.
	$N_k = \begin{cases} n_k/\cos\theta_k & for \quad TE \\ n_k\cos\theta_k & for \quad TM \end{cases}$, for k = 2, 3, ..., l.
Step 2:	Obtain η_2 by setting $\eta_l = N_l$.
	For k = l-1 to 2
	$\eta_k = N_k\frac{\eta_{k+1}\cos\beta_k + jN_k\sin\beta_k}{N_k\cos\beta_k + j\eta_{k+1}\sin\beta_k}$.
Step 3:	Compute $\rho = (N_1 - \eta_2)/(N_1 + \eta_2)$.
Output:	ρ

Here, θ_k is the angle of an incident light passing from k^{th} to k+1th layer. The refractive index of kth layer is denoted as n_k. d_k denotes the thickness of the kth layer. Given a wavelength vector $\boldsymbol{\lambda} = (\lambda_1\,\lambda_2\,\ldots\,\lambda_N) \in \mathbb{R}^{1\times N}$ in the range of interest, i.e., $\lambda_{\max} - \lambda_{\min}$. Let $\Delta\lambda = \frac{\lambda_{\max}-\lambda_{\max}}{N}$. Then, evaluating the function at the integer multiple of $\Delta\lambda$, i.e., $T(\lambda = \lambda_{\min} + n\Delta\lambda)$ for $n = 0, 1, \cdots, N-1$, we obtain the vector of transmission spectrum $\mathbf{T}_m \in \mathbb{R}^{1\times N}$ for the wavelength range. Then, the sensing pattern matrix of optical structures $\mathbf{T} \in \mathbb{R}^{M\times N}$ is obtained by repeating the calculation of $\mathbf{T}_m$ for $m = 1, 2, \cdots, M$.

We have used SiNx and SiO_2 for high- and low-refractive index materials, respectively. We numerically generated thin-film filters by alternately stacking high- and low- refractive index materials, changing the number of layers, and varying the thickness of each layer. The number of layers in each filter is in the interval of (19, 24), and the thickness (nm) of each layer is in the interval of (50, 300). Initially, we randomly generated reference filters and compute the mutual coherence among the filters. Then, new filters were generated by changing thicknesses of the layers and the mutual coherence of the filters is compared to the mutual coherence of reference filters. Filters with a smaller mutual coherence then became the new reference filters. This process is repeated until reasonable reference filters with the required small mutual coherence are obtained.

Figure 1 shows the heatmap for the transmission spectra of the reference filters and two selected transmission spectra. In Figure 1a, each of the transmission spectra shows a unique sensing pattern because of the iterative modeling process of the reference filters based on mutual coherence. Figure 1b shows two transmission spectra that correspond to the 15th and 30th rows in the heatmap of reference filters. The transmission spectrum reveals a deep spectral modulation depth and various features such as broadband backgrounds, multiple peaks with a small full width at half maximums (FWHMs), and irregular fluctuations.

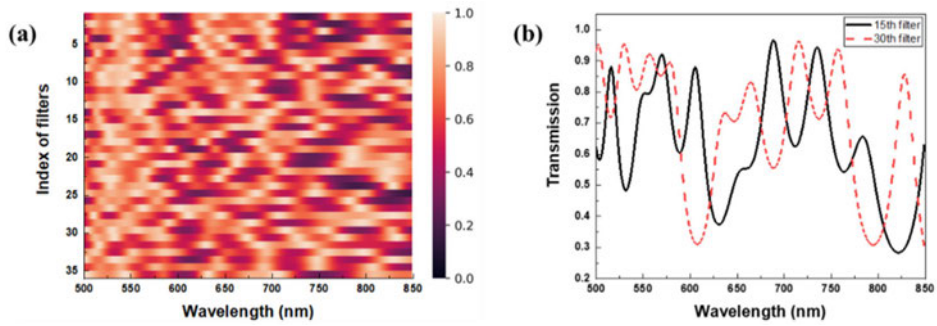


Figure 1. (**a**) Heatmap of the sensing matrix: each row represents the transmission spectrum of the designed thin-film filter. (**b**) Two transmission spectra corresponding to the 15th and 30th rows in the sensing matrix.

3. Compressive Sensing (CS) Spectrometers Using the Proposed Residual Convolutional Neural Network (ResCNN)

3.1. CS Spectrometers

In CS spectroscopy, the measurement column vector $\mathbf{y} \in \mathbb{R}^{M\times 1}$ is represented using the following relation:

$$\mathbf{y} = \mathbf{T}\mathbf{x}, \tag{2}$$

where $\mathbf{x} \in \mathbb{R}^{N\times 1}$ is the spectrum column vector of incident light and $\mathbf{T} \in \mathbb{R}^{M\times N}$ is the sensing matrix of the optical structure. Each row of $\mathbf{T}$ represents a transmission spectrum. Because the length of the measurement vector is smaller than the length of the spectrum vector ($M < N$), the system is underdetermined. Conventionally, if $\mathbf{x}$ is a sparse signal or can be sparsely represented in a certain basis, i.e., $\mathbf{x} = \mathbf{\Phi s}$, reconstruction algorithms can determine a unique sparse solution $\hat{\mathbf{S}}$ from the following optimization problem:

$$\min_{\mathbf{s}} \left\| \mathbf{T\Phi s} - \mathbf{y} \right\|_2^2 + \tau \|\mathbf{s}\|_1, \tag{3}$$

where $\mathbf{\Phi} \in \mathbb{R}^{N\times N}$ is a sparsifying basis and τ is a regularization parameter. Here, $\mathbf{s}$ is a sparse signal whose components are zero except for a small number of non-zero components. Then, the recovered spectrum $\hat{\mathbf{x}}$ is $\mathbf{\Phi}\hat{\mathbf{s}}$. In this paper, we refer to the methods of solving the optimization problem using Equation (3) as sparse recovery.

Typically, except for narrow-band spectra, a spectrum is not a sparse signal, and a fixed sparsifying basis cannot transform all spectra into sparse signals. Clearly, the use of a fixed basis may lead the sparse recovery to struggle, as no fixed basis will transform every signal into a sparse signal. In addition, the sparse recovery is time-consuming and takes a high computational cost.

Our goal is to overcome the limitations of the sparse recovery in CS spectroscopy and recover various kinds of spectra using ResCNN. Figure 2 shows the schematic of the CS spectroscopy system using ResCNN. This system consists of two parts: compressive sampling and dimension extension, and the reconstruction using ResCNN. In the compressive sampling and dimension extension, the measurement vector $\mathbf{y}$ is obtained from Equation (1), which then transforms into $\widetilde{\mathbf{x}} \in \mathbb{R}^{N\times 1}$ using a linear transformation. A transform matrix $\mathbf{A} \in \mathbb{R}^{N\times M}$ extends the M dimension of $\mathbf{y}$ to N dimension of $\widetilde{\mathbf{x}}$, where $\widetilde{\mathbf{x}}$ is a representative spectrum corresponding to $\mathbf{x}$. We used $\widetilde{\mathbf{x}}$ as the input for the reconstruction. ResCNN learnt a non-linear mapping between $\widetilde{\mathbf{x}}$ and $\mathbf{x}$, and afforded a reconstructed spectrum $\hat{\mathbf{x}} \in \mathbb{R}^{N\times 1}$. The dimension extension by the transform matrix was used to make it easier for ResCNN to extract features and reconstruct spectra from the non-linear mapping.

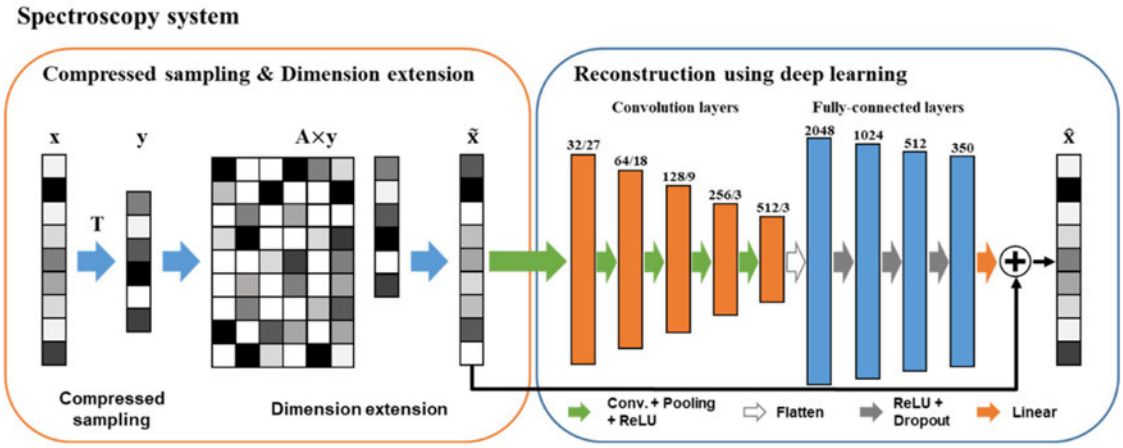


Figure 2. Overview of compressive sensing (CS) spectrosocopy system including the proposed residual convolutional neural network (ResCNN): An input spectrum is compressively sampled by the sensing matrix, and the dimension of measurements is extended by the transform matrix. ResCNN is trained to recover the original spectrum from the extented measurements.

3.2. The Proposed ResCNN

As depicted in Figure 2, ResCNN comprises nine learnable layers, five of which are convolution layers, four are fully-connected layers, and one is a residual connection. Convolution layers are used for the feature extraction in the non-linear mapping between $\widetilde{\mathbf{x}}$ and $\mathbf{x}$. Fully-connected layers are used for the spectra reconstruction. Each of the convolution layers has a set of one-dimensional learnable kernels with specific window sizes. The number of kernels and the window sizes are indicated in Figure 2. After every convolutional layer, the rectified linear unit (ReLU) is used as an activation function, and the subsampling is then applied. We use non-overlapping max-pooling to down-sample the output of the activation function. We stack the convolutional layer, the ReLU, and the subsampling five times. The output of the last subsampling is flattened and then fed into the subsequent four fully-connected layers. The first three layers are followed by the ReLU and dropout in sequence. The dropout is introduced to reduce the overfitting of ResCNN. The output of the last fully-connected layer is fed into a linear activation function. The number of units in each of the fully-connected layers is noted in Figure 2. Unlike CNN [28] in which learnable layers are simply stacked, we make the residual connection that the representative spectrum $\widetilde{\mathbf{x}}$ and the output of the linear activation function are added up to the reconstructed spectrum $\hat{\mathbf{x}}$. Consequently, $\hat{\mathbf{x}}$ is trained to become $\mathbf{x}$. Given training data $\left\{\mathbf{x}_t{}^i\right\}_{i=1}^k$, we train ResCNN to minimize a loss function L. We use the mean squared error between the original $\mathbf{x}_t$ and recovered $\hat{\mathbf{x}}_t$ as the loss function:

$$L = \frac{1}{k}\sum_{i=1}^{k}\left\|\mathbf{x}_t{}^i - \hat{\mathbf{x}}_t{}^i\right\|_2^2. \tag{4}$$

The non-linear mapping that $\widetilde{\mathbf{x}}$ becomes $\mathbf{x}$ can be defined as $H(\widetilde{\mathbf{x}}) = \mathbf{x}$. Because of the residual connection in ResCNN, $H(\widetilde{\mathbf{x}})$ can be rewritten as $H(\widetilde{\mathbf{x}}) = F(\widetilde{\mathbf{x}}) + \widetilde{\mathbf{x}}$, where $F(\widetilde{\mathbf{x}})$ is the mapping of the learnable layers. The representative spectrum $\widetilde{\mathbf{x}}$ is referenced by the residual connection, and then,$F(\widetilde{\mathbf{x}}) = H(\widetilde{\mathbf{x}}) - \widetilde{\mathbf{x}}$. In particular, the mapping of $F(\widetilde{\mathbf{x}})$ is called a residual mapping; therefore, the learnable layers learn the residual of $\mathbf{x}$ and $\widetilde{\mathbf{x}}$.

The previous researches [25,29] have used numerous residual connections in very deep neural networks in order to make networks converge faster by avoiding vanishing gradient problems. We use one residual connection between input and output of the moderate depth network. Figure 3 depicts the manner in which a spectrum is recovered in CNN and ResCNN. The learnable layers of CNN directly reconstruct the spectrum from the representative spectrum $\widetilde{\mathbf{x}}$. Alternatively, ResCNN reconstructs the

spectrum by passing the representative spectrum $\tilde{\mathbf{x}}$ through the residual connection shown in Figure 3b. Consequently, the learnable layers of ResCNN learn to reconstruct residuals.

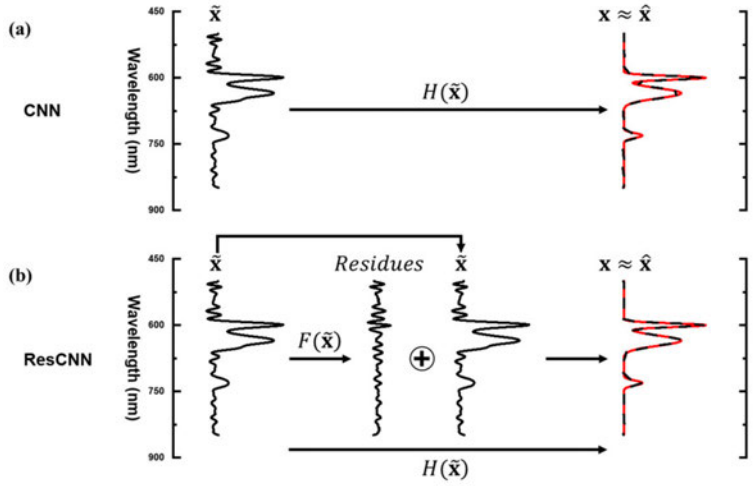


Figure 3. Descriptions of the spectrum recovery process: (**a**) convolutional neural network (CNN), (**b**) ResCNN.

4. Simulated Experiments

We reconstructed 350 spectral bands (N = 350) using 36 thin-film filters (M = 36) whose sensing patterns have a spacing of 1 nm for wavelengths from 500 to 850 nm. We determined the sensing matrix $\mathbf{T}$, assuming that the incident light falls onto the filters with normal incidence. As the transform matrix $\mathbf{A}$, we used the Moore–Penrose inverse of the sensing matrix $\mathbf{T}$, i.e., $\mathbf{A} = \mathbf{T}^T\left(\mathbf{T}\mathbf{T}^T\right)^{-1}$.

4.1. Spectral Datasets

To evaluate the performance of ResCNN, we used two synthetic spectral datasets and two measured spectral datasets. The first synthetic dataset is composed of Gaussian distribution functions while the other is composed of Lorentzian distribution functions. These two synthetic datasets were selected as generally these types of functions are used to represent spectral line shapes. As shown in Figure 4, component functions are added to produce the spectra. We generated 12,000 spectra for each dataset. For each spectrum, the number of component functions was generated using a geometric distribution with the probability parameter p set to 0.3. We added one to the number of component functions to prevent the number of component functions from becoming zero. Then, we randomly set a location, a height, and an FWHM of each peak. To set a peak location (nm), an integer number was randomly selected from a uniform distribution with the interval (500, 849). A random number from a uniform distribution in the interval (0, 1) was used for the height. An integer number for an FWHM (nm) was randomly drawn from a uniform distribution with the interval (2, 50). Finally, all of the component functions were summed to generate the spectrum. The height of each generated spectrum was normalized such that it was mapped from zero to one.

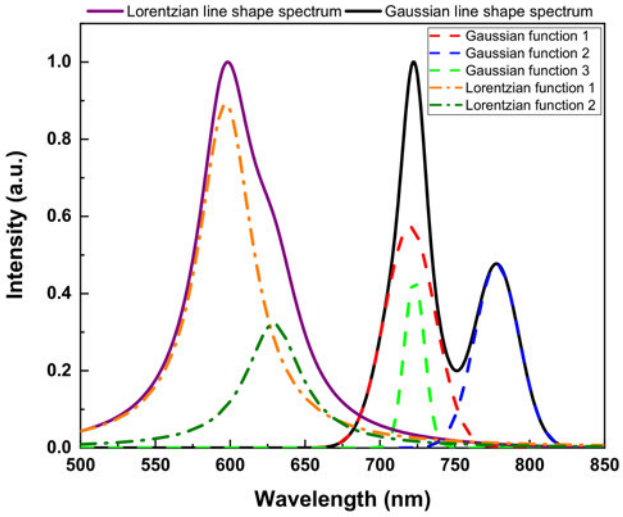


Figure 4. Examples of two synthetic spectra: the solid purple line is composed of two Lorentzian distribution functions (dash-dotted orange and olive lines), and the solid black line is composed of three Gaussian distribution functions (dashed red, blue, and green lines).

As measured datasets, we used the US Geological Survey (USGS) spectral library version 7 [32], and the glossy Munsell colors spectral dataset [33]. The USGS spectral library provides reflectance spectra for artificial materials, coatings, liquids, minerals, organic compounds, soil mixtures, and vegetation. We discarded any spectrum that has missing spectral bands. Then, we extracted the spectrum in the wavelength range of interest (500 to 849 nm) from the wavelength range of the original spectrum (350 to 2500 nm). The measured wavelength range for the glossy Munsell colors spectral dataset, which contains the reflectance spectra of the glossy Munsell color chips, was 380 to 780 nm. The wavelength range of the original spectrum was different from the wavelength range of interest. We decided to use the wavelength range from 400 to 749 nm to ensure each spectrum was set to 350 spectral bands. This selection of wavelengths is reasonable because the wavelengths were located in the center of the wavelength range of the original spectrum, and showed different spectral features with respect to each spectrum. In addition, our aim was to show the reconstruction performance with respect to various kinds of spectra. Finally, each spectrum was normalized such that the height varies from 0 to 1. Overall, 1473 spectra from USGS spectral dataset and 1600 spectra from Munsell color spectral dataset were used for our simulated experiments. Table 1 lists the details of each of the spectral datasets.

Table 1. Description of the spectral datasets.

Dataset	Training/Validation/Test	Avg. Number of Nonzero Values	Description
Gaussian dataset	8000/2000/2000	336.8/350	FWHM (nm) on the interval [2, 50], Height on the interval [0, 1]
Lorentzian dataset	8000/2000/2000	349/350	FWHM (nm) on the interval [2, 50], Height on the interval [0, 1]
US Geological Survey [32]	982/246/245	348.9/350	350–2500 nm, 2151 spectral bands (we use 350 spectral bands in 500–849 nm)
Munsell colors [33]	1066/267/267	349/350	380–780 nm, 401 spectral bands (we use 350 spectral bands in 400–749 nm)

4.2. Data Preprocessing and Training

Given the sensing matrix, the spectral data are compressively sampled as the measurement vector $\mathbf{y}$ shown in Equation (1), and then transformed into the representative spectrum $\widetilde{\mathbf{x}}$ by multiplying the transform matrix $\mathbf{A}$ and $\mathbf{y}$.

In each spectral dataset, the number of training, validation, and test spectra are randomly assigned using a ratio of 4:1:1 for the synthetic and measured data sets, respectively. The validation spectra are used for estimating the number of epochs and tuning the hyper-parameters. To train ResCNN, we used the Adam optimizer [34] implemented in Tensorflow with the batch size of 16 and 250 epochs. The experiments were conducted on an NVIDIA GeForce RTX 2060 graphics processing unit (GPU). Training the architecture can be done in half an hour for each dataset.

4.3. Sparsifying Bases for Spare Recovery

Using sparse recovery, we evaluated the performance of conventional CS reconstructions to benchmark the performance of ResCNN. As shown in Table 1, the spectra for both the synthetic and measured datasets are dense spectra. Therefore, we must transform the spectra into sparse signals to solve Equation (3). In this section, we considered methods to make a sparsifying basis $\mathbf{\Phi}$.

First, we considered a Gaussian line shape matrix as a sparsifying basis. Each column of the matrix comprises a Gaussian distribution function whose length is N. A collection of N Gaussian functions works as a sparsifying basis $\mathbf{\Phi} \in \mathbb{R}^{N \times N}$. We generate two Gaussian line shape matrices. Figure 5 a shows the heatmap images for two Gaussian line shape matrices. Seven different FWHMs are used to generate the Gaussian distributions. Given an FWHM, Gaussian distributions are generated by shifting the peak location using uniform spacing. To create a small dissimilarity between the two Gaussian line shape matrices, two of the seven FWHMs in Gaussian 1 were replaced with other FWHMs, thus producing Gaussian 2, as shown in Figure 5a.

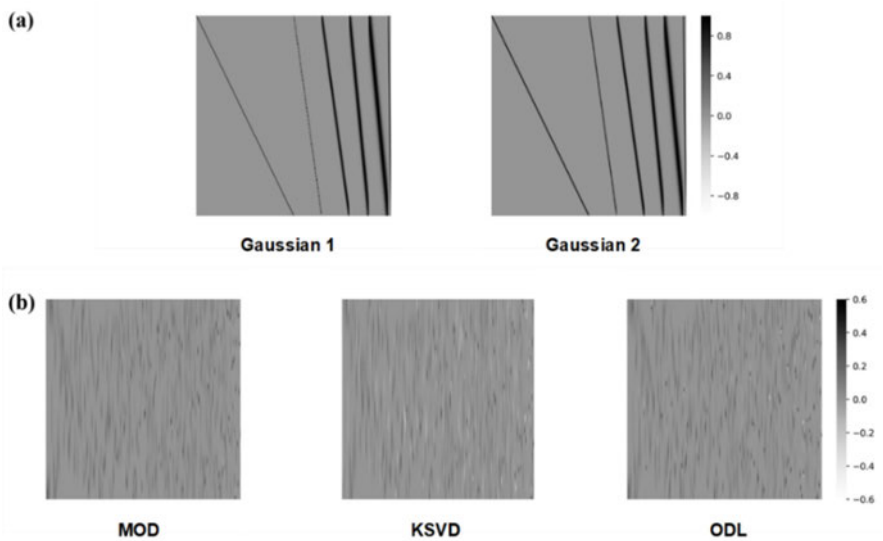


Figure 5. Heatmap images of sparsifying bases that were used in simulated experiments: (**a**) Gaussian line shape matrices, (**b**) the learned dictionaries which are from the Gaussian training dataset.

Second, a learned dictionary [35–38] is used as a sparsifying basis. Given a training dataset $\left\{\mathbf{x}_t{}^i\right\}_{i=1}^k$, we can derive a learned dictionary $\mathbf{\Phi}$ that sparsely represents the training data $\mathbf{x}_t$ by solving the following optimization problem, known as the dictionary learning problem:

$$\min_{\mathbf{\Phi}, \mathbf{s}_t{}^1, \ldots, \mathbf{s}_t{}^k} \sum_{i=1}^{k} \left\| \mathbf{x}_t{}^i - \mathbf{\Phi} \mathbf{s}_t{}^i \right\|_2^2 + \tau \left\| \mathbf{s}_t{}^i \right\|_1, \tag{5}$$

where τ is a regularization parameter and $\mathbf{s}_t{}^i$ is ith sparse signal over the training dataset. By fixing an initial guess for the dictionary $\mathbf{\Phi}$ in Equation (5), we obtain a solution for the sparse signals $\left\{\mathbf{s}_t{}^i\right\}_{i=1}^{k}$. The dictionary is then updated by solving Equation (5) using the sparse signals obtained. This process is iteratively repeated until convergence is reached and we derive the learned dictionary. We used three dictionary learning methods: method of optimal directions (MOD) [36], K-SVD [37], and online dictionary learning (ODL) [38]. The learned dictionaries are generated for each of the training datasets, and the reconstruction performances are evaluated for each test dataset. Figure 5b shows learned dictionaries identified using the Gaussian training dataset. The learned dictionaries clearly depend on the dictionary-learning methods used. Nevertheless, each column of the dictionaries shows a learned spectral feature from the training dataset.

5. Results

To demonstrate the ability of ResCNN to reconstruct spectra, we evaluated its performance using three different datasets: Synthetic datasets, noisy synthetic datasets, and measured datasets. We used the same hyper-parameters of ResCNN for each of these datasets. Moreover, we adopted *l1_ls* [39] as the fixed reconstruction algorithm in the sparse recovery. We compared the recovered signal with the original signal by calculating the root mean squared error (RMSE) and the peak signal to noise ratio (PSNR). In addition, the performance of five conventional sparse recovery methods, described in Section 4.3 and CNN was calculated.

5.1. Synthetic Datasets

The two synthetic data sets described in Table 1 were used to perform the signal recovery using sparse recovery and deep learning. Table 2 shows the average RMSE and PSNR for each of the seven methods evaluated. ResCNN shows the smallest average RMSE for both the Gaussian and Lorentzian datasets of 0.0094 and 0.0073, respectively. Moreover, ResCNN shows the largest average PSNR of 49.0 dB for the Lorentzian dataset. For the Gaussian dataset, the sparse recovery method with Gaussian 2 shows the largest average PSNR, 49.7 dB, which is slightly higher than the 47.2 dB for ResCNN. Note that the minor difference between the two Gaussian line shape matrices results in considerable performance difference. However, reconstruction using the learned dictionaries show similar performance across all of the synthetic datasets.

Table 2. Average root mean squared errors (RMSEs) and peak signal to noise ratios (PSNRs) over synthetic datasets.

	Sparse Recovery					**Deep Learning**	
Dataset	Gaussian 1	Gaussian 2	K-SVD	MOD	ODL	CNN	ResCNN
Gaussian dataset	0.0226 (43.1 dB)	0.0112 (**49.7 dB**)	0.0172 (40.3 dB)	0.0174 (40.3 dB)	0.0161 (41.1 dB)	0.0132 (40.5 dB)	**0.0094** (47.2 dB)
Lorentzian dataset	0.0146 (44.9 dB)	0.0094 (47.5 dB)	0.0136 (42.3 dB)	0.0137 (42.3 dB)	0.0127 (42.9 dB)	0.0101 (42.8 dB)	**0.0073** **(49.0 dB)**

Figure 6 shows the reconstructed test spectra from each of the synthetic datasets. The solid red line (i) is the input spectra from each dataset. ResCNN is shown in dashed black line (ii), while CNN is shown in solid orange lines (iii). The reconstructed spectra using sparse recovery with Gaussian 1 (iv), Gaussian 2 (v), and ODL (vi) are shown in solid green, blue, and purple lines in respectively. Because of the similar performance from each of the learned dictionaries, only the ODL method is shown. The RMSE and PSNR of ResCNN are 0.0138 (37.2 dB) for the spectrum from the Gaussian dataset and 0.0096 (40.4 dB) for the spectrum from the Lorentzian dataset. For the selected spectra, ResCNN achieves superior reconstruction performance compared with the other four reconstructions.

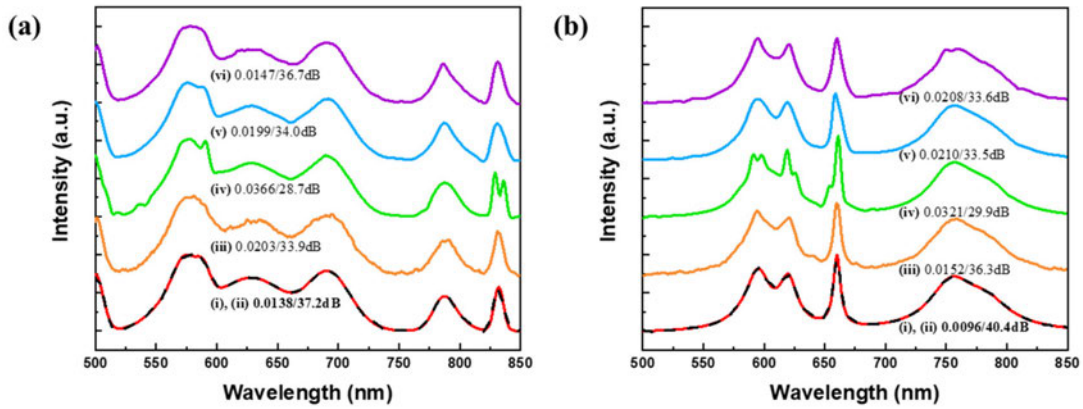


Figure 6. Spectral reconstructions of test spectra in synthetic datasets, (**a**) Gaussian dataset, (**b**) Lorentzian dataset. An input spectrum (solid red (i)) is compared with ResCNN (dashed black (ii)), CNN (orange (iii)), sparse recovery: Gaussian 1 (green (iv)), Gaussian 2 (blue (v)), and online dictionary learning (ODL) (purple (vi)). The baselines are shifted for clarity.

Only sparse recovery with Gaussian 1 fails to recover the fine details of the input spectrum. One example of the poor ability of sparse recovery with Gaussian 1 to resolve the signal is the recovery of the peak at ~830 and 590 nm being recovered as two neighboring peaks in Figure 6a,b, respectively. CNN was unable to capture the smoothness of the spectral features compared to the other methods.

5.2. Noisy Synthetic Datasets

To verify the stability of ResCNN, we evaluated the accuracy of the reconstruction at various noise levels. Gaussian white noise was added to the measurement vector $\mathbf{n} \in \mathbb{R}^{M\times 1}$ to Equation (2), i.e., $\mathbf{y = Tx+n}$. We considered six different noise levels whose signal-to-noise ratios (SNRs) are 15, 20, 25, 30, 35, and 40 dB. The SNR (dB) is defined as $10 \cdot \log_{10}\left(\|\mathbf{x}\|_2^2 / N\sigma^2\right)$, where σ is the standard deviation of the noise. Using Gaussian and Lorentzian datasets, we compared the reconstruction performance of ResCNN with the sparse recovery using Gaussian 2, which shows the best reconstruction performances among sparse recovery methods in synthetic datasets. ResCNN was evaluated with the same hyper-parameters that were used for the noise-free datasets. The average RMSE and PSNR for each of the six noise levels are shown in Table 3. While ResCNN was trained using noise-free data, it outperformed the sparse recovery with Gaussian 2 at every noise level, which indicates that ResCNN remains stable even with noisy datasets.

Table 3. Average RMSE and PSNR under various signal-to-noise ratios (SNRs, dB) with synthetic datasets.

		SNR (dB)					
Dataset	Method	15 dB	20 dB	25 dB	30 dB	35 dB	40 dB
Gaussian Dataset	Sparse recovery + Gaussian 2	0.0796 (22.7 dB)	0.0482 (27.1 dB)	0.0308 (31.2 dB)	0.0215 (34.8 dB)	0.0166 (37.9 dB)	0.0138 (40.7 dB)
	ResCNN	**0.0671 (24.2 dB)**	**0.0401 (28.7 dB)**	**0.0251 (32.9 dB)**	**0.0171 (36.6 dB)**	**0.0130 (39.8 dB)**	**0.0110 (42.4 dB)**
Lorentzian Dataset	Sparse recovery + Gaussian 2	0.0817 (22.6 dB)	0.0483 (27.1 dB)	0.0300 (31.2 dB)	0.0201 (35.0 dB)	0.0147 (38.5 dB)	0.0119 (41.4 dB)
	ResCNN	**0.0689 (24.1 dB)**	**0.0404 (28.7 dB)**	**0.0243 (33.1 dB)**	**0.0157 (37.1 dB)**	**0.0113 (40.6 dB)**	**0.0091 (43.4 dB)**

5.3. Measured Datasets

ResCNN was trained using the two measured datasets listed in Table 1, USGS and Munsell colors, and its reconstruction performance was evaluated. In addition, the signal reconstruction was performed using CNN and sparse recovery with five different sparsifying bases. Table 4 reports the average RMSE and PSNR for each of the seven methods. ResCNN achieves the smallest average RMSE and the largest average PSNR for both datasets. In the USGS dataset, the average RMSE and PSNR of ResCNN are 0.0048 and 52.4 dB, respectively. In addition, ResCNN achieves 0.0040 for the average RMSE and 50.0 dB for the average PSNR in the Munsell colors dataset. Similar to synthetic datasets, all of the learned dictionaries provided similar reconstruction performances. In addition, the small differences between Gaussian 1 and 2 show large differences in the RMSE and PSNR. The average RMSE and PSNR of the learned dictionary methods approach the values of ResCNN for Munsell colors dataset because the Munsell colors dataset has simpler spectral features than the other datasets.

Table 4. Average RMSEs and PSNRs for the measured datasets.

	Sparse Recovery					**Deep Learning**	
Dataset	Gaussian 1	Gaussian 2	K-SVD	MOD	ODL	CNN	ResCNN
USGS [32]	0.0081 (45.3 dB)	0.0061 (48.4 dB)	0.0070 (48.5 dB)	0.0081 (47.4 dB)	0.0074 (47.6 dB)	0.0116 (40.8 dB)	**0.0048 (52.4 dB)**
Munsell colors [33]	0.0068 (44.6 dB)	0.0050 (47.5 dB)	**0.0040** (49.8 dB)	**0.0040** (49.9 dB)	0.0042 (49.5 dB)	0.0076 (43.0 dB)	**0.0040 (50.0 dB)**

Figure 7 shows the reconstruction results of one test spectra from each of the measured datasets. The spectrum for the organic compound dibenzothiophene in the USGS dataset is reconstructed in Figure 7a. The spectrum of Munsell color 5 PB 2/2 is shown in Figure 7b. The solid red lines are the input spectra (i). ResCNN are shown in dashed black lines (ii), and CNN are shown in solid black lines (iii). The spectra of (iv) to (vi) are reconstructed spectra using the sparse recovery with Gaussian 1, Gaussian 2, and K-SVD. Because of the best performance of the K-SVD among the learned dictionaries only the K-SVD method is shown.

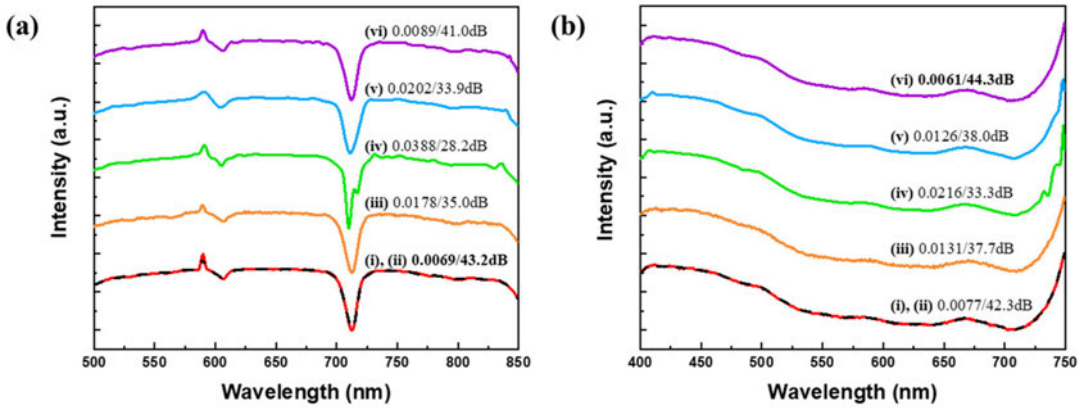


Figure 7. Spectral reconstructions of test spectra in measured datasets: (**a**) spectrum of organic compound dibenzothiophene in USGS dataset, (**b**) spectrum of Munsell color 5PB 2/2. The input spectrum (solid red line (i)) is compared with ResCNN (dashed black (ii)), CNN (orange (iii)), sparse recovery: Gaussian 1 (green (iv)), Gaussian 2 (blue (v)), and K-SVD (purple (vi)). The baselines are shifted for clarity.

The RMSE and PSNR for ResCNN are 0.0069 (43.2 dB) for the spectrum from the USGS dataset and 0.0077 (42.3 dB) for the spectrum from the Munsell colors dataset. ResCNN outperforms other approaches for the spectrum from USGS dataset. However, for the spectrum from Munsell colors

dataset, the sparse recovery with K-SVD outperforms ResCNN. ResCNN achieves slightly larger RMSE and smaller PSNR.

The performances of sparse recovery with Gaussian 2 is degraded for measured datasets compared with the performance for synthetic datasets. The measured datasets have rough spectral features unlike the smooth spectral features observed in the synthetic datasets. As a result, the sparse recovery with Gaussian 2 performs worse, because of its inability to represent rough spectral features using Gaussian distribution functions. The performance of sparse recovery with dictionary learning methods are improved for measured datasets compared with the performance of synthetic datasets. Because the number of spectra in measured datasets are smaller than the number of spectra in synthetic datasets. Therefore, finding the best-fit sparsifying basis for measured datasets is easier than finding the best-fit sparsifying basis for synthetic datasets using dictionary-learning methods. Meanwhile, ResCNN shows superior reconstruction performances regardless of spectral features of datasets and the size of datasets.

