Privacy Preservation for eHealth Big Data in Cloud Accessed Using Resource-Constrained Devices: Survey

Kittur Philemon Kibiwott[1], Zhang Fengli[1], Kimeli Victor K.[3], Omala A. Anyembe[2], Eugene Opoku-Mensah[1]

(Corresponding author: Kittur Philemon Kibiwott)

School of Information and Software Engineering, University of Electronic Science and Technology of China[1]
No.4 Section 2 Jianshe Rd North, Chengdu 610054, China
School of Computer Science and Engineering, University of Electronic Science and Technology of China[2]
Chengdu 611731, China
University of Eldoret[3]
Address: P. O. Box 1125-30100 Eldoret, Kenya
(Email: phkibiwott@gmail.com)

(Received Sept. 18, 2017; revised and accepted Jan. 10, 2018)

Abstract

Mobile technology is proving to offer unprecedented advantage to health professionals by providing a more efficient transmission and access to health services. However, mobile devices are resource-constrained. This is setback whenever storage and computation are required on ehealth big data. To mitigate this drawback, mobile computing is integrated with scalable cloud computing. While this is an advantage on mobile user's side by enlarging the limited resources of the device, it also gives rise to security and privacy concerns. In order to overcome these challenges associated with security and privacy, the data owner (hospital) encrypts data using Attribute Based Encryption (ABE) primitive due to the fine-grained access control advantage it offers then sends ciphertext to the cloud. To realize fast data access, the resource-constrained device securely outsources heavy computations to resource abundant cloud server on its behalf with the guarantee that the server cannot learn anything about plaintext. In this paper, a survey of ABE with outsourced decryption of the existing works that is applicable to resource-constrained device for accessing eHealth big data is provided.

Keywords: ABE; Big Data; Cloud Computing; eHealth; Mobile; Outsource; Resource-Constrained

1 Introduction

Advances in technology have led to the generation of variety of massive data from diverse sources. Consequently, traditional techniques of storage and sharing is difficult to implement. This is because the data is enormous, complex and some in unstructured format. Cloud computing is used to fill this gap. This is due to the scalability advantage it has over its traditional storage counterpart. Therefore, this means a third party will be the custodian of the data, for example a hospital can outsource its eHealth big data to the cloud to be shared with the users like government, insurance companies, patients, doctors and other hospitals. Data access control is an effective way to protect and preserve the privacy of eHealth big data and achieve confidentiality. In this case the owner of the data for instance has to encrypt the data using public key encryption (PKE), outsources it to the cloud and the intended receiver with valid decryption key corresponding to the encryption key recovers the data. But there is a limitation when encrypted data is to be decrypted by many categories of users especially in eHealth big data setting.

Identity Based Encryption (IBE) [54] scheme which regards identities as string of characters was proposed as an alternative to PKE to simplify certificate management process, hence decrease communication overhead. Drawback to this scheme is that when the volume of data is large and complex for example in the case of healthcare big data, the computation cost is high and also time consuming [42]. In the year 2005, another new kind of IBE known as Fuzzy Identity-Based Encryption (FIBE) was proposed by Sahai and Waters [53]. In FIBE, identities is regarded as set of descriptive attributes where a user that has a private key for a given set of attributes can recover the message. This was Attribute-Based Encryption (ABE) at its cradle stage. Attribute-Based Encryption

one of the public-key encryption flavors, proves to be applicable in securely accessing eHealth big data. It has emerged to be a promising access control primitive for cloud computing in the recent years. To access data in cloud [60], the data owner has to provide expressive fine-grained access control on how data is to be exchanged with the users.

In ABE scheme, the receiver's private key can decrypt a certain ciphertext only if the associated attributes and access policy correlate. There are two kinds of ABE schemes: Key-Policy ABE (KP-ABE) [4, 21, 31, 47] and Ciphertext-Policy ABE (CP-ABE) [6,20,62]. In KP-ABE scheme, ciphetexts are labeled with sets of attributes and access policies of this attributes are associated with end user's private keys. While in CP-ABE scheme, every ciphertext is associated with an access policy, and every end user's private key is associated with a set of attributes. CP-ABE is regarded as the suitable technology for data access control in cloud storage system because the the data owners defines the access policy [33]. ABE is one of the powerful and most important technology for realizing fine-grained access control of data in the cloud. However, in the majority of ABE schemes, the major drawback is the inefficiency as the size of ciphertext and decryption overhead grows with the complexity of the access policy. This is a setback to users using resource-constrained devices. To overcome this problem and therefore accommodate these devices, secured partial decryption should be carried out by the cloud.

1.1 Motivation

The emergence of smartphones and social media have further extended the usage of mobile devices that people carry these devices everywhere they go. Users *e.g.* patients can communicate with the physicians and obtain the information they require anywhere anytime without being in hospital physically. This not only saves time but also serves well during emergencies. While this is an advantage, the device lack in abundant resources. Big data is voluminous and therefore cannot be accommodated by mobile device which calls for assistance from untrusted or semi trusted unlimited resource cloud server platform for efficient storage and processing of the data. This poses security and privacy challenges as the data is stored by a third party which is outside the data owner's view.

To protect the data from leakage, the owner has to encrypt and define expressively how the data is to be accessed and shared by various users before he offloads to the cloud. When the mobile user with the required credentials wants to recover the message from the cloud, he has to borrow power from the cloud server to perform computational-intensive tasks on his behalf without the server jeopardizing the privacy of data. The overhead on end user's side is reduced significantly [22, 41]. In this paper, extensive review of secure ABE technologies with outsourced decryption suitable for resource-constrained devices to preserve the privacy of eHealth big data in cloud is provided. We will first discuss the characteristics that forms eHealth big data, then define eHealth and security issues, finally we will give comprehensive outsourced decryption technologies, proposed future work and conclusion. This work can serve as a guide to the beginners in understanding the fundamental issues in security and performance of ABE primitive with outsourced decryption.

2 eHealth

eHealth paradigm envisages the transfer of health resources and health care by electronic means. It includes information and data sharing between patients and health service providers, hospitals, health professionals and health information networks, electronic health records, telemedicine services, portable patient-monitoring devices and operating room scheduling softwares [50]. eHealth an implementation of information communication technology is currently one of the major sectors where big data explosion is experienced [38].

2.1 eHealth Big Data

Big Data is defined using 5V's: Volume, Velocity, Variety, Value and Veracity. This data originates from different multiple sources such as networked sensors, mobile devices, web logs, call centers, smartphones and social media sites such as facebook. A forecast by IDC Digital Universe for 2012-2020 reveals that digital data will swell by almost half, that is from 0.8 zettabytes to 40 zettabytes [24]. According to [26], it anticipates that by the year 2020 80% of the US healthcare service providers will have implemented Electronic Health Record (EHR) systems, 80% of the general population will have adopted Personal Health Record (PHR) systems, while 80% of PHR and EHR systems will be connected using Health Information Exchange (HIE) systems hence voluminous amount of the data will be generated. Health big data according to [37] stem from pre-hospital, in-hospital and post-hospital.

The healthcare system data volume in USA hit 150 exabytes in 2011 [1] and its projected to increase more and the drift is due to real-imaging, wearable computing devices etc. Genomic-driven study, probe-driven treatment and health management are the two major sources that generates massive amount of ehealth big data. Due to huge volume of healthcare big data (terabytes to petabytes) and its complexity, it becomes hard to store them in traditional storage. An effective alternative is to store them in cloud owing to the elastic scalability and computation advantage provided by cloud. This means a third party will be the custodian of the data. According to Arora *et al.* [2], since the cloud servers and data owners are not within the same trusted domain, then the biggest concern when big data is stored in third party is security and privacy as cloud storage is untrusted or semi-trusted.

2.2 Big Data Characteristics in eHealth

According to Gartner [17], big data is high-volume, high-velocity and high-variety information assets demanding cost effective, innovative forms of information processing to enhance insight and decision making. IBM added the fourth "V" (veracity), while Oracle added the fifth "V" (value).

1) *Volume*: Refers to massive amount of data generated from different sources in healthcare industry [46]. For example the ehealth big data can be generated by medical sensors, doctors, other hospitals, insurance companies, government etc. Utilizing Electronic Health Records (EHR) and its significant growth of the correlated healthcare related data generate increasing volume of health information. In 2012, the digital healthcare data in the entire universe was estimated to be equivalent to 500 Petabytes and its projected to attain 25,000 Petabytes by 2020 [28].

