Research Article

Blockchain-Based Access Control Model to Preserve Privacy for Personal Health Record Systems

Thein Than Thwin and Sangsuree Vasupongayya

Department of Computer Engineering, Faculty of Engineering, Prince of Songkla University, P.O. Box 2, Kho Hong, Hat Yai, Songkhla 90112, Thailand

Correspondence should be addressed to Sangsuree Vasupongayya; vsangsur@coe.psu.ac.th

Received 3 February 2019; Accepted 9 June 2019; Published 25 June 2019

Academic Editor: Prosanta Gope

Copyright © 2019 Thein Than Thwin and Sangsuree Vasupongayya. 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.

Personal health record system (PHR system) stores health-related information of an individual. PHR system allows the data owner to manage and share his/her data with selected individuals. The originality or tamper resistance feature is crucial for PHR system because of the irreversible consequence of incorrect information. Blockchain technology becomes a potential solution due to its immutability and irreversibility properties. Unfortunately, some technical impediments such as limited storage, privacy concern, consent irreversibility, inefficient performance, and energy consumption exist. This work aims to handle these blockchain drawbacks and propose a blockchain-based PHR model. The proposed model is built using the blockchain technology to support a tamper resistance feature. Proxy reencryption and other cryptographic techniques are employed to preserve privacy. Features of the proposed model include fine-grained and flexible access control, revocability of consent, auditability, and tamper resistance. A detailed security analysis shows that the proposed model is provably secure for privacy and tamper resistance. The performance analysis shows that the proposed model achieves a better overall performance compared with the existing approach in the literature. Thus the proposed model is more suitable for the PHR system usage.

1. Introduction

The personal health record system could be seen as a promising solution for preventive care of the PHR owners. PHR system enables the exchange of information with healthcare provider and it can help to foresee the health issues. Personal health record (PHR) stores the health-related personal data and usually contains highly sensitive information. Some incorrect modification or alteration of any PHR data may cause an irreversible harmful consequence. Thus, privacy becomes a key ingredient for any PHR system. In particular, a tamper resistance property is the most important feature for PHR system. PHR system would significantly provide the high-quality preventive personal healthcare if the lifelong health-related information of an individual can be securely captured and stored on tamper resistant storage. Immutability, cryptographic verifiability, and backup characteristics of blockchain can be an effective tamper resistant storage mechanism for PHR system.

Even though blockchain technology possesses many good properties such as immutability and irreversibility [1, 2], blockchain technology also contains some potential drawbacks to using in a PHR system development. The irreversibility property of blockchain becomes a barrier to a consent revocation feature—allowing the users to eliminate a permission of a certain action on the data for specified individuals. The transparent property of blockchain—allowing all participants on the network to view all data—can cause a confidentiality issue. The limited storage of blockchain becomes an availability issue for the explosive growth of the diverse medical related data. Although the private blockchain such as Hyperledger Fabric [3] can regulate the participation in its blockchain network [4], the PHR system still requires allowing only some certain members to access a specific part of the PHR system. As a result, a privacy leakage is still an issue. Other issues include performance and energy consumption of blockchain underlying mechanisms. The original blockchain technology in Bitcoin shows that Bitcoin's
minimum block creation time is 10 minutes and the maximum performance is 7 transactions per second and consumes a large amount of computational power and electricity for block creating process [5]. Thus, how to use blockchain technology as an underlying system for PHR system is still an important consideration.

In the premise of safeguarding the privacy, blockchain technology gains more attention from the healthcare community. As a result, many blockchain companies and other enterprises such as Factom [6], HealthNautica [7], Gem [8], and Capital One [9] are collaboratively trying to use blockchain technology for storing their medical data. Moreover, many healthcare data management system applications, such as [10–27], are emerging. Each application offers different solutions for different conditions. Most of them still suffer from some issues of blockchain. For instance, the blockchain-based access control layer is added to an existing database to preserve the privacy of the system [10–13]. These systems can support immutable access log and user-centric access control. However, a consent revocation and the confidentiality of data cannot be supported. Some systems [14–23] use attribute-based encryption scheme to provide an access control feature. However, the attribute-based encryption causes the growing operation time linearly with the number of unrevoked users, and the irreversibility property of blockchain becomes a barrier to revocation of consent. Systems in [24–27] tried to store the data on blockchain. However, the medical data can be large and the blockchain is not optimized for storing a large size data. The main objective of this work is to propose a blockchain-based PHR model that provides fine-grained access control, guarantees the tamper resistance, supports the verification of the integrity of data conveniently, and ensures the privacy.

In our previous work [28], the issues of using blockchain in PHR system are identified and a general secret data sharing scheme for PHR system is proposed as a potential solution. However, the previous work presented only the preliminary investigation of using blockchain technology for PHR system. The workflow for the model, the analysis for security, and analysis for usability are still lacking. In this work, blockchain technology will be used to support nonreputation, accountability, and tamper resistance properties; the proxy reencryption technique will be used to propose an access control mechanism that can support fine-grained access control and consent revocation properties; and the cloud storage will be used to support an availability property. The detailed model for access control of blockchain-based PHR system is proposed to show the workflow by using AFGH proxy reencryption algorithm [29]. The PHR data will be encrypted with the proxy cryptography technique and stored on a cloud storage. The related metadata will be stored on a private blockchain. In particular, the characteristics of PHR data will be stored forever in the blockchain. As a result, all data tampering will be detected and validated. The cryptographic authentication technique and an access control list are used to verify the users in order to support the accountability and revocability features. At the same time, the PHR owner is supported with a revocable fine-grained access control feature. The prototype is implemented with the Hyperledger blockchain as an underlying system. To ensure the protection of privacy, the security analysis is performed with four cases of adversary attempts including a tampering attack, a collusion attack, a replay attack, and a malicious access attack. The usability of the proposed model is ensured by comparing it with the existing system [14].

