Privacy Preserving Multi Sharing Control for Big Data Storage

K.Archana[1], Dr. Krishna mohanta[2]
[1]PG Student, [2]Professor
[1 2] Department of CSE, [1 2] Sri Ramanujar Engineering College, TamilNadu, India

***Abstract*--**security is a prime concern for any service that provides big data storage. The data of an individual should remain confidential and should be accessed only by any authenticated person. One of the aspects of security that is considered prior storing data is the anonymity of the service clients. The service that is used for storage should provide practical and fine-grained encrypted data sharing in such a way that only a ciphertext of data is shared among others by the data owner under some specified conditions. The required features are obtained by introduction a new technique for providing big data storage i,e, a privacy preserving ciphertext multi sharing mechanism. In this technique the advantages of proxy re-encryption technique are employed that enables ciphertext to be securely and conditionally shared multiple times and it also ensures that the knowledge of underlying message and the identity information of ciphertext senders and recipients is not leaked. The technique is also vulnerable to the chosen ciphertext attacks.

Keywords: Encryption, Decryption, Privacy-Anonymity.

I. INTRODUCTION

To date many individuals and companies choose to upload their data to clouds since the clouds supports considerable data storage service but also efficient data processing capability. Accordingly, it is unavoidable that trillions of personal and industrial data are flooding the Internet. For example, in some smart grid scenario, a governmental surveillance authority may choose to supervise the electricity consumption of a local living district. A great amount of electricity consumed data of each family located inside the district will be automatically transferred to the authority via Internet period by period. The need of big data storage, therefore, is more desirable than ever. Manuscript received; revised; accepted. Date of publication; date of current version. The work of K. Liang was supported in part by the Privacy- Aware Retrieval and Modeling of Genomic Data under Grant 13283250 and in part by the Academy of Finland, Finland. The work of W. Susilo was supported by the Australian Research Council Discovery Project under Grant ARC DP130101383. The work of J. K. Liu was supported by the National Natural Science Foundation of China under Grant 61472083. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Liqun Chen. (Corresponding author: Joseph K. Liu.) K. Liang is with the Department of Computer Science, Aalto University,

J. K. Liu is with the Faculty of Information Technology, Monash University, Melbourne, VIC 3800, Australia Color versions of one or more of the figures in this paper are available online at A basic security requirement of big data storage is to guarantee the confidentiality of the data. Fortunately, some existing cryptographic encryption mechanisms can be employed to fulfill the requirement. For instance, Public Key Encryption (PKE) allows a data sender to encrypts the data underthe public key of receiver such that no one except the valid recipient can gain access to the data. Nevertheless, this does not satisfy all the requirements of users in the scenario of big data storage. Consider the following scenario. We suppose a hospital. stores its patients' medical records in a cloud storage system and meanwhile, the records are all encrypted so as to avoid the cloud server from accessing to any patient's medical information. After a record is encrypted and further uploaded to the cloud, only those specified doctors can gain access to the record. By using some traditional PKE, Identity-Based Encryption (IBE), or Attribute-Based Encryption (ABE), the confidentiality of the record can be protected effectively. By trivially employing traditional encryption mechanisms (to guarantee the confidentiality of medical record), nevertheless,

We cannot prevent some sensitive personal information from being leaked to the cloud server but also the public. This is because traditional encryption systems do not consider the anonymity of a cipher text sender/receiver. Accordingly, someone, could be anyone with capability of obtaining a cipher text may know whose public key the cipher text is encrypted under, namely who is the owner of the cipher text, such that the patient associated with the cipher text can be easily identified. Similarly, the recipient/destination of

The ciphertext, e.g., Cardiology Dept., can be known from the ciphertext without any difficulty as well. This seriously disgraces the privacy of patient.Moreover, a patient might be transferred to more than one medical department in different treatment phases. The corresponding medical record then needs to be converted to the ciphertexts corresponding to various receivers so as to be shared among the departments. Therefore, the update of ciphertext recipient is desirable. Precisely speaking, a fine-grained ciphertext update for receivers is necessary in the sense that a ciphertext can be conditionally shared with others. The medical record owner, e.g., the patient, has rights to decide who can gain access to the record, and which kinds of data are allowed for access. For example, the patient can choose to specify that only the medical record described with "teeth" can be read by a dentist. This fine-grained control prevents a data sharing mechanism from being limited to the "all-or-nothing" share mode.

