

# Blockchain in Wireless Communications and Computing: Security Threats and Applications

Guest Editors: Weizhi Meng, Wenjuan Li, and Reza Malekian





---

# **Blockchain in Wireless Communications and Computing: Security Threats and Applications**

Wireless Communications and Mobile Computing

---

**Blockchain in Wireless  
Communications and Computing:  
Security Threats and Applications**

Guest Editors: Weizhi Meng, Wenjuan Li, and Reza  
Malekian




---


Copyright © 2021 Hindawi Limited. All rights reserved.

This is a special issue published in “Wireless Communications and Mobile Computing.” All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

# Chief Editor

Zhipeng Cai , USA

## Associate Editors

Ke Guan , China  
Jaime Lloret , Spain  
Maode Ma , Singapore

## Academic Editors

Muhammad Inam Abbasi, Malaysia  
Ghufran Ahmed , Pakistan  
Hamza Mohammed Ridha Al-Khafaji ,  
Iraq  
Abdullah Alamoodi , Malaysia  
Marica Amadeo, Italy  
Sandhya Aneja, USA  
Mohd Dilshad Ansari, India  
Eva Antonino-Daviu , Spain  
Mehmet Emin Aydin, United Kingdom  
Parameshchhari B. D. , India  
Kalapraveen Bagadi , India  
Ashish Bagwari , India  
Dr. Abdul Basit , Pakistan  
Alessandro Bazzi , Italy  
Zdenek Becvar , Czech Republic  
Nabil Benamar , Morocco  
Olivier Berder, France  
Petros S. Bithas, Greece  
Dario Bruneo , Italy  
Jun Cai, Canada  
Xuesong Cai, Denmark  
Gerardo Canfora , Italy  
Rolando Carrasco, United Kingdom  
Vicente Casares-Giner , Spain  
Brijesh Chaurasia, India  
Lin Chen , France  
Xianfu Chen , Finland  
Hui Cheng , United Kingdom  
Hsin-Hung Cho, Taiwan  
Ernestina Cianca , Italy  
Marta Cimitile , Italy  
Riccardo Colella , Italy  
Mario Collotta , Italy  
Massimo Condoluci , Sweden  
Antonino Crivello , Italy  
Antonio De Domenico , France  
Floriano De Rango , Italy

Antonio De la Oliva , Spain  
Margot Deruyck, Belgium  
Liang Dong , USA  
Praveen Kumar Donta, Austria  
Zhuojun Duan, USA  
Mohammed El-Hajjar , United Kingdom  
Oscar Esparza , Spain  
Maria Fazio , Italy  
Mauro Femminella , Italy  
Manuel Fernandez-Veiga , Spain  
Gianluigi Ferrari , Italy  
Luca Foschini , Italy  
Alexandros G. Fragkiadakis , Greece  
Ivan Ganchev , Bulgaria  
Óscar García, Spain  
Manuel García Sánchez , Spain  
L. J. García Villalba , Spain  
Miguel Garcia-Pineda , Spain  
Piedad Garrido , Spain  
Michele Girolami, Italy  
Mariusz Glabowski , Poland  
Carles Gomez , Spain  
Antonio Guerrieri , Italy  
Barbara Guidi , Italy  
Rami Hamdi, Qatar  
Tao Han, USA  
Sherief Hashima , Egypt  
Mahmoud Hassaballah , Egypt  
Yejun He , China  
Yixin He, China  
Andrej Hrovat , Slovenia  
Chunqiang Hu , China  
Xuexian Hu , China  
Zhenghua Huang , China  
Xiaohong Jiang , Japan  
Vicente Julian , Spain  
Rajesh Kaluri , India  
Dimitrios Katsaros, Greece  
Muhammad Asghar Khan, Pakistan  
Rahim Khan , Pakistan  
Ahmed Khattab, Egypt  
Hasan Ali Khattak, Pakistan  
Mario Kolberg , United Kingdom  
Meet Kumari, India  
Wen-Cheng Lai , Taiwan



Jose M. Lanza-Gutierrez, Spain  
Pavlos I. Lazaridis , United Kingdom  
Kim-Hung Le , Vietnam  
Tuan Anh Le , United Kingdom  
Xianfu Lei, China  
Jianfeng Li , China  
Xiangxue Li , China  
Yaguang Lin , China  
Zhi Lin , China  
Liu Liu , China  
Mingqian Liu , China  
Zhi Liu, Japan  
Miguel López-Benítez , United Kingdom  
Chuanwen Luo , China  
Lu Lv, China  
Basem M. ElHalawany , Egypt  
Imadeldin Mahgoub , USA  
Rajesh Manoharan , India  
Davide Mattera , Italy  
Michael McGuire , Canada  
Weizhi Meng , Denmark  
Klaus Moessner , United Kingdom  
Simone Morosi , Italy  
Amrit Mukherjee, Czech Republic  
Shahid Mumtaz , Portugal  
Giovanni Nardini , Italy  
Tuan M. Nguyen , Vietnam  
Petros Nicolitidis , Greece  
Rajendran Parthiban , Malaysia  
Giovanni Pau , Italy  
Matteo Petracca , Italy  
Marco Picone , Italy  
Daniele Pinchera , Italy  
Giuseppe Piro , Italy  
Javier Prieto , Spain  
Umair Rafique, Finland  
Maheswar Rajagopal , India  
Sujan Rajbhandari , United Kingdom  
Rajib Rana, Australia  
Luca Reggiani , Italy  
Daniel G. Reina , Spain  
Bo Rong , Canada  
Mangal Sain , Republic of Korea  
Praneet Saurabh , India

Hans Schotten, Germany  
Patrick Seeling , USA  
Muhammad Shafiq , China  
Zaffar Ahmed Shaikh , Pakistan  
Vishal Sharma , United Kingdom  
Kaize Shi , Australia  
Chakchai So-In, Thailand  
Enrique Stevens-Navarro , Mexico  
Sangeetha Subbaraj , India  
Tien-Wen Sung, Taiwan  
Suhua Tang , Japan  
Pan Tang , China  
Pierre-Martin Tardif , Canada  
Sreenath Reddy Thummaluru, India  
Tran Trung Duy , Vietnam  
Fan-Hsun Tseng, Taiwan  
S Velliangiri , India  
Quoc-Tuan Vien , United Kingdom  
Enrico M. Vitucci , Italy  
Shaohua Wan , China  
Dawei Wang, China  
Huaqun Wang , China  
Pengfei Wang , China  
Dapeng Wu , China  
Huaming Wu , China  
Ding Xu , China  
YAN YAO , China  
Jie Yang, USA  
Long Yang , China  
Qiang Ye , Canada  
Changyan Yi , China  
Ya-Ju Yu , Taiwan  
Marat V. Yuldashev , Finland  
Sherali Zeadally, USA  
Hong-Hai Zhang, USA  
Jiliang Zhang, China  
Lei Zhang, Spain  
Wence Zhang , China  
Yushu Zhang, China  
Kechen Zheng, China  
Fuhui Zhou , USA  
Meiling Zhu, United Kingdom  
Zhengyu Zhu , China





# Contents

---




## **A Privacy-Preserving Blockchain Supervision Framework in the Multiparty Setting**

Baodong Wen, Yujue Wang, Yong Ding , Haibin Zheng, Hai Liang , and Huiyong Wang  
Research Article (9 pages), Article ID 5236579, Volume 2021 (2021)




## **Sharing Pandemic Vaccination Certificates through Blockchain: Case Study and Performance Evaluation**

José L. Hernández-Ramos , Georgios Karopoulos , Dimitris Geneiatakis , Tania Martin, Georgios Kambourakis , and Igor Nai Fovino  
Research Article (12 pages), Article ID 2427896, Volume 2021 (2021)

## **Proof of Engagement: A Flexible Blockchain Consensus Mechanism**

Yuntao Xu, Xingyu Yang, Jiale Zhang , Junwu Zhu , Maosheng Sun, and Bing Chen   
Research Article (10 pages), Article ID 6185910, Volume 2021 (2021)

## **A Novel User Collusion-Resistant Decentralized Multi-Authority Attribute-Based Encryption Scheme Using the Deposit on a Blockchain**

Siwan Noh , Donghyun Kim, Zhipeng Cai , and Kyung-Hyune Rhee   
Research Article (15 pages), Article ID 9506796, Volume 2021 (2021)

## **Trustworthy Image Fusion with Deep Learning for Wireless Applications**

Chao Zhang, Haojin Hu, Yonghang Tai , Lijun Yun , and Jun Zhang   
Research Article (9 pages), Article ID 6220166, Volume 2021 (2021)

## Research Article

# A Privacy-Preserving Blockchain Supervision Framework in the Multiparty Setting

Baodong Wen,<sup>1</sup> Yujue Wang,<sup>2</sup> Yong Ding ,<sup>1,3</sup> Haibin Zheng,<sup>2</sup> Hai Liang ,<sup>1</sup> and Huiyong Wang<sup>4</sup>

<sup>1</sup>Guangxi Key Laboratory of Cryptography and Information Security, School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin, China

<sup>2</sup>Hangzhou Innovation Institute, Beihang University, Hangzhou, China

<sup>3</sup>Cyberspace Security Research Center, Pengcheng Laboratory, Shenzhen, China

<sup>4</sup>School of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin, China

Correspondence should be addressed to Yong Ding; [stone\\_dingy@126.com](mailto:stone_dingy@126.com)

Received 2 July 2021; Revised 8 August 2021; Accepted 20 August 2021; Published 24 September 2021

Academic Editor: Weizhi Meng

Copyright © 2021 Baodong Wen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Data supervision is an effective method to ensure the legality of user data on blockchain. However, the massive growth of data makes it difficult to achieve data supervision in existing blockchain applications. Also, data supervision often leads to problems such as disclosure of transaction data and user privacy information. To address these issues, this paper proposes a privacy-preserving blockchain supervision system (BSS) in the multiparty setting, where a supervision chain is introduced to realize data supervision on blockchain. All sensitive information such as user information in the supervising data is encrypted by the attribute-based encryption (ABE) technology, so that both privacy protection and access control on user data can be achieved. Theoretical analysis and comparison show that the proposed BSS scheme is efficient, and experimental analysis indicates the practicality of our BSS scheme.

## 1. Introduction

Blockchain is featured with the characteristics of decentralization, autonomy, and immutability [1]. As the key technology in the construction of trust systems, it is envisioned as an effective technology to address security issues faced in finance, property rights, smart cities, government affairs, supply chain, and other fields [2]. With the rapid development and wide application of blockchain, due to its open and transparent characteristics, more and more transaction data, user information, network node address, and other information face the risk of privacy leakage [3]. Unlike the traditional centralized architecture, blockchain does not rely on a central node; thus, it can effectively avoid the single point of failure. However, in order to reach consensus by all blockchain nodes, the data has to be disclosed to all of them, which also brings the risk of privacy leakage.

The lack of centralized entities makes it difficult for relevant government regulators to supervise the blockchain. Lack of regulation will seriously restrict the healthy and sustainable development of the entire blockchain industry. However, there remain some problems in the existing supervision methods on blockchain [4]. Due to the autonomous and decentralized features of blockchain, it is difficult to guarantee the legality of data on blockchain; that is, the data on blockchain cannot be well supervised. Moreover, supervision may bring the issue of privacy protection [5, 6]. Due to the significant difference between the blockchain technology and traditional system architecture, many traditional privacy protection methods are not applicable to blockchain. Therefore, it is necessary to design an effective supervision mechanism with privacy protection on user data in blockchain.

The multilayer structure has been used to achieve supervision on blockchain. Yang et al. [7] realized the monitoring



of user behavior and the verification of blocks by employing the multilayer structure. The block verification is executed by the system supervisor, which improves the performance and security of the system. Li et al. [8] designed a two-layer adaptive blockchain-based supervision framework (TABS) to address the supervision issues in off-site modular housing production, where the first layer contains the adaptive private sidechains of participants, and the second layer is the main blockchain for communication and transactions. TABS can effectively prevent the main blockchain from tampering with records, and it can also prompt users to quickly publish their transaction records without revealing their privacy. Note that a single supervisor is easily corrupted by an adversary, which can cause irreparable losses. The problem of single point of failure can be effectively avoided by setting multisupervision in the supervision process. Yang et al. [7] and Li et al. [8] put forward their schemes in the agricultural machinery scheduling and off-site modular housing production scenarios, respectively. However, there was no universal blockchain supervision framework for the application scenarios that need to be regulated. In addition, there is no blockchain supervision framework that allows multiple parties to participate in a multilayer blockchain structure.

*1.1. Our Contributions.* To address the above mentioned issues, this paper proposes a multiparty blockchain supervision system (BSS) framework. In BSS, a dual-chain architecture is introduced, which contains two types of blockchains, namely, business chain (BC) and supervision chain (SC). SC consists of the regulatory authorities and supervisors, which provides the supervision service for the data on BC. By deploying transaction information and supervision information separately, the scalability of the BSS can be improved.

ABE is employed to realize flexible access control on data; that is, the regulatory authority can set access control policies, so that different supervisors have different permissions. Regulatory authority can encrypt data for multiple supervisors at the same time and build communication channels without obtaining each supervisor's public key in advance. This process can reduce the computing overhead caused by encrypting data for each recipient. Data information will go through two rounds of supervision by regulatory authorities and supervisors. Smart contract in BSS can realize verification and upload supervision information to the blockchain. Thus, BSS supports the management and control on data in BC and also protects the data privacy, which offers trade-off between supervision and privacy protection. Through security and theoretical analysis, it is shown that the proposed BSS framework is suitable for different ABE and application scenarios.

*1.2. Related Works.* Yong et al. [9] designed a blockchain supervision system to supervise the supply of vaccines through smart contracts and machine learning. Their scheme not only supports the query on individual vaccination records and tracking the vaccine operation records through the smart contract but also allows the regulatory agencies to manage expired vaccines. Meng et al. [10] proposed a security mechanism to

build trust-based filtering. This mechanism processes and reduces malicious traffic by using traffic fusion and aggregation. Yin et al. [11] provided an approach using supervised machine learning to implement system supervision, where the gradient enhancement algorithm was used to predict the type of entity. A classifier was established to distinguish 12 categories by using 957 entities as sample data for authentication. The gradient boosting algorithm with default parameters was used to improve the accuracy of average cross validation. Ma et al. [12] proposed a traceable blockchain scheme, SkyEye, which enables the regulatory authority to track the identity of users. In [13], Ma et al. designed a blockchain traceable scheme with oversight function based on SkyEye, from a distributed multikey generation protocol and some other cryptographic primitives. Note that the supervisor must obtain the consent of committees when tracing some users. Meng et al. [14] proposed a blockchain-enabled single character frequency-based exclusive signature matching scheme to secure the security of smart IoT environment.

In terms of privacy protection, many information hiding mechanisms have been proposed for transaction contents, including Monero and Zcash. Monero is mainly based on the Cryptonote protocol, which uses one-time random address and ring signature to randomize the sender and other node information so as to realize the sender's anonymity. Encryption is used to realize the anonymity of the receiver; that is, only the receiver has the private key of the ciphertext. In Monero coin, the anonymity of the sender is determined by the size of the anonymous set. The stronger the anonymity is, the larger the anonymous set is, but the time complexity of encryption and decryption will also increase. Zcash [15] is a cryptocurrency embedded with noninteractive zero-knowledge proof, which divides the address into transparent address and hidden address. The hidden address is used to realize anonymity for users. Unlike Monero, Zcash can authenticate transactions without disclosing transaction data. However, users may not make transparent transactions due to the computational cost of zero-knowledge proof. Blockchain platforms such as Monax and Multichain [16] provide multichain solutions that enable privacy protection of transaction data through inter-chain isolation.

As a kind of computer protocol, smart contract [17] can realize automatic verification, programmable execution, irreversible, and other functions. The security and privacy of smart contract can be guaranteed by formal verification [18], decompilation [19], etc. Cheng et al. [20] introduced Eki-den, which combines blockchain with trusted hardware. The Eki-den system separates consensus and execution, which offers the high system performance and scalability. In the initialization phase, the smart contract is encrypted and stored on the blockchain after verification. The corresponding public key and private key should be provided when the smart contract is called and acquired. The privacy protection for smart contract is realized by storing encrypted contract.

ABE is developed on the basis of identity-based encryption (IBE) proposed by Boldyreva et al. [21]. Compared with the previous encryption methods, ABE realizes a one-to-many encryption mode, provides fine-grained access control

on data, and also supports certain fault tolerance [22]. Goyal et al. [23] designed a key policy attribute-based encryption scheme, where the access policy and attribute set are embedded in the key and ciphertext, respectively. Bethencourt et al. [24] proposed a ciphertext policy attribute-based encryption scheme, where the access policy and user attributes are, respectively, embedded in ciphertext and key, which can be used in access control applications such as private data sharing [25].

*1.3. Organization.* The remainder of this paper is organized as follows. Section 2 describes the system model, dual-chain architecture, and security requirements. Section 3 introduces the preliminaries for the proposed BSS scheme. A description of our BSS scheme is presented in Section 4. In Section 5, the security and performance of our BSS scheme are evaluated and compared. Section 6 concludes the paper.

## 2. System Model and Requirements

*2.1. System Model.* As shown in Figure 1, a BSS system consists of five types of entities, namely, regulatory authority, supervisor, key generation center (KGC), business chain (BC), and supervision chain (SC).

- (i) *Regulatory Authority.* For the data to be supervised, they can be judged by the regulatory authority according to some rules. The supervision results are encrypted by the regulatory authority and broadcasted to supervisors.
- (ii) *Supervisor.* The supervisors with decryption permission are able to decrypt the information sent by the regulatory authority and supervise the related data.
- (iii) *KGC.* The KGC is responsible for generating system public parameters and registering the attributes of supervisor.
- (iv) *BC.* BC is mainly used to maintain the business data information.
- (v) *SC.* SC mainly manages the supervision information. The supervision information processed by the regulatory authority and the supervisor is transmitted to SC.

The data to be supervised in the BSS system is first delivered to the regulatory authority. They are supervised according to the supervision rules by the regulatory authority and then encrypted by employing the hybrid encryption technology, where the access policy can be set during hybrid encryption. Only the supervisor satisfying the access policy has relevant authority to conduct supervision. The supervisors need to perform the second round of supervision on data. If the supervision results of two rounds are consistent, then the smart contract uploads the supervision information to SC.

*2.2. Dual-Chain Architecture.* In order to realize data supervision, this paper introduces a dual-chain architecture composed of BC and SC and uses the ABE method to protect the privacy of user data. The dual-chain architecture is shown in Figure 2.

- (i) *SC.* SC consists of the regulatory authority nodes (RAN) and the supervisor nodes (SUN). In real world applications, RAN can be the regulatory authority, while SUN may comprise legal departments. RAN is able to perform supervision and encryption on data. Data submitted to the RAN for supervision is reviewed in the first round to detect illegal information. SUN can complete the decryption and supervision on data, so that the supervisor can further supervise the relevant data after decryption. The supervision results in the two rounds of supervision can be confirmed and uploaded to SC through smart contract.
- (ii) *BC.* Different types of data are stored on BC, which should have been delivered to SC for supervision before being written to BC. Also, when the data is retrieved from BC, it should be first delivered to SC for supervision. In addition, the data in BC may also be taken out for supervising whenever necessary.

*2.3. Security Requirements.* A secure BSS system in the multiparty setting has to satisfy the following requirements.

- (i) *Anticollusion Attack.* Even when the keys of multiple supervisors are combined, these supervisors cannot obtain the valid ciphertext. Supervisors cannot obtain plaintext data that exceeds their regulatory capability.
- (ii) *Multiparty Supervision.* Data on the BC can be supervised by multiple parties. It avoids the problem of excessive concentration of power under the supervision of an individual or separate agency and reduces the security risk caused by the breach of one party.
- (iii) *Privacy Protection.* Sensitive data submitted to the SC should be encrypted to guarantee their privacy.
- (iv) *Access Control.* The supervisors are not allowed to access data that is not authorized. Different supervisors have their own permission and have different decryption capabilities for the data sent from the regulatory authority.

## 3. Preliminaries

*3.1. Attribute-Based Encryption.* A ciphertext-policy attribute-based encryption scheme  $A$  consists of four algorithms

$$A = (\text{Setup}, \text{KeyGen}, \text{Enc}, \text{Dec}). \quad (1)$$

- (i) *Setup( $d$ )*  $\rightarrow$  ( $PK, MK$ ). With input the security parameter  $d$ , the system setup algorithm outputs public parameter  $PK$  and master key  $MK$
- (ii) *KeyGen( $MK, U$ )*  $\rightarrow$   $SK$ . With input the master key  $MK$  and the attribute set  $U$ , the key generation

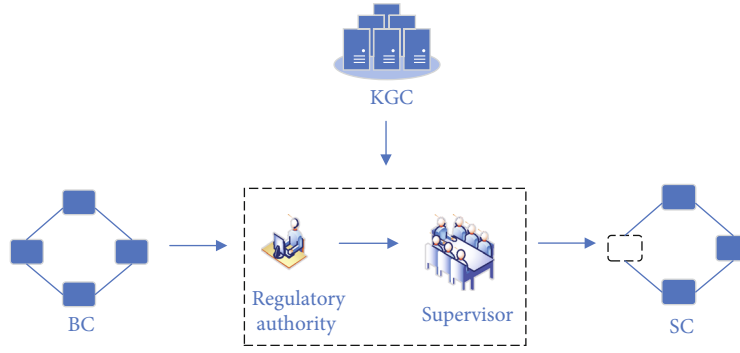


FIGURE 1: System model of BSS.

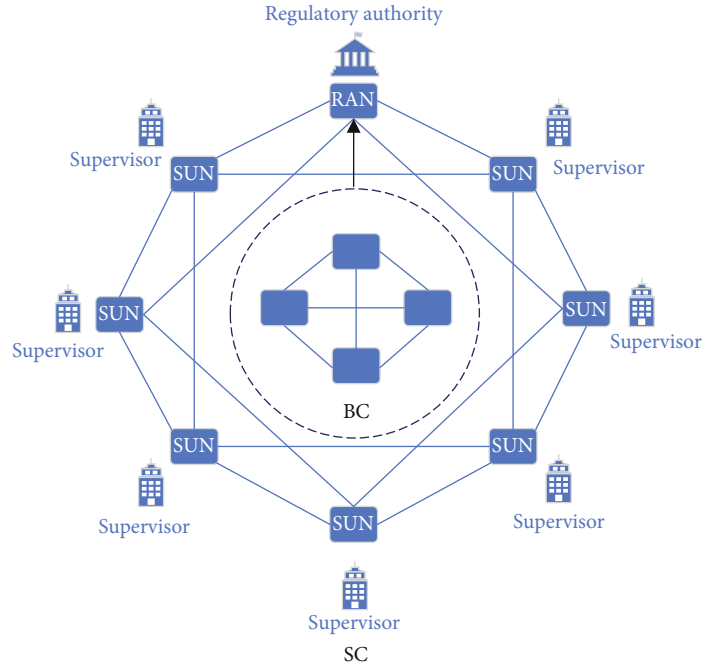


FIGURE 2: Dual-chain architecture.

algorithm outputs the private key  $SK$  associated with the user attribute set  $U$

(iii)  $Enc(PK, M, T) \rightarrow CT$ . With input the public parameter  $PK$ , a plaintext message  $M$ , and an access structure  $T$ , the encryption algorithm outputs a ciphertext  $CT$

(iv)  $Dec(CT, SK, PK) \rightarrow M$ . With input the ciphertext  $CT$ , private key  $SK$ , and public parameter  $PK$ , if the attributes in the user key match the access policy required by the ciphertext, then the decryption algorithm outputs the corresponding plaintext  $M$

**3.2. Bilinear Groups.** Let  $G_1$  and  $G_T$  be two cyclic groups of prime order  $p$  and  $g$  be a generator of  $G_1$ . A bilinear map  $e : G_1 \times G_1 \rightarrow G_T$  satisfies the following conditions:

(i) *Bilinearity.* For  $a, b \in_{\mathbb{R}} \mathbb{Z}_p$ , we have

$$e(g^a, g^b) = e(g, g)^{ab}. \quad (2)$$

(ii) *Nondegeneracy.* There exists  $r, s \in G_1$  such that

$$e(r, s) \neq 1_{G_T}, \quad (3)$$

where  $1_{G_T}$  is the identity of  $G_T$ .

(iii) *Computability.* For  $r, s \in_{\mathbb{R}} G_1$ , there is an efficient algorithm to compute  $e(r, s)$ .

## 4. BSS Construction

Our BSS framework consists of five procedures, namely, system setup, registration, regulatory authority supervision, supervisor

second-round supervision, and data processing. The frequently used notations are summarized in Table 1, and the process of supervision is shown in Figure 3.

**4.1. System Setup.** KGC selects a secure symmetric encryption scheme  $F = (\text{KeyGen}, \text{Enc}, \text{Dec})$  and a ciphertext-policy attribute-based encryption scheme  $A = (\text{Setup}, \text{KeyGen}, \text{Enc}, \text{Dec})$ . With input the security parameter  $d$ , KGC runs the algorithm  $A.\text{Setup}(d)$  to generate the public parameter PK and the master key MK. Then, KGC uploads the encryption scheme  $F$  and the public parameter PK to the blockchain, while MK is kept secret and not allowed to be accessed by other users.

**4.2. Registration.** The supervisor submits its own attribute set  $U$  to KGC for registration. KGC first searches the database; if the supervisor has already registered, the registration request is rejected. Otherwise, the attribute set  $U$  is added to the local database of KGC. Then, KGC generates a registration record as follows:

$$R \leftarrow (U \parallel \text{id}_S \parallel N \parallel t_{\text{Reg}}), \quad (4)$$

which contains the user's attribute set  $U$ , identity document of supervisor  $\text{id}_S$ , KGC's signature  $N$ , and registration time  $t_{\text{Reg}}$ . The registration record  $R$  is written to the blockchain. Then, KGC runs the  $A.\text{KeyGen}$  algorithm with the master key MK and user attribute  $U$  to generate the private key SK, which is sent to the corresponding supervisor to complete the registration process. The registration process is shown in Algorithm 1.

**4.3. Regulatory Authority Supervision.** The data  $m$  to be supervised is first uploaded to RAN, so that the regulatory authority can perform supervision based on its own rules. The regulatory authority generates supervision record

$$\beta_{\text{RA}} \leftarrow (m \parallel I \parallel J_1 \parallel t_{\text{RA}} \parallel \text{id}_{\text{RA}}), \quad (5)$$

where  $m$  is the supervised data,  $I$  is the user information,  $t_{\text{RA}}$  is the regulatory authority supervision time, and  $\text{id}_{\text{RA}}$  is the identity document of regulatory authority. Also,  $J_1 \leftarrow (\text{id}_m \parallel \text{id}_{\text{BC}} \parallel V \parallel \lambda_{\text{RA}})$  denotes the supervision result of the regulatory authority, where  $\text{id}_m$  is the identity document of  $m$ ,  $\text{id}_{\text{BC}}$  is the identity document of BC,  $V$  is the rule that  $m$  violates, and  $\lambda_{\text{RA}}$  is the judgment of regulatory authority.

The regulatory authority generates a symmetric key  $k \leftarrow F.\text{KeyGen}(d)$  and calculates

$$\begin{aligned} C_1 &= A.\text{Enc}(\text{PK}, k, T), \\ C_2 &= F.\text{Enc}(k, \beta_{\text{RA}}), \end{aligned} \quad (6)$$

where PK and  $T$  are the public parameters and access structure in ABE, respectively. The RAN outputs the corresponding ciphertext  $\langle C_1, C_2 \rangle$  and broadcasts it to supervisors.

**4.4. Supervisor Second-Round Supervision.** For the received ciphertext tuple  $\langle C_1, C_2 \rangle$ , the supervisor executes  $A.\text{Dec}(C_1, \text{SK}, \text{PK})$  with its private key SK. If SK satisfies the

TABLE 1: Notations.

Notations	Descriptions
$F$	Secure symmetric encryption algorithm
$A$	Ciphertext-policy attribute-based encryption scheme
$R$	Registration record
$U$	Attribute set
$m$	The data to be supervised
$I$	User information
$J_1$	The supervision result of the regulatory authority
$J_2$	The supervision result of the supervisor
$N$	The signature of KGC
$t$	Timestamp
$\beta$	Supervision record
$L$	Supervision information

access policy  $T$  in  $C_1$ , then the supervisor is allowed to get the symmetric key  $k$  through decryption. Moreover, the supervisor runs the algorithm  $F.\text{Dec}(C_2, k)$  to get the corresponding plaintext tuple  $\beta_{\text{RA}}$ , which contains data  $m$  on BC. The supervisor is then able to run a second round of supervision on data  $m$  and outputs the corresponding supervision record

$$\beta_S \leftarrow (m \parallel I \parallel J_2 \parallel t_S \parallel \text{id}_S), \quad (7)$$

where  $t_S$  is the supervisor supervision time. Here,  $J_2 \leftarrow (\text{id}_m \parallel \text{id}_{\text{BC}} \parallel V \parallel \lambda_S)$  is the supervision result of the supervisor, where  $\lambda_S$  is the judgment of supervisor.

**4.5. Data Processing.** The smart contract will compare the supervision results generated in two rounds of supervision by the regulatory authority and the supervisor, respectively. If they are consistent, then the smart contract generates the following supervision information

$$L \leftarrow (\beta \parallel t_{\text{Pro}}), \quad (8)$$

and adds it to SC, which consists of data information  $\beta \leftarrow (\beta_{\text{RA}} \parallel \beta_S)$  and data processing time  $t_{\text{Pro}}$ . If the two supervision results are inconsistent, a new round of supervision should be performed by RAN. The procedure of data processing is shown in Algorithm 2.

## 5. Analysis and Comparison

### 5.1. Security Analysis

**Theorem 1.** *If the symmetric encryption scheme  $F$  and ABE scheme  $A$  are secure, then the proposed BSS framework can resist collusion attacks.*

*Proof.* In ABE schemes, SK is associated with a random polynomial  $q(x)$  or a random number  $r$ . Different random polynomial  $q(x)$  or random number  $r$  will be selected when generating private key SK for different users. The Lagrange

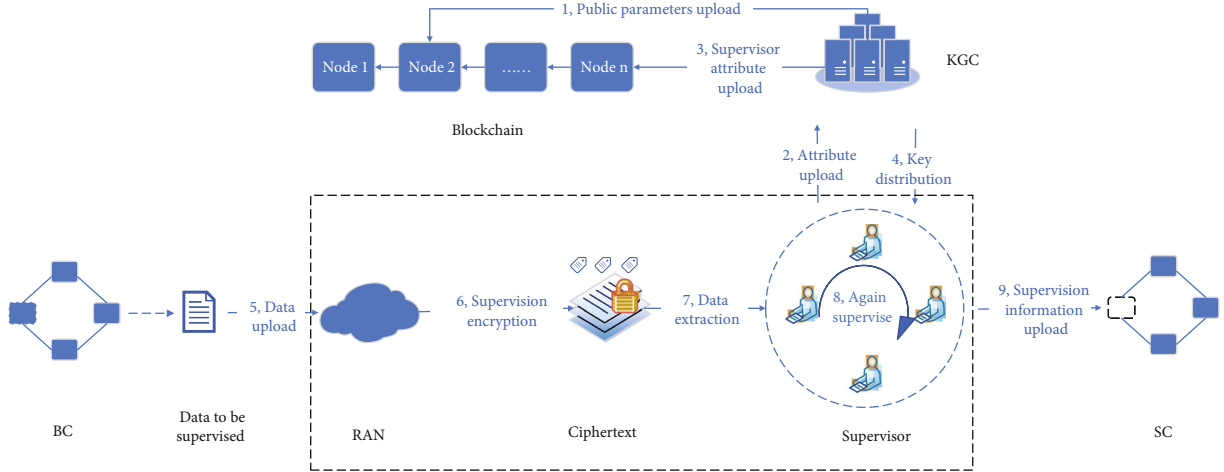


FIGURE 3: A procedure of BSS.

**Require:**  $U, id_S, MK$   
**Ensure:** *success/failure*  
 $KGC \leftarrow U$   
**if**  $id_S$  already registered **then**  
  **return failure**  
**else**  
   $KGC \leftarrow id_S, U$   
  Register  $id_S, U$  to local database  
  Generate  $N, t_{Reg}$   
   $R \leftarrow (U || id_S || N || t_{Reg})$   
  Send  $R$  to blockchain  
   $SK \leftarrow A.KeyGen(MK, U)$   
  Send  $SK$  to  $id_S$ .  
**else if**

ALGORITHM 1: Registration.

