

# **Blockchain Technology Based Information Classification Management Service**

Gi-Wan Hong<sup>1</sup>, Jeong-Wook Kim<sup>1</sup> and Hangbae Chang<sup>2,\*</sup>

<sup>1</sup>Department of Security Convergence, Graduate School, Chung-Ang University, Seoul, 06974, Korea 
<sup>2</sup>Department of Industrial Security, Chung-Ang University, Seoul, 06974, Korea 
<sup>\*</sup>Corresponding Author: Hangbae Chang. Email: hbchang@cau.ac.kr
Received: 03 August 2020; Accepted: 11 November 2020

**Abstract:** Hyper-connectivity in Industry 4.0 has resulted in not only a rapid increase in the amount of information, but also the expansion of areas and assets to be protected. In terms of information security, it has led to an enormous economic cost due to the various and numerous security solutions used in protecting the increased assets. Also, it has caused difficulties in managing those issues due to reasons such as mutual interference, countless security events and logs' data, etc. Within this security environment, an organization should identify and classify assets based on the value of data and their security perspective, and then apply appropriate protection measures according to the assets' security classification for effective security management. But there are still difficulties stemming from the need to manage numerous security solutions in order to protect the classified assets. In this paper, we propose an information classification management service based on blockchain, which presents and uses a model of the value of data and the security perspective. It records transactions of classifying assets and managing assets by each class in a distributed ledger of blockchain. The proposed service reduces assets to be protected and security solutions to be applied, and provides security measures at the platform level rather than individual security solutions, by using blockchain. In the rapidly changing security environment of Industry 4.0, this proposed service enables economic security, provides a new integrated security platform, and demonstrates service value.

**Keywords:** Information classification; data integrity; document security; blockchain; CIA

### 1 Introduction

In Industry 4.0, people and people, people and things, things and things are becoming hyperconnected rapidly and closely, unlike before. The advancement of hardware and the emergence of new Information Technology (IT) technologies such as cloud, 5G, blockchain, and artificial intelligence (AI) are creating new values and improving the quality of life based on big data collection in the existing services [1,2]. However, these developments are expanding the scope of



security for organizations and increasing rapidly the number of assets to be protected. As the number of security solutions implemented by organizations increases due to the assets that need to be protected, the economic cost of information security increases while it is becoming difficult to manage various security solutions.

Within this security environment, organizations must identify and classify assets by the value of data and the security perspective and apply appropriate security measures upon classification for effective security management. Differentiated security activities that apply strong protection measures to assets with important values and relatively weak protection measures to assets that are less important can enable organizations to perform effective security activities in terms of cost and time. In performing differentiated security activities by identifying and classifying assets, the targets of assets are variously identified, such as organization members, products, and IT devices.

Since the introduction of the organization's information system in the 2000's, most of the important assets possessed by the organization have been generated and distributed in the form of electronic document (files). As a measure to protect documents that contain sensitive information, the solutions of Digital Rights Management (DRM), Data Loss Prevention (DLP), document centralization, etc. exist, and several security solutions are used together heuristically in accordance with organizations. It leads organizations to increase security costs and makes managing security more complicated.

To cope with these security environments effectively, we need to enable economical security for an increasing number of assets and security solutions, and integrated security management from a platform perspective. In this paper, we propose information classification management service based on blockchain which presents and utilizes a model to classify the assets based on the data and security perspective. It records transactions of classifying assets and managing assets by each class in a distributed ledger of blockchain. The proposed service reduces digital content assets to be protected and security solutions to be applied by classifying based on the given model and using blockchain to provide security measures at the platform level rather than individual security solutions level.

The major contributions of this paper are listed below:

- It enables economical security by classifying assets depending on the presented model and reducing security cost by decreasing the assets to be protected through differentiated security activities by each class.
- It provides security of an integrated platform perspective rather than an individual solution perspective by utilizing blockchain technology to provide traceability and accountability of classification management rather than adding a new security solution.
- By incorporating blockchain technology into information security, it proves its value as a new service model based on blockchain.