6. Discussion

As shown in the results, we demonstrate empirically that ResCNN outperforms the sparse recovery methods and the CNN over all datasets. The sparse recovery shows unstable performance because it is highly dependent on the sparsifying basis and spectral features of dataset. This is a direct result of being unable to identify a fixed sparsifying basis that can transform any spectra into a sparse signal, which means the *a priori* structural information such as line shapes and FWHMs is required to select a consistent sparsifying basis. Learned dictionaries are used to cope with the problem of identifying a consistent sparsifying basis. The columns of learned dictionaries are composed of learned spectral features from the training dataset. While this shows an improvement in measured datasets, a learned dictionary is still limited to representing all the spectral features in the large dataset (i.e., synthetic datasets) using linear combinations of columns of the learned dictionary.

Compression approaches for summarizing information with a small number of sensors were proposed in [40]. These approaches can be exploited to generate a sparsifying basis by reducing the loss of spectral information in large datasets.

To improve the reconstruction performance in sparse recovery, pre-defined structure information and side information of unknown target signals were used in [41,42]. The reconstruction of three-dimensional electrical impedance tomography was improved by updating three-dimensional structural correlations using pre-defined structured signals [41]. To recover multi-modal data, a reconstruction framework is proposed in [42] that uses side information in unrolled optimization. Unrolled optimization approaches using deep learning were proposed in [43,44]. Deep-learning architectures were used to train hyper-parameters, such as a gradient regularizer and a step size. Using learned hyper-parameters, it was shown optimized solutions can be obtained within a fixed number of iterations. These proposed approaches for image reconstruction have assumed random sensing matrix and structured or sparse signals. In this work, however, we consider dense spectra and the sensing matrix from thin-film filters for the real implementation. Moreover, the reconstruction performance may change to a sparsifying basis as shown in results because a reconstructed spectrum $\hat{\mathbf{x}}$ should be represented as a linear combination of columns of a fixed sparse basis $\mathbf{\Phi}$ as $\mathbf{\Phi}\hat{\mathbf{s}}$.

For recovering spectra, ResCNN does not require the *a priori* knowledge of a sparsifying basis or prior information of spectral features. During training, ResCNN learns the spectral features using learnable layers, which enable it to recover the fine details for various kinds of spectra without identifying a sparsifying basis.

ResCNN is directly compared with CNN for the synthetic Gaussian dataset in Figure 8a where the mean squared error (Equation (4)) is plotted with respect to the epoch. The mean squared error for CNN and ResCNN are shown in solid black line and solid red line with square symbols, respectively. ResCNN shows a lower mean squared error than that of CNN. Moreover, ResCNN converges faster than CNN, indicating that ResCNN optimizes the learnable layers quicker, as expected based on

previous research using residuals [25,29]. In contrast to the previous research that numerous residual connections were used in very deep neural networks to converge networks faster by avoiding vanishing gradient problem, we achieve spectral reconstruction improvements even with one residual connection in a moderate depth CNN.

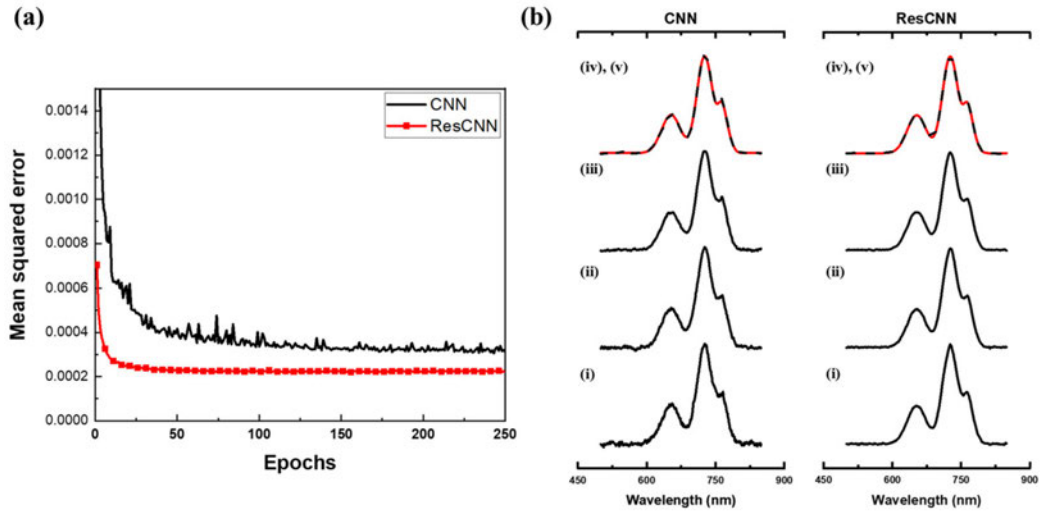


Figure 8. (**a**) Mean squared error of Gaussian dataset with respect to epochs. Solid black line denotes validation error of CNN, and solid red line with square symbols denotes validation error of ResCNN. (**b**) Reconstructions of a spectrum with respect to epochs where (i) to (iv) are epochs 1, 50, 150, and 250, respectively. Red line (v) denotes the original spectrum.

The reconstruction of an example spectrum with respect to the number of epochs is shown in Figure 8b. Black lines ((i) to (iv)) are the reconstructed spectra at 1, 50, 150, and 250 epochs, respectively. The solid red line (v) is the original spectrum, and the series of reconstructed spectrum for ResCNN show that the reconstruction converged earlier than CNN. The increased rate of convergence is because of the residual connection in ResCNN. Overall, the reconstruction performance of ResCNN is an improvement over CNN.

Note that both ResCNN and dictionary learning for sparse recovery require a training dataset and an optimization process to learn the spectral features. While this is a time-consuming process, remember that when using a learned dictionary to recover spectra, an iterative reconstruction algorithm is required, which needs additional time and incurs a high computational cost. The benefit of ResCNN is that it gives a reconstructed spectrum immediately once the training is completed.

7. Conclusions

In this paper, we propose a novel ResCNN for recovering the input spectrum from the compressed measurements in CS spectroscopy. As the optical structure for CS spectroscopy, we numerically generated multilayer thin-film filters which have a small mutual coherence. Therefore, we could compressively measure input spectra with unique sensing patterns. To reconstruct the input spectra from the compressively sampled measurements, we modeled ResCNN, which has a moderate-depth of learnable layers and a residual connection. We stacked nine learnable layers: five convolutional layers and four fully-connected layers with a single residual connection between the input and output of the learnable layers. The measurements were extended by a linear transformation and then fed into ResCNN. Finally, ResCNN reconstructed the input spectra. We demonstrated the empirical reconstruction results for ResCNN using synthetic and measured datasets. We compared the reconstruction performance of ResCNN with sparse recovery using five different sparsifying bases and CNN. Compared with sparse recovery methods, ResCNN shows better reconstruction performance without the *a priori* knowledge of either a sparsifying basis or any spectral features of the spectral datasets. On the other hand, the

sparse recovery methods show deviation of reconstruction performances to sparsifying bases and spectral datasets, meaning that a fixed sparsifying basis cannot represent all spectral features of input spectra. Furthermore, ResCNN shows stable reconstruction performances under noisy environments. Compared with CNN, ResCNN shows significant improvement in reconstruction performance and converges faster than CNN. In future work, we will explore compression approaches [40] and unrolled optimization approaches [43,44] for generating a sparsifying basis $\mathbf{\Phi}$ from the training dataset to fully represent spectra without loss of spectral features.

Author Contributions: Conceptualization, C.K. and H.-N.L.; methodology, C.K.; software, C.K. and D.P.; formal analysis, C.K. and D.P.; investigation, C.K.; data curation, C.K.; writing—original draft preparation, C.K.; writing—review and editing, C.K., D.P. and H.-N.L.; project administration, H.-N.L.; funding acquisition, H.-N.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (NRF-2018R1A2A1A19018665).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Clark, R.N.; Roush, T.L. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *J. Geophys. Res. Solid Earth* **1984**, *89*, 6329–6340. [CrossRef]
2. Izake, E.L. Forensic and homeland security applications of modern portable Raman spectroscopy. *Forensic Sci. Int.* **2010**, *202*, 1–8. [CrossRef] [PubMed]
3. Kim, S.; Cho, D.; Kim, J.; Kim, M.; Youn, S.; Jang, J.E.; Je, M.; Lee, D.H.; Lee, B.; Farkas, D.L.; et al. Smartphone-based multispectral imaging: System development and potential for mobile skin diagnosis. *Biomed. Opt. Express* **2016**, *7*, 5294–5307. [CrossRef] [PubMed]
4. Donoho, D.L. Compressed sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306. [CrossRef]
5. Eldar, Y.C.; Kutyniok, G. *Compressed Sensing: Theory and Applications*; Cambridge University Press: Cambridge, UK, 2012.
6. Candes, E.J.; Eldar, Y.C.; Needell, D.; Randall, P. Compressed sensing with coherent and redundant dictionaries. *Appl. Comput. Harm. Anal.* **2011**, *31*, 59–73. [CrossRef]
7. Oiknine, Y.; August, I.; Blumberg, D.G.; Stern, A. Compressive sensing resonator spectroscopy. *Opt. Lett.* **2017**, *42*, 25–28. [CrossRef]
8. Kurokawa, U.; Choi, B.I.; Chang, C.-C. Filter-based miniature spectrometers: Spectrum reconstruction using adaptive regularization. *IEEE Sens. J.* **2011**, *11*, 1556–1563. [CrossRef]
9. Cerjan, B.; Halas, N.J. Toward a Nanophotonic Nose: A Compressive Sensing-Enhanced, Optoelectronic Mid-Infrared Spectrometer. *ACS Photonics* **2018**, *6*, 79–86. [CrossRef]
10. Oiknine, Y.; August, I.; Stern, A. Multi-aperture snapshot compressive hyperspectral camera. *Opt. Lett.* **2018**, *43*, 5042–5045. [CrossRef]
11. Kim, C.; Lee, W.-B.; Lee, S.K.; Lee, Y.T.; Lee, H.-N. Fabrication of 2D thin-film filter-array for compressive sensing spectroscopy. *Opt. Lasers Eng.* **2019**, *115*, 53–58. [CrossRef]
12. Oliver, J.; Lee, W.-B.; Lee, H.-N. Filters with random transmittance for improving resolution in filter-array-based spectrometers. *Opt. Express* **2013**, *21*, 3969–3989. [CrossRef] [PubMed]
13. August, Y.; Stern, A. Compressive sensing spectrometry based on liquid crystal devices. *Opt. Lett.* **2013**, *38*, 4996–4999. [CrossRef] [PubMed]
14. Huang, E.; Ma, Q.; Liu, Z. Etalon Array Reconstructive Spectrometry. *Sci. Rep.* **2017**, *7*, 40693. [CrossRef] [PubMed]
15. Wang, Z.; Yu, Z. Spectral analysis based on compressive sensing in nanophotonic structures. *Opt. Express* **2014**, *22*, 25608–25614. [CrossRef] [PubMed]
16. Wang, Z.; Yi, S.; Chen, A.; Zhou, M.; Luk, T.S.; James, A.; Nogan, J.; Ross, W.; Joe, G.; Shahsafi, A. Single-shot on-chip spectral sensors based on photonic crystal slabs. *Nat. Commun.* **2019**, *10*, 1020. [CrossRef]
17. Pati, Y.C.; Rezaiifar, R.; Krishnaprasad, P.S. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In Proceedings of the 27th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 1–3 November 1993; pp. 40–44.

18. Dai, W.; Milenkovic, O. Subspace pursuit for compressive sensing signal reconstruction. *IEEE Trans. Inf. Theory* **2009**, *55*, 2230–2249. [CrossRef]
19. Chen, S.S.; Donoho, D.L.; Saunders, M.A. Atomic decomposition by basis pursuit. *SIAM Rev.* **2001**, *43*, 129–159. [CrossRef]
20. Candes, E.; Tao, T. Decoding by linear programming. *arXiv* **2005**, arXiv:math/0502327. [CrossRef]
21. Oliver, J.; Lee, W.; Park, S.; Lee, H.-N. Improving resolution of miniature spectrometers by exploiting sparse nature of signals. *Opt. Express* **2012**, *20*, 2613–2625. [CrossRef]
22. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436. [CrossRef]
23. Mousavi, A.; Baraniuk, R.G. Learning to invert: Signal recovery via deep convolutional networks. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 2272–2276.
24. Li, Y.; Xue, Y.; Tian, L. Deep speckle correlation: A deep learning approach toward scalable imaging through scattering media. *Optica* **2018**, *5*, 1181–1190. [CrossRef]
25. Lee, D.; Yoo, J.; Ye, J.C. Deep residual learning for compressed sensing MRI. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, VIC, Australia, 18–21 April 2017; pp. 15–18.
26. Mardani, M.; Gong, E.; Cheng, J.Y.; Vasanawala, S.S.; Zaharchuk, G.; Xing, L.; Pauly, J.M. Deep Generative Adversarial Neural Networks for Compressive Sensing MRI. *IEEE Trans. Med. Imaging* **2019**, *38*, 167–179. [CrossRef] [PubMed]
27. Jin, K.H.; McCann, M.T.; Froustey, E.; Unser, M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.* **2017**, *26*, 4509–4522. [CrossRef] [PubMed]
28. Kim, C.; Park, D.; Lee, H.-N. Convolutional neural networks for the reconstruction of spectra in compressive sensing spectrometers. In *Optical Data Science II*; International Society for Optics and Photonics: Bellingham, WA, USA, 2019; Volume 10937, p. 109370L.
29. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
30. Macleod, H.A. *Thin-Film Optical Filters*; CRC Press: Boca Raton, FL, USA, 2010.
31. Barry, J.R.; Kahn, J.M. Link design for nondirected wireless infrared communications. *Appl. Opt.* **1995**, *34*, 3764–3776. [CrossRef] [PubMed]
32. Kokaly, R.F.; Clark, R.N.; Swayze, G.A.; Livo, K.E.; Hoefen, T.M.; Pearson, N.C.; Wise, R.A.; Benzel, W.M.; Lowers, H.A.; Driscoll, R.L. *USGS Spectral Library Version 7 Data: US Geological Survey Data Release*; United States Geological Survey (USGS): Reston, VA, USA, 2017.
33. University of Eastern Finland. Spectral Color Research Group. Available online: http://www.uef.fi/web/spectral/-spectral-database (accessed on 2 August 2019).
34. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
35. Chen, G.; Needell, D. Compressed sensing and dictionary learning. *Finite Fram. Theory* **2016**, *73*, 201.
36. Engan, K.; Aase, S.O.; Husoy, J.H. Method of optimal directions for frame design. In Proceedings of the 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258), Phoenix, AZ, USA, 15–19 March 1999; Volume 5, pp. 2443–2446.
37. Aharon, M.; Elad, M.; Bruckstein, A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal. Process.* **2006**, *54*, 4311–4322. [CrossRef]
38. Mairal, J.; Bach, F.; Ponce, J.; Sapiro, G. Online dictionary learning for sparse coding. In Proceedings of the 26th Annual International Conference on Machine Learning, Montreal, QC, Canada, 14–17 June 2009; pp. 689–696.
39. Koh, K.; Kim, S.-J.; Boyd, S. An interior-point method for large-scale l1-regularized logistic regression. *J. Mach. Learn. Res.* **2007**, *8*, 1519–1555.
40. Martino, L.; Elvira, V. Compressed Monte Carlo for distributed Bayesian inference. *arXiv* **2018**, arXiv:1811.0505.
41. Liu, S.; Wu, H.; Huang, Y.; Yang, Y.; Jia, J. Accelerated Structure-Aware Sparse Bayesian Learning for 3D Electrical Impedance Tomography. *IEEE Trans. Ind. Inform.* **2019**. [CrossRef]
42. Tsiligianni, E.; Deligiannis, N. Deep coupled-representation learning for sparse linear inverse problems with side information. *IEEE Signal. Process. Lett.* **2019**, *26*, 1768–1772. [CrossRef]

43. Diamond, S.; Sitzmann, V.; Heide, F.; Wetzstein, G. Unrolled optimization with deep priors. *arXiv* **2017**, arXiv:1705.08041.
44. Gilton, D.; Ongie, G.; Willett, R. Neumann Networks for Linear Inverse Problems in Imaging. *IEEE Trans. Comput. Imaging* **2019**. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Optics and Lasers in Engineering 115 (2019) 53–58

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

journal homepage: www.elsevier.com/locate/optlaseng

Fabrication of 2D thin-film filter-array for compressive sensing spectroscopy

Cheolsun Kim, Woong-Bi Lee, Soo Kyung Lee, Yong Tak Lee, Heung-No Lee*

School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea

ARTICLE INFO

Keywords:
Spectroscopy
Thin films
Inverse problems
Compressive sensing

ABSTRACT

We demonstrate 2D filter-array compressive sensing spectroscopy based on thin-film technology and a compressive sensing reconstruction algorithm. To obtain different spectral modulations, we fabricate a set of multilayer filters using alternating low- and high-index materials and reconstruct the input spectrum using a small number of measurements. Experimental results show that the fabricated filter-array provides compatible spectral resolution performance with a conventional spectrometer in monochromatic lights and LEDs. In addition, the fabricated filter-array covers a wide range of wavelengths with a single exposure.

1. Introduction

The demand for spectrum information is increasing not only in research and development but also in the private sector. In response to this demand, researchers are trying to make spectrometers that are both small and inexpensive. These spectrometers could be used in various fields, such as medical systems, mobile applications, and remote sensing [1–3]. In particular, optical filter-based spectrometers do not need motorized or dispersive elements, and their filter-array can be directly attached to the detectors so that they can be easily miniaturized. However, there is a trade-off between size (for integrating filters) and spectral resolution with miniaturized spectrometers.

Over the years, numerous approaches to applying compressive sensing (CS) techniques have been proposed to reduce the size of spectroscopes without reducing spectral resolution, or potentially even improving it. These approaches [4–7] include the following: band pass filters [4], random transmittance filters [5], photonic crystal slabs [6], and liquid crystal phase retarders [7]. Recently, Fabry–Perot (FP)-based CS spectroscopy methods have been presented [8,9]. To acquire differently modulated spectral measurements, a 2D array of FP resonators with different cavity depths has been tried [8] as well as a piezo-actuated device that changes the distance between two FP mirrors has been tried [9]. A hundred FP resonators are used to recover the input spectrum in [8], and the operational range of the piezo-actuator imposes mechanical limitations in [9].

The CS framework [10–12] is an efficient sampling and reconstruction scheme that requires fewer samples to reconstruct the signal than that required by conventional sampling. The CS framework can be applied to filter-based spectroscopy, offering the advantage of reducing the number of filters and detectors required and allowing the system to be miniaturized.

In spectroscopy, the relation between the spectral components of the input light source $\mathbf{x} \in \mathbb{R}^{N\times 1}$ and the modulated signal $\mathbf{y} \in \mathbb{R}^{M\times 1}$ can be expressed as follows:

$$\mathbf{y} = \mathbf{T}\mathbf{x}, \tag{1}$$

where $\mathbf{T} \in \mathbb{R}^{M\times N}$ is the sensing matrix. Each row of the sensing matrix is to represent the transmission function (TF) of i-th filter, $\boldsymbol{T}_m \in \mathbb{R}^{1\times N}$ for $m = 1, 2, \ldots, M$. In order to achieve miniaturization of spectroscope without degradation of spectral resolution, the CS framework is utilized in spectroscopy, where the number of filters is set to be smaller than the number of spectral components ($M < N$). Then, Eq. (1) becomes an underdetermined linear system. A sparse signal reconstruction algorithm with $L1$ norm minimization can be used to solve Eq. (1), if the input spectrum is either naturally sparse or can be sparsely represented in some basis $\mathbf{\Phi} \in \mathbb{R}^{N\times N}$, i.e., $\mathbf{x} = \mathbf{\Phi}\mathbf{s}$, where $\mathbf{s} \in \mathbb{R}^{N\times 1}$ is a sparse vector. Then, Eq. (1) becomes

$$\mathbf{y} = \mathbf{T}\mathbf{\Phi}\mathbf{s} \tag{2}$$

The sparse signal $\mathbf{s}$ can be estimated by solving the following $L1$ norm minimization problem:

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \|\mathbf{s}\|_1 \text{ subject to} \|\mathbf{y} - \mathbf{T}\mathbf{\Phi}\mathbf{s}\|_2 \le \varepsilon \tag{3}$$

where ε is a small non-negative constant. The reconstructed input spectrum $\hat{\mathbf{x}}$ is then $\mathbf{\Phi}\hat{\mathbf{s}}$.

In this paper, we demonstrate 2D filter-array CS spectroscopy. This uses a multilayer thin-film filter-array for spectral modulation, where each filter modulates the input spectrum using different sensing patterns. A CMOS image camera reads out the modulated signals with a

* Corresponding author.
E-mail address: heungno@gist.ac.kr (H.-N. Lee).

https://doi.org/10.1016/j.optlaseng.2018.10.018
Received 22 February 2018; Received in revised form 2 October 2018; Accepted 31 October 2018
Available online 24 November 2018
0143-8166/© 2018 Elsevier Ltd. All rights reserved.

Table 1
Recursion for calculating reflection coefficients.

Input: $\lambda, \theta_1 = 0$, $\boldsymbol{n} = \{n_1, n_2, \cdots, n_{l-1}, n_l\}$, $\boldsymbol{d} = \{d_2, d_3, \cdots, d_{l-1}, d_l\}$.

Step 1: Obtain θ_k, β_k, and N_k
$\theta_k = \sin^{-1}(\frac{n_{k-1}}{n_k} \sin\theta_{k-1})$, $for\ k = 2,\ 3,\ \cdots,\ l$.
$\beta_k = 2\pi \cos(\theta_k) n_k d_k / \lambda$, $for\ k = 2,\ 3,\ \cdots,\ l$.
$N_k = \begin{cases} n_k/\cos\theta_k & for\ TE \\ n_k \cos\theta_k & for\ TM \end{cases}$, $for\ k = 1,\ 2,\ \cdots,\ l$.
Step 2: Set $\eta_l = N_l$
Step 3: Obtain η_2
Decrement k by 1 from $l-1$ to 2
$\eta_k = N_k \frac{\eta_{k+1}\cos\beta_k + jN_k \sin\beta_k}{N_k \cos\beta_k + j\eta_{k+1}\sin\beta_k}$
return η_2
Step 4: Compute $\rho = (N_1 - \eta_2)/(N_1 + \eta_2)$.

Output: ρ

single exposure, and then a reconstruction algorithm is applied that depends on the modulated signals and the sensing matrix, allowing the input spectrum to be recovered.

The research focus has been given to fabrication of the multilayer thin-film filters for actual CS spectroscopy implementation and verification experiments. For fabricating as a 2D filter-array, we use commonly available materials SiNx and SiO2 for high and low refractive index materials which are deposited alternately on the substrate with varying thicknesses. Furthermore, we come up with a practical way that set of filters can be deposited on a single substrate with different thicknesses of layers.

2. 2D filter-array

2.1. Multilayer thin-film filter

Thin films are a basic component that have been applied in a variety of areas, including semiconductor devices, optical coatings, and solar cells [13]. The theoretical TF of a multilayer thin-film filter is given by [14]

$$T(\lambda, \theta_1) = 1 - \tfrac{1}{2}\left(|\rho_{TE}|^2 + |\rho_{TM}|^2\right), \tag{4}$$

where ρ_{TE} and ρ_{TM} are the reflection coefficients. Given a wavelength λ and the incident angle θ_1, TF can be calculated using recursive routines shown in Table 1.

In Table 1, given the input of a wavelength λ, a vector of l refractive indices $\boldsymbol{n} = (n_1, n_2, \cdots, n_{l-1}, n_l)$ and a vector of $l-1$ layer thicknesses $\boldsymbol{d} = (d_2, d_3, \cdots, d_{l-1}, d_l)$, a reflection coefficient ρ is generated. Note that there are l layers considered in total. The first one is the layer of the air and the last one is the layer of the substrate. The light is assumed to be arriving from the air to the second layer in normal incidence. The first index n_1 in the vector $\boldsymbol{n}$ represents the refractive index of the air. The last one n_l in the vector $\boldsymbol{n}$ represents the refractive index of the substrate. The refractive indices of the intermediate thin-film layers are denoted by n_2 to n_{l-1}. The thickness of the air does not need to be considered. The thickness of the substrate is denoted by d_l.

The thicknesses of the intermediate thin-film layers are denoted by d_2 to d_{l-1}. The incidence angle of the light passing from the kth to the $k+1$th layer is θ_k, and η_k is the effective complex-valued index of the kth layer. A TF for a single filter is obtained by considering all wavelengths in the range of interest. An array of TFs for the M filters can be obtained by repeating this process where each filter $\boldsymbol{T}_m \in \mathbb{R}^{1\times N}$ for $m = 1,\ 2, \cdots, M$ in Eq. (1) is generated from a unique set of refractive index and thickness vectors.

2.2. Numerical design of 2D filter-array

To implement the proposed 2D filter-array, we numerically modeled the proposed spectroscopy method with reference to [14–16], and according to the following steps. (i) Generate the reference vector of layer thicknesses, i.e. $\boldsymbol{d} = (d_2, d_3, \cdots, d_{l-1}, d_l)$, for the reference filter. (ii) Generate a vector of thicknesses for the other filter by randomly removing one to five layer thicknesses from the reference vector. (iii) Repeat the step (ii) 35 times to create a total of 36 vectors of thicknesses. (iv) Use the recursion Table 1 and Eq. (4) to calculate the TFs for a new filter-array (sensing matrix). (v) Use the mutual coherence μ to quantify the goodness of the sensing matrix of the designed filter-array. Mutual coherence μ is defined as $\mu \triangleq \max_{i,j} |o_{ij}|$, where o_{ij} is the (i, j)th off-diagonal element of the Gram matrix, $\mathbf{T}^*\mathbf{T} \in \mathbb{R}^{N\times N}$. $\mathbf{T}^*$ denotes the conjugate transpose of $\mathbf{T}$. With these steps, we can generate a single set of 36 filters. By repeating these steps, multiple sets of 36 filters can be obtained. Among these sets of filter-arrays, the set of filters with a smallest mutual coherence is selected.

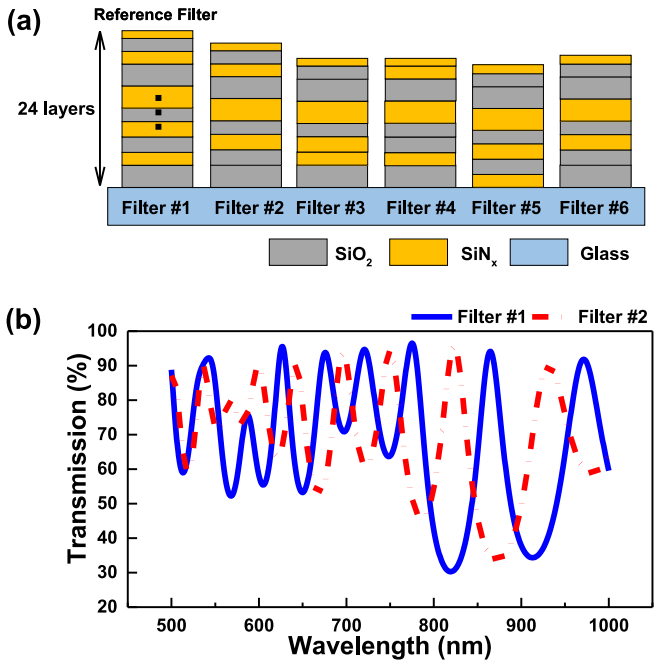


Fig. 1. (a) Schematic of the thin-film filter-array. (b) Example of two transmission functions for thin-film filters.

In CS framework, a sensing matrix with a smaller mutual coherence is better than the one with a higher mutual coherence to capture the information of input signal to be reconstructed [5,17]. A schematic of the proposed filter-array is shown in Fig. 1(a). Each time a layer is removed, the layers above and those below come together to form a single layer with two thicknesses added up. We consider two materials, SiN_x and SiO_2 for the high- and low-refractive index materials with refractive indices of 2.02 for SiN_x and 1.45 for SiO_2. The thickness range of each layer is from 50 to 150 nm. Through the numerical design, we empirically found that removal of up to five layers from the 24-layer reference filter was possible to create a 6×6 filter-array with a low coherence.

Fig. 1(b) shows the TFs for two designed filters as examples. In conventional spectroscopy, the TFs with a large spectral depth and a narrow spectral peak are preferred in order to prevent interference among measurements. In compressive sensing spectroscopy, Each TF of the filter should be wide enough so that the set of the small number of filters fully senses the spectral information in the given wavelength range [9].

Each filter shows several spectral peaks and rapid changes of transmission value with respect to wavelength. Therefore, each filter has a high optical throughput that the energy (intensity) which passes through the filter is higher than that with the conventional bandpass filter approach. In addition, fewer filters can be used to cover the entire wavelength range with the proposed method. For example, suppose 250 bandpass filters are used to cover the wavelength range from 500 to

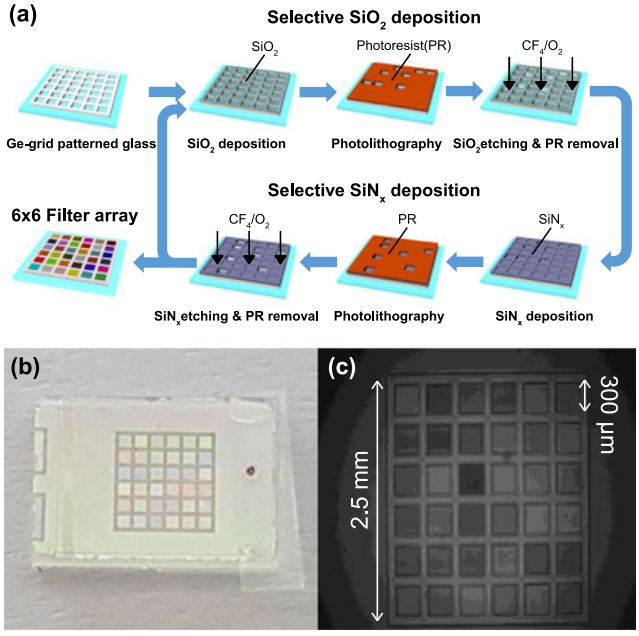


Fig. 2. (a) Schematic of the thin-film filter-array fabrication process. (b) Photograph of a fabricated thin-film filter-array. (c) Monochrome image of the thin-film filter-array taken at a wavelength of 700 nm.

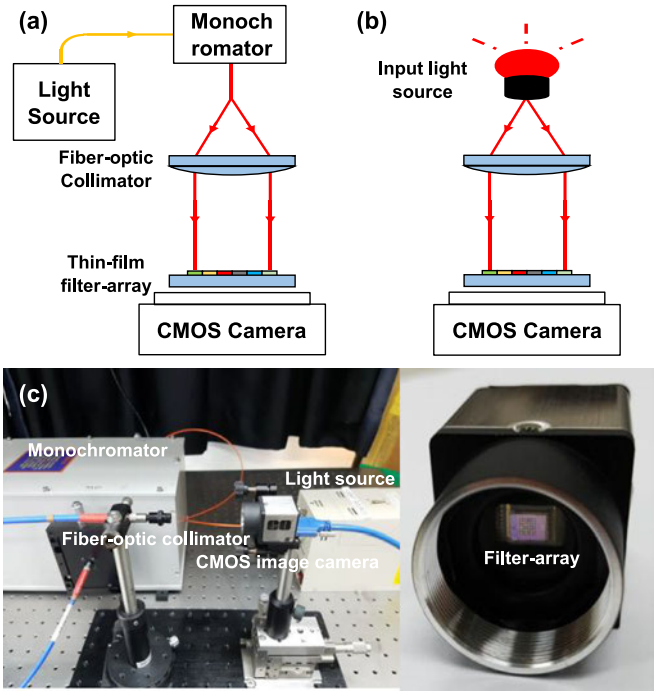


Fig. 3. (a) Schematic of the optical setup for measuring the sensing matrix. (b) Schematic of the optical setup for testing the performance of the proposed spectroscopy system. (c) Photographs of the optical setup and the CMOS image camera with the thin-film filter-array.

1000 nm. Then, the bandwidth of TF is 2 nm, according to the conventional bandpass filter design. In the proposed approach, the same range of wavelength can be covered with only 36 proposed filters, subject to the use of a recovery algorithm present at the reconstruction end.

2.3. Filter-array fabrication

Fig. 2(a) shows the process in which a thin-film filter-array is fabricated. This comprises two main parts; one is SiO_2 film deposition and the other is SiN_x film deposition according to the specified thicknesses. Prior to depositing an SiO_2 film, a 6×6 germanium (Ge) grid with elements of size 300 μm and spacing 100 μm was formed on the glass using an e-beam evaporator to separate the filters. In this grid, SiO_2 and SiN_x layers were deposited with the width of 300 μm in each filter. Then, selective deposition was done as follow: An intentionally thick SiO_2 film was deposited on the glass patterned with the Ge grid using plasma- enhanced chemical vapor deposition. The regions where the film should not be deposited were then removed by conventional photolithography, namely CF_4/O_2 reactive ion etching. The process pressure and radio- frequency power were maintained at 50 mTorr and 50 W, respectively. The SiN_x film deposition process was performed in the same manner as for SiO_2. Finally, these two main steps, SiO_2 and SiN_x film deposition, were repeated 12 times each to lay down 24 layers. Fig. 2(b) and (c) show a photograph of a fabricated thin-film filter-array and a monochrome image of the filter-array, respectively. Each filter is composed of a different number of layers each with different thicknesses; therefore, each one has unique color due to its different TF, as shown in Fig. 1(b).

3. Experiments

3.1. Experimental setup

Optical setups for experimental verification of the proposed spectroscopy system are shown in Fig. 3. Fig. 3(a) depicts the optical setup for measuring TFs of a filter-array. The setup for testing the performance of the proposed system is shown in Fig. 3(b). The photographs of the optical setup and the CMOS image camera with the thin-film filter-array are shown in Fig. 3(c). During the optical experiments, we set the incident angle to filter-array as normal incidence. Using a linear stage, a rotational stage and optical mounting posts, we aligned the optical fiber with the CMOS image camera (E0-1312, Edmund Optics) for the normal incidence.

In Fig. 3(a), a halogen lamp (KLS-150H-LS-150D, Kwangwoo) was used to provide a continuous light spectrum. It was put into a monochromator (MMAC-200, Mi Optics) to produce a specific narrow wavelength band. Then, a fiber-optic collimator was used to form a beam of parallel light. The beam was fed into the CMOS image camera through the fabricated thin-film filter-array. With a single exposure, each filter modulated the light in a different pattern. The modulated light was read out by pixels of CMOS image camera, yielding $M = 36$ distinct output signals $\mathbf{y}$ in Eq. (1). Each output signal was taken by summing up the modulated values of the pixels underneath the pertinent filter.

To apply CS reconstruction algorithms to the proposed system, the sensing matrix $\mathbf{T}$ must be pre-determined. Let us denote the intensity which passes through the filter-array as $IF(m, \lambda)$ and the intensity without the filter-array as $IWF(m, \lambda)$, where m is the filter index and λ is the wavelength. The sensing matrix is then given by

$$T(m,\lambda) = \frac{IF(m,\lambda) - BI(m,\lambda)}{IWF(m,\lambda) - BI(m,\lambda)}, \tag{5}$$

where $BI(m, \lambda)$ is the background intensity. We took 500 wavelength samples, spaced 1 nm apart, in the range from 500 to 1000 nm. The measured sensing matrix $\mathbf{T} \in \mathbb{R}^{36 \times 500}$ obtained from the fabricated thin-film filter-array is shown in Fig. 4. Each TF of the filters, a row of the color map, is shown as a combination of colors, i.e., red (high transmission value) and blue (low transmission value). Different TFs show different places of high and low transmission values indicating mutual uncorrelation. As a set of 36 filters, the filter-array covers the entire wavelength range with high optical throughput.