**Require:**  $J_1, J_2$   
**Ensure:** *success/failure*  
  Smart contract  $\leftarrow J_1, J_2$   
  **if**  $J_1 = J_2$  **then**  
   Generate  $t_{Pro}$   
    $\beta \leftarrow (\beta_{RA} || \beta_S)$   
   Smart contract  $\leftarrow \beta, t_{Pro}$   
    $L \leftarrow (\beta || t_{Pro})$   
    $SC \leftarrow L$   
   **return success**  
  **else**  
   **return failure**  
  **end if**

ALGORITHM 2: Supervision data processing.

polynomials cannot be combined to obtain the corresponding plaintext information. Thus, the private keys of different users cannot be combined, which means the proposed BSS framework can resist collusion attacks of multiple users.  $\square$

**Theorem 2.** *The proposed BSS framework can support the multiparty supervision of data on BC.*

*Proof.* In the proposed BSS framework, SC is used to implement review and supervision on data  $m$ , including the data uploaded to the BC, retrieved from the BC, and remain existed in the BC. Users are only allowed to upload and retrieve data that meets certain rules and conditions. Multiple parties are allowed to participate in the supervision process, where only the authorized supervisors are able to jointly supervise certain data. Only when the supervision results  $J_1$  and  $J_2$  are consistent, the supervision information  $L$  will be uploaded to SC, which can reduce the risk of privacy leakage caused by the concentration of power in the single supervision authority setting.  $\square$

**Theorem 3.** *If the symmetric encryption scheme  $F$  and ABE scheme  $A$  are secure, then the proposed BSS framework can provide privacy protection of data on BC.*

*Proof.* In the proposed BSS framework, the user information  $I$ , timestamp  $t$ , and data on BC  $m$  are encrypted by the hybrid encryption technology. Data information can only be decrypted and viewed by the user who has the corresponding private key, which can reduce the risk of privacy leakage caused by supervision. In addition, SC in the proposed scheme is realized by the consortium chain, so that only the licensed users can join SC. Compared to the public chain, the management of the consortium chain can provide better protection for data privacy and the accountability after privacy leakage.  $\square$

interpolation method requires that only when all values come from the same polynomial, the value of the target point can be solved. Therefore, when the keys of multiple users are combined, different random numbers and random

TABLE 2: Theoretical comparison.

Scheme	Supervision method	Access control	Application scenarios
Yong et al.'s scheme [9]	Machine learning	—	Vaccine supply
Yin et al.'s scheme [11]	Machine learning	—	—
Sun et al.'s scheme [26]	Multiblockchain model	—	Central bank digital currency
Peng et al.'s scheme [27]	Double-layer blockchain	—	Vaccine production
Our BBS framework	Dual-chain architecture	√	Financial trade, etc.

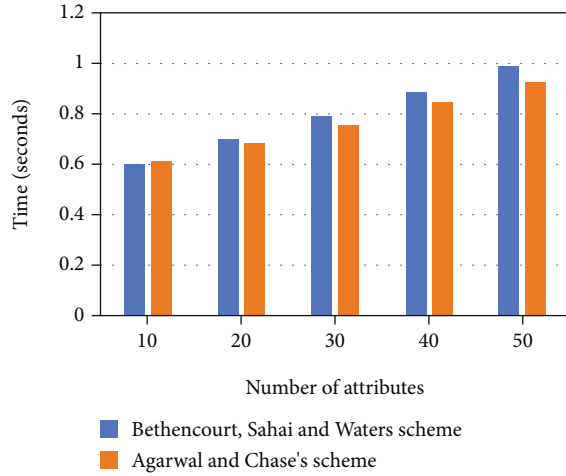


FIGURE 4: Time cost in the registration phase.

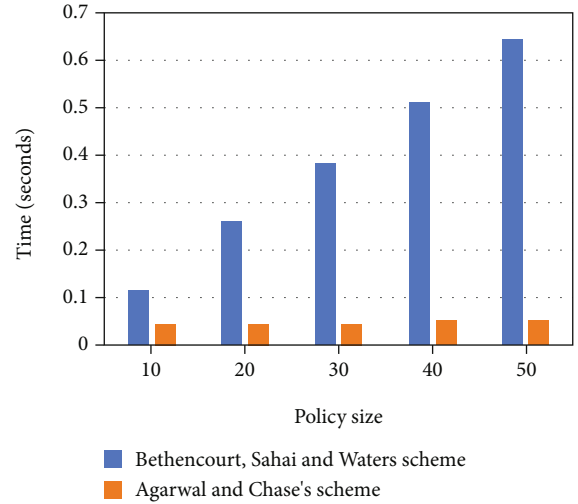


FIGURE 6: Time cost of decryption.

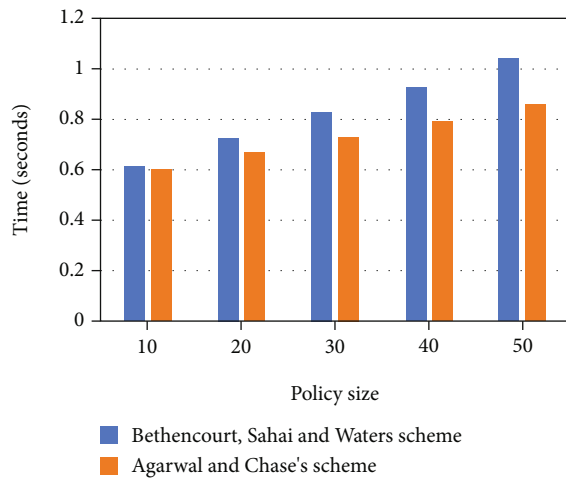


FIGURE 5: Time cost of encryption.

**Theorem 4.** *If the chosen ABE scheme  $A$  is secure, then the proposed BSS framework can support access control on data.*

*Proof.* In the proposed BSS framework, the ABE scheme  $A$  is used to control the permissions. By embedding access policy  $T$  in the encryption process, the specific supervisor is assigned to decrypt and access the data. That is, when the attribute set  $U$  of the supervisor satisfies the access policy  $T$ , the supervisor has the permission to supervise data  $m$ ;

otherwise, it does not have the permission to decrypt the data. Thus, different supervisors have different supervision permissions for different data.  $\square$   $\square$   $\square$

**5.2. Theoretical Analysis.** As shown in Table 2, the performance of our BSS framework is theoretical compared with existing supervision schemes. Yong et al.'s scheme [9] and Yin et al.'s scheme [11] mainly use the machine learning method to achieve supervision. Sun et al.'s scheme [26] introduces a multichain structure to complete supervision. Peng et al.'s scheme [27] achieves the supervising through a double-layer blockchain. In our BSS framework, the supervision of data in the blockchain is realized by designing a dual-chain architecture.

In addition, the schemes [9, 11, 26, 27] cannot control the permission of supervisors during the supervision process, whereas our BSS realizes the control on the supervisor's permission through the ABE technology. In terms of application scenarios, Yong et al.'s scheme [9] and Peng et al.'s scheme [27] are suitable to the supervision in the supply and production of vaccine, respectively, while Sun et al.'s supervision scheme [26] can be applied to the central bank digital currency. Our BSS framework is suitable for financial trade information supervision, industrial equipment maintenance information supervision, etc. In different application scenarios, regulatory authorities and supervisors use different rules to supervise different data. Our BSS framework is also suitable to other application scenarios that require multiparty supervision.

Transaction Info	Transaction Receipt
Block Hash:	0x467f6ebb24448a5f177484c14a5f7f054655e6f97da7dbde11f47251e32c114e
Block Height:	614
Gas:	30000000
From:	0x95198b93705e394a916579e048c8a32ddf900f7
To:	0x00
nonceRaw:	
Hash:	0x18ea723abd12a6207e559875693ebe8ada2d3dcfa7412716eddeb9f5b2502965
Timestamp:	2021-05-21 5:20:20

FIGURE 7: A part information of transaction.

**5.3. Experimental Analysis.** We conducted the experiments using Python and Solidity programming languages, on a platform with Ubuntu 16.04 operating system and 4 GB memory. The machine is with an AMD Ryzen 5 4600H at 3.00 GHz and 16 GB in memory. FISCO BCOS 2.0 was adopted as the underlying framework of consortium blockchain. In the Setup phase, 256-bit AES-CBC was chosen as the symmetric encryption algorithm  $F$ , which is implemented by the Crypto library. For ABE scheme, both Bethencourt, Sahai and Waters scheme [24] and Agrawal and Chase's scheme [28] were employed to process the data on the same chain. A 224-bit asymmetric elliptic curve MNT224 was chosen to realize bilinear mapping.

The performance of our BSS framework is compared by two instantiations from two ABE schemes [24] and [28], respectively. The number of supervisor attributes is a key factor for the timing of registration phase. Figure 4 shows the effect of the registration time when the number of supervisor attributes changes from 10 to 50. It can be seen from Figure 4 that the time in the registration phase enjoys a linear relationship with the number of attributes of the supervisor. When the number of attributes is 10, both instantiations take roughly the same registration time. Although the registration time grows as the number of attributes increases, the overall time increase of Agrawal and Chase's scheme [28] is lower than that of Bethencourt, Sahai and Waters scheme [24].

In the phases of regulatory authority supervision and supervisor second-round supervision, different access policies would affect the efficiency of encryption and decryption of data by regulatory authority and supervisors. Figures 5 and 6, respectively, show the encryption and decryption time of schemes [24] and [28] under different policy sizes. It can be seen from Figure 5 that the increase of the strategy will lead to the increase of encryption time of the system. While the encryption time of Agrawal and Chase's scheme [28] is lower than that of Bethencourt, Sahai and Waters scheme [24]. As shown in Figure 6, in the decryption phase, with the increase of policy size, the decryption time of Agrawal and Chase's scheme [28] does not have significant changes, while the scheme of Bethencourt, Sahai and Waters [24] changes greatly.

In the data processing phase, the smart contract will compare the supervision results  $J_1$  and  $J_2$  generated by the regulatory authority and supervisor, respectively. Then, the smart contract uploads the supervision information  $L$  to the SC. The system will output the transaction information when the supervision information  $L$  is uploaded successfully. Part of the transaction information in the data processing

phase is shown in Figure 7, which includes block hash, transaction hash, contract address, and other information. Here, the block hash is the hash value with regard to the current block, the transaction hash is the hash value generated at the end of supervision, and the contract address shows the address of the invoked contract.

## 6. Conclusion

This paper studied the problems of difficult supervision in BC, privacy leakage during supervision, and overconcentration of rights. To address these issues, a supervision system architecture BSS for data in BC is proposed. Through SC and the ABE technology, both data supervision and privacy protection can be realized. The supervisor is granted certain permission, and only the supervisor satisfying the relevant authority permission can supervise the data on BC. The proposed BSS framework also supports access control on supervisors and allows multiple supervisors to participate in supervision at the same time. The designed dual-chain architecture can effectively improve the scalability of the BSS system. Theoretical and experimental analysis shows that the BSS instantiations with different ABE schemes are suitable for real world applications.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This article is supported in part by the National Key R&D Program of China under project 2020YFB1006004; the National Natural Science Foundation of China under projects 61772150, 61862012, and 61962012; the Guangdong Key R&D Program under project 2020B0101090002; the Guangxi Natural Science Foundation under grants 2018GXNSFDA281054, 2019GXNSFFA245015, and 2019GXNSFGA245004; the Zhejiang Soft Science Research Program 2021C35044; the Guangxi Young Teachers' Basic Ability Improvement Program under Grant 2021KY0214; the Peng Cheng Laboratory Project of Guangdong Province PCL2018KP004; and the Open Program of Guangxi Key Laboratory of Cryptography and Information Security under Grant GCIS201930.

## References

- [1] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, "An overview of blockchain technology: architecture, consensus, and future trends," in *2017 IEEE International Congress on Big Data (Big-Data Congress)*, pp. 557–564, Honolulu, HI, USA, June 2017.
- [2] W. Yang, S. Garg, A. Raza, D. Herbert, and B. Kang, "Blockchain: Trends and Future," *Knowledge Management and Acquisition for Intelligent Systems*, K. Yoshida and M. Lee, Eds., pp. 201–210, Springer, Nanjing, China, 2018.

- [3] Q. Feng, D. He, S. Zeadally, M. K. Khan, and N. Kumar, "A survey on privacy protection in blockchain system," *Journal of Network and Computer Applications*, vol. 126, pp. 45–58, 2019.
- [4] G. Peters, E. Panayi, and A. Chapelle, "Trends in cryptocurrencies and blockchain technologies: a monetary theory and regulation perspective," *Journal of Financial Perspectives*, vol. 3, no. 3, pp. 92–113, 2015.
- [5] P. Pandey and R. Litoriya, "Implementing healthcare services on a large scale: challenges and remedies based on blockchain technology," *Health Policy and Technology*, vol. 9, no. 1, pp. 69–78, 2020.
- [6] J. X. Jiang and G. Bai, "Evaluation of causes of protected health information breaches," *JAMA Internal Medicine*, vol. 179, no. 2, pp. 265–267, 2019.
- [7] H. Yang, S. Xiong, S. A. Frimpong, and M. Zhang, "A consortium blockchain-based agricultural machinery scheduling system," *Sensors*, vol. 20, no. 9, p. 2643, 2020.
- [8] X. Li, L. Wu, R. Zhao, W. Lu, and F. Xue, "Two-layer adaptive blockchain-based supervision model for off-site modular housing production," *Computers in Industry*, vol. 128, p. 103437, 2021.
- [9] B. Yong, J. Shen, X. Liu, F. Li, H. Chen, and Q. Zhou, "An intelligent blockchain-based system for safe vaccine supply and supervision," *International Journal of Information Management*, vol. 52, p. 102024, 2020.
- [10] W. Meng, W. Li, and J. Zhou, "Enhancing the security of blockchain-based software defined networking through trust-based traffic fusion and filtration," *Information Fusion*, vol. 70, pp. 60–71, 2021.
- [11] H. H. Sun Yin, K. Langenheldt, M. Harlev, R. R. Mukkamala, and R. Vatrupu, "Regulating cryptocurrencies: a supervised machine learning approach to de-anonymizing the bitcoin blockchain," *Journal of Management Information Systems*, vol. 36, no. 1, pp. 37–73, 2019.
- [12] T. Ma, H. Xu, and P. Li, "Skyeye: a traceable scheme for blockchain," *Cryptology ePrint Archive: Report 2020/034, International Association for Cryptologic Research (IACR)*, vol. 2020, 34 pages, 2020.
- [13] T. Ma, H. Xu, and P. Li, "A Blockchain Traceable Scheme with Oversight Function," in *International Conference on Information and Communications Security*, pp. 164–182, Springer, Copenhagen, Denmark, 2020.
- [14] W. Meng, W. Li, S. Tug, and J. Tan, "Towards blockchain-enabled single character frequency-based exclusive signature matching in IoT-assisted smart cities," *Journal of Parallel and Distributed Computing*, vol. 144, pp. 268–277, 2020.
- [15] E. B. Sasson, A. Chiesa, C. Garman et al., "Zerocash: decentralized anonymous payments from bitcoin," in *2014 IEEE Symposium on Security and Privacy*, pp. 459–474, Berkeley, CA, USA, May 2014.
- [16] C. Cachin, "Architecture of the hyperledger blockchain fabric," in *Workshop on distributed cryptocurrencies and consensus ledgers*, vol. 310, pp. 1–4, Chicago, IL, 2016.
- [17] C. D. Clack, V. A. Bakshi, and L. Braine, "Smart contract templates: foundations, design landscape and research directions," 2016, <https://arxiv.org/abs/1608.00771>.
- [18] J. Woodcock, P. G. Larsen, J. Bicarregui, and J. Fitzgerald, "Formal methods," *ACM Computing Surveys*, vol. 41, no. 4, pp. 1–36, 2009.
- [19] T. Chen, X. Li, X. Luo, and X. Zhang, "Under-optimized smart contracts devour your money," in *2017 IEEE 24th International Conference on Software Analysis, Evolution and Reengineering (SANER)*, pp. 442–446, Klagenfurt, Austria, February 2017.
- [20] R. Cheng, F. Zhang, J. Kos et al., "Ekiden: a platform for confidentiality-preserving, trustworthy, and performant smart contracts," in *2019 IEEE European Symposium on Security and Privacy (EuroSecP)*, pp. 185–200, Stockholm, Sweden, June 2019.
- [21] A. Boldyreva, V. Goyal, and V. Kumar, "Identity-based encryption with efficient revocation," in *Proceedings of the 15th ACM Conference on Computer and Communications Security - CCS '08*, pp. 417–426, Alexandria, Virginia, USA, 2008.
- [22] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in *Proceedings of 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 457–473, Aarhus, Denmark, 2005.
- [23] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryption for finegrained access control of encrypted data," in *Proceedings of the 13th ACM Conference on Computer and Communications Security - CCS'06*, A. Juels, R. N. Wright, and S. De Capitani di Vimercati, Eds., pp. 89–98, ACM, Alexandria, VA, USA, 2006.
- [24] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *2007 IEEE Symposium on Security and Privacy (SP '07)*, pp. 321–334, Berkeley, CA, USA, May 2007.
- [25] B. Waters, "Ciphertext-policy attribute-based encryption: an expressive, efficient, and provably secure realization," in *Proceedings of International Workshop on Public Key Cryptography*, D. Catalano, F. Fazio, R. Gennaro, and A. Nicolosi, Eds., pp. 53–70, Springer, Berlin, Heidelberg, 2011.
- [26] S. He, H. Mao, X. Bai, Z. Chen, K. Hu, and W. Yu, "Multi-blockchain model for central bank digital currency," in *2017 18th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT)*, pp. 360–367, Taipei, Taiwan, December 2017.
- [27] S. Peng, X. Hu, J. Zhang et al., "An efficient double-layer blockchain method for vaccine production supervision," *IEEE Transactions on NanoBioscience*, vol. 19, no. 3, pp. 579–587, 2020.
- [28] S. Agrawal and M. Chase, "Fame: fast attribute-based message encryption," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security - CCS'17*, pp. 665–682, New York, NY, USA, 2017.



## Research Article

# Sharing Pandemic Vaccination Certificates through Blockchain: Case Study and Performance Evaluation

José L. Hernández-Ramos , Georgios Karopoulos , Dimitris Geneiatakis , Tania Martin, Georgios Kambourakis , and Igor Nai Fovino

*European Commission, Joint Research Centre, Ispra 21027, Italy*

Correspondence should be addressed to José L. Hernández-Ramos; [jose-luis.hernandez-ramos@ec.europa.eu](mailto:jose-luis.hernandez-ramos@ec.europa.eu)

Received 3 June 2021; Accepted 2 August 2021; Published 26 August 2021

Academic Editor: Wenjuan Li

Copyright © 2021 José L. Hernández-Ramos et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

During 2021, different worldwide initiatives have been established for the development of digital vaccination certificates to alleviate the restrictions associated with the COVID-19 pandemic to vaccinated individuals. Although diverse technologies can be considered for the deployment of such certificates, the use of blockchain has been suggested as a promising approach due to its decentralization and transparency features. However, the proposed solutions often lack realistic experimental evaluation that could help to determine possible practical challenges for the deployment of a blockchain platform for this purpose. To fill this gap, this work introduces a scalable, blockchain-based platform for the secure sharing of COVID-19 or other disease vaccination certificates. As an indicative use case, we emulate a large-scale deployment by considering the countries of the European Union. The platform is evaluated through extensive experiments measuring computing resource usage, network response time, and bandwidth. Based on the results, the proposed scheme shows satisfactory performance across all major evaluation criteria, suggesting that it can set the pace for real implementations. *Vis-à-vis* the related work, the proposed platform is novel, especially through the prism of a large-scale, full-fledged implementation and its assessment.

## 1. Introduction

The World Health Organization (WHO) declared COVID-19 a pandemic on March 11th, 2020. This disease is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) which was initially detected at the end of 2019 in the city of Wuhan, China [1]. Since then, the disease has spread unhindered worldwide. Besides the obvious health consequences, the socioeconomic impact is already notable in many countries globally. Indeed, the drastic—and sometimes controversial—measures to curb the spread, including social distancing and curfew, have already changed our daily behavior. Furthermore, recent economic analysis [2] predicts that many countries will not recover their economic levels of 2019 until 2022. These forecasts may vary based on the evolution of the pandemic during 2021.

To defend against the COVID-19 pandemic, several initiatives and actions have been hitherto undertaken, includ-

ing rapid diagnosis and isolation of infected people, as well as the creation of digital contact tracing frameworks [3, 4]. However, the second COVID-19 wave during the fall of 2020 and the successive outbreaks during 2021 showed that these measures are insufficient, especially when they are abruptly relaxed. Therefore, the sheer objective has been the development of effective and safe vaccines to be rolled out globally. Indeed, numerous efforts were initiated during 2020 involving medical institutions, pharmaceutical companies, and research centers worldwide to get a vaccine at unprecedented speed. At the end of May 2021 [5], there were 101 and 184 vaccines in clinical and preclinical development, respectively.

While the realisation of vaccines represents currently the main objective to terminate the pandemic, their manufacturing, distribution, and deployment are also associated with important challenges. First, logistics, storage, and transport requirements, say, regarding the temperature of

preservation, impose strong pressure on the supply chain to ensure global access to vaccines in a timely manner [6, 7]. Therefore, data transparency is the key to foster a secure monitoring of the epidemiological and vaccination situation in a certain region. Second, the vaccination process is being prioritised for certain population groups according to different aspects, such as age, health condition, and profession. Furthermore, the rate of vaccinations varies depending on the country [8]. Hence, immune and vulnerable people will live together during a certain period of time. Such a situation could be prolonged in case the virus that causes COVID-19 becomes endemic [9]. In this context, the use of digital vaccination certificates could help alleviate the burden on health systems, as vaccinated people would not need to perform viral tests, which are currently required to, say, travel to different countries. Unlike the current paper version of vaccination certificates, namely, the International Certificate of Vaccination or Prophylaxis (ICVP), these digital documents would allow a far more scalable solution along with a faster and more secure verification process [10].

Blockchain technology has been already identified as a promising approach to combat the pandemic in distinct scenarios, such as early detection of outbreaks, medical supply chain, or donation tracking [11, 12]. In the same mindset, the creation of a blockchain platform to share information about the pandemic would increase transparency, interoperability, and accountability, so that potential discrepancies among data from different sources, say, medical centers or governments, could be avoided. This would foster a more trustworthy reporting and monitoring of the pandemic evolution considering diverse territories and countries. Furthermore, such a platform would increase citizens' trust in the vaccination process, as the information related to vaccines could be publicly available [13].

The work at hand analyses the key requirements to build a scalable platform for sharing vaccination data and the advantages of blockchain for the realisation of such a platform. We focus on the scenario of vaccination certificates that can be generated after a citizen is vaccinated and how blockchain could aid in maintaining such information towards enabling a secure and privacy-aware verification process. Furthermore, unlike existing approaches that do not offer experimental results or consider small-scale deployment scenarios [14–17], we provide a comprehensive performance evaluation of the proposed platform by considering the vaccination of the EU population and 27 blockchain nodes, representing each member state (MS) in the EU. We meticulously assess our platform under different realistic network conditions, including latency and bandwidth, in an emulated infrastructure. To our knowledge, this is the first work to offer an estimation of the performance requirements associated with a blockchain-based platform for vaccination data in a large scale. Furthermore, we discuss practical aspects and security considerations for a large-scale deployment of the intended platform, along with potential regulatory implications of vaccination certificates.

The structure of this paper is as follows: the next section describes other works using blockchain technology for COVID-19-related certificates. Section 3 elaborates on the

needs for sharing COVID-19 information, as well as the advantages provided by blockchain for this purpose. Furthermore, Section 4 provides insights into the definition of digital vaccination certificates. The proposed blockchain platform for the registration and validation of digital vaccination certificates is described in Section 5. The results derived from the platform's evaluation are described in Section 6. Finally, Section 7 concludes our work with an outlook of potential future research directions.

## 2. Related Work

Since the beginning of the COVID-19 pandemic, different initiatives have been proposed for the implementation of COVID-19 certificates, so that individuals granted with such credentials could be exempt from physical restrictions to carry out certain activities in their daily life [18]. Indeed, based on our analysis of existing literature, three different types of COVID-19-related certificates can be identified: (a) *vaccination certificates*, referring to whether a person has received the vaccine or not; (b) *diagnostic test certificates*, demonstrating that a person has undergone a test; and (c) *immunity certificates* or *immunity passports*, implying that a person has developed antibodies after being infected. As shown in Table 1, some proposals support more than one type of certificates, while only a few of them provide an actual implementation, although in a small scale.

In the case of vaccination certificates, the authors of [19] focus on privacy aspects and propose a hashing algorithm that enables users to store the information on the blockchain anonymously using an ID that is created from their iris. In this case, the vaccination certificate data and a hash of the user ID are stored on the blockchain. This could imply a potential issue since it would demand a very high storage requirements of the blockchain nodes. This could be exacerbated in the case of populous or multiple countries using the same blockchain.

Furthermore, other works address several kinds of certificates. In particular, both vaccination and immunity certificates are considered by [15], which is based on Verifiable Credentials (VC) [22] as digital IDs, the decentralised data storage platform *Solid* [23], and a consortium Ethereum-based blockchain [24]. In a similar direction, [16] uses Ethereum smart contracts, Self-Sovereign Identity (SSI), and InterPlanetary File System (IPFS) to store medical tests and travel history in a decentralised manner. In addition, [14] addresses all the different types of certificates by integrating the use of VCs in a blockchain implementation called *uPort* [25], which provides SSI aspects on top of the Ethereum platform.

The authors of [20] introduce the concept of digital health passports, which is similar to the diagnostic test results required for travelers in certain cases. It is based on a private blockchain using the proof-of-authority consensus mechanism, where the test results are registered and stored.

For immunity certificates, the work of [17] presents SecureABC, a privacy-oriented protocol based on public key cryptography. This proposal does not use blockchain, and the certificates can be either paper- or app-based. As a

TABLE 1: Related work on COVID-19 certificates.

Scheme	Vaccination	Diagnostic test	Immunity	Blockchain	Benchmarks
[19]	✓			✓	—
[15]	✓		✓	✓	Small scale
[16]	✓		✓	✓	Small scale
[14]	✓	✓	✓	✓	Small scale
[20]		✓		✓	—
[17]			✓		Small scale
[21]			✓	✓	—
[10]		✓	✓		—

consequence, if the paper certificate or the mobile device is lost, so are the respective certificates. In [21], the concept of COVID-19 immunity certificates is based on a government-run blockchain, in which the information related to testing facilities and hospitals is also included. Furthermore, [10] proposed the use of VCs and Decentralized Identifiers (DID) [26] to link individuals' identity with their certificates. However, further details about implementation/-deployment aspects are not given.

In spite of recent efforts, only a few of these works present technical details or proof-of-concept implementation including evaluation results. For instance, they rather provide simple short high-level descriptions of the proposed solutions, or unconvincing benchmarks, limited to a small number of simultaneous requests, thus being far from real-world deployment scenarios. In contrast, our work tackles this problem through a comprehensive evaluation of a benchmark that includes 27 blockchain nodes (one node for each EU country) by considering different aspects, such as computing resource usage, network response time, and bandwidth. As highlighted by [12], even if the potential of blockchain to combat the COVID-19 pandemic has been reported by several works, there is a lack of studies related to latency and scalability aspects, which are key aspects for the deployment of this technology. Furthermore, our work concentrates on vaccination certificates, influenced by the views of WHO on immunity passports: "...there is not enough evidence about the effectiveness of antibody-mediated immunity to guarantee the accuracy of an 'immunity passport'." Another aspect driving us to this direction is that vaccination certificates will incite people to get vaccinated, while immunity certificates could motivate individuals get infected for possessing the necessary antibodies.

### 3. Managing COVID-19 Information through Blockchain

The global deployment of COVID-19 vaccines sets out unprecedented challenges to be addressed in the period ahead, including an efficient supply chain and effective monitoring of vaccination coverage in a certain region. Indeed, in the case of two-shot vaccines, more than 15 billion vaccines would be required to be distributed and deployed worldwide. Furthermore, the distribution of additional shots could be required depending on the immunity period provided by a certain vaccine or in case the virus that causes COVID-19

becomes endemic [9]. In this context, the WHO established the COVAX program together with *Gavi* and the Coalition for Epidemic Preparedness Innovations to facilitate equitable access and distribution of future vaccines, while those people most at risk are prioritized. COVAX is part of the global ACT Accelerator initiative that is designed to enhance the resources for COVID-19 tests, treatments, and vaccines.

At the European level, the commission published in Oct. 2020 a document on COVID-19 vaccination strategies and vaccine deployment for the 27 MS [27]. This document established the need to define a common strategy for the vaccination process, promoting coordination and collaboration among EU countries. One of the main goals of this strategy is to increase the acceptance of COVID-19 vaccines. Actually, recent studies reveal that a significant part of the population would not be willing to be vaccinated against the COVID-19 disease [13]. To address this issue, there is a need for an effective, consistent, and transparent communication of information related to COVID-19 vaccines and the vaccination process itself. As described in [27], the sharing of pandemic-related information among MS would cater for a better monitoring of the different vaccines under development, including data on possible side effects, which would be made readily available to the relevant authorities. Furthermore, this information could include data on the transport and distribution of vaccines to enable real-time monitoring and improve the supply chain process by considering the specific needs of each vaccine.

Moreover, vaccination campaigns have been carried out by considering different aspects (e.g., age or medical condition) established by organisations such as the WHO's Strategic Advisory Group of Experts on Immunization to prioritize the vaccination for certain groups of people. Therefore, currently a large part of the population is still vulnerable to the COVID-19. This situation is especially exacerbated in developing countries [8]. Furthermore, depending on the immunity period of each vaccine, the immunity of a certain person could come to an end at a certain point in time. Beyond the information on vaccines, the easy sharing of these vaccination data would improve the monitoring of the epidemiological situation of a territory and the vaccination coverage among different population groups. In fact, monitoring these aspects can make the vaccination strategy more flexible to be adapted in a certain region or country [27].

For the realisation of this COVID-19 data sharing platform, blockchain technology has been postulated in different related scenarios, including contact tracing and outbreaks, where information sharing is essential [28]. Blockchain is based on a distributed ledger that is shared by a set of entities. The ledger contains a list of immutable transactions that are validated by the participating entities through a consensus mechanism. Furthermore, a blockchain can be permissionless (any entity can participate) or permissioned (participation is limited to a set of entities). The development of a blockchain-based platform offers a high degree of transparency and accountability, fostering a trustworthy environment for the sharing of COVID-19 data.

Thus far, the use of blockchain to fight against the COVID-19 pandemic has been proposed for several use cases, including the distribution and delivery of vaccines, recording of patients' data, preventing fake news, registration of testing and reporting, and the distribution of medicines and healthcare equipment [11, 28–31]. While a blockchain platform for sharing pandemic data could help in distinct scenarios, we focus on the registration and verification process of the data associated with a vaccinated citizen. The envisioned platform will enable a trusted ecosystem to track the deployment of vaccines in a certain region and consider priority groups. That is, blockchain inherently supports decentralisation and data replication (data from all countries are replicated to all other countries), deterring issuance of fraudulent vaccination certificates as well. For this purpose, we examine the concept of digital vaccination certificates that could be demonstrated by citizens to carry out certain activities without the need of diagnostic tests. The following sections describe the design and architecture of a blockchain platform for digital vaccination certificates, as well as a thorough evaluation where each MS is represented by a blockchain node.

#### 4. Digital Vaccination Certificates

Digital vaccination certificates can be viewed as a digital version of the ICVP certificates created by WHO that show a person's vaccines and the date they were received. For the representation of such a certificate, there is a need to identify which specific information should be included, so that they can be used across the world. Such certificates should be interoperable globally, as well as supported by identity management techniques to unequivocally link the vaccination of citizens with their identity; in this way, the resulting certificate will be verifiable, scalable, and privacy-preserving.

The European Commission proposed a Digital Green Certificate in March 2021 [32] to facilitate safe and free movement inside the EU during the COVID-19 pandemic. Furthermore, the eHealth Network, which provides a platform of EU MSs' competent authorities dealing with eHealth, has recently described a set of guidelines on verifiable vaccination certificates, including trust and interoperability aspects. Precisely, [33] identifies a minimum dataset with the essential pieces of information to be embedded in the certificate, including person identification (e.g., citizen

ID), vaccination information (e.g., vaccine manufacturer), and certificate metadata, such as issuer and validity period.