2) *Velocity*: Velocity is needed by healthcare providers and consumers for timely and proper decision making. It refers to the rate at which ehealth big data is generated, stored, analyzed and apportioned to different healthcare providers and consumers. The system should be efficient and secure as patient's data is critical.

3) *Variety*: It refers to different forms of ehealth big data. This can be in structured form *e.g.* EHR which can be easily stored by machine, unstructured, or semi-strucured can be inform of prescription, doctor's notes, images, x-ray etc.

4) *Value*: Refers to extracting meaningful data from eHealth big data. This takes place during processing of healthcare data. Extracting desired data can be used for example in research to curb the future outbreak of diseases.

5) *Veracity*: Refers to ehealth data with different quality, relevance and meaning. Since we have different forms of eHealth big data it follows that we will also have different quality of data. The quality of data has direct implication on the life of the patient. For that matter, eHealth big data quality should be reliable.

2.3 Sources of Big Data in eHealth

According to Iroju *et al.* [46] healthcare data is generated from:

- Biometric Generated Data: Biometric data is the record of data that uniquely identifies people. It originates from individuals' bodily characteristics such as facial scans, genetics, retinal scans, heart rate, blood pressure.

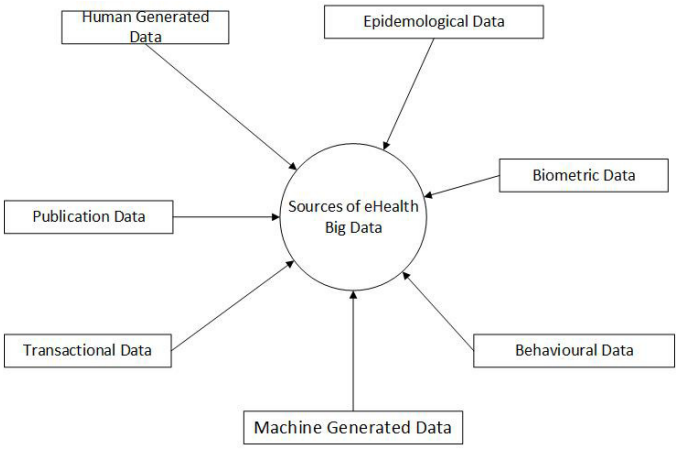

Figure 1: Sources of eHealth big data

- Transactional Generated Data:These include data emanating from healthcare individual claims and billing records. Examples include charges levied records on patients. They can be in semi structured or structured format.
- Publication Generated Data: Refers to data from health researches, medical science reference materials and government proposals data. Health research include exploratory research, descriptive research, explanatory research, and emancipatory research.
- Machine Generated Data: Refers to data that are generated by machines used in the healthcare system. Examples include data coming from remote sensors, wearable devices, x-ray machines, ECG machine, anesthesia machines etc.
- Human Generated Data: Refers to data produced by human beings in the healthcare system. It comprises unstructured and semi structured clinical data such as case prescription notes, hospital admission records, laboratory results, discharge summaries and electronic mails. Digitization of health records such as the use of structured Electronic Health Record (EHR) has also resulted to voluminous data.
- Epidemological Generated Data: These data include vital statistical data, causes of diseases, impacts, identify disease risk factors, health surveys and targets for preventive healthcare, patterns such as disease distribution in population and probe these disease causes.
- Behavioural Generated Data: Refers to data generated from social intercourses and communication tools such as websites and social media sites such as Twitter, Facebook and WhatsApp.

Figure 1 provides a summary of the sources of eHealth big data.

2.4 Information Security and Privacy for eHealth Big Data

2.4.1 eHealth Privacy

Safeguarding personal health information from disclosure, loss, unauthorized access, modification or used without patient's consent [49]. eHealth systems should be build with privacy as a priority [42]. Normally privacy is judged from the harm it causes if an individual information goes public [43]. eHealth data that demands privacy is divided into:

1) Highly risk data: Data that can cause harm to an individual.
2) Restricted data: Data covered by state or federal legislation.
3) Confidential Data: Example is patient's ID.

2.4.2 Confidentiality

Confidentiality makes certain that healthcare information provided to healthcare professions cannot leak to the third party without owner's consent. Confidentiality seeks "non inteference of information and protective actions of information such as security measures" [13].

2.4.3 Integrity

This ensures that the data/information remains unchanged. Data should only be modified by those who are authorized to do so [48]. Loss of data, data breach and the correctness of the data are the concerns as far as integrity of data is concerned when data is outsourced to the cloud.

2.4.4 Availability

The presence of eHealth information to authorized users when it's needed [45]. This permits health professionals to obtain accurate and timely health information that will add value to the patient treatment [48].

2.4.5 Authentication

Sources of eHealth data needs to be determined before its used to confirm its true originality [14]. The moment authentication process is established, it should be clearly stated what data is permitted to be accessed and the requirements for one to access them.

2.5 Access Structures [5]

These are set of qualified families that can construct a secret. Let $\mathbb{A}$ be the universe of attributes. A collection $\mathbb{P} \subseteq 2^{\mathbb{A}} \setminus \{\emptyset\}$ is monotone if $\forall \mathbb{B}, \mathbb{C}$: if $\mathbb{B} \in \mathbb{P}$ and $\mathbb{B} \subseteq \mathbb{C}$,then $\mathbb{C} \in \mathbb{P}$. An access structure is a collection $\mathbb{P}$ of non-empty subsets $\mathbb{P} \subseteq 2^{\mathbb{A}} \setminus \{\emptyset\}$. The sets in $\mathbb{P}$ are called authorized sets, and the sets not in $\mathbb{P}$ are unauthorized sets.

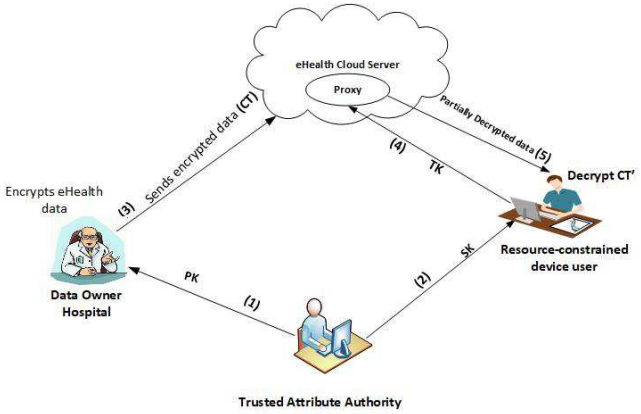

Figure 2: Secure eHealth big data access for resource-constrained devices in cloud computing

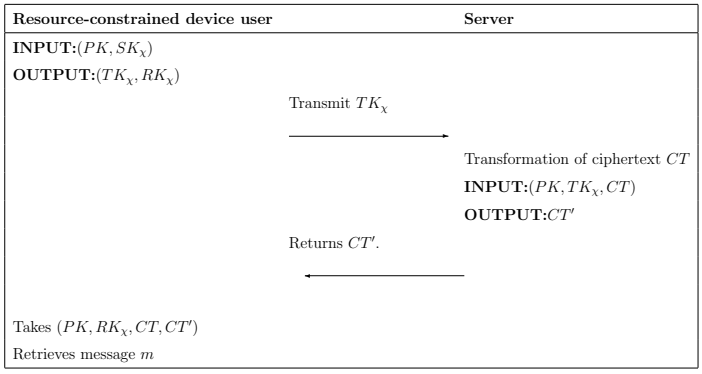

Figure 3: Outsourced decryption primitive

2.6 Bilinear Maps

Let $\mathbb{G}$,$\mathbb{G}_\tau$ be two multiplicative cyclic groups of prime order p. Let g, g_1 be $\mathbb{G}$ generators and e be a bilinear map; e: $\mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$. Bilinear map e has the following properties:

1) *Bilinearity* : $\forall g, g_1 \in \mathbb{G}$ and $a, b \in \mathbb{Z}_p^*$ we have $e(g^a, g_1^b) = e(g, g_1)^{ab} = e(g^b, g_1^a)$.
2) $Non - degeneracy : e(g, g_1) \neq 1$.
3) *Computability*: There is an efficient algorithm to compute $e(g, g_1)$.

