Securing patient data in the cloud using Attribute Based Encryption

Chiedza Hwata, R.Subburaj Professor, Gladman Jekese

Abstract — Cloud computing has attracted attention worldwide in all industries, including the medical field leading to the rise of electronic healthcare systems. Although it has brought about an improvement in the provision of healthcare in terms of information management, it also poses a lot of security and privacy concerns to the patients. This is due to the fact that personal and highly sensitive data is outsourced to a third party (Cloud Service Provider) for processing and storage. This paper seeks to improve security of cloud-based patient data in healthcare organizations by employing a Ciphertext Policy Attribute Based Encryption (CP-ABE) scheme. The proposed scheme provides data confidentiality and allows the patient to control who accesses her personal health data by encrypting it under a specified access policy alongside with her key. It also provides collusion-resistance, flexible and immediate revocation of users who are no longer allowed to access a patient's data.

Index Terms— cloud computing, electronic healthcare, security, encryption, revocation, privacy, attribute based encryption

1 INTRODUCTION

A lot of healthcare organizations are on the move to adopt electronic healthcare systems, the case for cloud data storage which is compelling for deploying Electronic Healthcare systems: not only is it inexpensive but it also provides the flexible, wide-area mobile access that is increasingly needed in the modern world. It has become essential to support individual process activities and to satisfy collaboration and coordination needs by providing ready access to patient and operational information regardless of location and time. Filling this information gap by enabling the provision of the right information, to the right people, at the right time fosters new challenges, including the specification of a common information format, the interoperability among heterogeneous institutional information systems or the development of new, ubiquitous trans-institutional systems [1]. In the cloud, computing resources including storage is provided by a third party service provider [2] and [3].

With healthcare providers looking at automating processes of health information manipulation at lower cost and higher gains, cloud computing has been viewed as an appropriate platform to deploy standard medical information systems for its scalable and cost-effective services delivered by cloud service providers [4], [5]. Despite the increased usage of cloud-based data sharing platforms, the privacy and security related problems have prevented their adoption in the healthcare domain [6],[7].

- *Chiedza Hwata is an M.Tech student in Information Technology at SRM University, Chennai, India, E-mail: chiedza11@gmail.com*
- *Dr. R Subburaj is a Professor and Consultant in the Department of Information Technology at SRM University, Chennai., India. E-mail: subburaj.spr@gmail.com*
- *Gladman Jekese is an M.Tech student in Information Technology at SRM University, Chennai, India, E-mail: jgman86@gmail.com*

Lately there has been a high rate of cloud-based healthcare systems deployment yet a lot of patients' health information has continually leaked in the past years [8]. Considering sensitivity of health data, patients start to worry because they realize that they would completely lose control over their personal information once it enters the cyberspace. On the other hand, the healthcare organizations also become skeptical of the security offered by Cloud Service providers.

There are good reasons to be cautious in keeping medical data private and limiting the access because some employer may decide not to hire someone with a certain disease [9]. The proposed secure patient healthcare system is inspired by the need for revocation flexibility, fine-grained access control, and cost efficiency of the cloud-based patient data outsourcing paradigm. The Ciphertext Policy Attribute Based Encryption variation, proposed in this paper allows revocation of users who are no longer entitled to access patient data due to various reasons for example a medical doctor who has been let go by the health organization. It also allows a patient to selectively revoke a physician based on attributes, such that they can continue to access some less sensitive data after revocation. CP-ABE has been said to have the key escrow problem as described in [10], [11], the system proposed in this paper is collusion resistant by using a minimally trusted party for storage provision and computations.

The rest of the paper is organized as follows; Section 2 discusses the background, section 3 presents the proposed system, section 4 outlines the preliminaries of this design while section 5 defines the system design and finally 6 concludes the study.

2 BACKGROUND

2.1 Challenges associated with cloud adoption in healthcare organizations

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly

provisioned and released with minimal management effort or service provider interaction [12]. It conveys all resources and services as a service over the internet based on user demand with the adoption of public cloud services, a large part of the network, system, applications, and data will move under third-party provider control [13]. In the healthcare environment it may just be outsourcing storage, i.e. databases. Since these databases are often filled with valuable data, they are high value targets for attackers and security breaches in such systems are not uncommon, especially by insiders. In addition, organizations with access to extremely sensitive data might not want to give an outside server any access to their information at all [14]. Also cloud data typically resides in a shared environment, users will neither know the exact location of their data nor the other sources of the data collectively stored with theirs [6], [7] and [15]. Therefore, in this Internet-based computing paradigm, users are universally required to accept the underlying premise of trust [5] and [7]. However, when considering standards such as HIPAA (Health Insurance Portability and Accountability in Act) [16], it is crucial for health-related data to be kept confidential from anyone unless authorized by the patient or some emergency regulations.

The Cloud is generally susceptible to many privacy and security attacks. [17], discusses some of the legal and ethical issues that ought to be followed in Medical health data handling, together with consequences of breaches. As a result, many hospitals and health organizations are reluctant to adopt Cloud technology as a privacy breach in regards to its patient information can be devastating, especially in terms of cost [18]. This has led to researches and publications of the possible attacks and a number of solutions have been proposed. Cloud Security Alliance describes some of the top threats that are associated with adopting cloud computing [19]. Some of the threats include; Abuse and nefarious use of cloud computing, Insecure interfaces and APIs, Malicious insiders, Shared technology issues, Data loss or leakage, Account or service hijacking and Unknown risk profile. Some service providers may not provide a clear Service Level Agreement, which provides services at different levels [20], [21]. For health organizations it is very crucial to provide privacy and confidentiality since it involves personal information which is sensitive when revealed. An employer may decide not to hire someone with certain diseases. An insurance company may refuse to provide life insurance knowing the disease history of a patient. In healthcare, data leaks risk patients' health as well as their identity [22]. Some other problems with cloud data are associated with access control [23]. Due to the fact that the data can be accessed from anywhere at any time, it is of utmost importance that the level of access be monitored because not everyone will have the right intentions always. In cases where data is protected from the outside world, even the insiders need to be controlled. Imagine a situation where situation where a dentist can easily retrieve HIV and AIDS patient's data, they can publish it without much suspicion. Or a case where a heart surgeon that has been let go, they may still try to access patient data that they were able to retrieve when they were still employed there. They may also try to access data that was added after their access was revoked, whether encrypted or not. Users, out curiosity or with a malicious intend can collude with a trusted party or some other users that are compromised in the system so that they can access information that they are not privileged for.