Other worldwide initiatives have been established for the development of digital vaccination certificates. In particular, the WHO Smart Vaccination Certificate Working Group [34] is intended to define standard specifications for digital vaccination certificates based on an architecture linking national and crossborder digital systems. Furthermore, the IATA Travel Pass Initiative [35] provides a mobile app to be used by travelers to store and manage their verified certifications for COVID-19 tests or vaccines. Another relevant effort is represented by the *Certify.health* initiative [36], which concentrates on the development of a privacy-by-design COVID-19 status certificate that will be extended into vaccination certificates.

For the representation of digital vaccination certificates, several formats could be considered. For example, [33] mentions QR codes and Verifiable Credentials (VC), which have been also considered by recent research proposals, as described in Section 2. The use of VC (together with DIDs) is intended to realize the vision of Self-Sovereign Identity (SSI), which has emerged as a decentralised alternative to traditional centralised identity management (IdM) systems. A VC represents a digital version of a paper certificate in which a certain entity (issuer) asserts certain information (claims) about a subject in a way that can be verified by other entities (verifiers). A VC is usually employed together with a DID, which is an identifier under the control of a DID subject that indicates a DID method and a specific identifier of such method. DIDs are registered in a Verifiable Data Registry (VDR), such as blockchain, and are intended to foster a decentralised authentication process.

It should be noted that the use of VCs in the context of the COVID-19 crisis has been fostered by the COVID-19 Credentials Initiative [37], which groups around 100 organisations to support efforts of using VCs to mitigate the spread of the virus.

While it is not the focus of our work, Figure 1 shows an example of VC that includes certain claims based on ongoing discussions about the use of VCs for vaccination certificates. In our example, we have considered that the validity of the certificate is associated with the period during which this vaccination is effective, taking into account that two shots are required. In particular, the *context* establishes a common language for referring to the attributes and values contained in the VC. Also, for our example shown in the figure, the URI <https://covid-19-vaccination-certificate.org/v1> indicates that the communication is about vaccination certificates. Furthermore, the *id* and *type* fields are used to identify the VC and indicate its type. Moreover, the *issuer* represents the entity that issued the VC and it makes reference to the medical center, which provided the vaccine. In this case, it is described through a DID that could be included in the blockchain, so that verifiers can use this information to validate the VC. This field can also indicate the type and name of the issuer, as well as its URL for more information. Besides, the *issuanceDate* and *expirationDate* indicate the validity of the certificate that is associated with the immunity period provided by the vaccine. Also, the *CredentialSubject* represents the entity on which the claims are made, i.e., the

```

{
  "@context": [
    "https://www.w3.org/2018/credentials/v1",
    "https://covid-19-vaccination-certificate.org/v1"
  ],
  "id": "https://covid-19-vaccination-certificate.org/
    credentials/JohnDoe",
  "type": [
    "VerifiableCredential",
    "VaccinationCertificate"
  ],
  "issuer": {
    "id": "did:web:vc.brussels.vaccination.centre",
    "location": {
      "type": "MedicalCenter",
      "name": "BrusselsVaccinationCentre",
      "url": "https://brussels-vaccination-centre.org/"
    }
  },
  "issuanceDate": "2020-01-31T14:30:23",
  "expirationDate": "2020-07-31T14:30:23",
  "name": "VaccinationCertificate",
  "description": "Electronic document certifying that the subject
    fulfilled the COVID-19 vaccination procedure.",
  "credentialSubject": {
    "id": "did:key:subject_key_value",
    "type": "VaccinationCertificateSubject",
    "givenName": "John",
    "familyName": "Doe",
    "birthDate": "1979-05-28",
    "image": "data:image/png;base64, image_value",
  },
  "injection": {
    "id": "injection_id",
    "type": "VaccinationCertificateInjection",
    "name": "vaccine_name",
    "issuanceDate": "2020-01-10T11:15:46",
  },
  "proof": {
    "type": "Ed25519Signature2018",
    "created": "2020-01-31T14:30:23",
    "jws": "JSON_Web_signature_value",
    "proofPurpose": "assertionMethod",
    "verificationMethod": "did:web:vc.brussels.vaccination.
      centre # additional_id_value"
  }
}

```

FIGURE 1: Example of a potential vaccination certificate based on VCs [22].

individual getting vaccinated, that includes the personal data about the user. In addition, the claim *injection* is used to describe which specific injection is being provided, including the vaccine and vaccination date. This information could be used to track the vaccines and injections being provided and may help with the management of the supply chain. Finally, the field *proof* makes reference to the cryptographic technique (typically a digital signature) that is used by the issuer to make the VC tamper-resistant.

While the design of an interoperable approach for the definition of digital vaccination certificates is still under dis-

cussion, in our approach, only a hash digest of such a credential will be stored in the blockchain platform. In this way, the proposed platform will be agnostic both of the vaccination certificate presentation format and of the data format being considered. The details of such a platform are described in the subsequent section.

## 5. Vaccination Certificate Scenario

For the development of the proposed blockchain platform, we consider the architecture in Figure 2. Naturally, the

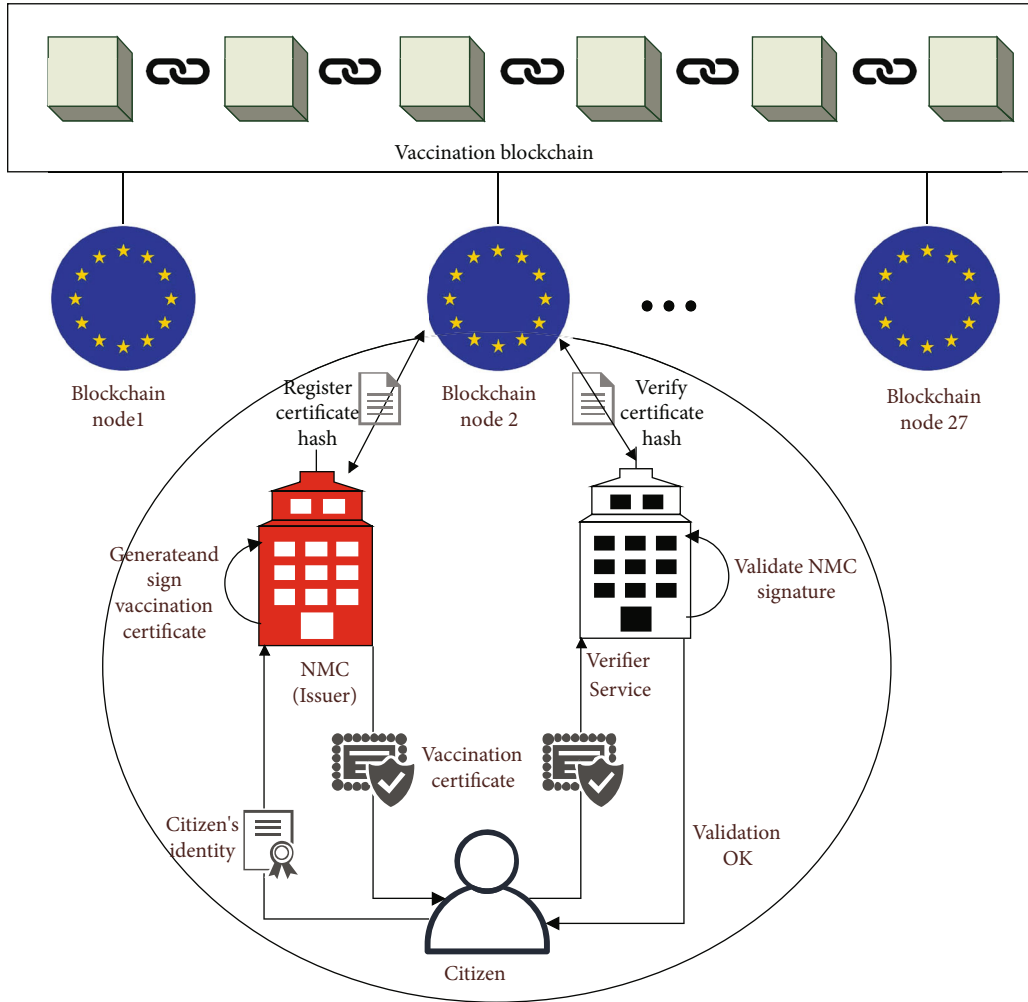


FIGURE 2: Overview of the proposed blockchain-based vaccination certificate platform.

depicted architecture does not reflect the reality of any decision made at the EU level, but it solely serves as a proof of concept for evaluation purposes. The architecture includes the vaccination permissioned blockchain where blockchain nodes in each MS are the only authorised entities to store vaccination information.

Each of the 27 MSs can designate a blockchain client node (which can be represented by a national health authority, say, the Ministry of Health) to interact with the blockchain. This entity is also responsible for designating a set of national medical centers (NMC) to generate vaccination certificates associated with already vaccinated people. These certificates will be validated by *verification centers*, which represent any organisation, public or private (e.g., airport or public administration building), that needs to verify the vaccination status of an individual.

The blockchain is used to store all relevant information about the vaccination process, including the registration of NMC. The registration of these entities can be performed by the national health authorities, which represent the blockchain nodes of their MS, by using smart contracts. Furthermore, the blockchain will simply contain a hash digest of the vaccination certificate per citizen that will be generated dur-

ing the registration process and used later to facilitate the process of verifying the vaccination status of them. It is noteworthy that the registration and verification processes analysed below are only illustrative examples of how our scheme can be used to manage vaccination certificates, while these processes are considered for evaluation purposes in the next section.

During the *registration* process, citizens go to an NMC, where they present a valid identity document, to get vaccinated. For this purpose, the citizen may use a VC through a digital wallet app on their smartphone (as proposed by [10]) or other more traditional approaches based on X.509 certificates. A physician performs the vaccination, and the corresponding certificate is generated. As described in Section 4, this certificate may contain information about the vaccine itself, as well as data about the specific dose to facilitate the management of the supply chain. Furthermore, the citizen's identity shown at the beginning of the process can be embedded in the credential. Assuming a two-shot vaccine, this credential may demonstrate that the citizen received the first shot, so it can be used in the process of administering the second one, or that they are immune as they already received both shots. The NMC, or the physician

on its behalf, digitally signs the certificate to guarantee its validity and sends this credential to the citizen, say, through a smartphone app, so that they can maintain the control of how the certificate is used. It should be noted that the process of sending the vaccination certificate is done through a secure channel by using well-known approaches, such as Transport Layer Security (TLS). Furthermore, the certificate could be encrypted before being stored in the user’s smartphone to protect the credentials while at rest. Moreover, a hash digest (e.g., by using SHA-256) of the certificate is generated and stored on the blockchain. The NMC sends this hash to the MS’s blockchain node that is responsible for registering it on the EU vaccination blockchain. Additionally, an encrypted version can be stored in the InterPlanetary File System (IPFS) [38] or another repository, so that the vaccination certificate can be recovered by the citizen in case of losing their smartphone. Again, such processes are carried out by using renowned approaches, such as TLS to protect the data in transit.

After citizens have received a certificate, they can use it to access certain places that require proof of citizens’ vaccination status, such as an airport or public administration building. During the *verification* process, citizens present their certificate to a verifier service. This service creates a hash digest of the provided certificate that is verified against the hash stored in the blockchain. For this process, the verifier service contacts the country’s blockchain node that is in control of performing the verification on the EU vaccination blockchain. Like in the issuance procedure, this process is performed through well-known security mechanisms, such as TLS. Furthermore, the verifier service validates the signature created by the NMC to confirm that the credential was generated by an approved entity. Additionally, it checks the validity of the citizen’s identity to ensure that they are indeed the person associated with the credential presented.

Alternatively, citizens are empowered to show a subset of their identity attributes by using zero-knowledge proofs (ZKPs) to access certain places that only require confirmation of a person’s vaccination status but do not need personal data. For example, in the case of VCs, the holder of a certain credential is enabled to combine several VCs from different issuers and selectively disclose specific claims composing a certain VC. However, this aspect is outside the scope of this work. Indeed, as described in the next section, the evaluation of our platform is focused on the performance requirements from the perspective of the blockchain implementation to register and verify vaccination certificates. Nevertheless, it should be noted that the proposed blockchain platform is intended to serve as a decentralized approach to manage vaccination information and to be integrated with SSI approaches, such as VCs, for the sake of providing privacy-preserving features. Furthermore, as already mentioned, only a hash of the vaccination certificate is stored on the blockchain, and an encrypted version of such certificate is stored on an off-chain repository (IPFS), so that users’ sensitive data is never disclosed to external entities. Therefore, citizens are enabled with the ownership of their data to manage their vaccination certificates. The integration of the proposed platform with SSI approaches, such as VCs and DIDs, will enable a more advanced privacy-preserving

TABLE 2: Network evaluation of registering and verifying vaccination certificates using blockchain.

Step	TPS	Response time (msec)	Peer bandwidth (kB)	Ordering bandwidth (kB)
Register	1	84	395	636
	2	81	419	825
	4	78	457	2019
	8	87	516	3644
	16	109	588	4938
	28	133	700	6019
	1	91	394	701
Verify	2	87	415	1123
	4	83	447	1788
	8	94	495	2153
	16	117	553	5069
	28	153	639	8122
	50	168	671	5919
	100	189	804	12109

approach for the issuance and verification processes through the integration of ZKPs in the whole ecosystem.

## 6. Evaluation

**6.1. Testbed.** To evaluate our proposal, we rely on the Experimental Platform for Internet Contingencies (EPIC) [39]. EPIC is an emulation testbed based on the Deter software [40, 41] for studying the security and stability of distributed systems. The use of emulation-based testbeds in cybersecurity is well established [42–44] and ensures repeatability and measurement accuracy. Furthermore, this approach was chosen for the sake of overcoming the major difficulties that arise while trying to simulate the behaviour of ICT components under stress, attacks, or failures. The infrastructure of EPIC comprises 356 experimental nodes, 8 switches, and a few special equipment, such as programmable logical controllers.

Overall, the setup relies on the deployment of Hyperledger Fabric on an emulated network in EPIC and implements the proposed architecture shown in Figure 2. It is assumed that the European health authorities, which are considered trusted, provide the “ordering” services, while each MS is a “peer” node in the Hyperledger Fabric terminology. This emulated 1 Gbps blockchain network comprises 27 nodes corresponding to the current EU MS with a network latency of 3 msec.

The ordering services comprise the following: ZooKeeper (3 instances), Kafka (4 instances), and orderer (3 instances). Their main purpose is to sort the messages/requests exchanged among the participants. Each instance of a given service runs on a different machine for supporting failover of the ordering services. This setup ensures ordering service availability if at maximum one instance of each service is in the fail status. The peer nodes are managed by the MSs for endorsing the transactions proposed by the clients. They also receive the ordered blocks of transactions

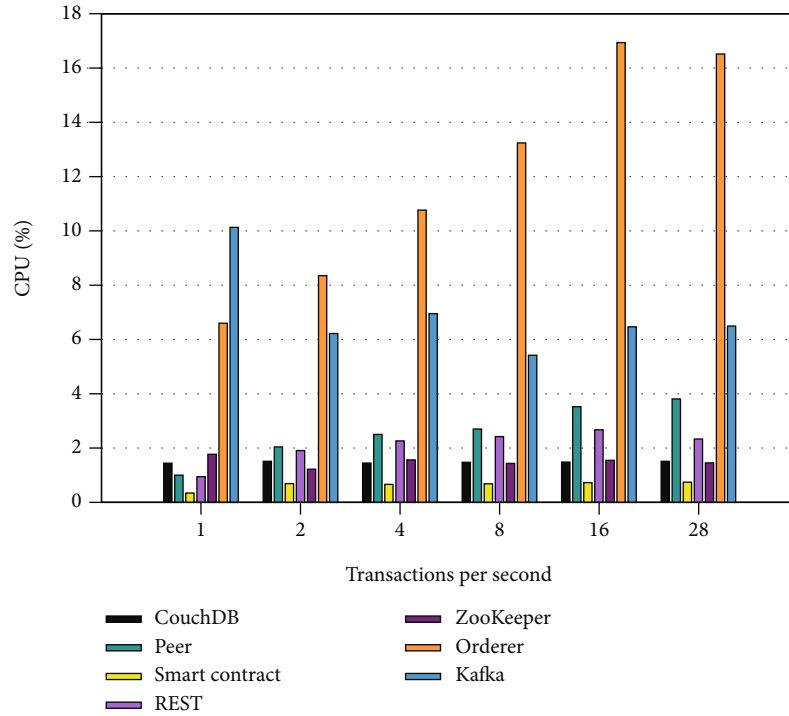


FIGURE 3: Dockerised services' CPU utilisation considering different TPS for registering new vaccination certificates in a blockchain system.

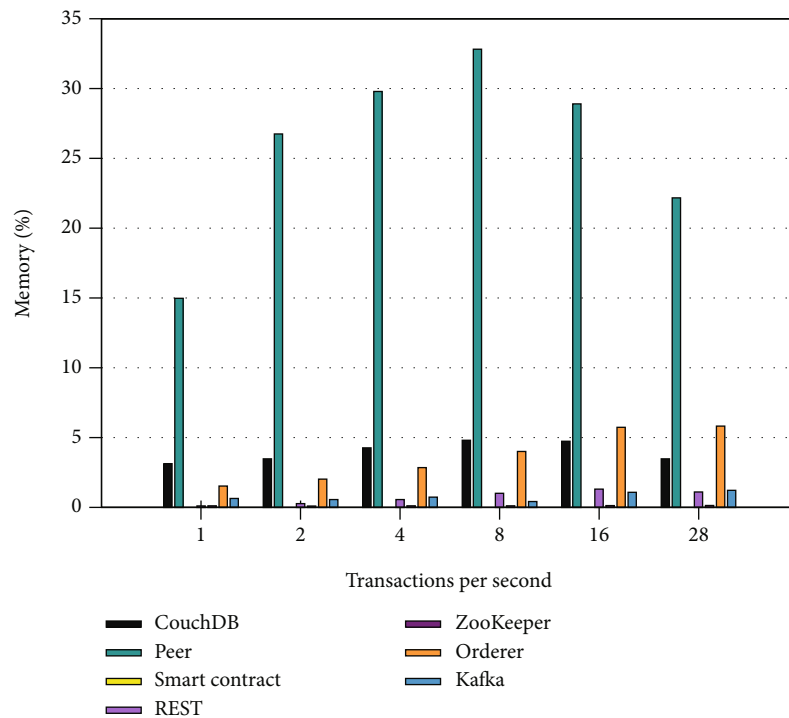


FIGURE 4: Dockerised services' memory utilisation considering different TPS for registering new vaccination certificates in a blockchain system.

from the ordering service to maintain their local copy of the ledger. The following services of a MS node are hosted on a single machine:

- (1) *CouchDB*: a database that maintains the valid transactions of the blockchain and allows content-based JSON queries



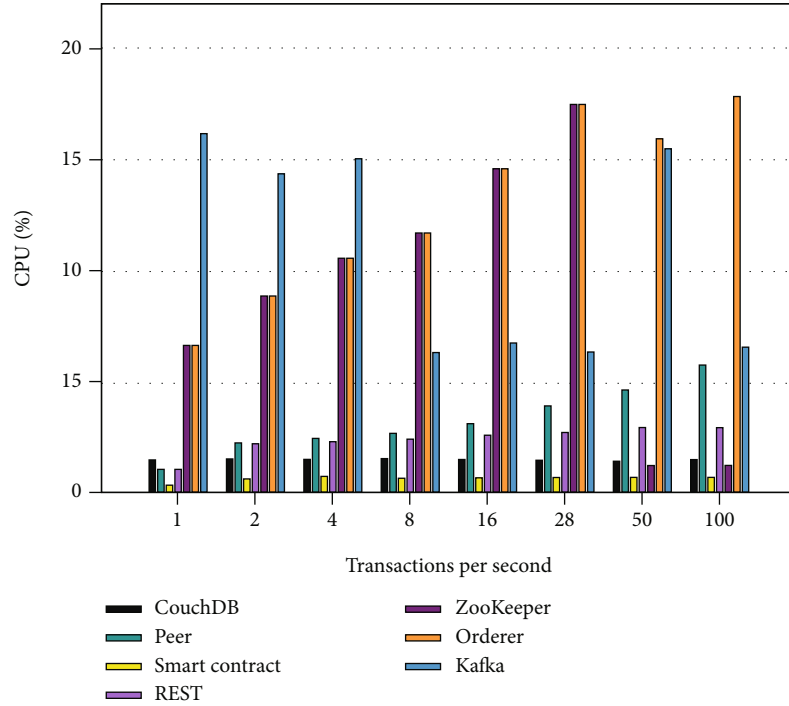


FIGURE 5: Dockerised services' CPU utilisation considering different TPS for verifying vaccination certificates in a blockchain system.

- (2) *Peer*: a core service in the Hyperledger Fabric architecture storing the ledger and validating the transactions
- (3) *Certificate authority*: this provides digital certificates to the participants of the MS node
- (4) *Smart contract*: this implements basic functionalities such as user access control and message conformity
- (5) *Application interface*: this interacts with the blockchain. It is implemented as a representational state transfer (REST) service and accomplishes all the interactions on behalf of the national health centers for committing a transaction in the blockchain network

As each MS acts independently, we deploy a single disjunctive (“OR”) policy among the participants, meaning that a transaction originating from a MS is only validated by the originating MS. What the system checks is whether the submitted transaction bears a valid digital signature from the MS blockchain node. This also means that any transaction stemming from a MS on behalf of another MS will be rejected by the blockchain.

All ordering and peer services are configured and executed using the corresponding Docker images with the standard deployment options. Moreover, all the underlying network communications among the participants (clients, peers, and the ordering service) are securely protected by Transport Layer Security (TLS). The certificates and private keys for both TLS and the blockchain services are generated during the blockchain network initialisation procedure, according to the Hyperledger Fabric specifications.

**6.2. Results.** We evaluate the adequacy of deploying our proposal in a real, large-scale architecture, concentrating on two fundamental provisioned services, namely, vaccination registration and verification. The focus is on user experience in terms of request round-trip time, i.e., the time required for receiving a response after submitting a request, and the utilisation of system resources, i.e., CPU, memory, and network bandwidth.

For the registration process, we consider the maximum number of transactions required to get all European citizens vaccinated in one year. According to Eurostat, the EU-27 population is  $\approx 447.5$  M inhabitants [45]. Thus, assuming that a vaccine requires two doses, that is, two blockchain transactions, a total of 28 transactions per second (TPS) will be required in the worst case. Table 2 summarises the average latency perceived when registering or verifying a vaccination certificate in the blockchain, as well as the bandwidth consumed by both the peer and the ordering nodes. As observed, the response time for registration ranges between 83 and 133 msec. Moreover, at the peer side, the bandwidth utilisation increases from 500 to 700 kB. Overall, both these numbers can be characterised as absolutely tolerable. On the other hand, the bandwidth consumed by the ordering service demonstrates a significant augmentation among the different TPS values, reaching  $\approx 6000$  kB in the most demanding case.

CPU and memory utilisation for registering new vaccination certificates under different traffic conditions per service are illustrated in Figures 3 and 4. Particularly, considering the worst case, CPU and memory utilisation for the peer services remain under 4 and 35%, respectively, while the ordering services' utilisation is under 17 and 7%. In any case, these requirements for both services are

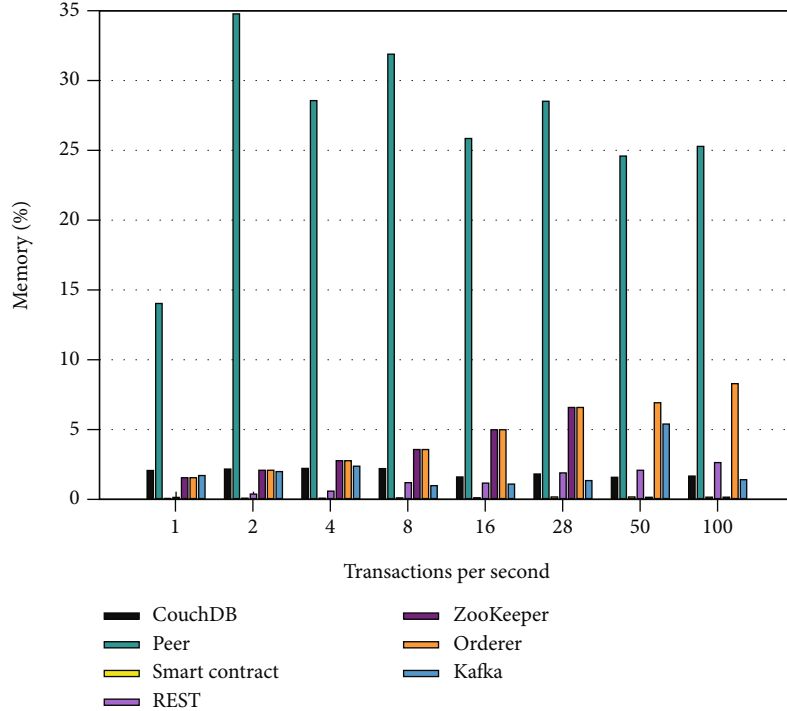


FIGURE 6: Dockerised services' memory utilisation considering different TPS for verifying vaccination certificates in a blockchain system.

manageable. It is also perceived that, when TPS increase from 8 to 16 and above, memory utilisation for the peer services starts to decrease. This can be explained by the fact that, along with TPS, the response time augments, having transactions submitted to the system at a lower rate. Interestingly and also on the positive side, CPU utilisation for the smart contract remains almost constant under different TPS, consuming less than 1% of the available CPU cycles. Overall, the registration process is more demanding in terms of CPU on the orderer and secondly on the Kafka services, while in terms of memory on the peer service.

Regarding vaccination certificate verification, we used data from Eurostat to calculate realistic requirements in terms of TPS. Specifically, we calculated the total number of air, marine, rail, and bus passengers for 2018, which is the latest year with data for all these categories. As verification transaction requests are forwarded to the national node of each MS, we consider the worst case, that is, the MS with the highest combined number of passengers in one year (3.2 billion); this gives us  $\approx 100$  TPS. Regarding the search operation, the worst case scenario is again followed; that is, the correct record is the last one. Similar to vaccination registration, the response time increases proportionally to the number of TPS, demonstrating a similar pattern. Overall, with reference to Table 2, the response time and bandwidth utilisation at a MS blockchain node fluctuate between 91 and 189 msec and 394 and 804 kB, respectively. However, for ordering, the utilised network bandwidth reaches up to 12,109 kB.

Figures 5 and 6 depict CPU and memory utilisation for vaccination certificate verification per blockchain service. As observed, CPU utilisation for both the peer and REST

services increases proportionally to TPS, while it is relatively stable for couchDB and smart contract. The orderer service initially increases and then stabilises, while the Kafka service fluctuates between 6 and 16%. However, in all cases, the CPU load remains under 18%.

As expected and similar to registration, memory usage for the peer service ranges between  $\approx 14\%$  and  $35\%$ , demonstrating that it is memory intensive. For the rest of the services, memory requirements are low, that is, under 8%. In summary, the verification process is more demanding in terms of CPU on the ordering services, while in terms of memory on the peer service.

## 7. Conclusions

The work at hand sheds light on the timely and intriguing issue of managing digital vaccination certificates on a large scale. After arguing that under the prism of COVID-19 and future epidemics, this need is rather a sine qua non, we specifically attempt to answer two key questions: how such an endeavour can be realistically organised with a focus on reducing complexity, and if so, would it be smooth-running under pragmatic conditions or even stress in terms of performance? For the first matter, we scrutinised on an envisaged wide-scale deployment capable of covering the needs of EU-27 and elaborated on a practical vaccination certificate scenario. For the second, we relied on the EPIC platform.

Specifically, based on the performance results obtained, including scalability aspects and challenges for the deployment of such platform, it is demonstrated that, for both registration and verification operations, the system achieves

satisfactory results even under stress. This strongly suggests that even a network decreased by one order of magnitude (100 Mbps) would be more than enough. Regarding CPU requirements, the ordering nodes need to be more powerful than MS ones, while the peer nodes necessitate more memory. Also, it is shown that, at least in a similar setup as our testbed, 100 TPS is the boundary, considering that above this limit, the system is saturated, producing errors and experiencing inconsistencies. This indicates that in most populated European countries, the MS node specifications should be carefully devised to support such a large number of TPS or even greater, if necessary.

Future work will concentrate more on the security, privacy, and ethical aspects associated with the registration and verification process of digital vaccination certificates. Also, an appealing direction is to investigate if this kind of platform could cater for the needs of the vaccine supply chain, ensuring efficient vaccine warehousing, handling, and stock administration.

### Data Availability

The data used to support the findings of this study are included within the article.

### Disclosure

A preliminary version of this paper can be found at [46].

### Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

### References

- [1] World Health Organization (WHO), *Timeline of WHO's Response to COVID-19*, 2020, <https://www.who.int/news/item/29-06-2020-covid-timeline>.
- [2] OECD Economic Outlook, *Interim Report September 2020*, OECD, 2020.
- [3] T. Martin, G. Karopoulos, J. L. Hernández-Ramos, G. Kambourakis, and I. N. Fovino, "Demystifying COVID-19 digital contact tracing: a survey on frameworks and mobile apps," *Wireless Communications and Mobile Computing*, vol. 2020, 29 pages, 2020.
- [4] V. Kouliaridis, G. Kambourakis, E. Chatzoglou, G. Dimitrios, and H. Wang, "Dissecting contact tracing apps in the android platform," *PLoS ONE*, vol. 16, no. 5, p. e0251867, 2021.
- [5] World Health Organization, "Draft landscape and tracker of COVID-19 candidate vaccines," WHO, 2021, <https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines>.
- [6] DHL, *DHL White Paper-Delivering Pandemic Resilience-How to Secure Stable Supply Chains for Vaccines and Medical Goods during the COVID-19 Crisis and Future Health Emergencies*, 2020, <https://www.dhl.com/content/dam/dhl/global/core/documents/pdf/glo-core-delivering-pandemic-resilience-2020.pdf>.
- [7] O. J. Wouters, K. C. Shadlen, M. Salcher-Konrad et al., "Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment," *The Lancet*, vol. 397, no. 10278, pp. 1023–1034, 2021.
- [8] *Tracking COVID-19 Vaccinations Worldwide* <https://edition.cnn.com/interactive/2021/health/global-covid-vaccinations/>.
- [9] N. Phillips, "The coronavirus is here to stay — here's what that means," *Nature*, vol. 590, no. 7846, pp. 382–384, 2021.
- [10] D. Gruener, *Immunity Certificates: If We Must Have Them, We Must Do It Right*, 2020.
- [11] A. Kalla, T. Hewa, R. A. Mishra, M. Ylianttila, and M. Liyanage, "The role of blockchain to fight against COVID-19," *IEEE Engineering Management Review*, vol. 48, no. 3, pp. 85–96, 2020.
- [12] A. A. Abd-alrazaq, M. Alajlani, D. Alhuwail et al., "Blockchain technologies to mitigate COVID-19 challenges: a scoping review," *Computer Methods and Programs in Biomedicine Update*, vol. 1, 2020.
- [13] J. V. Lazarus, S. C. Ratzan, A. Palayew et al., "A global survey of potential acceptance of a COVID-19 vaccine," *Nature Medicine*, vol. 27, no. 2, 2021.
- [14] A. Abid, S. Cheikhrouhou, S. Kallel, and M. Jmaiel, "Novid-chain: blockchain-based privacy-preserving platform for COVID-19 test/-vaccine certificates," *Software: Practice and Experience*, 2021.
- [15] M. Eisenstadt, M. Ramachandran, N. Chowdhury, A. Third, and J. Domingue, "COVID-19 antibody test/vaccination certification: there's an app for That," *IEEE Open Journal of Engineering in Medicine and Biology*, vol. 1, pp. 148–155, 2020.
- [16] H. R. Hasan, K. Salah, R. Jayaraman et al., "Blockchain-based solution for COVID-19 digital medical passports and immunity certificates," *IEEE Access*, vol. 8, pp. 222093–222108, 2020.
- [17] C. Hicks, D. Butler, C. Maple, and J. Crowcroft, "SecureABC: secure antibody certificates for COVID-19," 2020.
- [18] L. Alexandra, "COVID-19 immunity passports and vaccination certificates: scientific, equitable, and legal challenges," *The Lancet*, vol. 395, no. 10237, pp. 1595–1598, 2020.
- [19] S. Chaudhari, M. Clear, and H. Tewari, "Framework for a DLT based COVID-19 passport," in *Intelligent Computing. Lecture Notes in Networks and Systems*, K. Arai, Ed., vol. 285, Springer, Cham, 2021.
- [20] C. M. Angelopoulos, A. Damianou, and V. Katos, "DHP framework: digital health passports using blockchain-use case on international tourism during the COVID-19 pandemic," 2020, <http://arxiv.org/abs/2005.08922>.
- [21] A. Bansal, C. Garg, and R. P. Padappayil, "Optimizing the implementation of COVID-19 immunity certificates using blockchain," *Journal of Medical Systems*, vol. 44, no. 9, p. 140, 2020.
- [22] World Wide Web Consortium (W3C), "Verifiable credentials data model 1.0," 2019, <https://www.w3.org/TR/vc-data-model/>.
- [23] A. V. Sambra, E. Mansour, S. Hawke et al., "Solid: a platform for decentralized social applications based on linked data," *MIT CSAIL & Qatar Computing Research Institute, Tech. Rep.*, 2016.
- [24] G. Wood, "Ethereum: a secure decentralised generalised transaction ledger," *Ethereum project yellow paper*, vol. 151, 2014.
- [25] C. Lundkvist, R. Heck, J. Torstensson, Z. Mitton, and M. Sena, "Uport: a platform for self-sovereign identity," 2017, [https://whitepaper.uport.me/uPort\\_whitepaper\\_DRAFT20170221.pdf](https://whitepaper.uport.me/uPort_whitepaper_DRAFT20170221.pdf).

