J. Internet Comput. Serv.
ISSN 1598-0170 (Print) / ISSN 2287-1136 (Online)
http://www.jics.or.kr

하이퍼레저 패브릭 기반의 안전한 헬스케어 데이터 관리 및 공유 플랫폼 개발 연구☆

Secure Healthcare Data Management and Sharing Platform Based on Hyperledger Fabric

최 예 진[1] 김 경 진[2*]
Ye-Jin Choi Kyoung-jin Kim

요 약

본 논문은 허가형 블록체인 시스템 기반의 헬스케어 데이터 통합 관리 및 공유 플랫폼을 소개한다. 여기서 허가형 블록체인 시스템은 하이퍼레저 패브릭을 활용한다. 하이퍼레저 패브릭을 이용하여, 환자가 본인의 데이터에 쉽게 접근할 수 있고, 기관들은 그들이 필요로 할 때 동의하에 환자 데이터를 공유할 수 있으며, 데이터 공유를 해준 환자가 보상받는 구조이다. 헬스케어 데이터는 비식별처리 통해 블록체인에 저장되며, 저장된 데이터에 대한 자세한 접근 권한을 설정하여 환자의 개인 정보를 보호한다.

본 논문에서 제안한 모델은 퍼블릭 블록체인에서 사용되는 다른 모델에 비해 강화된 보안을 제공한다. 또한, 기존 데이터를 저장하는 방법과 본 연구에서 블록체인에 저장된 데이터를 비교함으로써 환자 데이터를 보다 안전하게 저장할 수 있음을 확인한다.

☞ 주제어 : 블록체인, 하이퍼레저 패브릭, 헬스케어, 허가형 블록체인

ABSTRACT

In this paper, we present a healthcare data integration management and sharing platform based on a permissioned blockchain-based system called the Hyperledger fabric. The Hyperledger fabric allows patients to easily access their data, share the data with agencies that need it, and also reward participants. The healthcare data is stored in the blockchain by a de-identification process. Privacy is protected by setting detailed access rights to the stored data. The proposed model provides higher security than other models using a public blockchain. This study confirms that patient data can be stored more securely, by comparing the data stored in the blockchain with that from existing information storage methods.

☞ keyword : Blockchain, Hyperledger Fabric, Healthcare, Permissioned Blockchain

1. Introduction

Medical care in the past has focused on post-treatment, rather than systems to manage health and prevent diseases. In the future, there will be increased demand for preventive care and personalized health care systems. Rising health care costs will be soon reduced through digital healthcare systems utilizing information and communications technology (ICT). Based on a variety of healthcare data, future health care systems will provide customized medical services utilizing personal data, and precision medical services utilizing big data and artificial intelligence technologies[1].

Healthcare data is the key to future health care systems. Advances in ICT allow healthcare data to be collected easily and at a low cost. Healthcare data consists of clinical data, DNA data, research data, and personal health record (PHR) data[2]. Since the 2000s, electronic medical records (EMR) have been used by most hospitals and government agencies to collect and store large amounts of healthcare data. However, healthcare data has not been fully utilized, owing

1 Dept. of Future Convergence Technology Engineering, Sungshin University, Seoul, 02844, Korea.
2 Dept. of Convergence Security Engineering, Sungshin University, Seoul, 02844, Korea.
* Corresponding author (kyongjin@sungshin.ac.kr)
[Received 28 June 2019, Reviewed 15 July 2019(R2 28 August 2019), Accepted 23 December 2019]
☆ 이 성과는 2018년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2018R1C1B5039199).
☆ A preliminary version of this paper was presented at ICONI 2018.

(Table 1) Type of Blockchain

Type	Permissionless	Permissioned	
	Public	Private	Consortium
Network	Decentralized	Partially decentralized	Partially decentralized
Access	Anyone	Selected organizations	Participants in the consortium
Participants	Permissionless	Permissioned	Permissioned
Identification	Anonymous	Known Identities	Known Identities

to fragmented data and institutional legal problems[3].

In addition, there are problems regarding the property and access rights of healthcare data providers. Healthcare data is managed by hospitals and is related to the institution, not to data providers. Data providers are not guaranteed the right to view and own their data, nor are they aware of who has access to their data. Thus, it is difficult for the data provider to know, e.g., if a medical information leakage event has occurred. The advent of the digital healthcare era will require data providers to have easy access to their healthcare data, and the control power to share and utilize their data easily.