2.7 Complexity Assumptions

The Decisional Bilinear Diffie-Hellman (DBDH) assumption: Let $x, y, z, c \in \mathbb{Z}_p$ be randomly chosen and $g \in \mathbb{G}$ be a generator. The DBDH assumption [53] holds in $\mathbb{G}$ if no probabilistic polynomial-time algorithm can distinguish the tuples $(g, g^x, g^y, g^z, e(g, g)^{xyz})$ from the tuple (g, g^x, g^y, g^z, g^c) with non-negligible advantage.
Discrete Logarithm (DL) Assumption: Let $(p, \mathbb{G}_1, \mathbb{G}_\tau, e)$ be a prime order bilinear group system. Given $(p, \mathbb{G}_1, \mathbb{G}_\tau, e,\ g, g^x)$, where $g \in \mathbb{G}$ and $x \in \mathbb{Z}_p^*$ are uniformly chosen randomly, the Discrete Logarithm problem

is to compute x. The Discrete Logarithm assumption [14] in the prime order bilinear group system $(p, \mathbb{G}_1, \mathbb{G}_\tau, e)$ is that no probabilistic polynomial-time (PPT) algorithm $\mathcal{A}$ can solve the DL problem with non-negligible advantage. The advantage of $\mathcal{A}$ is defined as

$$\Pr[\mathcal{A}(p, \mathbb{G}_1, \mathbb{G}_\tau, e, g, g^x) = x],$$

where probability space is over g, x chosen randomly and random bits consumed by $\mathcal{A}$.

Decisional Bilinear Diffie-Hellman Exponent (DBDHE) assumption: Let α, t be randomly chosen and $g \in \mathbb{G}$ be a generator. The decisional q-DBDHE assumption [8] is that no probabilistic polynomial-time algorithm $\mathbb{AD}$ can distinguish the tuple $e(g,g)^{\alpha^{q+1}} \in \mathbb{G}_\tau$ from a random element $C \in \mathbb{G}_\tau$ with more than a non-negligible advantage provided ϵ=(g, g_q, g_{q+2},...,g_{2q}, g^t), where g_I is denoting g^{α^I}. Advantage of algorithm $\mathbb{BD}$ solving decisional q-BDHE is:

$$|Pr[\mathbb{BD}(\epsilon, V = e(g,g)^{\alpha^{q+1}}) = 0] - Pr[\mathbb{BD}(\epsilon, V = C) = 0]| \geq \epsilon .$$

Decisional modifed Bilinear Diffie-Hellman (MBDH) assumption: Suppose a challenger $\mathbb{BD}$ randomly selects $\alpha, \beta, \gamma, z \in \mathbb{Z}_p$. Decisional modified bilinear Diffie-Hellman (MBDH) [65] is that no probabilistic polynomial-time algorithm $\mathbb{AD}$ can distinguish the tuple $(g, g^\alpha, g^\beta, g^\gamma, (g,g)^{\alpha\beta/\gamma})$ from $(g, g^\alpha, g^\beta, g^\gamma, (g,g)^z)$ with non-negligible advantage.

2.8 Formal Structure for Generic ABE

The intuition to this basic ABE is that the intended user also referred to as legitimate user with the given set of attributes specified in the access policy at the time of encryption can access the data.

An ABE consists of four basic algorithms [53] as follows:

- *Setup*: This is a probabilistic algorithm, executed by trusted attribute authority. Takes as input security parameter γ and outputs a pair (PK, MSK). Where PK is public parameter while MSK is master secret key.
$$(PK, MSK) \leftarrow Setup(1^\gamma).$$
- *KeyGen*: This algorithm is executed by trusted attribute authority to produce secret/private key. The input to the key generation algorithm is a set of attributes χ, master secret key MSK and public key PK. It yields private/decryption key SK_χ.
$$SK_\chi \leftarrow KeyGen(\chi, MSK, PK).$$
- *Encrypt*: This is a probabilistic algorithm, executed by data owner (sender). Takes as input the message m to be encrypted, the set of attributes χ and public key PK. It yields as output ciphertext CT.
$$CT \leftarrow Encrypt(PK, \chi, m).$$
- *Decrypt*: This algorithm is deterministic and its is executed by the intended user/decryptor. Takes as input ciphertext CT, public key PK, and private key SK satisfied by the set of attributes. It returns as an output a message m.
$$m \leftarrow Decrypt(PK, SK_\chi, CT).$$

Correctness.

$$m \leftarrow Decrypt(Encrypt(PK, \chi, m), PK, SK_\chi).$$

3 ABE with Outsourced Decryption for eHealth Big Data

3.1 Formal Structure for ABE with Outsourced Decryption

The intuition of the ABE with outsourced decryption primitive is that an authorized mobile user that possesses a given set of attributes satisfying the access structure can securely access the data. The hospital encrypts the data and specifies the access policy then sends it to the cloud. When a mobile device user having required set of attributes wants to access the data it will first sends a transformation key to the proxy to perform heavy computation overhead such as compute pairings on his/her behalf. Transformed ciphertext will then be sent to the end user. Limited-resource device user will perform final decryption. Using this primitive improves performance of resource-constrained device. The complete system is as shown in Figure 2.

Attribute Based Encryption with outsourcing decryption for eHealth big data system has the following five entities:

1) *Trusted Attribute Authority*: This is the only trusted entity. It generates public and private keys and parameters for Attribute Based Encryption scheme. Public key is used for encryption of data and private key is used for decryption to recover the original message. Resource-constrained device users also receive from trusted attribute authorities their attributes that corresponds to decryption keys.

2) *Hospital*: It's is the owner of the data. Prior to outsourcing the data to the cloud, it defines how data is to be accessed by authorized users then encrypt it under given access policy.

3) *eHealth cloud*: It is semi-trusted or untrusted entity. it has unlimited resources and therefore provides storage facilities, high computation power and access for eHealth big data.

4) *Proxy*: It interacts with resource-constrained device users. It transforms efficiently using blinding key the encrypted data into a simple ciphertext without learning the plaintext of the data.

5) *Resource-constrained device users*: Authorized entities possessing a set of attributes that satisfies the access policy of the encrypted data can decrypt and recover the message. They receive decryption keys that corresponds to their attributes from Trusted Attribute Authorities. In our setting users can be government, medical research organizations, insurance companies, other hospitals etc.

An ABE with outsourced decryption has the following basic algorithms [22, 36]:

- *Setup*: This is a randomized algorithm, executed by trusted attribute authority. Takes security parameter η as input and produces as output a pair (PK, MSK). Where PK is public parameter while MSK is master secret key.

$$(PK, MSK) \leftarrow Setup(1^{\eta}).$$

- *KeyGen*: This is randomized algorithm executed by trusted attribute authority to yield private key. The input to the key generation algorithm is a set of attributes χ, master secret key MSK and public parameter PK. It produces as output private/decryption key SK_{χ}.

$$SK_{\chi} \leftarrow KeyGen(\chi, MSK, PK).$$

- *Encrypt*: This is a probabilistic algorithm, executed by the owner of the data (sender). Takes a message m to be encrypted, the set of attributes χ and public key PK as inputs. It yields as ciphertext CT output.

$$CT \leftarrow Encrypt(PK, \chi, m).$$

- *Decrypt*: This algorithm is deterministic and its is executed by the intended data user/decryptor. Takes ciphertext CT, public key PK, and private key SK satisfied by the set of attributes as inputs. Returns as output message m.

$$m \leftarrow Decrypt(PK, SK_{\chi}, CT).$$

- $Transform_{out}$: This algorithm is executed by resource-constraint user as shown in figure 3. It takes transformation key TK_{χ} and ciphertext CT as input. It returns CT'.

$$CT' \leftarrow Transform_{out}(TK_{\chi}, CT).$$

- $Decrypt_{out}$: The algorithm is executed by intended resource-constrained device user. It takes retrieving key RK_{χ} and transformed ciphertext CT$'$ as input. Returns message m.

$$m \leftarrow Decrypt_{out}(RK_{\chi}, CT').$$

Correctness.

1) $Decrypt(PK, SK_{\chi}, Encrypt(PK, m, \mathbb{A})) = m$
2) $Decrypt_{out}(PK, RK_{\chi}, Transform_{out}(Encrypt(PK, M, \mathbb{A}), PK, \mathrm{TK}_{\chi})) = m$

3.2 Adversary Models

While proposing any system, security is the primary consideration. Users of the system should be convinced that it is secure enough to trust their data into it. Parties involved in the system, rarely trust each other but all have one common thing, they all trust the protocol proposed. Accessing the contents to which they are not authorized to is the main objective of any adversary. They may collude with others or work independently [15]. Some of the system threats in ABE with outsourced decryption originate from [15, 22, 29, 32];

1) *Semi-trusted/untrusted cloud servers colluding with unauthorized users*: It is assumed that the servers provides their services smoothly but may at times be curious of leaking sensitive information such as a patient data to illegitimate users.