- [26] World Wide Web Consortium (W3C), “Decentralized identifiers (dids) v1.0- core architecture, data model, and representations,” 2021, <https://www.w3.org/TR/did-core/>.
- [27] European Commission, “Communication from the Commission to the European Parliament and the Council- preparedness for COVID-19 vaccination strategies and vaccine deployment,” 2020, [https://ec.europa.eu/health/sites/health/files/vaccination/docs/2020\\_strategies\\_deployment\\_en.pdf](https://ec.europa.eu/health/sites/health/files/vaccination/docs/2020_strategies_deployment_en.pdf).
- [28] D. Marbough, T. Abbasi, F. Maasmi et al., “Blockchain for COVID-19: review, opportunities and a trusted tracking system,” 2020, [https://www.techrxiv.org/articles/preprint/Blockchain\\_for\\_COVID-19\\_Review\\_Opportunités\\_and\\_a\\_Trusted\\_Tracking\\_System/12609344](https://www.techrxiv.org/articles/preprint/Blockchain_for_COVID-19_Review_Opportunités_and_a_Trusted_Tracking_System/12609344).
- [29] A. Musamih, R. Jayaraman, K. Salah, H. Hasan, I. Yaqoob, and Y. Al-Hammadi, “Blockchain-based solution for distribution and delivery of COVID-19 vaccines,” *IEEE Access*, 2021.
- [30] M. Chang and D. Park, “How can blockchain help people in the event of pandemics such as the COVID-19?,” *Journal of Medical Systems*, vol. 44, no. 5, p. 102, 2020.
- [31] D. Nguyen, M. Ding, P. N. Pathirana, and A. Seneviratne, “Blockchain and AI-based solutions to combat coronavirus (COVID-19)-like epidemics: a survey,” *TechRxiv Preprint*, vol. 4, 2020.
- [32] European Commission, “Proposal for a Regulation of the European Parliament and of the Council on a framework for the issuance, verification and acceptance of interoperable certificates on vaccination, testing and recovery to facilitate free movement during the COVID-19 pandemic (Digital Green Certificate),” 2021, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0130>.
- [33] EHealth Network, “Guidelines on verifiable vaccination certificates - basic interoperability elements release 2,” 2021, [https://ec.europa.eu/health/sites/default/files/ehealth/docs/vaccination-proof\\_interoperability-guidelines\\_en.pdf](https://ec.europa.eu/health/sites/default/files/ehealth/docs/vaccination-proof_interoperability-guidelines_en.pdf).
- [34] *Smart Vaccination Certificate Working Group* <https://www.who.int/groups/smart-vaccination-certificate-working-group>.
- [35] *IATA Travel Pass Initiative* <https://www.iata.org/en/programs/passenger/travel-pass/>.
- [36] *Certify.health* <https://eithealth.eu/project/certify-health/>.
- [37] *COVID-19 Credentials Initiative* <https://www.covidcreds.org/>.
- [38] J. Benet, *IPFS - Content Addressed, Versioned, P2P File System*, CoRR, 2014, <http://arxiv.org/abs/1407.3561>.
- [39] C. Siaterlis, B. Genge, and M. Hohenadel, “EPIC: a testbed for scientifically rigorous cyber-physical security experimentation,” *IEEE Transactions on Emerging Topics in Computing*, vol. 1, no. 2, pp. 319–330, 2013.
- [40] J. Mirkovic, T. V. Benzel, T. Faber, R. Braden, J. T. Wroclawski, and S. Schwab, “The DETER Project: advancing the science of cyber security experimentation and test,” in *In 2010 IEEE International Conference on Technologies for Homeland Security (HST)*, pp. 1–7, 2010.
- [41] T. Benzel, “The science of cyber security experimentation: the DETER Project,” in *In 27th Annual Computer Security Applications Conference*, 2011.
- [42] C. M. Davis, J. E. Tate, H. Okhravi, C. Grier, T. J. Overbye, and D. Nicol, “Scada cyber security testbed development,” in *In 2006 38th north American power symposium*, pp. 483–488, 2006.
- [43] T. C. Eskridge, M. M. Carvalho, E. Stoner, T. Toggweiler, and A. Granados, “Vine: a cyber emulation environment for MTD experimentation,” in *In Proceedings of the Second ACM Workshop on Moving Target Defense, MTD 15*, pp. 43–47, New York, NY, USA, 2015.
- [44] K. E. Stewart, J. W. Humphries, and T. R. Anandel, “Developing a virtualization platform for courses in networking, systems administration and cyber security education,” in *in Proceedings of the 2009 Spring Simulation Multiconference, ser, SpringSim '09*. San Diego, CA, USA: Society for Computer Simulation International, 2009.
- [45] Eurostat, “Population and population change statistics 2021,” 2021, [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population\\_and\\_population\\_change\\_statistics](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_and_population_change_statistics).
- [46] J. L. Hernández-Ramos, G. Karopoulos, D. Geneiatakis, T. Martin, G. Kambourakis, and I. N. Fovino, “Sharing pandemic vaccination certificates through blockchain: case study and performance evaluation,” 2021, <http://arxiv.org/abs/2101.04575>.

## Research Article

# Proof of Engagement: A Flexible Blockchain Consensus Mechanism

Yuntao Xu,<sup>1</sup> Xingyu Yang,<sup>1</sup> Jiale Zhang ,<sup>1</sup> Junwu Zhu ,<sup>1</sup> Maosheng Sun,<sup>1</sup>  
and Bing Chen <sup>2</sup>

<sup>1</sup>College of Information Engineering, Yangzhou University, Yangzhou 225127, China

<sup>2</sup>College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

Correspondence should be addressed to Jiale Zhang; [jialezhang@yzu.edu.cn](mailto:jialezhang@yzu.edu.cn)

Received 2 July 2021; Accepted 10 August 2021; Published 20 August 2021

Academic Editor: Weizhi Meng

Copyright © 2021 Yuntao Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Consensus mechanism plays an important role in blockchain. At present, mainstream consensus mechanisms include proof of work (PoW), proof of stake (PoS), and delegated proof of stake (DPoS). PoW, as is widely used in virtual currency, results in significant energy consumption; PoS and DPoS are proposed to reduce energy waste caused by PoW, but their disadvantage is that they tend to create Matthew Effect (ME): “the rich get richer.” In order to balance the discourse power of new nodes and elder ones, this paper proposes a flexible consensus mechanism called proof of engagement (PoE), based on the activity and contribution of network nodes. We analyze the incentive compatibility of PoE from the perspective of mechanism design. In our simulation experiments, we tested the profit changes under PoW, PoS, and PoE. The results illustrate it is easier for new nodes to accumulate their profits under PoE than under PoW or PoS, so as to reduce the negative impacts of ME.

## 1. Introduction

Bitcoin [1], since it was proposed in 2008, has been considered the most successful application of blockchain. Its ability to work properly on a distributed system relies on the genius consensus algorithms, proof of work (PoW) [2]. Bitcoin is also commonly called Blockchain 1.0. Ethereum [3], also based on PoW, is called Blockchain 2.0 for its Turing-complete smart contract system.

Despite its widespread use, PoW still has several much-criticized problems [4, 5], and one of the most serious is its energy consumption. According to [digiconomist.net](http://digiconomist.net) [6], as of October 2020, Bitcoin has consumed electrical energy 74.38 TWh per year, which is comparable to the power consumption of Venezuela. The carbon footprint of Bitcoin has reached 35.33 Mt CO<sub>2</sub> per year, comparable to the carbon footprint of New Zealand. The reason for such a disappointing situation is that under PoW, network nodes need to run the SHA256 algorithm repeatedly until they successfully find the hash solution; then, they will be rewarded by many digital currencies; this process completely depends on the computing

power of the devices. Another problem is the centralization of computing power. To increase the chances of getting a reward, the users either buy more powerful computing devices and keep them running at full capacity or join some huge mining pools [7], which causes a shift in the computing power from decentralized back to centralized [8] and greatly threatens the security of the blockchain network.

Proof of stake (PoS) [9] was originally designed to solve the energy consumption problem. Under PoS, the probability of getting a reward is affected not only by the computing power but also by the length of time a node holding the coins (coinage). Thus, to some extent, PoE reduces energy consumption and weakens the absolute control of the full-time miners and mining pools over the blockchain network. However, this easily leads to the Matthew Effect [10, 11]: the richer always gains more profits than those who are not that rich. Unfortunately, it will be hard for new users to gain their profits under PoS, and the discourse power will gradually be centralized in the hands of a few rich ones. The centralization problem is still not effectively solved; this reduces the incentive of the whole system.

To increase the incentive of the system and reduce the negative effect of ME, in the following we propose a new consensus mechanism named proof of engagement (PoE). A blockchain system under PoE is more like a work-based society, where nodes can accumulate their profits by contributing computing power to maintain the security and creating high-quality smart contracts to maintain the autonomy. In this way, new nodes will be able to gain a voice more easily; the flexibility of the whole system is also increased.

*Our contributions.* Our contributions are summarized as follows:

- (i) We propose PoE, a flexible consensus mechanism for a smart contract system based on blockchain, which shows the ability to reduce the negative impacts of ME
- (ii) We build the static and dynamic evaluation models of smart contracts and calculate the contract quality according to the evaluation results. We calculate the activity of a node based on its transaction volume and computing power contribution during recent periods of blocks. From the perspective of mechanism design, we analyze the incentive compatibility of PoE
- (iii) We propose a method to study the flexibility of different consensus mechanisms. In our simulation experiments, we test the profit changes of those 3 nodes under PoW, PoS, and PoE. The results show that PoE is more flexible than PoW and PoS. At last, we discuss the main application directions of PoE

*Paper organization.* We organize the remainder of this paper as follows. Some related works are listed in Section 2; models and algorithms are presented in Section 3; the incentive compatibility of PoE is analyzed in Section 4; our simulation experiments and discussions are presented in Section 5; our conclusions are summarized in Section 6.

## 2. Related Work

Yuan et al. presented abstract models of PoW and PoS [12]. They modeled the rules of block production into simple inequalities.

PoW is modeled as follows:

$$F_{\text{diff}}(\text{blockheader}) \longrightarrow \text{SHA256}(\text{SHA256}(\text{blockheader})) < \frac{\text{MaxTarget}}{\text{diff}} \quad (1)$$

In this model, MaxTarget represents the maximum target value, and diff represents the degree of difficulty, which is used to control the production interval of each block.

PoS is modeled as follows:

$$\text{SHA256}(\text{SHA256}(\text{timestamp})) < \text{target} \times \text{CoinAge}. \quad (2)$$

In this model, CoinAge means the holding time of coins.

In recent years, researchers have been constantly improving or redesigning consensus mechanisms to make up for the deficiency of mainstream mechanisms and get adaption to different applications of blockchain.

To reduce the energy waste of PoW, various consensus mechanisms have been proposed. PoS [13, 14] is recognized as an excellent improvement. Besides, proof of luck (PoL) [15] and proof of elapsed time (PoET) [16], based on the trusted execution environments (TEE), are also practical alternatives. In such TEE, the node who becomes the book-keeper is decided by the waiting time generated by a random number generator, according to a presupposed probability. The introduction of TEE greatly reduces energy consumption and improves the output efficiency of blocks. Proof of useful work (PoUW) [17], proposed in 2017, gets rid of the meaningless SHA256 operation in PoW and replaces it with valuable operations in the actual scene, such as computing orthogonal vector problem, 3SUM problem, and shortest path problem. PBFT [18] is completely different from the concept of PoW: rather than choosing a winner to lead but to make sure that everyone performs the same action. It does not need to consume a lot of computing power.

Several studies are committed to solving the centralization problem brought by PoW and PoS. Based on the combination of PoW and PoS, researchers proposed proof of burn (PoB) [19], proof of activity (PoA) [20], etc. PoB enforces the miners to send their coins to a specific address that cannot be found, that is, to compete for the bookkeeping right by “burning” their coins, which alleviates the Matthew Effect to a certain extent.

## 3. Model and Process

*3.1. Generic Model.* Consensus mechanisms like PoW and PoS are usually classified as the “proof-class” mechanisms, in which once the calculated value is within the target range, and a new block is produced. PoE is also a kind of proof-class mechanism. In order to model PoE, we first present the abstract definition of a consensus mechanism for the blockchain.

*Definition 1.* A consensus mechanism  $M := \langle \mathbb{R}, \times B, f_M \rangle$  is a triple, which consists of the following:

- (i) A real number set  $\mathbb{R}$
- (ii) A blockchain  $\times B = \{B_{[0]}, B_{[1]}, \dots, B_{[k]}\}$
- (iii) A mapping  $f_M : \mathbb{R} \longrightarrow B_{[k+1]}$  of block production

Based on Definition 1, we present the generic model of the proof-class consensus mechanisms, as is shown in Figure 1. In Figure 1, the consensus mechanism extracts

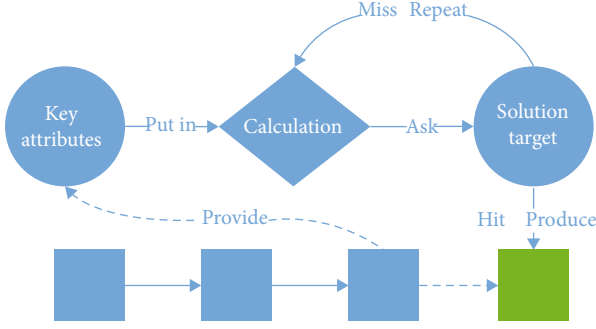


FIGURE 1: The generic model of a proof-class mechanism.

the key attributes from the existing blocks as the basis for the computation, and when the computation results in a certain target solution, a new block is produced.

Under a proof-class consensus mechanism  $M_p$ , a new block is produced if and only if  $f_{M_p}$  follows the inequality:

$$\text{Calculation}() < \text{target} \times \text{attribute}(). \quad (3)$$

$\text{Calculation}()$  represents the calculation for a solution, target represents the solution target, and  $\text{attribute}()$  represents the key attributes of network nodes for the “proof,” such as computing power in PoW and coinage in PoS; note that some of the attributes are recorded in the blockchain ledger.

*Main idea.* Our goal is to design a consensus mechanism with the following:

- (i) More flexible discourse power system
- (ii) Incentives for smart contract creators
- (iii) Preference support for different work areas

To achieve our goals, we introduce the contract quality and the node activity into PoE. Based on Definition 1, we present the generic model of PoE.

*Definition 2.*  $\text{PoE} := \langle \mathbb{R}, \times B, I, \times C, f_{\text{PoE}} \rangle$  is a quintuple, which consists of the following:

- (i) A real number set  $\mathbb{R}$
- (ii) A blockchain  $\times B = \{B_{[0]}, B_{[1]}, \dots, B_{[k]}\}$
- (iii) A set  $I = \{i_1, \dots, i_n\}$  of network nodes
- (iv) A set  $\times C = \times C_i \cup \times C_{-i}$  of smart contracts
- (v) A mapping  $f_{\text{PoE}} : \mathbb{R} \longrightarrow B_{[k+1]}$  of block production

Under PoE, a new block is produced if and only if  $f_{\text{PoE}}$  follows the inequality:

$$\text{Calculation}() < \text{target} \times \text{contractquality} \times \text{activity}. \quad (4)$$

TABLE 1: Parameters in PoE model.

Symbol	Description
$\Omega_i^a$	Node $i$ 's engagement in work area $a$
$C_i^a$	Smart contract applied in work area $a$ , created by node $i$
$Q_C$	Quality of smart contract $C$ , $Q_C \in [0, 1]$
$\text{Sta}(C)$	Static evaluation value of smart contract $C$
$\text{Dyn}(C)$	Dynamic evaluation value of smart contract $C$
$\theta$	Adjustable parameter to control the value of $Q_C$
$H_i^a$	Node $i$ 's activity in work area $a$
$k$	Block periods in a PoE-based blockchain
$T_i^a$	Node $i$ 's transaction volume in work area $a$
$\gamma_i^a$	Node $i$ 's computing power consumption in work area $a$
$P_i$	Node $i$ 's expected profits in a PoE-based blockchain

*3.2. Calculation Model.* Based on the generic model presented above, to be more specific, this section presents the calculation model of PoE. The parameters needed in our calculation model are shown in Table 1.

The engagement  $\Omega_i^a$  reflects the discourse power of node  $i$  in work area  $a$ , which is calculated by the following:

$$\Omega_i^a = \left(1 + \sum Q_{C_i^a}\right) \times H_i^a. \quad (5)$$

The quality  $Q_C$  of smart contract  $C$  is as follows:

$$Q_C = \text{Sta}(C) \times \text{Dyn}(C) \times \theta. \quad (6)$$

To show the professionalism of a node, the quantity and quality of smart contracts it creates must be considered. As we know, a new smart contract needs to be locally validated by other nodes (miners). Not only its feasibility, the security and algorithm complexity of a contract can also be evaluated during validation. We propose a model to evaluate the static properties of a smart contract.

The static evaluation  $\text{Sta}(C)$  reflects the code quality of smart contract  $C$ ; it is related to some properties of the code, such as extensibility, reusability, readability, security, and scalability, and the calculation of its value is a complex software engineering problem, which is not discussed more in this paper because of the limitation of space. However,  $\text{Sta}(C)$  only reflects the static property of a contract from the perspective of software engineering; it cannot reflect the practicality, so it is necessary to introduce the dynamic evaluation  $\text{Dyn}(C)$ .

The dynamic evaluation  $\text{Dyn}(C)$  reflects the popularity of smart contract  $C$ , which is indicated by the following:

$$\text{Dyn}(C) = \frac{\sum_{k=\bar{k}-\Delta k}^{\bar{k}} \text{CallCount}_{C[k]}}{\Delta k}. \quad (7)$$

In function (8),  $\bar{k}$  represents the current period of block, and  $\text{CallCount}_{C[k]}$  represents the number of times contract  $C$

was called during the period of  $B_{[k]}$ .  $\text{Dyn}(C)$  reflects how popular contract  $C$  was in the last  $\Delta k$  blocks.

Thanks to the Turing-complete programming language, smart contracts can perform almost any known computation, which makes them be applied to different work areas. We believe that the work area will be an important property of smart contracts in the future.

Work area  $a$  is a static property of a smart contract; it is determined by the application direction of the contract. The activity of node  $i$  in work area  $a$  is calculated by the following:

$$H_i^a = \frac{\sum_{k=\bar{k}-\Delta k}^{\bar{k}} (|T_i^a|_{[k]} + \gamma_{i[k]}^a)}{\Delta k}. \quad (8)$$

In function (9), node  $i$ 's transaction volume and consumption of computing power during the last  $\Delta k$  blocks show its recent activity.  $\gamma$  represents the computing power consumption for maintaining the blockchain, which includes contract validation and transaction creation and verification.

As we know, a node will be rewarded with a bonus once it successfully "dig out" a new block. In PoW and PoS, the production of new blocks has certain degrees of randomness, which makes the mechanism fairer. In PoE, randomness is preserved, so the expected profits of a node in a PoE-based blockchain should be calculated to evaluate its benefits.

The expected profit  $P_i$  of node  $i$  in a PoE-based blockchain is as follows:

$$P_i = \text{reward} \times \frac{\sum_{a=a_1}^{a_m} \Omega_i^a}{\sum_{x=1}^n \Omega_x}. \quad (9)$$

In function (10), a node's total engagement  $\Omega_x$  in a PoE-based blockchain is the sum of its engagement in all  $m$  different work areas.

**3.3. The Process of PoE.** The process of consensus is divided into 2 main stages according to the different types of input and output: one is the leader election among nodes, and the other is the chain update among transactions.

**3.3.1. Leader Election.** The leader election stage determines which node will become the bookkeeper.

**Definition 3.** The leader election stage consists of the following:

- (i) A set  $I = \{i_1, \dots, i_n\}$  of network nodes
- (ii) A set  $\Omega = \{\Omega_{i_1}, \dots, \Omega_{i_n}\}$  of nodes' total engagement
- (iii) A unique bookkeeper  $b$
- (iv) A encryption function  $\text{calculation}()$
- (v) A range target of solutions

```

Input:  $I, \Omega$ , target
Output:  $b$ 
1. while True do
2.   for all  $i \in I$  do
3.     if  $\text{calculation}().\text{result} < \text{target} \times \Omega_i$  then
4.        $b \leftarrow i$ 
5.       break
6.     else
7.       continue
8.   end if
9. end for
10. end while
11. return  $b$ 

```

ALGORITHM 1: Leader election.

In a PoE-based blockchain, the process of the leader election stage is shown as Algorithm 1. In Algorithm 1, the engagement is introduced as a key attribute to gain the range of the computing target, so a node with high-level engagement is more likely to be a bookkeeper.

**3.3.2. Chain Updation.** The chain updation stage determines which of the validated smart contracts will be updated onto the blockchain.

**Definition 4.** The chain updation stage consists of the following:

- (i) A set  $\times \tilde{C} = \{C_1, \dots, C_j\}$  of validated smart contracts
- (ii) A set  $\times \tilde{C} = \{C_1, \dots, C_i\}$  of smart contracts to be updated
- (iii) A set  $A = \{a_1, \dots, a_m\}$  of different work areas
- (iv) A set  $I = \{i_1, \dots, i_n\}$  of network nodes
- (v)  $n$  sets  $\times \Omega_i^a = \{\Omega_i^{a_1}, \dots, \Omega_i^{a_m}\}$  of  $n$  nodes' engagement in different work areas
- (vi) Block capacity  $V$ , the maximum number of contract-addresses stored in a block
- (vii) A function  $\text{getWorkArea}()$  for getting the work area of a smart contract

In a PoE-based blockchain, the validation process of smart contracts is shown as Algorithm 2 and the chain updation stage is shown as Algorithm 3.

In Algorithm 2,  $\hat{S}$  represents a threshold for the static evaluation of smart contracts. In other words, a smart contract is "qualified" when its static evaluation reaches  $\hat{S}$ . In Algorithm 3, network nodes vote for the validated smart contracts, and if a node agrees to contract  $C$ , he/she will use his/her engagement in the same field to endorse the contract by adding  $\Omega_i^a$  to  $\Omega_C$ . The mechanism will determine which contracts are eventually recorded on the blockchain, based on their engagement  $\Omega_C$  and the block capacity.



```

Input:  $\times C = \{C_1, \dots, C_n\}, \widehat{S}$ 
Output:  $\times \widetilde{C} = \{C_1, \dots, C_m\}$ 
1.  $\times \widetilde{C} \leftarrow \text{null}$ 
2. for  $C \in \times C$  do
3.    $S \leftarrow \text{Sta}(C)$  in function(3)
4.   if  $S \geq \widehat{S}$  then
5.     add  $C$  to  $\times \widetilde{C}$ 
6.   end if
7. end for
8. return  $\times \widetilde{C}$ 

```

ALGORITHM 2: Validation process of smart contracts.

```

Input:  $\times \widetilde{C}, A, I, \times \Omega_i^a, V$ 
Output:  $\times \widetilde{C}$ 
1.  $\times \widetilde{C} \leftarrow \text{null}$ 
2.  $L \leftarrow \text{null}$ 
3.  $v \leftarrow 0$ 
4. for  $C \in \times \widetilde{C}$  do
5.    $\Omega_C \leftarrow 0$ 
6.   for  $a \in A$  do
7.     if  $a == \text{getWorkArea}(C)$  then
8.       for  $i \in I$  do
9.         if agree then
10.           $\Omega_C \leftarrow \Omega_C + \Omega_i^a$ 
11.          add  $\Omega_C$  to  $L$ 
12.        else
13.          continue
14.        end if
15.      end for
16.    end if
17.  end for
18. end for
19. while  $v < V$  do
20.    $v \leftarrow v + 1$ 
21.   for  $\Omega_{C_i} \in L$  do
22.     if  $\Omega_{C_i} == \max(L)$  then
23.       add  $C_i$  to  $\times \widetilde{C}$ 
24.     end if
25.   end for
26. end while
27. return  $\times \widetilde{C}$ 

```

ALGORITHM 3: Chain updation of smart contracts.

#### 4. Game Analysis

In this section, the incentive compatibility of our mechanism is discussed. In a PoE-based blockchain, in order to gain more profits, a node must keep active and concentrated and try to increase its contract quality.

**Theorem 5.** Suppose there are average activity values  $\overline{H}_{i1}$ ,  $\overline{H}_{i2}$  of node  $i$ , which satisfy  $\overline{H}_{i1} > \overline{H}_{i2}$ , and all the other variables are constant, there is  $P_{i1} > P_{i2}$ .

*Proof.*  $P_i$  is estimated to be a function of the average activity  $\overline{H}_i$  of node  $i$ :

$$\begin{aligned}
 P_i &= \text{reward} \times \frac{\overline{H}_i \times \left(m + \sum_{j=1}^m Q_{C_i}\right)}{\overline{H}_i \times \left(m + \sum_{j=1}^m Q_{C_i}\right) + \overline{H}_{-i} \times \left(m + \sum_{j=1}^m Q_{C_{-i}}\right)} A \\
 &\leftarrow m + \sum_{j=1}^m Q_{C_i}, B \leftarrow \overline{H}_{-i} \times \left(m + \sum_{j=1}^m Q_{C_{-i}}\right) \\
 &\Rightarrow P_i = \text{reward} \times \frac{\overline{H}_i \times A}{\overline{H}_i \times A + B}.
 \end{aligned} \tag{10}$$

Calculate derivative of function (11) with respect to  $\overline{H}_i$ , and the outcome is as follows:

$$P_{i, \overline{H}_i} = \frac{\text{reward} \times A \times B}{\left(\overline{H}_i \times A + B\right)^2}. \tag{11}$$

Obviously, there is  $P_i, \overline{H}_i > 0$ , so function (11) is *strictly increasing*. It is proved that if a node is not active enough in a PoE-based blockchain, its expected profits will strictly reduce.  $\square$

**Theorem 6.** In the same period of block, suppose there are contract quality values  $Q_{i1}, Q_{i2}$  of node  $i$ , which satisfy  $Q_{i1} > Q_{i2}$ , and all the other variables are constant, there is  $P_{i1} > P_{i2}$ .

*Proof.* From function (11), it is obvious that  $P_i$  can also be estimated to be a function  $P_{i,A}$  of  $A$ . Calculate derivative of  $P_{i,A}$  with respect to  $A (A = m + Q)$ , and the outcome is as follows:

$$P_{i,A} = \frac{\text{reward} \times \overline{H}_i \times B}{\left(\overline{H}_i \times A + B\right)^2}. \tag{12}$$

Similarly there is  $P_{i,A} > 0$ , so the function  $P_{i,A}$  is *strictly increasing*. It is proved that if a node succeeds in increasing its contract quality in the next period of block, it will be likely to gain more profits.  $\square$

According to the theory of mechanism design [21], a direct mechanism  $(q, t)$  is incentive-compatible if and only if

- (1)  $q$  is increasing
- (2) For every  $\theta \in [\underline{\theta}, \overline{\theta}]$ , we have

$$t(\theta) = t(\underline{\theta}) + (\theta q(\theta) - \underline{\theta} q(\underline{\theta})) - \int_{\underline{\theta}}^{\theta} q(x) dx \tag{13}$$

It is given in function (9) that the computing power consumption  $\gamma$  of a node for maintaining the blockchain is positively correlated with its activity. In a PoE-based blockchain,

the computing power consumption must be considered as the “mining cost” of a node.

The utility  $u_i$  of node  $i$  is calculated by the following:

$$u_i = P_i - \Gamma_i - O(q_i)\Gamma_i = \sigma \times \gamma_i. \quad (14)$$

In function (15),  $\Gamma_i$  represents the average “mining cost” in the recent  $\Delta k$  periods of blocks.  $\sigma$  ( $\sigma > 0$ ) represents the unit cost.  $O(q_i)$  represents the cost for increasing the contract quality; it is small enough compared to  $\Gamma_i$ . Function (15) can also be estimated to be a function of  $\gamma_i$ :

$$u(\gamma_i) = P(\gamma_i) - \Gamma(\gamma_i) - O(q_i). \quad (15)$$

**Lemma 7.** *When  $\sigma$  is within a reasonable range, there exists a threshold  $\hat{\gamma}_i$  which satisfies the following:*

$$u(\hat{\gamma}_i) = \max(u(\gamma_i)). \quad (16)$$

*Proof.* Calculate derivative of function (16):

$$u'(\gamma_i) = P'(\gamma_i) - \Gamma'(\gamma_i). \quad (17)$$

Same as function (12),  $P'(\gamma_i)$  decreases strictly and approaches to 0.  $\Gamma'(\gamma_i) = \sigma$  ( $\sigma > 0$ ), when  $\sigma$  satisfies the following:

$$0 < \sigma < \frac{\text{reward} \times A}{B}. \quad (18)$$

Function (16) exhibits increasing and then decreasing. Lemma 7 stands.  $\square$

In a PoE-based blockchain, the strategy space of node  $i$  is as follows:

$$s_i = \begin{cases} \gamma_i \times q_i \gamma_i \in [0, \hat{\gamma}_i), q_i \in [0, 1), \\ q_i \gamma_i = \hat{\gamma}_i, q_i \in [0, 1), \\ \gamma_i \gamma_i \in [0, \hat{\gamma}_i), q_i = 1, \\ \hat{s}_i \gamma_i = \hat{\gamma}_i, q_i = 1. \end{cases} \quad (19)$$

In function (20),  $\gamma_i, q_i$  represents increment space of the computing power consumption and the contract quality. Now function (15) can be expressed as follows:

$$u_i = \begin{cases} u(\gamma_i, q_i) \gamma_i \in [0, \hat{\gamma}_i), q_i \in [0, 1), \\ u(q_i) \gamma_i = \hat{\gamma}_i, q_i \in [0, 1), \\ u(\gamma_i) \gamma_i \in [0, \hat{\gamma}_i), q_i = 1, \\ u(\hat{s}_i) \gamma_i = \hat{\gamma}_i, q_i = 1. \end{cases} \quad (20)$$

**Theorem 8.** *When  $s_i$  satisfies function (20) and  $\sigma$  satisfies function (17), the PoE mechanism is incentive-compatible.*

*Proof.* According to Theorem 5 and Theorem 6, it is obvious that  $P_i$  is increasing. According to function (14), function

TABLE 2: Attributes of nodes at  $B_{[0]}$  in Exp1.

Attribute	$a$	$b$	$c$	Remarks
Computing power	100	10	1	Constant
Coin age	10	100	1	Variable
Activity	10	1	10	Constant
Contract quality (total)	1	10	$q$	Variable

(15), and function (21), it is easy to infer that  $u_{s_i}$  satisfies the following:

$$u_{s_i} \geq u_{s'_i} \Leftrightarrow, \quad (21)$$

$$u_{s_i} \geq s_i \int_{s'_i}^{s_i} \frac{\delta P_i}{\delta s'_i} dx - \left( \int_{s'_i}^{s_i} \frac{\delta \Gamma_i}{\delta s'_i} dx + \int_{s'_i}^{s_i} \frac{\delta O(q_i)}{\delta s'_i} dx \right) \Leftrightarrow, \quad (22)$$

$$u_{s_i} \geq s_i \vartheta(s'_i) - s'_i \vartheta(s'_i) + s'_i \vartheta(s'_i) - \tau(s'_i) \Leftrightarrow, \quad (23)$$

$$u_{s_i} \geq s_i \vartheta(s'_i) - s'_i \vartheta(s'_i) + u(s'_i) \Leftrightarrow, \quad (24)$$

$$u_{s_i} - u(s'_i) \geq (s_i - s'_i) \vartheta(s'_i) \Leftrightarrow, \quad (25)$$

$$\int_{s'_i}^{s_i} \vartheta(x) dx \geq \int_{s'_i}^{s_i} \vartheta(s'_i) dx \Leftrightarrow. \quad (26)$$

Theorem 8 stands.  $\square$

It can be seen from the establishment of Theorem 5, Theorem 6, and Theorem 8 that as a consensus mechanism, PoE is incentive-compatible. It means that under PoE, rational nodes are more likely to choose to improve the quality of their contracts and remain active to improve their profits. This explains the feasibility of PoE.