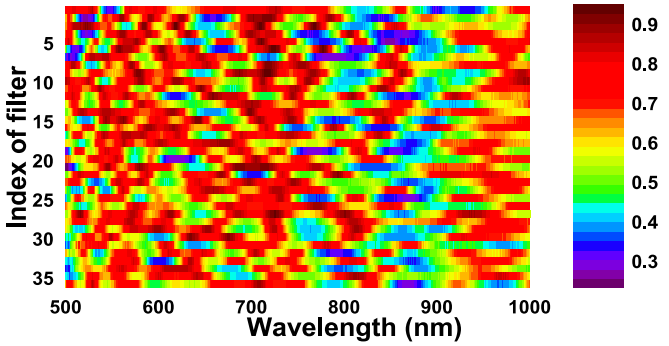


Fig. 4. Color map of the measured sensing matrix for the thin-film filter-array. Each row represents the TF of a filter with respect to wavelength.

3.2. Computational experiments

To quantify the performance and explore the two-point resolution of the fabricated filter-array, we conducted computational experiments. The two-point resolution is the ability to distinguish the spectral peaks which are closely spaced. For the experiments, we generated mono-peak spectra and two-peak spectra as input spectra using the Gaussian function. A generated input spectrum $\mathbf{x}$ was numerically modulated by multiplying the measured sensing matrix $\mathbf{T}$ as shown in Eq. (1). Then, using the M-modulated signals (measurements) and the sensing matrix $\mathbf{T}$, a reconstruction algorithm is used to recover the input spectrum. In the experiments, we considered that the input spectrum was a directly sparse signal. The mean-squared error (MSE) between the input spectrum $\mathbf{x}$ and the reconstructed spectrum $\hat{\mathbf{x}}$ was calculated. The MSE is defined as $\|\mathbf{x}-\hat{\mathbf{x}}\|_2^2/N$.

We firstly tested the spectral reconstruction performance of the fabricated filter-array with changing the full width at half maximum (FWHM) of the generated input signals. We made three noisy environments by adding the additive noise $\mathbf{n}$ to Eq. (1) as $\mathbf{y}=\mathbf{T}\mathbf{x}+\mathbf{n}$ whose the signal to noise ratios (SNRs) were 20, 25, 30 dB. The SNR in decibels is defined as $10\cdot\log_{10}(\|\mathbf{x}\|_2^2/N\sigma^2)$, where σ is the standard deviation of the noise.

The spectral reconstruction performances with respect to the FWHMs are shown in Fig. 5(a). For the two-peak spectrum, the distance between two peaks was determined as $[1.5\cdot\mathrm{FWHM}]$, where $[\cdot]$ is the nearest integer function. We averaged all the MSEs of the spectrum over the peak-locations from 500 to 999 nm in a given FWHM. As shown in Fig. 5(a), the mono-peak spectrum is reconstructed better than two-peak spectrum. As the FWHM increased, the performance of spectral reconstruction is degraded. This is due to the increased sparsity of the spectrum.

Second, we verified the stability of noise along the SNR conditions for the fabricated filter-array. As shown in Fig. 5(b), the reconstruction performance on mono-peak spectrum is better than that of the two-peak spectrum. In addition, when the FWHM is 1 nm, the reconstruction performance is better than the FWHM with 2 nm. Despite the additive noise, the results show that the fabricated filter-array is robust to the noisy environments.

As depicted in Fig. 5, the reconstruction performance of the fabricated filter-array depends on the FWHM and the SNR. For the two-point resolution, the MSE has the smallest value when the FWHM of the two-peak spectrum is 1 nm. The overall MSEs are small enough to use the fabricated filter-array to conduct the optical experiments.

3.3. Optical experiments

Optical experiments were then conducted to evaluate the performance of the proposed system, as shown in Fig. 3(b). Narrow-band monochromatic lights and LEDs were used as input light sources. To generate narrow-band light, a supercontinuum white light source (SuperK COMPACT, NKT Photonics) was placed in the monochromator, making a narrow band of light with a full width at half maximum (FWHM) of approximately 1 nm. These light sources were fed into the CMOS image camera through the filter-array, simultaneously capturing the M differently modulated signals. The M-modulated signals and the measured sensing matrix $\mathbf{T}$ were then used to solve Eq. (3). We used a Gaussian kernel matrix as the sparsifying basis $\mathbf{\Phi}$. The spectral waveform can be represented as a linear combination of Gaussian kernels, and a Gaussian kernel can be easily generated with two parameters, namely the peak location and the FWHM value [4,18]. The *l1_ls_noneg* algorithm [19] was used as a reconstruction algorithm to solve Eq. (3) with non-negativity constraints.

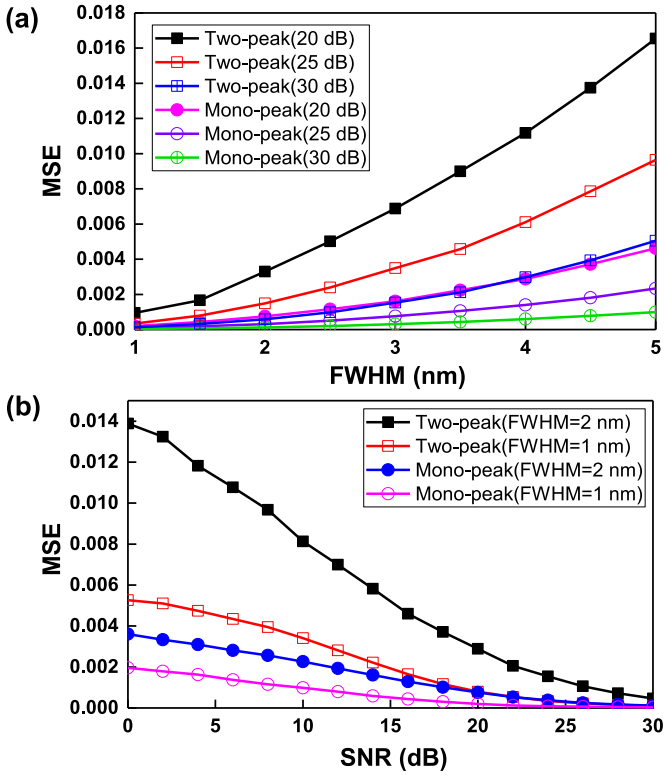


Fig. 5. (a) Computational reconstruction performance of the fabricated thin-film filter-array with respect to the FWHM. (b) Computational spectral reconstruction performance of the fabricated thin-film filter-array with respect to the SNR.

Fig. 6 shows the reconstruction results for monochromatic lights and LEDs. For comparison, the reference spectrum and the reconstructed spectrum were normalized to the range between zero and one.

The optical experimental results for monochromatic lights are shown in Fig. 6(a). In our optical experiment, depicted in Fig. 3(c), we use four different monochromatic spectra, with spectral peaks located at 600, 700, 800, and 900 nm, respectively. The reference spectra are measured using an optical spectrum analyzer (AQ-6315B, Ando) which indicate actual spectral peak locations at 598.7, 700.4, 800.5, and 900.4 nm, respectively. Using the fabricated filter-array CS spectroscopy with the reconstruction algorithm, the spectral peak locations are reconstructed at 599, 699, 799, and 901 nm, respectively. The mean FWHM of the reference spectra is approximately 1 nm, and the mean FWHM for the reconstructed spectra is approximately 1.4 nm.

Fig. 6(b) shows the spectral reconstructions of green (527 nm) and red (635 nm) LEDs. For the reference spectra, we measure the LEDs using a grating spectrometer (QE65000, Ocean Optics). The spectral peak locations for the reference LEDs are 527.6 nm (green LED) and 634.9 nm

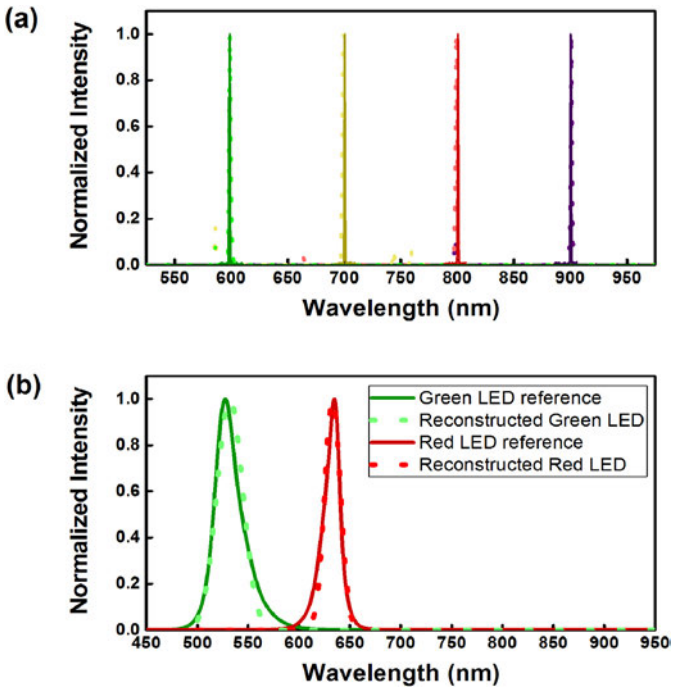


Fig. 6. Spectral reconstructions of several different input light sources. (a) Spectral reconstructions of monochromatic lights (dots) compared with reference spectra (solid lines): 600 nm (green), 700 nm (yellow), 800 nm (red), and 900 nm (purple). (b) Spectral reconstructions of LEDs (dots) compared with reference spectra (solid lines): green LED (527 nm), and red LED (635 nm).

(red LED), and the reconstructed spectral peak locations are 531 nm (green LED) and 633 nm (red LED). The peak signal-to-noise ratios are 28.3 dB (green LED) and 31.7 dB (red LED).

Discussing Fig. 6, the spectra of reconstructed monochromatic lights show several negligible spikes. This is probably due to background noise in the optical experiments. But overall, the reconstruction results of the proposed CS spectroscopy system for monochromatic lights and LEDs are similar to those of the grating spectrometer. Furthermore, the number of modulated signals is significantly small ($M = 36$) that the measurement to wavelength sample ratio is 36:500 (ratio between M and N).

To further explore the performance of the proposed CS spectroscopy, we conducted the computational experiment on the fabricated filter-array using a continuous light source, halogen lamp. For the experiment, we used the measured sensing matrix**T**. The conventionally measured spectrum of the halogen lamp was used as the input spectrum **x**. The modulated signals were generated by numerically multiplying the sensing matrix and the input spectrum. By solving Eq. (3), we reconstructed the continuous spectrum of light. In Fig. 7, we present computational spectral reconstruction of the halogen lamp. The peak signal-to-noise ratio is 43.8 dB. Due to the limitations of our optical components to reject the spectrum of the halogen lamp except for the wavelength range from 500 to 1000 nm, we could not perform the optical experiment on the continuous source. However, the computational reconstruction result of the halogen lamp indicates that the fabricated filter-array can be utilized for recovering the various kinds of spectra in the given wavelength range without limitations of the optical components.

Fabricating the proposed filter-array can be more difficult than fabricating Fabry–Perot structure due to the large number of layers for the proposed filter-array. However, the proposed spectroscope is compact and it does not need motorized components which were used with the Fabry–Perot structure [9]. In addition, thanks to the 2D array structure, the proposed spectroscope captures all measurements in a single exposure. But the Fabry–Perot spectroscope [9] required a number of

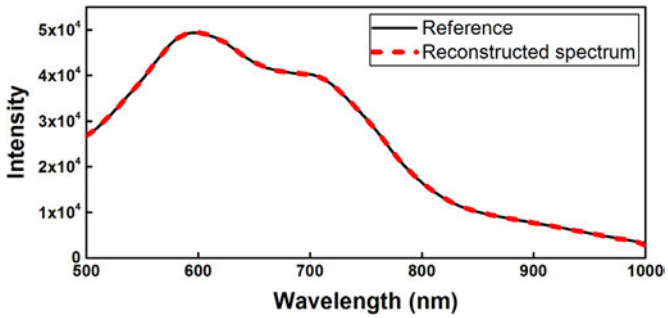


Fig. 7. Computational spectral reconstruction of a halogen lamp (red dash line) compared with the reference spectrum (black solid line) measured by a conventional spectrometer.

exposures as many times as the number of measurements. Compared to Fabry–Perot spectroscope [8], the proposed spectroscope utilizes 36 filters to cover the wavelength range from 500 to 1000 nm, but 100 filters were used in [8] to cover the range from 500 to 750 nm.

4. Conclusion

We have demonstrated a 2D array CS spectroscope based on thin-film technology. A 2D thin-film filter-array is fabricated based on array processing. Using the fabricated filter-array, measurements are obtained to which the CS reconstruction algorithm is applied. Finally, demonstration of input spectrum reconstruction is successfully made. The proposed system is compact, portable, and obtains the necessary measurements in a single exposure thanks to its structural advantages. Moreover, it works over a wide spectral range, from the visible light region to the near-infrared region. Compared with conventional spectrometers (non-CS spectrometers), the proposed system has a high optical throughput and compatible spectral resolution performance in monochromatic lights and LEDs with significantly less number of measurements.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (NRF-2018R1A2A1A19018665).

References

[1] Bacon CP, Mattley Y, DeFrece R. Miniature spectroscopic instrumentation: applications to biology and chemistry. Rev Sci Instrum 2004;75:1–16.
[2] Kim S, Cho D, Kim J, Kim M, Youn S, Jang JE, et al. Smartphone based multispectral imaging: system development and potential for mobile skin diagnosis. Biomed Opt Express 2016;7:5294–307.
[3] Clark RN, Roush TL. Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. J Geophys Res Solid Earth 1984;89:6329–40.
[4] Oliver J, Lee W, Park S, Lee H-N. Improving resolution of miniature spectrometers by exploiting sparse nature of signals. Opt Express 2012;20:2613–25.
[5] Oliver J, Lee W-B, Lee H-N. Filters with random transmittance for improving resolution in filter-array-based spectrometers. Opt Express 2013;21:3969–89.
[6] Wang Z, Yu Z. Spectral analysis based on compressive sensing in nanophotonic structures. Opt Express 2014;22:25608–14.
[7] August Y, Stern A. Compressive sensing spectrometry based on liquid crystal devices. Opt Lett 2013;38:4996–9.
[8] Huang E, Ma Q, Liu Z. Etalon array reconstructive spectrometry. Sci Rep 2017;7.
[9] Oiknine Y, August I, Blumberg DG, Stern A. Compressive sensing resonator spectroscopy. Opt Lett 2017;42:25–8.
[10] Donoho DL. Compressed sensing. IEEE Trans Inf Theory 2006;52:1289–306.
[11] Baraniuk RG. Compressive sensing [lecture notes]. IEEE Signal Process Mag 2007;24:118–21.
[12] Qaisar S, Bilal RM, Iqbal W, Naureen M, Lee S. Compressive sensing: from theory to applications, a survey. J Commun Netw 2013;15:443–56.

[13] Macleod HA. Thin-film optical filters. CRC press; 2001.
[14] Barry JR, Kahn JM. Link design for nondirected wireless infrared communications. Appl Opt 1995;34:3764–76.
[15] Pedrotti FL, Pedrotti LS. Introduction to optics. 2nd ed. Prentice Hall; 1993.
[16] Topasna DM, Topasna GA. Numerical modeling of thin film optical filters. Proc. SPIE 2009;9666:96661P.
[17] Candes EJ, Eldar YC, Needell D, Randall P. Compressed sensing with coherent and redundant dictionaries. Appl Comput Harmon Anal 2011;31:59–73.
[18] Kurokawa U, Choi BI, Chang C-C. Filter-based miniature spectrometers: spectrum reconstruction using adaptive regularization. IEEE Sens J 2011;11:1556–63.
[19] Koh K, Kim S-J, Boyd S. An interior-point method for large-scale l1-regularized logistic regression. J Mach Learn Res 2007;8:1519–55.

Filters with random transmittance for improving resolution in filter-array-based spectrometers

J. Oliver, Woong-Bi Lee, and Heung-No Lee*

School of Information and Communications, Gwangju Institute of Science and Technology, South Korea
*heungno@gist.ac.kr

Abstract: In this paper, we introduce a method for improving the resolution of miniature spectrometers. Our method is based on using filters with random transmittance. Such filters sense fine details of an input signal spectrum, which, when combined with a signal processing algorithm, aid in improving resolution. We also propose an approach for designing filters with random transmittance using optical thin-film technology. We demonstrate that the improvement in resolution is 7-fold when using the filters with random transmittance over what was achieved in our previous work.

©2013 Optical Society of America

OCIS codes: (300.6320) Spectroscopy, high-resolution; (120.6200) Spectrometers and spectroscopic instrumentation; (100.6640) Super-resolution.

References and links

1. D. J. Brady, *Optical Imaging and Spectroscopy* (John and Wiley Sons, 2009).
2. S. W. Wang, C. Xia, X. Chen, W. Lu, M. Li, H. Wang, W. Zheng, and T. Zhang, "Concept of a high-resolution miniature spectrometer using an integrated filter array," Opt. Lett. **32**(6), 632–634 (2007).
3. W. L. Wolfe, *Introduction to Imaging Spectrometers* (SPIE, 1997).
4. H. N. Lee, *Introduction to Compressed Sensing* (Lecture notes; Spring Semester, GIST, Korea, 2011). *http://infonet.gist.ac.kr/?page_id=843*
5. J. Oliver, W. B. Lee, S. J. Park, and H. N. Lee, "Improving resolution of miniature spectrometers by exploiting sparse nature of signals," Opt. Express **20**(3), 2613–2625 (2012).
6. C. C. Chang, N. T. Lin, U. Kurokawa, and B. I. I. Choi, "Spectrum reconstruction for filter-array spectrum sensor from sparse template selection," Opt. Eng. **50**(11), 114402 (2011).
7. C. C. Chang and H. N. Lee, "On the estimation of target spectrum for filter-array based spectrometer," Opt. Express **16**(2), 1056–1061 (2008).
8. U. Kurokawa, B. I. Choi, and C.-C. Chang, "Filter-based miniature spectrometers: spectrum reconstruction using adaptive regularization," IEEE Sens. J. **11**(7), 1556–1563 (2011).
9. C. Bendjaballah, "Information rates in optical channels," Opt. Commun. **17**(1), 55–58 (1976).
10. Y. Aizu, K. Ogino, and T. Asakura, "A laser velocimeter using a random pattern," Opt. Commun. **64**(3), 205–210 (1987).
11. J. Ojeda-Castañeda and A. Sauceda, "Random gratings as correlator sensors," Opt. Lett. **22**(5), 257–258 (1997).
12. D. L. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory **52**, 1289–1306 (2006).
13. R. Baraniuk, "Compressive sensing," IEEE Sig. Proc. Mag. **24**(4), 118–121 (2007).
14. E. Candes and J. Romberg, "11-magic: Recovery of sparse signals via convex programming," Technical report (2005). *http://users.ece.gatech.edu/~justin/l1magic/*
15. S. Park and H. N. Lee, "Designing an algorithm to solve basis pursuit denoising with a nonnegative constraint," IEEE Sig. Proc. Letters. (submitted to).
16. R. Tibshirani, "Regression shrinkage and selection via the lasso," J. R. Stat. Soc., B **58**, 267–288 (1996).
17. A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," SIAM J. Ima. Sciences **2**(1), 183–202 (2009).
18. A. Juditsky and A. Nemirovski, "First Order Methods for Nonsmooth Convex Large-Scale Optimization, I: General Purpose Methods," in Optimization for Machine Learning, S. Sra, S. Nowozin, and S.J. Write, eds. (MIT Press, 2011), pp. 1–28.
19. S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," Founda. and Tren. Mach. Learn. **3**(1), 1–122 (2010).
20. M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," IEEE Sig. Proc. Mag. **25**(2), 83–91 (2008).
21. R. Dikpal, A. Veeraraghavan, and R. Chellappa, "P2C2: Programmable pixel compressive camera for high speed imaging," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (IEEE, 2011), pp. 329–336.

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

22. C. Li, T. Sun, K. Kelly, and Y. Zhang, "A compressive sensing and unmixing scheme for hyperspectral data processing," Technical report. (http://www.caam.rice.edu/~zhang/reports/tr1101.pdf).
23. A. Rajwade, D. Kittle, T.-H. Tsai, D. Brady, and L. Carin, "Coded hyperspectral imaging and blind compressive sensing," submitted (2012).
24. K. Madanipour and M. T. Tavassoly, "Determination of modulation transfer function of a printer by measuring the autocorrelation of the transmission function of a printed Ronchi grating," Appl. Opt. **48**(4), 725–729 (2009).
25. J. R. Barry and J. M. Kahn, "Link design for non-directed wireless infrared communications," Appl. Opt. **34**(19), 3764–3776 (1995).
26. H. A. Macleod, *Thin-Film Optical Filters* (Institute of Physics Publishing, 2002).
27. Z. B. Haim, Y. C. Eldar, and M. Elad, "Coherence-based performance guarantees for estimating a sparse vector under random noise," IEEE Trans Sig. Proc. **58**, 5030–5043 (2010).
28. C. Z. Microscopy, "Fundamentals of mercury arc lamps," *http://zeiss-campus.magnet.fsu.edu/articles/lightsources/mercuryarc.html.*

1. Introduction

Miniature spectrometers are key instruments that are required in various academic and industrial applications such as bio-medical, chemical, and environmental engineering [1]. These spectrometers, built with integrated-filter arrays, provide superior portability, flexibility, and cost-effectiveness [2]. The spectrometers also provide fine details about the various spectral components of the incident light. These spectral components reveal a wealth of information concerning the composition and the structure of the various objects being observed [3]. The ability of a spectrometer to reveal fine information is usually attributed to its resolution. Thus, in recent years, researchers have focused on improving the resolution of spectrometers.

State-of-the-art filter-array-based spectrometers rely on digital signal processing (DSP) algorithms. These algorithms may run on a general-purpose computing platform of cellular phones and laptop computers, as shown in Fig. 1, before the spectrum estimate is displayed on the screen. They can also run on a specialized DSP chip integrated into the spectrometer unit in the case of stand-alone spectrometers. These algorithms process the raw spectrum acquired by the spectrometer to reduce noise and distortions and thereby aid in improving the resolution.

In the filter-array-based spectrometers, the factors that limit the resolution are the number of filters in the filter array and the shapes of the transmittance functions of these filters [4]. In practical miniature spectrometers, the number of filters is fixed. Thus, in order to improve the resolution, it is often not a realistic option to increase the number of filters. The second factor that affects the resolution is the transmittance functions of the filters. While the transmittance function (TF) is usually designed to meet certain criteria, in practice, owing to low-cost integrated-array fabrication, the TF of each of the filters is non-ideal. These non-ideal filters highly distort the raw spectrum and thereby necessitate the digital processing of the raw spectrum to determine information concerning the spectral components of the input signal spectrum. DSP algorithms or estimators have been reported in the literature [5–8]. The algorithms in [5, 6] are centered on L_1-norm minimization, which is a modern optimization tool in DSP for solving underdetermined system of linear equations, whereas the algorithms in [7, 8] are based on conventional least squares and are designed with an aim to reconstruct the signal spectrum but not to improve resolution.

In this paper, we assume the use of the L_1-norm-minimization-based DSP algorithm in [5] as the spectral estimator and aim to address the following relevant questions: Is it possible to enhance the resolution of the spectrometers by shaping the TFs in a particular way? Namely, is there a certain shape of the TF that is suitable for use with the L_1-norm-minimization algorithm at the back end? If yes, what is that shape, and how can such a TF be designed? Is there a practical way to implement such a TF? Also of great importance is the discussion of suitable performance metrics to compare different TFs. This is necessary to reflect the use of the L_1-norm-based spectral estimator we aim to use. With the new metric, one should be able to investigate the maximum possible resolution achieved by employing different TFs.

As per the discussion of the questions above, the role of TFs in improving the resolution of the spectrometers under the DSP-based spectral estimator needs to be

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

carefully examined. We observe that the TFs whose auto-covariance functions resemble that of a Dirac delta function acquire independent information about the closely spaced spectral components of the input signal spectrum under the DSP-based spectral estimator. Filters with random-TFs possess such auto-covariance. Based on our study, we show that it is possible to design filters with random-TFs using thin-film technology and improve the resolution up to 7-fold compared to the filters with non-ideal transmittances in [5].

The concept of random-TFs goes back to the 1970s. Different aspects of the random-TFs have been studied in the past [9–11]. In [9], the author has studied the maximum information transmission rate of an optical filter with random transmittance. A laser-based velocimeter is designed in [10] based on random transmittance patterns. In [11], the authors have designed a holographic-based sensor using random transmittances that are generated by using random gratings. As per our knowledge, however, no one has designed or reported filters with random transmittance for improving the resolution of a spectrometer.

This paper is organized as follows: Section 2 presents the system model, and Section 3 discusses in detail the need for designing filters with random transmittances for improving resolution. Section 4 discusses the design of random-transmittance filters using thin-film technology, and Section 5 presents the concept of the resolution of a spectrometer. Experimental results are discussed in Section 6, and Section 7 concludes the paper.

2. System description

We consider a spectrometer that consists of a planar filter array with M filters arranged in a 2D fashion as shown in Fig. 1. Each of these filters acts as a wavelength-selective device. Each filter is specified in terms of a transmittance function (TF), $T_i(\lambda)$, $i=1,2,\cdots,M$, which indicates the fraction of input light that the filter allows at each unit of wavelength. We note that $T_i(\lambda)$ is a continuous (analog) function of wavelength. Each filter is attached to a CCD element that converts light into electrical energy. The output of each of the CCD elements is sampled to form an $M\times 1$ vector $\boldsymbol{y}$, which is called a raw spectrum.

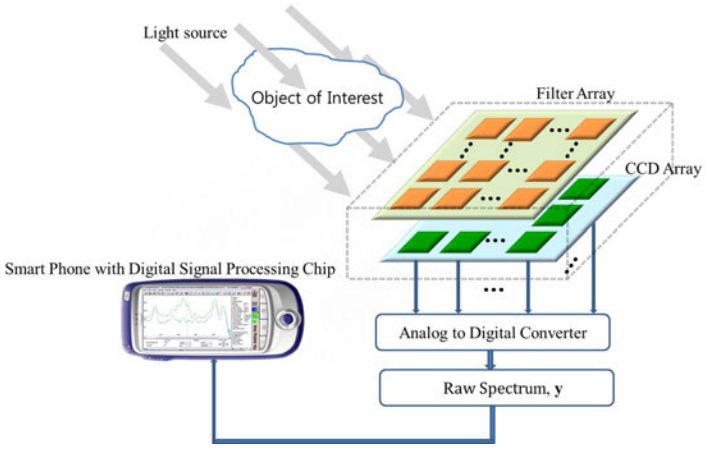


Fig. 1. Schematic of the proposed filter-array-based spectrometer.

Each sample of the raw spectrum denotes the projection (inner product) of the input signal spectrum onto a TF. That is, the ith sample of the raw spectrum is obtained as

$$y_i = \int x(\lambda) T_i(\lambda)\, d\lambda + w_i,\ i=1,2,\cdots,M, \tag{1}$$

where $x(\lambda)$ is the input signal spectrum and w_i denotes a Gaussian noise sample with zero-mean and variance σ^2. Thus, each raw spectrum sample contains information about

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

the input signal spectrum acquired by a single filter in the filter array. We observe that the model in Eq. (1) for the raw spectrum can be used to describe the raw spectrum of a grating-based spectrometer as well. Grating-based spectrometers are a more conventional way of implementing spectrometers [1]. The raw-spectrum model for the grating-based spectrometer is given [1, p. 339, Eq. (9.5)] by

$$g_i = \int x(\lambda)\, h(\lambda - i\Delta)\, d\lambda,\ i = 1,2,\cdots,M, \tag{2}$$

where g_i is the detected spectral intensity at the *i*th CCD element, $x(\lambda)$ is again the input signal spectrum, $h(\lambda)$ is the system impulse response or the instrumental function, and Δ is the sampling interval. The system impulse response $h(\lambda)$ denotes the overall impulse response of the complete system from the slit to the CCD sensor, including the transmittances of the slit and the grating. We note that the raw-spectrum models of the modern filter-based and conventional grating-based spectrometers are similar in their integral form input-output relationship. The role of $h(\lambda)$ in the grating-based spectrometer is taken by the TF, $T_i(\lambda)$, in the filter-array-based spectrometers.

Modern filter-array-based spectrometers [4–8] process the raw spectrum using a DSP algorithm or an estimator to recover the input signal spectrum. The data model for the raw spectrum, $\boldsymbol{y}$, which is an input to the DSP algorithm can be represented [4] as a system of linear equations:

$$\boldsymbol{y} = D\boldsymbol{x} + \boldsymbol{w} \quad \text{where} \quad \boldsymbol{x} \geq 0. \tag{3}$$

In Eq. (3), the $N \times 1$ vector $\boldsymbol{x}$ contains the samples of the original signal spectrum having an operating bandwidth of W_λ, and the matrix D is an $M \times N$ TF matrix. Each row of D is the TF of a filter. We observe that the matrix D contains only non-negative values as light transmittance is always non-negative.

In this paper, since we are dealing with improving resolution, the value of N is set to be greater than M, so $N > M$. That is, the number of samples of $\boldsymbol{x}$ is greater than the number of filters. The reason for setting $N > M$ is that resolution in filter-array-based spectrometers is limited by the number of filters M (as will be discussed in detail in Section 3.1). Thus, in order to increase resolution beyond this limit set by the fixed number of filters M, the value of N should be greater than M. By setting $N > M$, we note that the noise-free system of equations in Eq. (3) is underdetermined. We also note then that in Eq. (3), the spacing between the samples of $\boldsymbol{x}$ is $\Delta\lambda_N = \frac{W_\lambda}{N}$, while the spacing between the samples of the raw spectrum $\boldsymbol{y}$ is $\Delta\lambda_M = \frac{W_\lambda}{M}$.

Recently, sparse representation of the input signal spectrum, $\boldsymbol{x}$, has drawn considerable attention [4–6]. The aim of sparse representation is to decompose the signal spectrum into its constituent spectral (wavelength) components. For example, the signal spectrum of a low-power mercury lamp can be decomposed into its seven wavelength components, whereas the rest of the wavelength components are zero-valued. In general, any signal spectrum vector $\boldsymbol{x}$ in Eq. (3) can be equivalently represented as a sparse vector $\boldsymbol{s}$ using a linear transform matrix G, i.e., $\boldsymbol{x} = G\boldsymbol{s}$. The vector $\boldsymbol{s}$ is called a K-sparse vector. In this paper, we refer to $\boldsymbol{s}$ as a *sparse spectrum*. It contains K non-zero values and $N - K$ zero-valued components, $K \ll N$. The K non-zero values of $\boldsymbol{s}$ indicate the intensities of the spectral (wavelength) components that constitute $\boldsymbol{x}$. The set of K wavelengths corresponding to those intensities is called the support set. Thus, each element of the support set indicates the wavelengths (in nm) that are present in the signal spectrum. The intensities of the remaining N - K wavelength components of $\boldsymbol{s}$ are zero.

The matrix G is called the kernel matrix, and it is of size $N \times N$. The columns of G are chosen as functions that preserve the natural shape of the signal spectrum. Gaussian kernels are widely used for signal spectrum modeling [8]. The other possible non-Gaussian kernels are secant hyperbolic and Lorentzian [1]. To construct G, we need to know only a single kernel function, the width of which is controlled by a parameter called

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

full-width at half-maximum (FWHM). When the FWHM is small, the width of the kernel function is also small, and vice versa. Once the kernel function is chosen, it is sampled, and the sampled version of the kernel function is considered as the first column of G. The remaining $N-1$ columns of G are just shifted versions of the first column [6]. We observe that the Gaussian kernel has a uniform width that models narrowband light sources. For modeling broadband light sources, one may require kernels with varying widths, as discussed in [6]. In this paper, we consider only narrowband sources. With the sparse modeling of the input signal spectrum, Eq. (3) can be written as

$$\boldsymbol{y} = \underbrace{DG}_{A}\boldsymbol{s} + \boldsymbol{w} = A\boldsymbol{s} + \boldsymbol{w} \quad \text{where} \quad \mathrm{s} \geq 0\,. \tag{4}$$

The goal of the DSP algorithm or the spectral estimator is to obtain an estimate $\hat{\boldsymbol{s}}$ f $\boldsymbol{s}$ from the raw spectrum $\boldsymbol{y}$, given the matrices D and G.

3. Proposed transmittance functions

3.1. Introduction: Transmittance functions

The transmittance function (TF) of an optical filter is a waveform that has a shape that dictates the fraction of input light that the filter transmits at various wavelengths. The conventional approach is to attempt to design TFs with an ideal shape, i.e., a sharp brick wall as shown in Fig. 2(a). In Fig. 2(a), we show only three filters out of 40 in a filter array. Ideal filters allow light only for desired wavelengths, called the passband, and completely stop the remaining wavelengths, called the stopband. In addition, from Fig. 2(a), we note that the passbands of the ideal TFs do not overlap with each other.

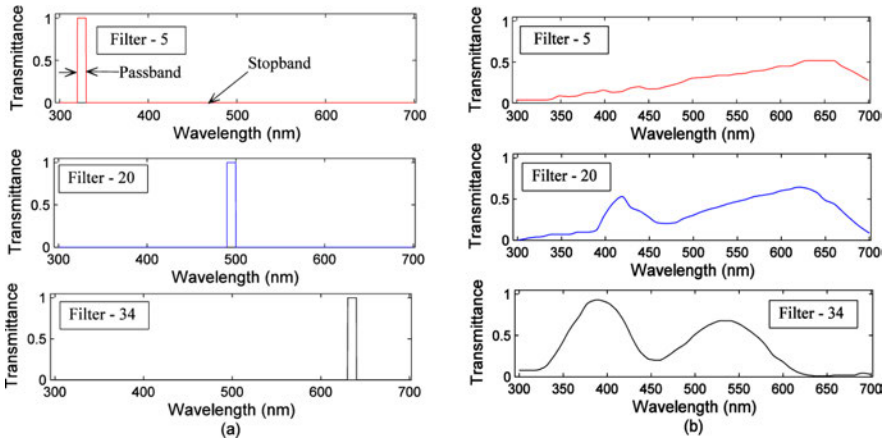


Fig. 2. Transmittances of (a) ideal filters (b) non-ideal filters.

Ideal filters are desired for two reasons. First, the raw spectrum (Eq. (3)) obtained by employing the ideal TFs is a direct estimate of the signal spectrum, and hence, there is no need for a DSP algorithm. This is because each sample of the raw spectrum is a projection (Sec. 2) of the input signal spectrum with an ideal TF. Since ideal TFs are non-overlapping by design, each sample of the raw spectrum only contains information about the spectral (wavelength) components that correspond to a single passband. Second, any two spectral components of the input signal spectrum that are one passband spacing apart are distinguishable (resolvable) in the raw spectrum. If there are any two spectral components separated by less than a passband width, they cannot be resolved. In order to resolve such spectral components, the passband width of each filter should be decreased. Consequently, this increases the number of TFs (number of filters) in order to cover the entire operating bandwidth. Thus, the resolution of the filter-array-based spectrometers with ideal TFs is limited by the number of filters, M. Since the spacing between the samples of the raw spectrum is $\Delta\lambda_M = \frac{W_\lambda}{M}$, it was believed to be the limit on the

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

resolution of filter-array-based spectrometers. This limit is the same for the more conventional grating-based spectrometer with a CCD array consisting of a total of M pixels. That is, by using the ideal system impulse response ($h(\lambda) = \delta(\lambda)$, the Dirac delta), the spacing between the samples of the raw spectrum of a grating-based spectrometer is also $\Delta\lambda_M = \frac{W_\lambda}{M}$.

However, filters with ideal TFs are not easy to realize in practice [2]. Therefore, non-ideal TFs are inevitable in spectrometers. The shapes of the practical non-ideal TFs employed in a typical filter-array spectrometer [7] are shown in Fig. 2(b). These TFs are not purposely designed in such shapes but are instead accidentally obtained in the micro-level processing of implementing the ideal filters. Unlike the ideal TFs, the waveforms of non-ideal TFs are smooth functions of wavelength. In addition, as seen from Fig. 2(b), the passband of each filter's TF leaks into the passbands of neighboring filters. Hence, non-ideal TFs have considerable overlap with each other. Non-ideal TFs are considered worse for the signal spectrum estimation process in conventional wisdom. The reason is that each sample of the raw spectrum now contains information about the spectral components not only from its own band but also from the neighboring filters' bands, *viz*. interferences. The interference from the use of non-ideal TFs causes severe distortion in the raw spectrum. Therefore, for the spectrometers with non-ideal TFs, post-processing of the raw spectrum using DSP algorithms is inevitable. These algorithms are designed with the aim of extracting the constituent spectral components (sparse spectrum) of the input signal spectrum from the raw spectrum $\boldsymbol{y}$, via Eq. (4).