This research work aims to solve the above problems. To preserve anonymity, some well-known encryption mechanisms are proposed in the literature, such as anonymous IBE . By employing these primitives, the source and the destination of data can be protected privately. However, the primitives cannot support the update of ciphertext receiver. There are some naive approaches to update ciphertext's recipient. For instance, data owner can employ the decryptthen- re-encrypt mode. Nonetheless, this is applicable to the scenario where there is only a small amount of data.

If the encrypted data is either a group of sequences of genome information or a network audit log, the decryption and re-encryption might be time consumed and computation costly.Moreover, this mode also suffers from a limitation that the data owner has to be on-line all the time.

A. Our Contributions

In this paper, we aim to propose a ciphertext sharing mechanism with the following properties:

• Anonymity: given a ciphertext, no one knows the identity information of sender and receiver.

• Multiple receiver-updates: given a ciphertext, the receiver of the ciphertext can be updated in multiple times. In this paper, we refer to this property as "multi-hop".

• Conditional sharing: a ciphertext can be fine-grained shared with others if the pre-specified conditions are satisfied.

Achievements:We investigate a new notion, AMH-IBCPRE.

• The security model of MH-IBCPRE is the basic one, in which a challenger plays the game with the adversary to launch Chosen-Ciphertext Attacks (CCA) to the original ciphertext and re-encrypted ciphertext in order to solve a hard problem.

• We also consider the case where a proxy colludes with delegatee to compromise the underlying message and the secret key of delegator. Here, the protection of themessage is very difficult to achieve as the delegatee can always decrypt the corresponding ciphertext for the proxy. The secret key of the delegator, however, is possible to be secured.

• For the definition of collusion attacks model, we allow an adversary to acquire all re-encryption keys, and the adversary wins the game if it outputs a valid secret key of an uncorrupted user. We note that our definition is in the selective model in which the adversary has to output a target identity at the outset of the game.

• As to the security model of anonymity, it is complicated in the sense that we categorize the game into two subgames: one is the anonymity for delegator (i.e. given the original ciphertext an adversary cannot output the identity of delegator), the other is the anonymity ofre-encryption key (i.e. an adversary cannot distinguish a valid re-encryption key from a random one belonging to re-encryption key space).We next propose a concrete construction for unidirectional.

B. Related Work

Here, we compare our work with the some related systems, and summarize the comparison of properties in Table I. While multiple ciphertext receiver update (denoting as M.U.), conditional (data) share, collusion resistance (denoting as C.R.), anonymity, and without random oracle (denoting as W.R.O.),have all five been partially achieved by previous schemes, 1We refer to multiple ciphertext receiver update to a notion called Multi-Hop (MH) in this paper

II. SYSTEM DEFINITION AND THREAT MODELS

A. System Definition

Definition 1: A unidirectional Multi-Hop Identity-Based Conditional Proxy Re-Encryption (MH-IBCPRE) scheme consists of the following algorithms:

