ICCEPM 2022
The 9th International Conference on Construction Engineering and Project Management
Jun. 20-23, 2022, Las Vegas, NV, USA

Automatic Linkage Method Between Email and Block Structure to Store Construction Project Documents in The Blockchain

Eu Wang Kim[1], Min Seo Park[2], Jong Inn Kim[3], Ameng Wei[4], Kyoungmin Kim[5], Kyong Ju Kim[6]*

[1] *Department of Smart cities, Chung-Ang University, Seoul 06974, Republic of Korea,* E-mail address: kew1013@cau.ac.kr
[2] *Department of Smart cities, Chung-Ang University, Seoul 06974, Republic of Korea,* E-mail address: garfieldyy@cau.ac.kr
[3] *Department of Smart cities, Chung-Ang University, Seoul 06974, Republic of Korea,* E-mail address: sscolgy2@naver.com
[4] *Department of Civil & Environmental Engineering, Chung-Ang University, Seoul 06974, Republic of Korea,* E-mail address: weiameng@naver.com
[5] *Department of Civil & Environmental Engineering, Chung-Ang University, Seoul 06974, Republic of Korea,* E-mail address: kmkim75@cau.ac.kr
[6] *Department of Civil & Environmental Engineering, Chung-Ang University, Seoul 06974, Republic of Korea,* E-mail address: kjkim@cau.ac.kr

Abstract: In construction projects, it is common to exchange documents using email because of convenience. In this study, a method extracting and organizing block information automatically based on email was developed. This method is composed of document exchange and archiving processes, which are difficult to manage and vulnerable to loss. Therefore, this study aims to develop a solution that can automatically link email and block information. The block data components are designed to derive from email exchange and user-additional input information. Also, automatically generating blocks process including extraction and conversion of information was proposed. This solution can lead to promote the convenience of project document management in terms of identifying the document flow and preventing loss of information.

Key words: document storing, document retrieval, email, blockchain, block data

1. INTRODUCTION

A massive number of documents of various types are generated during long-term construction project. The management of these documents is difficult, given the different types of stakeholders involved, such as employers, engineers, contractors, subcontractors, and vendors. Email has become a common tool for communication in construction projects because it is ease to use. The importance of email usage has increased in the area of communicating and storing major information, given that email can be employed as a technique to track, store, and extract progress data [1]. However, since email is not a specialized way to store, manage, and track documents, construction organizations often encounter troubles with regard to lossless storage and efficient management. In order to deal with these issues, it is necessary to introduce blockchain technology that has features in security, decentralization, and transparency.

This study is a basic study for developing a document management system with reliability based on blockchain and email. The purpose of this study is to develop a conceptual model that automatically extracts and organizes block information for blockchain based on email.

2. RESEARCH BACKGROUND

2.1 Document Management in Construction Project

Information technologies (IT) software for construction projects contains problems due to centralized governance and fragmented structure. On most projects, the IT softwares are managed exclusively by the general contractors [2]. It means the existing systems are centralized and have transparency issues and vulnerability of malicious manipulation and forgery. Poor maintenance of the centralized IT infrastructure can lead to data loss [3-6]. Lumineau et al. [7] pointed out that it is difficult to verify authenticity in a centralized system and can require a lot of time and cost. Contracts are used in the construction industry to bind parties together for the purpose of pursuing a construction project together. In today's construction megaprojects, the communication and change management are often done manually through email. This approach is vulnerable and prone to human errors [8]. This means that difficult situations may arise in the fact-checking process of potential disputed situations. Thus, it is required to improve the document management method to minimize the manual work and to store the email information automatically.