A recent recovery algorithm, reported in [5], is tailor-made for spectrometers. This algorithm is based on a DSP optimization tool called L_1-norm minimization. L_1-norm minimization is a recent approach more suitable for solving a noise-corrupted underdetermined system of linear equations such as Eq. (4) than a classical least-squares approach [12–14]. The minimization problem for the recovery of the sparse signal spectrum $\boldsymbol{s}$ in Eq. (4) can be expressed as:

$$\hat{\boldsymbol{s}} = \min_{s} \|\boldsymbol{s}\|_1 \quad \text{subject to} \quad \|DG\boldsymbol{s} - \boldsymbol{y}\|_2 \le \varepsilon, \ \boldsymbol{s} \ge 0 \tag{5}$$

where ε is a small positive constant.

Note that the problem in Eq. (5) is a bit different from the standard L_1-norm minimization approach with the inclusion of the additional non-negative constraint, $\boldsymbol{s} \ge 0$. In order to find the non-negative spectral estimate $\hat{\boldsymbol{s}}$, the authors in [5] have used a DSP algorithm that was reported in [15]. In [15], the authors have derived an algorithm using the modern interior point method. They have shown that the performance of their algorithm is best when compared to the standard L_1-norm minimization algorithms such as L_1 magic [14] and LASSO [16], which do not use the additional information. In addition, they have shown that their algorithm is robust against the observation noise, stable, and competitive with the existing algorithms. Using the algorithm in [15] for spectrometers, the authors in [5] have shown an improvement in resolution to about a factor of six below the resolution limit. In addition, the authors in [5] have provided a hint that the recovery performance (and hence the resolution) of the algorithm depends on the design of the TF matrix D.

Recently, fast algorithms for L_1-norm minimization such as FISTA [17], FOM [18], and ADMM-based algorithms [19] have gained considerable interest among the imaging community. In imaging systems [20–23], the number of measurements (typically more than a few million) is much more than the number of measurements used in the miniaturized spectrometers. Consequently, these imaging systems should be capable of handling large amounts of data efficiently; for spectrometers, however, the accuracy is more important than the efficiency. It is well known that the interior-point-method-based algorithms provide good accuracy. Thus, in this paper, we use the same algorithm that was used in [5]. Readers who may be interested in implementing fast algorithms for spectrometers can refer to the approaches in [17–19].

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

In this paper, our central goal is to design the TF matrix D in order to enhance the resolution. That is, we use the spectral estimation algorithm in [5] and show how to design the TF matrix D (which in turn determines the filter transmittances) that helps the algorithm to improve resolution. We show that by using our proposed TF matrix, it is possible to enhance resolution 7-fold compared to what was achieved in [5].

3.2. Motivations for new TF design

In this paper, we aim to present a new approach to further enhance (as compared to what was achieved in [5]) the resolution of the spectrometers by designing TFs. The motivations behind designing new TFs are based on the following observations:

1. We recall that each sample of the raw spectrum is modeled as a projection of the input signal spectrum onto a filter TF. In spectrometers with non-ideal TFs, the projection captures not only the in-band but also the out-of-band information about spectral components in each sample of the raw spectrum. Traditionally, when no DSP is used, unintentional capturing of the out-of-band information that is mixed with the information of the desired band is considered as pure distortion of the spectral components. In a completely different viewpoint, however, it can be considered as additional source from which useful information can be extracted. Namely, because of the shapes of the non-ideal filters, each sample of the raw spectrum contains extra information about the entire signal spectrum rather than the only about a particular band like the sample of an ideal TF does. That is, each non-ideal filter collects information from the entire signal spectrum and maps it into a single sample of the raw spectrum. Since the shape of each non-ideal filter is different from that of all other non-ideal filters, we can get many such independent "holistic" views of the entire signal spectrum from each sample of the raw spectrum. This led us to the very natural question: What kind of TFs should provide more holistic and independent information about the spectral components in each sample of the raw spectrum?

2. The modern, L_1-norm-minimization-based DSP spectral estimator [5] recovers the spectral components very well from the distorted raw spectrum, prompting us to investigate the reason for this unexpected level of performance. We found that the problem of resolving (identifying) the spectral components essentially reduces to the unique identification of $\boldsymbol{s}$ from $\boldsymbol{y}$. The success of this identification depends on the matrices D and G in Eq. (4). For a given class of signal spectrum, the matrix G is fixed. However, the matrix D can be determined from the design and fabrication of the filter TFs. Hence, we conclude that a good design for the TFs is important in order for the DSP algorithm to recover the spectral components from the raw spectrum.

Thus, it is crucial to design TFs that 1) capture holistic and independent information about the constituent spectral components and 2) aid in recovering those spectral components.

3.3. Design of new TFs

We observe from the previous section that the amount of additional information acquired in each sample of the raw spectrum is solely proportional to the shape of the TFs. By shaping the TFs through a proper design procedure, it is possible to extract much information about the input signal spectrum. Therefore, the TF matrix D that senses the maximum amount of information about the signal spectrum appears to be the best choice for the spectrometer case, and hence, it is our design goal. We note here that a similar design notion arises in the emerging signal processing research area called compressive sensing (CS) where a sensing matrix [12, 13] takes the role of the TF matrix in acquiring a signal.

In CS, the raw signal samples acquired using the matrix D is modeled as $\boldsymbol{y} = DG\boldsymbol{s} + \boldsymbol{w} = A\boldsymbol{s} + \boldsymbol{w}$, which is the same as the raw spectrum model in Eq. (4). In CS,

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

the matrix D is referred to as the sensing matrix; the matrix G, the sparsifying basis; the vector $\boldsymbol{s}$, a K-sparse signal; and $A\boldsymbol{s}$ is called the compressed signal. In CS, the sensing matrix D is chosen such that it captures enough information for the unique identification of the signal $\boldsymbol{s}$ with as few samples of $A\boldsymbol{s}$ (the compressed signal) as possible. In order to quantify the ability of the sensing matrix to acquire enough information about the signal in a minimum number of raw samples, the mutual coherence metric, μ, is used [12, 14]. The mutual coherence of the matrix $A = DG$ is defined as follows:

$$\mu = \max_{i \neq j} \left| \left\langle a_i, a_j \right\rangle \right| \quad i, j = 1, \cdots, N \tag{6}$$

where a_i is the ith column of A. The μ is a measure of the maximum possible correlation among pairs of columns of A. The smaller μ is, the smaller the correlation among the columns of A is. The smaller the correlation is, the better the reconstruction accuracy of $\boldsymbol{s}$ from $\boldsymbol{y}$ is. Thus in CS, it is desirable to design the sensing matrix D, such that the matrix $A = DG$ has low coherence.

It is well known in CS that sensing matrices D, the entries of which are drawn from i.i.d. samples of a random variable, exhibit low coherence. Such matrices are called random sensing matrices. These matrices are capable of capturing enough information about the signal $\boldsymbol{s}$ to perform reconstruction from a small number of samples of $\boldsymbol{y}$. Therefore, random sensing matrices are widely employed in CS-based applications.

It appears that the goal of designing a sensing matrix in CS resembles that of TF matrix design. That is, both the TF and the sensing matrix should be designed to capture sufficient information to permit the faithful recovery of the signal. Since both of these goals are met by random matrices [12], in this paper, we consider a random matrix as the TF matrix rather than using the non-ideal TF matrix as in [5–8]. Since each row of the TF matrix denotes a transmittance function of a filter, a row of the random TF matrix is termed a random transmittance function. That is, a filter with a transmittance that exhibits random fluctuations as opposed to possessing a pre-designed transmittance shape is termed a random transmittance function. Thus, our goal of improving resolution reduces to the design of a set of M filters with random TFs. A set of TFs $\{T_i(\lambda)\}$ is referred to as random when

1) The auto-covariance function (ACF) of each TF in the set is close to the Dirac-delta function. That is,

$$\delta(\Delta\lambda) \approx \int \left[T_i(\lambda) - m_i\right]\left[T_i(\lambda + \Delta\lambda) - m_i\right] d\lambda, \ \ i = 1, 2, \cdots, M, \tag{7}$$

where $\delta(.)$ is the Dirac-delta function, m_i is the mean of $T_i(\lambda)$, and $\Delta\lambda$ is the wavelength difference, and,

2) when the cross-covariance function (CCF) between each pair of TFs in the set is very small. That is, $\varepsilon_{i,j} = \int \left[T_i(\lambda) - m_i\right]\left[T_j(\lambda) - m_j\right] d\lambda, \ i \neq j$, where $\varepsilon_{i,j}$ is a small number close to zero. We discuss in detail the design of filters with such random transmittance in Section 4 using thin-film technology.

3.4. Significance of the proposed approach

The ideas in this paper mirror some of the prior work in CS with regard to the use of random matrices for signal acquisition and recovery. We briefly compare our approach to other well-known CS imaging approaches such as the single-pixel camera (SPC) [20], the P2C2 video system [21], the hyperspectral design (HSD) [22], and the CASSI system [23].

First, we note that the SPC is a well-known CS-based imaging system [20]. In SPC, in order to compressively sample a scene by optical random linear projection, a digital micromirror device (DMD) was used together with a biconvex lens. The DMD is programmed with random 0/1 Walsh patterns [20]. With the availability of such optical devices, random projections of a scene were made almost effortlessly without going

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

through a pixel-by-pixel scanning, and as a result, it was possible to realize the concept of SPC in practice. The P2C2 video system [21] extends the SPC architecture to video acquisition, and hence it also employs the random 0/1 Walsh patterns. We observe that in both the systems the measurement matrices are 0/1 Walsh matrices.

HSD [22] and CASSI [23] are spectral imaging systems. These systems collect a scene as a set of spatial images. Each image represents a range of an electromagnetic spectrum also known as a spectral band. These spatial images are arranged along various spectral bands to form a three-dimensional structure called a hyperspectral data cube. In HSD, the data cube is modeled as a linear combination of the spectral signatures of endmembers (that is, objects in a scene such as a road, grass, trees, or a roof), where each spectral signature can be obtained from a publicly accessible database such as HYDICE Urban hyperspectral data as used in [22]. Then, the HSD exploits the spatial and spectral redundancies of the data cube by using CS. In particular, HSD adopts the concept of SPC for acquisition of the data cube, and hence they use the DMD with 0/1 Walsh patterns. In contrast to HSD, the CASSI system models the data cube as a linear combination of two-dimensional wavelets and then acquires the data cube using the coded aperture masks. Thus, the measurement matrices in the CASSI system are also random binary 0/1 code matrices. We observe that both the HSD and CASSI systems compressively acquire the spectral data by means of spatial domain random measurement matrices, each element of which is either 0 or 1.

In contrast to the above imaging systems, we note that in our proposed spectrometer, we acquire the spectral data or signal spectrum using random spectral filters (not spatial). That is, as discussed in Section 2, the sensing of the signal spectrum is performed using random analog optical TFs, which are non-negative continuous functions of wavelength. We observe that the set of optical TFs play a sensing role as significant as that of the DMD in single-pixel cameras or the coded aperture masks. In addition, the TFs have values that can be any real number between 0 and 1, in contrast to the 0/1 patterns in the imaging applications. Table 1 summarizes the measurement matrices and their implementation in various imaging applications.

Second, in optical applications the sensing of signals and their subsequent recovery is accomplished with a sensing matrix stored beforehand. For example, in [20, p. 85], the DMD array stores Walsh, Hadamard, or noiselet patterns. The best matrices or the matrices of interest were designed in the digital domain. They were then used to dither the micromirrors back and forth in the DMD. We call this the *digital design first* approach here, as depicted in Fig. 3. In spectrometers, the actual mechanism of sensing is done using the analog TFs, as per Eq. (1), and then the signal spectrum recovery is done using its digital model. We note here that the digital TFs can be losslessly obtained from the TFs of analog thin-film filters (discussed in the next section). Thus, once the analog TFs are designed, they can be digitally represented in the TF matrix in Eq. (4) for spectrum recovery. We call this the *analog design first* approach in this paper (see Fig. 3), in contrast to the *digital design first* approach. The comparison of the *analog design first* approach with the *digital design first* approach may unearth a noteworthy characteristic of the spectrometer designs. Let us elaborate a little further.

Table 1. Measurement Matrices and Their Implementation in Various Imaging Applications

Application	Measurement matrix / Implementation	Page number
Single-pixel camera (2008) [20]	Random 0/1 Walsh matrix DMD array	87 84
P2C2 system for sensing videos (2011) [21]	Random 0/1 Walsh matrix DMD array	331 335
Hyperspectral data processing (2009) [22]	Random 0/1 Walsh matrix DMD array	9 13
Coded aperture snapshot spectral imaging (CASSI) system for hyperspectral video (2012) [23]	Binary CASSI code matrix Binary coded mask	6 15

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

For spectrometers, it is desirable to design a set of random analog TFs with the ACF of each TF shaped as close as possible to the Dirac delta function. The narrower the width of the ACF is, the better the set is at resolving the fine details of the input spectrum. Thus, the resolution of a spectrometer is closely related to the width of the ACF of the TF. In the *analog design first* approach, due to the optical nature of the thin-film filters, there is an inevitable practical limit on how narrow the width of an ACF can be made.

We note that there is no such limit for the digital design. That is, in the digital case, the width of the ACF can be made as arbitrarily narrow as one wishes. For example, this can be achieved by dividing the operating wavelength range by the number of samples, say N, and hence obtaining any resolution we wish to have, and then by digitally generating a TF of length N by drawing independent samples from a probability distribution such as the 0/1 Walsh matrix with $N = 65{,}536$ as in [20] or as in P2C2 video systems [21]. Since the samples of the TF are independently and identically generated from the given distribution, the ACF of the TF is always as close as possible to the Dirac delta function. This means that one can increase the resolution just by increasing N. This may work for image recovery applications, subject to the creation of a DMD array with so many more elements, but not for spectrometers. For spectrometers, the inherent width of the ACF of the analog TFs must be preserved in its digital TFs. If we first design a random (white) digital TF with an ACF very close to the Dirac delta, the digital TF may not be realizable as the TF of an analog filter. Thus, the resolution in the digital domain is a useless concept unless a way of constructing the practical analog counterpart of the digital TF is given.

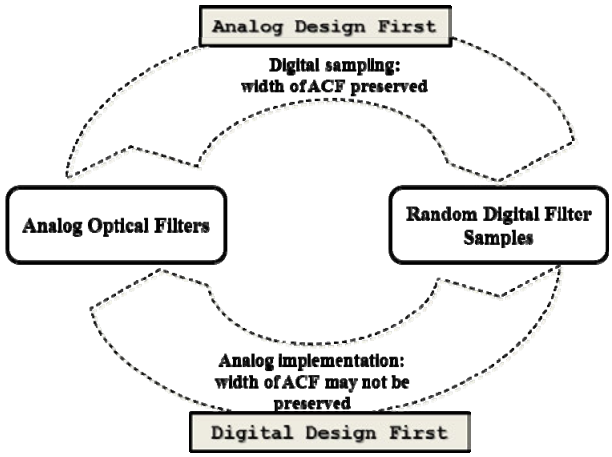


Fig. 3. Summary of *analog* and *digital design first* approaches.

Our main contribution is twofold. The first part shows a specific way to design a set of optical analog filters with TFs that are white and uncorrelated with each other. The second part illustrates that significant improvement in resolution can be achieved with such a filter design.

3.5. Characterizing random TFs

Usually, the degree of correlation of a function over an interval is characterized in terms of the auto-covariance function (ACF). A function can be correlated with a time lagged (or advanced) version of itself. Thus, the ACF is a function of the time difference (time lag). The ACF reveals the correlation among the samples of the function. Similarly, a random TF can be characterized in terms of their ACF, which is now a function of wavelength difference. The shape of the ACF reveals the degree of similarity (correlation) between the light intensities at two different wavelengths of the random TF. If the shape of the ACF is wide with slowly decaying tails, then the TF is said to be highly correlated. On the other hand, if the TF has low correlation at different wavelengths, then the shape of the ACF is narrow with rapidly decaying tails, resembling

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

a Dirac delta function. The shape of the Dirac delta function is ideal for an ACF. The width of the ACF measures the degree of correlation. That is, an ACF with narrow width exhibits low correlation, and vice versa. A Dirac-delta-like ACF is required in applications such as determining the modulation transfer function of an optical system [24] or detecting displacements or phase gradients in holography [11]. This evidence shows that filters with a Dirac-delta-like ACF are crucial in optical systems and can be designed in practice.

Filters having ACFs that resemble that of a Dirac delta are preferred for L_1-norm-minimization-based spectrometers. The reason is that a Dirac-delta-like ACF has narrow width, and hence two spectral components that are separated by more than this width are sensed differently by the TF. This independent information about the spectral components aids the L_1-norm-based algorithms to resolve closely spaced spectral components. Thus, the resolution of a spectrometer has a close connection to the width of the ACF. By using random TFs, we can reduce the width of the ACF and hence the spacing between two spectral components that are independently sensed. On the other hand, non-ideal TFs exhibit high correlation, and hence the widths of the ACFs are wider than those of the random TFs. Hence, information gathered about the two nearby spectral components is not too independent. This reduces the ability of the spectrometer to distinguish the two spectral components. Thus, our goal in this paper is to design the transmittances of filters with Dirac-delta-like ACFs. We call them *random* TFs, while the TFs in [5], *correlated* TFs, in the subsequent discussions.

4. Design of the proposed random transmittance

In the previous section, we introduced random TFs with the goal of improving the resolution of a spectrometer system. Towards this end, we ask ourselves the following question: is it possible, in practice, to design and implement *random* TFs in the analog domain? If yes, how should one be designed? We have found that filters with random transmittance can be designed and implemented by using thin-film optical filters [25, 26] in the way explained in this section.

A thin-film optical filter consists of multiple layers of high- and low-refractive index materials (dielectrics) deposited on a substrate [26]. Each layer has a thickness usually of one quarter-wavelength. Thin-film filters work based on the interference of light transmitted or reflected at the boundaries between the layers. This interference is wavelength-dependent; that is, depending on the number of layers, the index of refraction, and the thicknesses of the layers, the transmission (or reflection) of light through the filter changes with wavelength, and hence the overall transmittance changes. By controlling the thickness of each layer and the number of layers, optical engineers design various filter transmittances such as band-pass filters. In this section, we show how varying the thickness of each layer generates random transmittances using the design methodology in [25].

4.1 Thin-film-based random transmittance filter design

In this section, we first briefly discuss the generic design of thin-film filters to generate band-pass-like transmittances. We then motivate and show how to design random transmittances using the same thin-film structure by varying only the thicknesses of the layers. Let us consider a thin-film filter with m dielectric layers between the input medium (air) and the output medium (substrate). Let n_a and n_s denote the refractive indices of the input and output medium, while n_H and n_L refer to the refractive indices of the high-refractive-index and low-refractive-index intermediate dielectric layers, respectively, and d_k is the thickness of layer k. Assuming no dielectric losses, the transmittance of the thin-film filter as a function of wavelength (λ) and the angle of the incident light (θ) is given [25] by

$$T(\lambda,\theta) = 1 - \tfrac{1}{2}\left(\left|\rho_{TE}\right|^2 + \left|\rho_{TM}\right|^2\right) \tag{8}$$

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

where ρ_{TE} and ρ_{TM} are power parameters that measure the powers that are contained in the TE and TM polarization modes of the input light.

The power parameters in Eq. (8) are calculated using a recursive set of equations shown in Table 2. We note from Table 2 that the power parameters (and hence the transmittance) depend on the thickness and the refractive index of the different media, as well as on the wavelength and angle of the incident light. Optical engineers use these recursive equations to design the transmittance shapes, such as the band-pass profile, that are required in practice. For example, the typical band-pass transmittances designed using Eq. (8) for various angles of incidence are provided in [25, p. 3765, Fig. (1)]. Also, it is shown in [25, p. 3767, Fig. (3)] that the band-pass transmittances designed using the equations in Table 2 are accurate and are close to the actual transmittances obtained in practice. In Table 2, θ_k denotes the angle that the incident light makes with the normal as it travels from layer k to $k+1$, η_k is the effective complex-valued refractive index seen by the light as it enters layer k. In order to calculate ρ_{TE} and ρ_{TM}, we first start with $\eta_m = N_m$, and we apply the equation for η_k recursively until we arrive at η_2, which when substituted into the equation for ρ in Table 2 yields either ρ_{TE} or ρ_{TM}.

In this paper, we adopt the design equations in Table 2 to generate the random transmittances. We are able to generate the random transmittances based on the following observation: We first observed that thin-film filters can provide band-pass-like transmittance because of the principle of interference of light beams. That is, the incident light, when traveling through the thin-film filter, suffers different phase shifts as a result of propagating through different layers. The phase shifts depend on the refractive indices, the thicknesses of the layers, and the wavelength of the incident light. After transmission through the layers, the different beams that reappear at the filter output combine (add) either constructively or destructively depending on their path lengths. The constructive addition leads to increased transmission, and destructive addition leads to decreased transmission of the incident light. Since these constructive and destructive additions are wavelength-dependent, optical engineers usually design the thicknesses of the layers such that constructive interference occurs only for certain wavelengths and thereby achieve band-pass-like transmittances.

Table 2. Recursive Equations for Calculating Power Parameters

$$\rho = \frac{N_1 - \eta_2}{N_1 + \eta_2}$$

$$N_k = \begin{cases} n_k / \cos\theta_k & for\ \text{TE} \\ n_k \cos\theta_k & for\ \text{TM} \end{cases}, \quad k \in \{2, \ldots, m\}$$

$$\eta_k = N_k \frac{\eta_{k+1}\cos\beta_k + jN_k \sin\beta_k}{N_k \cos\beta_k + j\eta_{k+1}\sin\beta_k}, \quad k \in \{2, \ldots, m\}$$

$$\theta_k = \sin^{-1}\left(\frac{n_{k-1}}{n_k}\sin\theta_{k-1}\right), \quad k \in \{2, \ldots, m\}$$

$$\beta_k = 2\pi \cos(\theta_k) n_k d_k / \lambda$$

To avoid boundary reflections (and hence to obtain efficient transmission), it is preferred to have a 180° phase shift between the reflected beams at the boundary between the layers. This phase difference directly corresponds to a $\lambda/2$ phase shift of a sinusoidal wave, which can best be accomplished by setting the optical thickness of the layers to $\lambda/4$. Hence, in practice, stacks of quarter-wavelength layers are used as a basic building block [26] for many types of thin-film filters such as band-pass, long-wave-pass. For example, thin-film layers with thickness $\lambda_{\text{normal}}/4n$ are very common in practice, where λ_{normal} is the center wavelength of the band-pass filter with normal incidence and n is the

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

refractive index of the layer. Thus, we note that the filter transmittances have a direct relationship with the thicknesses of the layers. Based on these observations, we inferred that it is possible to create variations in the filter transmittances by just changing the thicknesses of the layers.

In this paper, in order to generate random transmittances, we propose to vary the thicknesses of the filters, in contrast with the conventional fixed quarter-wavelength thickness. Specifically, we vary the thicknesses of the layers as $\lambda_{\text{normal}}/U$, where U is a uniform random variable that has a mean and variance that are chosen such that the average thickness of a layer is on the order of μm, similar to the thickness of the quarter-wavelength layer. We note that once we generate the random thicknesses ($\lambda_{\text{normal}}/U$), they are inserted into the design, and hence the thicknesses are no longer random after they are generated. We use the same design equations shown in Table 2. The only parameter that changes in the design is β_k (in Table 2), which depends on the thickness. The other recursive equations stay the same. We note that we can produce M independent random filters by generating M different sets of random thicknesses.

4.2. Designed random transmittances and their ACF and CCF

Figure 4(a) shows the random TFs of three filters (5, 20, and 34) generated by the thin-film method for the specifications given in Section 6. It is apparent from Fig. 4(a) that it is possible to generate a random transmittance just by varying the thickness of the layers in the thin-film filters. Figure 4(b) shows the ACF of the transmittance of Filter-5. It is evident from Fig. 4(b) that the shape of the ACF of thin-film-based transmittances is Dirac-delta-like with near-zero correlations other than the zeroth lag, which is what we desire in the proposed spectrometers. Figure 4(c) shows the CCF between Filter-5 and Filter-20. We infer from Fig. 4(c) that the CCF values are near zero for all lags, as expected. We obtained similar ACFs and CCFs for the other filters as well. An additional advantage of using the random TFs is that no stringent filter design constraint, like keeping an exact quarter-wavelength thickness, has to be followed, and this opens the way for mass production. In Section 6, we demonstrate the improvement in resolution obtained by using the proposed random transmittances over the non-ideal transmittances in [5–8]. The implementation of the random transmittance filter array is currently underway in our laboratory.

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

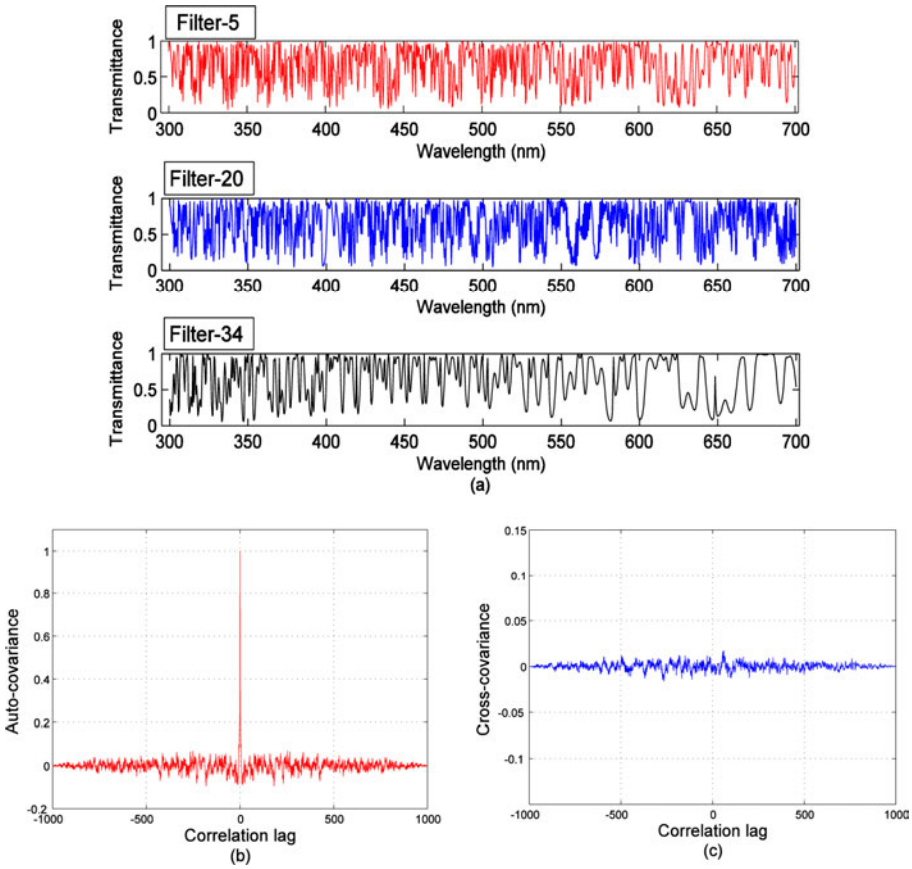


Fig. 4. (a) Random transmittances produced by the thin-film method. (b) ACF of Filter-5. (c) CCF between Filter-5 and Filter-20.

5. Resolution of filter-array-based spectrometers

Traditionally, the resolution of a spectrometer characterizes its ability to distinguish the peaks of two closely spaced spectral components of the input signal spectrum. The spectrometers that identify the closely spaced spectral components reveal fine details about the input signal spectrum. Thus, the quality of the spectrometers is usually specified in terms of this resolution. This was perhaps good enough for the conventional spectrometers but not appropriate for those that use a DSP algorithm. In this section, therefore, we first introduce a metric useful for defining the resolution that can be achieved by using various TF matrices introduced in Section 3. We then provide a new definition of resolution suitable for DSP-based filter-array spectrometers.

5.1. Performance metric useful for defining resolution

As discussed in Section 3.2, the problem of resolving distinct spectral components is equivalent to the exact recovery of the sparse spectral components of Eq. (4) by DSP algorithms. There are infinitely many different sparse spectra **s**. Thus, to determine the resolution, it is required to test the recovery of as many sparse spectra **s** as possible for a given TF matrix. Usually, a Monte Carlo simulation is used to evaluate if, on average, a specified resolution can be achieved by a TF matrix. The drawback of Monte Carlo simulation is that it is time-consuming. Therefore, in this section, we introduce a metric called genie-aided mean square error (g.MSE). It is a closed-form mean square error expression obtained assuming that the estimator knows the support set, i.e., a set of constituent wavelengths of the input signal spectrum. Thus, g.MSE can be used to indicate the minimum MSE obtainable by the L_1-norm-based DSP estimator. We will

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

provide a comparison based on the more practical MSE as well and discuss in Section 6 how well the g.MSE predicts the MSE that is achieved in practice.

Consider again the data model $\boldsymbol{y} = DG\boldsymbol{s} + \boldsymbol{w} = A\boldsymbol{s} + \boldsymbol{w}$ given by Eq. (4). Given $\boldsymbol{y}$ and A, the aim of the L_1-norm-based DSP algorithms is to find an estimate $\hat{\boldsymbol{s}}$ of $\boldsymbol{s}$. The DSP algorithm should find: 1) a set of constituent wavelengths (the support set) and 2) the non-zero intensity values corresponding to those constituent wavelengths. Let us suppose that the estimator exactly knows the support set. Then, by removing all the non-support set elements from the data model, the raw spectrum can be written as $\boldsymbol{y} = A_k\boldsymbol{s}_k + \boldsymbol{w}$, where $\boldsymbol{s}_k$ is a $K \times 1$ vector that contains the K non-zero intensity values of $\boldsymbol{s}$, and A_k is the $M \times K$ submatrix, the columns of which are the K corresponding columns of A. Note that this is an over-determined problem now, and finding $\hat{\boldsymbol{s}}$ essentially reduces to estimating the intensity vector $\boldsymbol{s}_k$ from $\boldsymbol{y}$.

An estimator that has complete knowledge of the support set is called an oracle estimator [27]. By making use of this prior information, the oracle estimator determines $\boldsymbol{s}_k$ from $\boldsymbol{y}$ using the conventional least-squares approach. The minimum MSE of the oracle estimator is termed a genie-aided MSE (g.MSE). This g.MSE provides a tight lower bound on the accuracy of an estimator that finds $\boldsymbol{s}_k$ from $\boldsymbol{y}$. In this paper, we use the g.MSE to predict the maximum achievable resolution, because

1. It is often used as a golden standard against which the performance of practical algorithms can be compared [27].
2. It can be pre-computed, with just the knowledge of A.

We now define the g.MSE. Let $\hat{\boldsymbol{s}}_k$ be the estimate of $\boldsymbol{s}_k$, the $K \times 1$ intensity vector. Let $\boldsymbol{e}_s = (\boldsymbol{s}_k - \hat{\boldsymbol{s}}_k)$ be the error vector. The co-variance of the error vector $\boldsymbol{e}_s$ is then given by $C_s = \mathbb{E}\left[\boldsymbol{e}_s\boldsymbol{e}_s^T\right] = \sigma^2\left(A_k{}^T A_k\right)^{-1}$ where A_k denotes an $M \times K$ submatrix of A obtained by taking only those K columns that correspond to K wavelengths in the support set. We note that the *expectation* operation is taken over the probability distribution of the noise vector. The matrix C_s is of size $K \times K$. Now, the average MSE, $\mathbb{E}\left[\left\|(\boldsymbol{s}_k - \hat{\boldsymbol{s}}_k)\right\|_2^2\right]$, is calculated as $\mathbb{E}\left[\left\|(\boldsymbol{s}_k - \hat{\boldsymbol{s}}_k)\right\|_2^2\right] = \mathbb{E}\left[\boldsymbol{e}_s^T\boldsymbol{e}_s\right] = tr\left(C_s\right)$, where $tr\left(C_s\right)$ is the trace of the matrix C_s. Since the input signal spectrum is $\boldsymbol{x} = G_k\boldsymbol{s}_k$ and $\hat{\boldsymbol{x}} = G_k\hat{\boldsymbol{s}}_k$, the minimum MSE between $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ with the known support set can be calculated. We define the error vector $\boldsymbol{e}_x = (\boldsymbol{x} - \hat{\boldsymbol{x}}) = G_k(\boldsymbol{s}_k - \hat{\boldsymbol{s}}_k)$. The co-variance of the error vector $\boldsymbol{e}_x$ is then given by $C_x = \mathbb{E}\left[\boldsymbol{e}_x\boldsymbol{e}_x^T\right] = \sigma^2 G_k\left(A_k{}^T A_k\right)^{-1} G_k{}^T$. Now, the average genie-aided MSE is calculated as

$$\text{g.MSE}(A, \sigma^2, K) \triangleq \mathbb{E}\left[\left\|(\boldsymbol{x} - \hat{\boldsymbol{x}})\right\|_2^2\right] = \mathbb{E}\left[\boldsymbol{e}_x^T\boldsymbol{e}_x\right] = tr\left(C_x\right) = tr\left(\sigma^2 G_k\left(A_k{}^T A_k\right)^{-1} G_k{}^T\right). \quad (9)$$

For simplicity, we drop the arguments of g.MSE in the subsequent discussions. We note that the MSE that we would obtain by using any spectrum estimator is always bounded below by the g.MSE.

We note from Section 2 that a sparse spectrum $\boldsymbol{s}$ contains only K non-zero wavelength components, and the remaining N-K wavelength components are zero. In addition, these K non-zero wavelength components can be present anywhere among the N possible wavelength components. That is, for a given N and K, there are $\binom{N}{K}$ possible sets of locations (support sets) where each set contains K wavelengths. Therefore, we can obtain $\binom{N}{K}$ g.MSEs. Each g.MSE is for a signal spectrum with K wavelengths. We note that for a fixed K, increasing N increases the number of support sets exponentially. In the next section, we define the resolution of a DSP-based spectrometer based on the g.MSE. In this paper, we always express the g.MSE in decibels (dB) for convenience.

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

5.2. Resolution: A DSP-based definition

As discussed in Section 3.1, when no DSP is used, the resolution is determined by the number of filters, M, in the filter array. That is, $\Delta\lambda_M = \frac{W_\lambda}{M}$ is the minimum distance between two spectral components that can be distinguished by the spectrometer. The larger the number of filters is, the better the resolution is. However, in miniature spectrometers, the number of filters is fixed, and hence the resolution is limited. Thus, $\Delta\lambda_M$ is called the resolution limit of the filter-array-based spectrometers. In order to further improve the resolution (beyond the resolution limit), DSP estimators are used [5]. These estimators exploit the sparse nature of the signal spectrum (Section 2) and employ L_1-norm-based estimators to improve the resolution. In this section, we introduce a new definition to measure the resolution that can be obtained by using the DSP estimators for the filter-array-based spectrometers.

We recall from Section 2 that the spacing between the samples of the input signal spectrum $\boldsymbol{x}$ (or equivalently $\boldsymbol{s}$) is $\Delta\lambda_N = \frac{W_\lambda}{N}$ while the spacing between the samples of the raw spectrum is $\Delta\lambda_M = \frac{W_\lambda}{M}$. Since $M < N$, $\Delta\lambda_M > \Delta\lambda_N$, and therefore, the raw spectrum is in a low-resolution state ($\Delta\lambda_M$), while the signal spectrum $\boldsymbol{x}$ is in a high-resolution ($\Delta\lambda_N$) state. The high-resolution state can be realized and determined when a quality estimate of $\boldsymbol{x}$ is obtained for all possible $\boldsymbol{x}$.