The rest of the paper is organized as follows. Some related works and the background information on blockchain technology, proxy reencryption, and personal health record system on a cloud environment are presented in Section 2. The proposed model is described in Section 3. The security and privacy of the proposed model are analyzed in Section 4, and the performance of the proposed model is analyzed in Section 5. The discussion is presented in Section 6. Finally, the paper is concluded in Section 7.

2. Background and Related Works

Similar to Hao Wang et al. [14], the healthcare data of our proposed system is encrypted and stored on a cloud storage for availability, and the metadata is stored on blockchain for tamper resistance. However, the access control mechanism based on an attribute-based encryption as used in [14] has some drawbacks because the change or modification to an access policy is difficult since the modification of an access policy requires an extra computational cost to perform an attribute revocation and a reencryption process on the data. The append-only storage of blockchain also disallows the modification of the attribute-based encrypted data to update the access control. An additional process must be created in order to work around it. In our model, however, an access control list, which is controlled by a proxy reencryption scheme, is used in order to support a consent revocation.

The blockchain-based data sharing systems, in which the blockchain-based access control layer is added to the existing databases of the providers, are utilized in [10]–[13]. These systems stored only the metadata to describe the real data and its permissions on the blockchain. These systems can also keep an immutable access log and support a user-centric access control. However, a consent revocation and the confidentiality of the user data may be needed to support PHR system purpose. Moreover, some of these systems involved transaction fees, and the users are required to be involve in mining activities. Similar to these systems, the blockchain in our model is not used for storing the entire healthcare data. The blockchain in our model is also used to store the metadata while the access control model is constructed to support the consent revocation and confidentiality. To reduce the mining problem, the private blockchain (Hyperledger) is used in our model.

The attribute-based encryption scheme and the semitrusted servers are used to store the PHR data in [15–23]. Under these systems, an access control is created by directly encrypting the real PHR data. The real PHR data may contain several large files and the attribute-based encryption can cause a growing computational cost linearly with the number of unrevoked users. Nevertheless, the functionality is limited because the encrypted data cannot support a searching feature. Therefore, such system will need a special
mechanism such as a searchable encryption scheme which is computationally expensive [30]. To ensure the validity of electrical health data, blockchain is used as a storage technology and an attribute-based signature scheme is proposed by using multiple authorities in [24]. Under such system, all healthcare information of a user is grouped in a single block. Similarly, a blockchain-based PHR data preservation system is also proposed in [25]. The idea that the medical data are broken into pieces and stored on multiple blocks is proposed in [26, 27]. However, the medical data can include large data such as images while blockchain might not be optimized for storing massive data. Even though the data can be broken into many smaller pieces in order to store on blockchain, such action may cause some performance issues when encrypting the data.

Blockchain is also used in an access control for a transaction processing in [31–33]. These systems are not intended for a data storage purpose. Blockchain and other mechanisms are restricted to authorizing some people on sensitive transactions and achieving the undeniability on such transactions. In this work, blockchain will be used to propose a privacy-preserving personal health record system to support non-repudiation, accountability, and tamper resistance properties. An access control mechanism will be proposed by using proxy reencryption technique to support fine-grained access control and consent revocation features while the cloud storage will be used to support an availability feature.

2.1. Blockchain Technology. Blockchain technology has become popular along with Bitcoin—a cryptocurrency [34]. Blockchain is a system that is composed of nodes, communicating with each other through a protocol. The communication protocol is defined by blockchain developers. A node can be a physical machine or a virtual machine. The IP address is used to identify the node in the blockchain system. The public key is used as a reference of the user in the blockchain system. The private key is used for a cryptographic signing process on all messages. As a result, each user can log in from any node in the system. The data stored on blockchain is replicated on every full node and a synchronization must be performed [35]. Information on blockchain is digitally signed to guarantee authenticity and accuracy properties. Blockchain technology can support an immutable storage and a fraud protection property. Figure 1 represents a general idea of how blockchain technology works to support the tasks required by Bitcoin.

According to Figure 1, blockchain system stores its transactional data into a specific structure called “block.” The blocks under the blockchain are cryptographically linked together to form a chain of blocks. Each block inside the blockchain stores a hash code of the previous block. Thus, the chain of blocks is grouped or linked in a chronological order. As a result, the data, stored on the blockchain, cannot be altered without a notification by all nodes inside the system. With this property, the data, stored on the blockchain, provide a tamper resistance feature. Thus, this property of blockchain is suitable for storing some sensitive data such as medical data [36].

Even though the blockchain technology can offer opportunities for the PHR system, the blockchain technology also contains some issues. The primary hurdle is the storage cost. The storage cost is very high because the data is stored after a verification process; the data must be replicated; the synchronization must be done on every full node inside

![Blockchain and block structure.](image-url)
the blockchain network. The append-only nature of any blockchain storage is suitable for storing the data to support a tamper resistance feature. However, storing an access permission of the data may cause difficulty in providing a consent revocation feature. Blockchain transparently shows all transactions of each public key [37]. Thus, the public nature of the blockchain means the private data flow through every node inside the network. Furthermore, the processing speed of blockchain is slower than that of the traditional database because the blockchain has to perform some extra tasks such as signature verification, a consensus mechanism, and redundancy. The block creating process of a public blockchain consumes a large amount of computational power and a large amount of electricity. Thus, the model proposed in this work must consider a way to handle these limitations in order to employ the blockchain technology for a PHR system.