Owing to its high reliability and security, blockchain technology is being offered as a solution in the healthcare sector, as it deals with sensitive personal information. Blockchain technology will enable transactions to be performed without third parties, eliminating many procedures and reducing cost.

This paper suggests a platform for solving the existing healthcare data management problems and safely sharing and utilizing digital healthcare data, by utilizing blockchain technology.

The remainder of the paper is organized as follows. Section 2 introduces the research and case studies on blockchain technology. Section 3 addresses the problems of the healthcare platform. In Section 4, we clarify the healthcare data management and sharing platform based on blockchain. In Section 5, we verify the security of the proposed platform Finally, Section 6 presents the conclusions.

2. Related Work

2.1 Overview of Blockchain

Blockchain is a distributed ledger technology, as all users of the network jointly own the transaction records, and all the network participants own the ledger. Thus, blockchain ensures transparency and integrity, in contrast to existing centralized systems[4,5].

Blockchains are classified into permissionless or permission blockchains, based on the way network users participate. A permissionless blockchain can be freely engaged without any permission. This public blockchain is freely available to anyone, and therefore, anyone can access the data. A permission blockchain only allows authorized users to join the network. In addition, the permission blockchain is provided with control of the data, and the ability to control access to the data. Permission blockchains can be further classified into consortium blockchains or private blockchains[6]. Table 1 lists the types of blockchains.

2.2 Blockchain-Based Healthcare Platform

Yue et al. proposed a healthcare data gateway (HGD) architecture utilizing blockchain technology to enable patients to easily and securely own, control, and share healthcare data without infringing privacy[7]. Li et al. proposed a data preservation system (DPS) to prevent the theft, manipulation, and deletion of healthcare data, and proved its efficiency through an Ethereum-based prototype[8]. To maintain the confidentiality of healthcare data, research has been conducted on using blockchain technology to control access and provide integrity [9,10,11].

Gordon et al. applied blockchain technology for a patient-centric shared system of healthcare data. In addition, incentive-based participation mechanisms were studied to increase the voluntary sharing of healthcare data [12]. Kim at el. studied systems that enhance security by encrypting healthcare data through blockchain functions, providing pre-approval processes and roll-based access control of data

subjects[6].

Xia et al. used blockchain to track the usage of medical data stored in cloud and study the access control functions through smart contact[13].

Liang et al. took ownership of the data and utilized blockchain for the interoperability of various groups such as doctors, patients, and insurance companies[14]. Vishal Patel also investigated the secure sharing of medical imaging data using blockchain technology[15].

Mamoshina et al. proposed a highly-distributed storage system (HDS) that used blockchain technology to apply artificial intelligence to healthcare data[16], Wang et al. proposed a blockchain parallel healthcare systems framework to improve the accuracy and treatment methods of medical diagnosis through artificial intelligence[17]. Blockchain technology has also been studied for remote care and real-time healthcare data monitoring, storage, and utilization [18,19].

3. Problem Statement

Blockchain technology has been applied to address three issues in the healthcare field.

The first is the secure storage and management of healthcare data. The second is the provision of easy access for healthcare data providers to their data, and the strengthening of the provider's authority. The third is the efficient sharing and utilization of health care data. However, most studies focus solely on the problems of providers or users of healthcare data.

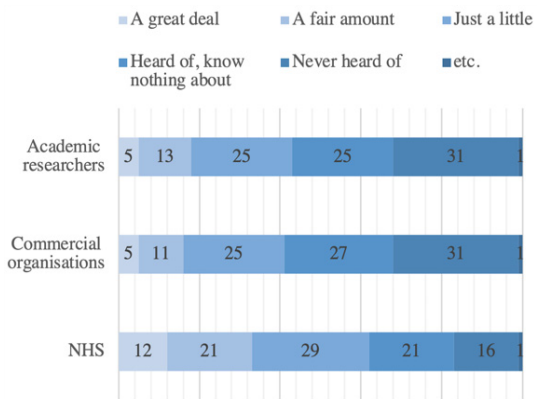
Therefore, in this paper, we address all the three abovementioned issues in healthcare data, discussed as follows.

First, we address the management and sharing of healthcare data. In Korea, more than 90% of medical institutions have introduced EMR, but less than 1% share medical information among medical institutions [20].

There is a significant amount of data, and not enough technology to securely share and manage it. There are also threats to the integrity and reliability of the healthcare data. Existing centralization methods have a risk of forgery or alteration of malicious data that is difficult to recognize. In addition, there is a possibility of data loss, owing to data damage caused by ransomware and attacks by hackers.