A way to define the resolution of a DSP-based spectrometer therefore is to find the minimal distance between the spectral components that can be accurately detected by the DSP estimator. Since the minimal distance is the distance between any two adjacent spectral components, viz. $\Delta\lambda_N = \frac{W_\lambda}{N}$, as N increases, the minimal distance decreases, and hence the resolution increases. We may come to the point of diminishing returns when going beyond a certain N. We call this $N_{\max}$. In addition, we define the minimum distance to be $\Delta\lambda_{\min} = \frac{W_\lambda}{N_{\max}}$. Thus, to find this "digital resolution," it is sufficient to find $N_{\max}$.

We obtain $N_{\max}$ using the g.MSE measure introduced in Section 4.1 as follows: We first fix the parameters N, M, K, G, σ^2, and a TF matrix, which are required to calculate the g.MSE. We then check if the condition g.MSE $\le \delta$ is met for a given percentage of the total $\binom{N}{K}$ possible support sets, where δ is an MSE specified by the user. If the condition is met, it indicates that the DSP estimator is able to detect all the spectral components. We then increase the value of N further and check the condition again. We keep repeating this process (that is, increasing the value of N and checking the condition) until the condition is no longer met. We finally declare the largest value of N that met the condition as $N_{\max}$. The minimum spacing and hence the maximum achievable resolution of the DSP-based spectrometer is then $\Delta\lambda_{\min} = \frac{W_\lambda}{N_{\max}}$. In this paper, we increase the value of N starting from M that corresponds to the resolution limit $\Delta\lambda_M = \frac{W_\lambda}{M}$ of filter-array-based spectrometers. We note that it is not possible to increase N indefinitely with the hope of obtaining a minimum spacing. There is a limit for increasing N. This limit depends only on the TF matrix for the fixed parameters M, K, G, and σ^2. Since the g.MSE captures all these parameters, it is enough to define the resolution based on the g.MSE measure.

For a signal spectrum with K spectral components, we define the maximum resolution ($\Delta\lambda_{\min}$) of the spectrometer as follows:

$$N_{\max} := \max\left\{N \in \{M, M+1, \cdots\} \ : \ \Pr\{\text{g.MSE} \le \delta\} \ge \rho\right\}$$
$$\Delta\lambda_{\min} := \frac{W_\lambda}{N_{\max}} \tag{10}$$

where the user-defined MSE δ is specific to an application. It specifies the minimum MSE that should be guaranteed by the DSP algorithm. The notation $\Pr\{\text{g.MSE} \le \delta\}$

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

denotes the probability of the event $\text{g.MSE} \le \delta$. The sample space for the event is the set of all possible support sets. Each support set is mapped into a g.MSE value. The variable ρ, $0 \le \rho \le 1$, denotes the probability of the event $\text{g.MSE} \le \delta$. That is, the variable ρ represents the fraction of the total support sets that satisfy the event. We express ρ in terms of percentage as $100 \times \rho\%$. If $\rho = 1$, then 100%; that is, all the $\binom{N}{K}$ support sets should satisfy the event. If $\rho = 0.95$, then 95% of the total number of support sets should satisfy the event. We say that the K spectral components are resolvable if we can find an N such that the probability of the event $\text{g.MSE} \le \delta$ is greater than or equal to a specified ρ. We note that $\Delta\lambda_{\min}$ in Eq. (10) is the resolution of the DSP-based spectrometer whereas $\Delta\lambda_M = \frac{W_\lambda}{M}$ is the limit on the resolution of filter-array spectrometers.

The definition of the resolution given by Eq. (10) is general in nature in that it can include prior information. For example, if we know beforehand that the signal spectral components occupy only a certain range of the spectrum, then this definition automatically incorporates this prior information. Since prior information reduces the number of support sets and therefore the change is only in the event $\text{g.MSE} \le \delta$, the definition of the resolution remains unaffected. Since we use g.MSE, the minimum possible MSE, the corresponding resolution obtained is the maximum possible by employing any kind of DSP-based estimator. In the next section, we evaluate the resolution that can be achieved by the correlated TF matrix in [5] and the proposed random TF matrix.

6. Results and discussion

In this section, we aim to investigate how much gain in terms of resolution improvement can be obtained by using random TFs in comparison with the correlated, non-ideal TFs in [5]. We first investigate the maximum possible resolution by using the definition in Eq. (10). We then use the L_1-based DSP algorithm reported in [5] and see how much of this resolution can be realized in practice. We set the number of filters in the array to $M = 40$. The non-ideal TFs, for the 40 filters, have been obtained from [7], which are also used in [5]. The random TFs are generated using thin-film filters with $m = 4$ layers. We used the following refractive indices: $n_a = 1$, $n_s = 3$, $n_H = 4$, and $n_L = 1.2$ (Section 4.1). The angle of incidence of the light at the input medium is 45°. The center wavelength $\lambda_{\text{normal}} = 850$ nm. The thicknesses of the four layers are chosen as explained in Section 4.1, with a different thickness for each filter. After we generate the random TFs, we sample them, and arrange them in rows to form a random TF matrix.

We choose a Gaussian kernel with an FWHM of 1 nm to generate the kernel matrix G. The operating wavelength range of the spectrometer is from 300 nm to 700 nm, corresponding to a W_λ of 400 nm. We vary N in order to find the maximum possible resolution $\Delta\lambda_{\min}$ as per Eq. (10). The value of N begins with M (which corresponds to the resolution limit) and is set to increase in steps of 40, i.e., N = 40, 80, 120, etc. For $N = 40$, the minimum spacing between two adjacent spectral components is $\Delta\lambda_N = \frac{400\,\text{nm}}{40}$ = 10 nm (from Section 3), which is the resolution limit (maximum possible resolution without using a DSP estimator). We define the signal-to-noise ratio (SNR) for the data model in Eq. (4) as $\text{SNR} = \frac{\|DGs\|^2}{M\sigma^2}$. We keep the SNR at 30 dB.

We consider a sparse signal spectrum with K = 5 spectral (wavelength) components. The locations of these spectral components within the spectrum are random. Thus, the total number of possible locations (support sets) is $\binom{N}{5}$ for a given N. We choose 50,000 support sets at random from this total population. For each of these support sets, we calculate the g.MSE as per Eq. (9). We set $\rho = 0.95$. As we vary N, we note the value of the g.MSE below which the g.MSEs of 95% of the support sets lie. Figure 5 shows the

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

plot of the 95th-percentile MSE versus N obtained for the TF matrix in [5] and for the random TF matrix.

We now set $\delta = -5$ dB and find $N_{\max}$. That is, we aim to find the maximum N such that the 95th-percentile MSE is less than -5 dB . We observe from Fig. 5 that for the TF matrix in [5], the 95-percentile MSE is less than -5 dB when $N_{\max}$ = 62. Thus, the maximum resolution of the DSP-based spectrometer obtained by using the TF matrix in [5] is $\Delta\lambda_{\min} = \frac{W_\lambda}{N_{\max}} = \frac{400\text{ nm}}{62} = 6.5\text{ nm}$. Therefore, the improvement in resolution over the resolution limit is 1.54 (10 nm /6.5 nm). In the case of random TFs, $N_{\max}$ = 405 and the maximum resolution obtained is $\Delta\lambda_{\min} = \frac{400\text{ nm}}{405} = 0.99\text{ nm}$. The improvement in resolution in this case is 10.1 (10 nm/0.99 nm). We observe that the random TFs provide a resolution improvement of approximately a factor of 7 over that of the TFs in [5].

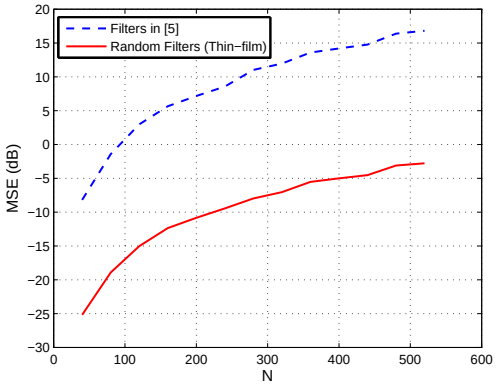


Fig. 5. g.MSE against resolution (N) for K = 5 and 50,000 support sets with $\rho = 0.95$.

Table 3 summarizes the resolution obtained by using various design approaches and algorithms for the same filter and signal settings as before. From Table 3, we note that the proposed random TFs achieve 7-fold resolution improvement compared to the TFs in [5] and 10-fold resolution improvement when compared to the ideal TFs.

We have also considered the non-Gaussian kernels such as Lorentzians and secant hyperbolic (mentioned in Section 2) for the same FWHM. Here, we report that we have obtained a similar resolution improvement to those obtained by using the Gaussian kernel. In addition, we note that comparing the resolution obtained by using ideal TFs with that of the random TFs is equivalent to comparing the resolution of ideal grating-based spectrometers with the random filter-array-based spectrometer. This is because, as discussed in Section 3.1, the resolution limit of both of the spectrometers is the same. Thus, we are in one sense indirectly comparing our proposed approach with the best possible conventional grating-based spectrometers.

Table 3. Summary of Filter Design Approaches and Corresponding Resolution

	Methods	Design approach	Resolution	Recovery method
1	Conventional	Ideal brick-wall	10 nm	No DSP
2	Correlated TFs in [5]	Holistic design by accident	10 nm	Least-squares and adaptive regularization [7, 8]
3	Correlated TFs in [5]	Holistic design by accident	6.5 nm (a 1.5-fold improvement compared to the ideal brick wall filters)	L_1 DSP estimator [5]
4	Proposed Random TFs	Random design by purpose	0.99 nm (an improvement of 7-fold compared to the TFs in [5] and 10-fold compared to the ideal brick wall filters)	L_1 DSP estimator [5]

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

We now investigate the practical spectral estimate obtained by using the DSP estimator reported in [5] when both the non-ideal and the random TFs are used in the filter-array spectrometer. For this purpose, we consider a signal spectrum generated by a mercury arc lamp, commonly used in fluorescence microscopy applications. The original signal spectrum of the mercury arc lamp has five (K = 5) prominent spectral components located at 365 nm, 405 nm, 436 nm, 546 nm, and 579 nm [22]. We consider N = 360 and M = 40. Figure 6 shows the original signal spectrum of the mercury arc lamp and an estimate of the signal spectrum obtained by using the correlated TF matrix in [5]. Figure 7 shows the estimated signal spectrum obtained by using the proposed random TF matrix.

From Figs. 6 and 7, we observe the following: First, it is evident that the estimate obtained by using the random TFs is closer to the original signal spectrum than that of the estimate using the non-ideal, correlated TFs. Second, the five spectral components of the original signal spectrum are determined correctly by the random TFs. However, the non-ideal TFs miss all five of the spectral components. Both the non-ideal and random TFs detect additional spectral components that are not present in the original signal spectrum. These additional spectral components appear as small bumps in Figs. 6 and 7.

In order to investigate how much resolution can be obtained using the DSP algorithm, we computed the MSE of the spectral estimates shown in Figs. 6 and 7 and compared them with the theoretical MSE shown in Fig. 5. The MSE of the spectral estimate by the random TFs (in Fig. 7) is about –5.5 dB which is just 1.3 dB away from the theoretical 95th-percentile MSE of –6.8 dB (in Fig. 5) at N = 360. For the non-ideal filters, the calculated MSE of the spectral estimate is 14 dB, which is 1.1 dB away from the MSE (12.9 dB) shown in Fig. 5. Thus, we observe that the resolution improvement deduced from Fig. 5 is actually achievable by using the practical DSP algorithms. The difference in MSE between the theoretical and the practical estimates is because of the detection of additional spectral components by the DSP algorithm. The difference in the practical MSE between the non-ideal TF in [5] and the random TFs using the DSP algorithm is about 18 dB.

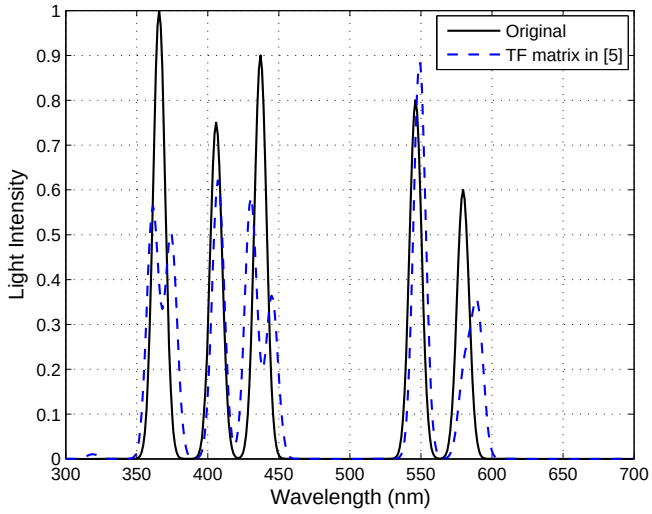


Fig. 6. Reconstruction of a mercury arc signal spectrum by a non-ideal TF matrix in [5].

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

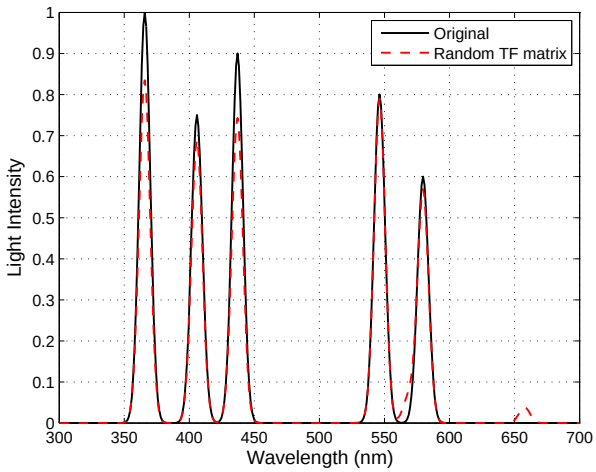


Fig. 7. Reconstruction of a mercury arc signal spectrum by the proposed random TF matrix.

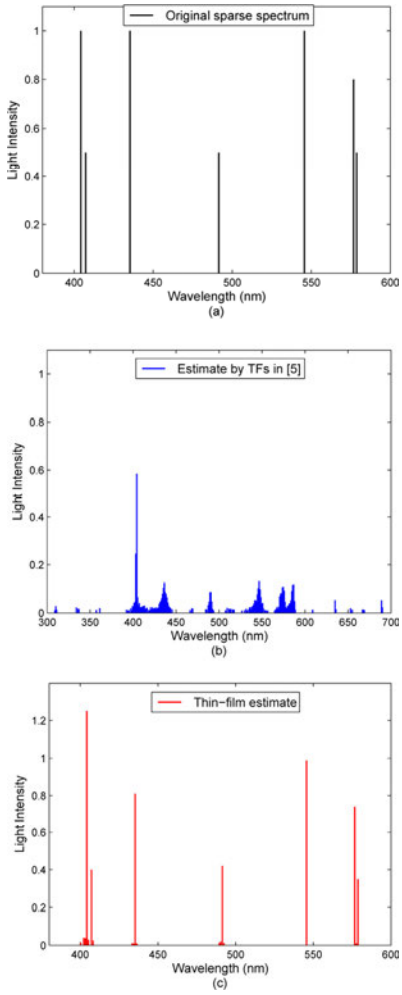


Fig. 8. (a) Original sparse spectrum of the mercury lamp. (b) Estimated sparse spectrum by using TFs in [5]. (c) Estimated sparse spectrum by thin-film-based random TFs.

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

In order to show that a resolution of 0.99 nm (see Table 3) is achievable by using random TFs, we consider an original mercury lamp spectrum [28] with prominent and weak spectral lines as follows: 404.656 nm, 407.781 nm, 435.835 nm, 491.604 nm, 546.074 nm, and a pair at 576.959 nm and 579.065 nm. Thus, the sparse spectrum of the mercury lamp contains seven spectral lines or components as shown in Fig. 8(a). We note that the least separation among these spectral lines is 2.106 nm, which is between the pair 576.959 nm and 579.065 nm. We aim to resolve these components in the sparse domain. We consider the case of $N = 400$. Figures 8(b) and 8(c) show the estimate of the sparse spectrum using the TFs in [5] and using the proposed thin-film-based random TFs, respectively. It is apparent from Fig. 8(c) that random TFs correctly estimate the seven dominant wavelength components in the sparse spectrum. In addition, random TFs clearly resolve the pair of two closely spaced spectral components at 576.959 nm and 579.065 nm. This evidence supports the fact shown in Table 3 that random TFs are capable of resolving any two spectral components that are more than 0.99 nm apart. In contrast, the estimate obtained by using the TFs in [5] misses all seven of the dominant spectral components as shown in Fig. 8(b). In addition, the estimate contains spectral components that were not originally present in the signal spectrum. Based on these results and observations, we conclude that the filters with random TFs outperform the non-ideal TFs in terms of resolution and reconstruction performance.

7. Summary and conclusions

In this paper, we have proposed random-transmittance-based filters to improve the resolution of a spectrometer. We have shown that the optical filters with random TFs are designed to acquire *holistic* information about the signal spectrum, rather than the *localized* information that was the target of traditional designs. The holistic information captured by each filter when put together at the spectral estimator provides multiple independent views and helps the DSP algorithm to differentiate the details in the spectral components. We have shown via examples that spectrometers with random TFs provide 7-fold resolution improvement when compared to the use of filters that aim at implementing the ideal TFs. Rather than trying to design filters with the ideal brick wall TF, which has been the design paradigm in the past, our study shows that designing filters with random TFs should be the design paradigm for modern DSP-based spectrometers. This shift in the design paradigm not only brings forth an order-of-magnitude improvement in terms of resolution and MSE but also may relax the difficulty of the filter engineering process, resulting in less stringent design requirements.

Acknowledgments

The authors would like to thank the anonymous reviewers for their thorough review of the manuscript, for useful references, and for their role in improving the original manuscript. This work was supported by the National Research Foundation of Korea (NRF) grant [redacted] 2012-0005656).

#177709 - $15.00 USD Received 11 Oct 2012; revised 7 Dec 2012; accepted 13 Jan 2013; published 11 Feb 2013
(C) 2013 OSA

No	구분	발명의 명칭	출원의명칭	출원국	권리구분	연구책임자	등록번호	등록일	출원번호	출원일	패밀리국가
1	영지식증명EVM	스택기반 가상머신 프로그램을 위한 효율적인 유니버셜 영지식 증명 시스템	스택 머신 프로그램을 위한 범용 영 지식 스나크 증명 시스템 및 방법	미국	특허	이흥노			18/070,497	2022-11-29	US, KR
2	블록체인	블록체인 안전거래 속도 최적화 알고리즘	블록체인의 거래검증시스템, 및 블록체인의 거래검증방법	미국	특허	이흥노			17/611,506	2021-11-15	US, WO, KR
3	영지식증명투표	검표자가 필요 없는 영지식 증명과 스마트 컨트랙트 기반 블록체인 투표 시스	블록체인 전자투표시스템, 그 시스템의 운용방법	미국	특허	이흥노			17/517,563	2021-11-02	US, KR
4	영지식증명투표	검표자가 필요 없는 영지식 증명과 스마트 컨트랙트 기반 블록체인 투표 시스	검표자가 필요 없는 영지식 증명과 스마트 컨트랙트 기반 블록체인 투표 시스템	미국	특허	이흥노			63/112723	2020-11-12	US, WO, KR
5	ECC-POW	새로운 블록체인 및 암호화폐 작업증명 생성, 증명, 검증 시스템	새로운 블록체인 및 암호화폐 작업증명 생성, 증명, 검증 시스템	미국	특허	이흥노			62/699,752	2018-07-18	US
6	ECC-POW	새로운 블록체인 및 암호화폐 작업증명 생성, 증명, 검증 시스템	새로운 블록체인 및 암호화폐 작업증명 생성, 증명, 검증 시스템	미국	특허	이흥노			62/699751	2018-07-18	US, KR
1	영지식증명EVM	스택기반 가상머신 프로그램을 위한 효율적인 유니버셜 영지식 증명 시스템	스택 머신 프로그램을 위한 범용 영 지식 스나크 증명 시스템 및 방법	대한민국	특허	이흥노			2022-0018393	2022-02-11	US, KR
2	유일ID합의블록체인	유일 ID 기반 블록체인 거버넌스 및 합의 메카니즘	채굴 장치	대한민국	특허	이흥노			2021-7033618	2021-10-18	WO, KR
3	영지식증명투표	검표자가 필요 없는 영지식 증명과 스마트 컨트랙트 기반 블록체인 투표 시스	검표자가 필요 없는 영지식 증명과 스마트 컨트랙트 기반 블록체인 투표 시스템	대한민국	특허	이흥노			2020-0146369	2020-11-04	US, WO, KR
4	ECC-POW난이도조절	Fine Control Apparatus of ECCPoW Difficulty	Scalable DeSecure ECCPoW Blockchains 난이도 조절 Algorithm	대한민국	특허	이흥노			2019-0099153	2019-08-13	KR
5	유일ID합의블록체인	유일 ID 기반 블록체인 거버넌스 및 합의 메카니즘	블록체인거버넌스	대한민국	특허	이흥노			2019-0084800	2019-07-12	WO, KR
6	블록체인안전거래	블록체인 안전거래 속도 최적화 알고리즘	블럭체인의 거래 컨펌횟수결정시스템	대한민국	특허	이흥노			2019-0061493	2019-05-24	KR
7	MAC-ZKP원천기술	Spatially Concatenated Channel-Network Code for Underwater Wireless Sen	센싱 데이터 처리장치 및 데이터 처리방법	대한민국	특허	이흥노			2016-0066621	2016-05-30	WO, KR
8	무선통신	무선 다중 접속망 관리 시스템 및 방법 Control Method and System for wirele	무선 다중 접속망 관리시스템 및 방법	대한민국	특허	이흥노			2013-0151397	2013-12-06	KR
9	무선통신	무선 다중 접속 망에서 노드 감시 시스템 및 방법 Node Observation Method	무선 다중 접속망을 위한 노드 감시 시스템 및 방법	대한민국	특허	이흥노			2013-0151395	2013-12-06	KR
10	뇌파BCI	뇌파(EEG) 신호의 sparse 표현을 통한 움직임 상상 기반의 BCI 분류법	뇌파(ＥＥＧ) 신호의 ｓｐａｒｓｅ표현을 통한 움직임 상상 기반의 ＢＣＩ 분류법	대한민국	특허	이흥노			2011-0043567	2011-05-09	KR
1	유일ID합의	유일 ID 기반 블록체인 거버넌스 및 합의 메카니즘	채굴 장치	PCT	특허	이흥노			PCT/KR2020/009222	2020-07-13	WO, KR
2	블록체인안전거래	블록체인 안전거래 속도 최적화 알고리즘	블록체인의 거래검증시스템, 및 블록체인이 거래검증방법	PCT	특허	이흥노			PCT/KR2019/017571	2019-12-12	US, WO, KR
3	ECC-POW	새로운 블록체인 및 암호화폐 작업증명 생성, 증명, 검증 시스템	새로운 블록체인 및 암호화폐 작업증명 생성, 증명, 검증 시스템	PCT	특허	이흥노			PCT/KR2019/008904	2019-07-18	WO, KR
4	MAC-ZKP원천	Spatially Concatenated Channel-Network Code for Underwater Wireless Sen	센싱 데이터 처리장치 및 데이터 처리방법	PCT	특허	이흥노			PCT/KR2017/005653	2017-05-30	WO, KR
5	지능형카메라	곤충눈 구조의 고해상도 이미징 장치	다수의 렌즈를 이용한 촬상장치	PCT	특허	이흥노			PCT/KR2016/010547	2016-09-21	WO
6	초분광기원천	Apparatus for Improving Spectral Resolution using Random Transmittances	분광장치 및 분광방법	PCT	특허	이흥노			PCT/KR2013/009014	2013-10-08	WO
7	ECC-POW	유한체에서 압축 센싱 복호를 위한 메시지 전달 알고리즘	유한체의 희소신호 복구방법, 유한체의 희소신호 복구장치, 및 이 방법을 기록되는 기	PCT	특허	이흥노			PCT/KR2013/004875	2013-06-03	WO
8	뇌파BCI	뇌파(EEG) 신호의 sparse 표현을 통한 움직임 상상 기반의 BCI 분류법	뇌-컴퓨터 접속 장치, 그리고 그의 분류 방법	PCT	특허	이흥노			PCT/KR2012/003572	2012-05-07	WO

No	구분	발명의 명칭	출원의명칭	출원국	권리구분	연구책임자	등록번호	등록일	출원번호	출원일	패밀리국가
1	지능형카메라	곤충눈 구조의 고해상도 이미징 장치	다수의 렌즈를 이용한 촬상장치	미국	특허	이흥노	10,605,962	2020-03-31	15/761,855	2018-03-21	US
2	지능형카메라	디지털신호처리 기반 곤충눈 구조의 이미징 장치의 거리 감지 기법	촬상장치, 촬상방법, 거리측정장치, 및 거리측정방법	미국	특허	이흥노	10,484,585	2019-11-19	15/669554	2017-08-04	US, KR
3	초분광기	Apparatus for Improving Spectral Resolution using Random Transmittances	SPECTROMETRY APPARATUS AND SPECTROMETRY METHOD	미국	특허	이흥노	10,458,843	2019-10-29	14/434757	2015-04-09	US
4	초분광기	불규칙 매체를 통한 이미지 전송 및 복구 방법 (Image transmission and reco	MICROSCOPE	미국	특허	이흥노	10,082,659	2018-09-25	15/166,278	2016-05-27	US
5		불규칙 매체로 입력단이 코팅된 다중모드 광섬유를 통한 이미지 전송 및 복구	CROSS REFERENCE TO RELATED APPLICATION	미국	특허	이흥노	10,080,485	2018-09-25	15/166,276	2016-05-27	US
6	ECC-POW	유한체에서 압축 센싱 복호를 위한 메시지 전달 알고리즘	METHOD FOR RECONSTRUCTING SPARSE SIGNAL IN FINITE FIELD, APPARAT	미국	특허	이흥노	9,734,128	2017-08-15	14/405234	2014-12-03	US
7	VCA(검증가능연산알고	압축 센싱을 위한 신뢰 전파 기반의 희소 신호 복구 알고리즘	METHOD AND APPARATUS FOR TRANSMITTING SPARSE SIGNAL, AND METHO	미국	특허	이흥노	9,160,398	2015-10-13	13/420176	2012-03-14	US
8	지능형센서	Apparatus for Improving Spectral Resolution Beyond the Conventional Sens	METHOD AND APPARATUS FOR PROCESSING OPTICAL SIGNAL OF SPECTRO	미국	특허	이흥노	9,030,662	2015-05-12	13/711930	2012-12-12	US
9	MIMO 복호	다중안테나 시스템 복구를 위한 초월평면수치 열거법기반 초월평면 구형 검	ORTHOTOPE SPHERE DECODING METHOD AND APPARATUS FOR SIGNAL RE	미국	특허	이흥노	8983006	2015-03-17	13/970774	2013-08-20	US
10	MIMO 복호	다중안테나 시스템 신호 복구를 위한 복잡도 예상 초월평면 구형검출기	ORTHOTOPE SPHERE DECODING METHOD AND APPARATUS FOR SIGNAL RE	미국	특허	이흥노	8,798,209	2014-08-05	13/670025	2012-11-06	US
11	VCA검증가능연산알고	다중센서 시스템에 대한 압축센싱 방법 및 기구	Signal Acquisitio and Method for Distributed Compressive Sensing and Joint Signal R	미국	특허	이흥노	8,391,800	2013-03-05	13/250,082	2011-09-30	US
1	초분광기	Apparatus for Improving Spectral Resolution using Random Transmittances	분광장치 및 분광방법	일본	특허	이흥노	6290905	2018-02-16	2015-536712	2015-04-10	JP
2	ECC-POW	유한체에서 압축 센싱 복호를 위한 메시지 전달 알고리즘	유한체의 희소신호 복구방법, 유한체의 희소신호 복구장치, 및 이 방법을 기록되는 기	일본	특허	이흥노	5914755	2016-04-08	2015-514928	2014-11-27	JP
1	초음파이미징	무작위 간섭을 이용한 초음파 이미징 기술 (Ultrasound Imaging Using Rande	무작위 간섭을 이용한 초음파 이미징 장치 및 그 방법	대한민국	특허	이흥노	10-2550262	2023-06-27	2021-0023473	2021-02-22	KR
2	영지식증명투표	검표자가 필요 없는 영지식 증명과 스마트 컨트랙트 기반 블록체인 투표 시스	블록체인 전자투표시스템, 그 시스템의 운용방법	대한민국	특허	이흥노	10-2430835	2022-08-04	2020-0171424	2020-12-09	US, WO, KR
3	전파탐지	전파 신호 송출원 식별 기술의 모델 확장을 위한 다중 레이블 (multi-label) 비	다중 레이블 아웃라이어 검출 방법 및 신호 송출원 식별 모델 확장 방법	대한민국	특허	이흥노	10-2415975	2022-06-28	2021-0097234	2021-07-23	KR
4	무선네트워크	A Cooperative Wireless Sensor Network for Indoor Industrial Monitoring	무선 센서 네트워크의 데이터 처리장치 및 데이터 처리 방법	대한민국	특허	이흥노	10-2412058	2022-06-17	2016-0066625	2016-05-30	KR
5	RF 핑거프린팅	상승 천이상태, 정상상태, 하강 천이상태 신호를 동시에 활용하는 앙상블 분류	앙상블 기반 무선 핑거프린팅 장치 및 이를 이용한 송출원 식별 방법	대한민국	특허	이흥노	10-2347174	2021-12-30	2021-0095830	2021-07-21	KR
6	ECC-POW 난이도미세	Fine Control Apparatus of ECCPoW Difficulty	부호-암호 화폐 시스템	대한민국	특허	이흥노	10-2288769	2021-08-05	2019-0151246	2019-11-22	KR
7	블록체인안전거래	블록체인 안전거래 속도 최적화 알고리즘	블록체인의 거래검증시스템, 및 블록체인이 거래검증방법	대한민국	특허	이흥노	10-2288776	2021-08-05	2019-0120655	2019-09-30	US, WO, KR
8	부호-암호화폐시스템	새로운 블록체인 및 암호화폐 작업증명 생성, 증명, 검증 시스템	부호-암호 화폐 시스템	대한민국	특허	이흥노	10-2231257	2021-03-17	2018-0097677	2018-08-21	US, WO, KR
9	초분광이미징	디지털신호처리 기반 초분광 이미징 장치	하이퍼스펙트럼 이미지 장치	대한민국	특허	이흥노	10-1986998	2019-06-03	2017-0077480	2017-06-19	KR
10	난수발생기	랜덤 나노 파티클을 이용한 난수 발생기	물리적 복제방지 장치 및 이를 이용한 난수 생성 방법	대한민국	특허	공득조	10-1975106	2019-04-26	2017-0101207	2017-08-09	KR
11	MIMO 복호	다중안테나 시스템 복구를 위한 초월평면수치 열거법기반 초월평면 구형 검	다중 안테나 시스템의 신호 복구를 위한 초월 평면 스피어 디코딩 방법 및 이를 위한 장	대한민국	특허	이흥노	10-1959039	2019-03-11	2012-0091394	2012-08-21	KR
12	생각스펠러	Thought Speller	BCI 시스템에 사용되는 스마트 키보드 및 이의 입력 방법	대한민국	특허	이흥노	10-1959049	2019-03-11	2017-0081462	2017-06-27	KR
13	초분광기	디지털 신호처리를 이용한 필터 기반 분광기의 각도 추정 기법	분광장치 및 분광방법	대한민국	특허	이흥노	10-2030735	2019-10-02	2018-0010046	2018-01-26	KR
14	지능형카메라	디지털신호처리 기반 곤충눈 구조의 이미징 장치의 거리 감지 기법	촬상장치, 촬상방법, 거리측정장치, 및 거리측정방법	대한민국	특허	이흥노	10-1865126	2018-05-31	2016-0099298	2016-08-04	US, KR
15	초분광기	Apparatus for Improving Spectral Resolution using Random Transmittances	분광장치 및 분광방법	대한민국	특허	이흥노	10-1854815	2018-04-27	2012-0112453	2012-10-10	KR
16	현미경	불규칙 매체를 통한 이미지 전송 및 복구 방법 (Image transmission and reco	현미경	대한민국	특허	이흥노	10-1766328	2017-08-02	2015-0075286	2015-05-28	KR
17	곤충눈카메라	곤충눈 구조의 고해상도 이미징 장치	다수의 렌즈를 이용한 촬상장치	대한민국	특허	이흥노	1692428	2016-12-28	2016-0050015	2016-04-25	KR
18	곤충눈카메라	곤충눈 구조의 고해상도 이미징 장치	다수의 렌즈를 이용한 촬상장치	대한민국	특허	이흥노	1638022	2016-07-04	2015-0132810	2015-09-21	KR
19	곤충눈내시경	불규칙 매체로 입력단이 코팅된 다중모드 광섬유를 통한 이미지 전송 및 복구	내시경	대한민국	특허	이흥노	1638016	2016-07-04	2015-0075285	2015-05-28	KR
20	MIMO 복호	미모 시스템의 신호 복구를 위한 스피어 디코딩 방법 및 그 시스템	미모 시스템의 신호 복구를 위한 스피어 디코딩 방법 및 그 시스템	대한민국	특허	이흥노	1499448	2015-03-02	2013-0151399	2013-12-06	KR
21	MIMO 복호	미모 시스템의 신호 복구를 위한 스피어 디코딩 방법 및 그 시스템	미모 시스템의 신호 복구를 위한 스피어 디코딩 방법 및 그 시스템	대한민국	특허	이흥노	1498267	2015-02-25	2013-0151398	2013-12-06	KR
22	초분광기	램덤 광학 필터의 구현 및 투과율 검출 (Apparatus and methods for implemen	랜덤필터모듈의 투과율 검출방법	대한민국	특허	이흥노	1572080	2015-11-20	2015-0029193	2015-03-02	KR
23	초분광기	램덤 광학 필터의 구현 및 투과율 검출 (Apparatus and methods for implemen	랜덤필터모듈, 랜덤필터모듈의 투과율 검출방법, 및 랜덤필터모듈을 이용하는 분광기	대한민국	특허	이흥노	1526870	2015-06-02	2014-0052394	2014-04-30	KR
24	MIMO 복호	다중안테나 시스템 신호 복구를 위한 복잡도 예상 초월평면 구형검출기	다중 안테나 시스템의 신호 복구를 위한 초월 평면 스피어 디코딩 방법 및 이를 위한 장	대한민국	특허	이흥노	1423965	2014-07-22	2012-0093627	2012-08-27	KR
25	초분광기	Apparatus for Improving Spectral Resolution Beyond the Conventional Sens	신호의 희소 특성을 이용한 분광계의 광 신호 처리 방법 및 그 장치	대한민국	특허	이흥노	1423964	2014-07-22	2012-0079171	2012-07-20	KR
26	뇌파분석BCI	뇌파(EEG) 신호의 sparse 표현을 통한 움직임 상상 기반의 BCI 분류법	뇌-컴퓨터 접속 장치, 그리고 그의 분류 방법	대한민국	특허	이흥노	1380964	2014-03-27	2012-0039497	2012-04-17	KR
27	통신	A method for using sparse tones as a means to achieve reliable communica	제한된 전력 범위의 선형 증폭기가 장비된 수신 장치에서의 안정적인 통신을 위해 희소	대한민국	특허	이흥노	1352618	2014-01-10	2012-0083577	2012-07-31	KR
28	ECC-POW	유한체에서 압축 센싱 복호를 위한 메시지 전달 알고리즘	유한체의 희소 신호 복구 방법 및 장치	대한민국	특허	이흥노	1284569	2013-07-04	2012-0060225	2012-06-05	KR
29	신호처리	임플랜트 디바이스에 적용 가능한 상관관계 있는 데이터의 저전력 무선 전송	상관관계 있는 신호의 전송 방법과 이를 구현한 송신기, 그리고 상관관계 있는 신호의	대한민국	특허	이흥노	1270238	2013-05-27	2011-0127905	2011-12-01	KR

30	VCA	압축 센싱을 위한 신뢰 전파 기반의 희소 신호 복구 알고리즘	희소 신호 전송 방법 및 장치, 그리고 희소 신호 복구 방법 및 장치	대한민국	특허	이흥노	1209908	2012-12-03	2011-0077836	2011-08-04	KR
31	통신	다중 패킷 수신 환경에서의 다중 접근 통신을 위한 전송파 위상 조절 방법	다중 패킷 수신 환경에서의 다중 접근 통신을 위한 전송파 위상 조절 장치 및 방법	대한민국	특허	임혁	1117791	2012-02-10	2010-0101450	2010-10-18	KR
32	VCA	다중센서 시스템에 대한 압축센싱 방법 및 기구	분산적 압축 센싱 및 협력 복구를 수행하는 신호 취득 장치 및 그 방법	대한민국	특허	이흥노	1112746	2012-01-30	2010-0096128	2010-10-01	KR