The rest of this paper is organized as follows: Section 2 reviews the related works. Section 3 discusses in detail the possibility of applying a new information classification model and blockchain technology to the existing security environment and proposes a research model. Section 4 shows performance analysis of the proposed research model. Finally, conclusions and future works of the paper are presented in Section 5.

## 2 Related Works

## 2.1 Current Document Security Solutions

Due to political, economic, and social factors, the number of industrial technology leaks is rapidly increasing every year, but measures against countermeasures and recovery are insufficient, and punishment is low. In this situation, an organization is conducting document security activities by treating most documents as 1st class documents for the purpose of strengthening fragmentary document security activities or prevailing in legal disputes due to technology leakage accidents that may occur later.

In the technical aspect, various security technologies have been commercialized for document security. The document centralization technology is to automatically (or forcibly) move files created on a Personal Computer (PC) to a central server after a certain period. Through these procedures, document centralization has the advantage of systematically classifying and storing files to prevent illegal leakage of files and to increase the utilization of documents [3]. However, in order to classify and store files, the inconvenience of having to register the information of the corresponding document also coexists.

DRM technology is a technology that encrypts files and allows them to be decrypted and used only by authorized persons. It has the advantage that files can be encrypted and protected even when leaked to the outside. However, there is a disadvantage that it is necessary to set the access authority of files by classifying all the various departments and members of the company. If it needs to share a file with a person who does not have permission to use it, the security must be sent to the decrypted file. At this time, it is accompanied by inconvenience such as the need to additionally acquire the approval of the upper-ranking person.

DLP is a technology that prevents data breaches and detects confidential data transmitted, used, or stored at endpoints based on specific detection techniques such as data matching, rule and regular expression matching, and keywords [4]. DLP technology prevents personal identification information such as customer or employee records, corporate information, and confidential information including intellectual property from leaking out of the company through the combination of company members, processes, and technology [5]. There are some cases of reluctance to introduce DLP because there is a perception that DLP decreases productivity, has no effect for Small and Medium-sized Enterprises (SMEs), or it is difficult to implement and takes a long time [6].

# 2.2 Information Classification of Documents

For individuals and organizations, information security is a critical issue because it causes financial losses. Therefore, companies must accurately identify and focus on relatively important assets. The purpose of classification information in a company is to perform cost-effective security activities through differential security policies based on importance after evaluation.

There is confidentiality, integrity and availability (CIA) based classification methods. Confidentiality refers to keeping information confidential, integrity keeping information immutable, and availability refers to using information immediately, regardless of geographic or temporal constraints [7]. The security class of the document is evaluated for each area of confidentiality, integrity and availability, and the total impact is calculated to reflect this, and the classes are classified. This approach has limitations that do not reflect the importance of information and various values and business processes [8].

Based on the understanding of corporate information, the information classification standard enables an evaluation by defining the concept of corporate information, calculating scope, and integrating types [9]. This study derived the evaluation factors of information assets such as information development and maintenance cost, information grade, information sharing utilization, internal damage and external risk in case of information leakage.

# 2.3 Blockchain Technology and Service

Blockchain is a technology in which network participants record and share transaction data in a distributed ledger [10]. Unlike the existing centralized method, network participants share data in a decentralized method, and data is stored in blocks and connected like a chain through a hash value, making it almost impossible to modify or delete the recorded data. Blockchain technology consists of distributed network communication technology, consensus algorithm, smart contract, digital signature, hash algorithm, and distributed ledger technology.

Blockchain networks can be classified into public, permissioned (private), and consortium blockchains according to their characteristics, and policies. The public blockchain is a network model in which all participants can trade and share data, such as Bitcoin and Ethereum. Permissioned blockchain is a network model in which only authorized participants can transact and share data. Consortium blockchain is a network model in which public and permissioned types are mixed.