2.2. Proxy Reencryption Scheme. The proxy reencryption scheme is an asymmetric cryptosystem that enables its users to share their decryption capabilities with others [38]. Under the proxy reencryption scheme, the ciphertext—encrypted with the user public key—can be reencrypted in such a way that another user can decrypt it by using his/her private key although the ciphertext is not originally encrypted with his/her public key. The data will not be fully decrypted during the transmission. Thus, the scheme will be a useful method to create a secure data sharing scheme. To share the data under the proxy reencryption scheme, the data owner must send the reencryption key to the proxy. However, the proxy will not be able to gain any information on the original data from the reencryption key. The reencryption key is generated from the combination of the owner’s secret key and the intended-user’s reencryption key. There encryption key is generated from the proxy reencryption scheme, the data owner must send the reencryption key to the proxy. However, the proxy will not be able to gain any information on the original data from the proxy reencryption scheme. The data owner must send the reencryption key to the proxy. However, the proxy will not be able to gain any information on the original data from the reencryption key. The reencryption key is generated from the combination of the owner’s secret key and the intended-user’s public key. Thus, the proxy reencryption scheme is flexible to create an access control management system in our proposed model.

The proxy reencryption scheme is introduced by the work of Blaze, Bleumer, and Strauss (BBS) [39]; however, the BBS scheme contains some weakness such as being bidirectional and prone to collusion attacks. To solve such issue, Ateniese, Fu, Green, and Hohenberger (AFGH) proposed an improved proxy reencryption scheme [29] and remedied the weaknesses of BBS. The following properties are supported in AFGH:

(i) Unidirectional (reencryption from $A \rightarrow B$ does not allow a reencryption from $B \rightarrow A$).

(ii) Noninteractive (no trusted third party or interaction is needed to generate a reencryption key).

(iii) Original-access (the delegator can decrypt reencrypted cipher-texts).

(iv) Key optimal (the size of the secret remains constant, regardless of how many delegations is accepted).

(v) Nontransitive (the proxy alone cannot redelegate the decryption rights).

These properties are suitable for our proposed model. Thus, the AFGH reencryption scheme will be used to develop the access control mechanism in our model.

2.3. Personal Health Record System on a Cloud Environment. Deploying the PHR system on a cloud environment offers several opportunities such as ubiquitous accessibility, elastic computation resource, high degree of fault tolerance, and interoperability with other systems [40, 41]. According to HIPPA [42], the cloud service providers are considered noncovered entities. Thus, the cloud service provider has no obligation to ensure the confidentiality and proper access to the consumers PHR [43]. Consequently, the privacy concern becomes one of the most important issues to adapt the PHR system to a cloud environment.

PHR data is managed and controlled by the PHR owner, unlike the other digital medical record systems [44, 45]. The PHR owners can share their health data selectively with others while keeping some parts private. The cloud environment allows accessing the PHR data anytime and anywhere. The cloud can also support the PHR system to prepare for medical appointments and to maintain more complete picture of personal health for sharing, collaborating, and engaging.

The data sources of the PHR system range from the data produced by some devices used by the PHR owner to the health data such as Electronic Health Record (EHRs) data. On the other hand, the cloud may have business interest in analyzing the PHRs, and it may also have malicious employees or the cloud can even be hacked. As a result, PHR system will interact with various types of users and the employed access control mechanism is needed to support accountability (traceability of which user performs what action within the system) and consent revocation (ability to support the PHR owners to eliminate their consent or permission of a certain action on the PHR for specified individuals) features. Thus, the PHR system must provide a tamper resistance feature and protect the PHR owner privacy. In our model, the underlying cloud infrastructure of the PHR system is defined to be semitrusted and the blockchain and other cryptographic mechanisms are used as added security.

2.4. Cryptographic Primitives. The main workflow of our model is based on pairing-based cryptography and bilinear maps which can provide a cyclic group [46]. Therefore, the properties of these cryptographic primitives which are used in our model will be presented.

A group $G$ which is a set of elements with an abstract binary operation, $(\cdot)$, $(G, \cdot)$, can satisfy the following axioms:

(i) Closure which means that the result of applying the operation on any two elements in the set is another element in the set $(a, b \in G \text{ for all } a, b \in G)$.

(ii) Associativity which means that it does not matter in which order the operation on more than two elements is applied ($(a, b, c = a, (b, c) \text{ for all } a, b, c \in G)$).

(iii) Existence of identity element which means that $e \in G$ such that $a \cdot e = a = e \cdot a$ for all $a \in G$.

(iv) Existence of inverse element which means that $a^{-1} \in G$, such that $a \cdot a^{-1} = e = a^{-1} \cdot a$ for all $a \in G$.

The abstract binary operation of the group can be mapped, addition or multiplication.
For $G_1$ and $G_T$, which are cyclic groups of prime order $q$, and $g$, which is a generator of $G_1$, the bilinear map $e : G_1 \times G_2 \rightarrow G_T$ satisfies the following axioms:

(i) **Bilinearity** ($e(g_1^a, g_2^b) = e(g_1, g_2)^{ab}$ for all $g_1 \in G_1$, $g_2 \in G_2$, $a, b \in \mathbb{Z}_q$).

(ii) **Nondegeneracy** (if $G_1 = \langle g_1 \rangle$, $G_2 = \langle g_2 \rangle$, then $G_T = \langle e(g_1, g_2) \rangle$).

(iii) **Computability** ($e$ can be efficiently computed).

### 3. The Proposed Model

The proposed model will be discussed in this section. Firstly, the overall data flow and the system component will be introduced. Next, the system operation will be described. Finally, the data format and the access control protocol will be discussed.