2) *Attribute authorities (AA)*: AA may willingly violate data owners by conspiring with the outsiders where they issue them with keys to enable them access unauthorized data.

3) *Legitimate users colluding with each other*: Authorized users can combine their attributes to access unauthorized data which individually could not access.

4) *Replay attacks*: Legitimate users can re-submit the previous tokens to the servers to obtain unauthorized data.

5) *Active attacks*: Unauthorized users may introduce malicious data into the cloud to harvest some data or corrupt them.

6) *Servers colluding with authorized users*: Cloud server may collude with authorized users to obtain unauthorized data for the purpose of using them for their malicious end.

3.3 Security Models

Definition of security of any cryptosystem is based on what is to be achieved and a particular attack. There are three known security models common to all cryptosystems. In order of security strength they are adaptive chosen ciphertext attack (CCA2) [52] (stronger), non-adaptive chosen ciphertext attack (CCA1) [44] and chosen plaintext attack (CPA) [19]. Where CCA2 security model is more secure.

Security goals: To realize fully the benefits of ABE with outsourced decryption for eHealth big data, the following security goals should be met [30, 41, 55]:

1) *Fine-grained access control*: Hospital (owner of data) should be in a position to safeguard its sensitive information using strong security measures. Only legitimate users with the set of attributes defined in the

access policy, manages to retrieve the stored data in the cloud.

2) *Efficient encryption/decryption*: Legitimate users should be able to access the stored information without delay as the health big data is so sensitive as it deals with people's lives.

3) *Collusion resistance*: Since data custodian is the third party which is semi-trusted / untrusted in nature, collusion among different unauthorized parties should be thwarted for example collusion between cloud and illegitimate users to acquire private keys to access unauthorized data which individually could not access should be prevented.

4) *Confidentiality of data*: eHealth big data should not get leaked to the third party without owner's approval.

5) *Convenience*: With the proliferation of mobile devices which is resource-constrained in nature, and which are incapable to finish decryption independently or consumes much time to finish decryption, utilizing outsourcing decryption enables the legitimate users to access a bulk of eHealth data anywhere anytime since heavy computations is offloaded to cloud.

6) *Unidirectional*: The main property of outsourced decryption is unidirectionality. This means the server has capability only to transform original (*e.g.* alice's) ciphertext to another (*e.g.* Bob's) ciphertext and not in reverse direction.

7) *Public verifability*: Information should be available that enables involved parties i.e users to confirm/verify the genuineness of original ciphertext and the transformed ciphertext.

8) *Immediate revocation*: Mischievous/malicious users should easily and completely be revoked from all future data access.

9) *User/ciphertext anonymity*: Disclosure of user's identity or key privacy also referred to as ciphertext anonymity should not be revealed. It should hide users/ciphertext identities.

10) *Scalability*: With the increase of legitimate users, the system efficiency is still guaranteed. The performance of the system cannot be affected by number of legitimate users.

4 State-of-the-art of ABE with Outsourced Decryption for Resource-Constrained Devices

The leading efficiency drawback in the vast majority of ABE is the increase in size of the ciphertext and the decryption cost (computational cost) with the increase in complexity of the access policy. The applications executing in mobile devices which are resource-constrained in nature in terms of battery life, computational resources, storage, and bandwidth may have to hold on for a long time or even abort before execution to finality. Using these devices to access eHealth big data is not faster enough as it is costly due to bilinear pairing operations involved. To curb this limitations, the remedy is to adopt mobile cloud computing where heavy computations are offloaded to the cloud [18, 64].

The first ciphertext-policy attribute-based PRE (CP-ABPRE) scheme, in which a cloud server is authorized to transform a ciphertext under a specified access structure (represented only as AND gates on positive and negative attributes) into the one under another access structure was proposed in 2009 [35]. In this scheme, the user successfully decrypts the ciphertext if and only if the set of positive and negative attributes are embedded in the access structure.

To minimize the number of pairing operations on end users side and hence reduce decryption overhead, ABE with outsourced decryption schemes is proposed where intensive computational tasks is outsourced to cloud service providers [33, 34, 61]. The schemes proposed by Green *et al.* [22] and De *et al.* [12] provides fine-grained access control solution to the lightweight devices such as mobile phones with constrained computing resources which independently cannot successfully execute basic encryption/decryption while protecting sensitive data outsourced to the public cloud [32]. Scheme of [22] achieved CPA-security which was later extended to achieve the stronger RCCA-security in random oracle model. With the provision of outsourced decryption, heavy computations and storage can be offloaded and the light computations be performed by resource-constrained devices effectively and efficiently. Generally in [22], the key generating algorithm is designed to output two key pair to the data user as follows:

1) A short El Gamal kind private key known as retrieving key rk.

2) Its paired key known as transformation key tk, which is send to the server and its publicly known.

In this scheme a key for blinding (ie transformation key) tk is sent to the third-party (server) for translation of any ciphertext CT satisfied by end user's attributes or access policy into a simpler ciphertext CT$'$. The end user incurs minimal overhead to recover plaintext from transformed ciphertext. The major drawback to this technique is that a ciphertext can be mutated on the transit therefore making the user unable to realize and detect the change. For the case of eHealth it can lead to wrong diagnosis hence cause fatal consequences.

In addressing the same ABE computational problem, where cost grows linearly with respect to complexity of the number of attributes or ciphertext policy and which

Table 1: Comparison of the schemes characteristics with outsourced decryption

Scheme	Characteristics							
	Fine-grained access	Efficiency	Collusion resistance	Confidentiality	Unidirectionality	Verifiability	Immediate revocation	Scalability
Zhou [64]	√	√	√	√	√	×	×	√
Li [34]	√	√	√	√	√	×	√	√
Lai [29]	√	√	√	√	√	√	×	√
Green [22]	√	√	×	√	√	×	×	√
Li [32]	√	√	√	√	√	√	×	√
Jiguo 16 [33]	√	√	√	√	√	×	√	√
Lin [36]	√	√	×	√	√	√	×	√
Mao [41]	√	√	√	√	√	√	×	√
Jiguo 17 [27]	√	√	×	√	√	√	×	√
Zechao [39]	√	√	×	√	√	√	×	√

Abbreviations: √: Scheme supports the corresponding characteristic, ×: Scheme does not support the corresponding characteristic.

Table 2: Comparison of the security models of ABE schemes with outsourced decryption

Scheme	Security	Model	Complexity assumption
Zhou [64]	IND-CPA	adaptive	Co-DBDH
Li [34]	RCCA	selective	DBDH
Lai [29]	RCCA	selective	DL
Green [22]	CPA	selective	Decisional q-BDHE
Li [32]	RCCA	selective	DBDH
Jiguo 16 [33]	CPA	selective	DCDH
Lin [36]	IND-CPA	adaptive	DL
Mao [41]	IND-CPA	selective	Generic group
Jiguo 17 [27]	IND-CPA	selective	DL
Zechao [39]	IND-CPA	selective	DON'T EXIST

Abbreviations: $DBDH$: Decisional Bilinear Diffie-Hellman, DL: Discrete Logarithm, $q-BDHE$: Decisional Bilinear Diffie-Hellman Exponent, $Co-DBDH$: Co-Decisional Bilinear Diffie-Hellman, $DCDH$-Divisible Computation Diffie-Hellman, $IND-CPA$:Indistinguishable Chosen Plaintext Attack, $RCCA$:Replayable Chosen Ciphertext Attack.

is a bottleneck to resource-constrained devices such as mobile devices, Zechao *et al.* [39] proposed a new CP-ABE scheme known as Offline/online attribute-based encryption with verifiable outsourced decryption protocol. In this scheme, offline/online technique is combined with the outsourced verifiable computation technique using bilinear groups, which supports both offline/online generation of key and encryption, as well as the verifying outsourced decryption. Heavy computation during key generation is executed offline, and encryption can be split into two phases offline and online where heavy tasks is executed offline and lightweight tasks are executed online efficiently. On the other hand, decryption workload is offloaded to the server. To overcome the disadvantage of complexity, Song *et al.* [40] extended the scheme of Emura *et al.* [16] scheme. In [40], an alternative technique is proposed where decryption process is made faster by making use of only a constant number of bilinear operations. The decryption cost and ciphertext length are decreased significantly in comparison with previous protocols. While addressing the same issue of defining the access structure policy and allowing the owner to outsource intensive computation tasks to cloud server providers, Yu *et al.* [61] integrated techniques of Attribute-Based Encryption (ABE), proxy re-encryption and lazy re-encryption. In this scheme, the data owner enforces a unique access structure on each user. While executing this protocol, the cloud servers is prevented from learning about the underlying plaintext.