Blockchain 1.0 appeared as Bitcoin and grew greatly, but it had limitations such as limited application to the financial field, slow transaction speed, high social cost, etc. [11]. Blockchain 2.0 has developed a smart contract function and has begun to be applied to other fields as well as finance through Ethereum [12]. In transactions or contracts, a smart contract allows transactions to be automatically executed when conditions are met through the smart contract function without intermediary intervention by a central agency or a third party. In Blockchain 3.0, various platforms such as EOS, ADA, Hyperledger, and Quorum have appeared, and various consensus algorithms have been developed to solve the hard fork problem and low Transaction Per Second (TPS) problem that occurred in Blockchain 2.0.

Blockchain technology is currently in Blockchain 4.0 and is being applied to various industries with new service development. In addition, global IT companies are offering Blockchain as a service (BaaS) that combines a blockchain platform and cloud services [13]. BaaS facilitates the development and deployment of decentralized applications (dApps) in a blockchain network configuration, and facilitates network resource management, monitoring and policy management. Blockchain-based services have business characteristics such as decentralization, irreversibility, traceability, accountability, integrity etc. depending on the form of network and platform [14].

The goal of information security is to ensure the CIA of valuable information assets that can be sensitive information. Blockchain technology is composed of distributed networks, distributed ledgers, consensus algorithms, smart contracts, public key encryption, digital signatures, hash algorithms, etc., thereby enhancing information security.

Blockchain is designed to transparently share data between participants through a distributed ledger, but Confidentiality can also be ensured using zero-knowledge proofs and smart contracts [15]. In addition, a service can be implemented to maintain confidentiality from a system or protocol point of view [16–18]. The principle of information storage of the blockchain is that the data to be recorded on the blockchain is stored in a block, and the block is connected like a chain in the form of a double linked list using the hash value of the block, and the connected chain is

shared among participants through a distributed network. Based on this process, the blockchain guarantees a high level of trust in integrity. Blockchain is composed of a distributed network, so even if one node fails, the service can be used through another node, so availability can be guaranteed. Using these technical characteristics of the blockchain, there is a research shows the feasibility of applying public services from the perspective of CIA standards [19].

# 3 Proposed Model

# 3.1 Consideration about Information Classification and Blockchain Technology

In the information security environment, security costs and difficulty in managing security solutions are augmented with the rapidly increasing protection assets through hyper-connectivity. We aim to apply differential protection measures by each class for effective classification management with classifying assets in terms of data value and security. Furthermore, we intend to ensure CIA of information security more strongly and traceability for classification management, by utilizing blockchain technology from a platform point of view.

We use the information classification standard for the classification of a document. The details of the information classification are given in Tab. 1. The weights for each item shown in Tab. 1 were randomly calculated. The class of information is calculated based on the total score derived from each item. Since this information classification is performed in considerations of each item, the value of the information can be reflected worthily in terms of security. This also enables organizations to classify information security assets in more effective way unlike CIA, and thus can have real value in a business environment.

Information items Evaluation content Weight Evaluation score 1. Information creation Committed to developing and 10 (A) maintaining information and maintenance cost Evaluate required efforts (personnel, time, funds, etc.) as a whole Evaluate the quality and level of 2. Output information 15 (B) information (accuracy, novelty, grade usability, etc.) Evaluate the frequency and extent of 3. Information 10 (C) use of the information utilization 4. Internal utilization Evaluate the extent to which the 35 (D) effect information has contributed to the company's value (revenue) creation 5. Risk of external leak Evaluate the possibility of operating 30 (E) loss due to information leakage to other companies Total score A + B + C + D + E

**Table 1:** Information classification model

In order to use blockchain technology for information classification management services by each class, the CIA triad of information security must be guaranteed. In the perspective of confidentiality, since the blockchain guarantees data transparency by sharing a distributed ledger among network participants, a separate method for maintaining confidentiality must be needed.