#### 3.1. System Architecture

The overall architecture of the proposed model is shown in Figure 2. In our model, the real PHR data will be encrypted using the public key (master key) of the PHR owner and stored on a cloud storage to ensure the confidentiality. The PHR will be shared via a proxy reencryption process. Thus, the reencryption keys and other information needed for an authentication process will be stored on a proxy called the gateway server. The metadata of the PHR will be stored on the private blockchain to support search and tamper resistance features. The PHR will be accessed by the PHR owner or others such as healthcare providers, e.g., doctors, nurses. Five main entities will be included in our model as follows:

(i) The PHR owner (PO) is an entity who owns the PHR data and wishes to access or store their PHR data. PO has a full control over his/her PHR data. PO must define an access control policy on his/her PHR data and PO can allow or disallow some access permissions on his/her PHR data to others at will. Therefore, PO has to generate the reencryption keys and the metadata for his/her PHR. PO can also allow some users to add more data on his/her behalf and PO can recontrol the access to the newly added PHR data.

(ii) The user (U) is an entity who requests an access to the PHR data with the permission from the corresponding PO. Typical U can be from several sections such as doctors, nurses, etc.
as healthcare providers (e.g., doctors, nurses), health insurance providers, and caregivers. U can search and get the metadata via the blockchain and can later request the PHR data from the gateway server. U with a delegated authority can create/upload the new PHR data into the system.

(iii) The gateway server (GS) is responsible for verifying the authenticity of all actions inside the system. The actions include reencrypting the PHR data, storing the metadata, and accessing the log on the chain. GS is also responsible for administering the private blockchain. All communications between GS and other entities will be done via an SSL/TLS secure channel. In this work, GS is considered a semitrusted server—the server obeys the procedure defined in the work but it is curious to know the data.

(iv) Cloud storage (CS) is responsible for storing the actual encrypted PHR data. CS is also considered a semitrusted server.

(v) The private blockchain (BC) is responsible for storing the metadata and the access log of the system. BC can be accessed by a predefined group of users.

The detailed workflow of the model will be presented in the next section.

3.2. Workflow of the Proposed Model. The workflow of the proposed model will be presented in this section. Firstly, the system setup phase will be introduced to explain how the initial agreement of the system is set up. Next, the detailed procedure of accessing the PHR data in the system will be described. Finally, how to revoke the access right on the PHR data will be discussed.

3.2.1. System Setup. In the setup phase, all users including the PHR owner must register with the gateway server to make an initial agreement required for the operations of our model. Firstly, each user generates the private/public key pair. The key pair generation is based on $G_1$ and $G_2$ which are cyclic groups of a prime order $q$ and bilinear map $e : G_1 \times G_1 \rightarrow G_2$. The random generators $g \in G_1$ and $(g, g) \in G_2$ are used as the system parameters. A random number $a \in \mathbb{Z}_q$ is randomly selected and used as the secret key ($SK = a$). Then, the public key is computed as $PK = g^a$. Next, the secret key (SK) is locally kept as the secret and the public key (PK) and the identity information is sent to the gateway server for a registering process. The gateway server stores the registered information and issues an access right to the private blockchain. This action is considered as the end of the registration process.

3.2.2. Storing a PHR. The PHR data storing process begins with creating the PHR data as shown in Figure 3. Once a new record (m) is created, the message digest (md) is calculated
to support the data integrity checking in the system. The message digest (md) will be calculated as $md = H(m)$ by using the hashing algorithm (SHA-2).

The data is encrypted with the public key (PK_A) of the PHR owner (A) to support the privacy and confidentiality properties. To encrypt the message $m \in G_2$, the ciphertext $C = (mZ^k, g^ak)$ is computed by using “k” which is a randomly selected element in $z_q$ and “g^ak” which is the public key PK_A. Then, the metadata is created to support the search feature within the PHR, and the digital signature (s) is created by signing on the message digest (md) using the PHR owner private key.

The reencryption keys are generated for each person who is allowed to access the PHR data, and the person will be included in the access list (al) for sharing purpose. Each reencryption key is produced by using the private key of the PHR owner (SK_A) and the corresponding public keys PK of the user. For instance, the PHR owner allows the user B to access the data by publishing the reencryption key $RK_{A \rightarrow B} = g^{b/a} \in G_1$ where a is the secret key of PHR owner and g^b is the public key of user B. Finally, the encrypted PHR (C), the metadata, the access list (al), the message digest (md), and the signature (s) are sent to the gateway server.

The gateway server verifies the PHR owner signature (s) for authenticity. Next, the encrypted PHR (C) will be stored on the cloud storage. The link to the encrypted data (l) is collected. Then, the gateway server assigns the data-id and maps the data-id to the link (l). All the three information items, namely, the data-id, the link (l), and the access list (al), are stored locally on the gateway server. The gateway server then creates its signature on the data-id. Finally, the metadata, the owner signature, the message digest, the data-id, and the signature of the gateway server are stored on the private blockchain.

3.2.3. Retrieving a PHR. To retrieve the PHR data, the user can get the information of the requested PHR data via the metadata from the private blockchain as shown in Figure 4. The user verifies the PHR data using the owner signature and the signature of the gateway server. If the PHR data is correct, the user appends the timestamp to the data-id and signs on it. Then, the user sends the resulting signed data-id to the gateway server in order to request the actual PHR data.

The gateway server checks the user authenticity using the user signature. Next the gateway server checks the time stamp on the request and validates the request life time. If the user is authorized, the gateway server stores the request on the private blockchain for auditing purpose. Then, the gateway server retrieves the information related to the data from the local storage using the data-id, and the requested encrypted PHR data is retrieved from the cloud storage.