NO	과제상태	연구계정	과제명	직급	연구책임자	연구부서	연구기간 해당연도연구시작일	연구기간 해당연도연구종료일	연구비 총연구시작일	연구비 총연구종료일	연구비 최종협약금액	연구비 현금	연구비 현물	연구비 총연구비	지원기관	연구사업	소관부처
1	진행	GK01270	연구보전비	교수	이흥노	정보통신융합연구센터								2370362	광주과학기술원	자체(정부	광주과학기술원
2	완료	NN05850	부호 이론적 다중 압축 센싱 시스템 개발		이흥노	정보통신공학부								0	한국연구재단	국가연구	교육과학기술부
3	완료	NN06790	네트워크 코딩 기반의 무선자원 최적화 연구(1/3)		이흥노	정보통신공학부								0	한국연구재단	국가연구	교육과학기술부
4	완료	GE03360	Thought Decoding을 위한 신호처리 방법 연구		이흥노	정보통신공학부								0	광주과학기술원	자체(정부	광주과학기술원
5	중단	GP26360	잔고계정		이흥노	전기전자컴퓨터공학부								0	광주과학기술원	자체(정부	광주과학기술원
6	진행	GS60673	인건비풀링제-이흥노		이흥노	정보기전공학부								0	광주과학기술원	국가연구	광주과학기술원
7	완료	NN07050	부호 이론적 다중 압축 센싱 시스템 개발		이흥노	정보통신공학부								0	한국연구재단	국가연구	교육과학기술부
8	완료	NN08290	네트워크 코딩 기반의 무선자원 최적화 연구(2/3)		이흥노	정보통신공학부								0	한국연구재단	국가연구	교육과학기술부
9	완료	GE03770	뇌-컴퓨터 접속 인터페이스를 위한 특징신호 추출 알고리즘 연구		이흥노	정보통신공학부								0	광주과학기술원	자체(정부	광주과학기술원
10	완료	GG10010	방송대역 이 기종 TVBD 간 채널 점유 fairness에 최적화 된 자원 할당 알고리즘		이흥노	정보통신공학부								0	한국전자통신연구원	공공수탁	공공수탁
11	완료	GK0271A	무선 뇌 생각 디코더를 이용한 재활 공학 분야의 응용		이흥노	AI-의과학 연구센터								0	광주과학기술원	자체(정부	광주과학기술원
12	완료	NN08660	부호 이론적 다중 압축 센싱 시스템 개발		이흥노	정보통신공학부								0	한국연구재단	국가연구	교육과학기술부
13	완료	NN09960	네트워크 코딩 기반의 무선자원 최적화 연구(3/3)		이흥노	정보통신공학부								0	한국연구재단	국가연구	교육과학기술부
14	완료	GG10441	전자전특화연구센터:전파탐지 및 분석연구실(EW-11)		이흥노	정보통신공학부								320363720	국방과학연구소	공공수탁	공공수탁
15	완료	NN10360	부호 이론적 다중 압축 센싱 시스템 개발		이흥노	정보통신공학부								0	한국연구재단	국가연구	미래창조과학부
16	진행	NN12030	초강력 레이저과학 연구단		이흥노	연구정책팀								0	기초과학연구원	국가연구	미래창조과학부
17	완료	GK0403F	랜덤 투과율을 갖는 랜덤 필터를 이용한 고해상도 소형 분광 시스템		이흥노	과학기술혁신사업단								0	광주과학기술원	자체(정부	광주과학기술원
18	완료	NN12120	부호 이론적 다중 압축 센싱 시스템 개발		이흥노	정보통신공학부								0	한국연구재단	국가연구	미래창조과학부
19	완료	GK04627	4th GIST-Caltech Workshop on Innovative Research		이흥노	연구관리팀								25000000	광주과학기술원	자체(정부)	
20	진행	GK04628	2013, 2014 Caltech 공동연구비		이흥노	연구처								533957270	광주과학기술원	자체(정부	광주과학기술원
21	진행	GK04629	2013, 2014 GIST-CALTECH 국제공동연구		이흥노	연구처								290000000	광주과학기술원	자체(정부	광주과학기술원
22	완료	GK05660	첨단산업기술기초연구 (기타)		이흥노	연구처								22000000	광주과학기술원	자체(정부)	
23	진행	NN14150	부호 이론적 다중 압축 센싱 시스템 개발		이흥노	정보통신공학부								0	한국연구재단	국가연구	미래창조과학부
24	진행	NN14590	해상도 향상 DSP 알고리즘을 이용한 초소형 고해상도 분광기 모듈 개발		이흥노	정보통신공학부								0	피큐브(주)	국가연구	미래창조과학부
25	진행	GG11420	전라남도 청색기술 산업화 기본계획 연구용역	교수	이흥노	미래연구센터								43873000	전라남도	공공수탁	공공수탁
26	진행	GI05150	다중채널 디지털 수신기용 압축센싱 알고리즘 연구 및 RTL 로직 개	교수	이흥노	정보통신공학부								300000000	한화탈레스(주)	민간수탁	산업체
27	진행	GK05550	2016 Caltech 공동연구소 설치 운영	교수	이흥노	GIST연구원								1243518572	광주과학기술원	자체(정부	광주과학기술원
28	진행	GK05553	2015년 선정 Caltech 공동연구비(신규-3단계)	교수	이흥노	연구처								540000000	광주과학기술원	자체(정부	광주과학기술원
29	진행	GK055SC	5th GIST-Caltech Workshop on Innovative Research	교수	이흥노	연구관리팀								30000000	광주과학기술원	자체(정부	광주과학기술원
30	진행	GK05580	첨단산업기술기초연구	교수	이흥노	GIST연구원								748000000	광주과학기술원	자체(정부	광주과학기술원
31	진행	GK06310	차세대 뉴로네비게이션 소프트웨어-전기적 뇌자극술의 뇌졸중 치료	교수	이흥노	창업진흥센터								45000000	광주과학기술원	자체(정부	광주과학기술원
32	진행	GK06320	2016 기술사업화 플랫폼 고도화 미래 신산업창출 기획과제	교수	이흥노	기술사업화센터								500000000	광주과학기술원	자체(정부	광주과학기술원
33	진행	GK06323	디지털 신호 처리 기반 저비용 휴대용 초분광 영상 장치 개발	교수	이흥노	과학기술혁신사업단								50000000	광주과학기술원	자체(정부	광주과학기술원
34	진행	GK07050	GIST-Caltech 공동연구 과제 해외송금	교수	이흥노	GIST연구원								958481428	광주과학기술원	자체(정부	광주과학기술원
35	진행	GK07060	GIST-Caltech 공동연구 과제 해외송금(II)		이흥노	GIST연구원								225000000	광주과학기술원	자체(정부	광주과학기술원
36	진행	NN15860	부호 이론적 다중 압축 센싱 시스템 개발	교수	이흥노	전기전자컴퓨터공학부								837000000	한국연구재단	국가연구	미래창조과학부
37	진행	NN16060	해상도 향상 DSP 알고리즘을 이용한 초소형 고해상도 분광기 모듈 개발		이흥노	전기전자컴퓨터공학부								97000000	피큐브(주)	국가연구	미래창조과학부
38	진행	NN16242	다중생체정보 기반 치매 중증도 모니터링 기술 개발	교수	이흥노	최적치매관리기술연구								1768000000	한국연구재단	국가연구	정보통신부
39	진행	GG11760	분석체계 선행연구 용역	교수	이흥노	전기전자컴퓨터공학부								45000000	국군 제3707부대	공공수탁	국방부
40	진행	GK07430	GIST-Caltech 공동연구 과제 해외송금(나머지)		이흥노	GIST연구원								52000000	광주과학기술원	자체(정부	광주과학기술원
41	진행	GK07440	GIST-Caltech 공동연구 과제 해외송금(나머지) 2		이흥노	GIST연구원								150000000	광주과학기술원	자체(정부	광주과학기술원
42	진행	GR04620	과학기술기반 창업중심대학 추진 계획 수립	교수	이흥노	전기전자컴퓨터공학부								60000000	한국연구재단	국가연구	미래창조과학부
43	진행	GR04750	인공지능(AI) 중심 창업단지 조성	교수	이흥노	미래연구센터								5000000	(재)광주테크노파크	공공수탁	공공수탁
44	진행	NN17530	초강력 레이저과학 연구단	교수	이흥노	연구정책센터								0	기초과학연구원	국가연구	미래창조과학부
45	진행	NN17642	다중생체정보 기반 치매 중증도 모니터링 기술 개발	교수	이흥노	최적치매관리기술연구								1748000000	한국연구재단	국가연구	미래창조과학부
46	진행	NN18200	부호 이론적 다중 압축 센싱 시스템 개발		이흥노	전기전자컴퓨터공학부								837000000	한국연구재단	국가연구	미래창조과학부
47	진행	GK09830	공공서비스 (Police, 코인가치교환) 블록체인 시스템 설계 연구	교수	이흥노	전기전자컴퓨터공학부								50000000	광주과학기술원	자체(정부	광주과학기술원
48	진행	GK09900	차세대 체감형 탁구 (하키) 로봇 게임기 인공지능 기술 개발	교수	이흥노	전기전자컴퓨터공학부								80000000	광주과학기술원	자체(정부	광주과학기술원
49	진행	GRE0630	시험분석료 수입	교수	이흥노	정보통신융합연구센터								0	광주과학기술원	자체(정부	기타
50	진행	GRE0720	시험분석료 수입	교수	이흥노	정보통신융합연구센터								0	광주과학기술원	자체(정부	광주과학기술원
51	진행	NN19882	다중생체정보 기반 치매 중증도 모니터링 기술 개발	교수	이흥노	최적치매관리기술연구								1748000000	한국연구재단	국가연구	과학기술정보통신
52	진행	NN21320	레이다, 분광기, 뇌컴퓨터 인터페이스 시스템에 특화된 지능형 신호	교수	이흥노	센서지능화연구센터								838428000	한국연구재단	국가연구	과학기술정보통신
53	진행	GE10560	[한국과학기술원 이흥노]잡음통신기술을 적용한 적응형 다중모드	교수	이흥노	전기전자컴퓨터공학부								418619000	국방과학연구소	국가연구	국방부
54	진행	GK10390	공공서비스 (Police, 코인가치교환) 블록체인 시스템 설계 연구	교수	이흥노	블록체인인터넷경제								35000000	광주과학기술원	자체(정부	광주과학기술원
55	진행	GK11230	연구보전비	교수	이흥노	센서지능화연구센터								0	광주과학기술원	자체(정부	광주과학기술원
56	진행	GK11380	암호-부호 작업증명(PoW) 비트코인/이더리움 하드포크 1.0 개발	교수	이흥노	블록체인인터넷경제								100000000	광주과학기술원	자체(정부	광주과학기술원
57	진행	NN22242	다중생체정보 기반 치매 중증도 모니터링 기술 개발	교수	이흥노	최적치매관리기술연구								700000000	한국연구재단	국가연구	과학기술정보통신
58	진행	NN23400	레이다, 분광기, 뇌컴퓨터 인터페이스 시스템에 특화된 지능형 신호	교수	이흥노	센서지능화연구센터								838428000	한국연구재단	국가연구	과학기술정보통신
59	진행	NN24010	확장가능한 탈중앙화 보안성 ECCPoW 블록체인	교수	이흥노	전기전자컴퓨터공학부								1200000000	정보통신산업진흥원	국가연구	정보통신부
60	완료	GG12510	도약주파수 천이상태 신호기반 통신망 식별을 위한 RF Fingerprint	교수	이흥노	전기전자컴퓨터공학부								220000000	LIG넥스원(주)	공공수탁	국방부
61	진행	NN23970	3D 영상센서와 초소형 라이다(LiDAR)를 결합한 차세대 3D인식 복	교수	이흥노	전기전자컴퓨터공학부								1700000000	한국항공우주연구원	국가연구	과학기술정보통신
62	진행	NN25052	다중생체정보 기반 치매 중증도 모니터링 기술 개발	교수	이흥노	최적치매관리기술연구								700000000	한국연구재단	국가연구	과학기술정보통신
63	진행	NN25780	레이다, 분광기, 뇌컴퓨터 인터페이스 시스템에 특화된 지능형 신호	교수	이흥노	센서지능화연구센터								838428000	한국연구재단	국가연구	과학기술정보통신
64	진행	NN27150	일반 연산을 검증하고 트랜잭션 검증량과 저장 공간을 줄여주는 유	교수	이흥노	블록체인인터넷경제								300000000	정보통신산업진흥원	국가연구	과학기술정보통신
65	진행	GK14340	연구보전비	교수	이흥노	블록체인인터넷경제								0	광주과학기술원	자체(정부	광주과학기술원
66	진행	GK14520	대학ICT연구센터육성사업 연구관리지원비(연구보전비)	교수	이흥노	블록체인 지능 융합선								298233923	광주과학기술원	자체(정부	광주과학기술원
67	진행	GR06620	위치정보에 대한 블록체인 기술	교수	이흥노	전기전자컴퓨터공학부								11000000	주식회사 얍컴퍼니	민간수탁	산업체

68	진행	NN28050	일반 연산을 검증하고 트랜잭션 검증량과 저장 공간을 줄여주는 유니버셜 영지		이흥노	블록체인인터넷경제					300000000	정보통신산업진흥원	국가연구	과학기술정보통
69	진행	GR06760	2021년도 국가인적자원개발컨소시엄(K-Digital Platform)사업	교수	이흥노	인공지능연구소					1000000000	한국산업인력공단	국가연구	고용노동부
70	진행	NN27842	다중생체정보 기반 치매 중증도 모니터링 기술 개발	교수	이흥노	최적치매관리기술연구					758334000	한국연구재단	국가연구	과학기술정보통
71	진행	NN28260	전파신호식별 및 초분광이미징에 특화된 지능형 신호 복원 및 분류	교수	이흥노	센서지능화연구센터					887871000	한국연구재단	국가연구	과학기술정보통
72	진행	NN29950	대규모 노드를 위한 탈중앙화 합의체 구성 기술 개발	교수	이흥노	블록체인인터넷경제					500000000	한국전자통신연구원	국가연구	과학기술정보통
73	진행	NN30620	영지식 센싱, 암호인증, 블록체인 기반 클라우드 서비스 융합 기술	교수	이흥노	블록체인 지능 융합선					6800000000	정보통신기획평가원	국가연구	과학기술정보통
74	진행	NN30740	AI 융합대학 지원사업	교수	이흥노	인공지능연구소					4701680000	인공지능산업융합시	국가연구	과학기술정보통
75	진행	NN30860	영지식 센싱, 암호인증, 블록체인 기반 클라우드 서비스 융합 기술 개발		이흥노	블록체인 지능 융합선					6800000000	정보통신기획평가원	국가연구	과학기술정보통
76	진행	NN30930	대규모 노드를 위한 탈중앙화 합의체 구성 기술 개발		이흥노	전기전자컴퓨터공학부					500000000	한국전자통신연구원	국가연구	과학기술정보통
77	진행	NN31550	전파신호식별 및 초분광이미징에 특화된 지능형 신호 복원 및 분류	교수	이흥노	전기전자컴퓨터공학부					887871000	한국연구재단	국가연구	과학기술정보통
78	진행	NN33140	노지작물 드론촬영 데이터	교수	이흥노	인공지능연구소					150000000	한국정보화진흥원	국가연구	과학기술정보통
79	진행	GK16210	Pervasive Healthcare Monitoring System for the Elderly based on	교수	이흥노	정보통신융합연구센터					25000000	광주과학기술원	자체(정부	광주과학기술원
80	중단	GK16560	블록체인 및 클라우드 상에서 동작하는 연합학습 플랫폼 개발	교수	이흥노	블록체인 지능 융합선					40000000	광주과학기술원	자체(정부	광주과학기술원
82	진행	NN34720	영지식 센싱, 암호인증, 블록체인 기반 클라우드 서비스 융합 기술 개발		이흥노	블록체인 지능 융합선					6800000000	정보통신기획평가원	국가연구	과학기술정보통
83	진행	NN34960	대규모 노드를 위한 탈중앙화 합의체 구성 기술 개발		이흥노	전기전자컴퓨터공학부					500000000	한국전자통신연구원	국가연구	과학기술정보통
84	진행	NN36080	전파신호식별 및 초분광이미징에 특화된 지능형 신호 복원 및 분류 시스템 연구		이흥노	전기전자컴퓨터공학부					887871000	한국연구재단	국가연구	과학기술정보통

[redacted]

총연구비 **₩17,173,774,738**

붉은색 표시는 Worldland 블록체인 원천 기술 개발 관련 사업

₩5,660,233,923

No	*Payroll No	Name	School/Dept.	*Sequence	*Title	Name of Authors including	Name of Publication/Conference	Vol.(No.), Page (First-Last)	Published Date	*Type of Paper	Number of Author	*Type of Author	Number of Same Type Author	ImpactFactor	JCR Category	JCR Version	Rank with Recent Category	%.within Categories	ISBN/ISSN	Research Account	SCIE 등재여부	USE_YN
1	60673	이흥노	School of Electrical Engineering and Computer Science	1	Extraction of sources of tremor in hand movements of patients with movement disorders	R. Vinjamuri, D. Crammond, D. Kondziolka, Heung-No Lee, Zhi-Hong Mao	IEEE Trans. on Inf. Tech. in Biomedicine	12, 49-56	2009-01	International journal SCI	5	ETC		1.436	Bio. Med. Eng.		15/52	29.00	1089-7771	NSF		
2	60673	이흥노	School of Electrical Engineering and Computer Science	2	Information Capacity of the Thumb and the Index Finger in Communication	Zhi-Hong Mao, Heung-No Lee, Robert J. Sclabassi, Mingui Sun	IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING	56_5_1535-1545	2009-05	International journal SCI	4	ETC		2.496	ENGINEERING, BIOMEDICAL		15/52	29.00	0018-9294			
3	60673	이흥노	School of Electrical Engineering and Computer Science	3	Schnorr-Euchner Sphere Decoder with Statistical Pruning for MIMO Systems	Junil Ahn, Heung-No Lee, and K. Kim	Prof. 2009 International Workshop on Stellite and Space Communications		2009-09	International prominent proceeding	3	ETC		0.00				0.00				
4	60673	이흥노	School of Electrical Engineering and Computer Science	4	Ordering Aided Schnorr-Euchner Sphere Decoding with Statistical Pruning Based on IRA for MIMO Systmes	Junil Ahn, Heung-No Lee, and K. Kim	Proc. Of the 15th Asia-Pacific Conference on Communications		2009-10	International prominent proceeding	3	ETC		0.00				0.00				
5	60673	이흥노	School of Electrical Engineering and Computer Science	5	Power control for multiple access communication systems with multi-packet reception capability	Choi Wooyeol, D. Jung, Heung-No Lee, Hyuk Lim	Prof. IEEE Conference on Local Computer Networks		2009-10	International prominent proceeding	4	ETC		0.00				0.00				
6	60673	이흥노	School of Electrical Engineering and Computer Science	6	Correlation 과 CS를 이용한 MS문제 해결	박상준, 신영학, 이준호, 송현우, 강재욱, 김기선, 이흥노	정보통신 합동학술대회	113-116	2009-11	Domestic prominent proceeding	7	CA		0.00				0.00				
7	60673	이흥노	School of Electrical Engineering and Computer Science	7	MB iterative decoding algorithm on systematic LDSM codes: Performance evaluation	Cheng-Chun Chang, Zhi-Hong Mao, Heung-No Lee Cheng-Chun Chang, Zhi-Hong Mao, Heung-No Lee	SIGNAL PROCESSING	90_1_373-377	2010-01	International journal SCI	3	CA		1.135	ENGINEERING, ELECTRICAL & ELECTRONIC		107/246	43.00	0165-1684	140317		
8	60673	이흥노	School of Electrical Engineering and Computer Science	8	New Compressive Sensing Model for Multiple Sensor System	Junho Lee/ Sangjun Park and Heung-No Lee	제 23회 신호처리합동학술대회	pp58-61	2010-01	Domestic prominent proceeding	3	CA		0.00				0.00		N05850		
9	60673	이흥노	School of Electrical Engineering and Computer Science	9	Information-Theoretical Analysis for Multiple Sensor System based on Compressive Sensing	Junho Lee/ Sangjun Park and Heung-No Lee	제 23회 신호처리합동학술대회	pp55-57	2010-01	Domestic prominent proceeding	3	CA		0.00				0.00		N05850		
10	60673	이흥노	School of Electrical Engineering and Computer Science	10	General random coding bounds: AWGN channels to MIMO fading channels	Heung-No Lee, Jingqiao Zhang and Cheon Won Choi	Annals of Transplantation	65_1_87-99	2010-02	International journal SCI	3	FA/CA		0.969	SURGERY		101/167	0.00	1425-9524	K140317		
11	60673	이흥노	School of Electrical Engineering and Computer Science	11	Dimensionality Reduction in Control and Coordination of the Human Hand	Ramana Vinjamuri, Mingui Sun, Cheng-Chun Chang, Heung-No Lee, Robert J. Sclabassi and Zhi-Hong Mao	IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING	57_2_284-295	2010-02	International journal SCI	6	ETC		2.154	ENGINEERING, BIOMEDICAL		22/59	0.00	ISSN: 0018-9294	X		
12	60673	이흥노	School of Electrical Engineering and Computer Science	12	Energy-Efficient Utility Maximization for Wireless Networks with/without Multipath Routing	J. Zhang and Heung-No Lee	AEU-INTERNATIONAL JOURNAL OF ELECTRONICS AND COMMUNICATIONS	64_2_99-111	2010-02	International journal SCI	2	CA		0.508	TELECOMMUNICATIONS		45/77	75.00	1434-8411	140317		
13	60673	이흥노	School of Electrical Engineering and Computer Science	13	Reduced-complexity orthotope sphere decoding for multiple-input multiple-output antenna system	Hwanchol Jang/ Heung-No Lee and Nooshabadi	IEEE International Midwest Symposium on Circuits and Systems (MWSCAS	pp.660-663	2010-04	International prominent proceeding	3	CA		0.00				0.00				
14	60673	이흥노	School of Electrical Engineering and Computer Science	14	A simple Security Mechanism using Network Codes for Wireless Multip Access Network	Woong-Bi Lee and Heung-No Lee	대한전자공학회/ 하계종합학술대회	pp139-141	2010-06	Domestic prominent proceeding	2	CA		0.00				0.00				
15	60673	이흥노	School of Electrical Engineering and Computer Science	15	Performance analysis of MIMO System on LDPC codes for AWGN channels using Union-Bound	Jin-Taek Seong and Heung-No Lee	대한전자공학회/ 하계종합학술대회	pp. 999-1001	2010-06	Domestic prominent proceeding	2	CA		0.00				0.00				
16	60673	이흥노	School of Electrical Engineering and Computer Science	16	A Study On Communication in the Underwater Acuostic Channel Using OFDM	Su-Je Lee/ Hyeong-Won Jeon/ Le Tran Tien Khanh and Heung-No Lee	대한전자공학회/ 하계종합학술대회	pp.1015-1018	2010-06	Domestic prominent proceeding	4	CA		0.00				0.00				
17	60673	이흥노	School of Electrical Engineering and Computer Science	17	OFDM System Design Based on Underwater Acoustic Channel	Hyeong-Won Jeon/ Su-Je Lee/ Le Tran Tien Khanh and Heung-No Lee	대한전자공학회/ 하계종합학술대회	pp1015-1018	2010-06	Domestic prominent proceeding	4	CA		0.00				0.00				
18	60673	이흥노	School of Electrical Engineering and Computer Science	18	Analysis of Joint Typicality Based Information-theoretic Limits on Multiple Sensor System Based on Compressive Sensing	Sangjun Park/ Junho Lee and Heung-No Lee	대한전자공학회/ 하계종합학술대회	pp.1005-1007	2010-06	Domestic prominent proceeding	3	CA		0.00				0.00				
19	60673	이흥노	School of Electrical Engineering and Computer Science	19	Temporal Postural Synergies of the Hand in Rapid Grasping Tasks	Ramana Vinjamuri, Mingui Sun, Cheng-Chun Chang, Heung-No Lee, Robert J Sclabassi, Zhi-Hong Mao	IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE	14_4_986-994	2010-07	International journal SCI	6	CA		1.694	COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS		27/95	28.00	1089- 7771			

20	60673	이흥노	School of Electrical Engineering and Computer Science	20	Carrier phase adjustment for multiple access communication systems with multi-packet reception	Wooyeol Choi/ Taewoon Kim/ Heung-No Lee and Hyuk Lim	IEEE Wireless Communications and Networking Conference (WCNC 2010)/	pp.1-6	2010-08	International prominent proceeding	4	ETC	0.00			0.00				
21	60673	이흥노	School of Electrical Engineering and Computer Science	21	Per-Sensor Measurements Behavior of Compressive Sensing System for Multiple Measurements	Sangjun Park/ JunhoLee/ Heung-No Lee	44th Annual Asilomar Conference on Signals/ Systems and Computers		2010-11	International prominent proceeding	3	CA	0.00			0.00				N05850
22	60673	이흥노	School of Electrical Engineering and Computer Science	22	Throughput behavior of link adaptive 802.11 DCF with MUD capable access node	Mir Hamza Mahmooda, Cheng-Chun Changc, Daewon Jungb, Zhi-Hong Maod, Hyuk Limb, and Heung-No Leeb	AEU-INTERNATIONAL JOURNAL OF ELECTRONICS AND COMMUNICATIONS	64_11_1031-1041	2010-11	International journal SCI	6	CA	0.508	TELECOMMUNICATIONS	45/77	0.00	1434- 8411			K140317
23	60673	이흥노	School of Electrical Engineering and Computer Science	23	A fast BER evaluation method for LDGM codes	Cheng-Chun Changa, Zhi-Hong Maob, Heung-No Lee	JOURNAL OF THE FRANKLIN INSTITUTE-ENGINEERING AND APPLIED MATHEMATICS	347_7_1368-1373	2010-12	International journal SCI	3	CA	1.130	ENGINEERING, MULTIDISCIPLINARY	25/79	0.00	0016- 0032			F00810
24	60673	이흥노	School of Electrical Engineering and Computer Science	24	INTELLIGENT IMPLEMENTATION OF SCHNORR-EUCHNOR SPHERE DECODING FOR MIMO SYSTEMS	Junil Ahn/ Donghun Lee/ Heung-No Lee/ Kiseon Kim	ITC-CSCC 2010	pp.700-701	2010-12	International prominent proceeding	4	ETC	0.00			0.00				
25	60673	이흥노	School of Electrical Engineering and Computer Science	25	A Study on the Dual-mode Antenna Selection for Spatial Multiplexing Systems with Linear Receivers	Woong-Bi Lee and Heung-No Lee	한국통신학회/ 춘계종합학술발표회	pp.139-141	2010-12	Domestic prominent proceeding	2	CA	0.00			0.00				
26	60673	이흥노	School of Electrical Engineering and Computer Science	26	Comparison of Selection-Based Cooperation techniques for Cooperative Networks	Jaewook Kang/ Donghun Lee/ Heung-No Lee and Kiseon Kim	한국통신학회/ 춘계종합학술발표회	pp.187-189	2010-12	Domestic prominent proceeding	4	ETC	0.00			0.00				
27	60673	이흥노	School of Electrical Engineering and Computer Science	27	압축센싱소개	이흥노,박상준,박순철	전자공학회지	Vol 38, issue 1, pp. 1-12	2011-01	Domestic etc.	3	CA	0.00			0.00				N07050
28	60673	이흥노	School of Electrical Engineering and Computer Science	28	An EEG signal Analysis Technique for Brain Computer Interface(뇌전도 신호를 이용한 실시간 1차원 커서제어 실험)	이흥노, 이승찬, 신영학	한국통신학회 2011 동계종합학술발표회	vol.44, pp 65	2011-02	The other domestic proceeding	3	CA	0.00			0.00				N07050
29	60673	이흥노	School of Electrical Engineering and Computer Science	29	OFDM system design over frequency/time selective underwater acoustic channels	Hyeong-Won Jeon, Su-Je Lee and Heung-No Lee	한국통신학회, 2011 동계종합학술발표회	pp. 319-320	2011-02	Domestic prominent proceeding	3	CA	0.00			0.00				N07050,N08290
30	60673	이흥노	School of Electrical Engineering and Computer Science	30	Concatenation of LDPC codes with Golden Space-Time Block Codes over the Block Fading MIMO Channel: System Design and Performance Analysis	Jin-Taek Seong, Heung-No Lee	45th Annual Conference on Information Science and Systems	1 - 4	2011-03	International prominent proceeding	2	CA	0.00			0.00				N07050,N08290
31	60673	이흥노	School of Electrical Engineering and Computer Science	31	A Research on Reliable Coded OFDM System in Underwater Acoustic Channels	Hyeong-Won Jeon, Su-Je Lee and Heung-No Lee	JCCI2011 (제 21회 통신정보 합동학술대회)	pp. 31	2011-05	Domestic prominent proceeding	3	CA	0.00			0.00				N07050,F01110
32	60673	이흥노	School of Electrical Engineering and Computer Science	32	Underwater Channel Analysis and Transmission Method Research via Coded OFDM	Hyeong-Won Jeon, Su-Je Lee and Heung-No Lee	한국통신학회논문지	Vol. 36, No. 5, pp. 573-581	2011-05	Domestic etc.	3	CA	0.00			0.00				N07050,F01110
33	60673	이흥노	School of Electrical Engineering and Computer Science	33	Motor Imagery based BCI Classification via Sparse Representation of EEG Signals	Younghak Shin, Seungchan Lee, Minkyu Ahn, Sung Chan Jun and Heung-No Lee	8th Int'l Symp on Noninvasive Functional Source Imaging of the Brain and Heart and the 8th Int'l Conf on Bioelectromagnetism	93-97	2011-05	International prominent proceeding	5	CA	0.00			0.00				N07050
34	60673	이흥노	School of Electrical Engineering and Computer Science	34	무선 다중 접속 망에서의 네트워크 부호를 이용한 보안 매커니즘	이웅비, 최재건, 이흥노	한국통신학회, 하계종합학술발표회		2011-06	The other domestic proceeding	3	CA	0.00			0.00				N08290F01110
35	60673	이흥노	School of Electrical Engineering and Computer Science	35	A New BCI Classification Method based on EEG Sparse Representation	Younghak Shin, Seungchan Lee and Heung-No Lee	Sigan! Processing with Adaptive Sparse Structured Representation 2011	104	2011-06	International prominent proceeding	3	CA	0.00			0.00				N07050
36	60673	이흥노	School of Electrical Engineering and Computer Science	36	연접된 LDPC 코드와 Golden 시공간 코드를 이용한 블록 페이딩 MIMO 채널에서 코딩 시스템 설계	Jin-Taek Seong, Heung-No Lee	한국통신학회 2011 하계종합학술발표회	1 - 2	2011-06	Domestic prominent proceeding	2	CA	0.00			0.00				N07050,N08290
37	60673	이흥노	School of Electrical Engineering and Computer Science	37	Performance Limits of the Measurements on Compressive Sensing for Multiple Sensor System	Sangjun Park, Hwanchol Jang, Heung-No Lee	Sigan! Processing with Adaptive Sparse Structured Representation 2011	123	2011-06	International prominent proceeding	3	CA	0.00			0.00				N07050
38	60673	이흥노	School of Electrical Engineering and Computer Science	38	Performance Verification of LDPC coded OFDM System in Underwater Acoustic Channels	Hyeong-Won Jeon, Su-Je Lee and Heung-No Lee	ITC-CSCC2011 (The 26th International Technical Conference on Circuits/Systems)	pp. 686-687	2011-06	International prominent proceeding	3	CA	0.00			0.00				N07050,F01110
39	60673	이흥노	School of Electrical Engineering and Computer Science	39	Improved Stopping Set Analysis of Regular LDPC code	Jeongmin Ryu, Sangjun Park, Heung-No Lee	한국통신학회 2011	Vol.45 pp. 322	2011-06	The other domestic proceeding	3	CA	0.00			0.00				N08290

40	60673	이흥노	School of Electrical Engineering and Computer Science	40	An EEG signal Analysis Technique for Brain Computer Interface(뇌-컴퓨터 접속기술을 위한 뇌전도 신호 분석법)	이흥노, 이승찬, 신영학	한국통신학회 2011 하계종합학술발표회	vol.45, pp 113	2011-06	The other domestic proceeding	3	CA	0.00			0.00		N07050
41	60673	이흥노	School of Electrical Engineering and Computer Science	41	LDPC matrix's spark value Research by using Litsyn's Distance Sepctrum	Su-Je Lee, Hyeong-Won Jeon and Heung-No Lee	전남, 광주 춘계종합학술발표회	pp.32-33	2011-06	The other domestic proceeding	3	CA	0.00			0.00		N07050
42	60673	이흥노	School of Electrical Engineering and Computer Science	42	Message Passing Aided Least Square Recovery for Compressive Sensing	강재욱, 이흥노, 김기선	SPARS'11 Workshop: Signal Processing with Adaptive Sparse Structured Representations		2011-06	International prominent proceeding	3	ETC	0.00			0.00		N08290
43	60673	이흥노	School of Electrical Engineering and Computer Science	43	Compressive Sensing RIP condition Research by using Distance Spectrum of the LDPC Matrix	Su-Je Lee, Hyeong-Won Jeon and Heung-No Lee	대한전자공학회 하계종합 학술대회	pp.211-212	2011-06	The other domestic proceeding	3	CA	0.00			0.00		N07050,F01110
44	60673	이흥노	School of Electrical Engineering and Computer Science	44	Reconstruction of binary signal using specially designed measurement matrix	Soonchul Park, Hwanchol Jang and Heung-No Lee	한국통신학회 2011년도 하계종합학술발표회	pp.25-25	2011-06	Domestic prominent proceeding	3	CA	0.00			0.00		N07050
45	60673	이흥노	School of Electrical Engineering and Computer Science	45	A Realistic Distributed Compressive Sensing framework for Multiple Wireless Sensor Networks	J. Oliver, Heung-No Lee	Sigan! Processing with Adaptive Sparse Structured Representation 2011	105 (1 Page)	2011-06	International prominent proceeding	2	CA	0.00			0.00		N07050
46	60673	이흥노	School of Electrical Engineering and Computer Science	46	Sorted Orthotope Sphere Decoding for MIMO Detection	Hwanchol Jang, Saeid Nooshabadi, Sangjun Park, Heung-No Lee	5th JCITCT & YES-ICuC		2011-08	International prominent proceeding	4	CA	0.00			0.00		N07050
47	60673	이흥노	School of Electrical Engineering and Computer Science	47	A Near-ML Decoding with Improved Complexity over Wider Ranges of SNR and System Dimension in MIMO Systems	Junil Ahn, Heung-No Lee, Kiseon Kim	IEEE Trans. on Wireless Communications		2011-09	International journal SCI	3	ETC	2.152	TELECOMMUNICATIONS	9/80	11.00	1536-1276	
48	60673	이흥노	School of Electrical Engineering and Computer Science	48	On the Derivation of RIP for Random Gaussian Matrices and Binary Sparse Signals	Sangjun Park, Heung-No Lee	International Conference on ICT Convergence	120-124	2011-09	International prominent proceeding	2	CA	0.00			0.00		N07050
49	60673	이흥노	School of Electrical Engineering and Computer Science	49	검출-추정 구조 기반의 희소 신호 복구	강재욱, 이흥노, 김기선	제 24회 신호처리 합동학술대회	pp.217-220	2011-09	International prominent proceeding	3	ETC	0.00			0.00		N07050,N08290
50	60673	이흥노	School of Electrical Engineering and Computer Science	50	Compressed Sensing over Galois Fields: Sensing Bounds and Recovery Algorithms	Heung-No Lee, Jin-Taek Seong, Su-Je Lee	2011 대한수학회 가을 연구발표회	1 - 5	2011-10	Domestic prominent proceeding	3	FA	0.00			0.00		N07050
51	60673	이흥노	School of Electrical Engineering and Computer Science	51	User-satisfaction based bandwidth allocation for transmission of multiple sources of human perceptual data	heng-Chung Chang, T.-Y. Kuo, Y.-C. Lo, Heung-No Lee, D. Askey, Zhing-Hong Mao	Journal of the Franklin Institute		2011-11	International journal SCI	6	ETC	1.508	ENGINEERING, MULTIDISCIPLINARY	14/87	16.00	0016-0032	
52	60673	이흥노	School of Electrical Engineering and Computer Science	52	Study on Performance Behavior of Compressive Sensing Measurements for Multiple Sensor System	Sangjun Park, Hwanchol Jang, Heung-No Lee	45th Asilomar Conference on Signals, Systems and Computes, Asilomar		2011-11	International prominent proceeding	3	CA	0.00			0.00		N07050
53	60673	이흥노	School of Electrical Engineering and Computer Science	53	LDPC Coded OFDM System Design and Performance Verification on a Realistic Underwater Acoustic Channel Model	Hyeong-Won Jeon, Su-Je Lee and Heung-No Lee	MILCOM 2011 (The 30th anniversary of the premier international conference for military communications ())	pp. 2200-2204	2011-11	International prominent proceeding	3	CA	0.00			0.00		N07050,F01110
54	60673	이흥노	School of Electrical Engineering and Computer Science	54	Predicting the Pruning Potential in Sphere Decoding for Multiple-Input Multiple-Output Detection	Hwanchol Jang, Saeid Nooshabadi, Heung-No Lee	45th Asilomar Conference on Signals, Systems and Computes, Asilomar		2011-11	International prominent proceeding	3	CA	0.00			0.00		N07050,N08290
55	60673	이흥노	School of Electrical Engineering and Computer Science	55	Improving resolution of miniature spectrometers by exploiting sparse nature of signals	J. Oliver, WoongBi Lee , SangJun Park , and Heung-No Lee	Optics Express	Vol. 20, Issue 3, pp. 2613-2625	2012-01	International journal SCI	4	ETC	3.587	OPTICS	6/79	7.60	1094-4087	N07050
56	60673	이흥노	School of Electrical Engineering and Computer Science	56	Efficient Joint Source-Channel Decoding of Multi-State Markov Sequences	H. Kim, D. Har, Z.H. Mao, M. Sun, Heung-No Lee	IET Communications	Volume 6, Issue 9 Pages 1038-1044	2012-06	International journal SCI	5	ETC	0.829	ENGINEERING, ELECTRICAL & ELECTRONIC	148/245	60.40	1751-8628	-
57	60673	이흥노	School of Electrical Engineering and Computer Science	57	Analysis and Design of Coding and Interleaving in a MIMO-OFDM Communication System	Zafar Iqbal, Saeid Nooshabadi, and Heung-No Lee	IEEE Transactions on Consumer Electronics	Vol58, No3 pp758~766	2012-08	International journal SCI	3	ETC	0.941	TELECOMMUNICATIONS	39/79	49.40	0098-3063	N07050 N08290
58	60673	이흥노	School of Electrical Engineering and Computer Science	58	Computation of an Equilibrium in Spectrum Markets for Cognitive Radio Networks	Sang-Seon Byun, Ilangko Balasingham, Athanasios V. Vasilakos, and Heung-No Lee	IEEE Transactions on Computers	issue 99	2012-08	International journal SCI	4	CA	1.103	COMPUTER SCIENCE, HARDWARE & ARCHITECTURE	19/50	38.00	0018-9340	N08660 N08290
59	60673	이흥노	School of Electrical Engineering and Computer Science	59	Sparse representation-based classification scheme for motor imagery-based brain-computer interface systems	Younghak Shin, Seungchan Lee, Junho Lee and Heung-No Lee	Journal of Neural Engineering	Vol9. issue5, art number 056002(12pp)	2012-10	International journal SCI	4	ETC	3.837	ENGINEERING, BIOMEDICAL	5/72	6.90	1741-2560	N08660