Depending on the type of blockchain network, public blockchain provides data transparency to everyone participating in the transaction, but permissioned and consortium blockchain provides limited data transparency only to trusted participants and ensures confidentiality to non-participants. In addition, if the contents of a sensitive asset are not stored in the distributed ledger of the blockchain, it is not necessary to guarantee the confidentiality of the asset.

Integrity is to keep the data recorded on the blockchain as original and not forged. It is difficult to forge data stored in a distributed ledger due to the irreversibility of the blockchain. It also can be verified that the data has not been changed by comparing the hash value of the block [20].

Availability means that even if a node participating in the blockchain network fails, the service should be available. As the distributed network of the blockchain operates based on several nodes, even if a failure occurs, it can secure availability by providing services to users through other nodes that are operating [19].

Blockchain-based information classification management service can establish a permissioned blockchain network to secure confidentiality, and to record only the hash value, not the content of sensitive assets, on the distributed ledger. It can provide traceability for security activities by including only the hash value of assets. Furthermore, for integrity and availability, the service can be implemented in a way to compare the hash value on the distributed ledger and prevent service interruption due to failure by configuring multiple nodes of the blockchain.

# 3.2 Designing Blockchain Technology Based Information Classification Service

In order to grasp the importance of documents when an organization performs business, and apply protection measures accordingly, this paper classifies the documents based on the information classification model presented in Fig. 1. First, the assets produced by the business activities of organizations are uploaded to the system and are classified according to the information classification model. Users can classify documents by the classification model. There exist 1st, 2nd, and 3rd classes within the information classification model. Second, documents corresponding to 1st class enable classification management traceability by recording the metadata of assets and the class information to the blockchain. Third, document classification management is performed by applying security solutions to 1st and 2nd class documents. 3rd class is public and does not carry out any security measures. Fourth, access to classified assets. Fifth, if the accessing asset is 1st class, the integrity of the asset is verified by checking the hash value.

# 3.3 Blockchain Technology Based Classification Management Method

Through the information classification model, if the score is 75 points or higher, it is the 1st class, if it is 50 or higher, the 2nd class, and the rest is the 3rd class. The differential classification management plan by the class is as shown in Tab. 2. The documents corresponding to the 1st class confidential information upload the security class and meta information of documents to the blockchain system to enable classification management traceability, and the information recorded on the blockchain guarantees integrity and verifies security activities through a shared distributed ledger. The DRM solution is applied to the 1st and 2nd class documents for document security. 2nd class documents are not recorded on the blockchain. 3rd class documents are public information and do not carry out security measures.

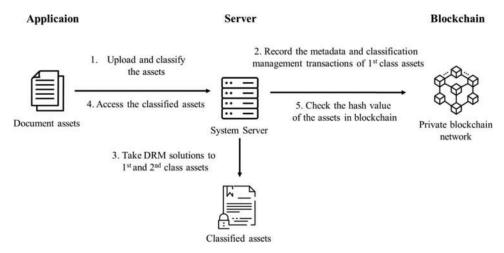


Figure 1: Blockchain-based information classification management service flow

Class Score Differential classification management 1st Confidential 75 points or more Document meta data information record on blockchain, DRM solution 2nd Internal use 50 points or more DRM solution information 3rd Public information Less than 50 points Nothing (Public)

Table 2: Differential classification management method

# 3.4 Blockchain Technology Service Structure

The block diagram of the blockchain service based on the designed service is shown in Fig. 2 below. Each organization's system consists of a document management server, web application, DBMS and DRM solution, and the entire system consists of blockchain network and organization's system. The system of each organizations implements a security service for each class, providing document upload, download, information classification, DRM encryption, decryption, blockchain record, inquiry, and monitoring functions. Organizations participating in the blockchain network can verify document security activity by sharing the 1st class document traceability as block information recorded in the distributed ledger. Organization, member information, document information, classification, download, DRM transaction, etc. are stored in DBMS.