The gateway server is then performing the reencryption process in order to send the encrypted PHR data to the requester. To reencrypt the original encrypted PHR of A to a ciphertext that user B (the requester) can decrypt, the gateway server must obtain the reencryption key $RK_{A \rightarrow B} = g^{b/a}$ from the access list and change the ciphertext $C_A = (mZ^k, g^ak)$ into $C_B = (mZ^k, Z^bk)$, where $Z^bk = e(g^{ak}, g^{b/a})$. 

**Figure 4:** The process for retrieving a PHR.
The newly created ciphertext is then sent to the requested user (B). User B can decrypt the ciphertext \( C = (\alpha, \beta) \) with his secret key \( \text{SK}_B = b \) by computing \( m = \alpha/e(\beta; g^{ib}) \). Moreover, user B can check the data integrity from the message digest (md), since the owner signature is created on the message digest.

3.2.4. Revoking a User. The case of revoking a user is also considered in our PHR model. The cloud storage is provided to store the actual encrypted PHR data. All users query the encrypted PHR through the gateway server. The gateway server verifies each request in terms of the authority of the requester according to the predefined access list generated by the PHR owner. After successfully verifying the request, the gateway server stores the transaction log information on the blockchain and reencrypts the PHR for the authorized users with the corresponding reencryption key. As a result, the PHR owner can revoke any access to his/her PHR data by updating the access list and maintaining the ownership on the PHR. Section 4 provides more information on how our model can resist the collusion attack. The time-consuming encryption process does not need to be performed as long as nobody performs the auditable access of the data. If gateway server violates the assumption that the gateway server follows the procedure defined in this work (semitrusted), it can be audited on blockchain. If some suspicious action of the revoked user is found on blockchain, the PHR owner must update his/her data and encrypt the data again. Thus, the encryption process will be necessary for totally untrusted gateway server.

3.3. Access Control Protocol. Under our proposed model, the access control protocol by the way of proxy reencryption reduces the requirement of the gateway server. As such, the gateway server can be viewed as a semitrusted entity in our proposed model. The actual PHR data is securely encrypted with the PHR owner public key, and the ciphertext can be accessed by a group of authorized users according to the access list as shown in Table 1(a). The access to the actual PHR data can be easily revoked by updating the access list. The delegated user can also add new PHR data and create the corresponding metadata on behalf of the PHR owner as shown in Table 1(b). The access control list is stored in the local database of the gateway server. However, the corresponding secret keys are protected because the secret key belongs to the PHR owner. The reencryption keys used by the gateway server can only be generated by the PHR owner. Moreover, the reencryption keys only allow the gateway server to reencrypt the original ciphertext for the authorized user. Thus, the gateway server cannot gain access to the actual PHR data, because the actual PHR data will never be decrypted at the gateway server.

4. Security and Privacy Analysis

In this section, the security and privacy of the proposed model are evaluated. Some security properties of the proposed model are also discussed. The privacy of the PHR owner can be achieved by controlling who will be allowed to access the PHR under what conditions. The security of the PHR data can be achieved by securing the PHR data from unauthorized disclosure, alteration, or deletion [47]. The privacy and security analysis are performed according to two assumptions as follows.

Assumption 1. Assume that \( g \) is the generator of a cyclic group \( G \) of order \( q \) and \( a \in Z_q^n \) cannot be computed from \( (g, g^a) \) with a nonnegligible probability.

Assumption 2. The servers that are used in this model are semitrusted so that the servers follow the procedure defined in this work but the servers are curious to know the data.

Case 1. The proposed PHR model is secure against a security attack such as tampering data by an adversary.

Threat model: The PHR system contains the medical data such as diagnosis results or medical records, and the adversary aims to either modifying some data or replace the original data with the new one.

Argument: The uploaded PHR data is actually encrypted and stored on a cloud server. The link to the encrypted PHR is known only by the gateway server. The adversary cannot modify the real encrypted PHR data. Even if the adversary can modify or replace the encrypted PHR data, the message digest on the blockchain will be able to detect such actions. If the adversary wants to modify the metadata on the blockchain, an extensive work is required to construct a new main chain. This situation is nearly impossible because the blockchain characteristic ensures that the stored data will be very difficult to modify or delete once confirmed.

<table>
<thead>
<tr>
<th>Access list</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) User ID:</td>
</tr>
<tr>
<td>(ii) Write:</td>
</tr>
<tr>
<td>(iii) Read:</td>
</tr>
<tr>
<td>(iv) Public Key:</td>
</tr>
<tr>
<td>(v) Re-encryption Key:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Meta-data</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) Message digest:</td>
</tr>
<tr>
<td>(ii) Creator:</td>
</tr>
<tr>
<td>(iii) Signature:</td>
</tr>
<tr>
<td>(iv) Subject:</td>
</tr>
<tr>
<td>(v) Type:</td>
</tr>
<tr>
<td>(vi) Date:</td>
</tr>
<tr>
<td>(vii) Format:</td>
</tr>
<tr>
<td>(viii) Description:</td>
</tr>
<tr>
<td>(ix) Data ID:</td>
</tr>
</tbody>
</table>
Case 2. The proposed PHR model is secure against a security attack such as a collusion between the gateway server and the adversary.

Threat model: The reencryption keys are included in the access list which is stored at the gateway server, and the gateway server can perform a reencryption process. The adversary and the gateway server may collusively try to obtain the PHR data or reencrypt the PHR data for the adversary.

Argument: Although the access control list is locally stored at the gateway server, the information on the access control list does not include the corresponding secret key. As a result, the gateway server cannot gain an access to the encrypted PHR data. Since, the secret key is in the care of the PHR owner, the reencryption keys used by the gateway server can only be generated by the PHR owner. The gateway server cannot create a reencryption key for the adversary from its existing reencryption. For example, \( R_{A \rightarrow C} = g^{cb} \) cannot be produced from \( R_{A \rightarrow B} = g^{da} \) and \( R_{B \rightarrow C} = g^{db} \).