60	60673	이홍노	School of Electrical Engineering and Computer Science	60	Filters with random transmittance for improving resolution in spectrometers	J. Oliver, WoongBi Lee and Heung-No Lee	Optics Express	Vol. 21, No. 4, pp. 3969?3989	2013-02	International journal SCI	3	ETC	3.546	OPTICS	5/80	6.25	-	
61	60673	이홍노	School of Electrical Engineering and Computer Science	61	A novel method for controlled synthesis of nanosize hematite (α-Fe2O3) thin film on liquidvapor interface	Pawan Kumar, Raj Kumar Singh, Nitin Rawat, Partha Bir Barman, Subhash Chander Katyal, Hwanchol Jang, Heung-No Lee*, and Rajesh Kumar*	Journal of nano particle research	15:1532	2013-04	International journal SCI	8	ETC	2.175	CHEMISTRY, MULTIDISCIPLINARY	57/241	23.65	-	
62	60673	이홍노	School of Electrical Engineering and Computer Science	62	An Inventory Model-based Spectrum Pooling for Wireless Service Provider and Unlicensed Users	Sang-Seon Byun, Ilangko Balasingham, and Heung-No Lee	Computer Communications	Vol. 36, issues. 10-11, pp. 118671191	2013-06	International journal SCI	3	ETC	1.079	COMPUTER SCIENCE, INFORMATION SYSTEMS	58/132	43.94	-	
63	60673	이홍노	School of Electrical Engineering and Computer Science	63	Necessary and Sufficient Conditions for Recovery of Sparse Signals over Finite Fields	Jin-Taek Seong, Heung-No Lee	IEEE Communications Letters	vol. 17, no. 10, pp. 1976-1979	2013-10	International journal SCI	2	ETC	1.160	Telecommunications	30/78	38.46	-	
64	60673	이홍노	School of Electrical Engineering and Computer Science	64	Design of Dry Electrode for EEG based BCI systems	Seungchan Lee, Younghak Shin, Soogil Woo, Kiseon Kim and Heung-No Lee	5th International BCI Meeting		2013-12	International prominent proceeding	5	CA	0.00		/	0.00		
65	60673	이홍노	School of Electrical Engineering and Computer Science	65	두 노드의 협력 네트워크에서 비이진 네트워크 코딩을 이용한 Outage Probability	Jin-Taek Seong and Heung-No Lee	2013년 동계종합학술발표회, 한국통신학회(KICS)	pp. 635-636	2013-12	Domestic prominent proceeding	2	CA	0.00		/	0.00		
66	60673	이홍노	School of Electrical Engineering and Computer Science	66	Coordinating Transmit Power and Carrier Phase for Wireless Networks with Multi-Packet Reception Capability	Wooyeol Choi, Taewoon Kim, Daeyoung Park, Heung-No Lee and Hyuk Lim	EURASIP Journal on Wireless Communications and Networking	doi:10.1186/1687-1499-2013-1	2013-12	International journal SCI	5	ETC	0.540	TELECOMMUNICATIONS	56/78	71.79		
67	60673	이홍노	School of Electrical Engineering and Computer Science	67	Enhancing the numerical aperture of lenses using ZnO nanostructures-based turbid media	Richa Khokhra, Manoj Kumar, Nitin Rawat, Partha Bir Barman, Hwanchol Jang, Rajesh Kumar and Heung-No Lee	Journal of Optics	pp.1-6	2013-12	International journal SCI	7	ETC	1.990	OPTICS	22/80	27.50	-	
68	60673	이홍노	School of Electrical Engineering and Computer Science	68	압축 센싱 신호 복구를 위한 AMP(Approximate Message Passing) 알고리즘 소개 및 성능 분석	Hyeong-ho Baek, Jae-wook Kang, Ki-sun Kim, Heung-no Lee	한국통신학회논문지	Vol. 38C, no. 11, pp. 1029-1043	2013-12	Domestic etc.	4	CA	0.00		/	0.00		
69	60673	이홍노	School of Electrical Engineering and Computer Science	69	Noisy Behavior of MAP-based Sparse Support Detection	Jaewook Kang, Heung-No Lee, and Kiseon Kim	SPARS 2013, Lausanne in Switzerland		2013-12	International prominent proceeding	3	ETC	0.00		/	0.00		
70	60673	이홍노	School of Electrical Engineering and Computer Science	70	Random Transmittance Based Filter Array Spectrometers: Sparse Spectrum Recovery And Resolution Improvement	J. Oliver, WoongBi Lee and Heung-No Lee	IEEE Global Conference on Signal and Information Processing (GlobalSIP), Austin, Texas, U.S.A		2013-12	International prominent proceeding	3	CA	0.00		/	0.00		
71	60673	이홍노	School of Electrical Engineering and Computer Science	71	Exact Outage Probability of Two Nodes for Cooperative Networking using GF(4)	Jin-Taek Seong and Heung-No Lee	14th IEEE International Workshop Signal Processing Advances in Wireless Communications, Darmstadt, Germany		2013-12	International prominent proceeding	2	CA	0.00		/	0.00		
72	60673	이홍노	School of Electrical Engineering and Computer Science	72	Design of Dry Electrode for Wireless BCI systems	Seungchan Lee, Younghak Shin, Soogil Woo, Kiseon Kim and Heung-No Lee	35th IEEE EMBC 2013		2013-12	International prominent proceeding	5	CA	0.00		/	0.00		
73	60673	이홍노	School of Electrical Engineering and Computer Science	73	Random optical scatter filters for spectrometers: Implementation and Estimation	Woong-Bi Lee, J. Oliver, Seung-Chul Kim, and Heung-No Lee	Optics and photonics congresses: Applied Industrial Optics ? Spectroscopy, Imaging, & Metrology, Arlington, Virginia		2013-12	International prominent proceeding	4	CA	0.00		/	0.00		
74	60673	이홍노	School of Electrical Engineering and Computer Science	74	압축 센싱을 이용한 초음파 이미지의 신호 복원	Pavel Ni, Jin-Taek Seong, Hwanchol Jang, Sangjun Park, Jeong-Min Ryu, Heung-No Lee	2013년도 대한전자공학회 하계종합학술대회	pp.769-771	2013-12	Domestic prominent proceeding	6	CA	0.00		/	0.00		
75	60673	이홍노	School of Electrical Engineering and Computer Science	75	협력적 무선 다중 접속 망에서 네트워크 부호를 이용한 릴레이 공격 감지 방법	Woong-Bi Lee, Jusung Kang, J. Oliver, and Heung-No Lee	2013 하계학술대회, 한국통신학회	pp.758-759	2013-12	Domestic prominent proceeding	4	CA	0.00		/	0.00		
76	60673	이홍노	School of Electrical Engineering and Computer Science	76	BCI 시스템을 위한 Dry 전극의 설계 및 임피던스 성능 비교	Seungchan Lee, Younghak Shin, Soogil Woo, and Heung-No Lee	HCI Korea 2013	pp. 773-775	2013-12	Domestic prominent proceeding	4	CA	0.00		/	0.00		
77	60673	이홍노	School of Electrical Engineering and Computer Science	77	A new signal processing technique for improving resolution of spectrometers	J. Oliver, WoongBi Lee , SangJun Park , and Heung-No Lee	2013년 동계종합학술발표회, 한국통신학회 (KICS)	pp. 62-63	2013-12	Domestic prominent proceeding	4	CA	0.00		/	0.00		
78	60673	이홍노	School of Electrical Engineering and Computer Science	78	압축 센싱신호 복구를 위한 AMP(Approximate Message Passing)의 재분석	Hyeong-ho Baek, Sang-Jun Park, Jeong-Min Ryu, and Heung-No Lee	2013 하계학술대회, 한국통신학회	pp.701-702	2013-12	Domestic prominent proceeding	4	CA	0.00		/	0.00		
79	60673	이홍노	School of Electrical Engineering and Computer Science	79	Performance Enhancement by Sparse Representation of EEG Signals for Motor Imagery Based BCI systems	Younghak Shin, Seungchan Lee, Soogil Woo and Heung-No Lee	5th International BCI Meeting		2013-12	International prominent proceeding	4	CA	0.00		/	0.00		

80	60673	이종노	School of Electrical Engineering and Computer Science	80	Exact Outage Probability and Power Allocation of Two Nodes in Cooperative Networks	Jin-Taek Seong and Heung-No Lee	2013 IEEE WCNC(2013 IEEE Wireless Communications and Networking conference)	pp. 2751-2756	2013-12	International prominent proceeding	2	CA		0.00			/	0.00				
81	60673	이종노	School of Electrical Engineering and Computer Science	81	Dry Electrode Design and Performance Evaluation for EEG based BCI systems	Seungchan Lee, Younghak Shin, Soogil Woo, Kiseon Kim and Heung-No Lee	2013 IEEE International Winter Workshop on Brain-Computer Interface	pp. 2-31 – 2-32	2013-12	International prominent proceeding	5	CA		0.00			/	0.00				
82	60673	이종노	School of Electrical Engineering and Computer Science	82	Performance Increase by using a EEG Sparse Representation based Classification Method	Younghak Shin, Seungchan Lee, Soogil Woo and Heung-No Lee	2013 IEEE ICCE(2013 IEEE International Winter Workshop on Brain-Computer Interface)	pp. 205–207	2013-12	International prominent proceeding	4	CA		0.00			/	0.00				
83	60673	이종노	School of Electrical Engineering and Computer Science	83	BCI 시스템의 건식전극 설계 및 성능 분석	Soogil Woo, Younghak Shin, Seungchan Lee, Jongmok Shin, and Heung-No Lee	2013년도 대한전자공학회 하계종합학술대회		2013-12	Domestic prominent proceeding	5	CA		0.00			/	0.00				
84	60673	이종노	School of Electrical Engineering and Computer Science	84	Evaluation of resolution improvement capability of a DSP technique for filter-array-based spectrometers	J. Oliver, Woong-Bi Lee, Sanglun-Park, and Heung-No Lee	한국통신학회	vol. 38C, no. 6, pp. 497-502	2013-12	Domestic etc.	4	CA		0.00			/	0.00				
85	60673	이종노	School of Electrical Engineering and Computer Science	85	Sparse or Dense-Message Passing (MP) or Approximate Message Passing (AMP) for Compressde Sensing Signal Recovery	Asad Mahmood, Jaewook and Heung-No Lee	2013 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, University of Victoria, Victoria, B.C., Canada		2013-12	International prominent proceeding	3	CA		0.00			/	0.00				
86	60673	이종노	School of Electrical Engineering and Computer Science	86	Ultra wideband channel estimation based compressive sensing and approximate message passing (AMP)	Hyeongho Baek, Jongmok Shin, Jusung Kang and Heung-No Lee,	한국통신학회, 2014년도 동계종합학술발표회		2014-01	Domestic prominent proceeding	4	CA		0.00			/	0.00				
87	60673	이종노	School of Electrical Engineering and Computer Science	87	Synthesis of phase pure iron oxide polymorphs thin films and their enhanced magnetic properties	Pawan Kumar, Heung-No Lee, Rajesh Kumar	J Mater Sci: Mater Electron	Vol.25, PP.4553-4561	2014-01	International journal SCI	3	ETC		1.966	MATERIALS SCIENCE, MULTIDISCIPLINARY		80/251	31.87	0957-4522	N12120		
88	60673	이종노	School of Electrical Engineering and Computer Science	88	Issues and Resolution Efforts Pertaining to TV Whitespace Usage	Muhammad A. Raza, Zafar Iqbal, Seungchan Lee, Haeung Choi, and Heung-No Lee	한국통신학회, 2014년도 동계종합학술발표회		2014-01	Domestic prominent proceeding	5	CA		0.00			/	0.00				
89	60673	이종노	School of Electrical Engineering and Computer Science	89	Design of Active Dry Electrodes for BCI systems and Performance Evaluation	Seungchan Lee, Younghak Shin, Jeongmin Ryu, and Heung-No Lee	HCI Korea 2014		2014-02	Domestic prominent proceeding	4	CA		0.00			/	0.00				
90	60673	이종노	School of Electrical Engineering and Computer Science	90	압축 센싱 기술과 신호간의 상관관계를 이용하여 전력 소모를 줄인 무선 센서 네트워크 시스템의 설계	Haeung Choi, Jaegun Choi, Jongmok Shin, Seungyun Choe, and Heung-No Lee	대한전자공학회 2014년도 하계종합학술발표회		2014-06	Domestic prominent proceeding	5	CA		0.00			/	0.00				
91	60673	이종노	School of Electrical Engineering and Computer Science	91	Speckle suppression via sparse representation for wide-field imaging through turbid media	Hwanchol Jang, Changhyeong Yoon, Euiheon Chung, Wonshik Choi, and Heung-No Lee	Optics Express	Vol. 22, No. 13,?pp. 16619?16628	2014-06	International journal SCI	5	CA		3.525	OPTICS		6/83	7.23	1094-4087	N12120		
92	60673	이종노	School of Electrical Engineering and Computer Science	92	Editorial: Special Issue On Compressed Sensing: Seeing through Computation	이종노	대한전자공학회, 전자공학회지	Vol. 41, No. 6	2014-06	Domestic etc.	1	FA		0.00			/	0.00				
93	60673	이종노	School of Electrical Engineering and Computer Science	93	Narrow band Interference Cancellation In Wide band Communications Using Compressed Sensing	신종목, 강주성, 이종노	전자공학회지	2014.6월, 547-555	2014-06	Domestic etc.	3	CA		0.00			/	0.00				
94	60673	이종노	School of Electrical Engineering and Computer Science	94	Predicting the Performance of Cooperative Wireless Networking Schemes with Random Network Coding	Jin-Taek Seong and Heung-No Lee	IEEE Transactions on Communications	vol. 62, no. 8, pp. 2951-2964	2014-08	International journal SCI	2	ETC		1.979	ENGINEERING, ELECTRICAL & ELECTRONIC		61/248	24.60	0090-6778	N12120		
95	60673	이종노	School of Electrical Engineering and Computer Science	95	Design of Active Dry Electrodes for EEG based BCI systems	Seungchan Lee, Younghak Shin and Heung-No Lee	6th International Brain-Computer Interface Conference	pp.16-19	2014-09	International prominent proceeding	3	CA		0.00			/	0.00				
96	60673	이종노	School of Electrical Engineering and Computer Science	96	Controlling band gap and refractive index in dopant free α-Fe2O3 films	Pawan Kumar, Nitin Rawat, Da-Ren Hang, Heung-No Lee and Rajesh Kumar	Electronic Materials Letters	vol. 11, no. 1, pp. 13-23	2015-01	International journal SCI	5	ETC	1	1.980	MATERIALS SCIENCE, MULTIDISCIPLINARY	2014 JCR	88/260	33.80	1738-8090	N14150		
97	60673	이종노	School of Electrical Engineering and Computer Science	97	Bayesian Hypothesis Test using Nonparametric Belief Propagation for Noisy Sparse Recovery	Jaewook Kang, Heung-No Lee, Kiseon Kim	IEEE Transactions on Signal Processing	Vol. 110, pp. 1781-1785	2015-02	International journal SCI	3	ETC	1	2.787	ENGINEERING, ELECTRICAL & ELECTRONIC	2014 JCR	27/249	10.80	1053-587X	N14150		
98	60673	이종노	School of Electrical Engineering and Computer Science	98	Holistic random encoding for imaging through multimode fibers	Hwanchol Jang, Changhyeong Yoon, Euiheon Chung, Wonshik Choi, and Heung-No Lee	Optics Express	Vol. 23, No. 5, pp. 6705-6721	2015-03	International journal SCI	5	CA	1	3.488	OPTICS	2014 JCR	10/87	11.50	1094-4087	N14150		
99	60673	이종노	School of Electrical Engineering and Computer Science	99	Noise Robustness Analysis of Sparse Representation based Classification Method for Non-stationary EEG Signal Classification	Younghak Shin, Seungchan Lee, Minkyu Ahn, Hohyun Cho Sung Chan Jun and Heung-No Lee	Biomedical Signal Processing and Control	Vol.21, pp. 8-18	2015-08	International journal SCI	6	CA	1	1.419	MEDICAL LABORATORY TECHNOLOGY	2014 JCR	15/30	50.00	1746-8094	N14150		

100	60673	이홍노	School of Electrical Engineering and Computer Science	100	Magnetic field induced one-dimensional nano/micro structures growth on the surface of iron oxide thin film	Pawan Kumar, Rajesh Kumar* and Heung-No Lee	Thin Solid Films	Vol.592, pp.155-161	2015-10	International journal SCI	3	ETC	1	1.759	MATERIALS SCIENCE, COATINGS & FILMS	2014 JCR	6/17	35.30	0040-6090	N14150
101	60673	이홍노	School of Electrical Engineering and Computer Science	101	Simple Adaptive Sparse Representation based Classification Schemes for EEG based Brain-Computer Interface Applications	Younghak Shin, Seungchan Lee, Minkyu Ahn, Hohyun Cho, Sung Chan Jun and Heung-No Lee	Computers in Biology and Medicine	Vol.66, pp.29-38	2015-11	International journal SCI	6	CA	1	1.240	BIOLOGY	2014 JCR	49/85	57.60	0010-4825	N14150
102	60673	이홍노	School of Electrical Engineering and Computer Science	102	불규칙 매체를 통한 컴퓨테이셔널 이미징의 최근 연구 동향	장환철, 윤창형, 정의헌, 최원식, 이홍노	한국통신학회	Vol. 39A, No. 12, pp. 764 ? 770	2015-12	Domestic etc.	5	CA		0.00				0.00		-
103	60673	이홍노	School of Electrical Engineering and Computer Science	103	상용 Walkie-Talkie에 대한 천이상태 신호 기반 Radio Frequency Fingerprinting 시스템	Jusung Kang, Haewoong Choi, Jaewon Bang, Rohit Thakur, Cheolsun Kim and Heung-No Lee	2016년도 한국통신학회 동계종합학술발표회	59, 555-556	2016-01	The other domestic proceeding	6	CA	1	0.00			/	0.00	2383-8302	G10441
104	60673	이홍노	School of Electrical Engineering and Computer Science	104	COMPU-EYE: a high resolution computational compound eye	Woong-Bi Lee, Hwanchol Jang, Sangjun Park, Young Min Song, Heung-No Lee	Optics Express	24, 2013-2026	2016-01	International journal SCI	5	CA	1	3.148	OPTICS	2015	14/90	15.60	1094-4087	N15860
105	60673	이홍노	School of Electrical Engineering and Computer Science	105	안드로이드 모바일 기기들을 위한 뇌-컴퓨터 인터페이스 문자 입력기 어플리케이션	Seungyun Choe, Seungchan Lee, Younghak Shin, Jehyuk Jang and Heung-No Lee	2016년도 한국통신학회 동계종합학술발표회	59, 647-648	2016-01	The other domestic proceeding	5	CA	1	0.00			/	0.00	2383-8302	N14150
106	60673	이홍노	School of Electrical Engineering and Computer Science	106	희소한 가중 그래프 상에서의 압축센싱 복구방법	Haeung Choi and Heung-No Lee	JCCI 2016	26, 384-385	2016-04	The other domestic proceeding	2	CA	1	0.00			/	0.00		N14150
107	60673	이홍노	School of Electrical Engineering and Computer Science	107	Modulated Wideband Converter를 위한 혼합 시퀀스의 조건	Jehyuk Jang, Nam Yul Yu, and Heung-No Lee	JCCI 2016	26, 382-383	2016-04	The other domestic proceeding	3	CA	1	0.00			/	0.00		X
108	60673	이홍노	School of Electrical Engineering and Computer Science	108	군집드론의 전자전 활용	Cheolsun Kim, Woong-Bi Lee, Heung-No Lee	제 26회 통신정보 합동학술대회	26, 381	2016-04	The other domestic proceeding	3	CA	1	0.00			/	0.00		N14150
109	60673	이홍노	School of Electrical Engineering and Computer Science	109	Design of thin-film filters for resolution improvements in filter-array based spectrometers using DSP	Woong-Bi Lee, Cheolsun Kim, Gun Wu Ju, Yong Tak Lee, and Heung-No Lee	SPIE Defense+Commercial Sensing 2016	9855, 0-0	2016-04	International prominent proceeding	5	CA	1	0.00			/	0.00		N15860
110	60673	이홍노	School of Electrical Engineering and Computer Science	110	Circular Sphere Decoding: A Low Complexity Detection for MIMO Systems with General Two-dimensional Signal Constellations	Hwanchol Jang, Saeid Nooshabadi, Kiseon Kim, and Heung-No Lee	IEEE Transactions on Vehicular Technology	0,0	2016-05	International journal SCI	4	CA	1	2.243	ENGINEERING, ELECTRICAL & ELECTRONIC, TELECOMMUNICATIONS, TRANSPORTATION SCIENCE & TECHNOLOGY	2015	55/257	21.40	0018-9545	N15860
111	60673	이홍노	School of Electrical Engineering and Computer Science	111	Spatially concatenated channel-network code for underwater wireless sensor networks	Zafar Iqbal, Heung-No Lee	IEEE Transactions on Communications	64, 3901-3914	2016-06	International journal SCI	2	CA	1	2.298	ENGINEERING, ELECTRICAL & ELECTRONIC, TELECOMMUNICATIONS	2015	48/257	18.70	0090-6778	N15860
112	60673	이홍노	School of Electrical Engineering and Computer Science	112	압축센싱 신호처리 기법을 적용한 형재명/저피탐 통신 시스템	Jae-Won Bang, Hae-Ung Choi, Chul-Sun Kim, Heung-No Lee	한국통신학회 2016년도 하계종합학술발표회	60, 1289-1290	2016-06	The other domestic proceeding	4	CA	1	0.00			/	0.00	3283-8302	N15860
113	60673	이홍노	School of Electrical Engineering and Computer Science	113	A self-organizing wireless sensor network for industrial monitoring	Zafar Iqbal and Heung-No Lee	31st Int. Conf. on Circuits/Systems, Computers, and Communications (ITC-CSCC 2016)	0, 351-354	2016-07	The other international proceeding	2	CA	1	0.00			/	0.00		N15860
114	60673	이홍노	School of Electrical Engineering and Computer Science	114	Formation of oxygen vacancies and Ti3+ state in TiO2 thin film and enhanced optical	Bandna Bharti, Santosh Kumar, Heung-No Lee, Rajesh Kumar	Scientific Reports-Nature	6, 0-0	2016-08	International journal SCI	4	ETC	2	5.228	MULTIDISCIPLINART SCIENCES	2015	7/63	11.10	2045-2322	N15860
115	60673	이홍노	School of Electrical Engineering and Computer Science	115	Design of Multiplierless Cosine Modulated Filterbank using Hybrid Technique in Sub-Expression Space	I. Sharma, A. Kumar and G. K. Singh, H.-N. Lee	21th IEEE International Conference	360-364	2016-10	International prominent proceeding	4	CA	1	0.00	ENGINEERING, ELECTRICAL & ELECTRONIC, TELECOMMUNICATIONS		/	0.00		사사없음
116	60673	이홍노	School of Electrical Engineering and Computer Science	116	압축센싱과 광대역 레이더 신호 획득 기술	Jeong Park, Jaehyuk Jang, Heung-No Lee	대한전자공학회, 전자공학회지	43(10), 809-814	2016-10	The other domestic proceeding	3	CA	1	0.00			/	0.00	1016-9288	I05150
117	60673	이홍노	School of Electrical Engineering and Computer Science	117	A novel gamma correction approach using optimally clipped sub-equalization for dark image enhancement	H. Singh, A. Kumar G. K. Singh, H.-N. Lee	21th IEEE International Conference	497-501	2016-10	International prominent proceeding	4	CA	1	0.00	ENGINEERING, ELECTRICAL & ELECTRONIC, TELECOMMUNICATIONS		/	0.00		사사없음
118	60673	이홍노	School of Electrical Engineering and Computer Science	118	A Study on Mixing Sequences in Modulated Wideband Converters	Jehyuk Jang, Nam Yul Yu, and Heung-No Lee	2016 IEEE Conference on Signal and Information Processing (Global SIP)	0, 1408-1412	2016-12	International prominent proceeding	3	CA	1	0.00			/	0.00		사사없음

119	60673	이종노	School of Electrical Engineering and Computer Science	119	컬러 영상 데이터베이스를 활용한 적외선 영상 객체 분류 신경망 학습 방법에 관한 연구	Jusung Kang, Hyunjun Han, Rohit Thakur, Nakwoo Kim, Byongtak Lee and Heung-No Lee	CEIC 2016		2016-12	The other domestic proceeding	6	CA	1	0.00			/	0.00			N16060
120	60673	이종노	School of Electrical Engineering and Computer Science	120	A method of incident angle estimation for high resolution spectral recovery in filter-array-based spectrometers	Cheolsun Kim, Woong-Bi Lee, Gun Wu Ju, Jeonghoon Cho, Seongmin Kim, Jinkyung Oh, Dongsung Lim, Yong Tak Lee, and Heung-No Lee	Proc. SPIE 10117	Vol. 10117 101170K- 1 - 6	2017-01	International prominent proceeding	9	CA	1	0.00			/	0.00			N15860
121	60673	이종노	School of Electrical Engineering and Computer Science	121	Design of multiplierless prototype filter for two-channel filter bank using hybrid method in FCSD space	I. Sharma, A. Kumar, and G. K. Singh, H.-N. Lee	IET Circuits Devices & Systems	Vol. 11 (No. 1), pp. 29 - 40	2017-01	International journal SCI	4	ETC	3	1.092	ENGINEERING, ELECTRICAL & ELECTRONIC	2016	187/262	71.00	1751-858X	-
122	60673	이종노	School of Electrical Engineering and Computer Science	122	Dual-hop cooperation protocol for spectrum sensing in cognitive radio networks	Zafar Iqbal and Heung-No Lee	16th Int. Conf. on Electron. Inf. and Commun. (ICEIC 2017)	pp. 409-410	2017-01	International prominent proceeding	2	CA	1	0.00			/	0.00			N15860
123	60673	이종노	School of Electrical Engineering and Computer Science	123	Compressive Sensing Based Secure Storage and Transmission of Ultrasound Images	Mohamed Yaseen.J and Heung-No Lee	16th Int. Conf. on Electron. Inf. and Commun. (ICEIC 2017)	pp. 395-396	2017-01	International prominent proceeding	2	CA	1	0.00			/	0.00			N15860
124	60673	이종노	School of Electrical Engineering and Computer Science	124	광대역 신호 수신기 잡음 중첩 영향 분석	Jeong Park, Jehyuk Jang and Heung-No Lee	2017년도 한국통신학회 동계종합학술발표회	pp. 888 - 889	2017-01	Domestic prominent proceeding	3	CA	1	0.00			/	0.00			I05150
125	60673	이종노	School of Electrical Engineering and Computer Science	125	Compression of Medical Image using Wavelet based Sparsification and Coding	Utpreksh Patbhaje, Ranjeet Kumar, A. Kumar, Heung-No Lee	IEEE Int. Conf. SPIN 2017		2017-02	International prominent proceeding	4	CA	2	0.00			/	0.00			N15860
126	60673	이종노	School of Electrical Engineering and Computer Science	126	Circular Sphere Decoding: A Low Complexity Detection for MIMO Systems with General Two-dimensional Signal Constellations	Hwanchol Jang, Saeid Nooshabadi, Kiseon Kim, and Heung-No Lee	IEEE Transactions on Vehicular Technology	Vol. 66 (No. 3), pp. 2085 - 2098	2017-03	International journal SCI	4	CA	1	4.066	ENGINEERING, ELECTRICAL & ELECTRONIC	2016	32/262	12.00	0018-9545	N14150
127	60673	이종노	School of Electrical Engineering and Computer Science	127	격자구조 및 그래프 모델 기반 무인이동체 자율주행 시스템 연구	Hyunjun Han, Jusung Kang, Rohit Thakur and Heung-No Lee	2017년도 항공우주시스템공학회 춘계학술대회		2017-04	Domestic prominent proceeding	4	CA	1	0.00			/	0.00			N16960
128	60673	이종노	School of Electrical Engineering and Computer Science	128	An Improved Design Method based on Polyphase Components for Digital FIR Filters	A. Kumar, B. Kuldeep, G. K. Singh and Heung No Lee	International Journal of Electronics	Vol. 104 (11), 1920-1934	2017-04	International journal SCI	4	ETC	2	0.729	ENGINEERING, ELECTRICAL & ELECTRONIC	2016	217/262	83.00	0020-7217	K07860
129	60673	이종노	School of Electrical Engineering and Computer Science	129	A cooperative wireless sensor network for indoor industrial monitoring	Zafar Iqbal, Kiseon Kim, and Heung-No Lee	IEEE Transactions on Industrial Informatics	Vol. 13 (No. 2), pp. 482 - 491	2017-04	International journal SCI	3	CA	1	6.764	ENGINEERING, INDUSTRIAL	2016	1/44	2.00	1551-3203	N15860
130	60673	이종노	School of Electrical Engineering and Computer Science	130	A Calibration for the Modulated Wideband Converter Using Sinusoids with Unknwon Phases	Jeong Park, Jehyuk Jang, and Heung-No Lee	9th Int. Conf. on Ubiquitos and Future Network (ICUFN 2017)		2017-04	International prominent proceeding	3	CA	1	0.00			/	0.00			N15860 I05150
131	60673	이종노	School of Electrical Engineering and Computer Science	131	Efficient compression technique based on temporal modelling of ECG signal using principle component analysis	Ranjeet Kumar, Anil Kumar, G.K. Singh, and Heung-No Lee	IET Science Measurement & Technology	Vol. 11 (No. 3), pp. 346 - 353	2017-05	International journal SCI	4	ETC	3	1.263	ENGINEERING, ELECTRICAL & ELECTRONIC	2016	167/262	63.00	1751-8822	-
132	60673	이종노	School of Electrical Engineering and Computer Science	132	Computer Aided Diagnostic System for Ultrasound Liver Images: A Systematic Review	Mohamed Yaseen J and Heung-No Lee	Optik – International Journal for Light and Electron Optics	Vol. 140, pp. 1114 - 1126	2017-05	International journal SCI	2	CA	1	0.835	OPTICS	2016	74/92	80.00	0030-4026	N15860
133	60673	이종노	School of Electrical Engineering and Computer Science	133	Low-latency and high-reliability cooperative WSN for indoor industrial monitoring	Zafar Iqbal and Heung-No Lee	IEEE 85th Vehicular Tech. Conf. (VTC-Spring)		2017-06	International prominent proceeding	2	CA	1	0.00			/	0.00			N18200
134	60673	이종노	School of Electrical Engineering and Computer Science	134	Evolutionary Channel Sharing Algorithm for Heterogeneous Unlicensed Networks	Muhammad Asif Raza, Sangjun Park and Heung-No Lee	IEEE Transactions on Wireless Communications	Vol. 16, 4378-4389	2017-07	International journal SCI	3	CA	1	4.951	ENGINEERING, ELECTRICAL & ELECTRONIC	2016	19/262	7.00	1536-1276	N15860
135	60673	이종노	School of Electrical Engineering and Computer Science	135	A Versatile Coexistence Decision-Making System for Efficient TV Whitespace Sharing among Whitespace Objects	M. asif raza, Zafar Iqbal, Sang-Seon Byun, Hyunduk Kang, Heung-No Lee,	Wireless Communications and Mobile Computing	DOI: 10.1155/2017/8498217	2017-08	International journal SCI	5	CA	1	1.899	COMPUTER SCIENCE, INFORMATION SYSTEMS ENGINEERING, ELECTRICAL & ELECTRONIC	2016	72/146	49.00	1530-8669	N18200
136	60673	이종노	School of Electrical Engineering and Computer Science	136	An Information-Theoretic Study for Joint Sparsity Pattern Recovery with Different Sensing Matrices	Sangjun Park, Nam Yul Yu, Heung-No Lee	IEEE Transactions on Information Theory	Vol. 63(9), 5559-5571	2017-09	International journal SCI	3	CA	1	2.679	ENGINEERING, ELECTRICAL & ELECTRONIC	2016	74/262	28.00	0018-9448	N18200
137	60673	이종노	School of Electrical Engineering and Computer Science	137	Visible and UV photo-detection in ZnO nanostructured thin films via simple tuning of solution method	Richa Khokhra, Bandna Bharti, Heung-No Lee, and Rajesh Kumar	Scientific Reports-Nature	Vol 7	2017-11	International journal SCI	4	CA	2	4.259	MULTIDISCIPLINARY SCIENCES	2016	10/64	15.00	2045-2322	N18200