2.2 Related work

There have been numerous publications in relation to security and privacy of cloud-assisted healthcare dating back to Medical Information Privacy Assurance (MIPA) [24]. MIPA was one of the early works done in electronic health which pointed out the importance and unique challenges of medical information privacy, and the devastating privacy breach facts that resulted from insufficient supporting technology and implemented a system that allows users to protect their data.
Since then, a number of asymmetric schemes have been put to use including Public-key cryptography, that uses a pair of keys for encryption where a private key is kept secret and a public key is widely distributed. If Alice wants to send a confidential message to Bob, she can encrypt the message with the public key of Bob and only Bob can decrypt the message using his private key. The problem in Public-Key Infrastructure (PKI) is that a public key must be obtained from, or at least be certified by the Trusted Third Party (TTP) of the PKI. As an improvement to PKI, Identity-Based Encryption (IBE), with its variations, was introduced. In IBE any string can be used to generate a public key without involving the TTP, thus creating a degree of flexibility. It basically allows any pair of users to communicate securely without exchanging any key (public or private) but based on the recipient's identity is used to encrypt the message. However if Alice does not fully know Bob's identity, except for his few attributes, then neither a PKI nor IBE will work. Sahai and Waters [25] introduced Attribute-Based Encryption (ABE) as a means for encrypted access control. Later on, a fine-grained access control ABE scheme propose by Goyal et al. [26] and Bethencourt et al. [27] further enhanced ABE practicality. Two categories of attribute-based encryption are distinguished in [28], where Key-Policy Attributed-Based Encryption (KP-ABE) , a ciphertext is associated with a set of attributes while a private key is issued as per a certain access control policy while Ciphertext-Policy Attribute-Based Encryption (CP-ABE) a ciphertext is generated according to some access control policies while private keys are issued in association with attributes. A secure and privacy preserving opportunistic computing framework termed CAM is proposed in [29]. The paper discusses about attribute-based access control and a new privacy preserving scalar product computation (PPSPC) technique. CAM promotes a user-centric access control by allowing each medical user to decide who can participate in the opportunistic computing, among the qualified helpers but its security model does not consider the possible side-channel attack due to the co-residency on shared resources either because it could be mitigated with either system level protection or leakage resilient cryptography.

In [30], a solution for patient data collection, which delivers an integrated telemedicine service that, automates the process

from data collecting to information deliver as a computing utility is developed. Sun et.al studied a Privacy-preserving health data storage system, where patients would have to encrypt their personal health data and get it stored it on a third-party server [31]. Tu and Niu [32] make use of CP-ABE in the context of enterprise applications and developed a revocation mechanism that simultaneously allows high adaptability, fine-grained access control and revocation. When a user is revoked access rights, the data is reencrypted in the Cloud rendering the revoked user's key useless. Even though, the re-encryption process is delegated to the Cloud, this is not efficient when considering very large data sizes. In [11], a scheme to do away with the key escrow problem that is found in CP-ABE was proposed, where two parties would be assigned the role to issue keys so that not one authority will get access to all the sensitive data. In [33], [34], [35], revocation schemes are proposed and proved to be efficient but required re-encryption of ciphertexts, which is computationally expensive. A proposal is made in [36], with fine-grained access control and flexible revocation scheme in MSNs. That CP-ABE scheme with revocation defined for MSNs is proposed healthcare environment in this paper.

3 PROPOSED SYSTEM

3.1 System overview

To achieve a fine-grained access control, data confidentiality, flexible and immediate revocation, the proposed system will employ the concept of ciphertext-policy attribute-based encryption (CP-ABE) as a basis for the encryption construction. In the CP-ABE scheme, a ciphertext is encrypted with an access policy chosen by an encryptor and a corresponding decryption key is created with respect to a set of attributes. As long as the set of attributes associated with a decryption key satisfies the access policy associated with a given ciphertext, the key can be used to decrypt the ciphertext. Due to the fact that patient health data consists of private and personal data, an access structure is defined first before a patient can share their information in the system. Therefore, based on the credentials assigned to each professional, they will be allowed access to patient information. For example anyone with the credentials of a medical doctor will be allowed to view a patient's medical history, past medications, family history and patient's diseases. On the other hand a pharmacist may just need to be allowed to view a patient's prescription because that is sufficient for him to provide the patient with the required medication.

The patient initializes the system, generates a master key for the patient and sends the public key to the other physicians who are meant to access patient data. Prior to storing their data on the cloud, a patient will encrypt their data, alongside the defined access structure. For the purpose of user revocation, she creates the revocation list, generates the corresponding revocation keys, and sends them to the service provider who will help her enforce access policy using these revocation keys. Whenever a physician wants to access the patient data, he first sends his transformation key to the service provider, which is part of his private key. Unless his private key has been regenerated, the physician only sends it once. The provider checks the revocation key, if the physician is not on the revocation list, transforms the encrypted data CT to CT', and transfers it to the physician, provided his attribute set S satisfies the access structure A. Receiving the transformed ciphertext CT', the physician can efficiently retrieve the plaintext using his secret value γ by executing only one exponent operation.