The bilinear pairings computation is considered to be the most costly operation experienced in pairing-based cryptographic protocols construction. In order to execute the protocol with pairing and hence accommodate devices with limited resources, Benot *et al.* [11] proposed secure delegation of elliptic-curve pairing by resource-constrained device to a more powerful device. Pairing $e(X,Y)$ for example is delegated to a more powerful device (for instance a PC). Delegation is done in such a way that a powerful device cannot learn about the points X and Y. To verify the output and confirm whether the terminal is cheating, the resource-constrained device either yields the correct output or nothing with overwhelming accuracy. However, the drawback to this scheme is that the resource-constrained device restricts itself to a simple

Table 3: Comparison of the storage overhead of schemes with outsourced decryption

Scheme	Length of the Key					Ciphertext	
	Public Key	Master Key	Private Key	Transform Key	Retrieval Key	Original ciphertext	Transformed ciphertext
Zhou [64]	$2\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$	$\|\mathbb{G}\|+\|\mathbb{Z}_p\|$	$(2N+3)\|\mathbb{G}\|$	$(2N+3)\|\mathbb{G}\|$	$\|\mathbb{Z}_p\|$	$(2N+1)\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$	$(4N+2)\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$
Li [34]	$(N+4)\|\mathbb{G}\|$	$\|\mathbb{Z}_p\|$	$(2N+5)\|\mathbb{G}\|$	$2N\|\mathbb{G}\|$	$2\|\mathbb{G}\|$	$(N+2)\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$	$2\|\mathbb{G}\|+2\|\mathbb{G}_\tau\|$
Lai [29]	$(N+5)\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$	$\|\mathbb{Z}_p\|$	$(N+2)\|\mathbb{G}\|$	$(N+2)\|\mathbb{G}\|$	$\|\mathbb{Z}_p\|$	$(4N+3)\|\mathbb{G}\|+2\|\mathbb{G}_\tau\|$	$\|\mathbb{G}\|+4\|\mathbb{G}_\tau\|$
Green [22]	$2\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$	$\|\mathbb{Z}_p\|$	$(N+2)\|\mathbb{G}\|$	$(N+2)\|\mathbb{G}\|$	$\|\mathbb{Z}_p\|$	$(2N+1)\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$	$2\|\mathbb{G}_\tau\|$
Li [32]	$(N+4)\|\mathbb{G}\|$	$\|\mathbb{Z}_p\|$	$(2N+3)\|\mathbb{G}\|$	$(2N+3)\|\mathbb{G}\|$	$\|\mathbb{Z}_p\|$	$(N+2)\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$	$\|\mathbb{G}_\tau\|$
Jiguo 16 [33]	$2\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$	$\|\mathbb{G}\|+\|\mathbb{Z}_p\|$	$(2N+3)\|\mathbb{G}\|$	$(2N+3)\|\mathbb{G}\|$	$\|\mathbb{Z}_p\|$	$(2N+4)\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$	$5\|\mathbb{G}_\tau\|$
Lin [36]	$(N+4)\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$	$\|\mathbb{Z}_p\|$	$(N+2)\|\mathbb{G}\|$	$(N+2)\|\mathbb{G}\|$	$\|\mathbb{Z}_p\|$	$(2N+2)\|\mathbb{G}\|$	$\|\mathbb{G}_\tau\|$
Mao [41]	$(N+4)\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$	$\|\mathbb{Z}_p\|$	$(N+2)\|\mathbb{G}\|$	$(N+2)\|\mathbb{G}\|$	$\|\mathbb{Z}_p\|$	$(2N+2)\|\mathbb{G}\|$	$\|\mathbb{G}_\tau\|$
Jiguo 17 [27]	$(N+5)\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$	$(N+1)\|\mathbb{Z}_p\|$	$2\|\mathbb{G}\|$	$2\|\mathbb{G}\|$	$\|\mathbb{Z}_p\|$	$5\|\mathbb{G}\|+2\|\mathbb{G}_\tau\|$	$\|\mathbb{G}\|+4\|\mathbb{G}_\tau\|$
Zechao [39]	$5\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$	$5\|\mathbb{G}\|+\|\mathbb{G}_\tau\|+\|\mathbb{Z}_p\|$	$(2k+5)\|\mathbb{G}\|+3k\|\mathbb{Z}_p\|$	$(2k+5)\|\mathbb{G}\|+3k\|\mathbb{Z}_p\|$	$\|\mathbb{Z}_p\|$	$(1+3l)\|\mathbb{G}\|+3l\|\mathbb{Z}_p\|$	$3\|\mathbb{G}\|+\|\mathbb{G}_\tau\|$

Abbreviations: N: Attribute size, $|\mathbb{G}|$: bit length of an element in $\mathbb{G}$, $|\mathbb{G}_\tau|$: bit length of an element in $\mathbb{G}_\tau$, $|\mathbb{Z}_p|$: bit length of an element in $\mathbb{Z}_p$.

curve or provided field operations. Not flexible enough to support complex curves.

Muhammad *et al.* [3], proposed attribute-based encryption with encryption and decryption outsourcing that reduces the computational load on both the host and the users using devices that are computationally resource-constrained (*e.g.*. mobile devices). The scheme is comprised of two proxies which are independent and cannot collude, one on the host side and the other one on the user's side. In the former, data owner is allowed to outsource cryptographic creation policy to semi-trusted proxy. The proxy is unable to learn about encrypted messages and is enforced to encrypt the messages based on the policy specified on the attributes. While in the latter, the heavy computation overhead during decryption is reduced by allowing a user to offload the verification policy onto another semi-trusted proxy where it borrows power from the proxy to verify the policy using the user's key transformation attributes. This scheme is provable secure under the generic group model.

To ensure the server legitimately executes outsourced decryption, a number of schemes have been proposed [27, 29, 36, 41, 59]. Lai *et al.* [29] and Mao *et al.* [41] separately introduced verifiability primitive in the outsourced decryption. To accomplish this, an extra instance is added to the existing ABE in encryption/decryption algorithm phases. A drawback to the scheme is that owner of data has to perform an extra work of encrypting the random message then compute checksum value corresponding to two messages. As a result, computation and communication overhead are duplicated. To overcome this drawback, Lin *et al.* [36] proposed a more efficient ABE with verifiable outsourced decryption based on an attribute-based key encapsulation mechanism, a symmetric-key encryption scheme and a commitment scheme in generic model. The scheme in [36] can be considered both in Key-Policy Attribute-Based Encryption (KP-ABE) and also in Ciphertext-Policy Attribute-Based encryption (CP-ABE) settings. Solution to checking the integrity of outsourced data while maintaining privacy and secrecy of the stored data was proposed by Yadav *et al.* [59]. In this scheme, secure operations in data storage can be augmented to provide remote integrity checking. It is carried out by computing just once the hash of data, and therefore mobile user does not need to posses outsourced data. However, from all this primitives the number of the attributes grows linearly with the length of the ciphertext and the size of costly pairing computations. This greatly affects outsourced CP-ABE scheme by limiting verifiability. To avoid this, Jiguo *et al.* [27] proposed Verifiable Outsourced Decryption of Attribute-Based Encryption with Constant Ciphertext Length that saves the communication cost.