2.2 Blockchain Application

Blockchain is a type of transparent database using a peer-to-peer (P2P) network, as opposed to a centralized server [9]. In other words, blockchain technology secures the reliability of shared documents within a construction project while simultaneously managing the data shared among multiple persons effectively. The blockchain owned by each node are synchronized after block generation due to transaction occurrence and block propagation using the P2P protocol through this blockchain, the participant of the construction project can record, verify, store, access information with mutual cooperation [10]. Several studies have been conducted on the storage and application of different types of project information (e.g., project documents, site information, schedule, costs, quality, materials, BIM files, and business compliance). Blockchain is a disruptive technology, given that it can build mutual trust in cooperation and enable different organizations of a project to achieve security and consistency in their information management [11]. Previous studies shows that blockchain technology has advantages in information storing and tracking. But, research gaps exist because no research has been conducted to store and utilize email information and construction documents on the blockchain.

3. RESEARCH FRAMEWORK

3.1 Design of block data structure

The block header and body consist of three types of information, such as email event information, additional user-input information, and information referring to existing blockchain. The concept for block generation is shown in Figure 1. First, the email structure contains header and body like the block structure. Header of email includes attributes such as "from", "to", "CC", "subject", "date", etc. Body includes "contents" and "attachment". The email event information includes these attribute values of email structure. Second, user additional information is directly input by users. It includes attributes such as "Document classification code", "Related keywords",

"Transaction ID", and "Attachment version", which are available for efficient search, versioning and filtering. Third, the others (i.e., "Block number", "Previous Hash", "Previous transaction ID", "Previous attachment version") are created referring to existing blockchain. The block header and body are reorganized by combining the three sources.

The details of the components of the header and body of the block are shown in Table 1 and 2, respectively. The block header is mainly composed of information related to email event and information referring to existing blockchain. In the case of the block body, it is composed based on information of attached document and additional user-input information. Among these, "encrypted original text" is a hex string type. This is the output value after extracting the text from the attached file and putting the extracted text as the input value of the sha-256 function.