3.2 Proposed architecture

There are four parties involved in this scheme: the Cloud Service Provider (or simply the cloud), the patient (data owner), physician (includes doctors and all the related medical staff), and a trusted authority (TA). In this paper an attempt is made to utilize the CP-ABE scheme in [36], making use of these four entities;

Cloud Service Provider (CSP): It is a minimally trusted entity which is responsible for storing patients' (encrypted) data in a database and performing searches for the physicians. It is basically a third-party service provider that offers the on-demand storage and computing services for the health organization. The CSP is not fully trusted by users in the domain, in this paper we assume that the service provider is honest-but-curious; that is, it will try to find out as much secret information from the outsourced data as possible, but it will honestly execute the tasks assigned to it by legitimate parties in the system.

Patient: This entity is the person who is being treated at the health organization, who wishes to securely store and share her private data with her doctors selectively. In this system, she can define the access policy by herself based on the physician's attributes and a specific access structure; enforce it by encrypting it under the policy before outsourcing it to the Cloud Service Provider.

Physicians: This is the group that includes all the hospital staff that will be handling patient data, for example the doctor who is treating a certain patient. The physicians obtain their private keys based on their professional responsibilities, and need to access the patient records for providing medical care.

Trusted Authority: This entity is in charge of cryptographically initializing each user's registration into the system. The TA assigns a Global identity called the Master Key as well as a global public key, which is published to all the other users to allow them to use it for encrypting or decrypting data relative to that user. Each patient and physician is assigned a key pair which is used to perform security operations such as authentication in this domain.

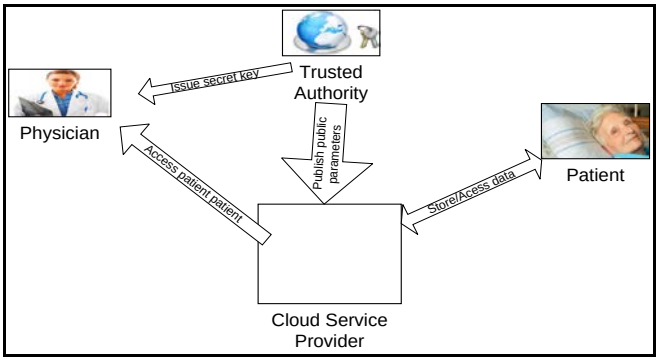

Figure. 1

4 PRELIMINARIES

This section gives formal definitions for access structures and relevant background on Linear Secret Sharing Schemes (LSSS), followed by the definitions of Ciphertext Policy Attribute Based Encryption (CP-ABE) then a brief description on revocation.

4.1 Linear Secret Sharing Scheme

In our context, the role of the parties described in Figure.2 is taken by the attributes. Thus, the access structure $\mathbb{A}$ will contain the authorized sets of attributes S and if a set of attributes is not in $\mathbb{A}$ then the owner is an unauthorized user.

Definition 1 (Access Structure [6]). *Let $\{P_1,P_2,\ldots,P_n\}$ be a set of parties. A collection $\mathbb{A} \subseteq 2^{\{P_1,P_2,\ldots,P_n\}}$ is monotone if $\forall B,C$: if $B \in \mathbb{A}$ and $B \subseteq C$ then $C \in \mathbb{A}$. An* access structure *(respectively, monotone access structure) is a collection (respectively, monotone collection) $\mathbb{A}$ of non-empty subsets of $\{P_1,P_2,\ldots,P_n\}$, i.e., $\mathbb{A} \subseteq 2^{\{P_1,P_2,\ldots,P_n\}}\backslash\{\emptyset\}$. The sets in $\mathbb{A}$ are called the* authorized sets, *and the sets not in $\mathbb{A}$ are called the* unauthorized sets.

Figure. 2

Definition 2 (Linear Secret-Sharing Schemes (LSSS)). *A secret-sharing scheme Π over a set of parties $\mathcal{P}$ is called* linear *(over $\mathbb{Z}_p$) if*

1. *The shares for each party form a vector over $\mathbb{Z}_p$.*
2. *There exists a matrix an M with ℓ rows and n columns called the share-generating matrix for Π. For all $i = 1,\ldots,\ell$, the i'th row of M we let the function ρ defined the party labeling row i as $\rho(i)$. When we consider the column vector $v = (s,r_2,\ldots,r_n)$, where $s \in \mathbb{Z}_p$ is the secret to be shared, and $r_2,\ldots,r_n \in \mathbb{Z}_p$ are randomly chosen, then Mv is the vector of ℓ shares of the secret s according to Π. The share $(Mv)_i$ belongs to party $\rho(i)$.*

Figure. 3

It is shown in [34] that every linear secret sharing-scheme, according to the Figure. 3, definition also enjoys the linear reconstruction property, defined as follows:

Suppose that Π is an LSSS for the access structure $\mathbb{A}$. Let S $\in \mathbb{A}$ be any authorized set, and let I $\subset \{1, 2,\ldots, l\}$ be defined as $I = \{i : \rho(i) \in S\}$. Then, there exist constants $\{\omega i \in Z_p\}_{i\in I}$ such that, if $\{\lambda_i\}$are valid shares of any secrets according to Π, then $\omega i\lambda i = \boldsymbol{S}.\sum_{i\in I}\omega_i\lambda_i = s$. Furthermore, it is shown in [34] that these constants $\{\omega_i\}$ can be found in time polynomial in the size of the share-generating matrix M [32].

4.2 Ciphertext-Policy Attribute-Based Encryption (CP-ABE)

Proposed by Bethencourt et al. in 2007, CP-ABE is a variation of Attribute Based Encryption in [27] that makes use of encrypted data (ciphertext) for its access policy. It is considered one of the most suitable schemes for data access in cloud storage because it provides data owners with more direct control and flexibility on access policies. In a ciphertext-policy attribute-based encryption scheme, ciphertexts are associated with access structures over the subsets of at most n attributes of the attributes set, for some specified $n \in \mathbb{N}$. Decryption works only if the attribute set ω associated with a certain secret key is authorized in the access structure $\mathbb{A}$ (i.e., $\omega \in \mathbb{A}$) [21]. Bethencourt et al.'s CP-ABE scheme consists of the following four algorithms: Setup, KeyGen, Encrypt, and Decrypt, as shown in Figure. 4. These algorithms are described in detail in [28].