1. *(mpk, msk)←Setup(1k):* on input a security parameter *k*, output a master public key *mpk* and a master secret key *msk*. For simplicity, we omit *mpk* in the expression of the following algorithms.
2.
   a. *skI D ← KeyGen(msk, ID):* on input *msk*, and an identity *I D* ∈ {0, 1}∗, output a secret key *skI D*.
   b. 3) *rkw,I Di→I Di_ ← ReKeyGen(I Di , skI Di , I Di_, w):* on input a delegator's identity *I Di* and the corresponding secret key *skI Di* , a delegatee's identity *I Di_* , and a condition *w* ∈ {0, 1}∗, output a re-encryption key *rkw,I Di→I Di_* from *I Di* to *I Di_* under condition *w*.
3. *C*1,*I Di*,*w* ← *Enc(IDi , w, m):* on input an identity *I Di*, a condition *w* and a message *m*, output a 1-level ciphertext *C*1,*I Di*,*w* under identity *I Di* and *w*.
4. 4) *Cl*+1,*I Di_* ,*w* ← *ReEnc(rkw,I Di→I Di_, Cl,I Di ,w):* on input *rkw,I Di→I Di_*, and an *l*-level ciphertext *Cl*,*I Di* ,*w* under identity *I Di* and *w*, output an *(l* + 1*)*-level ciphertext *Cl*+1,*I Di_* ,*w* under identity *I Di_* and *w* or ⊥ for failure, where *l* ≥ 1, *l* ∈ N.
5. 5) *m* ← *Dec(skI Di ,Cl,I Di ,w):* on input *skI Di* , and an *l*-level ciphertext *Cl*,*I Di* ,*w* under identity *I Di* and *w*, output a message *m* or ⊥ for failure, where *l* ≥ 1, *l* ∈ N. LIANG *et al.*:

B. Threat Models

We define four models in terms of the selective condition and selective identity chosen ciphertext security (IND-sConsID- CCA), collusion resistance, the anonymity of the original ciphertext and anonymity of the re-encryption key in this section. Before proceeding, we define some notations.

- *Delegation Chain:* There is a set of re-encryption keys $RK = \{rkw,I\,Di1 \rightarrow I\,Di2, \ldots, rkw,I\,Dil-1 \rightarrow I\,Di\}$ under the same condition w, for any re encryption key $rkw,I\,Di\,j \rightarrow I\,Di\,j+1$ in RK, $I\,Di\,j = I\,Di\,j+1$. We say that there exists a delegation chain under w from identity $I\,Di1$ to identity $I\,Dil$, denoted as $w|I\,Di1 \rightarrow \ldots \rightarrow I\,Dil$. Note this delegation chain includes the case where $I\,Di1 = I\,Dil$. Besides, we use $w|I\,D$ to indicate a ciphertext under w and $I\,D$, and for a single identity $I\,D$ we use $\perp |I\,D$ to denote it.
- *Uncorrupted/Corrupted Identity:* If the secret key of an identity is compromised by an adversary, the identity is a corrupted identity. Else, it is an uncorrupted identity.
- *Uncorrupted Delegation Chain:* Suppose there is a delegation chain under w from $I\,Di$ to $I\,Dj$ (i.e. $w|I\,Di \rightarrow \ldots \rightarrow I\,Dj$). If there is no corrupted identity in the chain, it is an uncorrupted delegation chain., it is corrupted.

Definition 2: A unidirectional MH-IBCPRE scheme is IND-sCon-sID-CCA-secure if no PPT adversary A can win the game below with non-negligible advantage. In the game, B is the game challenger and k is the security parameter.

i. **Init.** A outputs a challenge identity $I\,D* \in \{0, 1\}*$ and a challenge condition $w \in \{0, 1\}*$.
ii. **Setup.** B runs $setup(1k)$ and returns mpk to A.
iii. **Phase 1.** A is given access to the following oracles. a) *Osk(ID):* given an identity $I\,D$, output $skI\,D \leftarrow KeyGen(msk, I\,D)$.

b) *Ork (IDi , I Di_,w):* on input two distinct identities $I\,Di$ and $I\,Di\_$, and a condition w, output $rkw,I\,Di \rightarrow I\,Di\_ \leftarrow ReKeyGen(I\,Di , skI\,Di , I\,Di\_ , w)$, where $skI\,Di \leftarrow KeyGen(msk, I\,Di)$.

c) *Ore(IDi , I Di_,w,Cl,I Di ,w):* on input two distinct identities $I\,Di$ and $I\,Di\_$, a condition w, and an l-level ciphertext $Cl,I\,Di\,,w$ under $I\,Di$ and w, output $Cl+1,I\,Di\,,w \leftarrow ReEnc(rkw,I\,Di \rightarrow I\,Di\_ , Cl,I\,Di\,,w)$, where $rkw,I\,Di \rightarrow I\,Di\_ \leftarrow ReKeyGen(I\,Di , skI\,Di , I\,Di\_ , w)$, $skI\,Di \leftarrow KeyGen(msk, I\,Di)$.

d) *Odec(IDi ,Cl,I Di ,w):* on input an identity $I\,Di$, and an l-level ciphertext $Cl,I\,Di\,,w$, output $m \leftarrow Dec(skI\,Di\,,Cl,I\,Di\,,w)$, where $skI\,Di \leftarrow KeyGen(msk, I\,Di)$.