## 5. Experiment and Discussion

In this section, a method of evaluating the flexibility of a consensus mechanism is proposed. According to the generic models of PoW and PoS mentioned in Section 3.1 and Algorithm 1, we first test and compare the performances of 3 nodes under 3 different mechanisms in Exp1. Then, we test the impact of the contract quality on nodes' performances under PoE in Exp2. At last, we discuss the main application directions of PoE.

**5.1. Flexibility Comparison.** The 3 nodes have different characteristics and strategy preferences at the very beginning, as is shown in Table 2: node  $a$  performs as a full-time miner, so he has the highest computing power and always keeps active; node  $b$  performs as a lazy rich guy, so he has a high level of coinage and creates some smart contracts for transactions; node  $c$  performs as a new and hard-working developer, so he keeps active and creates a number of smart contracts for different applications.

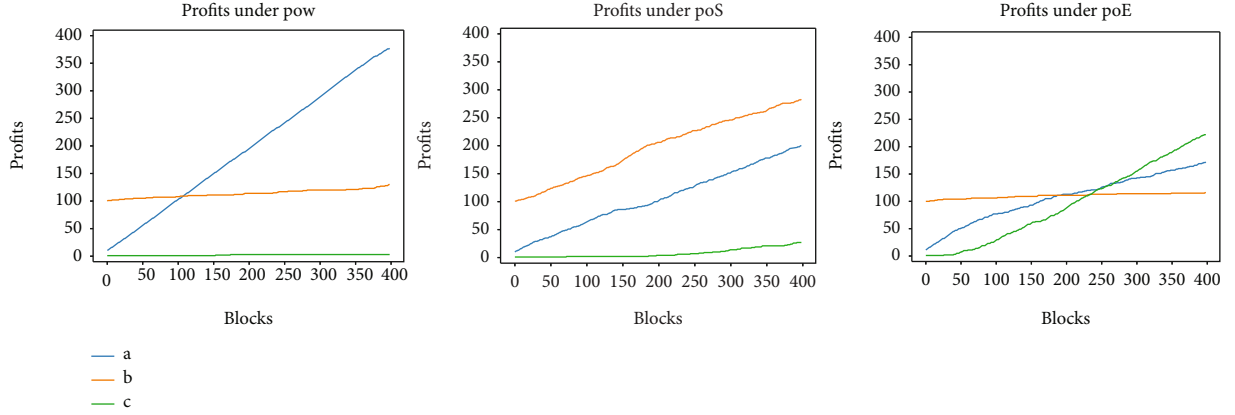


FIGURE 2: Profits under different consensus mechanisms in Exp1.

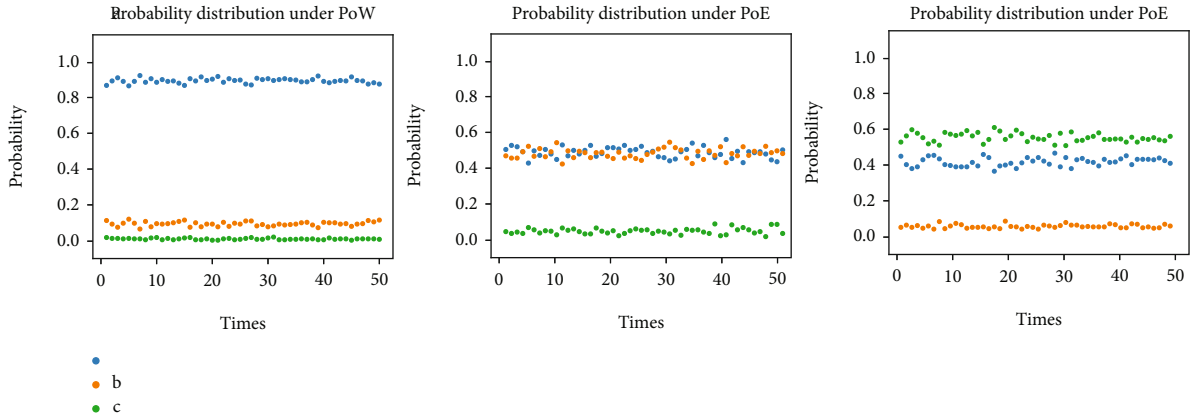


FIGURE 3: Probability distribution under different consensus mechanisms in Exp1.

5.1.1. *Profits.* The profit changes of  $a, b, c$  under PoW, PoS, and PoE during 400 periods of blocks are shown in Figure 2. It is directly shown in the line charts that compared to those under PoW or PoS, node  $a$  accumulates its profits more quickly under PoE and surpasses  $b$  and  $c$  in a short time.

5.1.2. *Probability Distributions.* We tested  $a, b, c$ 's probability distribution of becoming the bookkeeper during 400 periods of blocks for 50 times under PoW, PoS, and PoE. The probability distributions are shown in Figure 3. Figures 2 and 3 directly indicate that node  $a, b$ , and  $c$  have their own advantages under the 3 different mechanisms.

The flexibility  $\Phi_M$  of a consensus mechanism  $M$  is a property which reflects the novice-friendliness and the incentive of  $M$ .

**Proposition 9.** *The flexibility  $\Phi_M$  of a consensus mechanism  $M$  is proportional to the average probability  $\bar{\phi}_{new}$  that a new node successfully becomes the bookkeeper under  $M$  during the first  $k$  periods of blocks.*

We do not have to prove the simple proposition presented above. Note that node  $c$  is a new node in the block-

 TABLE 3: Attributes of nodes at  $B_{[0]}$  in Exp2.

Attribute	$a$	$b$	$c$	Remarks
Computing power	1	1	1	Constant
Coin age	1	1	1	Variable
Activity	10	10	10	Constant
Contract quality	$c$	$c$	$c$	Variable
Contract quality (average)	1	1	$q$	Variable

chain, so we calculate the average probability  $\bar{\phi}_c$  of node  $c$  under different mechanisms:

$$\begin{aligned}
 \bar{\phi}_c(\text{PoW}) &= 0.0097, \\
 \bar{\phi}_c(\text{PoS}) &= 0.0452, \\
 \bar{\phi}_c(\text{PoE}) &= 0.5462.
 \end{aligned} \tag{27}$$

Obviously, node  $c$  has excellent performance under PoE. Note that the cost of creating a smart contract (or improving the contract quality) is much less than increasing the computing power or increasing the coinage. Hence, it is not

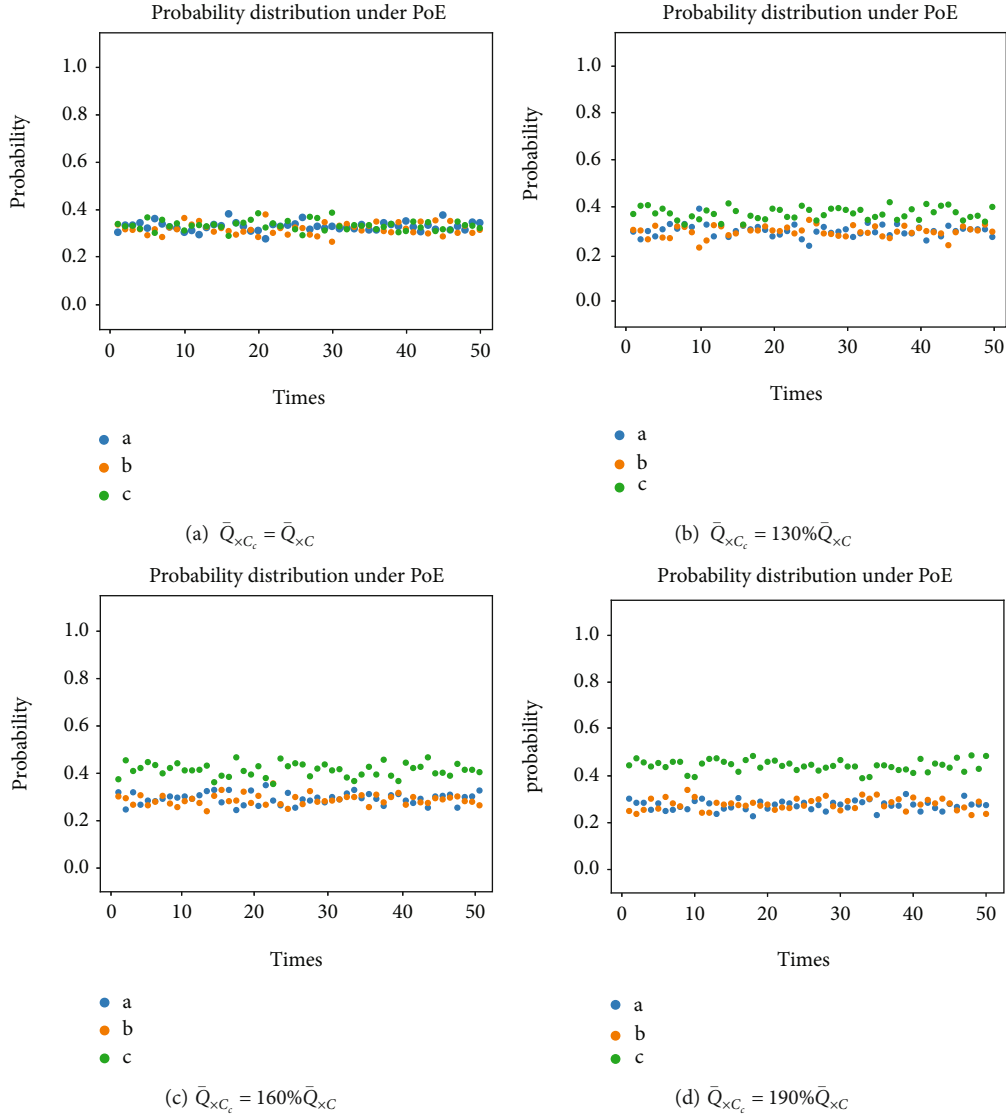


FIGURE 4: Probability distribution under PoE with contract quality increasing in Exp2.

difficult to summarize that PoE is more flexible than PoW and PoS.

**5.2. Impact of Contract Quality.** In Exp2, we assume that  $a$ ,  $b$ , and  $c$  are all hard-working developers, and their characteristics are shown in Table 3.

In Exp2, the basic rules are as follows:

- (i) At  $B_{[0]}$ ,  $P_a = 1$ ,  $P_b = 1$ ,  $P_c = 1$
- (ii)  $P_i = P_i + 1$  if  $b \leftarrow i$
- (iii) At  $B_{[0]}$ ,  $c = 1$ , and there is  $c = c + 1$  when  $B_{[k]} = B_{[k+1]}$

Different from Exp1, we considered the average quality of smart contracts in Exp2. We improved the average contract quality of node  $c$  by 0%, 30%, 60%, and 90%, respectively, and tested  $a$ ,  $b$ ,  $c$ 's probability distribution of becoming the bookkeeper under PoE, and the results are shown in Figure 4.

We calculate the average probability  $\bar{\phi}_c$  of node  $c$  when  $\bar{Q}_{x_{C_c}}$  is 0%, 30%, 60%, and 90% higher than the average value  $\bar{Q}_{x_C}$ :

$$\begin{aligned}
 \bar{\Phi}_c(\bar{Q}_{x_{C_c}} = 100\% \bar{Q}_{x_C}) &= 0.3384, \\
 \bar{\Phi}_c(\bar{Q}_{x_{C_c}} = 130\% \bar{Q}_{x_C}) &= 0.3820, \\
 \bar{\Phi}_c(\bar{Q}_{x_{C_c}} = 160\% \bar{Q}_{x_C}) &= 0.4144, \\
 \bar{\Phi}_c(\bar{Q}_{x_{C_c}} = 190\% \bar{Q}_{x_C}) &= 0.4444.
 \end{aligned} \tag{28}$$

Our experimental results directly indicate that when node  $c$ 's average contract quality improves by 30%, 60%, and 90%, and its average probability to become a bookkeeper is improved by about 12.9%, 22.5%, and 31.3%.

**5.3. Application Discussion.** According to the characteristics of PoE, we have made some assumptions about its application directions.

**5.3.1. Knowledge Payment Platforms.** Higher quality, more practical knowledge deserves a higher price.

Suppose in a Q&A community, users ask and answer questions in different areas, such as medicine, finance, and law. The community will reward users with excellent answers, and the questioner also gives the answerer some appreciation for solving his/her problems. A PoE-based blockchain is suitable for building a Q&A community, for example, a senior lawyer can quickly accumulate his fame and income by answering a number of questions related to law in such a community.

**5.3.2. Copyright Protection.** The data should be traceable and hard to be tampered with.

Thanks to the characteristics of a PoE-based blockchain, the publication time and author information of work can be traced back and it is very difficult to be tampered with. In addition, excellent works can be endorsed by experts in the industry, so that the author's hard work can be recognized by the industry.

**5.3.3. Distributed Social Networks.** High-quality content creators should be recommended to be followed by users. A more active user should have a louder voice.

Suppose on a video website, users can get coins by daily logging in and watching videos, which will be given to their favorite videos as "like." Videos with more coins will be recommended to each user's home page, and the video creators will be recommended to be followed. A PoE-based blockchain can meet the needs of such distributed social networks and greatly reduce the pressure of centralized storage. In addition, it can protect users' creations from usurpation.

## 6. Conclusion

This study mainly investigates a novel consensus mechanism called proof of engagement. Specifically, we first present a definition of consensus mechanism and propose a generic model of proof-class consensus mechanisms, and on this basis, establish the generic and calculation models of PoE. Secondly, we present the algorithms of the consensus process. Thirdly, from the perspective of mechanism design, we analyze the incentive compatibility of our mechanism. If a node keeps active and improves its contract quality, it will always gain more profits than performing idleness. At last, we test the flexibility of our mechanism through a series of simulation experiments and discuss the main application directions of PoE. The experimental results show that a new node is more easier to increase profits if he/she maintains a high level of engagement. Our mechanism has better incentives for contract creators than PoW and PoS. This is a strong indication that to a large extent, PoE weakens the Matthew Effect. Generally speaking, PoE is a flexible, fair, and novice-friendly mechanism. This work provides a new idea for the industrial combination of blockchain; it will bring benefits to the development

of smart contracts and distributed autonomous systems such as DApp, DAO, and DAS.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61872313, in part by the Key Research Projects in Education Informatization in Jiangsu Province under Grant 20180012, in part by the Yangzhou Science and Technology under Grant YZ2020174 and Grant YZ2019133, and in part by the Open Project in the State Key Laboratory of Ocean Engineering, Shanghai Jiao Tong University under Grant 1907.

## References

- [1] S. Nakamoto, "Bitcoin: a peer-to-peer electronic cash system," *Decentralized Business Review*, 2008, article 21260.
- [2] 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*, New York, NY, USA, 2016.
- [3] G. Wood, "Ethereum: a secure decentralised generalised transaction ledger," *Ethereum project yellow paper*, vol. 151, pp. 1–32, 2014.
- [4] J. Golosova and A. Romanovs, "The advantages and disadvantages of the blockchain technology," in *2018 IEEE 6th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE)*, Vilnius, Lithuania, 2018.
- [5] I. Bentov, A. Gabizon, and A. Mizrahi, "Cryptocurrencies without proof of work," in *International Conference on Financial Cryptography and Data Security*, Berlin, Heidelberg, 2016.
- [6] A. Vries, "Renewable energy will not solve bitcoin's sustainability problem," *Joule*, vol. 3, no. 4, pp. 893–898, 2019.
- [7] Y. Lewenberg, Y. Bachrach, Y. Sompolinsky, A. Zohar, and J. S. Rosenschein, "Bitcoin mining pools: a cooperative game theoretic analysis," in *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, Istanbul, Turkey, 2015.
- [8] L. Cong, Z. He, and J. Li, "Decentralized mining in centralized pools," *The Review of Financial Studies*, vol. 34, no. 3, pp. 1191–1235, 2021.
- [9] S. King and N. Scott, *Ppcoin: peer-to-peer crypto-currency with proof-of-stake*. Self-published paper, 2012.
- [10] R. Merton, "The Matthew effect in science: the reward and communication systems of science are considered," *Science*, vol. 159, no. 3810, pp. 56–63, 1968.
- [11] Y. Liu, K. Wang, Y. Lin, and W. Xu, "LightChain: a lightweight blockchain system for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3571–3581, 2019.

- [12] Y. Yuan, X. Ni, S. Zeng, and F. Wang, "Blockchain consensus algorithms: the state of the art and future trends," *Acta Automatica Sinica*, vol. 44, no. 11, pp. 2011–2022, 2018.
- [13] F. Saleh, "Blockchain without waste: proof-of-stake," *The Review of Financial Studies*, vol. 34, no. 3, pp. 1156–1190, 2021.
- [14] L. Chen, L. Xu, N. Shah, Z. Gao, Y. Lu, and W. Shi, "On security analysis of proof-of-elapsed-time (poet)," in *International Symposium on Stabilization, Safety, and Security of Distributed Systems*, Cham, Boston, MA, USA, 2017.
- [15] M. Milutinovic, W. He, H. Wu, and M. Kanwal, "Proof of luck: an efficient blockchain consensus protocol," in *Proceedings of the 1st Workshop on System Software for Trusted Execution*, Trento, Italy, 2016.
- [16] A. Baldominos and S. Yago, "Coin. AI: a proof-of-useful-work scheme for blockchain-based distributed deep learning," *Entropy*, vol. 21, no. 8, p. 723, 2019.
- [17] A. Shoker, "Sustainable blockchain through proof of exercise," in *2017 IEEE 16th International Symposium on Network Computing and Applications (NCA)*, Cambridge, MA, USA, 2017.
- [18] W. Chiu and W. Meng, "EdgeTC - a PBFT blockchain-based ETC scheme for smart cities," in *Peer-to-Peer Networking and Applications*, pp. 1–13, Springer, 2021.
- [19] K. Karantias, A. Kiayias, and D. Zindros, "Proof-of-burn," in *International Conference on Financial Cryptography and Data Security*, Cham, Kota, Kinabalu, Malaysia, 2020.
- [20] I. Bentov, C. Lee, A. Mizrahi, and M. Rosenfeld, "Proof of activity: extending bitcoin's proof of work via proof of stake," *ACM SIGMETRICS Performance Evaluation Review*, vol. 42, no. 3, pp. 34–37, 2014.
- [21] T. Børgers and J. Li, "Strategically simple mechanisms," *Econometrica*, vol. 87, no. 6, pp. 2003–2035, 2019.

## Research Article

# A Novel User Collusion-Resistant Decentralized Multi-Authority Attribute-Based Encryption Scheme Using the Deposit on a Blockchain

Siwan Noh <sup>1</sup>, Donghyun Kim,<sup>2</sup> Zhipeng Cai <sup>2</sup>, and Kyung-Hyune Rhee <sup>3</sup>

<sup>1</sup>Department of Information Security, Graduate School, Pukyong National University, Busan 48513, Republic of Korea

<sup>2</sup>Department of Computer Science, Georgia State University, Atlanta, GA 30303, USA

<sup>3</sup>Department of IT Convergence and Application, Pukyong National University, Busan 48513, Republic of Korea

Correspondence should be addressed to Kyung-Hyune Rhee; khrhee@pknu.ac.kr

Received 23 April 2021; Revised 23 May 2021; Accepted 18 June 2021; Published 7 July 2021

Academic Editor: Weizhi Meng

Copyright © 2021 Siwan Noh et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recently, the concept of a decentralized data marketplace is getting much attention to exchange user data. Multi-authority attribute-based encryption (ABE), which can provide flexibility and user-centric access control, is previously widely used in decentralized data sharing applications and also becoming a foundation to build decentralized data trading applications. It is known that users in a multi-authority ABE system can collude by sharing their secret information for malicious purposes. To address this issue, the collusion-resistant multi-authority ABE model was introduced in which a unique global identifier (GID) is issued by the central authority (CA) to each user. Unfortunately, such approach cannot be used directly to build a decentralized data marketplace as (a) such intervention of the CA is directly against the main motivation of the decentralized trading platform and, mostly importantly, (b) the CA can exploit its full knowledge on users' GID to launch various attacks against users. Motivated by these observations, this paper introduces a novel user collusion-resistant decentralized multi-authority ABE scheme for privacy preserving data trading systems. In the existing multi-authority ABE systems, users utilize his/her GID that is solely assigned by the CA to generate his/her secret keys throughout the collaboration with authorities and a user can compute multi-authority keys by combining the secret keys (stem from the same GID) in various ways. In the proposed system, the CA only has a partial knowledge of users' GIDs, and thus, users' privacy can be protected. On the other hand, we set the user's own partial GID as a secret which can be used to withdraw his/her deposit to discourage any possible collusion among users.

## 1. Introduction

The development of the Internet of Things (IoT) has led people to generate larger amounts of data in their daily lives. Experts predict that the amount of these data will explosively increase each year [1]. One major use of the massive amount of data is the training of various machine learning algorithms to build artificial intelligence-empowered applications for our daily lives [2–4]. As acquiring sufficient amounts of data from individuals to train the machine learning algorithms has many limitations, the concept of a data marketplace is introduced; a data marketplace is an online trading platform where people can trade data, and it can be considered suitable

for legally acquiring the data required for learning. Google [5] and Amazon [6] provide an online data marketplace service based on their cloud infrastructure. On these platforms, people can buy qualified data and analyze it on the cloud computing service. However, there is a significant issue that the most centralized marketplace does not support the user-to-user data trade and all rights to sales data are controlled by the central administrator. A decentralized marketplace platform [7–9] where the concept has recently been proposed can provide decentralized data trading among users without a centralized administrator based on the blockchain smart contract. A decentralized data trading platform achieves decentralization by excluding the participation of

trusted intermediaries. However, the decentralized platform does not support data sharing through trusted administrator among the data seller and multiple buyers. Therefore, considering the problem that sellers must always be online for data sharing, the most efficient way is for the seller to outsource sales data to the cloud server and the server to provide data only to users authorized by the seller. However, traditional one-to-one cryptographic schemes (e.g., symmetric key encryption and public key encryption) are not suitable for decentralized data trading. If a one-to-one scheme is used, the seller will need to transform (or re-encrypt) data stored on the cloud server into new ciphertexts for the buyer every time.

An attribute-based encryption (ABE) is a cryptographic scheme that can provide one-to-many encryptions, which satisfies the utmost requirements of data trading. The seller can specify the buyer's job or position in the process of generating a ciphertext, and buyers can also efficiently find the desired data based on specified attributes in the ciphertext. In the ciphertext-policy attribute-based encryption (CP-ABE) [10–13], the message sender defines the attributes required for the decrypting of the message as an access tree. The ciphertext can be decrypted by any user who has appropriate attributes. To decrypt the message, firstly, the recipient must prove his/her attribute set and gets corresponding secret keys from the key authority. However, in the multi-authority environment [14–17], authorities manage attributes of users belonging to them privately unless there is a predefined communication channel among them and users can belong to more than one authority with multiple secret keys. The message sender can define a set of attributes issued by different authorities as conditions for decrypting data. Therefore, in the multi-authority ABE system, recipients must be able to generate a secret key by combining his/her multiple attributes generated by multiple authorities.

However, if the combination of secret keys is allowed, the system must consider the following two issues: first, authorities do not know a secret key generated by other key authorities. If the key authority receives a request to combine secret keys from the user, the authority must be able to verify that the requested secret keys are all held by the same user. In other words, all authorities must be able to distinguish between the honest user's request and the malicious attacker's request. The second issue is a revocation of the secret key. If the user leaves the system or an attribute is updated, the corresponding user's secret key should no longer be used. If the system does not provide the key revocation, attackers may be able to attempt a collusion attack using a revoked user's secret key.

Chase [14] proposed a global identifier (GID) to distinguish between the honest request and collusion attack. The user uses his/her unique GID issued by the central authority as an input parameter for the key generation algorithm, and only combinations of secret keys generated from the same GID are allowed in the system. Subsequent works [15, 17–20] have improved Chase's approach, where instead of identifiers, the CA chooses a random secret value for each user and generates the secret key based on this random value with each authority. This random value is revealed only when the

user has enough secret keys. If the user attempts to combine the secret keys generated from different random values, it is not revealed and the algorithm does not return a valid computational result. However, these approaches assumed a trusted CA that issues a unique identifier for each user and there was a limit to the system being overly dependent on CA. A decentralized attribute-based encryption (DABE) [10, 18–20] was proposed to solve the concentration of secrets in the CA during the key generation process, but there remained a problem of the centralized GID. Even if the power of the CA has been weakened, the user's GID is still determined by CA, so the CA can collude with the user to violate the privacy of other users [21].

## 2. Related Work

Sahai and Waters [22] first proposed an ABE as an extension concept of identity-based encryption [23]. In ABE cryptosystems, a user is represented by a set of attributes instead of unique identities. Therefore, ABE can provide flexible access control based on the user's attributes. In [22], the ciphertext can be decrypted if the recipient has at least  $d$  attributes in the entire set of attributes. To generate the secret key, first, the user proves his/her identity with attributes to the trust authority. After validation of attributes is completed, the key authority uses its master key to generate the requester's secret key corresponding to the set of attributes that the requester has.

Many researchers have recently used ABE to achieve decentralized data sharing. Gao et al. [24] and Zhang et al. [25] proposed a decentralized data sharing system. They propose a method that combines the blockchain and CP-ABE to resolve problems caused by the untrusted cloud service provider in traditional data sharing systems. Instead of storing personal data on the untrusted cloud server, Gao et al. [24] suggested that InterPlanetary File System (IPFS) nodes store a chunk of encrypted data. In his data sharing scheme, only users with appropriate attributes can collect chunks from the decentralized file system and reconstruct the original decrypted data. Zhang et al. [25] used an attribute-based signature (ABS), CP-ABE, and the blockchain smart contract to share IoT sensor data on the untrusted cloud server. In [25], the data owner stores encrypted data on the cloud server and generates a smart contract that manages the access control table of the data. If the IoT device that wants access to data has appropriate attributes, it can submit ABS to the smart contract and receive the encrypted private key. Subsequently, the IoT device submits ABS to the cloud server and obtains encrypted data via a secure channel established through mutual authentication. However, [22, 24, 25] assumed a single attribute key authority in their system, so it failed to present available ABE in the multi-authority model in which several different authorities operate simultaneously. A single-authority ABE model violates our goal of mitigating the dependence of the centralized authority. Therefore, we require consideration of a multi-authority environment in which multiple authorities manage users' attributes individually.



Chase [14] proposed a multi-authority ABE model based on the global identifier. In Chase's ABE model, all users in the system have a unique GID issued by the central authority (CA) and they use it as a secret seed to issue secret keys from each attribute authority. In the process of the key generation, the requester's GID and his/her attributes and a pseudorandom function (PRF) that each key authority has uniquely are used. Upon receiving a request to combine several secret keys from the user, the authority reconstructs this identifier from the requested secret keys. If all secret keys are generated from the same GID, the authority combines the requested secret keys. Since the reconstruction of GID requires information about PRF that each authority uses to generate secret keys, Chase's model necessitates a fully trusted entity that maintains the state of all authorities in the system. However, this entity will have a significant impact on the entire system if it is compromised (known as a single point of failure).

DABE can mitigate the impact of the aforementioned issues by dividing the role of the fully trusted CA in traditional ABE systems into multiple entities in the system [10, 18–20, 26, 27]. Hur and Kang [19] proposed a DABE model that improves the model of Bethencourt et al. [12]. In [19], the CA and a group of attribute authorities  $A_i$  cooperatively generate user's secret keys. Each authority generates only a part of the user's secret key in the key generation phase, so, nobody knows the entire user secret key except the user. The CA securely generates a portion of the user's secret key while maintaining privacy for their input data by running a two-party computation protocol with key authorities in the system. The user completes his/her secret keys by combining portions of the secret key issued by the CA and key authorities. Wang et al. [20] proposed the DABE model in which a key authority (KA) and a cloud service provider (CSP) cooperatively issue user's secret keys in the single-authority environment based on Water's model [13]. Wang mitigated the key escrow problem due to the key generation of the single authority by splitting the key generation operations separately between KA and CSP, similar to [19]. Similar to the scheme in [19], Wang's scheme runs a two-party computation protocol between KA and CSP, so only the user knows his/her secret key. Lin et al. [10] proposed a collaborative key management protocol based on Water's model [13]. In [10], the key authority and cloud server issue a user's secret key and a decryption server helps the user's decryption process. In Lin's scheme, the key authority (KA), cloud server (CS), and decryption server (DS) generate secret keys for each user during the key generation process. Unlike previous studies, in Lin's scheme, no one in the system has a user's secret key. Instead of issuing a secret key to the user, encryption and decryption operations based on the user's secret key issued by the CA are divided by KA, CS, and DS. In [27], Rahulamathavan et al. proposed a strong privacy-preserving DABE model with an anonymous key-issuing protocol to prevent key authorities from tracing users' GIDs and violating users' privacy. In Rahulamathavan's scheme, users can get a secret key from the key authority without revealing their GID to key authorities. However, most of the proposed schemes still rely on the GID generated by the CA to prevent collusion attacks among users. The security

of the above systems has the disadvantage of relying entirely on a single fully trusted authority. Therefore, rather than simply distributing the running of traditional ABE algorithms across multiple entities, we need a method to efficiently prevent collusion attacks.

Blockchain is a variant of the distributed database. Users in the blockchain network (called a full node) manage a local copy of the blockchain ledger themselves without the database manager. In the blockchain network, a consensus protocol is used to synchronize ledgers even if malicious nodes are participating in the network. The blockchain can be divided into two types depending on the network model [28]. The first type is a public blockchain, where all nodes in the network are untrustworthy nodes. Therefore, the public blockchain uses consensus algorithms such as a Proof-of-Work (PoW), Proof-of-Stake (PoS), and Delegated Proof-of-Work (DPoS), which can provide strong network security. A strong consensus algorithm can protect a blockchain ledger from malicious nodes in the network; however, it leads to trade-offs between security and performance [29, 30]. The second type, a private blockchain, improves the performance of the blockchain network by constructing nodes in the network only as authorized users. The private blockchain network is managed by a network administrator, and only users authorized by the administrator can participate in the network. Therefore, the private blockchain can improve the network performance based on efficient consensus algorithms such as PBFT [31].

A security deposit means money held as an initial payment of the purchase process. It is used in various fields, such as leasing, and is used as a means of ensuring fairness in the contract. For example, if a lessee damages an apartment or cancels a lease contract during an apartment lease, the lessor will deduct the amount from the lessor's deposit. Poon and Dryja [32] proposed a method to share cryptographic proofs that allow the counterparty to withdraw their deposit whenever the channel is updated. If a malicious user attempts to close the channel abnormally, the counterparty will be able to withdraw the malicious user's deposit without the original owner's consent. McCorry et al. [33] proposed a monitoring solution that improves the problem of channel participants being always online and monitoring the network. Participants in the channel always have to monitor the blockchain network to know if the counterparty closes the channel abnormally. McCorry et al. have delegated these monitoring roles to third parties and proposed a monitoring solution that allows users to validate the results. A Hashed Timelock Contract (HTLC) [34] is a method to support a cross-chain transaction between heterogeneous blockchain networks (e.g., exchange Bitcoin and Ethereum). Unlike traditional cryptocurrency transactions that use the digital signature as a proof of ownership, HTLC uses knowledge of preimage  $r$  of hash values  $H(r)$  recorded in the blockchain ledger as proof of ownership. A preimage for the proof of ownership is automatically disclosed at the phase of consuming the counterparty's transaction after both sides send the transaction to the counterparty. As such, the security deposit is used as a motivation to induce honest behavior among untrusted users.

**2.1. Our Contribution.** In this paper, we propose a fair data trading system on the multi-authority ABE model. In the proposed system, we adopt a blockchain-based security deposit to prevent collusion attacks and a data trading protocol based on HTLC to guarantee fairness between the buyer and seller. To summarize, our contributions are listed as follows:

- (i) In the proposed model, any user can attempt a collusion attack. However, if anyone tries to attack, he/she will lost his/her security deposit. People will act as honestly as possible to keep their deposits.
- (ii) In the proposed system, the trading and sharing of sales data are done without the participation of a trusted administrator. We propose a decentralized data trading protocol that can guarantee reliability between untrusted sellers and buyers.
- (iii) The proposed system is controlled only by smart contracts operating on the blockchain. Nobody can control the blockchain network that is controlled only by nodes in the network, and the proposed system is secure unless a fatal attack on the blockchain network is known.