138	60673	이홍노	School of Electrical Engineering and Computer Science	138	Depth-estimation-enabled compound eyes	Woong-Bi Lee, Heung-No Lee	Optics Communications	온라인 게재	2017-12	International journal SCI	2	CA	1	1.588	OPTICS	2016	52/92	56.00	0030-4018	N18200
139	60673	이홍노	School of Electrical Engineering and Computer Science	139	Evaluating the effect of various speckle reduction filters on ultrasound liver cancer images	Jabarulla, M.Y.; Lee, Heung-No	International Conference on Electronics, Information and Communication, ICEIC 2018		2018-01	Domestic prominent proceeding	2	CA		0.00				0.00				
140	60673	이홍노	School of Electrical Engineering and Computer Science	140	A Sub-Nyquist Radar Electronic Surveillance System	Park, Jeong; Jang, Jehyuk; Im, Sanghun; Lee, Heung-No	IEEE Access	6,,10080–10091	2018-02	SCI	4	CA	1	3.557	COMPUTER SCIENCE, INFORMATION SYSTEMS	2017	24/148	16.22	2169-3536		
141	60673	이홍노	School of Electrical Engineering and Computer Science	141	An improved design technique of digital finite impulse response filter for notch filtering	Kumar, Anil; Baderia, Kuldeep; Singh, G.K.; Lee, Singh; Lee, H.-N.	Lecture Notes in Electrical Engineering		2018-03	International prominent proceeding	5	ETC		0.00				0.00				
142	60673	이홍노	School of Electrical Engineering and Computer Science	142	Depth-estimation-enabled compound eyes	Lee, Woong-Bi; Lee, Heung-No	Optics Communications	412,,178–185	2018-04	SCI	2	CA	1	1.887	OPTICS	2017	46/94	48.94	0030-4018	N18200	
143	60673	이홍노	School of Electrical Engineering and Computer Science	143	Piecewise Gamma Corrected Optimally Framed Grumwald-Letnikov Fractional Differential Masking for Satellite Image Enhancement	Singh, H.; Kumar, A.; Balyan, L.K.; Lee, Heung-No	Proceedings of the 2018 IEEE International Conference on Communication and Signal Processing, ICCSP 2018		2018-04	International prominent proceeding	4	ETC		0.00				0.00				
144	60673	이홍노	School of Electrical Engineering and Computer Science	144	Retinal Fundus Image Enhancement Using Piecewise Gamma Corrected Dominant Orientation Based Histogram Equalization	Reddy, P.S.; Singh, H.; Kumar, A.; Balyan, L.K.; Lee, Heung-No	Proceedings of the 2018 IEEE International Conference on Communication and Signal Processing, ICCSP 2018		2018-04	International prominent proceeding	5	ETC		0.00				0.00				
145	60673	이홍노	School of Electrical Engineering and Computer Science	145	Speckle Reduction on Ultrasound Liver Images Based on a Sparse Representation over a Learned Dictionary	Jabarulla, Mohamed Yaseen; Lee, Heung-No	Applied Sciences-basel	8,6,	2018-05	SCI	2	CA	1	1.689	PHYSICS, APPLIED	2017	77/146	52.74	2076-3417		
146	60673	이홍노	School of Electrical Engineering and Computer Science	146	Intentional Aliasing Method to Improve Sub-Nyquist Sampling System	Jang, Jehyuk; Im, Sanghun; Lee, Heung-No	IEEE Transactions on Signal Processing	66,12,3311–3326	2018-06	SCI	3	CA	1	4.203	ENGINEERING, ELECTRICAL & ELECTRONIC	2017	32/260	12.31	1053-587X		
147	60673	이홍노	School of Electrical Engineering and Computer Science	147	Learning Through Adverse Event for Collision Avoidance: A Self-Learning Approach	Han, H.; Kang, J.; Raza, M.A.; Lee, H.-N.	International Conference on Ubiquitous and Future Networks, ICUFN		2018-07	International prominent proceeding	4	CA		0.00				0.00				
148	60673	이홍노	School of Electrical Engineering and Computer Science	148	Reduced computational complexity orthogonal matching pursuit using a novel partitioned inversion technique for compressive sensing	Kim, S.; Yun, U.; Jang, J.; Seo, G.; Kang, J.; Lee, Heung-No; Lee, Minjae	Electronics (Basel)	7,9,	2018-09	SCI	7	ETC	5	2.110	ENGINEERING, ELECTRICAL & ELECTRONIC	2017	113/260	43.46	2079-9292	GI05150	
149	60673	이홍노	School of Electrical Engineering and Computer Science	149	Highly Reliable Decision-Making Using Reliability Factor Feedback for Factory Condition Monitoring via WSNs	Iqbal, Zafar; Lee, Heung-No; Nooshabadi, Saeid	Wireless Communications and Mobile Computing	2018,,	2018-10	SCI	3	CA	1	0.869	ENGINEERING, ELECTRICAL & ELECTRONIC	2017	211/260	81.15	1530-8669	NN21320	
150	60673	이홍노	School of Electrical Engineering and Computer Science	150	Fast Mixed Integer Quadratic Programming for Sparse Signal Estimation	Park, Sangjun; Lee, Heung-No	IEEE Access	6,,58439–58449	2018-10	SCI	2	CA	1	3.557	COMPUTER SCIENCE, INFORMATION SYSTEMS	2017	24/148	16.22	2169-3536	NN21320	
151	60673	이홍노	School of Electrical Engineering and Computer Science	151	Fuzzified Histogram Equalization based Gamma Corrected Cosine Transformed Energy Redistribution for Image Enhancement	Singh, Himanshu; Kumar,Anil; Balyan, L.K.; Lee, Heung-No	International Conference on Digital Signal Processing, DSP		2018-11	International prominent proceeding	4	ETC		0.00				0.00				
152	60673	이홍노	School of Electrical Engineering and Computer Science	152	Design of portable functional near-infrared spectroscopy-based brain monitoring system	Lee, Seungchan; Lee, Heung-no	ICEIC 2019 - International Conference on Electronics, Information, and Communication		2019-01	International prominent proceeding	2	ETC		0.00				0.00				
153	60673	이홍노	School of Electrical Engineering and Computer Science	153	Convolutional neural networks for the reconstruction of spectra in compressive sensing spectrometers	Kim, Cheolsun; Park, Dongju; Lee, Heung-No	Proceedings of SPIE - The International Society for Optical Engineering		2019-02	Domestic prominent proceeding	3	CA		0.00				0.00				
154	60673	이홍노	School of Electrical Engineering and Computer Science	154	A new design method for FIR notch filter using Fractional Derivative and swarm intelligence	Kumar, A.; Mustikovila, K. N.; Singh, G. K.; Lee, S.; Lee, H-N	Sadhana - Academy Proceedings in Engineering Sciences	44,3,	2019-03	SCI	5	ETC	4	0.769	ENGINEERING, MULTIDISCIPLINARY	2018	73/88	82.96	0256-2499		
155	60673	이홍노	School of Electrical Engineering and Computer Science	155	Fabrication of 2D thin-film filter-array for compressive sensing spectroscopy	Kim, C.; Lee, W.-B.; Lee, S.K.; Lee, Y.T.; Lee, H.-N.	Optics and Lasers in Engineering	115,,53–58	2019-04	SCI	5	CA	1	4.059	OPTICS	2018	14/95	14.74	0143-8166		
156	60673	이홍노	School of Electrical Engineering and Computer Science	156	Fractional-Order Integration Based Fusion Model for Piecewise Gamma Correction Along With Textural Improvement for Satellite Images	Singh, Himanshu; Kumar, Anil; Balyan, L. K; Lee, Heung-No	IEEE Access	7,,37192–37210	2019-04	SCI	4	CA	1	4.098	COMPUTER SCIENCE, INFORMATION SYSTEMS	2018	23/155	14.84	2169-3536		

157	60673	이종노	School of Electrical Engineering and Computer Science	157	Dry Electrode-Based Fully Isolated EEG/fNIRS Hybrid Brain-Monitoring System	Lee, Seungchan; Shin, Younghak; Kumar, Anil; Kim, Minhee; Lee, Heung-No	IEEE Transactions on Biomedical Engineering	66,4,1055–1068	2019-04	SCI	5	CA	1	4.491	ENGINEERING, BIOMEDICAL	2018	11/80	13.75	0018-9294 NN18200; NN17642	
158	60673	이종노	School of Electrical Engineering and Computer Science	158	Multimodal Sparse Representation-Based Classification Scheme for RF Fingerprinting	Yang, Kiwon; Kang, Jusung; Jang, Jehyuk; Lee, Heung-No	IEEE Communications Letters	23,5,867–870	2019-05	SCI	4	CA	1	3.457	TELECOMMUNICATIONS	2018	28/88	31.82	1089-7798	
159	60673	이종노	School of Electrical Engineering and Computer Science	159	Optimally sectioned and successively reconstructed histogram sub-equalization based gamma correction for satellite image enhancement	Singh, Himanshu; Kumar, Anil; Balyan, L. K; Lee, Heung-no	Multimedia Tools and Applications	78,14,20431–20463	2019-07	SCI	4	ETC	3	2.101	COMPUTER SCIENCE, THEORY & METHODS	2018	37/105	35.24	1380-7501	
160	60673	이종노	School of Electrical Engineering and Computer Science	160	Two-wired active spring-loaded dry electrodes for EEG measurements	Seungchan Lee; Younghak Shin; Anil Kumar; Kseon Kim; Heung-No Lee	Sensors (Switzerland)	19,20,	2019-10	SCI	5	CA	1	3.031	INSTRUMENTS & INSTRUMENTATION	2018	15/61	24.59	1424-8220	
161	60673	이종노	School of Electrical Engineering and Computer Science	161	Texture-Dependent Optimal Fractional-Order Framework for Image Quality Enhancement Through Memetic Inclusions in Cuckoo Search and Sine-Cosine Algorithms	Singh, Himanshu; Kumar, Anil; Balyan, L.K.; Lee, H.N.	Studies in Computational Intelligence	873,,19–45	2020-__	International prominent proceeding	4	ETC	3	0.00				0.00	1860-949X	
162	60673	이종노	School of Electrical Engineering and Computer Science	162	ASIC 저항성을 위한 ECCPoW 블록체인 구현 방법	정현준; 채중홍; 이종노	한국정보기술학회논문지	18,6,33–42	2020-__	KCI	3	CA	1	0.00				0.00	1598-8619	
163	60673	이종노	School of Electrical Engineering and Computer Science	163	Compressive sensing spectroscopy using a residual convolutional neural network	Kim Cheolsun; Park Dongju; Lee Heung-No	Sensors (Switzerland)	20,3,	2020-02	SCI	3	CA	1	3.275	INSTRUMENTS & INSTRUMENTATION	2019	15/64	23.44	1424-8220	
164	60673	이종노	School of Electrical Engineering and Computer Science	164	Hilbert Transform Design Based on Fractional Derivatives and Swarm Optimization	Kumar A.; Agrawal N.; Sharma I.; Lee S.; Lee, Heung-No	IEEE Transactions on Cybernetics	50,5,2311–2320	2020-05	SCI	5	CA	1	11.079	AUTOMATION & CONTROL SYSTEMS	2019	1/63	1.59	2168-2267	
165	60673	이종노	School of Electrical Engineering and Computer Science	165	ECCPoW: Error-Correction Code based Proof-of-Work for ASIC Resistance	Jung, Hyunjun; Lee, Heung-No	SYMMETRY-BASEL	12,6,	2020-06	SCI	2	CA	1	2.645	MULTIDISCIPLINARY SCIENCES	2019	29/71	40.85	2073-8994	
166	60673	이종노	School of Electrical Engineering and Computer Science	166	High-Resolution Ultrasound Imaging Using Random Interference	Ni, Pavel; Lee, Heung-No	IEEE TRANSACTIONS ON ULTRASONICS FERROELECTRICS AND FREQUENCY CONTROL	67,9,1785–1799	2020-09	SCI	2	CA	1	2.812	ACOUSTICS	2019	7/32	21.88	0885-3010	
167	60673	이종노	School of Electrical Engineering and Computer Science	167	High-Resolution Ultrasound Imaging Enabled by Random Interference and Joint Image Reconstruction	Ni, Pavel; Lee, Heung-No	SENSORS	20,22,	2020-11	SCI	2	CA	1	3.275	INSTRUMENTS & INSTRUMENTATION	2019	15/64	23.44	1424-8220	
168	60673	이종노	School of Electrical Engineering and Computer Science	168	Profitable double-spending attacks	Jang, J.; Lee, H.-N.	Applied Sciences (Switzerland)	10,23,1–23	2020-12	SCI	2	CA	1	2.474	ENGINEERING, MULTIDISCIPLINARY	2019	32/91	35.17	2076-3417	
169	60673	이종노	School of Electrical Engineering and Computer Science	169	이더리움 환경을 위한 ECCPoW 블록체인 구현 방법	정현준; 이종노	한국정보기술학회논문지	19,3,61–70	2021-__	KCI	2	CA	1	0.00				0.00	1598-8619	
170	60673	이종노	School of Electrical Engineering and Computer Science	170	Screening of COVID-19 Suspected Subjects Using Multi-Crossover Genetic Algorithm Based Dense Convolutional Neural Network	Singh, Dilbag; Kumar, Vijay; Kaur, Manjit; Jabarulla, Mohamed Yaseen; Lee, Heung-No	IEEE ACCESS	9,,142566–142580	2021-__	SCI	5	CA	1	3.367	ENGINEERING, ELECTRICAL & ELECTRONIC	2020	94/273	34.43	2169-3536	
171	60673	이종노	School of Electrical Engineering and Computer Science	171	Error-Correction Code Proof-of-Work on Ethereum	Kim, Hyoungsung; Jang, Jehyuk; Park, Sangjun; Lee, Heung-No	IEEE ACCESS	9,,135942–135952	2021-__	SCI	4	CA	1	3.367	ENGINEERING, ELECTRICAL & ELECTRONIC	2020	94/273	34.43	2169-3536	
172	60673	이종노	School of Electrical Engineering and Computer Science	172	Weighted Least Square Design Technique for Hilbert Transformer using Fractional Derivative	Agrawal, Nikhil; Kumar, Anil; Kuldeep, B.; Lee, S.; Lee, H. N.	SIGNAL IMAGE AND VIDEO PROCESSING	,,	2021-__	SCI	5	ETC	3	2.157	ENGINEERING, ELECTRICAL & ELECTRONIC	2020	159/273	58.24	1863-1703	
173	60673	이종노	School of Electrical Engineering and Computer Science	173	Blockchain-Based Distributed Patient-Centric Image Management System	Jabarulla, Mohamed Yaseen; Lee, Heung-No	APPLIED SCIENCES-BASEL	11,1,	2021-01	SCI	2	CA	1	2.679	ENGINEERING, MULTIDISCIPLINARY	2020	38/90	42.22	2076-3417	
174	60673	이종노	School of Electrical Engineering and Computer Science	174	Covert Anti-Jamming Communication Based on Gaussian Coded Modulation	Choi, Haeung; Park, Sangjun; Lee, Heung-No	APPLIED SCIENCES-BASEL	11,9,	2021-05	SCI	3	ETC	2	2.679	ENGINEERING, MULTIDISCIPLINARY	2020	38/90	42.22	2076-3417	
175	60673	이종노	School of Electrical Engineering and Computer Science	175	A Blockchain and Artificial Intelligence-Based, Patient-Centric Healthcare System for Combating the COVID-19 Pandemic: Opportunities and Applications	Jabarulla, Mohamed Yaseen; Lee, Heung-No	HEALTHCARE	9,8,	2021-08	SCI	2	CA	1	2.645	HEALTH POLICY & SERVICES	2020	40/88	45.46	2227-9032	

176	60673	이종노	School of Electrical Engineering and Computer Science	176	Radio Frequency Fingerprinting for Frequency Hopping Emitter Identification	Kang, Jusung; Shin, Younghak; Lee, Hyunku; Park, Jintae; Lee, Heungno	APPLIED SCIENCES-BASEL	11,22,	2021-11	SCI	5	CA	1	2.679	ENGINEERING, MULTIDISCIPLINARY	2020	38/90	42.22	2076-3417
177	60673	이종노	School of Electrical Engineering and Computer Science	177	Hybrid Genetic Firefly Algorithm-Based Routing Protocol for VANETs	Singh, Gagan Deep; Prateek, Manish; Kumar, Sunil; Verma, Madhushi; Singh, Dilbag; Lee, Heung-No	IEEE Access	10,,9142–9151	2022-01	SCI	6	CA	1	3.476	ENGINEERING, ELECTRICAL & ELECTRONIC	2021	105/276	38.04	2169-3536
178	60673	이종노	School of Electrical Engineering and Computer Science	178	Systematic Survey on Visually Meaningful Image Encryption Techniques	Himthani, Varsha; Dhaka, Vijaypal Singh; Kaur, Manjit; Singh, Dilbag; Lee, Heung-No	IEEE ACCESS	10,,98360–98373	2022-01	SCI	5	CA	1	3.476	ENGINEERING, ELECTRICAL & ELECTRONIC	2021	105/276	38.04	2169-3536
179	60673	이종노	School of Electrical Engineering and Computer Science	179	Privacy Protection Framework for Android	Mishra, B.; Agarwal, A.; Goel, A.; Ansari, A.A.; Gaur, P.; Singh, D.; Lee, H.-N.	IEEE Access	10,,7973–7988	2022-01	SCI	7	CA	1	3.476	ENGINEERING, ELECTRICAL & ELECTRONIC	2021	105/276	38.04	2169-3536
180	60673	이종노	School of Electrical Engineering and Computer Science	180	Efficient Stochastic Model for Operational Availability Optimization of Cooling Tower Using Metaheuristic Algorithms	Kumar, Ashish; Saini, Monika; Gupta, Nivedita; Sinwar, Deepak; Singh, Dilbag; Kaur, Manjit; Lee, Heung-No	IEEE ACCESS	10,,24659–24677	2022-01	SCI	7	CA	1	3.476	ENGINEERING, ELECTRICAL & ELECTRONIC	2021	105/276	38.04	2169-3536
181	60673	이종노	School of Electrical Engineering and Computer Science	181	Juice Jacking: Security Issues and Improvements in USB Technology	Singh, Debabrata; Biswal, Anil Kumar; Samanta, Debabrata; Singh, Dilbag; Lee, Heung-No	SUSTAINABILITY	14,2,	2022-01	SCI	5	CA	1	3.889	ENVIRONMENTAL STUDIES	2021	57/128	44.53	2071-1050
182	60673	이종노	School of Electrical Engineering and Computer Science	182	Systematic Review of Security Vulnerabilities in Ethereum Blockchain Smart Contract	Kushwaha, Satpal Singh; Joshi, Sandeep; Singh, Dilbag; Kaur, Manjit; Lee, Heung-No	IEEE Access	10,,6605–6621	2022-01	SCI	5	CA	1	3.476	ENGINEERING, ELECTRICAL & ELECTRONIC	2021	105/276	38.04	2169-3536
183	60673	이종노	School of Electrical Engineering and Computer Science	183	Data Security-Based Routing in MANETs Using Key Management Mechanism	Bondada, Praveen; Samanta, Debabrata; Kaur, Manjit; Lee, Heung-No	APPLIED SCIENCES-BASEL	12,3,	2022-02	SCI	4	CA	1	2.838	ENGINEERING, MULTIDISCIPLINARY	2021	39/92	42.39	2076-3417
184	60673	이종노	School of Electrical Engineering and Computer Science	184	6+: A Novel Approach for Building Extraction from a Medium Resolution Multi-Spectral Satellite	Dixit, M.; Chaurasia, K.; Mishra, V.K.; Singh, D.; Lee, H.-N.	Sustainability (Switzerland)	14,3,	2022-02	SCI	5	CA	1	3.889	ENVIRONMENTAL STUDIES	2021	57/128	44.53	2071-1050
185	60673	이종노	School of Electrical Engineering and Computer Science	185	Cross Channel Scripting and Code Injection Attacks on Web and Cloud-Based Applications: A Comprehensive Review	Indushree, M.; Shashidhara, R.; Raj, Manish; Kaur, Manjit; Lee, Heung-No	Sensors	22,5,	2022-03	SCI	5	CA	1	3.847	INSTRUMENTS & INSTRUMENTATION	2021	19/64	29.69	1424-8220
186	60673	이종노	School of Electrical Engineering and Computer Science	186	Two-Way Feature Extraction for Speech Emotion Recognition Using Deep Learning	Aggarwal, A.; Srivastava, A.; Agarwal, A.; Chahal, N.; Singh, D.; Alnuaim, A.A.; Alhadlaq, A.; Lee, H.-N.	Sensors	22,6,	2022-03	SCI	8	CA	1	3.847	INSTRUMENTS & INSTRUMENTATION	2021	19/64	29.69	1424-8220
187	60673	이종노	School of Electrical Engineering and Computer Science	187	ECG Data Analysis with Denoising Approach and Customized CNNs	Mishra, Abhinav; Dharahas, Ganapathiraju; Gite, Shilpa; Kotecha, Ketan; Koundal, Deepika; Zaguia, Atef; Kaur, Manjit; Lee, Heung-No	Sensors (Basel, Switzerland)	22,5,	2022-03	SCI	8	CA	1	3.847	INSTRUMENTS & INSTRUMENTATION	2021	19/64	29.69	1424-8220
188	60673	이종노	School of Electrical Engineering and Computer Science	188	Evolving Fusion-Based Visibility Restoration Model for Hazy Remote Sensing Images Using Dynamic Differential Evolution	Singh, D.; Kaur, M.; Jabarulla, M.Y.; Kumar, V.; Lee, H.-N.	IEEE Transactions on Geoscience and Remote Sensing	60,,	2022-03	SCI	5	CA	1	8.125	GEOCHEMISTRY & GEOPHYSICS	2021	5/87	5.75	0196-2892
189	60673	이종노	School of Electrical Engineering and Computer Science	189	Deep Learning-Based Computer-Aided Pneumothorax Detection Using Chest X-ray Images	Malhotra, P.; Gupta, S.; Koundal, D.; Zaguia, A.; Kaur, M.; Lee, H.-N.	Sensors	22,6,	2022-03	SCI	6	CA	1	3.847	INSTRUMENTS & INSTRUMENTATION	2021	19/64	29.69	1424-8220
190	60673	이종노	School of Electrical Engineering and Computer Science	190	Spatial Entropy Quartiles-Based Texture-Aware Fractional-Order Unsharp Masking for Visibility Enhancement of Remotely Sensed Images	Singh, Himanshu; Kumar, Anil; Balyan, L. K.; Lee, H. N.	IEEE TRANSACTIONS ON SYSTEMS MAN CYBERNETICS-SYSTEMS	54,4,2275–2288	2022-04	SCI	4	ETC	1	11.471	AUTOMATION & CONTROL SYSTEMS	2021	4/65	6.15	2168-2216
191	60673	이종노	School of Electrical Engineering and Computer Science	191	Classification and Analysis of Pistachio Species with Pre-Trained Deep Learning Models	Singh, D.; Taspinar, Y.S.; Kursun, R.; Cinar, I.; Koklu, M.; Ozkan, I.A.; Lee, H.-N.	Electronics (Switzerland)	11,7,	2022-04	SCI	7	CA	1	2.690	ENGINEERING, ELECTRICAL & ELECTRONIC	2021	139/276	50.36	2079-9292
192	60673	이종노	School of Electrical Engineering and Computer Science	192	Artificial Intelligence-Based Cyber-Physical System for Severity Classification of Chikungunya Disease	Singh, Dilbag; Kaur, Manjit; Kumar, Vijay; Jabarulla, Mohamed Yaseen; Lee, Heung-No	IEEE JOURNAL OF TRANSLATIONAL ENGINEERING IN HEALTH AND MEDICINE	10,,	2022-04	SCI	5	CA	1	2.890	ENGINEERING, BIOMEDICAL	2021	64/98	65.31	2168-2372
193	60673	이종노	School of Electrical Engineering and Computer Science	193	Detection and Classification of Colorectal Polyp Using Deep Learning	Tanwar, Sushama; Vijayalakshmi, S.; Sabharwal, Munish; Kaur, Manjit; AlZubi, Ahmad Ali; Lee, Heung-No	BioMed Research International	2022,,	2022-04	SCI	6	CA	1	3.246	BIOTECHNOLOGY & APPLIED MICROBIOLOGY	2021	90/159	56.60	2314-6133
194	60673	이종노	School of Electrical Engineering and Computer Science	194	Probabilistic Framework Allocation on Underwater Vehicular Systems Using Hydrophone Sensor Networks	Kshirsagar, Pravin R.; Manoharan, Hariprasath; Shitharth, S.; Alshareef, Abdulrhman M.; Singh, Dilbag; Lee, Heung-No	Water (Switzerland)	14,8,	2022-04	SCI	6	CA	1	3.530	WATER RESOURCES	2021	36/100	36.00	2073-4441

195	60673	이종노	School of Electrical Engineering and Computer Science	195	Efficient Fine Tuned Trapezoidal Fuzzy-Based Model for Failure Mode Effect Analysis Risk Prioritization	Subramanian, Rajesh; Taterh, Swapnesh; Singh, Dilbag; Lee, Heung-No	IEEE ACCESS	10,,50037–50046	2022-05	SCI		4	CA		1	3.476	ENGINEERING, ELECTRICAL & ELECTRONIC	2021	105/276	38.04	2169-3536	
196	60673	이종노	School of Electrical Engineering and Computer Science	196	Mining and Interpretation of Critical Aspects of Infant Health Status Using Multi-Objective Evolutionary Feature Selection Approaches	Piri, J.; Mohapatra, P.; Singh, D.; Samanta, D.; Singh, D.; Kaur, M.; Lee, H.-N.	IEEE Access	10,,32622–32638	2022-05	SCI		7	CA		1	3.476	ENGINEERING, ELECTRICAL & ELECTRONIC	2021	105/276	38.04	2169-3536	
197	60673	이종노	School of Electrical Engineering and Computer Science	197	Ethereum Smart Contract Analysis Tools: A Systematic Review	Kushwaha, S.S.; Joshi, S.; Singh, D.; Kaur, M.; Lee, Heung-No	IEEE Access	10,,57037–57062	2022-05	SCI		5	CA		1	3.476	ENGINEERING, ELECTRICAL & ELECTRONIC	2021	105/276	38.04	2169-3536	
198	60673	이종노	School of Electrical Engineering and Computer Science	198	Controlling Agronomic Variables of Saffron Crop Using IoT for Sustainable Agriculture	Kour, Kanwalpreet; Gupta, Deepali; Gupta, Kamal; Juneja, Sapna; Kaur, Manjit; Alharbi, Amal H.; Lee, Heung-No	SUSTAINABILITY	14,9,	2022-05	SCI		7	CA		1	3.889	ENVIRONMENTAL STUDIES	2021	57/128	44.53	2071-1050	
199	60673	이종노	School of Electrical Engineering and Computer Science	199	Perception Exploration on Robustness Syndromes With Pre-processing Entities Using Machine Learning Algorithm	Kshirsagar, Pravin R.; Manoharan, Hariprasath; Selvarajan, Shitharth; Alterazi, Hassan A.; Singh, Dilbag; Lee, Heung-No	FRONTIERS IN PUBLIC HEALTH	10,,	2022-06	SCI		6	CA		2	6.461	PUBLIC, ENVIRONMENTAL & OCCUPATIONAL HEALTH	2021	18/182	9.89	2296-2565	
200	60673	이종노	School of Electrical Engineering and Computer Science	200	Binary Drone Squadron Optimization Approaches for Feature Selection	Singh, Harpreet; Sharma, Suchita; Khurana, Manju; Kaur, Manjit; Lee, Heung-No	IEEE ACCESS	10,,87099–87114	2022-07	SCI		5	CA		1	3.476	ENGINEERING, ELECTRICAL & ELECTRONIC	2021	105/276	38.04	2169-3536	
201	60673	이종노	School of Electrical Engineering and Computer Science	201	An improved segmentation technique for multilevel thresholding of crop image using cuckoo search algorithm based on recursive minimum cross entropy	Kumar, Arun; Kumar, Anil; Vishwakarma, Amit; Lee, Heung-No	IET SIGNAL PROCESSING	16,6,630–649	2022-08	SCI		4	ETC		2	1.819	ENGINEERING, ELECTRICAL & ELECTRONIC	2021	195/276	70.65	1751-9675	
202	60673	이종노	School of Electrical Engineering and Computer Science	202	A Systematic Review on Acute Leukemia Detection Using Deep Learning Techniques	Raina, Rohini; Gondhi, Naveen Kumar; Chaahat; Singh, Dilbag; Kaur, Manjit; Lee, Heung-No	ARCHIVES OF COMPUTATIONAL METHODS IN ENGINEERING	30,1,251–270	2022-09	SCI		6	CA		1	8.171	MATHEMATICS, INTERDISCIPLINARY APPLICATIONS	2021	2/108	1.85	1134-3060	
203	60673	이종노	School of Electrical Engineering and Computer Science	203	Comprehensive Review of Orthogonal Regression and Its Applications in Different Domains	Pallavi; Joshi, Sandeep; Singh, Dilbag; Kaur, Manjit; Lee, Heung-No	Archives of Computational Methods in Engineering	29,6,4027–4047	2022-10	SCI		5	CA		1	8.171	MATHEMATICS, INTERDISCIPLINARY APPLICATIONS	2021	2/108	1.85	1134-3060	
204	60673	이종노	School of Electrical Engineering and Computer Science	204	Comparative performance assessment of deep learning based image steganography techniques	Himthani, Varsha; Dhaka, Vijaypal Singh; Kaur, Manjit; Rani, Geeta; Oza, Meet; Lee, Heung-No	SCIENTIFIC REPORTS	12,1,	2022-10	SCI		6	CA		1	4.997	MULTIDISCIPLINARY SCIENCES	2021	19/74	25.68	2045-2322	
205	60673	이종노	School of Electrical Engineering and Computer Science	205	A Systematic Review of Computational Image Steganography Approaches	Kaur, Sharanpreet; Singh, Surender; Kaur, Manjit; Lee, Heung-No	ARCHIVES OF COMPUTATIONAL METHODS IN ENGINEERING	29,7,4775–4797	2022-11	SCI		4	CA		1	8.171	MATHEMATICS, INTERDISCIPLINARY APPLICATIONS	2021	2/108	1.85	1134-3060	
206	60673	이종노	School of Electrical Engineering and Computer Science	206	Mass production-enabled computational spectrometers based on multilayer thin films	Kim, C.; Ni, P.; Lee, K.R.; Lee, H.-N.	Scientific Reports	12,1,	2022-12	SCI		4	CA		1	4.997	MULTIDISCIPLINARY SCIENCES	2021	19/74	25.68	2045-2322	
207	60673	이종노	School of Electrical Engineering and Computer Science	207	A Systematic Review on Acute Leukemia Detection Using Deep Learning Techniques	Raina, Rohini; Gondhi, Naveen Kumar; Chaahat; Singh, Dilbag; Kaur, Manjit; Lee, Heung-No	ARCHIVES OF COMPUTATIONAL METHODS IN ENGINEERING	30,1,251–270	2023-01	SCI		6	CA		1	8.171	MATHEMATICS, INTERDISCIPLINARY APPLICATIONS	2021	2/108	1.85	1134-3060	Y
208	60673	이종노	School of Electrical Engineering and Computer Science	208	Image reconstruction for the artificial compound eye based on deep learning	Lee, Jioh; Kim, Cheolsun; Choi, Youngin; Lee, Heung-No	Proceedings of SPIE - The International Society for Optical Engineering		2023-01	International prominent proceeding		4	CA			0.00				0.00		Y
209	60673	이종노	School of Electrical Engineering and Computer Science	209	TomFusioNet: A tomato crop analysis framework for mobile applications using the multi-objective optimization based late fusion of deep models and background elimination	Kaushik, H.; Khanna, A.; Singh, D.; Kaur, M.; Lee, Heung-No	Applied Soft Computing	133,,	2023-01	SCI		5	CA		1	8.70	COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS	2022	12/110	10.91	1568-4946	Y
210	60673	이종노	School of Electrical Engineering and Computer Science	210	ECG Signal Compression Based on Optimization of Wavelet Parameters and Threshold Levels Using Evolutionary Techniques	Singhai, Paridhi; Kumar, Anil; Ateek, A.; Ansari, Irshad Ahmad; Singh, G. K.; Lee, Heung No	CIRCUITS SYSTEMS AND SIGNAL PROCESSING	42,6,3509–3537	2023-01	SCI		6	ETC		4	2.30	ENGINEERING, ELECTRICAL & ELECTRONIC	2022	162/275	58.91	0278-081X	Y
211	60673	이종노	School of Electrical Engineering and Computer Science	211	Efficient Evolving Deep Ensemble Medical Image Captioning Network		IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS	27,2,1016–1025	2023-02	SCI		5	CA		1	7.70	MATHEMATICAL & COMPUTATIONAL BIOLOGY	2022	4/55	7.27	2168-2194	Y
212	60673	이종노	School of Electrical Engineering and Computer Science	212	An effective ECG signal compression algorithm with self controlled reconstruction quality	Pal, Hardev Singh; Kumar, A.; Vishwakarma, Amit; Singh, Girish Kumar; Lee, Heung-No	COMPUTER METHODS IN BIOMECHANICS AND BIOMEDICAL ENGINEERING	,,	2023-04	SCI		5	ETC		4	1.60	ENGINEERING, BIOMEDICAL	2022	84/96	87.50	1025-5842	Y

213	60673	이중노	School of Electrical Engineering and Computer Science	213	Social Network Analysis: A Survey on Measure, Structure, Language Information Analysis, Privacy, and Applications	Singh, Shashank Sheshar; Srivastava, Vishal; Kumar, Ajay; Tiwari, Shailendra; Singh, Dilbag; Lee, Heung-No	ACM Transactions on Asian and Low-Resource Language Information Processing	22,5,	2023-05	SCI	6	CA	1	2.00	COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	2022	113/145	77.93	2375-4699	Y
214	60673	이중노	School of Electrical Engineering and Computer Science	214	Smart Contract-Based Checkpoint for Initial PoW Network Security	Kim, Seungmin; Maeng, Gyeongdeok; Lee, Heung-No	International Conference on Ubiquitous and Future Networks, ICUFN		2023-07	International prominent proceeding	3	ETC		0.00				0.00		Y
215	60673	이중노	School of Electrical Engineering and Computer Science	215	Computational deep air quality prediction techniques: a systematic review	Kaur, Manjit; Singh, Dilbag; Jabarulla, Mohamed Yaseen; Kumar, Vijay; Kang, Jusung; Lee, Heung-No	ARTIFICIAL INTELLIGENCE REVIEW	56,SUPPL 2,2053–2098	2023-08	SCI	6	ETC	4	12.00	COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	2022	7/145	4.83	0269-2821	Y
216	60673	이중노	School of Electrical Engineering and Computer Science	216	Human Activity Recognition Using Self-Powered Sensors Based on Multilayer Bidirectional Long Short-Term Memory Networks	Su, Jian; Liao, Zhenlong; Sheng, Zhengguo; Liu, Alex X.; Singh, Dilbag; Lee, Heung-No	IEEE SENSORS JOURNAL	23,18,20633–20641	2023-09	SCI	6	ETC	5	4.30	INSTRUMENTS & INSTRUMENTATION	2022	15/63	23.81	1530-437X	Y
217	60673	이중노	School of Electrical Engineering and Computer Science	217	MLNet: Metaheuristics-Based Lightweight Deep Learning Network for Cervical Cancer Diagnosis	Kaur, Manjit; Singh, Dilbag; Kumar, Vijay; Lee, Heung-No	IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS	27,10,5004–5014	2023-10	SCI	4	CA	1	7.70	MATHEMATICAL & COMPUTATIONAL BIOLOGY	2022	4/55	7.27	2168-2194	Y
218	60673	이중노	School of Electrical Engineering and Computer Science	218	Electrocardiogram signal compression using adaptive tunable-Q wavelet transform and modified dead-zone quantizer	Pal, Hardev Singh; Kumar, A.; Vishwakarma, Amit; Lee, Heung-No	ISA Transactions	142,,335–346	2023-11	SCI	4	ETC	3	7.30	ENGINEERING, MULTIDISCIPLINARY	2022	6/90	6.67	0019-0578	Y