Case 3. The proposed PHR model is secure against a security attack such as a replay attack.

Threat model: The adversary may copy a transaction of an authorized user from blockchain or by the way of intercepting the messages sent by an authorized user and then replay the message on the gateway server in order to obtain the PHR data.

Argument: The gateway server verifies not only the signature of the requested user but also the timestamp on the request. Firstly, if the timestamp is not valid, the gateway server will not respond to the request. Second, if the timestamp is still valid, the adversary can obtain the reencrypted PHR. Since, the reencrypted PHR is generated for the intended authorized user, the adversary still cannot decrypt the ciphertext due to the lack of the corresponding private keys.

Case 4. The proposed PHR model is secure against a security attack such as a malicious access.

Threat model: A malicious user wants to read/write the PHR data without an authorization.

Argument: A user must meet the access control protocol in order to decrypt the data. Firstly, the user has to search the data-id on the private blockchain. If the adversary does not obtain a valid identity credential of an authorized user, the adversary cannot access the private blockchain. The proposed model is protected against malicious readers and writers by employing the gateway server. Before any read transaction, for example, the gateway server verifies that the requested read action is valid and the request is sent by a party that is listed as an authorized reader in the corresponding access list. Thus, the gateway server will only reencrypt the encrypted PHR data for the authorized reader with the corresponding reencryption key.

5. Performance Analysis

The performance analysis of the proposed model is conducted by comparing it with the existing work of Hao Wang et al. [14]. Both systems consist of multiple components so the experiments are separated into two primary categories: a cryptographic test and the time required for all operations of the system. Both systems mainly perform their cryptographic operations on the real data and store the characteristics of the data on the blockchain. The cryptographic operation of the proposed system is based on a proxy reencryption scheme while the cryptographic operation of the model presented in [14] is based on an attribute-based encryption scheme. The time requirement of blockchain in our model is presented at the end to estimate the total storing and retrieving operation of our model on the current 4G network.

5.1. Experimental Setup. For the cryptographic operation comparison, all experiments are conducted on a VMware workstation using Ubuntu OS. The host system is a machine with Intel(R) Core(TM) i7-4510U CPU, 2.60 GHz, 8 GB RAM, running Windows 8.1. The proxy reencryption scheme and the attribute-based encryption scheme were implemented using Eclipse IDE (oxygen), Java 1.8, Java security library, Oracle Commons Lang 3.6 (OCL), Java Pairing-Based Cryptography (jPBC) [48], and Bounty Castle libraries. For the blockchain service testing, the Hyperledger blockchain network is created in Docker environment [49] with node.js. The blockchain network simply contains endorsing node, order node, and two peer nodes.

The synthetic PHR data consisting of various data sizes are used in this experiment. To generate our synthetic PHR workload, various data sources are investigated. First, the plaintext PHR records supported by [50] are 169 KB. Most medical videos used in MedCram [51] are less than 30 MB. The MP4 video of length 19:52 of 1280 pixels frame width, 720 pixels frame height, 29 frames/second of frame rate 183 kbps bit rate, and 2 stereo audio channel is approximately 26.6 MB. To cover these data sizes, the synthetic PHR workload consists of 128 KB, 512 KB, 2 MB, 8 MB, 32 MB, and 128 MB size. Because the healthcare data management systems must react with several users, the performance of the system on different numbers of users is also needed to be examined. The experiment is conducted on sharing the PHR data with 1, 2, 4, 8, and 16 users.

To perform the comparative performance analysis, the proxy reencryption scheme (PRES) of the proposed model and the attribute-based encryption scheme (ABES) of the model presented in [14] are implemented. The proxy reencryption scheme is implemented using the AFGH algorithm [29] and AES [52]. The attribute-based encryption scheme is implemented using the CP-ABE [53] and AES. Under both schemes, the data is encrypted with a symmetric cipher AES and the symmetric key is also encrypted to control the access. Therefore, the ciphertext is the combination of the encrypted data and the encrypted symmetric key. The AFGH algorithm and CP-ABE algorithm are performed for the encryption and decryption processed on the symmetric key.

5.2. Experimental Results. The experiments are conducted by varying the size of the data, varying the number of users, and using a real scenario for evaluation. In each experiment, the cryptographic operation time is tested by running the proxy reencryption scheme and the attribute-based encryption
Table 2: The time requirements of the encryption/decryption operations for various data sizes.

<table>
<thead>
<tr>
<th>Data Size</th>
<th>Encryption time of PRES (milliseconds)</th>
<th>Decryption time of PRES (milliseconds)</th>
<th>Encryption time of ABES (milliseconds)</th>
<th>Decryption time of ABES (milliseconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>128 KB</td>
<td>91.18</td>
<td>3.19</td>
<td>366.46</td>
<td>161.69</td>
</tr>
<tr>
<td>512 KB</td>
<td>94.01</td>
<td>6.04</td>
<td>370.69</td>
<td>170.01</td>
</tr>
<tr>
<td>2 MB</td>
<td>101.19</td>
<td>16.62</td>
<td>375.85</td>
<td>179.67</td>
</tr>
<tr>
<td>8 MB</td>
<td>142.29</td>
<td>59.19</td>
<td>423.11</td>
<td>226.02</td>
</tr>
<tr>
<td>32 MB</td>
<td>303.37</td>
<td>238.33</td>
<td>593.05</td>
<td>405.03</td>
</tr>
<tr>
<td>128 MB</td>
<td>1828.21</td>
<td>1814.79</td>
<td>2242.42</td>
<td>1950.48</td>
</tr>
</tbody>
</table>

scheme 100 times, and then the average processing time is recorded. Thus, the first experimental result shows the time requirements of the cryptographic operations for sharing various data sizes with a single user, the second experimental result shows the time requirements of the cryptographic operations for sharing the same data size with different numbers of users, and the third experimental result shows the estimated storing and retrieving time for the whole system.