$\mathsf{Setup}(1^\lambda)$: Select a generator $g \in_R \mathbb{G}$ and an integer $x \in_R \mathbb{Z}_q$, and set $g_1 = g^x$. Then, pick elements $g_2, h_1, \ldots, h_n \in_R \mathbb{G}$. Finally, output the public key $pk = (g, g_1, g_2, d, h_1, \ldots, h_n)$ and the master secret key $msk = x$.

$\mathsf{KeyGen}(\omega, msk)$: Upon receiving a private key request on ω, randomly pick a $(d-1)$-degree polynomial $poly(\cdot)$ with $poly(0) = msk$. Then, for each attribute $i \in \omega$, compute $d_{i0} = g_2^{poly(i)} \cdot (g_1 h_i)^{r_i}$ and $d_{i1} = g^{r_i}$, where $r_i \in_R \mathbb{Z}_q$. Finally return $sk = (\{(d_{i0}, d_{i1})\}_{i\in\omega})$.

$\mathsf{Encrypt}(\hat{\omega}, m)$: To encrypt a message $m \in \mathbb{G}$ under $\hat{\omega}$, select an integer $s \in_R \mathbb{Z}_q$. Then, compute $c_0 = m \cdot e(g_1, g_2)^s$, $c_1 = g^s$ and $e_i = (g_1 h_i)^s$ for $i \in \hat{\omega}$. Finally, publish the ciphertext as $ct = (\hat{\omega}, c_0, c_1, \{e_i\}_{i\in\hat{\omega}})$.

$\mathsf{Decrypt}(sk, ct)$: Suppose that a ciphertext ct is encrypted under an attribute set $\hat{\omega}$ and physician is assigned with a private key sk for attribute set ω, which satisfies the restriction that $\gamma(\hat{\omega}, \omega) = 1$. Then, the decryption proceeds as follows. Firstly, an arbitrary d-element subset set $S \subseteq \hat{\omega} \cap \omega$ is selected. Then, the ciphertext is decrypted as $m = \frac{\prod_{i\in S} e(c_1, d_{i0})^{\Delta_{i,S}(0)}}{\prod_{i\in S} e(d_{i1}, e_i)^{\Delta_{i,S}(0)}}$, where $\Delta_{i,S}(0) = \prod_{j\in S, j\neq i} \frac{-j}{i-j}$ is the Lagrange polynomial at zero point.

Figure. 4

4.3 Ciphertext Policy Attribute Based Encryption with user revocation

Motivated by the fact that the professionals handling the patient data or the people in the social circles of the patient may change and the need for access control in cloud storage, the need for revocation arose. This was meant to address cases such as; how to disqualify a user and prevent him to access ciphertexts that were encrypted while the user still had rights and also ensuring that a newly encrypted data is not decryptable by a user whose key has been revoked already. The paradigm of revocation is categorized into two designs namely the central-control and another one is user-control. In a central-control design, system manager or a trusted third parties T centrally maintains revocation lists and this is the one that was incorporated by Fatos et.al. To implement revocation, an additional algorithm; **PubUpdate** is used which is periodically run such that using publicly available information, ciphertexts stored on the system can be updated periodically [37]. This will ensure that as soon as a user U has been revoked, all files become inaccessible to them, regardless of how old each file is. The algorithm is shown in Figure. 5.

$\mathsf{PubUpdate}(pk_{\text{old}}, \mathcal{S}, msk_{\text{old}})$: To update all the attribute public keys corresponding to the attributes in $\mathcal{S}$, the authority uses the old master secret key $msk_{\text{old}} = (x, \{t_i\}_{i\in\mathcal{U}})$ to proceed as follows. For each $i \in \mathcal{U}$, it computes $h_i' = h_i^{t_i'/t_i}$ where $t_i' \in_R \mathbb{Z}_q$ if $i \in \mathcal{S}$, and sets $h_i' = h_i$ and $t_i' = t_i$ if $i \notin \mathcal{S}$. Finally output the updated public keys $pk_{\text{new}} = (\{h_i'\}_{i\in\mathcal{U}})$, re-encryption key $rk = (\{\frac{x+t_i'}{x+t_i}\}_{i\in\mathcal{U}})$ and keep the adaptively updated master secret key $msk_{\text{new}} = (x, \{t_i'\}_{i\in\mathcal{U}})$ local.

$\mathsf{ReEncrypt}(ct, rk)$: In order to transform a ciphertext encrypted with an old version of public keys to that under the current public keys, the cloud server computes $e_i' = e_i^{\frac{x+t_i'}{x+t_i}}$ for all the $i \in \hat{\omega}$, $c_1' = c_1$ and $c_0' = c_0$. Finally the ciphertext is updated as $ct_{\text{new}} = (c_0', c_1', \{e_i'\}_{i\in\hat{\omega}})$.

Figure 5

5 SYSTEM DESIGN

The proposed system, using the CP-ABE as the underlying primitive, includes;

System setup: it is basically an initialization of the system. It is used to generate the master secret key and public parameters from an input of the security parameter $\mathcal{K}$ and the attribute universe U.

Key generation: the patient generates her private key and the revocation key for the physician which has to be revoked.

Data storage: data is to be stored in the cloud, but prior to its transmission to the Cloud Service Provider, it has to be encrypted.

Data access: for a physician to be able to access a patient's data, he sends a request message and his transformation key *tk* to the Cloud Service Provider, which will return the transformed ciphertext CT' or a reject response.

Decryption: If the physician has received a ciphertext from the Cloud Service Provider, then he can apply his secret key *sk* to decrypt the ciphertext.