In this phase the followings are forbidden to issue:

• *Osk (I D)* for any $I\,D$, if there is an uncorrupted delegation chain under $w*$ from $I\,D*$ to $I\,D$, or $I\,D* = I\,D$.

• *Ork (I Di , I Di_,w∗)* for any $I\,Di$, $I\,Di\_$, if there is an uncorrupted delegation chain under $w*$ from $I\,D*$ to $I\,Di$ or $I\,D* = I\,Di$, but $I\,Di\_$ is in a corrupted delegation chain.

4) Challenge**.** A outputs two equal length messages $m0$, $m1$, and a set of identities $\{I\,Di\,j\}$ $j=l*-1$ $j=1$ to B. B computes $Cl*,I\,D*,w*$ as $ReEnc(ReKeyGen(I\,Dil*-1 , skI\,Dil*-1 , I\,D* ,w* ReEnc(ReKeyGen(I\,Dil*-2 , skI\,Dil*-2, I\,Dil*-1,w*), . . . , ReEnc(ReKeyGen(I\,Di1 , skI\,Di1, I\,Di2,w*), Enc(I\,Di1,w*,mb))))$, where $l* \geq 2$, $l* \in \mathrm{N}$, $b \in_R \{0, 1\}$. Note that we here put $I\,D*$ to the $l*$ level of the ciphertext. This shows no difference from putting it in the first level of the ciphertext since the system supports multi-hop property.

5) **Phase 2.** Same as in **Phase 1** except the followings:

a) *Ore(IDi , I Di_,w*,Cl,I Di ,w*):* if *(I Di ,Cl,I Di ,w*)* is a derivative of *(I D*,Cl*,I D*,w*)*, and $I\,Di\_$ is in a corrupted delegation chain. As of [11], a derivative of*(I D* ,Cl*,I D*,w*)* is defined.

6) **Guess.** A outputs a guess $b\_ \in \{0, 1\}$. If $b\_ = b$, A wins. The advantage of A is defined as $\_ =$ $AdvI\,ND\text{-}sCon\text{-}s\,ID\text{-}CCA\,MH\text{-}I\,BCPRE,A\,(1k) = |Pr[b\_ = b] - 12|$. We now proceed to collusion resistance that guarantees that an adversary cannot compromise the entire secret key of a delegator even if it colludes with the delegatee. *Definition 3:* A unidirectional MH-IBCPRE scheme holds against selective collusion attacks if the advantage $AdvCR\,A\,(1k)$ is negligible for any PPT adversary A in the following experiment. Set $O1 = \{Osk, Ork\}$ and $AdvCR\,A\,(1k)$ as$Pr[skI\,D* \in _ : (I\,D*, State) \leftarrow A(1k);(mpk, msk) \leftarrow Setup(1k); skI\,D* \leftarrow AO1\,(mpk, State)]$ where k is the security parameter, $State$ is the state information, $I\,D*$ is the target and uncorrupted identity, Osk and Ork are the oracles defined in below with non-negligible advantage.

1) Init**.** A outputs a delegator's identity $I\,D\_$, a challenge delegatee's identity $I\,D*$, and a challenge condition $w*$.

2) Setup**.** Same as Definition 2.

3) Phase 1. A is allowed to issue queries to Osk,Ork , Ore and $Odec$ which are the oracles defined Definition 2 with the same restrictions.