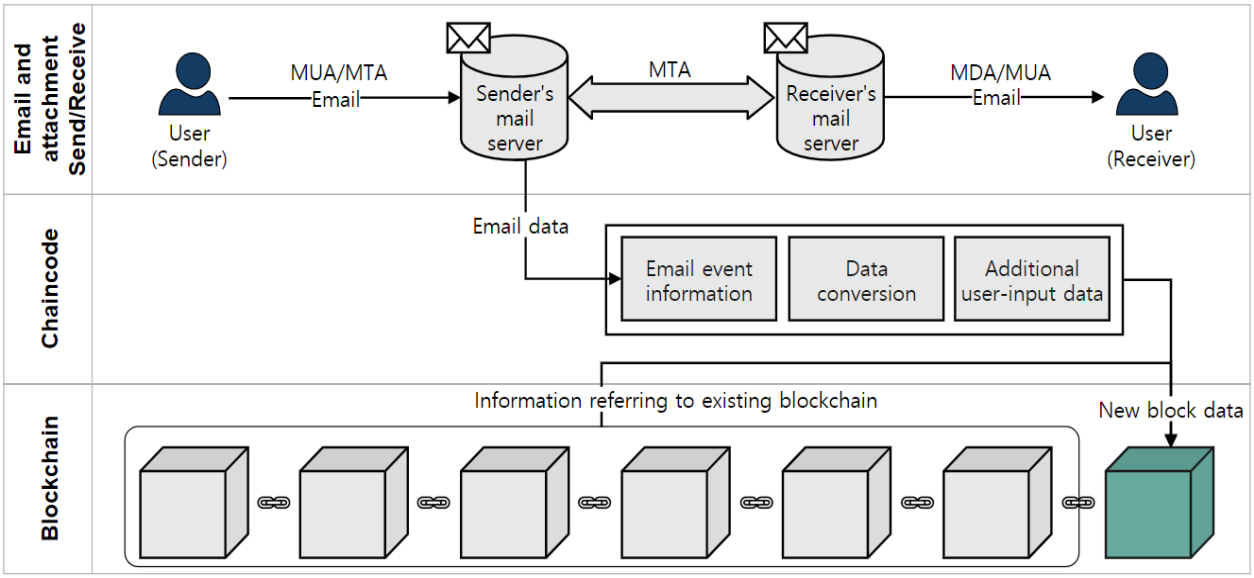

Figure 1. Concept of block generation model

Table 1. Component of the block header

Name	Type	Description
Block number	CHAR	Number assigned to each block
Block hash	CHAR	Hash value of the header data of new block
Previous Hash	CHAR	Hash value of the previous block
Merkle root	CHAR	Hash value of merkle tree
senderID	VARCHAR	Email Sender's unique ID
receiverID	VARCHAR	Email receiver's unique ID
Subject	VARCHAR	Subject of email
Content	VARCHAR	Content of the email
Timestamp	TIMESTAMP	Occurrence time of an event [Year, Month, Day, Hour, Minute, Second]

Table 2. Component of the block body

Name	Type	Description
transactionID	CHAR	Transaction (attached file) number
prevTransactionID	CHAR	Previous transaction number
attachmentVersion	SET	Version of the attachment
prevAttachmentVersion	SET	Previous version of the attachment
encryptedOriginalText	CHAR	Hash value of attachment text data (encrypted)
documentClassficationCode	SET	Document classification code
relatedKeywords	VARCHAR	Keywords associated with attachments

3.2 Block generation process

3.2.1 Email event

The block information generation process is shown in Figure 2. In the email environment, the mail user agent (MUA) is used to construct the screen view exposed to the user such as Outlook. The mail delivery agent (MDA) is used to bring the information of the email to mail server, and the mail transfer agent (MTA) is used to transmit the email between each email server .

The order in which users write and send emails on this application is as follows. The sender logs in using a application. The application serves as the role of MUA. The sender writes an email and requests to send an email. The MUA uses SMTP to send an email request to the sender's MTA. The sender-side MTA delivers mail information to the receiver-side MTA using the SMTP. The information stored in the mailbox can be accessed by the receiver through MUA. When the receiver requests confirmation, the receiver's MUA requests information from the receiver's mailbox and then delivers it to the receiver. When the receiver's MUA requests the recipient-side mailbox, POP3 or IMAP is used. If mail data is well transmitted to the receiver's MTA, it returns the reception result that the information was delivered to the sender's MTA. The sender's MTA transfers the reply message to the application. The sender's MTA transmits information about the mail sent to the sender's mailbox and the application starts the block generation process.

3.2.2 Email information extraction and conversion

The application requests and returns email data to the sender's mailbox, and then undergoes a data converting process of the email according to the block-data structure. Email information is converted to block information as follows; "from" to "senderID", "to" to "receiverID", "subject" to "Subject", "date" to "Timestamp" and attached files to "Encrypted original text". The original text of the attachment is extracted and encrypted using SHA-256 algorithm which is a one-way encryption method for business privacy. The encrypted original text is derived from each attached file, that is, transactions are generated by the number of attachments.

When the data conversion process is completed through this process, the sender can input additional information for document management such as classification, versioning, and correlation in the application. Next, when the input of additional information is terminated, a blockchain upload request is made to the application. The application then generates block information based on the result of data converting and additional input information.

The remaining information generated by referring to the existing blockchain includes "Block number", "Block Hash", "Previous Hash", "Merkle Root", and "Tx ID". "Block Number" and "Previous Hash" are formed based on the data of the last block currently stored in the blockchain.