Second, we address the issue of ownership of the health care information. In Korea, most medical institutions have the right to manage and use healthcare data. Under the Medical Act, individuals are entitled to inspect EMR data, but there is no legal basis for ownership and access.

Therefore, it is necessary to go through several steps to read an individual's medical records during an emergency, and a patient cannot confirm the access history of his/her medical records. Data providers are often unaware of how their data is used. Figure 1 illustrates the information regarding the use of healthcare data for each institution. Only 5% of the patients knew how their healthcare data was used in commercial institutions and laboratories [21].



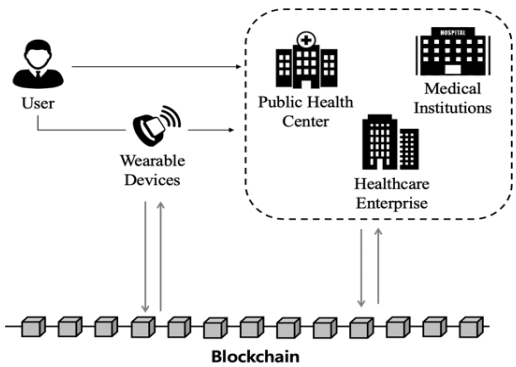
(Figure 1) Awareness of the use of healthcare data

Third, patient privacy issues must be addressed. Healthcare data stores a wide variety of data, including personal medical records, sensitive information, and life information. Recently, personal healthcare data have become more valuable to hackers than personal credit card data or social security numbers [22]. If healthcare data is leaked, there is a risk of a serious privacy invasion. However, in the present system, it is difficult for the patient, who is the subject of the data, to understand the data, even if the data is leaked owing to a hacker attack.

4. Hyperledger-Based Healthcare Data Management and Sharing Platform

In this paper, we propose a blockchain-based platform that can securely manage and share healthcare data and enhance patient rights using the Hyperledger fabric. The Hyperledger fabric shares the ledger with authorized users as an authorization blockchain. It also guarantees faster speeds and higher confidentiality than a public blockchain and provides privacy protection.

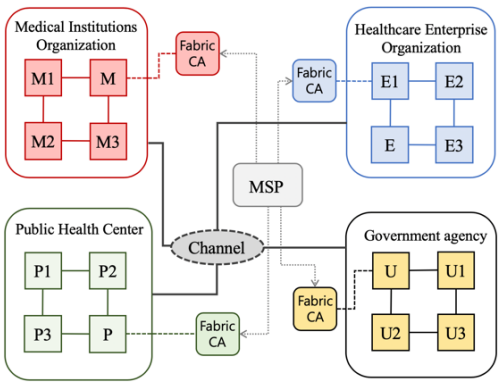
Figure 2 shows the composition of the "Hyperledger-Based Healthcare Data Management and Sharing Platform".



(Figure 2) Hyperledger-Based Healthcare Data Management and Sharing Platform

Hospitals, public health centers, healthcare enterprises, and wearable devices for managing and sharing healthcare data can all store and share data in the blockchain using the Chaincode program. The data provider and/or user can use an application to access their healthcare data recorded on the blockchain. Data providers can also share and sell data to the organizations. Figure 3 shows the Hyperledger fabric network.

Figure 3 depicts four organizations. They are: the medical institutions organization (formed by hospitals); the enterprise organization, for enterprises which require healthcare data such as research institutes, and pharmaceutical companies; the public health center organization, which is affiliated with the government; and the government agency, with general patients under the control of government regulatory agencies.



(Figure 3) Example of a healthcare platform network based on Hyperledger

- **Secure data management and sharing**

When data needs to be shared with other organizations or with only with a fraction of the members of an organization, authorized peers can create channels to create separate blockchains. This allows each agency to efficiently share data and protect the privacy of patients.

In the Hyperledger network, the Fabric certificate authority (Fabric-CA) and membership service providers (MSPs) provide secure identification and control over the access and participation of each peer. MSPs exist in networks, channels, and peers. For example, MSPs can be defined by the participating organization, and MSPs define which peers will create channels, which can create Chaincode instances, and so on. MSPs allow for the tracking of what each user has done and therefore prevent the denial of their individual behaviors. All of these processes are traceable, but the Fabric-CA protects each user's privacy.

Each peer can identify itself through the organization's Fabric-CA and can participate in each organization by obtaining a participation certification. Even after obtaining such a certification, transaction certificates continue to be received via Fabric-CA. Transaction certificates are not subject to identity disclosure. This allows users to track their behavior without disclosing their identity information.