4) Challenge. If the following queries

• *Osk(I Di)* for any $I\,Di$, if there is an uncorrupted delegation chain under $w*$ from $I\,D*$ to $I\,Di$, or $I\,D* = I\,Di$.

• *Ork (I Di , I Dj,w∗)* for any $I\,Di$, $I\,Dj$, if there is an uncorrupted delegation chain under $w*$ from $I\,D*$ to $I\,Di$ or $I\,D* = I\,Di$, but $I\,Dj$ is in a corrupted delegation chain. are never made, B flips a coin-toss for $b \in \{0, 1\}$.

5) Phase 2**.** Same as Phase 1 except the followings:

a) *Osk(I Di)* for any $I\,Di$, if there is an uncorrupted delegation chain under $w*$ from $I\,D*$ to $I\,Di$, or $I\,D* = I\,Di$;

b) *Ork (I Di , I Dj,w∗)* for any $I\,Di$, $I\,Dj$, if there is an uncorrupted delegation chain under $w*$ from $I\,D*$ to $I\,Di$ or $I\,D* = I\,Di$, but $I\,Dj$ is in a corrupted delegation chain;

c) *Ore(IDi , I Di_,w*,Cl,I Di ,w*):* if *(I Di ,Cl,I Di ,w*)* is a (derivative of) ciphertext generated by a re-encryption key in the delegation chain under $w*$ from $I\,D*$ to $I\,Di$, and $I\,Di\_$ is in a corrupted delegation chain.

III. PRELIMINARIES

A. Asymmetric Pairing

Let BSetup be an algorithm that on input the security parameter k, outputs the parameters of a bilinear map as (q, g, ˆg,G1,G2,GT , e), where G1, G2 and GT are multiplicative cyclic groups of prime order q, where |q| = k, and g is a random generator of G1, ˆg is a random generator of G2. The mapping e : G1 ×G2 → GT has three properties:

(1) Bilinearity: for all a, b ∈R Z ∗ q , e(ga, ˆ gb) = e(g, ˆg)ab;

(2) Non-degeneracy: e(g, ˆ g) = 1GT, where 1GT is the unit of GT; (3) Computability: e can be efficiently computed. Asymmetric Decisional BDH (ADBDH) Problem: Given a tuple (g, ga, gc, ˆ g, ˆ ga, ˆ gb)∈ G31 × G32 and T ∈ GT , decide whether T = e(g, ˆ g)abc. (Asymmetric) Decisional P-BDH Problem [14]: Given a tuple (g, ga, gab, gc, ˆ g, ˆ ga, ˆ gb)∈ G41 ×G32 and T ∈ GT , decide

whether T = e(g, ˆg)abc.

B. An Anonymous IBE and Its Extensions

Ducas [14] introduces an efficient anonymous IBE (Du-ANO-IBE) scheme in the standard model. We review its construction below, and omit the definition and security model of Du-ANO-IBE as the details can be found in [14].

• Setup(1k): run (q, g, ˆg,G1,G2,GT , e) ← BSetup(1k), choose random values α, β, γ, δ, η ∈ Z ∗ q, and set g1 = gα, g2 = gβ, h = gγ , f = gδ, t = gη, gˆ1 = gˆα, ˆ g2 = ˆ gβ, ˆh = gˆγ , fˆ = gˆδ, tˆ = gˆη. The master secret key msk = (gˆ0 = gˆαβ, fˆ, tˆ), the master public key mpk = (g, ˆ g, g1, h, f, t, ˆ g2, ˆh).

• Extract(msk, ID): given msk and an identity I D ∈

Z ∗ q , randomly choose r, R ∈ Z ∗ q , output skI D =

(skI D0 , skI D1 , skI D2) = (ˆ g0(ˆh I D fˆ)r tˆR, gˆr , gˆ R).

• Enc(mpk, ID, m): randomly choose s ∈ ∗ q, compute C1 = e(g1, ˆ g2)s · m, C2 = gs , C3 = (hI D f)s , C4 = ts , and output the ciphertext C = (C1,C2,C3,C4), where I D ∈ Z ∗ q , m ∈ GT .