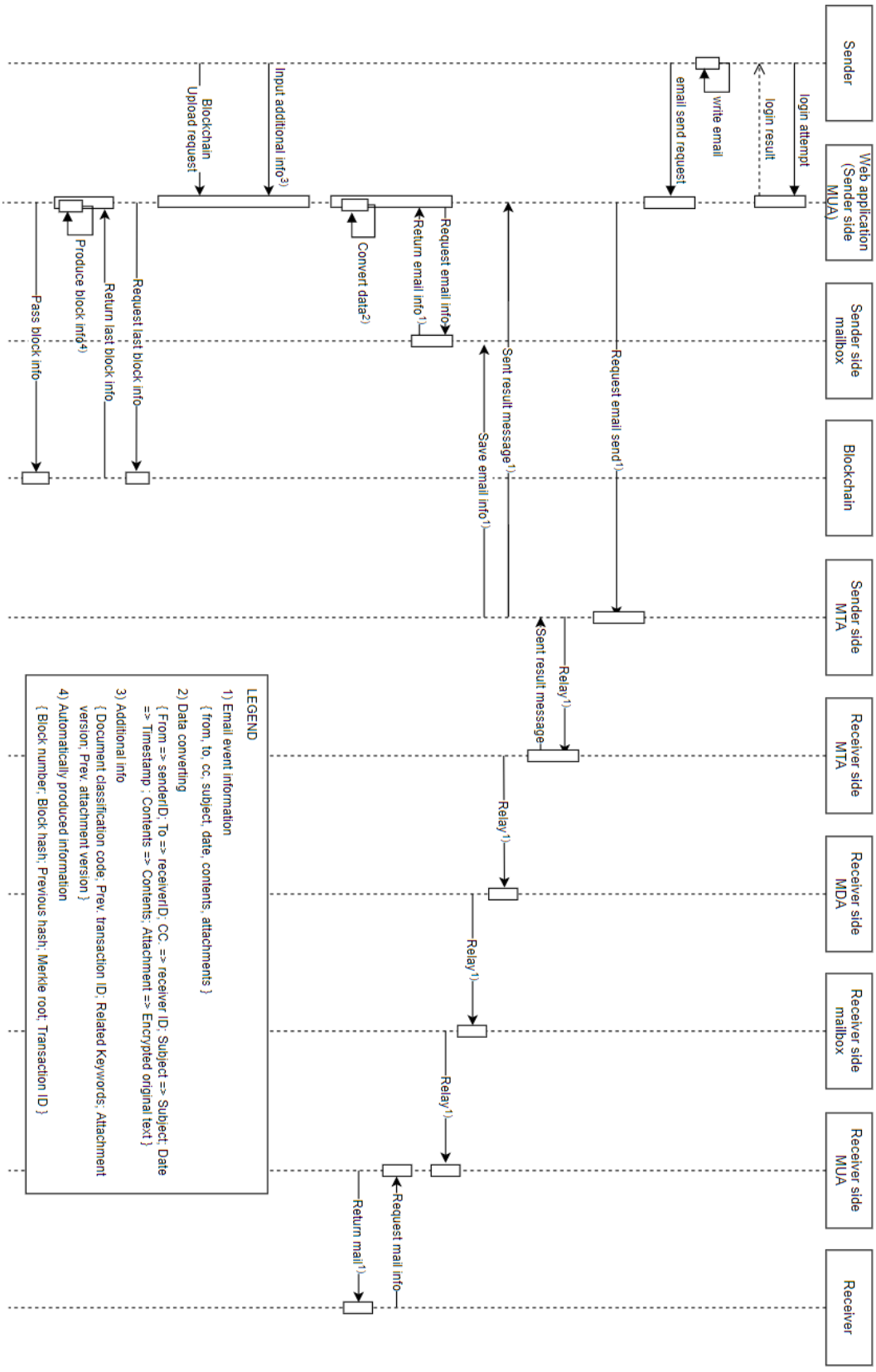

Figure 2. Flow of email information and automatic generation process of block data

"Block Number" is given by adding 1 to the the last block number (N) of the blockchain, that is N+1. “Block Hash” value of the last block is set to “Previous Hash” value of the new block. All of these are stored as the components of new block’s header. "Tx ID" refers to the order of “Encrypted original text”, which is also formed by the number of attached files. Each "Tx ID" is assigned to the leaf nodes and forms a binary tree, and the value of the top node is determined as "Merkle root". "Tx ID"s are stored in the body, and "Merkle Root" is stored in the header of the block.