Avoiding the distruption of the medical information system while simultaneously achieving fine-grained, privacy and confidentiality properties, Junbeom *et al.* [23] proposed a scheme that is key escrow resilient and which allows the partial decryption of the encrypted medical data by device controller without leaking any private information to the controller. This improves computational efficiency of the medical devices where most of the laborious decryption tasks is delegated to the device controller. However, the schemes do not achieve checkability on the output returned, therefore there is no guarantee of the accuracy of the partial decrypted ciphertext. To provide a solution to this, a fine-grained, multiparty access control with outsourcing decryption, was proposed by Qinlong *et al.* [51] where Cloud Service provider (CSP) can transform original ciphertext defined under access policy to another

Table 4: Comparison of the computation cost of schemes with outsourced decryption

Scheme	Computation cost				Access structure
	Encrypt	Transform$_{out}$	Decrypt	Decrypt$_{out}$	
Zhou [64]	$(2\|A_{ct}\|+1)\mathbb{G}+2\mathbb{G}_\tau$	$(2\|A_{ct}\|+1)\mathbb{G}$	DONT EXIST	$2\|A_{ct'}\|C_p+4\mathbb{G}_\tau$	Threshold
Li [34]	$C_p+(2\|A_{ct}\|+3)\mathbb{G}+2\mathbb{G}_\tau$	$2(\|A_{ct'}\|-1)C_p+2\|A_{ct'}\|\mathbb{G}_\tau$	DONT EXIST	$2C_p+3\mathbb{G}_\tau$	(t,n)-Threshold
Lai [29]	$(8\|A_{ct}\|+10)\mathbb{G}+4\mathbb{G}_\tau+2H$	$4(\|A_{ct'}\|-2)C_p+(4\|A_{ct'}\|-2)\mathbb{G}_\tau$	$4(\|A_{ct'}\|-1)C_p+4A_{ct'}\mathbb{G}_\tau$	$4\mathbb{G}_\tau$	LSSS
Green [22]	$(4\|A_{ct}\|+1)\mathbb{G}+3\mathbb{G}_\tau+\|A_{ct}\|H$	$(\|A_{ct'}\|+2)C_p+3(\|A_{ct'}\|-1)\mathbb{G}+(\|A_{ct'}\|+1)\mathbb{G}_\tau$	DONT EXIST	$2\mathbb{G}_\tau$	LSSS
Li [32]	$C_p+(2\|A_{ct}\|+3)\mathbb{G}+2\mathbb{G}_\tau$	$2(\|A_{ct'}\|-1)C_p+2\|A_{ct'}\|\mathbb{G}_\tau$	DONT EXIST	$2\mathbb{G}_\tau$	Threshold
Jiguo 16 [33]	$(2\|2A_{ct'}\|+6)\mathbb{G}+3\mathbb{G}_\tau+2H$	$2C_p+2\mathbb{G}_\tau$	DONT EXIST	$6\mathbb{G}_\tau$	Tree
Lin [36]	$4(\|A_{ct'}\|+6)\mathbb{G}+\mathbb{G}_\tau+2H$	$2(\|A_{ct'}\|-1)C_p+2(\|A_{ct'}\|-1)\mathbb{G}_\tau$	$2(\|A_{ct'}\|-1)C_p+2A_{ct'}\mathbb{G}_\tau$	$\mathbb{G}_\tau$	LSSS
Mao [41]	$(2\|A_{ct'}\|+8)\mathbb{G}+3H$	$2(\|A_{ct'}\|-1)C_p+2(\|A_{ct'}\|-1)\mathbb{G}_\tau$	$2(\|A_{ct'}\|-1)C_p+2A_{ct'}\mathbb{G}_\tau$	$\mathbb{G}_\tau$	LSSS
Jiguo 17 [27]	$(2\|A_{ct'}\|+6)\mathbb{G}+4\mathbb{G}_\tau+2H$	$4C_p+2\mathbb{G}_\tau$	$4C_p+4\mathbb{G}_\tau$	$4\mathbb{G}_\tau$	non-monotonic AND gate
Zechao [39]	$(3\|A_{ct'}\|+1)\mathbb{G}+2H$	$4C_p+\mathbb{G}_\tau$	$4C_p$	$\mathbb{G}_\tau$	LSSS

Abbreviations: $|A_{ct}|$: Attributes size that belongs to original ciphertext, $|A_{ct'}|$: Attributes size that belongs to transformed ciphertext, C_p: Bilinear pairing operation, $\mathbb{G}$: Group, $\mathbb{G}_\tau$: Target group, H: Hash function.

simpler ciphertext by making use of attribute-based proxy re-encryption. According to Qinlong *et al.* [51], the user utilizing symmetric encryption algorithm encrypts data with random data encryption key. The data encryption key is then encrypted by employing access policy. To reduce computation cost, most of decryption operations are delegated to the CSP. To ensure the output returned from the service provider is correct, checkability is provided to guarantee accurateness of the outsourced/partial decrypted ciphertext. A major drawback to this scheme is inefficiency and its inflexibility when the owner of data for example hospital needs to select some but not all of the data are to be published by particular users. To overcome this, Weng *et al.* [57] proposed a scheme that insures that only ciphertexts which satisfies a stated condition can be re-encrypted. Keywords are the only conditions utilized in this scheme, on the other hand it is not practical in real life applications.

Realizing data privacy is the primary focus when designing any cryptographic scheme. To achieve both forward security and backward security, Xiao *et al.* [58] proposed a scheme that supports efficient outsourced decryption, user revocation and dynamic entry/exit of attribute authorities. In their schemes, user revocation is only related to revoked user. To improve efficiency of accessing remote data and preserve the privacy of the user's identity, Wang *et al.* [56] proposed the first anonymous distributed fine-grained access control protocol with verifiable outsourced decryption in the public cloud which is multi-authority in nature. Later, Camenisch *et al.* [10] proposed scheme is employed which in addition to anonymity employs pseudonym technique where legitimate user's public/private key pair corresponding to pseudonym is generated.

Generation of private key for user's policy can also increase computation overhead. To reduce such local overhead, Li *et al.* in [32] proposed a scheme where attribute authority can outsource partial private key generation to a key generation service provider (KGSP). From this scheme, constant efficiency is achieved at both attribute authority and user's end side. In order to avoid incorrect output, checkability is performed on the outcome returned from KGSP.

ABE computational overhead from exponentiation at user's end side can be relieved by adopting the traditional approach to utilize server-aided techniques [7,25]. But the notable common drawback to these schemes is that by directly utilizing this schemes in ABE, it may not work efficiently [32], and to mitigate this challenge, Zhou *et al.* [64] proposed the ABE scheme which allows secure outsourcing of both the encryption and decryption to cloud service providers.

To bring data close to the user, fog computing [9] was proposed which is an extension of cloud computing. The work related to CP-ABE in fog computing with outsourced decryption have been proposed [63,65]. A system with both outsourced encryption and decryption capabilities in fog computing using CP-ABE was proposed in [63]. In this scheme the workload operations of encryption and decryption are offloaded to the fog nodes.Therefore the computation operations on the data owner's side during encryption and also on users side during decryption are not relevant to the attributes size in the access structure and private keys respectively. Since the update concentrates only on the ciphertext associated with the corresponding updated attribute, the cost incurred by at-

tribute update is minimal and hence efficient.

5 Performance and Security Analysis Comparisons

In this section, comparisons is made of the existing works to analyse the goals and the efficiency costs of the schemes. Table 1 features the comparisons of goals achieved by respective schemes, whereas Table 2, Table 3 and Table 4 highlights comparison of security models with their complexities, efficiency cost from storage perspective and computation efficiency cost for the corresponding schemes against one another respectively. We have used the following notations: N: Attribute size; $\mathbb{G}$: Group; $\mathbb{G}_\tau$: Target group; $|\mathbb{G}|$: bit length of an element in $\mathbb{G}$; $|\mathbb{G}_\tau|$: bit length of an element in $\mathbb{G}_\tau$; $|\mathbb{Z}_p|$: bit length of an element in $\mathbb{Z}_p$; H: Hash function; $|A_{ct}|$: Attributes size that belongs to original ciphertext; $|A_{ct'}|$: Attributes size that belongs to transformed ciphertext; C_p: Bilinear pairing operation.

From the output of Table 1 it can be seen that since each scheme is constructed to realize a given security goal, therefore none of the schemes can achieve all the goals concurrently. However, it shows that all the schemes supports fine-grained access, efficiency, unidirectionality, confidentiality and scalability characteristics while schemes [29, 32, 34, 41, 64] supports collusion resistance. Verifiability is supported by schemes [27, 29, 32, 36, 39, 41] whereas schemes [33, 34] supports immediate revocation.

Regarding the comparisons of security models as depicted by Table 2, it shows that more than half of the models are selective except schemes [36, 64] which are adaptive. Selective security means the initialization phase comes prior to setup algorithm. In this case, the adversary initially provides the challenger with access structure.

Consequently, in Table 3, compared to other schemes Jiguo *et al.*'s scheme [27] is ideal in terms of key length, while Lin *et al.*'s scheme [36] and Mao *et al.*'s scheme [41] are ideal in terms of ciphertext length. While in Table 4, half of the schemes do not have decrypt algorithm and among their counterparts which have decrypt algorithms, Zechao *et al.*'s scheme [39] has an ideal computation cost.