- **Enhance Rights of data subjects**

The platform proposed herein can safeguard the privacy of the patient by safely storing data. In addition, using Chaincode and other applications, patients' abilities to access

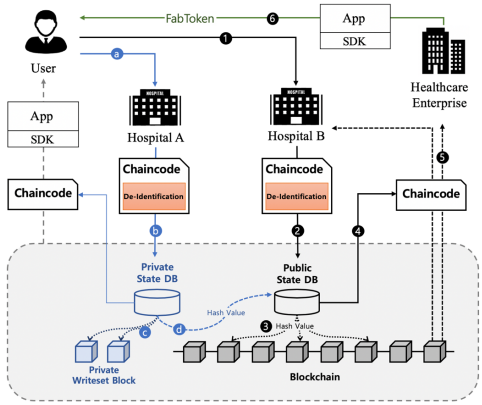
http://dx.doi.org/10.7472/jksii.2014.00.00

data and ownership are strengthened.

Using Chaincode, it is possible to conduct efficient and secure transactions of data. Data from patients recorded in the blockchain can be safely sold and shared through Chaincode. All of these processes can be verified by the data provider or patient through the application, by specifying the range of access to the data and tracking the access of each data user.

As the healthcare data of a de-identified patient is not stored in the blockchain as is, it is free from the data destruction problem inherent in the existing public blockchain. In addition, patients can receive compensation for data delivery using a Hyperledger "FabToken". The FabToken can be used as a payment tool or can be changed to a variety of assets from institutions participating in the blockchain network [23].

Figure 4 shows the safe storage and utilization processes for healthcare data on the proposed platform.



(Figure 4) Secure Healthcare Data Management and Sharing Platform

① The data provider visits the hospital and provides information. Hospitals use Chaincode to transmit patient clinical records. The patient's data will go through a de-identification process via smart contract.

② The execution results for health care data and transactions that have undergone the de-identification process are recorded in the "Public State DB".

③ After verifying the signature and endorsing policies, the transaction data and hash values of the Public State DB are stored in the blockchain.

④ To use healthcare data, the data provider and the data user enter into a contract through Chaincode.

⑤ The processes of accessing and using healthcare data after the contract is signed are also recorded in the blockchain.

⑥ Using Hyperledger's FabToken, data providers can be rewarded for providing data.

The proposed platform also provides privacy features using the "Private State DB" and "Private Writeset Block" to protect patient-sensitive data and privacy. The hash values of the content recorded in the Private State DB are recorded in the Public State DB and can ensure the reliability and integrity of the data. The Private Writeset Block is not connected to a blockchain, and shares data with authorized peers via a P2P method to help secure patient privacy.

ⓐ The data provider visits the hospital to provide healthcare data, and requests security for the data. The hospital transmits the patient's medical record data using Chaincode. The patient data is subjected to a de-identification process through a smart contract.

ⓑ The de-identification process results of the healthcare data and transactions are stored in the Private State DB, not the Public State DB.

ⓒ After verifying the signature and guarantee policy, the transaction data and Private State DB records are stored in the Private Writeset Block.

ⓓ The hash value of the Private State DB is recorded in the Public State DB to maintain the confidentiality of the data, and to provide transparency to the integrity and transaction.

5. Security Verification

5.1 Secure management and utilization of healthcare data

Bitcoin and Ethereum-based healthcare platforms allow all nodes to access data stored in a blockchain. However, the Hyperledger fabric only allows access from authorized users in the permission blockchain. It also protects personal

information and ensures confidentiality, by detailed control of access rights for authorized users through MSPs.

The Hyperledger fabric divides the different state DBs from the blockchain (which does not store data directly), making it suitable for handling sensitive healthcare data. It also ensures data integrity and reliability. As a structural feature of this Hyperledger fabric, the proposed platform can manage and store data more securely than a traditional permission blockchain.

5.2 Protecting the privacy of healthcare data providers

Medical records from hospitals and patient-generated data from patients are both stored as a distributed ledger based on blockchain. The stored data is defined as follows:

GD = {g: g is the set of general attributes, patient serial number (PSN), gender, age, and city}

MD = {m: m is the set of medical attributes, medical serial number (MSN), date, departments, doctor, diagnosis, and disease code}

Leveraging our proposed platform can enable the sharing of medical records across the Hyperledger, without compromising security. Chaincode returns results from patient-generated data with the hash function used. This key should be for a PSN on the data, i.e., Hash (GD). Chaincode similarly returns results from medical records with the hash function used. This key should be for an MSN on the data, i.e., Hash (MD).