### 3. System Architecture

In this section, we describe our proposed system architecture with our security considerations.

**3.1. System Description.** There are five system entities in our system, the central authority (CA), attribute authorities ( $A_i$ ), the user (buyer and seller), the blockchain network, and the cloud service provider (CSP) as shown in Figure 1.

- (i) *Central authority (CA)* generates a secret key to the user in cooperation with  $A_i$  that can be used to decrypt the ciphertext in the system. The CA periodically updates all secret keys to revoke a malicious user or attributes. Moreover, the CA deploys a security deposit contract (SC) for each user that motivates users to act honestly in the system
- (ii) *Attribute authorities ( $A_i$ )* are responsible for verifying the user's possession of attributes and issuing attributes to the user. In our system, a secret key means a secret value corresponding to a set of attributes and the user can obtain multiple attributes and its secret keys from multiple authorities
- (iii) *User* is classified as a seller or buyer according to their role in the trading protocol. All users must set up a security deposit as a constraint on their behavior and then participate in the system
- (iv) *Blockchain* is a decentralized P2P network without a network administrator. A smart contract is a program that runs on the blockchain network, which is controlled only by network nodes, making control by third parties practically difficult. Therefore, it is

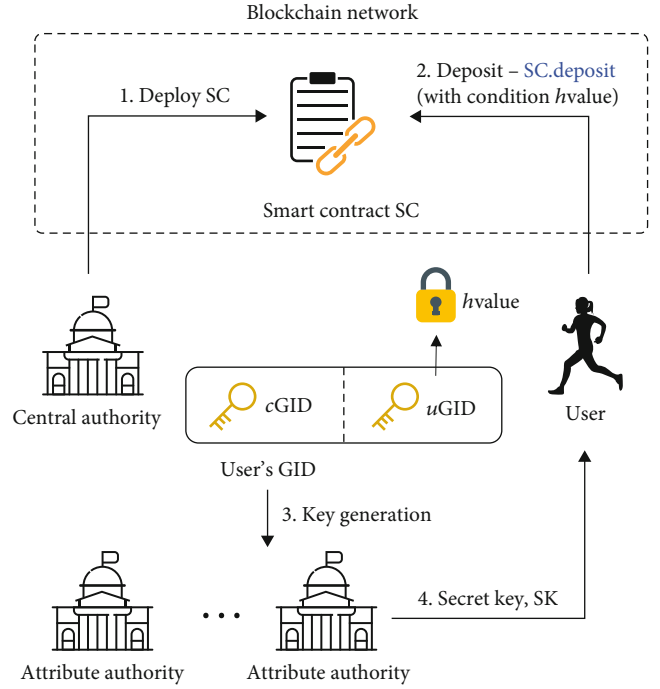


FIGURE 1: Proposed system overview.

practically difficult to control the operation of the program by malicious third parties

- (v) *Cloud service provider (CSP)* is a cloud server that stores ciphertexts, which in our system CSP serves to provide ciphertexts to buyers and deploys a trading support contract (TSC) that supports users' data trading. Furthermore, we assume that in the proposed system, CSP is a semitrusted entity and honest but curious

**3.2. Security Goals.** In this paper, we assume a collusion attack as a major threat to the proposed system. In the CP-ABE model, a user's secret key represents a set of attributes authenticated by the key authority. Even if users share their secret keys with the collusion attacker, there is no damage to colluders unless the system administrator detects and punishes the colluder who shared the secret key. The attacker may be able to get secret keys through a trusted colluder who wants to access the same data. However, realistically, it would be more desirable to assume that a third-party user who has no contact with the attacker has this secret key. Therefore, we assume that a group of collusion attackers in the proposed system does not have enough attributes to decrypt the target data (also, the attacker's group consists of reasonable users and has no trust among them). We also assume that an attacker pays a reward and gets a secret key to a third-party user who has insufficient attributes.

A collusion attack can be classified into several types of attacks according to the role of participants in the system [21], and we consider that the following two types of attacks are a threat to the proposed system.

- (i) *Among users*: an attacker's secret key (i.e., a set of attributes) cannot decrypt the target data that he wants to access. The attacker can obtain insufficient attribute keys from colluders to achieve the condition to decrypt the target data
- (ii) *Authority and user*: the CA does not issue secret keys to users. Instead, it can be the target of many attacks as an administrator of the system. A compromised CA can generate a secret key for any user without attribute authorities based on the information that can be obtained from the secret key of the user in the system

Under the threat model noted above, we consider the following security goals for a fair data trading system against collusion attacks on the multi-authority ABE model:

- (i) *Collusion resistance*: although collusion attacks can occur among entities in the system, resistance to collusion attacks should be guaranteed
- (ii) *User revocation*: the secret key of a user who has left the system or been revoked by the CA must no longer be valid on the system. Furthermore, the revoked secret key should not be used in the system even though it is still valid
- (iii) *Decentralization*: all data trading transactions occurring on the proposed system are made without the participation of the trusted intermediary. It is a trade between untrusted users, but no user should be able to harm the other party by malicious behavior

**3.3. Cryptocurrency Deposit.** The proposed model uses cryptocurrency as a deposit. However, cryptocurrency may not be appropriate to use as a deposit due to its unique floating exchange rate. For example, on January 31, 2020, Ethereum was priced at \$183.68 per dollar. However, a year later, the price increased sevenfold to \$1,317.58 on January 31, 2021. If the value of the cryptocurrency decreases, the deposit will not prevent collusion attacks, and if the value increases, it will be a factor that makes it difficult for new users to participate in the system. Therefore, cryptocurrency-based deposits can have a significant impact on the reliability of the system. To prevent issues caused by the volatility of cryptocurrencies, we assume the use of Tether (USDT) instead of ordinary cryptocurrencies such as Bitcoin (BTC) and Ethereum (ETH) as deposits used in the system. As of February 22, 2021, Tether's volatility over the past 30 days was 0.01, while Bitcoin was 0.88 and Ethereum was 0.94 [35]. Tether tokens exist as digital tokens built on Bitcoin (Omni and Liquid Protocol), Ethereum, EOS, Tron, Algorand, SLP, and OMG blockchains. Tether is implemented in the Ethereum blockchain as an ERC-20 token and supports smart contracts. Furthermore, the low variability of the tether is suitable for use in the proposed system.

## 4. Decentralized CP-ABE

A common framework of the CP-ABE scheme consists of the following four algorithms:

- (i) *Setup*: the authority runs a setup algorithm to generate public parameters for the system and then generates its master key (MK) and public key (PK)
- (ii) *Key generation*: The authority validates attributes that the user has and then issues the user's attribute key using its master key
- (iii) *Encryption*: the user (message sender) defines the access policy to decrypt the ciphertext. Then, the user uses the access policy and authority's public key to generate the ciphertext
- (iv) *Decryption*: the user (recipient) uses his/her attribute key and the authority's public keys to decrypt the ciphertext. If the user has enough attributes defined in the access policy of the ciphertext, he can acquire the plaintext, but otherwise, he will fail to decrypt it.

Hur and Kang [19] proposed a decentralized attribute-based encryption model that improves the CP-ABE model proposed by Bethencourt et al. [12]. He improved the key generation algorithm in the Bethencourt model to collaborate with the central authority and attribute authorities. This model is as follows.

### 4.1. Setup

- (i) *Global setup*: the trusted initializer chooses a bilinear group  $\mathbb{G}_0$  of prime order  $p$  with generator  $g$  and a cryptographic hash function  $H : \{0, 1\}^* \rightarrow \mathbb{G}_0$
- (ii) *Authority setup*: a central key authority (CA) chooses a random exponent  $\beta \in \mathbb{Z}_p$  as its master key and computes its public key  $g^\beta$ . CA's master private and public key pair is given by  $(MK_{CA} = \beta, PK_{CA} = g^\beta)$ . Each local key authority ( $A_i$ ) chooses a random exponent  $\alpha_i \in \mathbb{Z}_p$  as its master private key and computes its public key  $e(g, g)^{\alpha_i}$  (where  $e$  is a bilinear map and is denoted by  $e : \mathbb{G}_0 \times \mathbb{G}_0 \rightarrow \mathbb{G}_1$ ). Key authorities' master private and public key pairs are given by  $(MK_{A_i} = \alpha_i, PK_{A_i} = e(g, g)^{\alpha_i})$ . Then, it publishes a public parameter  $\text{param} = \{\mathbb{G}_0, g, H\}$

**4.2. Key Generation.** In Bethencourt's model, the CA generates all parts of the user's secret key by itself. However, in Hur and Kang's model, the CA and key authorities generate users' secret keys cooperatively shown in Figure 2. Therefore, none of them can acquire the entire part of the user's secret key. The key generation protocol is as follows:

- (1) The user  $u$  requests the CA, to generate a secret key
- (2) The CA chooses random exponents  $\gamma_i \in \mathbb{Z}_p^*$  for each key authority  $A_i$  and sets  $\text{GID}_u = \sum_{i=1}^m \gamma_i$ . Then, the CA and each key authority  $A_i$  run a secure two-party computation (2PC), where the private input of the CA is  $(\gamma_i, \beta)$ , authority  $A_i$ 's private input is  $\alpha_i$ , and the protocol returns the private output  $x = (\alpha_i + \gamma_i)\beta$  to  $A_i$

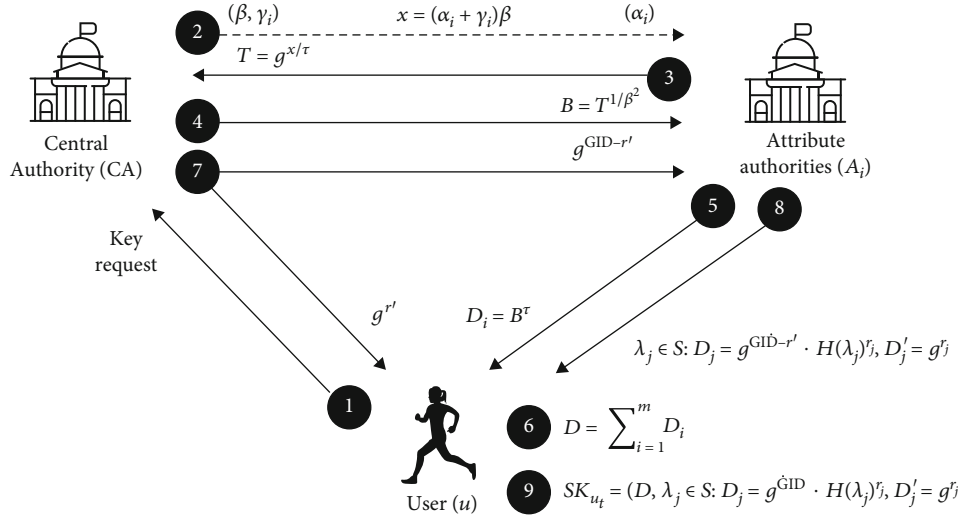


FIGURE 2: Key generation phase in the Hur and Kang model [19].

- (3)  $A_i$  randomly chooses  $\tau \in \mathbb{Z}_p^*$  and computes  $T = g^{x\tau} = g^{(\alpha_i + \gamma_i)\beta\tau}$  and then sends it to the CA
- (4) CA computes  $B = T^{1/\beta^2} = g^{\alpha_i + \gamma_i\tau\beta}$  and sends it to  $A_i$
- (5)  $A_i$  computes  $D_i = B^\tau = g^{(\alpha_i + \gamma_i)\beta}$  and sends it to the user  $u$
- (6) After receiving  $D_i$  from all key authorities, the user  $u$  computes the part of his/her secret key  $D = \prod_{i=1}^m D_i = g^{(\alpha_1 + \dots + \alpha_m) + \text{GID}_u \beta}$
- (7) The CA randomly chooses  $r' \in \mathbb{Z}_p$  and then sends  $g^{\text{GID}_u - r'}$  to  $A_i$  and  $g^{r'}$  to  $u$ .
- (8)  $A_i$  randomly chooses  $r_j \in \mathbb{Z}_p$  and sends the following secret value to the user  $u$  (value  $r_j$  related to the set of attributes  $\lambda_j$  issued to the user  $u$  by authority  $A_i$ ):

$$\forall \lambda_j \in S : D_j = g^{\text{GID}_u - r'} \cdot H(\lambda_j)^{r_j}, \quad D'_j = g^{r_j} \quad (1)$$

- (9) The user  $u$  computes  $g^{r'} \cdot D_j$  for all attribute keys that he has. The secret key that the user  $u$  obtains is as follows: (where  $D = g^{(\alpha_1 + \dots + \alpha_m) + \text{GID}_u \beta}$ )

$$\text{SK}_{u_t} = (D, \forall \lambda_j \in S : D_j = g^{\text{GID}_u} \cdot H(\lambda_j)^{r_j}, D'_j = g^{r_j}) \quad (2)$$

In the Hur and Kang model, a decentralized key generation is possible because each authority generates only a part of the user's secret key. Moreover, each user's secret key uses a different secret value  $\text{GID}_u$  randomly chosen by the CA, which prevents the collusion attack among users. Without

knowledge of this secret value  $\text{GID}_u$ , collusion attack among users is impossible.

However, if the CA is compromised, it is possible to generate a secret key for another user by colluding with the user without the participation of key authorities [21]. The user  $u$  gives the CA the value  $D$  that is part of his/her secret key  $S$   $K_u$ . The CA can compute the value  $g^\alpha$  required to generate another user's secret key from the value  $D$  received from the user  $u$  as follows:

$$\begin{aligned} \frac{D^\beta}{g^{\text{GID}_u}} &= \left( g^{(\alpha_1 + \dots + \alpha_m) + \text{GID}_u \beta} \right)^\beta \times g^{-\text{GID}_u} \\ &= g^{(\alpha_1 + \dots + \alpha_m) + \text{GID}_u + (-\text{GID}_u)} = g^{(\alpha_1 + \dots + \alpha_m)}. \end{aligned} \quad (3)$$

A simple solution for this vulnerability is to set the value  $\text{GID}_u$ , which the CA randomly determines for each user, to secret information that the CA does not know. However, if users choose the value  $\text{GID}_u$  themselves, it leads to a problem which is a collusion attack among users. And if the user gives a value  $\text{GID}_u$  as well as a value  $D$  to the CA, the solution no longer guarantees resistance to collusion attacks.

## 5. Proposed Scheme

In this section, we present a decentralized ABE model that improves the Hur and Kang model [19]. To solve the aforementioned problem, we propose a security deposit protocol to avoid collusion attacks. We use a security deposit on the blockchain as a precaution against malicious behavior by users in the attribute-based encryption system. In the proposed model, the user sets up a security deposit to the blockchain smart contract as a guarantee of his/her honest behavior to participate in the system. If the user colludes with another user, during the exchange of information for the attack, he exposes the secret value that is required to withdraw his/her security deposit from the smart contract. Table 1 shows notations used in our scheme.

TABLE 1: Notations and descriptions.

Notation	Description
$cGID_{u,t}$	CA-generated global identifier for user $u$ at period $t$
$uGID_{u,t}$	User-generated global identifier at period $t$
$\mathcal{D}_u$	Deposit from user $u$
$k_{u,t}$	Random value for proof of ownership of the deposit $\mathcal{D}_u$
$H$	Cryptographic hash function

**5.1. Deposit Setup.** All users must set up their security deposit to participate in the system. The security deposit is managed by a blockchain smart contract, and if a user shares his/her secret key to another user or authority for the collusion attack, the colluder who receives the secret key can withdraw the security deposit set by the original owner of the shared secret key. We allow users to choose a part of the global identifier  $uGID_u$ , which was previously determined uniquely by the CA for each user. And knowledge of this  $uGID_u$  is used as a condition for withdrawing the user's deposit. The user requests the CA to generate his/her secret key and then runs a deposit function shown in Figure 3.

CA deploys a deposit smart contract SC to the blockchain network for user  $u$  who requests key generation. At this time, the state of the smart contract is initialized (init). Then, the user  $u$  chooses random exponents  $\delta_i, k_{u,t_1} \in \mathbb{Z}_p^*$  for each key authority  $A_i$  and for proof of ownership (where  $uGID_{u,t_1} = \sum_{i=1}^m \delta_i$ ). The user  $u$  computes and sends hash results  $hvalue = H(uGID_{u,t_1})$  and  $hindex = H(k_{u,t_1})$  to the smart contract SC at the period  $t_1$  (where  $H$  is a cryptographic hash function). After a simple verification process, the state of SC transits from the init to the active.

Then, the user  $u$ , CA, and each key authority  $A_i$  run a key generation algorithm as shown in Figure 4.

- (1) The user  $u$  requests the CA, to generate a secret key
- (2) The CA chooses random exponents  $\gamma_i \in \mathbb{Z}_p^*$  for each key authority  $A_i$  and sets  $cGID_{u,t_1} = \sum_{i=1}^m \gamma_i$
- (3) The user  $u$  securely sends  $\delta_i$  for each key authority  $A_i$
- (4) The CA and each key authority  $A_i$  run a secure two-party computation (2PC), where the private input of the CA is  $(\gamma_i, \beta)$ , authority  $A_i$ 's private input is  $(\alpha_i + \delta_i)$ , and the protocol returns the private output  $x = (\alpha_i + \gamma_i + \delta_i)\beta$  to each  $A_i$

The subsequent key generation process is the same as Hur and Kang's model [19], and the encryption and decryption algorithms are the same. After completing the key generation algorithm, the user  $u$  gets a his/her secret key (where  $D = \prod_{i=1}^m D_i = g^{(\alpha_1 + \dots + \alpha_m) + cGID_{u,t_1} + uGID_{u,t_1} \beta}$ ):

$$SK_u = \left( D, \forall \lambda_j \in S : D_j = g^{(cGID_{u,t_1} + uGID_{u,t_1})} \cdot H(\lambda_j)^{r_j}, D'_j = g^{r_j} \right). \quad (4)$$

**5.2. Get Ciphertext.** After completing the deposit setup, users can participate in the trade with other users. In the proposed system, the seller stores his/her encrypted data on the cloud server, and after the transaction is completed, the cloud service provider provides the data to the buyer. However, since transactions do not go through trusted intermediaries, the trading between users on the external channel cannot guarantee the fairness of the trading. Therefore, we applied HTLC to the proposed system to ensure the fairness of the trading as shown in Figure 5. In the system, data trading and data sharing between the buyer and the seller are as follows:

- (1) The seller stores encrypted data  $CT$  on the cloud server
- (2) The buyer requests the seller to sell the data  $CT$
- (3) The seller randomly chooses  $R \in \mathbb{Z}_p^*$  and computes  $hproof = H(R)$  and then sends  $hproof$  to the CSP and the buyer. In the proposed system, the CSP manages lists for managing each user's data. The list consists of  $\{ID_{seller}, CT, txID, hproof\}$  and the CSP provides corresponding data only if the buyer has provided valid proof (i.e., preimage of  $hproof$ )
- (4) The buyer deposits the transaction amount to the trading support contract (TSC) with the  $hproof$
- (5) The seller submits his/her digital signature  $\sigma_{seller}$  and the preimage of  $hproof$  (i.e.,  $R$ ) to TSC and receives the transaction amount
- (6) To request access to the data  $CT$ , the buyer submits the value  $R$  exposed by the seller in the process of receiving the transaction amount to the CSP. The buyer generates the following message and sends it to the CSP to show that the state of his/her smart contract is active and that the trade with the seller has been completed

$$msg_{req} = \left\{ SC, txID, R', k_{u,t_1}, CT \right\} \quad (5)$$

- (7) If  $SC.state = active$ ,  $SC.hindex = H(k_{u,t_1})$ , and  $hproof = H(R')$ , the CSP returns the requested ciphertext  $CT$  to the buyer

**5.3. State Transition.** The activated SC can be transited to three states: active, close, and revoke. We have five state transition scenarios as shown in Figure 6. A detailed description of each state transition is as follows:

- (i) [active  $\rightarrow$  active]: for the user revocation, all users in the system must periodically update their secret keys. The update period  $t$  is determined by the CA, which is inserted into the user's contract SC during the deposit setup process. Therefore, all contracts in the system have the same timer of update period  $t$ .

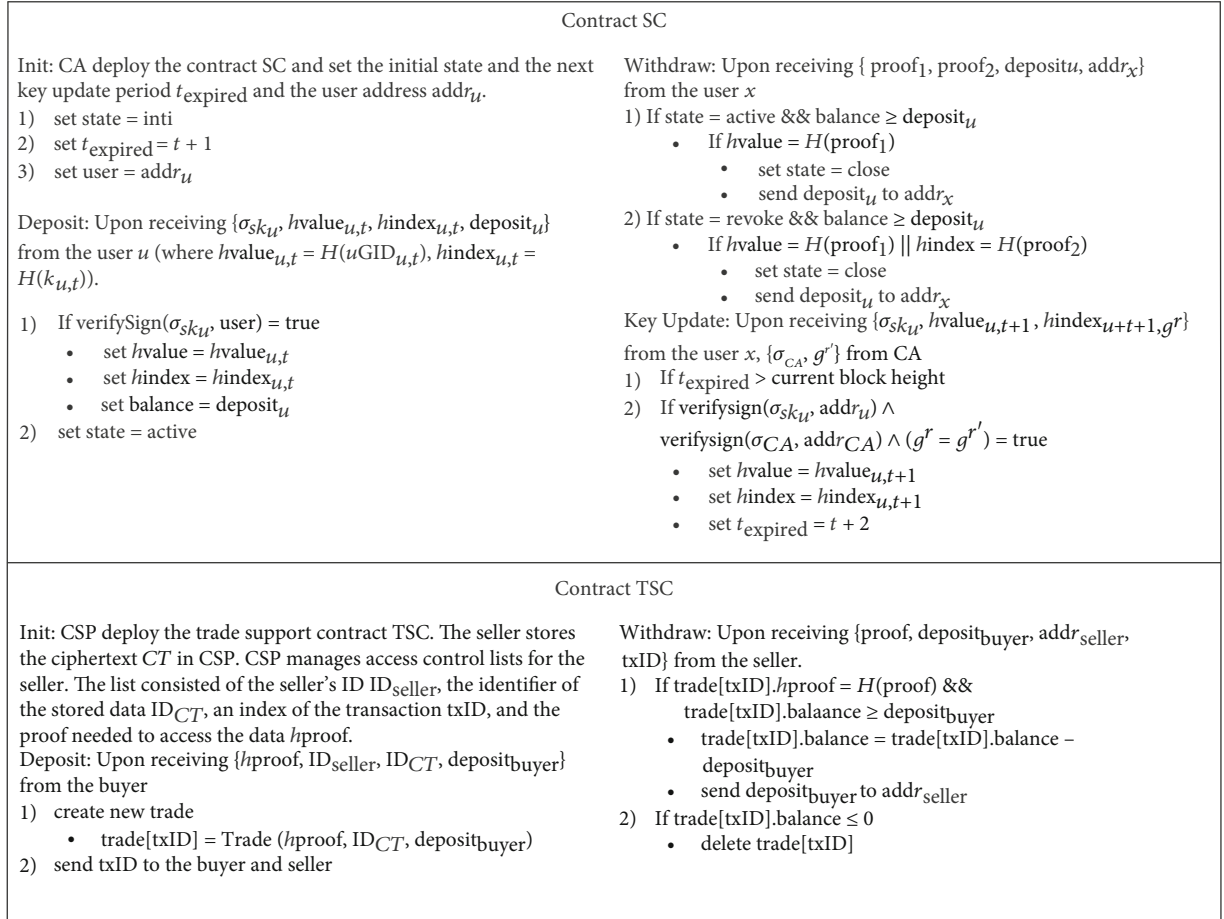


FIGURE 3: Contract design for SC and TSC.

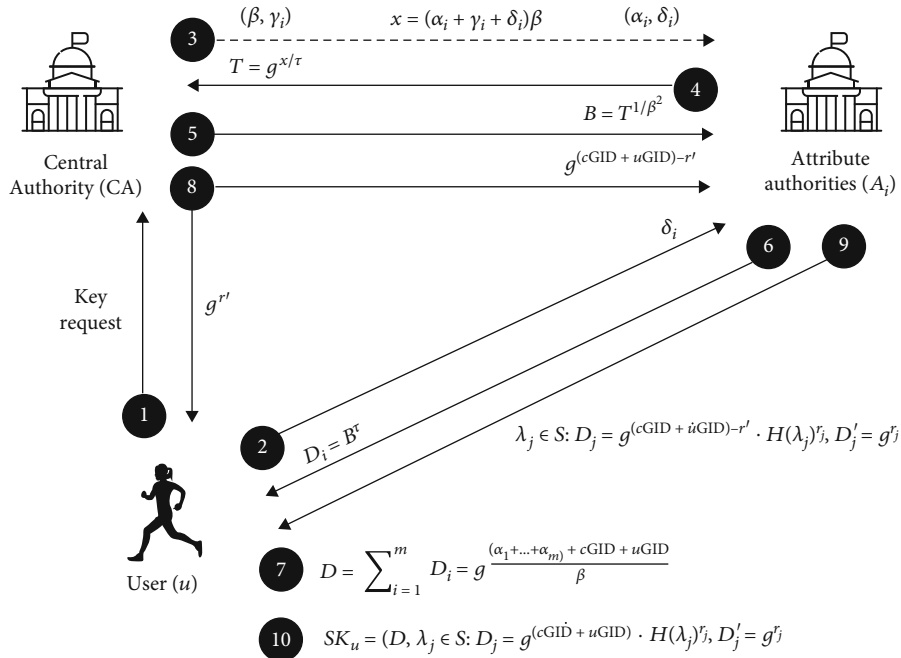


FIGURE 4: Key generation phase in the proposed system.

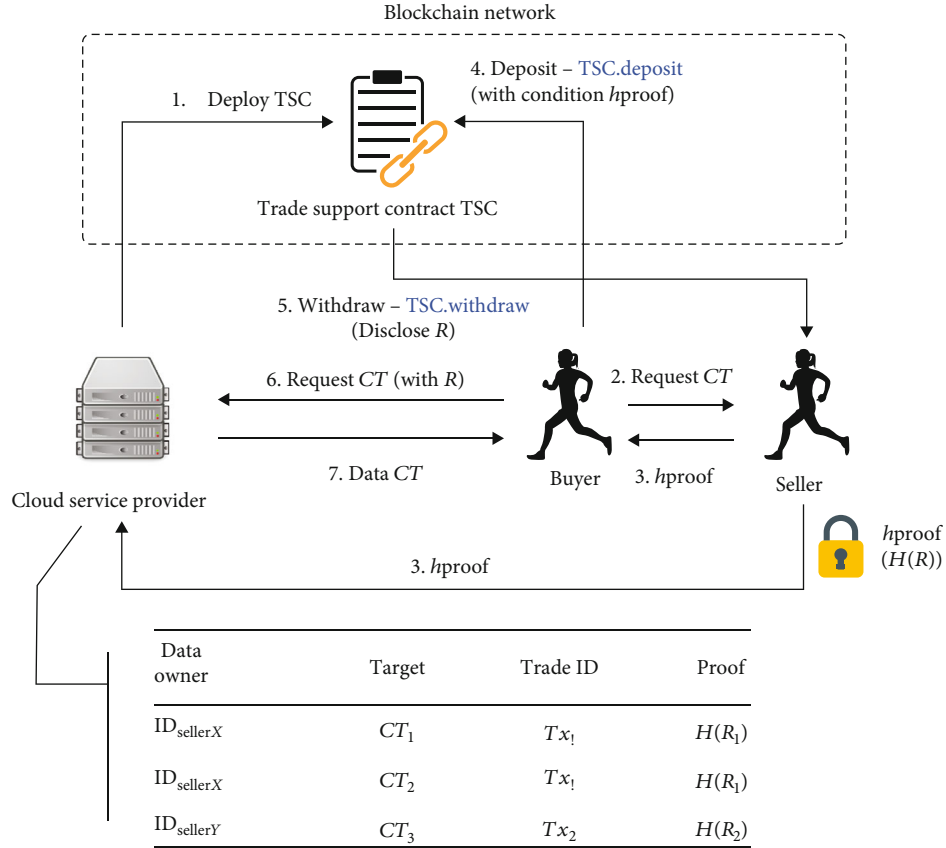


FIGURE 5: Data transaction overview.

Since implementing the synchronized timer in an asynchronous network is hard, we measure this period  $t$  as the  $n$ th block is appended to the blockchain (e.g., update every 100 blocks appended to the blockchain). If the user takes no action until the  $n$ th block is appended, his/her contract will automatically be transited to the revoke state. In the revoke state contract, the user loses ownership of the deposit locked in the contract, so to avoid this situation, all users will participate in the key update function honestly before the contract timer is expired as shown in Figure 6 (case 1). To update the secret key, the user chooses a new random value  $\delta_i, k_{u,t_2} \in \mathbb{Z}_p^*$  for each attribute authority and the new proof of ownership. Then, set  $uGID_{u,t_2} = \sum_{i=1}^m \delta_i$  and compute  $hvalue = H(uGID_{u,t_2}, hindex = H(k_{u,t_2}))$ . The user uses  $uGI D_{u,t_2}$  to generate a new secret key as in the deposit setup phase. At the end of the key generation process in Hur and Kang's model, the CA generates  $g^{r'}$  and sends it to the user. The user finally gets the secret key using  $g^{r'}$  received from the CA and the part of the secret key  $D_j$  received from the attribute authorities. The user and CA send this value  $g^{r'}$  to the contract after key generation is completed, and if the values submitted by the CA and the user are the same, the contract is transited to the active state of the new period  $t_2$ .

- (ii) [active  $\rightarrow$  close]: a deposit in the contract can be withdrawn by showing a preimage of  $hvalue$  (i.g.,  $uGID_{u,t}$ ) using the withdraw function shown in Figure 3. The user can withdraw their deposit before the contract timer  $t$  is expired himself (case 2). To withdraw the deposit, the user  $u$  submits the secret value  $uGID_{u,t}$  used to generate his/her secret key to the contract. When the deposit is withdrawn, the contract transits to a close state automatically. The withdrawal algorithm of the proposed model does not verify the original owner of the deposit but simply verifies that it has valid proof as shown in Figure 3. Therefore, anyone with valid proof can withdraw the deposit without the original owner's consent. As mentioned earlier, in the proposed model, the user who wants to join the collusion attack must disclose his/her secret value  $uGID_{u,t}$  to the colluder. That is, if the secret value  $uGID_{u,t}$  is disclosed to another user and the deposit is withdrawn, the contract is equally transited to the close state (case 3)
- (iii) [active  $\rightarrow$  revoke]: a transit to the revoke state occurs due to the user's revocation. As mentioned earlier, we use a revocation scheme that periodically regenerates a new secret key for all valid users in the system. However, if a user does not complete a key update (case 5) or is revoked from the system before the key update (case 4), the contract will be transited

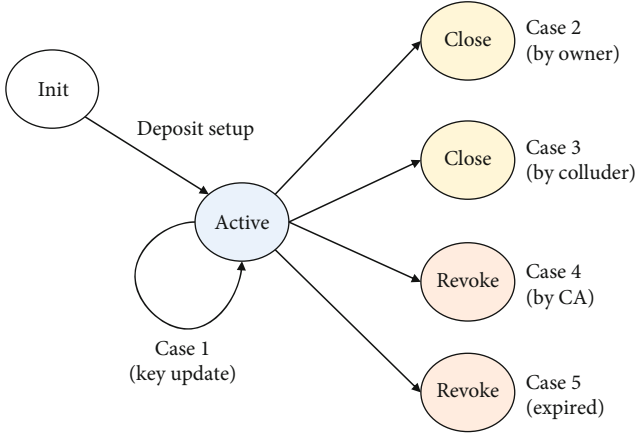


FIGURE 6: A state transition diagram of the proposed system.

to the revoke state. After the state of the contract is transitioned to revoke,  $hindex$  is added as an option to the conditions for deposit withdrawal. In other words, the deposit of the revoked user can be withdrawn with knowledge of the preimage of either  $h$  value or  $hindex$  as shown in Figure 3

## 6. Analysis and Evaluation

### 6.1. Security Analysis

- (1) *Collusion resistance*: our proposed system ensures resistance to collusion (a) *among users* and (b) *among a user and central authority*. All users in the system can share their secret keys with collusion attackers. An attacker will pay a certain reward price<sub>reward</sub>  $\times n$  to induce users to share their secret keys (where  $n$  is the number of colluders). They will be able to obtain unauthorized data through attacks and benefit price<sub>data</sub> from it. In the end, the benefit of the collusion attacker from the attack would be

$$\text{price}_{\text{attack}} = \text{price}_{\text{data}} - (\text{price}_{\text{reward}} \times n) \quad (6)$$

However, if the reward to colluders price<sub>reward</sub>  $\times n$  is greater than the benefits from the attack (i.e., price<sub>data</sub>  $<$  (price<sub>reward</sub>  $\times n$ )), the attackers will lose the motivation of the attack, because they will suffer financial damage even if the attack succeeds.