6 Proposed Future Work

6.1 Accelerating the Efficiency of Attribute-Based Encryption Schemes without Using Outsourced Decryption

Nearly all the existing ABE schemes utilize bilinear pairings as a building block for a useful algorithm construction. However, bilinear pairing has high computational overhead, which makes algorithms complex, costly and therefore inefficient. Building pairing free algorithms or reducing the bilinear pairing size operations improves efficiency by simplifying computation complexity in resource-constrained devices. In addition, employing technologies like lattice to build an ABE scheme can also improve the computational efficiency for resource-constrained device users.

6.2 Reducing Communication Cost

In ABE with outsourced decryption, the resource-constrained users usually sends a transformation key to the unlimited-resource server (proxy) to simplify the ciphertext. The original message is then recovered from the simple ciphertext by the user. This increases the communication overhead between user and proxy. To minimize the overhead, ABE schemes that does not require the user to send the blinding (transformation) key to the proxy before performing final decryption should be built.

7 Conclusion

With the proliferation of mobile devices and development of easy to use application softwares, ABE schemes with outsourced decryption is gaining popularity due to advantages it has that supports devices with limited resource capabilities. By utilizing this primitive, encrypted eHealth big data stored in the unlimited resource cloud can now be accessed by resource-constrained devices where the user has to request a cloud server to perform heavy computations overhead on his/her behalf without learning about the plaintext of the data stored. This paper provides a survey of ABE schemes with outsourced decryption by reviewing the characteristics of eHealth big data, sources of eHealth big data, design structure, investigating adversary and security models and finally comparing efficiency costs for various existing schemes.

Lastly, based on this survey, the future work was proposed to provide roadmap to the solution of the problem encountered during outsourcing decryption.

Acknowledgments

The authors acknowledges the support by National Natural Science Foundation of China under Grant No. 61602097, Sichuan Science-Technology Support Plan Program under Grant Nos. 2016GZ0065, 2015JY0178, 2016ZC2575 and 17ZA0322, Fundamental Research Funds for the Central Universities under Grant No. ZYGX2015J072.

References

[1] J. Andreu-Perez, C. C. Y. Poon, R. D. Merrifield, S. T. C. Wong, and G. Z. Yang, "Big data for health," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, pp. 1193–1208, July 2015.

[2] D. Arora and P. Malik, "Analytics: Key to go from generating big data to deriving business value," in *IEEE First International Conference on Big Data Computing Service and Applications*, pp. 446–452, Mar. 2015.

[3] M. Asim, M. Petkovic, and T. Ignatenko, "Attribute-based encryption with encryption and decryption outsourcing," in *Department of Mathematics and Computer Science*, pp. 21-28, 2014.

[4] N. Attrapadung, B. Libert, and E. de Panafieu, *Expressive Key-Policy Attribute-Based Encryption with Constant-Size Ciphertexts*, pp. 90–108, 2011.

[5] A. Beimel, *Secure Schemes for Secret Sharing and Key Distribution*, Ph.D. dissertation,Technion-Israel Institute of technology, Faculty of computer science, 1996.

[6] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *IEEE Symposium on Security and Privacy (SP'07)*, pp. 321–334, May 2007.

[7] K. Bicakci and N. Baykal, *Server Assisted Signatures Revisited*, pp. 143–156, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[8] D. Boneh, X. Boyen, and E. J. Goh, *Hierarchical Identity Based Encryption with Constant Size Ciphertext*, pp. 440–456. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[9] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in *Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing (MCC'12)*, pp. 13–16, 2012.

[10] J. Camenisch and A. Lysyanskaya, "An efficient system for non-transferable anonymous credentials with optional anonymity revocation," in *International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 93-118, 2001.

[11] B. Chevallier-Mames, J. S. Coron, N. McCullagh, D. Naccache, and M. Scott, "Secure Delegation of Elliptic-Curve Pairing," in *Smart Card Research and Advanced Application*, vol. 6035, pp. 24–35, 2010.

[12] S. J. De and S. Ruj, "Decentralized access control on data in the cloud with fast encryption and outsourced decryption," in *IEEE Global Communications Conference (GLOBECOM'15)*, pp. 1–6, Dec. 2015.

[13] P. Lobato, de Faria and J. V. Cordeiro, "Health data privacy and confidentiality rights: Crisis or redemption?," *Revista Portuguesa de Saude Publica*, vol. 32, no. 2, pp. 123–133, 2014.

[14] H. Delfs, H. Knebl, and H. Knebl, "Introduction to cryptography," *Principles and Applications*, vol. 2, 2002.

[15] X. Dong, J. Yu, Y. Luo, Y. Chen, G. Xue, and M. Li, "Achieving an effective, scalable and privacy-preserving data sharing service in cloud computing," *Computers & Security*, vol. 42, pp. 151–164, 2014.

[16] K. Emura, A. Miyaji, A. Nomura, K. Omote, and M. Soshi, "A ciphertext-policy attribute-based encryption scheme with constant ciphertext length," in *Proceedings of the 5th International Conference on Information Security Practice and Experience (ISPEC'09)*, pp. 13–23, 2009.

[17] Gartner, "Gartner's three-part definition of big data," *Data Education for Business and IT Professionals*, 2013. (`http://www.dataversity.net/gartners-three-part-definition-of-big-data/`)

[18] I. Giurgiu, O. Riva, D. Juric, I. Krivulev, and G. Alonso, *Calling the Cloud: Enabling Mobile Phones as Interfaces to Cloud Applications*, pp. 83–102, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[19] S. Goldwasser and S. Micali, "Probabilistic encryption and how to play mental poker keeping secret all partial information," in *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing (STOC'82)*, pp. 365–377, 1982.

[20] V. Goyal, A. Jain, O. Pandey, and A. Sahai, *Bounded Ciphertext Policy Attribute Based Encryption*, Springer Berlin Heidelberg, Berlin, Heidelberg", pp. 579–591, 2008.

[21] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in *Proceedings of the 13th ACM Conference on Computer and Communications Security (CCS'06)*, pp. 89–98, 2006.

[22] M. Green, S. Hohenberger, and B. Waters, "Outsourcing the decryption of abe ciphertexts," in *Proceedings of the 20th USENIX Conference on Security (SEC'11)*, pp. 34–34, 2011.

[23] J. Hur and K. Kang, "Dependable and secure computing in medical information systems," *Computer Communications*, vol. 36, no. 1, pp. 20-28, 2012.

[24] IDC, "Big data meets big data analytics," *SAS*, 2012. (`https://www.datanami.com/whitepaper/big_data_meets_big_data_analytics/`)

[25] M. Jakobsson and S. Wetzel, *Secure Server-Aided Signature Generation*, pp. 383–401. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[26] J. G. Peter, "Medical informatics 20/20," *Medicine & Health Sciences*, (`http://www.hoise.com/vmw/08/articles/vmw/LV-VM-01-08-1.html`)

[27] L. Jiguo, S. Fengjie, Z. Yichen, X. Huang, and S. Jian, "Verifiable outsourced decryption of attribute-based encryption with constant ciphertext length," *Security and Communication Networks*, vol. 2017, 2017.

[28] J. Sun and C. Reddy, "Big data analytics for healthcare," *Big Data Analytics in Healthcare*, (`https://www.siam.org/meetings/sdm13/sun.pdf/`)

[29] J. Lai, R. H. Deng, C. Guan, and J. Weng, "Attribute-based encryption with verifiable outsourced decryption," *IEEE Transactions on Information Forensics and Security*, vol. 8, pp. 1343–1354, Aug. 2013.

[30] C. C. Lee, P. S. Chung, and M. S. Hwang, "A survey on attribute-based encryption schemes of access control in cloud environments," *International Journal Network Security*, vol. 15, pp. 231–240, 2013.

[31] A. Lewko, A. Sahai, and B. Waters, "Revocation systems with very small private keys," in *IEEE Symposium on Security and Privacy*, pp. 273–285, May 2010.

[32] J. Li, X. Huang, J. Li, X. Chen, and Y. Xiang, "Securely outsourcing attribute-based encryption with checkability," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, pp. 2201–2210, Aug. 2014.

[33] J. Li, W. Yao, Y. Zhang, H. Qian, and J. Han, "Flexible and fine-grained attribute-based data storage in cloud computing," *IEEE Transactions on Services Computing*, vol. 10, no. 5, pp. 785-796, 2016.