5.1 System setup

To initialize the system, the patient runs the following steps:

(1) Set the security parameter $\mathcal{K}$ and the attribute universe description $U = \{1,2,\dots,|U|\}$ and choosing two multiplicative cyclic groups G and G_T of prime order p with an admissible bilinear map $e: G \times G \rightarrow G_T$ and a hash function $H: \{0,1\}^* \rightarrow G$;

(2) Randomly choosing $\propto, \beta, a \in Z_P$ and a polynomial P_x of degree t_x (t_x is the maximum number of revoked physicians for attribute x at a given time) over Z_P, for each attribute x such that $P_x(0) = \beta$. The public parameters are published as

$$params = \{G, g, e(g,g)^{\alpha}, g^{a}, H\}, \quad (1)$$

Where, g is a generator of G. The master secret key is set as $msk = (g^{\alpha}, \beta, \{P_x\}_{\forall_x})$.

5.2 Key generation and distribution

The patient generates her private key and the revocation key in the following steps:

(1) The private key generation process for patient involves taking as input the msk and calculating u_k's transformation key tk as follows:

$$K = g^{(\propto/y}\, g^{at\beta}, \quad L = g^{t\beta}, \quad (2)$$
$$\{K_x = H(x)^t\}_{\forall_x \in S}, \quad \{K'_x = H(x)^{tP_x(u_k)}\}_{\forall_x \in S}.$$

Then, set $sk = (\gamma, tk) = \left(\gamma, \left(K, L, \{K_x, K'_x\}_{\forall_x \in S}\right)\right)$.

(2) Take as input *msk* and $\{RL_x\}_{\forall_x \in U'}$, where $U' \subset U$ is the set of attributes that the patient decides to revoke and RL_x is a list of physicians $\{u_1, \dots, u_{t_x}\}$whose attribute x will be revoked by the patient. Then, create the revocation key $rk = (rk_1, rk_2)$ as follows:

$$rk_1 = \left\{\langle u_i, P_x(u_i)\rangle_{\forall_{u_i} \in RL_x}\right\}_{\forall x \in U'},$$

$$rk_2 = \left\{\langle x_i, P_x(x_i)\rangle_{\forall_{x_i} \in \{1,\dots,t_x\}}\right\}_{\forall x \in U - U'}, \quad (3)$$

Where; $\{x_i\}$ are chosen from Z_p and they are different from any user's identity.

Lastly, the patient sends the private key *sk* to the user u_k through a secure communication channel and updates the revocation key *rk* to the Cloud Service Provider.

5.3 Encryption

A patient's data includes his personal details, medical history and current diagnosis that are to be sent for storage in the cloud and have to be encrypted first. Encryption has to be done before the data is uploaded to protect it while in transit as well as at rest on the Cloud Service Provider's premises. The encryption process for the health data is done as follows.

As input parameters, a message *m*, and an access structure (M, ρ), are taken, with M denoting a matrix of *l* rows and *n* columns and the function ρ associating rows of *M* to the attribute universe U.

A random vector $v = (s, \gamma_2, \ldots, \gamma_n) \in Z_p^n$ is firstly chosen, where the values $\gamma_2, \ldots, \gamma_n$ will be used to share the encryption exponent s, and then calculate $\lambda_i = vM_i$ for $i = 1$ to l, where M_i is the ith row of M.

Randomly choose $r_1, \ldots, r_l \in Z_n$ at first, and then compute the ciphertext *CT* as follows:

$$C = m \cdot e(g,g)^{\alpha s}, \quad C' = g^s,$$
$$(C_i = g^{a\lambda_i} H(\rho(i))^{-r_i}, \; D_i = g^{r_i})_{\forall i \in \{1,\ldots,l\}}. \qquad (4)$$

Finally, the patient uploads the encrypted data;

$CT = ((M, \rho), C, C', (C_i, D_i)_{\forall i \in \{1,\ldots,l\}})$ to the Cloud Service Provider.

5.4 Data access

When a physician wants to access a patient's data, he sends the request message and his transformation key *tk* to the CSP. Then, the CSP transforms the encrypted data *CT* to CT' as follows.

(1) It takes as input *rk*, *tk*, u_k and *CT* and outputs ⊥, if u_k's attribute set S′ after revocation does not satisfy the access structure (M, ρ). Otherwise, let $I' \subset \{1, \ldots, l\}$ be defined as $I' \subset \{i : \rho(i) \in S'\}$ and let $\{\omega_i' \epsilon Z_p\}_{i \in I'}$ be a set of constants, such that $\sum_{i \in I'} \omega_i' \lambda_i = s$ if $\{\lambda_i\}$ are valid shares of any secret *s* according to *M*. First calculate D_i' as follows:

$$D_i' = D_i^{\sum_{j=1}^{t_{\rho(i)}} \mu_{\rho(i),j} P_{\rho(i)}(u_j)}, \forall i \in I', \qquad (5)$$

Where $\mu_{\rho(i),j} = \left(u_k / (u_k - u_j)\right) \Pi_{n \neq j} \left(\frac{u_n}{u_n - u_j}\right)$, for all

$j, n \in \{1, \ldots, t_{\rho(i)}\}$, $k \notin \{1, \ldots, t_{\rho(i)}\}$, and then compute $\mu_{\rho(i),k}$:

$$\mu_{\rho(i),k} = \prod_{n \neq k} \frac{u_n}{u_n - u_k}, \quad \forall n \in \{1, \ldots, t_{\rho(i)}\}, k \notin \{1, \ldots, t_{\rho(i)}\}. \qquad (6)$$