Even if stored medical data is leaked from the Private State DB, it could not be used to identify a specific individual's disease, as the medical data {city, age, code} is the data that has been subjected to a de-identification process with an ℓ-diversity privacy model.

{g': g' is the set of critical health attributes, gender, age, city} ← GD ∩ { Quasi-IDentifier }

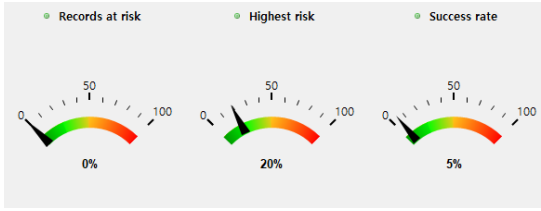
{m': m' is the set of critical health attributes, doctor, department, diagnosis, disease code} ← MD ∩ {Quasi-IDentifier, Sensitive}

We use the open-source program ARX to apply an optimal de-identification processing level with ℓ-diversity (ℓ=2). In our study, the respective processing levels of {date, departments, doctor, diagnosis, disease code} are {0, 1, 2, 6, 5}:

MD: level (0, 1, 2, 6, 5), applies de-identification processing → MD*

Classes of MD* are divided into five parts without a lost record (n = 100). This shows (a) the risk of re-identification on input general records, and (b) the risk of re-identification in de-identified data on the blockchain. As depicted in Figure 5, the risk has evidently been reduced.

(a) Risk of re-identification on input general records



(b) Risk of re-identification in de-identified data on blockchain

(Figure 5) Gap analysis between existing risk and proposed platform risk

6. Conclusion

In this paper, we proposed a platform that can solve problems in the existing healthcare ecosystem and can safely utilize healthcare data through the Hyperledger fabric platform.

The proposed model provides privacy protection via the Private State DB and Private Writeset Block. Patient data is securely stored on the blockchain through de-identification processing. To verify this, we compared the risk of re-identification of general records and de-identified data on the blockchain and confirmed that the risk of re-identification on the blockchain was clearly reduced.

http://dx.doi.org/10.7472/jksii.2014.00.00

The proposed platform will enable the safe and efficient management and use of sensitive data and personal information. It will also be able to enhance the authority of data providers and control data access protections for privacy.

As future work, we plan to study the privacy-enabled ledger of the Hyperledger fabric, and further research the ways in which the FabToken can be used.

Reference

[1] J. H. Choi, “ICT Technologies for Future Healthcare Industry”, TELCO Journal, Vol. 5, 2017, pp. 75-96.

[2] K.S Jeong and S. J. An, “Personal Health Record (PHR) Industry, Standards and Policy Trends”, TTA Journal, Vol. 164, pp.65-69, 2016.

[3] OECD, “OECD Health Policy Studies: Health Data Governance”, 2015.

[4] H. Kang, H. R. Kim and S. Hong, ″A Study on the De sign of Smart Contracts mechanism based on the Block chain for anti-money laundering,″ Journal of Internet C omputing and Services, vol. 19, no. 5, pp. 1-11, 2018. http://dx.doi.org/10.7472/jksii.2018.19.5.1.

[5] S. Hwang and H. Lee, ″Identification of Counterfeit Android Malware Apps using Hyperledger Fabric Blockchain,″ Journal of Internet Computing and Services, vol. 20, no. 2, pp. 61-68, 2019. http://dx.doi.org/10.7472/jksii.2019.20.2.61.

[6] K. Kim and S. Hong, ″A Trusted Sharing Model for Pa tient Records based on Permissioned Blockchain,″ Jour nal of Internet Computing and Services, vol. 18, no. 6, pp. 75-84, 2017. http://dx.doi.org/10.7472/jksii.2017.18.6.75.

[7] X. Yue, H. Wang, D. Jin, M. Li, and W. Jiang, “Healthcare Data Gateways: Found Healthcare Intelligence on Blockchain with Novel Privacy Risk Control,” J. Med. Syst., vol. 40, no. 10, 2016. http://dx.doi.org/10.1007/s10916-016-0574-6

[8] H. Li, L. Zhu, M. Shen, F. Gao, X. Tao, and S. Liu, “Blockchain-Based Data Preservation System for Medical Data,” J. Med. Syst., vol. 42, no. 8, pp. 1–13, 2018. http://dx.doi.org/10.1007/s10916-018-0997-3

[9] Guo, R., Shi, H., Zhao, Q., & Zheng, D., “Secure attribute-based signature scheme with multiple authorities for blockchain in electronic health records systems”, IEEE Access, vol. 6, pp.11676-11686. 2018.