- (a) *Collusion among users*: a collusion attack among users can be prevented by making it impossible to combine if the global identifier  $GID_u$  used for key generation is not the same. In the existing scheme, the CA chooses random exponents for each attribute authority  $A_i$  and sets  $GID_u = \sum_{i=1}^m \gamma_i$  (where  $GID_u$  is uniquely determined by each user). Then, each attribute authority  $A_i$  and CA cooperatively generate the user's secret key based on their master key and given  $\gamma_i$ . Even though the attribute authority generates a secret key corresponding to the same attribute set,

each user has a different secret key because the exponent  $\gamma_i$  is different. Therefore,  $GID_u$  used to generate a user's secret key can never be reconstructed from the secret keys of different users. Previous models used a CA-determined global identifier  $GID_u$  for secret key generation. Therefore, the CA knew the parameters that each attribute authority used to generate a user's secret key and could access the ciphertexts in the system without having enough attributes. In contrast, in our proposed system, the CA and user choose and combine the global identifier to be used for key generation separately. In more details, the CA generates a global identifier  $cGID_{u,t} = \sum_{i=1}^m \gamma_i$  the same for previously proposed models and the user  $u$  also generates a global identifier  $uGI_{u,t} = \sum_{i=1}^m \delta_i$  at period  $t$ . The user  $u$  keeps  $uGI_{u,t}$  secret from the CA and sends  $\delta_i$  to each attribute authority  $A_i$ . Each  $A_i$  performs a two-party computation with the CA using its master key and  $\delta_i$  received from the user as a private input. Finally, the user  $u$  receives  $D_i = g^{(\alpha_i + \gamma_i + \delta_i)\beta}$  from each attribute authority and computes a part of his/her secret key  $D = \prod_{i=1}^m D_i = g^{(\alpha_1 + \dots + \alpha_m) + cGID_{u,t} + uGI_{u,t}\beta}$  from it. In the proposed model,  $GID_{u,t} = cGID_{u,t} + uGI_{u,t} = \sum_{i=1}^m (\gamma_i + \delta_i)$  is computed from the random exponents chosen by the CA ( $\gamma_i$ ) and the user ( $\delta_i$ ). In other words, the CA only knows a part of the user's global identifier and the random parameters used by each attribute authority to generate secret keys. To achieve the collusion among users requires that attackers and colluders should know their global identifier used to generate their secret keys. If attackers know the global identifier used to generate secret keys, they can combine secret keys generated from the same global identifier. An attacker can combine colluders' secret keys generated from the same  $GID$  or create a new secret key using the same  $GID$  of the colluder to attempt the collusion attack. In other words, the attacker should be able to recover the colluder's global identifier. However,  $cGID_u$  chosen by the CA is not disclosed to the user on the key generation process, and in contrast, the colluder may disclose his/her  $uGI_u$  to the attacker but this will be shared with a trustworthy user, the ownership of his/her deposit until the next key update period  $t + 1$ . That is, the colluder has the risk of losing his/her deposit by the collusion attacker until the end of the period  $t$ , whether or not the collusion attack is successful. The colluder will only join the attack if the attacker pays at least more than the deposit he may lose (i.e., price<sub>reward</sub>  $\geq \mathcal{D}_u$ ). Theoretically, the attacker will not be able to obtain the  $GID$  for the attack, neither from the CA nor from colluders, so the proposed system can provide resistance to the collusion attack among users

- (b) *Compromised authority*: Hur's model had a vulnerability that a compromised CA could generate a secret key for the unauthenticated user from the secret key



of the colluder. In this attack scenario, the CA receives the part of the secret key  $D = g^{(\alpha_1 + \dots + \alpha_m) + \text{GID}_u \beta}$  from colluder  $u$  and computes secret information of the attribute authorities  $g^\alpha = g^{(\alpha_1, \dots, \alpha_m)}$  that required to generate a new secret key (where  $\alpha_i$  is a master secret key of the attribute authority  $A_i$ ). The CA can generate a new secret key alone without the cooperation of attribute authorities based on  $g^\alpha$ . However, in our proposed system, the CA needs the global identifier of the colluder  $\text{GID}_u = c\text{GID}_u + u\text{GID}_u$  to extract  $g^\alpha$  from his/her secret key. In other words, the CA must receive  $u\text{GI} D_u$  along with the secret key from the colluder. However,  $u\text{GID}_u$  is secret information needed to withdraw the deposit from the colluder's smart contract. If the colluder  $u$  shares his/her global identifier  $u\text{GID}_u$  with the CA, the CA may withdraw the colluder's deposit by submitting the preimage of  $h$ value,  $u\text{GID}_u$ , to the smart contract

- (2) *User revocation*: for situations where a user leaves the system (by themselves or by force), the system must provide a revocation of the user's secret key. If the system does not provide the key revocation, our deposit protocol will not guarantee resistance to collusion attacks. Our proposed deposit protocol relies on activated security deposits to prevent users from malicious behavior. All users of the system will always act rationally because they have the risk of losing their deposit due to their misbehavior. This causes an attacker to lose motivation for the attack by making it more costly for the attack. However, assuming the key revocation is not provided, the user's deposit will be withdrawn if the user leaves the system, while the secret key will remain valid in the system. In other words, the system cannot prevent a user who has already left the system from using his/her secret key in the collusion attack. A user revocation prevents the secret key of the user who has left the system from being used in the system. In our proposed system, we apply a user revocation to the system, which periodically regenerates and reencrypts secret keys and ciphertexts of all users in the system so that revoked user's secret keys can no longer be used. However, even if the user is revoked, his/her secret key is still valid until the next key update period. This can be solved by shortening the key update period, but it results in a high computational cost on the system. We adopt a method to prevent revoked users from using their secret keys until the next key update period based on the security deposit. In the proposed system, a user revocation can be divided into two cases:

- (i) *Revoked by the CA*: the CA can revoke a user from the system and prevent that user from participating in the system. the CA can transit the

state of the user's contract SC from active to revoke. However, even if the contract is transited to the revoke state, the revoked user still retains ownership of the deposit locked in the contract and his/her secret key can also be used until the next key update period. However, revoked users may ignore the revoked state and participate in data trading. The revoked user submits his/her knowledge of the preimage of SC.hindex to show ownership of the deposit while depositing the transaction amount in TSC for the data trading. In the active state, withdrawals of a security deposit require submission of SC.hvalue's preimage (i.e.,  $u\text{GID}$  at period  $t$ ), but in the revoke state, the deposit withdrawal can be made by submission of SC.hindex's preimage. TSC approves deposits of the transaction amount only when the buyer's SC.state is active. In other words, the revoked user is no longer able to proceed with the trading process. Furthermore, his/her secret information  $k$  (preimage of SC.hindex) is exposed to the blockchain network, so his/her security deposit could be withdrawn by a third-party user in the system

- (ii) *Expired*: all users must periodically update their secret keys. If the user did not update the secret key at period  $t$ , his/her SC is automatically transited to the revoke state as the next period  $t + 1$  begins. The revoked user can withdraw his/her deposit and participate in the system again with a new deposit
- (3) *Decentralization*: in the proposed system, data trading among users is made through a smart contract TSC. In each trade, the seller provides data to the buyer and the buyer pays the seller its price. However, on the decentralized trading platform where trusted intermediaries do not participate, fairness issues arise in the order of the trading protocol. Therefore, we adopt a method to ensure fairness of the decentralized trading based on HTLC. When the buyer requests the seller to sell the data  $CT$ , the seller chooses a random value  $R$ , then computes  $h\text{proof} = H(R)$ , and sends it to the buyer and CSP. In our proposed system, the CSP manages a special list for each seller. This list is an access control list of the seller's data. The CSP provides requested data only when the buyer submits a valid proof (i.e., preimage of  $h$  proof =  $H(R)$ ). The buyer sets the seller's digital signature  $\sigma_{\text{seller}}$  and the preimage of  $h\text{proof}$  as a withdrawal condition while depositing the transaction amount to TSC. The seller submits randomly selected  $R$  to TSC to withdraw the price, in which  $R$  is disclosed to all participants in the blockchain network. The buyer can get the purchased data by submitting the disclosed  $R$  to the CSP. In the proposed system, the random value  $R$  was used as a method to ensure the fairness of the trade transaction. In the trading protocol, two situations can be considered:

TABLE 2: Computational cost.

Phase	Entity	Operation	Cost
Phase 1 (deposit setup)	CA	Deploy SC	1tx
	User	Randomly choose $\delta_i, k \in \mathbb{Z}_p^*$	$(i + 1)$ RNG
		Compute $uGID = \sum \delta_i$ , $H(uGID), H(k)$	1sum 2hash
		Send Tx to SC (deposit)	1tx
Phase 2 (get ciphertext)	Seller	Randomly choose $R \in \mathbb{Z}_p^*$	1 RNG
		Compute $H(R)$ Generate signature $\sigma_{seller}$	1hash 1sig
	Buyer	Send Tx to TSC (deposit)	1tx
	CSP	Compute $H(k), H(R')$	2hash
Phase 3 (state transition: Update)	CA	—	—
	User	Randomly choose $\delta_{i,new}, k_{new} \in \mathbb{Z}_p^*$	$(i + 1)$ RNG
		Compute $uGID_{new} = \sum \delta_{i,new}$ , $H(uGID_{new}), H(k_{new})$	1sum 2hash
		Send Tx to SC (update)	1tx

TABLE 3: Computational cost notation.

Operation	Meaning
tx	Create blockchain transaction and broadcast it to the network
sig	Generate a digital signature (ECDSA)
hash	Cryptographic hash function (Keccak-256)
sum	Summation operation
RNG	Random number generate function

(i) *Malicious seller*: may not share the purchased data after receiving the price from the buyer

(ii) *Malicious buyer*: may not pay the price after receiving purchased data from the seller

In the proposed system, the seller and buyer set specific conditions  $hproof$  for data sharing and trade transactions, respectively. For more details, the seller sets specific conditions for data sharing, and then, the buyer also sets a copy of conditions (i.e., same conditions  $hproof = H(R)$  set by seller) for withdrawing price in TSC. A buyer can get purchase data by submitting valid proof to the CSP, and a seller can also submit valid proof to TSC to withdraw the price from the contract. When the seller submits the random value  $R$  as proof to withdraw the price, it is recorded in the blockchain and automatically disclosed to all network nodes. The seller cannot withdraw the price from TSC without disclosing  $R$ . The buyer also cannot find out the  $R$  randomly chosen by the seller and submit it to the CSP without the transaction being completed. The buyer also cannot find out the random value  $R$  chosen by the seller and submit it to the CSP without the seller completing the withdrawal of the price.

**6.2. Evaluation and Implementation.** Our approach is to mitigate the dependence of fully trusted CA in the system by

TABLE 4: Costs required to execute the proposed contract SC and TSC.

Phase	Command	Cost (gas)	Cost (\$)
Phase 1	Deploy contract SC	399816	63.84
	Deposit (SC.deposit)	69419	11.08
Phase 2	Deploy contract TSC	335323	53.54
	Deposit (TSC.deposit)	103479	16.52
	Withdraw (TSC.withdraw)	45322	7.23
Phase 3	Key update (SC.keyupdate)	65940	10.52
	Withdraw (SC.withdraw)	43661	6.97

issuing only a portion of GID by the CA and the user, instead of entirely issuing GID by the fully trusted CA. In the key generation phase, the user randomly selects a secret value  $u$  GID to be used as a portion of his GID and submits the deposit to the smart contract. Deposits locked in the smart contract can be withdrawn by submitting this secret value  $u$  GID to the contract. In summary, compared to previous researches, our approach can mitigate the dependence of the CA but the cost of user participation in the protocol is inevitable. Table 2 shows the computational cost of the system entities at each phase of the proposed protocol. Table 3 lists the notations to express the computational cost.

To evaluate the computational cost, we consider only the computational cost of each entity except for the operations performed by the smart contract. We also consider only the operations that each entity performs to communicate with smart contracts as evaluation targets.

In phase 1, the CA deploys a smart contract SC for each user. The deployment of smart contracts requires 1tx operation, and the CA keeps the addresses of all smart contracts deployed (each address is 20 bytes in size). The user submits a security deposit to SC for CP-ABE key generation. The user

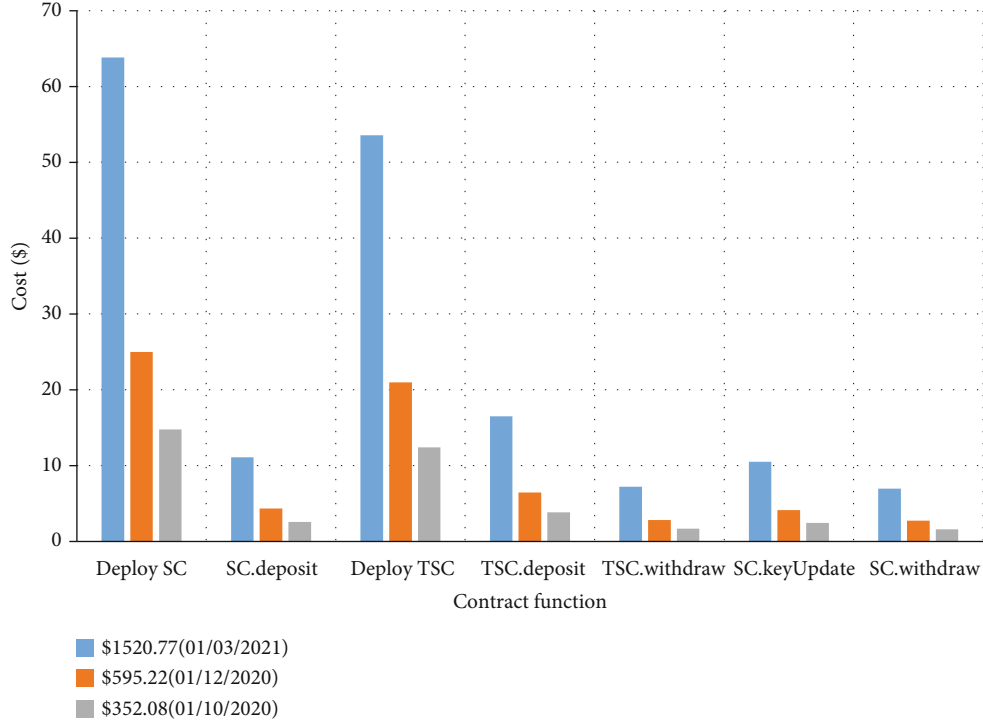


FIGURE 7: Cost of contract functions according to the exchange rate in the proposed system.

performs an  $i$  RNG operation that chooses the random number for  $i$  key authorities in the system to generate  $\mu$ GID, 1 RNG operation that generates a random number for the proof of ownership of the contract SC and 1tx operation that submits a deposit after 1sum operation for  $\mu$ GID generation.

In phase 2, the data seller performs 1 RNG operation that chooses a random  $R$  for data trading, 1 hash operation to generate an HTLC proof, and a digital signature to generate operation 1 sig for deposit acquisition. The buyer performs 1tx operation that submits a deposit, and the CSP performs 1hash operation for the seller's data table update and 1hash operation for the buyer's data request verification.

In phase 3, the user updates his smart contract SC by performing the same operations, as in phase 1, for new key generation.

In our protocol, users' deposits are controlled by HTLC based on simple hash operations instead of complex operations. Except for the  $(i + 1)$ RNG operation required in the key generation (and key update) process, other operations can be seen efficiently in practical terms with low-cost operations. However, since this can be resolved by properly adjusting the key update cycle, the computational cost of the user is reasonable even if the proposed protocol is applied.

We implemented the smart contract SC and TSC on the *Kovan* Ethereum test network. The price of the gas that we set in the test is  $1.6 \times 10^{-7}$  ether (160 Gwei) and the cost required to execute the contract is calculated as the gas price  $\times$  gas used. Unfortunately, however, it is not appropriate to show the execution cost of the contract at this point due to the soaring price of the cryptocurrency in 2021. Therefore, we present the costs with the current and previous exchange rates together for a more appropriate evaluation.

Table 4 shows the results of converting the costs required to execute our contract into gas and US dollars (exchange rate as of 01/03/2021: 1 ether = \$1520.77) at each phase in the proposed system. Figure 7 shows a recalculation of the execution cost of each function of the contract shown in Table 3 according to the previous exchange rate (exchange rate as of 01/10/2020: 1 ether = \$595.22 and 01/12/2020: 1 ether = \$352.08). In the proposed system, the honest user pays for the data trading (SC.deposit, SC.withdraw) and a periodic key update (SC.keyUpdate), and the cost of each function, excluding contract deployment costs, is reasonable because it is less than \$5 with the previous exchange rate.

## 7. Conclusions

In this paper, we propose a data trading protocol for decentralized user-to-user data trading and a security deposit protocol for preventing collusion attacks in the multi-authority ABE system. Our system is controlled by two smart contracts SC and TSC, and instead of centralized methods by a trusted third party, our system leads to honest behavior of users based on their security deposit. However, the management history of the user's deposit is transparent to everyone on the network. If a malicious third party can relate a user's identity in the ABE system to his/her deposit in the blockchain network, there is a possibility of a new attack resulting from this vulnerability. Applying many privacy-preserving techniques used in the blockchain, we think that it would be possible to further improve the proposed protocol in terms of security. Moreover, we expect the proposed system to be utilized as a way to ensure that the system operates

honestly without a central system administrator in the modern complex society.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Acknowledgments

This research was supported by the MSIT (Ministry of Science, ICT), Korea, under the High-Potential Individuals Global Training Program (2020-0-01596) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation) and partially supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2021-2020-0-01797) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

## References

- [1] D. Reinsel, J. Gantz, and J. Rydning, "The digitization of the world from edge to core[white paper]," 2018, <http://cloudcode.me/media/1014/idc.pdf>.
- [2] Z. Cai, Z. He, X. Guan, and Y. Li, "Collective data-sanitization for preventing sensitive information inference attacks in social networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 4, pp. 577–590, 2016.
- [3] Z. Xiong, H. Xu, W. Li, and Z. Cai, "Multi-source adversarial sample attack on autonomous vehicles," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 3, pp. 2822–2835, 2021.
- [4] K. Li, G. Luo, Y. Yang, W. Li, S. Ji, and Z. Cai, "Adversarial privacy-preserving graph embedding against inference attack," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6904–6915, 2020.
- [5] "What is Google cloud marketplace?," <https://cloud.google.com/marketplace/docs>.
- [6] "AWS marketplace," <https://aws.amazon.com/marketplace>.
- [7] B. Goertzel, S. Giacomelli, D. Hanson, C. Pennachin, and M. Argentieri, "SingularityNET: a decentralized, open market and inter-network for AIs[white paper]," 2017, <http://airesearch.com/ai-research-papers/singularitynet-a-decentralized-open-market-and-inter-network-for-ais/>.
- [8] O. Protocol, "A decentralized data exchange protocol to unlock data for AI[white paper]," 2020, <https://oceanprotocol.com/>.
- [9] X. Ren, P. London, J. Ziani, and A. Wierman, "Datum: managing data purchasing and data placement in a geo-distributed data market," *IEEE/ACM Transactions on Networking*, vol. 26, no. 2, pp. 893–905, 2018.
- [10] G. Lin, H. Hong, and Z. Sun, "A collaborative key management protocol in ciphertext policy attribute-based encryption for cloud data sharing," *IEEE Access*, vol. 5, pp. 9464–9475, 2017.
- [11] C. Li, J. He, L. Cheng, C. Guo, and K. Zhou, "Achieving privacy-preserving CP-ABE access control with multi-cloud," in *2018 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom)*, pp. 801–808, Melbourne, VIC, Australia, 2018.
- [12] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *2007 IEEE symposium on security and privacy*, pp. 321–334, Berkeley, CA, USA, 2007.
- [13] B. Waters, "Ciphertext-policy attribute-based encryption: an expressive, efficient, and provably secure realization," in *Public Key Cryptography – PKC 2011. PKC 2011*, D. Catalano, N. Fazio, R. Gennaro, and A. Nicolosi, Eds., vol. 6571 of Lecture Notes in Computer Science, pp. 53–70, Springer, Berlin, Heidelberg, 2011.
- [14] M. Chase, "Multi-authority attribute based encryption," *Proceedings of the Theory of cryptography conference*, pp. 515–534, 2007.
- [15] K. Yang and X. Jia, "Attributed-based access control for multi-authority systems in cloud storage," in *2012 IEEE 32nd International Conference on Distributed Computing Systems*, pp. 536–545, Macau, China, 2012.
- [16] A. Gao and Z. Li, "Free global ID against collusion attack on multi-authority attribute-based encryption," *Security and Communication Networks*, vol. 6, no. 9, p. 1152, 2013.
- [17] K. Riad, "Multi-authority trust access control for cloud storage," in *2016 4th International Conference on Cloud Computing and Intelligence Systems (CCIS)*, pp. 429–433, Beijing, China, 2016.
- [18] J. Hur, D. Koo, S. O. Hwang, and K. Kang, "Removing escrow from ciphertext policy attribute-based encryption," *Computers & Mathematics with Applications*, vol. 65, no. 9, pp. 1310–1317, 2013.
- [19] J. Hur and K. Kang, "Secure data retrieval for decentralized disruption-tolerant military networks," *Proceedings of the IEEE/ACM transactions on networking*, vol. 22, no. 1, pp. 16–26, 2012.
- [20] S. Wang, K. Liang, J. K. Liu, J. Chen, J. Yu, and W. Xie, "Attribute-based data sharing scheme revisited in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 8, pp. 1661–1673, 2016.
- [21] E. Meamari, H. Guo, C.-C. Shen, and J. Hur, "Collusion attacks on decentralized attributed-based encryption: analyses and a solution," 2020, <https://arxiv.org/abs/2002.07811>.
- [22] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in *Advances in Cryptology – EUROCRYPT 2005. EUROCRYPT 2005*, R. Cramer, Ed., vol. 3494 of Lecture Notes in Computer Science, pp. 457–473, Springer, Berlin, Heidelberg, 2005.
- [23] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Advances in Cryptology. CRYPTO 1984*, G. R. Blakley and D. Chaum, Eds., vol. 196 of Lecture Notes in Computer Science, pp. 47–53, Springer, Berlin, Heidelberg, 1984.
- [24] H. Gao, Z. Ma, S. Luo, Y. Xu, and Z. Wu, "BSSPD: a blockchain-based security sharing scheme for personal data with fine-grained access control," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 6658920, 20 pages, 2021.
- [25] Y. Zhang, D. He, and K.-K. R. Choo, "BaDS: blockchain-based architecture for data sharing with ABS and CP-ABE in IoT," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 2783658, 9 pages, 2018.

- [26] A. Lewko and B. Waters, "Decentralizing attribute-based encryption," in *Advances in Cryptology – EUROCRYPT 2011. EUROCRYPT 2011*, K. G. Paterson, Ed., vol. 6632 of Lecture Notes in Computer Science, pp. 568–588, Springer, Berlin, Heidelberg, 2011.
- [27] Y. Rahulamathavan, S. Veluru, J. Han, F. Li, M. Rajarajan, and L. Rongxing, "User collusion avoidance scheme for privacy-preserving decentralized key-policy attribute-based encryption," *IEEE Transactions on Computers*, vol. 65, no. 9, pp. 2939–2946, 2015.
- [28] K. Wüst and A. Gervais, "Do you need a blockchain?," in *2018 Crypto Valley Conference on Blockchain Technology (CVCBT)*, pp. 45–54, Zug, Switzerland, 2018.
- [29] K. Croman, C. Decker, I. Eyal et al., "On scaling decentralized blockchains," in *Financial Cryptography and Data Security. FC 2016*, J. Clark, S. Meiklejohn, P. Ryan, D. Wallach, M. Brenner, and K. Rohloff, Eds., vol. 9604 of Lecture Notes in Computer Science, pp. 106–125, Springer, Berlin, Heidelberg, 2016.
- [30] C. Decker and R. Wattenhofer, "Information propagation in the bitcoin network," in *IEEE P2P 2013 Proceedings*, pp. 1–10, Trento, Italy, 2013.
- [31] M. Castro and B. Liskov, "Practical byzantine fault tolerance," *Proceedings of the Third Symposium on Operating Systems Design and Implementation*, vol. 99, no. 1999, pp. 173–186, 1999.
- [32] J. Poon and T. Dryja, "The bitcoin lightning network: scalable off-chain instant payments[White paper]," 2016, <https://lightning.network/lightning-network-paper.pdf>.
- [33] P. McCorry, S. Bakshi, I. Bentov, A. Miller, and S. Meiklejohn, "Pisa: arbitration outsourcing for state channels," in *Proceedings of the 1st ACM Conference on Advances in Financial Technologies*, pp. 16–30, Zurich, Switzerland, 2019.
- [34] M. Herlihy, "Atomic cross-chain swaps," in *Proceedings of the 2018 ACM Symposium on Principles of Distributed Computing*, pp. 245–254, Egham, United Kingdom, 2018.
- [35] "Tether price index - CoinDesk 20," <https://www.coindesk.com/price/tether>.

## Research Article

# Trustworthy Image Fusion with Deep Learning for Wireless Applications

Chao Zhang,<sup>1</sup> Haojin Hu,<sup>1</sup> Yonghang Tai<sup>1</sup> ,<sup>1</sup> Lijun Yun<sup>2</sup> ,<sup>2</sup> and Jun Zhang<sup>1</sup> 

<sup>1</sup>School of Physics and Electronic Information, Yunnan Normal University, Kunming, China

<sup>2</sup>School of Information Science and Technology, Yunnan Normal University, Kunming, China

Correspondence should be addressed to Lijun Yun; [yunlj@163.com](mailto:yunlj@163.com) and Jun Zhang; [junzhang@ynnu.edu.cn](mailto:junzhang@ynnu.edu.cn)

Received 21 April 2021; Revised 28 May 2021; Accepted 9 June 2021; Published 1 July 2021

Academic Editor: Weizhi Meng

Copyright © 2021 Chao Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To fuse infrared and visible images in wireless applications, the extraction and transmission of characteristic information security is an important task. The fused image quality depends on the effectiveness of feature extraction and the transmission of image pair characteristics. However, most fusion approaches based on deep learning do not make effective use of the features for image fusion, which results in missing semantic content in the fused image. In this paper, a novel trustworthy image fusion method is proposed to address these issues, which applies convolutional neural networks for feature extraction and blockchain technology to protect sensitive information. The new method can effectively reduce the loss of feature information by making the output of the feature extraction network in each convolutional layer to be fed to the next layer along with the production of the previous layer, and in order to ensure the similarity between the fused image and the original image, the original input image feature map is used as the input of the reconstruction network in the image reconstruction network. Compared to other methods, the experimental results show that our proposed method can achieve better quality and satisfy human perception.

## 1. Introduction

It is a big research challenge to fuse infrared and visible images to provide high-quality images for wireless applications, such as target recognition, visual enhancement, and cyber surveillance. Infrared images are mapped by infrared sensors capturing thermal radiation as a grayscale image and can emphasise thermal targets in low-light situations, but infrared images have a low resolution and do not show more detail in the scene. In contrast, the visible light sensor collects visible images to represent rich texture details, usually with higher resolution. Still, it is easily affected by imaging conditions (such as weather conditions, and lighting) [1]. The thermal radiation information of the infrared image and the texture information of the visible image can be fused to obtain an image with better visual quality and more information, which is the primary purpose of the fusion of infrared and visible images. Device can analyse the image which are been fused with computer vision and processing.

In the last few decades, many algorithms have been designed to implement the fusion of infrared and visible

images, which get good fusion result. Fusion algorithms for infrared and visual images can be divided into general methods and deep learning-based methods. Various image processing techniques are used for feature extraction in available image fusion methods [2]. Different fusion rules are designed for multimodal images, making the design complex and the generalization of the fusion poor. Along with the continuous development of deep learning, numerous scholars have developed image fusion models based on deep learning models [3]. Liu et al. first proposed a convolutional neural network- (CNN-) based fusion algorithm for infrared and visible images [4], which provides better fusion results than traditional methods. Liu et al. [5] used CNN as a feature extraction model to achieve the fusion of multifocused images by rule-based fusion. Li and Wu [6] proposed an auto-encoder-based method for fusing infrared and visible images, which can use feature maps to obtain fused images eventually. The deep learning-based fusion method of infrared and visible images has the following drawbacks: (1) the method based on deep learning still cannot get rid of manual rule design, and the deep learning frames just as part of

the fusion architecture; (2) the fusion strategy cannot achieve the fusion of infrared images in the item. The information is balanced with the visible image, and the fusion image is only similar to the source image; (3) the extracted compelling features were largely lost in the transmission process, and the feature information used for the fusion image is reconstructed with only a small amount of feature information.

We proposed a framework for fusing infrared images with visible images based on a deep learning model to solve the above issues. Our model is composed of three parts: a feature extraction network, a fusion network, and a reconstruction network. To ensure effective extraction of feature information, the output of features extracted by the feature extraction network in each convolutional layer will be fed to the next layer together with the output of the previous layer; short direct connections are built between each layer and all layers in a feed-forward fashion, thus effectively reducing the loss of valid information. In the feature fusion process, we use point-to-point approach to merge the feature maps of different channels to obtain the fused feature maps. In reconstruction network, the fused feature maps are the input, and the source image pair also used for reconstruction of fusion image. Considering trustworthy is a critical issue in the real-world applications of image fusion [7–10], we also propose to apply blockchain technology to protect sensitive information.

## 2. Related Work

In this section, we briefly describe the infrared and visible image fusion methods based on general and deep learning that have been developed in recent years, in particular for wireless applications. Initially, signal processing algorithms were widely used in image fusion [11], using mean and median filtering to extract the fundamental and detail layers of features before using dominant features to obtain a weight map and then combining these three components to obtain a fused image. The existing traditional methods of image fusion mainly consist of multiscale transform-based methods and sparse representation-based fusion methods. The original input image is decomposed into scale components of different scales in a multiscale transform-based approach [12], and each scale component is then fused according to specific rules, and finally, the combined image is obtained by the corresponding inverse scale transform. The main multiscale transforms are the pyramid transform [13], the wavelet transform [14], and the nondown sampled contour wavelet transform [15]. Sparse representation-based fusion methods learn dictionaries from high-quality images and then use the learned dictionaries to sparse representations of the source images. The method first decomposes the source images into overlapping blocks by a sliding window strategy and learns dictionaries from high-quality images, using the dictionaries to encode each image path sparsely. The sparse representation coefficients are then fused according to the fusion rules, and finally, the fusion coefficients of the fused images are reconstructed using the dictionaries, such as the joint sparse representation [16], the directional gradient

histogram-based fusion method [17], and the cospase representation [18]. Traditional methods require the manual design of feature extraction rules, feature fusion rules, and image reconstruction rules, resulting in computationally intensive and challenging designs.

Deep learning in the field of digital image processing has shown advanced performance in recent years; for the complex relationship between data, it can model the context knowledge and automatically extraction the perform feature without human intervention. Liu et al. [4] designed a sparse convolutional representation- (CSR-) based image fusion method to overcome the cumbersome rules in manual design. In 2017, Liu et al. [19] proposed the fusion of medical images using convolutional neural networks, which uses convolutional neural networks to generate pixel weight maps, but the method did not achieve total neural network fusion but rather multiscale transform fusion using image pyramids. Masi et al. [20] propose a fusion method, which are entirely based on deep learning; the method based on deep learning can extract the feature from image and reconstruct the fused image. In ICCV2017, the unsupervised learning framework was used for multiexposure image fusion by Prabhakar et al. [21], which has an extraordinary fusion loss function. Li and Wu [6] add dense block fast to this structure and design a separate fusion strategy in the fusion layer. Xu et al. [22] proposed an unsupervised and unified densely connected network for different types of image fusion tasks. Mustafa et al. [23] use multilevel dense network multifocus image fusion. Ma et al. proposed FusionGAN [24], which uses adversarial networks for image fusion, using discriminators to distinguish differences in the fused image from the original image. A dual-discriminator conditional generative adversarial network called DDcGAN [25] proposed by Ma et al. used to fusion multimodality medical images of different resolutions.