5.2.1. Varying Data Sizes. This assessment is conducted by sharing various PHR data sizes to a single user. The proxy reencryption scheme consists of key pair generation, reencryption key generation, encryption, reencryption, and decryption processes. For all data sizes, the key pair generation time is 2.54 milliseconds; the reencryption key generation time is 16.98 milliseconds; and the proxy reencryption time is 17.41 milliseconds. The attribute-based encryption scheme consists of master key generation, user key generation, encryption, and decryption processes. For all data sizes, the master key generation time is 375.05 milliseconds, and the user key generation time is 375.46 milliseconds. The time for the encryption and decryption processes of PRES and IBES is shown in Table 2.

According to Table 2, the time required for both encryption and decryption processes of ABES is significantly larger than that of the PRES, while the incremental rate of each process of both schemes is similar because the experiment aim is to share the data with a single user and both schemes use the same encryption algorithm (AES). According to the result, the estimated time values for cryptographic operations on various data sizes for storing and retrieving processes of the proposed model and the model presented in [14] are compared in Figure 5. According to the data presented in Figure 5, the proposed model is more efficient than the model presented in [14] on various data sizes.

5.2.2. Varying the Number of Users. The medical data management systems need to support the data sharing because these systems usually contain multiple data uses. The data owner has to share his/her data with various users during the storing process by assigning it in the access control mechanism. In this assessment, the cryptographic operations are evaluated with different number of users with 8MB PHR data. The results are shown in Table 3.

According to the results shown in Table 3, the experiment is conducted on the 8MB sized data with 1, 2, 4, 8, and 16 users. Under the proxy reencryption scheme, the reencryption key generation time is increased with the number of users. The time for other processes including key pair generation time (2.54 milliseconds), proxy reencryption (17.41 milliseconds), encryption time (142.29 milliseconds), and decryption time (59.19 milliseconds) remains unchanged. Under the attribute-based encryption scheme, the values of pairing time significantly increased (i.e., the master key generation and the encryption time). That is, the time for the master key generation increases from 377.42 for 1 user to 737.31 for 2 users, and the time increases to 5602.93 for 16 users. The encryption time is 423.11, 784.47, 1,501.89, 2,844.20, and 5,639.91 for 1, 2, 4, 8, and 16 users, respectively. Thus, Figure 6 shows the estimated time required for cryptographic operations of both models on various numbers of users.

Under the proposed model, the data owner needs to perform the proxy reencryption key generation and the encryption processes for storing the PHR data. To retrieve the data, the proxy reencryption and the decryption processes must be performed. When the number of authorized users increases, an additional reencryption key process must be performed. According to the results shown in Table 3, the
<table>
<thead>
<tr>
<th>Number of users</th>
<th>Re-encryption key gen; time of PRES (milliseconds)</th>
<th>Encryption time of PRES (milliseconds)</th>
<th>Decryption time of PRES (milliseconds)</th>
<th>Master key gen; time of ABES (milliseconds)</th>
<th>User key gen; time of ABES (milliseconds)</th>
<th>Encryption time of ABES (milliseconds)</th>
<th>Decryption time of ABES (milliseconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16.62</td>
<td>142.29</td>
<td>59.19</td>
<td>377.42</td>
<td>375.46</td>
<td>423.11</td>
<td>226.02</td>
</tr>
<tr>
<td>2</td>
<td>33.01</td>
<td>142.29</td>
<td>59.19</td>
<td>737.31</td>
<td>375.46</td>
<td>784.47</td>
<td>226.02</td>
</tr>
<tr>
<td>4</td>
<td>67.66</td>
<td>142.29</td>
<td>59.19</td>
<td>1471.49</td>
<td>375.46</td>
<td>1501.89</td>
<td>226.02</td>
</tr>
<tr>
<td>8</td>
<td>134.52</td>
<td>142.29</td>
<td>59.19</td>
<td>2802.81</td>
<td>375.46</td>
<td>2844.20</td>
<td>226.02</td>
</tr>
<tr>
<td>16</td>
<td>272.89</td>
<td>142.29</td>
<td>59.19</td>
<td>5602.93</td>
<td>375.46</td>
<td>5639.91</td>
<td>226.02</td>
</tr>
</tbody>
</table>
reencryption key generation operation increases linearly with the number of users. Thus, the storing operation time will be affected while the retrieving operation time remains the same as shown in Figure 6.

Under the model presented in [14], the master key generation and the encryption processes need to be performed for a storing operation. For a retrieving operation, the user key generation and the decryption processes need to be performed. When the number of authorized users is increased, the attribute set and the policy must be changed. Thus, the master key generation and the encryption time will significantly increase because the identities of the users cannot be separated from the cryptographic operation. As a result, the storing operation time also significantly increases as shown in Figure 6.

5.2.3. Using Real Scenario. In this assessment, the time required for a storing operation and a retrieving operation is calculated on the 32 MB data to estimate the whole operation time. The time required for the cryptographic process of a storing operation is 322.64 milliseconds, and the time required for a cryptographic process of a retrieving operation is 254.65 milliseconds. The blockchain service time of a retrieving operation is 28.47 seconds on average and the system operation time for a retrieving operation is 28.72 seconds. By adding the upload/download time which is calculated on 4G network—theoretical peak speeds 100 Mbps to 1 Gbps [54] with the average download speeds of 18.6 Mbps and the average upload speeds of 9.0 Mbps—the estimated time of a storing operation for 32 MB data is 43.22 seconds (28.44 +14.78) and the estimated time of a retrieving operation for 32 MB data is 42.48 seconds (i.e., 13.76 second + 28.72 second).