[10] Dagher, G. G., Mohler, J., Milojkovic, M., & Marella, P. B., “Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology”. Sustainable cities and society, vol.39, pp.283-297.2018

[11] Al Omar, A., Rahman, M. S., Basu, A., and Kiyomoto, S., “Medibchain: A blockchain based privacy preserving platform for healthcare data”, In International conference on security, privacy and anonymity in computation, communication and storage. Springer, pp. 534-543, 2017.

[12] W. J. Gordon and C. Catalini, “Blockchain Technology for Healthcare: Facilitating the Transition to Patient-Driven Interoperability,” Comput. Struct. Biotechnol. J., vol. 16, pp. 224–230, 2018. http://dx.doi.org/10.1016/j.csbj.2018.06.003

[13] M. Guizani, E. B. Sifah, J. Gao, Q. Xia, K. O. Asamoa h, and X. Du, “MeDShare: Trust-Less Medical Data Sh aring Among Cloud Service Providers via Blockchain,” IEEE Access, vol. 5, pp. 14757–14767, 2017. http://d x.doi.org/10.1109/access.2017.2730843

[14] X. Liang, J. Zhao, S. Shetty, J. Liu, and D. Li, “Integrating blockchain for data sharing and collaboration in mobile healthcare applications,” IEEE Int. Symp. Pers. Indoor Mob. Radio Commun. PIMRC, vol. 2017-Octob, pp. 1–5, 2018.

[15] V. Patel, “A framework for secure and decentralized sharing of medical imaging data via blockchain consensus,” Health Informatics J., 2018.

[16] L. Ojomoko et al., “Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare,” Oncotarget, vol. 9, no. 5, pp. 5665–5690, 2017.

[17] S. Wang et al., “Blockchain-Powered Parallel Healthcare Systems Based on the ACP Approach,” IEEE Trans. Comput. Soc. Syst., vol. 5, no. 4, pp. 942–950, 2018.

[18] M. A. Uddin, A. Stranieri, I. Gondal, and V.

Balasubramanian, “Continuous Patient Monitoring with a Patient Centric Agent: A Block Architecture,” IEEE Access, vol. 6, pp. 32700 – 32726, 2018.

[19] S. Chakraborty, S. Aich, and H. C. Kim, “A Secure Healthcare System Design Framework using Blockchain Technology,” Int. Conf. Adv. Commun. Technol. ICACT, vol. 2019. pp. 260 – 264, 2019.

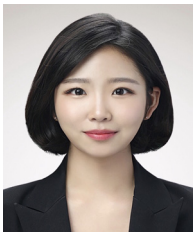
[20] D. Lee and S.K Kim, “Trends of Digital Healthcare Innovation and Policy Implications”, STEPT Science & Technology Policy, vol.48, 2018.

[21] Rumbold, J. M. M., and Pierscionek, B., “The effect of the general data protection regulation on medical research”. Journal of medical Internet research, vol.19, no.2, 2017.

[22] S.S. Baek. “Blockchain-based Electronic Medical Record Sharing Framework Using Ciphertext Policy Attribute-Based Cryptography for patient’s anonymity”, Convergence security journal, vol.19, no.1, pp.49-60, 2019.

[23] Hyperledger Fabric, “FABToken” https://hyperledger-fabric.readthedocs.io/en/latest/token/FabToken.html

◐ 저 자 소 개 ◑



최 예 진(Ye-Jin Choi)
2018년 2월 : 성신여자대학교 정보시스템공학과 (공학사)
2018년 3월~현재 : 성신여자대학교 미래융합기술공학과 석사과정
관심분야 : 블록체인, 개인정보보호, 정보보호 etc.
E-mail : hs1yejin@gmail.com

김 경 진(Kyoung-jin Kim)
2007년 성신여자대학교 컴퓨터정보학부 졸업 (공학사)
2009년 성신여자대학교 대학원 전산학과 (이학석사)
2013년 성신여자대학교 대학원 컴퓨터학과 (이학박사)
2013년 3월～2015년 8월 성신여자대학교 컴퓨터학과 박사후연구원
2015년 9월～2017년 2월 서강대학교 스마트 핀테크 연구센터 박사후연구원
2017년 3월～현재 성신여자대학교 융합보안공학과 교수
관심분야 : 블록체인, 접근제어, 프라이버시 보호 etc.
E-mail : kyongjin@sungshin.ac.kr