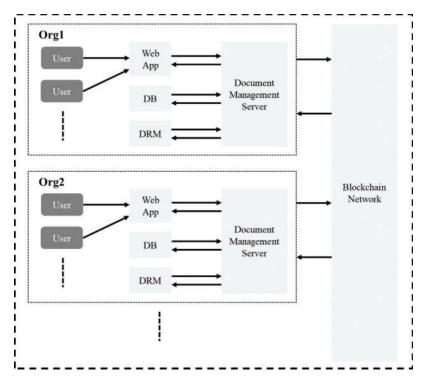


Figure 2: Blockchain service structure diagram

## 4 Performance Analysis

We intend to implement the blockchain based information classification management service proposed in this paper and discuss the results and the contributions.

#### 4.1 Experimental Environment

As a blockchain platform, the open source based Hyperledger fabric platform developed in the Hyperledger project of the Linux Foundation was used, and features such as distributed ledger, smart contract [21], consensus algorithm [22], blockchain network construction, and encryption were used. Hyperledger Fabric implements the configuration node as a docker image through a container. Smart contracts are developed through chaincode, chaincode can also be managed by version, and work through Docker containers.

The blockchain network of the security service for each blockchain-based information classification has built an environment on the Ubuntu 18.04 LTS operating system [23], and it is possible to allocate two logical business organizations (Org1, Org2) and two peers, each as a docker container. So that it was implemented. Order of node was set to 1, and CouchDB supported by Hyperledger Fabric was used as the distributed ledger. HFC and chaincode are implemented through node.js and Go programming languages. The client program that provides an interface to upload/download documents and centrally manage documents was implemented as a web application.

## 4.2 Results and Discussions

It is the main screen of the web application of the implemented service in Fig. 3 below. The ability to upload your organization's document assets is implemented on the left side of the screen.

On the right, you can see the documents uploaded in the same organization. Document number, document name, author, creation date, document security level, blockchain record status, DRM status can be checked, and document download is available.



Figure 3: Service main view

It is an information classification view in Fig. 4 below. The evaluation items were implemented based on the information classification model. Each evaluation item can be evaluated as good, average, or bad. The progress of the evaluation is indicated on the horizontal bar below.

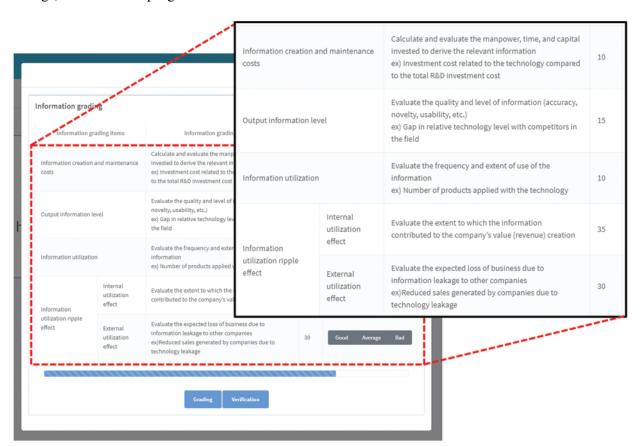


Figure 4: Information classification view

After completing the evaluation of the classification and the upload to the system, if the document is a 1st class, the meta information of the document and the hash value information of the document are recorded on the blockchain. It shows a block in the distributed ledger shared under the channel of the blockchain implemented using the Hyperledger fabric platform in Fig. 5 below. A block is composed of document meta information and hash value. Meta information includes access type, upload date, document description, file name, author, file size, upload result, and class.