(2) Terminate the transformation process if the output in step (1) is ⊥. Otherwise, let $I \subset \{1, \ldots, l\}$ be defined as $I = \{i: \rho(i) \in S\}$ and let $\{\omega_i \epsilon Z_p\}_{i \in I}$ be a set of constants, such that $\sum_{i \in I'} \omega_i \lambda_i = s$ if $\{\lambda_i\}$ are valid shares of any secret s according to *M* (there could be different ways to choose the values to satisfy this). Then, calculate TC_1 as follows:

$$TC_1 = \frac{e(C', K)}{\prod_{i \in I'} (e(L, C_i) e(D_i, K'_{\rho(i)})^{\mu_{\rho(i),k}} e(D_i', K_{\rho(i)}))^{\omega_i'}}$$

$$= \frac{e(C', K)}{\prod_{i \in I'} (e(g,g)^{at\lambda_i \beta} \, e(g, H(\rho(i)))^{-tr_i \beta})^{\omega_i'}}$$

$$\cdot \frac{1}{\prod_{i \in I'} (e(g, H(\rho(i)))^{tr_i \mu_{\rho(i),k} P_{\rho(i)}(u_k)})^{\omega_i'}}$$

$$\cdot \frac{1}{\prod_{i \in I'} (e\left(g, H(\rho(i))\right)^{tr_i \sum_{j=1}^{t_{\rho(I)}} \mu_{\rho(I),j} P_{\rho(i)}(u_j)})^{\omega_i'}}$$

$$= \frac{e(C', K)}{\prod_{i \in I'} e(g,g)^{at\lambda_i \omega_i' \beta}} = \frac{e(g,g)^{(\alpha/\gamma)s} \, e(g,g)^{at\beta s}}{e(g,g)^{at\beta \sum_{i \in I'} \lambda_i \omega_i}} \qquad (7)$$

$$= e(g,g)^{(\alpha s/\gamma)}.$$

In the end the Cloud Service Provider sends back the transformed ciphertext to the physician.

5.5 Decryption

Decryption is associated with patient data access. When a doctor, for example a heart surgeon wants to access a patient's medical history, they first have to download the file which will be in a ciphertext format. Depending on whether he has the right privileges, he can access the file. The decryption is a success if and only if; the attribute set S associated with the *CT* satisfies the access policy given by A, and the physician's identity specified by the secret key has not been revoked according to the revocation list

Receiving the transformed ciphertext CT', the authorized physician u_k can easily retrieve the patient's data in the following way:

(1) Take as input the transformed ciphertext CT' and his private key *sk* associated with an attribute set S which satisfies the access structure enforced on the encrypted data.

(2) Retrieve the message *m* by simply computing

$$\frac{TC_0}{TC_1^{\gamma}} = \frac{m \cdot e(g,g)^{\alpha s}}{e(g,g)^{\alpha s}} = m. \qquad (8)$$

6 Security Analysis

6.1 Confidentiality

The proposed system prevents unauthorised who have attributes that do not satisfy the access policy from learning the content of the private data encrypted under the policy. Revoked users should also be prevented from accessing patient data unless their remaining attributes satisfy the access policy.

For an unauthorized physician who has an attribute set S which does not satisfy the access policy, he cannot recover the desired value $e(g,g)^{\alpha s}$, which is needed for decryption in both

physician and attribute revocation. While trying to retrieve $e(g,g)^{at\beta s}$ from pieces $e(g,g)^{at\,\lambda_i\beta}$ during the decryption process, the physician is required to have a secret key associated to an attribute set satisfying the access structure (M,ρ) but when revoked, he cannot decrypt the ciphertext in the user revocation case. This is due to the fact that his secret key would be completely revoked; that is, $e(g,H(\rho(i)))^{tr_i\beta}$ cannot be retrieved using Lagrange interpolation for every attribute(i), $i \in I$. For the attribute revocation case, if the partial secret key corresponding to some attributes that satisfy the access policy is revoked, the physician cannot decrypt the ciphertext (unless the rest of his attributes still satisfy the policy). This is because $e(g,H(\rho(i)))^{tr_i\beta}$cannot be obtained for the revoked attributes $\rho(i), i \in I'$. Therefore, in both cases the physician cannot access $e(g,g)^{\propto s}$. Because the CSP is not fully trusted, it cannot decrypt any ciphertext, thought it possesses the revocation keys since it will not be having secret keys assigned to it. On the other hand, even if the service provider can help some physician transform the ciphertext CT to CT′ and obtain $e(g,g)^{\propto s}$ using the user's transformation key tk, he still cannot decrypt the ciphertext, because he does not know the secret key. Hence, data confidentiality against the service provider is also guaranteed.

6.2 Collusion resistance

Even if a group of physicians that are not entitled to access a specific ciphertext collude together by combining their attributes, the y still cannot decrypt the ciphertext. The proposed scheme is collusion-resistant, considering attacks from Cloud Service Provider (CSP), users who cannot decrypt ciphertext alone and users that are authorized but revoked.

Attack from the CSP: An authorized yet revoked physician u_k tries to decrypt a selected ciphertext by colluding with the CSP. However, the service provider just keeps the latest revocation key in its memory and the old one is erased each time the revocation takes place. Therefore, D_i' and $\mu_k$cannot be calculated hence the revoked user fails to decrypt the ciphertext even though his attributes meet the specified access policy. The proposed system is thus collusion-resistant from the revocation and CSP attack.

Attack from unrevoked yet unauthorized physician or revoked yet authorized physician: In the case of an authorized but revoked physician, their attribute set S meets the access structure but still cannot decrypt the ciphertext because his secret key is revoked completely. Due to the fact that he cannot get the coefficient μ_k and D_i' from the CSP, he cannot obtain $e(g,H(\rho(i)))^{tr_i\beta}$. Therefore he cannot retrieve $e(g,g)^{at\,\lambda_i\beta}$ neither can he reconstruct the required value $e(g,g)^{at\,\beta s}$for decryption. Even in a case where the authorized and revoked physician colludes with the unauthorized but unrevoked physician, he cannot retrieve $e(g,g)^{at\,\lambda_i\beta}$, since t is a random and unique exponent for each user. Furthermore, the latter can only obtain finite $e(g,H(\rho(i)))^{tr_i\beta}$ for his attribute set which does not satisfy the policy. For the same reasons, the former cannot help the latter retrieve $e(g,g)^{at\,\beta s}$ as well.