[34] J. Li, X. Chen, J. Li, C. Jia, J. Ma, and W. Lou, *Fine-Grained Access Control System Based on Outsourced Attribute-Based Encryption*, pp. 592–609, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[35] X. Liang, Z. Cao, H. Lin, and J. Shao, "Attribute based proxy re-encryption with delegating capabilities," in *Proceedings of the 4th International Symposium on Information, Computer, and Communications Security (ASIACCS'09)*, pp. 276–286, 2009.

[36] S. Lin, R. Zhang, H. Ma, and M. Wang, "Revisiting attribute-based encryption with verifiable outsourced decryption," *IEEE Transactions on Information Forensics and Security*, vol. 10, pp. 2119–2130, Oct. 2015.

[37] J. Liu, Y. Zhang, and C. Xing, "Medical big data web service management platform," in *IEEE 11th International Conference on Semantic Computing (ICSC'17)*, pp. 316–321, Jan. 2017.

[38] W. Liu and E. K. Park, "Big data as an e-health service," in *International Conference on Computing, Networking and Communications (ICNC'14)*, pp. 982–988, Feb. 2014.

[39] Z. Liu, Z. L. Jiang, X. Wang, X. Huang, S. M. Yiu, and K. Sadakane, "Offline/online attribute-based encryption with verifiable outsourced decryption," *Concurrency and Computation: Practice and Experience*, vol. 29, no. 7, pp. e3915, 2017.

[40] S. Luo, J. Hu, and Z. Chen, *Ciphertext Policy Attribute-Based Proxy Re-encryption*, pp. 401–415, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[41] X. Mao, J. Lai, Q. Mei, K. Chen, and J. Weng, "Generic and efficient constructions of attribute-based encryption with verifiable outsourced decryption," *IEEE Transactions on Dependable and Secure Computing*, vol. 13, pp. 533–546, Sep. 2016.

[42] A. Mehmood, I. Natgunanathan, Y. Xiang, G. Hua, and S. Guo, "Protection of big data privacy," *IEEE Access*, vol. 4, pp. 1821–1834, 2016.

[43] B. Mounia and C. Habiba, "Big data privacy in healthcare moroccan context," *Procedia Computer Science*, vol. 63, pp. 575–580, 2015.

[44] M. Naor and M. Yung, "Public-key cryptosystems provably secure against chosen ciphertext attacks," in *Proceedings of the Twenty-Second Annual ACM Symposium on Theory of Computing (STOC'90)*, pp. 427–437, 1990.

[45] HealthIT.gov, *Guide to Privacy and Security of Electronic Health Information*, ver. 2, 2015. (`https://www.healthit.gov/sites/default/files/pdf/privacy/privacy-and-security-guide.pdf/`)

[46] I. Olaronke and O. Oluwaseun, "Big data in healthcare: Prospects, challenges and resolutions," in *Future Technologies Conference (FTC'16)*, pp. 1152–1157, Dec. 2016.

[47] R. Ostrovsky, A. Sahai, and B. Waters, "Attribute-based encryption with non-monotonic access structures," in *Proceedings of the 14th ACM Conference on Computer and Communications Security (CCS'07)*, pp. 195–203, 2007.

[48] R. Pankomera and D. V. Greunen, "Privacy and security issues for a patient-centric approach in public healthcare in a resource constrained setting," in *IST-Africa Week Conference*, pp. 1–10, May 2016.

[49] H. K. Patil and R. Seshadri, "Big data security and privacy issues in healthcare," in *IEEE International Congress on Big Data*, pp. 762–765, June 2014.

[50] J. Pavolotsky, "Demystifying big data," *Telecommunications Policy*, vol. 40, no. 9, pp. 837-854, 2016. (`http://breakinggov.sites.breakingmedia.com/wp-content/uploads/sites/4/2012/10/TechAmericaBigDataReport.pdf/`)

[51] H. Qinlong, M. Zhaofeng, Y. Yixian, N. Xinxin, and F. Jingyi, "Improving security and efciency for encrypted data sharing in online social networks," *China Communications*, vol. 11, pp. 104–117, Mar. 2014.

[52] C. Rackoff and D. R. Simon, "Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack," in *Proceedings of the 11th Annual International Cryptology Conference on Advances in Cryptology (CRYPTO'91)*, pp. 433–444, 1992.

[53] A. Sahai and B. Waters, *Fuzzy Identity-Based Encryption*, pp. 457–473, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[54] Adi Shamir, *Identity-Based Cryptosystems and Signature Schemes*, pp. 47–53, Springer Berlin Heidelberg, Berlin, Heidelberg, 1985.

[55] J. Shao, Z. Cao, X. Liang, and H. Lin, "Proxy re-encryption with keyword search," *Information Sciences*, vol. 180, no. 13, pp. 2576–2587, 2010.

[56] H. Wang, D. He, and J. Han, "VOD-ADAC: Anonymous distributed fine-grained access control protocol with verifiable outsourced decryption in public cloud," *IEEE Transactions on Services Computing*, vol. 1, no. 99, pp. 1–1, 2017.

[57] J. Weng, R. H. Deng, X. Ding, C. K. Chu, and J. Lai, "Conditional proxy re-encryption secure against chosen-ciphertext attack," in *Proceedings of the 4th International Symposium on Information, Computer, and Communications Security (ASIACCS'09)*, pp. 322–332, 2009.

[58] M. Xiao, M. Wang, X. Liu, and J. Sun, "Efficient distributed access control for big data in clouds," in *IEEE Conference on Computer Communications Workshops*, pp. 202–207, Apr. 2015.

[59] H. Yadav and M. Dave, "Secure data storage operations with verifiable outsourced decryption for mobile cloud computing," in *International Conference on Recent Advances and Innovations in Engineering (ICRAIE'14)*, pp. 1–5, May 2014.

[60] K. Yang, X. Jia, K. Ren, B. Zhang, and R. Xie, "Dacmacs: Effective data access control for multiauthority cloud storage systems," *IEEE Transactions on Information Forensics and Security*, vol. 8, pp. 1790–1801, Nov. 2013.

[61] S. Yu, C. Wang, K. Ren, and W. Lou, "Achieving secure, scalable, and fine-grained data access control in cloud computing," in *Proceedings of IEEE*, pp. 1–9, Mar. 2010.

[62] J. Zhang and Z. Zhang, "A ciphertext policy attribute-based encryption scheme without pairings," in *Proceedings of the 7th International Conference on Information Security and Cryptology (INSCRYPT'11)*, pp. 324–340, 2012.

[63] P. Zhang, Z. Chen, J. K. Liu, K. Liang, and H. Liu, "An efficient access control scheme with outsourcing capability and attribute update for fog computing," *Future Generation Computer Systems*, vol. 78, no. 2, pp. 753-762, 2016.

[64] Z. Zhou and D. Huang, "Efficient and secure data storage operations for mobile cloud computing," in *Proceedings of the 8th International Conference on Network and Service Management (CNSM'12)*, pp. 37–45, 2013.

[65] C. Zuo, J. Shao, G. Wei, M. Xie, and M. Ji, "CCA-secure ABE with outsourced decryption for fog computing," *Future Generation Computer Systems*, vol. 78, no. 2, pp. 730-738, 2016.

Biography

Kittur Philemon Kibiwott is currently a Ph.D candidate at the University of Electronic Science and Technology of China (UESTC). He received Bsc degree in Computer Science from Periyar University (India) and Msc degree Computer Science from Bharathiar University (India). His research area include cloud computing, cryptography and information security.

Zhang Fengli received her Ph.D degree from the University of Electronic Science and Technology of China (UESTC) in 2007 and M.S. degree in 1986. She is currently a Professor at the University of Electronic Science and Technology of China (UESTC). She has published more than eighty papers in refereed international journals and conferences which more than 50 are indexed by SCI and EI. Her research area of interest include mobile data management and application, network security, database.

Kimeli Victor K. is currently Dean of the School of Science at the University of Eldoret (UoE). He obtained his Ph.D from IHIT, China and M.Sc degree from Essex. His research area of interest include Databases, E-commerce, information security and Sensor Networks.

Omala A. Anyembe is currently a Ph.D candidate in the school of Computer Science and Engineering at the University of Electronic Science and Technology of China (UESTC). His research area of interest include cryptography, IoT and information security.

Eugene Opoku-Mensah is currently a Ph.D candidate in the school of Information and Software Engineering at the University of Electronic Science and Technology of China (UESTC). His research area include privacy preservation in cloud computing, Web page Ranking, and data mining.