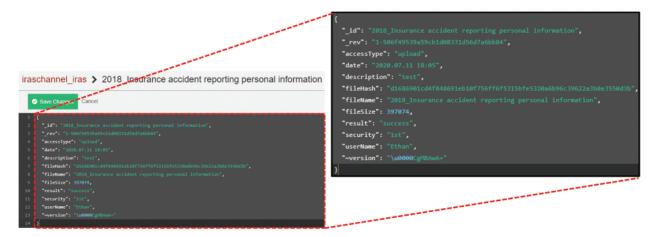


Figure 5: Document class record on blockchain

In addition, Hyperledger Fabric Explorer, which can monitor the blockchain network, is implemented as shown in Fig. 6 below. It shows the number of blocks recorded on the blockchain, the number of transactions issued, the number of nodes participating in the network, the number of chain codes, the peer name, the number of blocks per hour and per minute, the number of transactions per hour and per minute and etc.



Figure 6: Monitoring blockchain network

The proposed service provides traceability of classification management by recording the asset's meta information and class information in the blockchain for assets classified as 1st class in Figs. 4 and 5. By providing differentiated security services through classification considering the value and security factors of data in the increasing number of assets, it is possible to reduce the targets to be protected without managing all assets confidentially. In addition, it does not use various individual document security solutions, but provides security from a platform perspective by providing traceability of classification management using blockchain technology.

We consider that the proposed service needs to observe confidentiality, integrity, and availability (CIA), which can be referred to as the CIA requirement of information security. By implementing the service using the Hyperledger Fabric platform, a permissioned blockchain network, only trusted nodes shown in Fig. 6 can access the blockchain network. As distributed ledger data can be read only by trusted targets, it ensures confidentiality of the data. In addition, in the data of the distributed ledger, only the meta-information and the security level of the asset are recorded as shown in Fig. 5, so that sensitive contents of assets to be kept confidential are not recorded. We can check the hash values of assets, blocks, and transactions, because data contains the unique hash value of the asset and the block data hash value and transaction hash value are included in the blockchain as shown in Figs. 5 and 6. So, the service can prove integrity by verifying data authenticity through hash values for assets, blocks, and transactions. The blockchain network consists of two organizations and four peer nodes as shown in Fig. 6, so even if one node fails, the service can be provided through another node, thereby ensuring availability. Therefore, we can confirm that the proposed service is valuable as a security service by satisfying the three factors of confidentiality, integrity and availability of information security.

## 5 Conclusions and Future Works

In order to carry out the economical security activities of an organization, we proposed a service to classify the security class of documents based on the information contained in the documents, perform differential security activities and provide classification management traceability using blockchain technology. It is expected to enable economic security activities through the differential management method, and to provide a classification management platform by using blockchain technology. In addition, we examined the validity of the service from the viewpoint of the CIA element of information security. As a future study, we will research to develop an integrated security technique from a platform perspective by connecting endpoints between security solutions and the proposed system and verify effectiveness of the service. We will also research methods to detect abnormal behavior of classification management using the traceability of the blockchain for enable enterprise-wide classification management.

**Funding Statement:** This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2020-2018-0-01799) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation).