However, the total estimated time does not include other operation time such as indexing. The total estimated time is well within the standard emergency response time (i.e., 8 minutes) [55].

6. Discussion

In this work, we proposed a blockchain-based access control model to preserve privacy for PHR system. Our proposed model uses a private blockchain, cloud storage, and other cryptographic mechanisms including proxy reencryption, hashing, and digital signature, to provide tamper resistance and privacy properties for the PHR system. Our proposed model can support the following attractive features: (1) the PHR owner can securely store and share his/her PHR data; (2) the PHR owner can revoke an access right on any PHR data easily; (3) all users including the PHR owner can conveniently check the integrity of the data. To achieve the tamper resistance property of our model, the blockchain technology is used. Consequently, the issues of using blockchain in the PHR system such as availability, consent revocation, and confidentiality of stored data are handled.

To handle the availability issue of blockchain, our model makes blockchain maintain only the small metadata. The encrypted real PHR is kept on the cloud storage to grantee the availability. The user can search the PHR via the metadata on blockchain and the user can request the encrypted PHR from the gateway server. Consequently, all accesses to the PHR data will be collected on blockchain to support an immutable audit trail. To handle the confidentiality issue of blockchain, the actual PHR is encrypted with the public key of the PHR owner, and only the metadata is revealed on the blockchain. To handle the consent revocation issue, an access control list, which includes the reencryption keys of the authorized users, can restrict the users on a certain operation. The PHR owner can revoke the access to his/her PHR data by updating the access list (revoking the reencryption key) and maintaining the ownership of the data.

To ensure the accomplishment of our goals, the proposed model is analyzed from security, privacy, and performance perspectives. For a security analysis, the first case shows that the proposed model achieves a tamper resistance. The second case, the third case, and the fourth case show that the proposed model can ensure the privacy of the PHR owner via the revocable access control list. For the performance evaluation, the proposed model is compared with an existing system [14]. The PHR system may contain large medical data and interact with various users. Therefore, the performances are compared by conducting two types of experiments. First, the proposed model is evaluated on different sizes of data when interacting with a single user. The proposed model outperforms the existing system [14] for every data size; however, the operation time incremental rate on various data sizes is similar. Using the same underlying encryption algorithm AES 256 with CBC mode causes similar incremental rate in operation time. Second, the proposed model is evaluated on a constant data size of 8 MB when interacting with various numbers of users (i.e., 1, 2, 4, 8, and 16). The proposed model also outperforms the existing system [14] for all numbers of users, while the operation time incremental rate is significantly different. The operation time of the existing
system [14] increases explosively while the operation time of the proposed model increases linearly. Accordingly, the proposed system is not only more efficient but also more suitable for PHR system.

To benchmark the private blockchain (Hyperledger Fabric) performance for our model, the latency of the system is studied through the default parameter configuration. Latency is the response time per transaction and it can be measured as the time taken from the data being sent by the application until the transaction is committed. Latency can be used as blockchain service time in our work. After sending 5,000 transactions for each of the writing and querying requests, the average blockchain service time of a storing operation is 12.48 seconds, and the average blockchain service time of a retrieving operation is 0.2 second. The upload/download time for 32 MB data on the 4G network is used to estimate the total operation time. Finally, the estimated total time for a storing operation of 32 MB data is 43.22 seconds and for a retrieving operation is 42.48 seconds.

To handle the performance and energy consumption issue of blockchain, the private blockchain (Hyperledger) is used in our system. Hyperledger supports the best transaction of more than 10K transactions per second and can reduce the overhead of mining process [28]. Hyperledger Fabric uses various components and various processing phases. Hyperledger Fabric provides various configurable parameters. However, the default parameter is used in this work. Thus, a comprehensive study on the effect of Hyperledger Fabric configurable parameters must be performed in the future work.

7. Conclusion

The blockchain-based personal health record system to promote a privacy-preserving access control model is proposed in this work. The proposed solution addresses the recurrent requirements of PHR system and the issues of using the blockchain technology in PHR system development. To preserve the privacy in the PHR system, the qualitative requirement, namely, a tamper resistance storage, and the functional requirement, namely, a revocable access control, are necessary. The blockchain is suitable for a tamper resistance storage; however, it is difficult to provide a revocable access control mechanism on blockchain. Moreover, there are other issues such as limited storage and privacy of on-chain data for using blockchain in PHR development. Our access control model is designed to be implemented with existing cryptographic primitives and the private blockchain technology in such a way that it can handle the blockchain issues for PHR system development and demonstrates our prioritization of privacy and access control. Then the privacy and security of the proposed model are analyzed with four thread models, namely, a tampering attack, a collusion attack, a replay attack, and a malicious access attack, to ensure the accomplishment of our original goals. This model is proposed not only to remedy the PHR privacy problems, but also to look forward to continuing the research in the uses of blockchain for healthcare data management and protection of privacy. The proposed solution is compared with existing work in healthcare data management area on performance analysis to ensure that the proposed solution is usable with an efficient computational cost.

Data Availability

PHR workload in this study was generated from various data sources as described in Section 5.1 and the workload is available from the authors upon request.

Conflicts of Interest

There are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the Higher Education Research Promotion and the Thailand’s Education Hub for Southern Region of ASEAN Countries Project Office of the Higher Education Commission.

References


[34] A. Baird, F. North, and T. S. Raghu, “Personal Health Records (PHR) and the future of the physician-patient relationship,” in...


[47] University of Miami, “What is the difference between the privacy and security of health information?” in Privacy — Office of Privacy, Data Security at Miller School of Medicine, 2017.