Attack from multiple physicians who cannot decrypt the ciphertext alone: Colluding physicians can retrievee$(g,g)^{at\,\lambda_i\beta}$, as they try to recover the required value $e(g,g)^{at\,\beta s}$for decryption purposes. With enough shares $e(g,g)^{at\,\lambda_i\beta}$ of exponent S, according to LSSS it is quite easy to obtain $e(g,g)^{as}$ however, the value t is a random and unique exponent for each physician, so the reconstruction of s is prevented by the distinct exponents. Hence, attribute collusion attack can be precluded in the proposed schemes.

6.3 Forward and backward secrecy

In the healthcare environment, a patient will have to revoke a physician who is no longer providing her healthcare, for example if a patient had fever a week and Physician A was their doctor. After the fever is gone the patient can revoke Physician A because there is no reason for him to still continue accessing her data, concerning that sickness. If a patient revokes a physician u_k, the patient creates and sends a revocation list (RL) including the physician u_k that has to be revoked, and generates a revocation key *rk* to the CSP. The CSP cannot cancel the appearance of the random part $e(g,H(\rho(i)))^{-tr_i\beta}$in the transformation process because the CSP cannot calculate μ_k and $D_i'$for u_k since u_k belongs to the revocation list. Therefore the physician's secret key would be completely revoked and will not have access to the plaintext of subsequent data uploaded afterwards. For attribute revocation, the CSP would be unable to compute $\mu_{\rho(i),k}$ and D_i' if u_k's attributeρ(i) is revoked, hence it will not be able to cancel $e(g,H(\rho(i)))^{-tr_i\beta}$. Therefore the physician will be in no position to access the plain text of subsequent data, unless his remaining attributes can still satisfy the access structure, for example if he is also a specialist in another department.

With regards to backward secrecy, a revoked physician who wants to access the previous data will transmit u_k and tk to the CSP. However it cannot calculate μ_k/ $\mu_{\rho(i),k}$ and D_i' for u_k, since the revocation key rk is updated and u_k is included in the revocation list. Therefore the physician will not be able to access the previous data, except his remaining attributes still satisfy the access policy (in the case of attribute revocation). Even when μ_k/ $\mu_{\rho(i),k}$ for the revocation list RL/$RL_{\rho(i)}$ and D_i' for the ciphertext D_i have been calculated before his attributes were revoked, they will not be sufficient to decrypt subsequent or previous data, because *rk* is updated and D_i' for the new D_i needs to be calculated.

7 Conclusion

This paper discusses the challenges that are faced in the healthcare industry due to the adoption of cloud computing, CP-ABE and its variation with revocation. It proposes a secure healthcare system that inherits ciphertext policy attribute based encryption, with the aim to improve privacy, confidentiality and security by improving revocability at the same time ensuring collusion-resistance together with forward and backward secrecy.

REFERENCES

1 M. Poulymenopoulou & F. Malamateniou & G. Vassilacopoulos, Emergency Healthcare Process Automation Using Mobile Computing and Cloud Services Department of Digital Systems University of Piraeus, Greece, Springer Science and Business Media, 2011.

2 L. Wang, D. Chen, Y. Hu, Y. Ma, J. Wang, Towards enabling cyber infrastructure as a service in clouds, Computer. Electrical. Engineering, vol. 1, pp. 3–14, 2013.

3 L. Wang, G. Laszewski, A.J. Younge, X. He, M. Kunze, J. Tao, C. Fu, Cloud computing: a perspective study, New Generation. Computing, vol. 28 no.2, pp. 137–146, 2010.

4 Mu-Hsing Kuo, Opportunities and challenges of cloud computing to improve health care services, Journal of Medical Internet Research, 2011.

5 Syed A. Ahson, Mohammad Ilyas, Cloud Computing and Software Services Theory and Techniques, CRC Press, 2010.

6 L.M. Kaufman, Data security in the world of cloud computing, IEEE Security and Privacy. vol. 7, no. 4, pp. 61–64, 2009.

7 D. Zissis, D. Lekkas, Addressing cloud computing security issues, Future Generation Computer. Systems, pp. 583–592, 2012.

8 Managing cyber risks in an interconnected world Key findings from The Global State of Information Security Survey 2015, www.pwc.com/gsiss2015 , www.pwc.com/cybersecurity , 2014

9 Yue Tong, Jinyuan Sun, Sherman S. M. Chow, and Pan Li, Cloud-Assisted Mobile-Access of Health Data With Privacy and Auditability, IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 2, 2014.

10 S. Vimercati, S. Foresti, S. Jajodia, S. Paraboschi, P. Samarati, A data outsourcing architecture combining cryptography and access control, Computer security architecture, ACM, pp. 63-69, 2007.

11 Junbeom Hur, Dongyoung Koo, Seong Oun Hwang, Kyungtae Kang, Removing escrow from ciphertext policy attribute-based encryption, School of Computer Science and Engineering, Chung-Ang University, Department of Computer Science, KAIST, Department of Computer and Information Communications Engineering Hongik University, Department of Computer Science and Engineering, Hanyang University, Computers and Mathematics with Applications, Elsevier, 2012.

12 Peter Mell & Timothy Grance, The NIST Definition of Cloud Computing, Computer Security Division Information Technology Laboratory, National Institute of Standards and Technology Gaithersburg, NIST Special Publication 800-145, 2011.

13 Tim Mather, Subra Kumaraswamy, and Shahed Latif, Cloud Security and Privacy, O'Reilly Media, Inc., First Edition, September 2009.

14 Amit Sahai , Hakan Seyalioglu, Brent Waters, Dynamic Credentials and Ciphertext Delegation for Attribute-Based Encryption, UCLA, University of Texas at Austin, 2012.