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

### References

- [1] S. Namasudra, P. Roy, P. Vijayakumar, S. Audithan and B. Balamurugan, "Time efficient secure DNA based access control model for cloud computing environment," *Future Generation Computer Systems*, vol. 73, pp. 90–105, 2017.
- [2] S. Namasudra, "An improved attribute-based encryption technique towards the data security in cloud computing," *Concurrency and Computation: Practice and Exercise*, vol. 31, no. 3, pp. 1–15, 2019.
- [3] Y. Yoon and Y. Kim, "An operation of cloud system for the centralization services of business documents," *Journal of Information Technology Services*, vol. 13, no. 4, pp. 309–324, 2014.
- [4] S. Liu and R. Kuhn, "Data loss prevention," IT Professional, vol. 12, no. 2, pp. 10–13, 2010.
- [5] J. U. Choi, Y. J. Lee and J. M. Park, "E-DRM-based privacy protection technology for overcoming technical limitations of DLP-based solutions," *Journal of the Korea Institute of Information Security & Cryptology*, vol. 22, no. 5, pp. 1103–1113, 2012.
- [6] S. J. Yoo, "A study on DLP system for preventing internal information leakage," *Convergence Security Journal*, vol. 18, no. 5, pp. 121–126, 2018.
- [7] J. M. Anderson, "Why we need a new definition of information security," *Computers & Security*, vol. 22, no. 4, pp. 308–313, 2003.
- [8] M. R. McGurk and J. W. Lu, "Intersection of patents and trade secrets," *Hastings Science & Technology Law Journal*, vol. 7, pp. 189–214, 2015.
- [9] O. Na, L. W. Park, H. Yu, Y. Kim and H. Chang, "The rating model of corporate information for economic security activities," *Security Journal*, vol. 32, no. 4, pp. 435–456, 2019.
- [10] M. Nofer, P. Gomber, O. Hinz and D. Schiereck, "Blockchain," *Business & Information Systems Engineering*, vol. 59, no. 3, pp. 183–187, 2017.
- [11] M. Mainelli and M. Smith, "Sharing ledgers for sharing economies: An exploration of mutual distributed ledgers (aka blockchain technology)," *Journal of Financial Perspectives*, vol. 3, no. 3, pp. 38–58, 2015.
- [12] M. Xu, X. Chen and G. Kou, "A systematic review of blockchain," *Financial Innovation*, vol. 5, no. 1, pp. 27, 2019.
- [13] W. Zheng, Z. Zheng, X. Chen, K. Dai, P. Li *et al.*, "Nutbaas: A blockchain-as-a-service platform," *IEEE Access*, vol. 7, pp. 134422–134433, 2019.
- [14] S. Namasudra, G. C. Deka, P. Johri, M. Hosseinpour and A. H. Gandomi, "The revolution of blockchain: State-of-the-art and research challenges," *Archives of Computational Methods in Engineering*, vol. 6, no. 2, pp. 20632, 2020.
- [15] Y. Wang and A. Kogan, "Designing confidentiality-preserving blockchain-based transaction processing systems," *International Journal of Accounting Information Systems*, vol. 30, pp. 1–18, 2018.
- [16] S. Namasudra, R. Chakraborty, S. Kadry, G. Manogaran and B. S. Rawal, "FAST: Fast accessing scheme for data transmission in cloud computing," *Peer-to-Peer Networking and Applications*, vol. 4, no. 4, pp. 1–13, 2020.
- [17] S. Namasudra, R. Chakraborty, A. Majumder and N. R. Moparthi, "Securing multimedia by using DNA based encryption in the cloud computing environment," *ACM Transactions on Multimedia Computing Communications and Applications*, 2020.
- [18] S. Namasudra and P. Roy, "Time saving protocol for data accessing in cloud computing," *IET Communications*, vol. 11, no. 10, pp. 1558–1565, 2017.
- [19] M. Warkentin and C. Orgeron, "Using the security triad to assess blockchain technology in public sector applications," *International Journal of Information Management*, vol. 52, 102090, 2020.
- [20] A. Dorri, M. Steger, S. S. Kanhere and R. Jurdak, "Blockchain: A distributed solution to automotive security and privacy," *IEEE Communications Magazine*, vol. 55, no. 12, 119–125, 2017.

- [21] W. Hu, Z. Fan and Y. Gao, "Research on smart contract optimization method on blockchain," *IT Professional*, vol. 21, no. 5, pp. 33–38, 2019.
- [22] Y. Manevich, A. Barger and Y. Tock, "Endorsement in hyperledger fabric via service discovery," *IBM Journal of Research and Development*, vol. 63, no. 2/3, pp. 2–1, 2019.
- [23] J. H. Park and J. H. Park, "Blockchain security in cloud computing: Use cases, challenges, and solutions," *Symmetry*, vol. 9, no. 8, pp. 164, 2017.