3.2.3 Block generation

All values stored in the header of the block are collected by the concatenation method, and then it is encrypted by the sha-256 function. The output value is finally set to "Block hash". When the entire block information is generated through this process, it is broadcasted and connected to the existing blockchain as a new block by chain mechanism. Each block is sequentially connected to the existing blockchain whenever it is created.

4. CONCLUSION

The purpose of this study is to develop a conceptual model that automatically extracts and organizes block information about the blockchain based on email. In this study, three types of data sources were defined to link email structure and block structure; email event information, additional user-input information, and information referring to existing blockchain. Furthermore, a method of automatically generating blocks using sources was presented through information flow. Through the proposed method, it is possible for users to automatically generate blocks at the same time as sending an email.

The proposed method minimizes manual work for email based document management through automatic data linkage and prevents information loss because relevant information is immediately uploaded to the blockchain when email is distributed. In addition, since the generated blocks are sequentially connected to the blockchain, these information can be referenced when identifying the document flow and fact-checking. That is, it is possible to promote the convenience of construction project document management.

Future work will focus on the solution with extended functions by integrating the proposed conceptual model into the construction document management system.

ACKNOWLEGEMENTS

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (No. 2019R1A2C2010794) and the Chung-Ang University Graduate Research Scholarship in 2019.

REFERENCES

[1] T. Omar, M. L. Nehdi, “Data acquisition technologies for construction progress tracking”, Automation in Construction, vol. 70, pp. 143-155, 2016.

[2] P. Arnold, A. Javernick-Will, “Projectwide access: Key to effective implementation of construction project management software systems”, Journal of Construction Engineering and Management, vol. 139, no. 5, pp. 510-518, 2013.

[3] A. S. E. Pradeep, T. W. Yiu, Y. Zou, R. Amor,“Blockchain-aided information exchange records for design liability control and improved security”, Automation in Construction, vol. 126, no. 103667, pp. 510-518, 2021.

[4] M. Vacek,“Email overload: causes, consequences and the future”, International Journal of Computer Theory and Engineering, vol. 6, no. 2 pp. 170-176, 2014.

[5] R. A. Ara, K. Paardenkooper, R. van Duin,“A new blockchain system design to improve the supply chain of engineering, procurement and construction (EPC) companies–a case study in the oil and gas sector”, Journal of Engineering, Design and Technology, vol. ahead-of-print, no. ahead-of-print, 2021.

[6] A. M. Mahamadu, L. Mahdjoubi, C. A. Booth, “Determinants of Building Information Modelling (BIM) acceptance for supplier integration: A conceptual model”, In Proceedings 30th Annual ARCOM Conference, Portsmouth, UK, pp.723-732, 2014.

[7] F. Lumineau, W. Wang, O. Schilke, “Blockchain governance—A new way of organizing collaborations?.”, Organization Science, vol.32, no.2, pp.500-521, 2021.

[8] S. Shokri, K. Maloney, “Interface management integration for major capital projects”. In Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, UAE, pp. SPE-177943-MS, 2015.

[9] R. Yang, R. Wakefield, S. Lyu, S. Jayasuriya, F. Han, X. Yi, X., Yang, G. Amarasinghe, S. Chen, “Public and private blockchain in construction business process and information integration”, Automation in Construction, vol. 118, no.103276, 2020.

[10] K. Kim, G. Lee, S. Kim, “A study on the application of blockchain technology in the construction industry”, Ksce Journal of Civil Engineering, vol. 24, no. 9 pp. 2561-2571, 2020.

[11] B. Zhong, H. Wu, L. Ding, H. Luo, Y. Luo, X. Pan, “Hyperledger fabric-based consortium blockchain for construction quality information management”, Frontiers of Engineering Management, vol. 7, no. 4 pp. 512-527, 2020.