15 J. Zhao, L. Wang, J. Tao, J. Chen, W. Sun, R. Ranjan, J. Kolodziej, A. Streit, D. Georgakopoulos , A security framework in G-Hadoop for big data computing across distributed cloud data centers, Journal of Computer and Systems Sciences, vol. 80, no.5, pp. 994–1007, 2014.

16 HIPAA. U.S. Department of Health and Human Services. http://www.hhs.gov/ocr/privacy/index.html.

17 Kathryn A. Booth, Leesa G. Whicker, Terri D. Wyman, Donna J. Pugh, Sharon Thompson, Medical Assisting: Administrative and Clinical Procedures, Legal and Ethical Issues in Medical Practice, Including HIPAA, , McGraw-Hill Higher Education, 2009.

18 Juha Saarinen, UK health trust fined for privacy breach. Information Technology News, Aug 7, 2012. http://www.itnews.com.au/News/311079,uk-health-trust-fined-for-privacy-breach.aspx .

19 M. Li, S. Yu, Y. Zheng, K. Ren, and W. Lou, "Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption, IEEE Transactions, Parallel Distributed Systems., vol. 24, no. 1, pp. 131–143, 2013.

20 Bruce Jennings, Mary Ann Baily, Melissa Bottrell, Joanne Lynn, HealthCare Quality improvement: ethical and regulatory issues, The Hastings Center Garrison, New York, 2007.

21 Brussels, Cloud Service Level Agreement Standardisation Guidelines, ISO/IEC JTC 1/SC 38- http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=63902 , 2014.

22 M. Eric Johnson and Nicholas D. Willey, Usability Failures and Healthcare Data Hemorrhages, Dartmouth College, IEEE Computer and Reliability Societies, 2011.

23 Dan Hubbard, Michael Sutton, Zscaler, Top Threats to Cloud Computing, Websence, Cloud Security Alliance, 2010.

24 Ateniese, R. Curtmola, B. de Medeiros, and D. Davis, Medical information privacy assurance: Cryptographic and system aspects, 3rd Security in Communication Networks Conference, Amalfi, Italy, 2002.

25 A. Sahai and B. Waters. Fuzzy Identity Based Encryption. In Advances in Cryptology – Eurocrypt, Springer, vol. 3494 of LNCS, pp. 457–473, 2005.

26 V. Goyal, O. Pandey, A. Sahai, B.Waters, "Attribute -based encryption for fine-grained access control of encrypted data", ACMCCS 2006, pp. 89 -98, 2006.

27 Bethencourt J, Sahai A, et al. Ciphertext-Policy attribute-Based Encryption, IEEE, Security and Privacy, pp.321-334, 2007.

28 Cheng-Chi Lee, Pei-Shan Chung, and Min-Shiang Hwang, A Survey on Attribute-based Encryption Schemes of Access Control in Cloud Environments, Department of Library and Information Science, Fu Jen Catholic University, Department of Management Information Systems, National Chung Hsing University, Department of Computer Science and Information Engineering, Asia Universtiy, International Journal of Network Security, vol.15, no.4, pp. 231-240, 2013.

29 Huang Lin, Jun Shaoy, Chi Zhangz, and Yuguang Fang,Fellow, "CAM: Cloud-Assisted Privacy Preserving Mobile Health Monitoring, IEEE Tranasctions on image processing, vol 8, no. 6, 2013.

30 Carlos Oberdan Rolim, Fernando Luiz Koch, Carlos Becker Westphall, Jorge Werner, Armando Fracalossi, Giovanni Schmitt Salvador, A Cloud Computing Solution for Patient's Data Collection in Health Care Institutions, Network and Management Laboratory - LRG Federal University of Santa Catarina Florianopolis, Telemedicine, and Social Medicine, eTELEMED, 2010.

31 J. Sun, X. Zhu, C. Zhang, and Y. Fang, HCPP: Cryptography based secure EHR system for patient privacy and emergency healthcare, IEEE International Conference Distributed Computing Systems, pp. 373–382, 2011.

32 Tu S, Niu S, Li H, Xiao-ming Y, Li M (2012): "Fine-grained access control and revocation for sharing data on clouds", IEEE 26th international parallel and distributed processing symposium workshops and PhD forum (IPDPSW), pp. 2146–2155, 2012.

33 CHENG, Zhi-ying WANG, Jun MA, Jiang-jiang WU, Song-zhu MEI, Jiang-chun, "Efficient revocation in ciphertext-policy attribute-based encryption based cryptographic cloud storage", Yong, School of Computer, National University of Defense Technology, Changsha, China, Zhejiang University and Springer-Verlag Berlin Heidelberg, Computer and Electronics, vol. 14, no. 2, pp. 85-97, 2013.

34 D. Boneh, X. Ding, G. Tsudik, and C. Wong, "A method for fast revocation of public key certificates and security capabilities," in Proceedings of the 10th conference on USENIX Security Symposium, vol. 10. USENIX Association, p.

22, 2001.

35 Danan Thilakanathan, Shiping Chen, Surya Nepal, Rafael Calvo, Leila Alem, A platform for secure monitoring and sharing of generic health data in the Cloud, The University of Sydney, Australia, Future Generation Computer Systems, Elsevier, 2013.

36 Shi-Feng Sun, Chen Lyu,1 Dawu Gu, Yuanyuan Zhang, and Yanli Ren, Towards Efficient, Secure, and Fine-Grained Access Control System in MSNs with Flexible Revocations, Department of Computer Science & Engineering, Shanghai Jiao Tong University, School of Communication and Information Engineering, Shanghai University, China, International Journal of Distributed Sensor Networks Volume 2015, 2015.

37 Fatos Xhafa, Jingwei Li, Gansen Zhao, Jin Li, Xiaofeng Chen, Duncan S. Wong, Designing cloud-based electronic health record system with attribute-based encryption, Springer Science and Business Media New York, 2014.