## 3. Materials and Methods

We proposed a deep neural network for infrared image, and visible image fusion is described in detail in this section. With the consideration of zero trust security model, blockchain is used to protect feature information. A private blockchain is implemented to store, share, and transmit feature data. The network consists of three main parts: a feature extraction network, a feature map fusion network, and a reconstruction network; above description is shown in Figure 1.

*3.1. Feature Extraction Network.* Extracting useful feature information from images of different modalities is a critical process in image fusion, and a good feature extraction strategy can reduce redundant feature information and provide more complex scene clues for subsequent processing. Therefore, the way of the feature extraction network is designed directly determines the effectiveness of the fusion. The feature extraction network proposed in this paper consists of 5 convolutional layers; each convolutional layers can obtain 48 feature maps by  $3 \times 3$  filters. The first convolutional layer will extract the details and global information of the source image, and the subsequent convolutional layers are used for

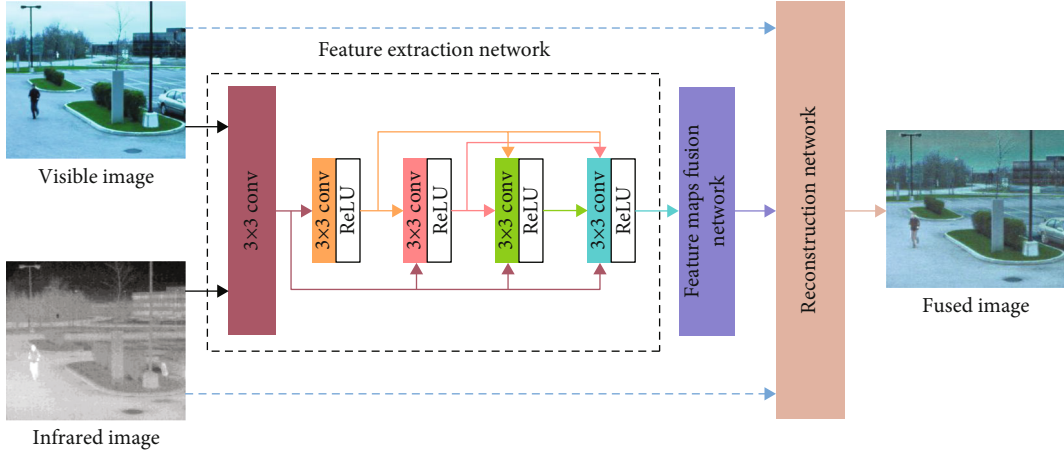


FIGURE 1: The overall architecture of the proposed network.

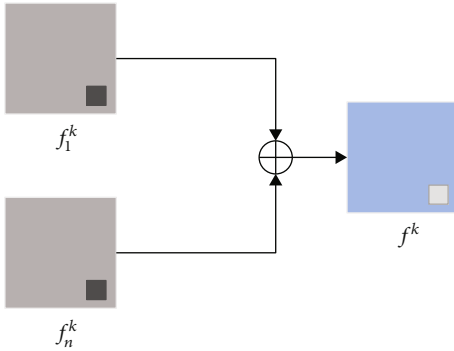


FIGURE 2: The architecture of combine with feature maps.

abstract feature generation. During the convolution process, the sequential sampling of the image makes the feature map gradually shrink, and a large number of valid information are lost. It cannot be repaired during the sampling process, this resulting in the disappearance of a large number of original features in the fused image. Therefore, we do not use pooling operations between the individual convolutional layers, but instead use the output of each layer, along with the output of the previous layer, as input to the next layer, model allowing valid information features to be passed throughout the convolutional network. Following Li and Wu [6], when the image of input is three-channel (RGB), each pair of the channel will be the input of the feature extraction network. To speed up convergence and avoid gradient sparsity, we use the ReLU activation function after each convolutional layer of the encoder.

**3.2. Features Fusion.** In DeepFuse [21], CNN is used to implement the fusion of exposure image pair; the feature maps obtained by the CNN are subjected to a point-to-point summation operation to get the final fused feature map; the same strategy was used in DenseFuse [6] by Li and Wu. Achieving accurate fusion is a difficult task, because the infrared and visible images are both come from different sensors. In this paper, following DeepFuse and DenseFuse,

we implement the pixel-level's point-to-point merging of the feature maps from the feature extraction network by using an addition strategy, which is shown in Figure 2.

The input image is extracted by the feature network to form a feature map;  $f_n^k(x, y)$  is the set composed of all feature maps, and  $f^k(x, y)$  represents the merged feature map. The merging strategy is shown in Equation (1).  $(x, y)$  is the corresponding position coordinates of the feature map and the fused feature map. The merged feature map will be used as the input to the reconstruction network reconstruct the fused image.

$$f^k(x, y) = \sum_1^n f_n^k(x, y). \quad (1)$$

**3.3. Reconstruction Network.** Image reconstruction is also an essential task for networks, and deconvolution is used typically to reconstruct images. In our network, we replace the deconvolution layers of the reconstructed network with regular convolution. The reconstructed network consists of four Conv layers, using a ReLU layer of  $3 \times 3$  kernel size. To feed the reconstruction network with more information, we use the input image as input to the reconstruction network, and the feature map and the original image were both used to reconstruct the fused image. The architecture of reconstruction network is shown in Figure 3. When the feature maps are calculated by feature extraction network and feature fusion layer, the source image pair and the fused maps are used for image reconstruction, the following equation defines this task:

$$\text{Fused}(x, y) = \sum_{i=1}^n f^i(x, y) \text{Source}_i(x, y), \quad (2)$$

where the  $\text{Fused}(x, y)$  is the fused image,  $f^i(x, y)$  represented the fused feature maps from feature maps fusion network,  $\text{Source}_i(x, y)$  is the source image pair, and  $(x, y)$  represents the corresponding pixel point.



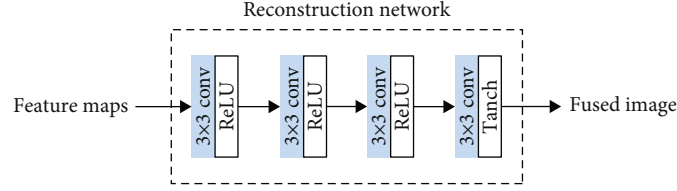


FIGURE 3: The architecture of combine with feature maps.



FIGURE 4: Source images of the VIFB dataset tested in our experiments, with the first row containing visible images and the second row containing infrared images.

## 4. Experimental Results and Discussion

In this section, first, we will describe the source images and the experimental environment. Secondly, we evaluate our fused images using subjective vision. Finally, the proposed algorithm is quantitatively assessed by using a variety of metrics. In order to validate the effectiveness of the deep learning model, we divided the comparison algorithms into general and deep learning-based methods in our experiments.

**4.1. Experimental Settings.** We train our proposed method with 5000 input images that we choose from MS-COCO dataset [26]; the learning rate is set to  $10^{-4}$ ; the batch size is set 24. Because there are no fully connected layers in our method, any same-scale infrared and visible image pairs can be fused using our model. Our experiment compares our model with even state-of-art image fusion methods in VIFB [27] with the particular consideration of wireless applications including object recognition and cyber resilience. VIFB is a visible and infrared image fusion benchmark, which consist 21 image pairs, and the size of each image pair is different. Four examples of VIFB are shown in Figure 4. The fusion methods used in this paper fall into two categories: general methods and methods based on deep learning. General methods include ADF [28], guided filter algorithm (GFF) [29], cross bilateral filter (CBF) [30], and VSMWLS [31]; methods based on deep learning include DenseFuse

[6], CNN [5], ResNet [32], and our method. DenseFuse, CNN, ResNet, and our model are implemented with Pytorch and trained with double Tesla V100, 16GB RAM GPUs. Other methods are implemented with MATLAB 2016B.

**4.2. Subjective Visual Evaluation.** In this section, subjective visual evaluations are used to assess the performance of various infrared and visible image fusion algorithms, which is based on the way of the human visual system. In order to validate the effectiveness of deep learning models, we classify the current fusion methods into categories: general and deep learning. We chose four images for the night environment and the daytime environment. In the daytime environment, the first image is darker in the evening, and the second chapter is better lit; in the dark night environment, the light is weaker in the first image than in the second image. All four images we selected contained thermal targets for verifying the algorithm's performance in highlighting thermal targets. The fusion results obtained by the different fusion algorithms are presented in Figure 5.

As shown in Figure 5, we use the red dashed line to divide the images into three groups. The first group shows the original input visible image with infrared image, the second group shows the image fusion results using the general methods, from top to bottom, ADF, GTF, CBF, and VSMWLS, respectively, and the third group offers the fused images based on deep learning methods, from top to bottom,

DenseFuse, CNN, ResNet, and our proposed methods. Among the general techniques, ADF and VSMMLS work better; the fused images obtained by GTF produce more significant artefacts, and the fused images contain more information about the infrared than the visible images; the CBF method achieves fusion, but a large number of blurred areas appear in the fused images. Deep learning-based methods achieve good image fusion with minimal visual discrepancies; CNN methods show coloured streaks when fusing images in daylight. DenseFuse and ResNet achieve better fusion results, and these methods achieve fused image images that contain more information about the original. Our fused images have three main advantages over other methods. Firstly, our results for hot tar (e.g., human portraits) have high contrast. Secondly, the images we obtain contain rich textural detail and more detailed information in the background. Thirdly, our method produces images that better balance the modalities of infrared and visible images and have a better visual perception, resulting in a more natural fusion.

**4.3. Quantitative Evaluation.** This section compares our approach with general methods and the approach base on the deep learning carried out in VIFB 21 for the quantitative analysis of images. We use ten metrics such as average gradient (AG) [33], correlation coefficient (CC), peak signal-to-noise ratio (PSNR) [34], information entropy (EN) [35], structural similarity of images (SSIM) [36], mutual information (MI) [37], image similarity metric based on edge information (Qabf) [38], pixel feature mutual information (FMI\_pixel) [39], discrete cosine characteristic mutual information (FMI\_dct) [39], and wavelet features mutual information (FMI\_w) [40] for evaluation.

- (i) *Average Gradient (AG).* This evaluation indicator reflects the sharpness of the image. The average gradient is calculated only necessary to consider the fused image, an evaluation metric that reflects the sharpness of the image and is defined by the following equation:

$$AG(F) = \frac{1}{(M-1)(N-1)} \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} \sqrt{(I(i+1, j) - I(i, j))^2 + (I(i, j+1) - I(i, j))^2} \quad (3)$$

where  $M$  and  $N$  are the fused image's width and height and  $I(x, y)$  is the pixel value of the image at that spot.

- (ii) *Correlation Coefficient (CC).* Correlation coefficient reflects the degree of correlation among the IR image and the visible image as well as the fused image. We calculated the correlation coefficients  $CC(I, F)$  and  $CC(V, F)$  for the infrared and visible images and the fused image, respectively, and finally obtained the overall correlation coefficient, which is defined by the following equation:

$$CC(I, V, F) = \frac{1}{2} \left( \frac{\sum_{i=1}^M \sum_{j=1}^N (I_{i,j} - \bar{I})(F_{i,j} - \bar{F})}{\sqrt{(\sum_{i=1}^M \sum_{j=1}^N (I_{i,j} - \bar{I})^2)(\sum_{i=1}^M \sum_{j=1}^N (F_{i,j} - \bar{F})^2)}} + \frac{\sum_{i=1}^M \sum_{j=1}^N (V_{i,j} - \bar{V})(F_{i,j} - \bar{F})}{\sqrt{(\sum_{i=1}^M \sum_{j=1}^N (V_{i,j} - \bar{V})^2)(\sum_{i=1}^M \sum_{j=1}^N (F_{i,j} - \bar{F})^2)}} \right), \quad (4)$$

where  $I$  and  $V$  represent the infrared image and the visible image;  $F$  represents the fused image;  $I(i, j)$ ,

$V(i, j)$ , and  $F(i, j)$  are the pixels corresponding to the pixel value of the pixel point; and  $\bar{I}$ ,  $\bar{V}$ , and  $\bar{F}$  are the mean values.

- (iii) *Peak Signal-to-Noise Ratio (PSNR).* This assessment measures whether the image is distorted or not. Its value is the ratio of valid information to noisy information in the image. Its formula is as follows:

$$PSNR(I, V, F) = \frac{1}{2} \left( 10 \log_{10} \left( \frac{L^2}{MSE(I, F)} \right) + 10 \log_{10} \left( \frac{L^2}{MSE(V, F)} \right) \right). \quad (5)$$

MSE represents the mean squared error,  $MSE(x, y) = (1/mn) \sum_{i=0}^m \sum_{j=0}^n \|x(i, j) - y(i, j)\|^2$ , and  $x(x, j)$  and  $y(i, j)$  are the pixels at the corresponding positions. When the peak signal-to-noise ratio is higher, the difference between the fused image and the original image is more minor.

- (iv) The information entropy (EN) can represent the average amount of information in an image, a metric does not need to take into account the input

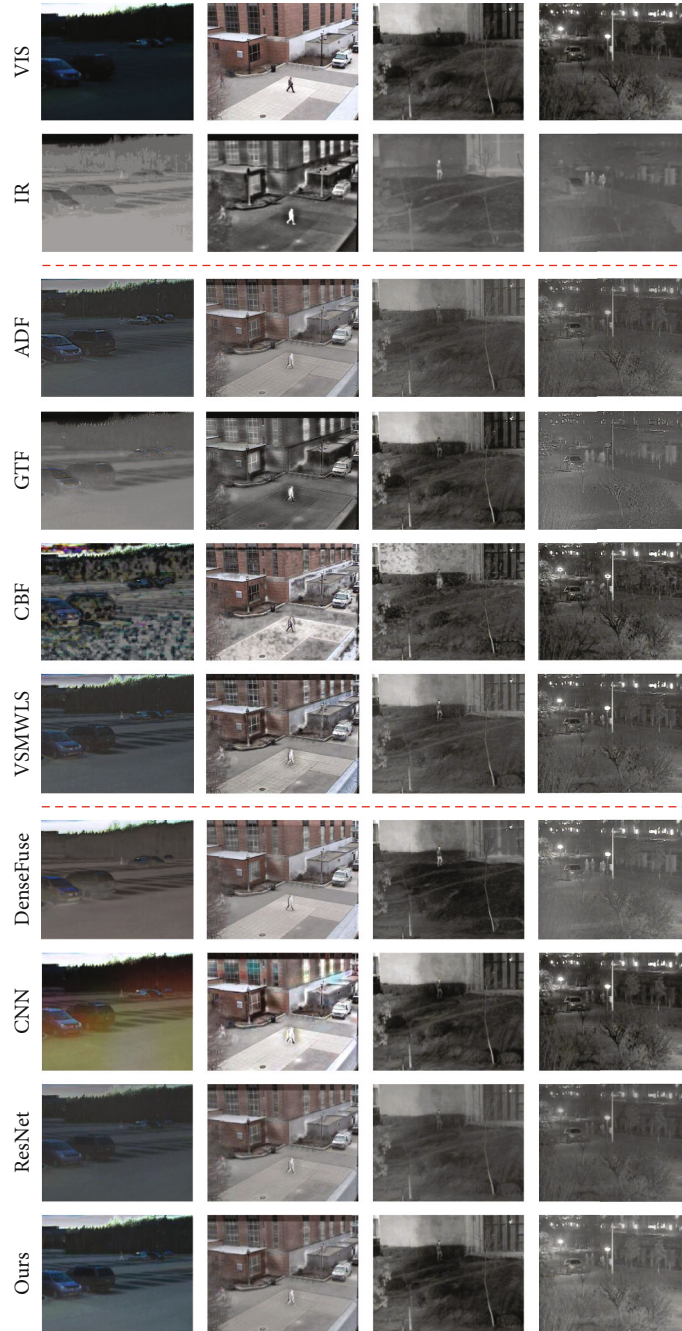


FIGURE 5: Visual fusion results of images from different scenes in the VIFB dataset. From top to bottom: infrared images, visible images, ADF, GTF, CBF, VSMWLS, DenseFuse, CNN, fusion results from ResNet, and our proposed method; the red dashed line divides the images into three parts: original images, fused images from the general method, and fused images based on the deep learning method.

image, is determined only from the fused image and is defined by the following equation:

$$EN(x) = \sum_{i=0}^{l-1} p(x_i) \log_b p(x_i), \quad (6)$$

where  $p(x_i)$  is the percentage of pixels within the grayscale image  $x$  with grayscale  $i$  and  $l$  is taken to be 256 and is the grayscale level; this equation is a

256 element entropy function; each element can be obtained with equal probability of occurrence as the maximum value; when the value of EN is larger, it means that there is more information in the image.

- (v) *Structural Similarity of Images (SSIM)*. The structural similarity of an image can be measured in terms of luminance, contrast, and structure, where the mean, standard deviation, and covariance are

used as estimates of the illumination, contrast, and structural similarity phases, which given by the following formula:

$$\text{SSIM}(I, V, F) = \frac{1}{2} \left( \frac{(2\mu_I\mu_F + c_1)(2\sigma_{IF} + c_2)}{(\mu_I^2 + \mu_F^2 + c_1)(\sigma_I^2 + \sigma_F^2 + c_2)} + \frac{(2\mu_V\mu_F + c_1)(2\sigma_{VF} + c_2)}{(\mu_V^2 + \mu_F^2 + c_1)(\sigma_V^2 + \sigma_F^2 + c_2)} \right), \quad (7)$$

where  $\mu_I, \mu_V$ , and  $\mu_F$  are the image mean;  $\sigma_I, \sigma_V$ , and  $\sigma_F$  are the standard deviation;  $\sigma_{IF}$  and  $\sigma_{VF}$  are the covariance; and  $c_1 = (k_1L^2)$  and  $c_2 = (k_2L^2)$ , where  $k_1=0.01, k_2=0.03$ , and  $L = 255$ .

- (vi) *Mutual Information (MI)*. Mutual Information measures the dependence between two domain variables. It measures the similarity between the fused image and the source image based on the amount of information retained by the combined image in the source image and is calculated as follows:

$$\text{MI}(I, V, F) = \sum_{i,j} p_{IF}(i, j) \log_2 \frac{p_{IF}(i, j)}{p_I(i)p_F(j)} + \sum_{i,j} p_{VF}(i, j) \log_2 \frac{p_{VF}(i, j)}{p_V(i)p_F(j)}. \quad (8)$$

- (vii) *Image Similarity Metric Based on Edge Information (Q\_abf)*. Xydeas et al. [34] argue that image quality is closely related to the integrity and sharpness of the edges and that the similarity between the fused image and the source image is measured from the edge perspective

- (viii) *Feature Mutual Information (FMI)*. FMI measures the quality of an image by calculating the mutual information of image features, and a higher value of FMI indicates better fusion quality:

$$\text{FMI}(I, V, F) = \frac{1}{2}(T(I; F) + T(V; F)), \quad (9)$$

where  $T(I; F) = (2/n) \sum_{i=1}^n (T_i(I; F)/(H_i(I) + H_i(F)))$  and  $T(V; F) = (2/n) \sum_{i=1}^n (T_i(V; F)/(H_i(V) + H_i(F)))$ , where  $H_i(I), H_i(V)$ , and  $H_i(F)$  are the entropy of the corresponding windows from the three images;  $n = M \times N$ ,  $n$  is the size of the image, and a more significant value of FMI indicates better image fusion performance. In the paper, we will calculate the pixel feature mutual information (FMI\_pix) and discrete cosine feature mutual information (FMI\_dct) and wavelet feature mutual information (FMI\_w) to evaluate our fusion performance.

The results of our quantitative analysis are shown in Figure 6, where the values are the average values of the differ-

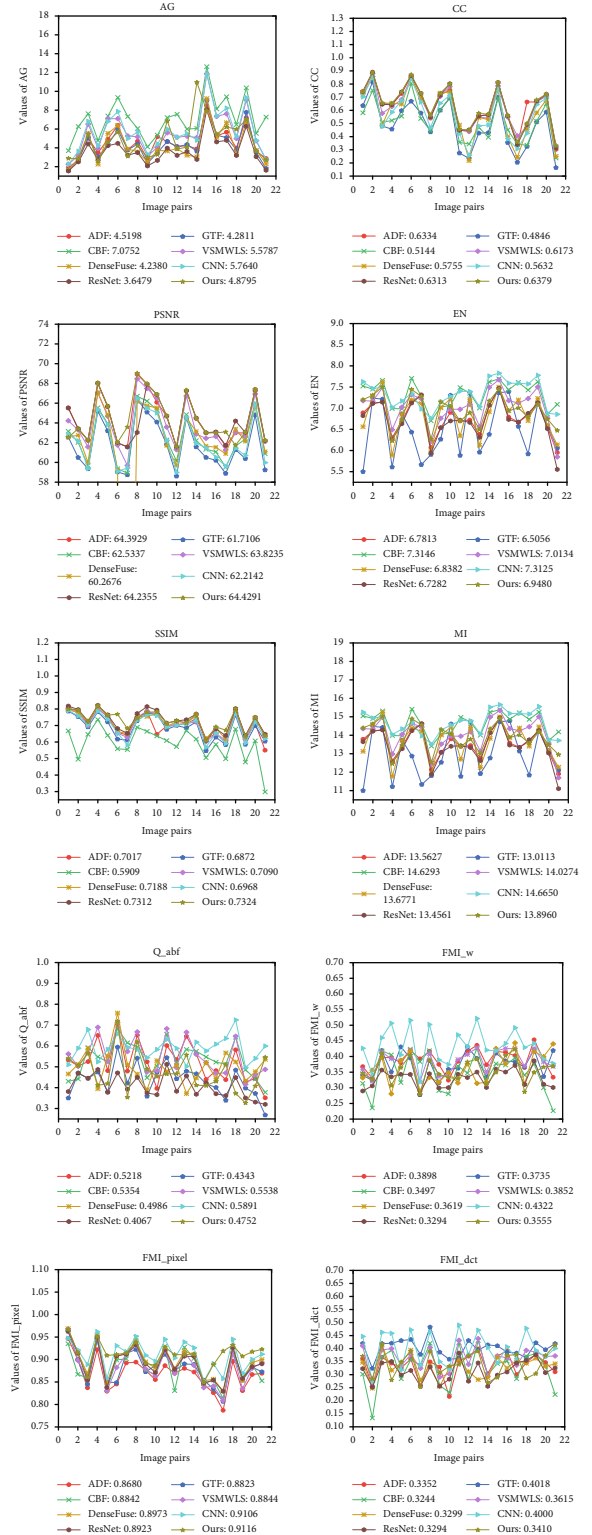


FIGURE 6: Comparative results of quantitative analysis of different fusion algorithms in ten fusion metrics.

ent evaluation metrics for the 21 pairs of images by the other algorithms. Overall, image fusion methods based on general methods achieved the best values on 3 metrics and fusion methods based on deep learning achieved the best values on 7 metrics. In the general method, ADF, CBF, and GTF

TABLE 1: Average runtime comparison of different methods on 21 testing image pairs.

Method	ADF	GTF	CBF	VSMWLS
Running time	1.32	0.37	20.58	3.42
Method	DenseFuse	CNN	ResNet	Ours
Running time	9.85	33.25	4.53	0.95

achieved the best values for AG, EN, and FMI\_dct, respectively. In the deep learning-based approach, CNN obtained the best values for MI, Qabf, and FMI\_w. The fused images generated by our method achieved the best values on four metrics: CC, SSIM, PSNR, and FMI\_pixel.

The average runtime of the 8 methods on the 21 testing image pairs is also reported in Table 1. It can be seen that the running times of the image fusion methods vary considerably. In our comparison method, our method is the fastest deep learning-based method; although the GPU is used to perform the computation, it still took an average of 0.95 seconds to fuse an image pair.

## 5. Conclusions

This proposes a novel and effective deep learning structure for wireless applications to implement the fusion of infrared and visible images. Our fusion structure consists of three main components: a feature extraction network, a feature map fusion network, and a reconstruction network. The feature output extracted by the feature extraction network of each convolutional layer will be fed to the next layer together with the previous layer output, and the original image is also involved in the reconstruction of the image, thus effectively reducing the loss of feature information. The images we obtain contain rich texture details and more background detail information, which can better balance the modality of infrared and visible images, have a better visual experience, and achieve a more natural fusion.

## Data Availability

We use the open dataset MS-COCO that is publicly available on <https://cocodataset.org/#home>.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] J. Ma, Y. Ma, and C. Li, "Infrared and visible image fusion methods and applications: a survey," *Information Fusion*, vol. 45, pp. 153–178, 2019.
- [2] Y. Miao, C. Chen, L. Pan, Q. L. Han, J. Zhang, and Y. Xiang, "Machine learning based cyber attacks targeting on controlled information: a survey," *ACM Computing Survey*, 2021.
- [3] Y. Liu, X. Chen, R. K. Ward, and Z. Jane Wang, "Image fusion with convolutional sparse representation," *IEEE Signal Processing Letters*, vol. 23, no. 12, pp. 1882–1886, 2016.
- [4] Y. Liu, X. Chen, H. Peng, and Z. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Information Fusion*, vol. 36, pp. 191–207, 2017.
- [5] Y. Liu, X. Chen, J. Cheng, H. Peng, and Z. Wang, "Infrared and visible image fusion with convolutional neural networks," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 16, no. 3, pp. 1850018–1850018:20, 2018.
- [6] H. Li and X.-J. Wu, "Densefuse: a fusion approach to infrared and visible images," *IEEE Transactions on Image Processing*, vol. 28, no. 5, pp. 2614–2623, 2018.
- [7] J. Qiu, J. Zhang, W. Luo, L. Pan, S. Nepal, and Y. Xiang, "A survey of android malware detection with deep neural models," *ACM Computing Survey*, vol. 53, no. 6, pp. 1–36, 2021.
- [8] M. Wang, T. Zhu, T. Zhang, J. Zhang, S. Yu, and W. Zhou, "Security and privacy in 6G networks: new areas and new challenges," *Digital Communications and Networks*, vol. 6, no. 3, pp. 281–291, 2020.
- [9] G. Lin, S. Wen, Q.-L. Han, J. Zhang, and Y. Xiang, "Software vulnerability detection using deep neural networks: a survey," *Proceedings of the IEEE*, vol. 108, no. 10, pp. 1825–1848, 2020.
- [10] R. Coulter, Q. Han, L. Pan, J. Zhang, and Y. Xiang, "Data-driven cyber security in perspective—intelligent traffic analysis," *IEEE Transactions on Cybernetics*, vol. 50, no. 7, pp. 3081–3093, 2020.
- [11] D. P. Bavirisetti and R. Dhuli, "Two-scale image fusion of visible and infrared images using saliency detection," *Infrared Physics & Technology*, vol. 76, pp. 52–64, 2016.
- [12] Y. Liu, S. Liu, and Z. Wang, "A general framework for image fusion based on multi-scale transform and sparse representation," *Information Fusion*, vol. 24, pp. 147–164, 2015.
- [13] G. Liu, Z. Jing, S. Sun, J. Li, Z. Li, and H. Leung, "Image fusion based on expectation maximization algorithm and steerable pyramid," *Chinese Optics Letters*, vol. 2, no. 7, pp. 386–389, 2004.
- [14] Y. Zou, X. Liang, and T. Wang, "Visible and infrared image fusion using the lifting wavelet," *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 11, no. 11, pp. 6290–6295, 2013.
- [15] F. Meng, M. Song, B. Guo, R. Shi, and D. Shan, "Image fusion based on object region detection and non-subsampled contourlet transform," *Computers & Electrical Engineering*, vol. 62, pp. 375–383, 2017.
- [16] J.-j. Zong and T.-s. Qiu, "Medical image fusion based on sparse representation of classified image patches," *Biomedical Signal Processing and Control*, vol. 34, pp. 195–205, 2017.
- [17] R. Gao, S. A. Vorobyov, and H. Zhao, "Image fusion with cosparsity analysis operator," *IEEE Signal Processing Letters*, vol. 24, no. 7, pp. 943–947, 2017.
- [18] Q. Zhang, Y. Fu, H. Li, and J. Zou, "Dictionary learning method for joint sparse representation-based image fusion," *Optical Engineering*, vol. 52, no. 5, article 057006, 2013.
- [19] N. Sun, J. Zhang, P. Rimba, S. Gao, L. Y. Zhang, and Y. Xiang, "Data-driven cybersecurity incident prediction: a survey," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 2, pp. 1744–1772, 2019.
- [20] Y. Liu, X. Chen, J. Cheng, and H. Peng, "A medical image fusion method based on convolutional neural networks," in *2017 20th International Conference on Information Fusion (Fusion)*, pp. 1–7, Xi'an, China, July 2017.

- [21] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neural networks," *Remote Sensing*, vol. 8, no. 7, p. 594, 2016.
- [22] K. R. Prabhakar, V. S. Srikanth, and R. V. Babu, "DeepFuse: a deep unsupervised approach for exposure fusion with extreme exposure image pairs," in *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 4724–4732, Venice, Italy, 2017.
- [23] H. Xu, J. Ma, Z. Le, J. Jiang, and X. Guo, "FusionDN: A Unified Densely Connected Network for Image Fusion," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 7, pp. 12484–12491, 2020.
- [24] H. T. Mustafa, M. Zareapoor, and J. Yang, "MLDNet: multi-level dense network for multi-focus image fusion," *Signal Processing: Image Communication*, vol. 85, pp. 110923–115965, 2020.
- [25] J. Ma, W. Yu, P. Liang, C. Li, and J. Jiang, "FusionGAN: a generative adversarial network for infrared and visible image fusion," *Information Fusion*, vol. 48, pp. 11–26, 2019.
- [26] J. Ma, H. Xu, J. Jiang, X. Mei, and X. P. Zhang, "DDcGAN: a dual-discriminator conditional generative adversarial network for multi-resolution image fusion," *IEEE Transactions on Image Processing*, vol. 29, pp. 4980–4995, 2020.
- [27] T.-Y. Lin, M. Maire, S. Belongie et al., "Microsoft coco: common objects in context," in *Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science*, vol. 8693, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., pp. 740–755, Springer, Cham, 2014.
- [28] X. Zhang, P. Ye, and G. Xiao, "VIFB: a visible and infrared image fusion benchmark," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 104–105, Seattle, WA, USA, June 2020.
- [29] D. P. Bavirisetti and R. Dhuli, "Fusion of infrared and visible sensor images based on anisotropic diffusion and Karhunen-Loeve transform," *IEEE Sensors Journal*, vol. 16, no. 1, pp. 203–209, 2015.
- [30] J. Ma, C. Chen, C. Li, and J. Huang, "Infrared and visible image fusion via gradient transfer and total variation minimization," *Information Fusion*, vol. 31, pp. 100–109, 2016.
- [31] B. K. S. Kumar, "Image fusion based on pixel significance using cross bilateral filter," *Signal, Image and Video Processing*, vol. 9, no. 5, pp. 1193–1204, 2015.
- [32] J. Ma, Z. Zhou, B. Wang, and H. Zong, "Infrared and visible image fusion based on visual saliency map and weighted least square optimization," *Infrared Physics & Technology*, vol. 82, pp. 8–17, 2017.
- [33] H. Li, X.-j. Wu, and T. S. Durrani, "Infrared and visible image fusion with ResNet and zero-phase component analysis," *Infrared Physics & Technology*, vol. 102, article 103039, 2019.
- [34] G. Cui, H. Feng, Z. Xu, Q. Li, and Y. Chen, "Detail preserved fusion of visible and infrared images using regional saliency extraction and multi-scale image decomposition," *Optics Communications*, vol. 341, pp. 199–209, 2015.
- [35] P. Jagalingam and A. V. Hegde, "A review of quality metrics for fused image," *Aquatic Procedia*, vol. 4, pp. 133–142, 2015.
- [36] J. W. Roberts, J. A. Van Aardt, and F. B. Ahmed, "Assessment of image fusion procedures using entropy, image quality, and multispectral classification," *Journal of Applied Remote Sensing*, vol. 2, no. 1, article 023522, 2008.
- [37] L. Liu, O. de Vel, Q. Han, J. Zhang, and Y. Xiang, "Detecting and preventing cyber insider threats: a survey," *IEEE Communications Surveys and Tutorials*, vol. 20, no. 2, pp. 1397–1417, 2018.
- [38] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electronics Letters*, vol. 38, no. 7, pp. 313–315, 2002.
- [39] C. S. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308–309, 2000.
- [40] M. Haghghat and M. A. Razian, "Fast-FMI: non-reference image fusion metric," in *2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT)*, pp. 1–3, Astana, Kazakhstan, October 2014.