• Dec(skI D, C): given a ciphertext C = (C1,C2,C3,C4), using the private key skI D to recover the plaintext m = C1 · e(C3, skI D1) · e(C4, skI D2)/e(C2, skI D0).

IV. PROPOSED WORK

We allow condition and identities to be arbitrary length, but they should be hashed by a Target Collision Resistant (TCR) hash function [13] H0 : {0, 1}∗ → Z ∗ q beforehand.

• Setup(1k): Given k, run (q, g, ˆg,G1,G2,GT , e) ←BSetup(1k). Let w ∈ Z ∗ q be a condition. Choose α, β,

γ , δ1, δ2, δ3, η ∈R Z ∗ q , and set g1 = gα, g2 = gβ,

h = gγ , f1 = gδ1 , f2 = gδ2 , f3 = gδ3 , t = gη, ˆ g1 = ˆgα, ˆ g2 = ˆgβ, ˆh = gˆγ , fˆ1 = gˆδ1 , fˆ2 = gˆδ2 , fˆ3 = gˆδ3 , tˆ = gˆη.

V. CONCLUSIONS

In this work, we introduced a novel notion, anonymous multi-hop identity-based conditional proxy re-encryption, to preserve the anonymity for cipher text sender/receiver, conditional data sharing and multiple recipient-update. We further proposed a concrete system for the notion. Meanwhile, we proved the system CCA-secure in the standard model under the decisional P-bilinear Differ-Hellman assumption. To the best of our knowledge, our primitive is the first of its kind in the literature.

REFERENCES

[1]. G. Ateniese, K. Benson, and S. Hohenberger, "Key-private proxy re-encryption," in Topics in Cryptology–CT-RSA (Lecture Notes in Computer Science), vol. 5473. Berlin, Germany: Springer-Verlag, 2009, pp. 279–294.

[2]. G. Ateniese, K. Fu, M. Green, and S. Hohenberger, "Improved proxy re-encryption schemes with applications to secure distributed storage," in Network and Distributed System Security. Berlin, Germany: Springer-Verlag, 2005, pp. 29–43.

[3]. G. Ateniese, K. Fu, M. Green, and S. Hohenberger, "Improved proxy re-encryption schemes with applications to secure distributed storage," ACM Trans. Inf. Syst. Secur., vol. 9, no. 1, pp. 1–30, 2006.

[4]. M. Bellare and S. Shoup, "Two-tier signatures, strongly unforgeable signatures, and Fiat–Shamir without random oracles," in Public Key Cryptography (Lecture Notes in Computer Science), vol. 4450. Berlin, Germany: Springer-Verlag, 2007, pp. 201–216.

[5]. M. Blaze, G. Bleumer, and M. Strauss, "Divertible protocols and atomic proxy cryptography," in Advances in Cryptology. Berlin, Germany: Springer-Verlag, 1998, pp. 127–144.

[6]. D. Boneh and X. Boyen, "Efficient selective-ID secure identity-based encryption without random oracles," in Advances in Cryptology–EUROCRYPT (Lecture Notes in Computer Science), vol. 3027. Berlin, Germany: Springer-Verlag, 2004, pp. 223–238.

[7]. D. Boneh, X. Boyen, and E.-J. Goh, "Hierarchical identity based encryption with constant size ciphertext," in Advances in Cryptology–EUROCRYPT (Lecture Notes in Computer Science), vol. 3494. Berlin, Germany: Springer-Verlag, 2005, pp. 440–456.

[8]. X. Boyen and B. Waters, "Anonymous hierarchical identity-based encryption (without random oracles)," in Advances in Cryptology–CRYPTO (Lecture Notes in Computer Science), vol. 4117. Berlin, Germany: Springer-Verlag, Aug. 2006, pp. 290–307.

[9]. J. Camenisch, M. Kohlweiss, A. Rial, and C. Sheedy, "Blind and anonymous identity-based encryption and authorised private searches onpublic key encrypted data," in Public Key Cryptography (Lecture Notes in Computer Science), vol. 5443. Berlin, Germany: Springer-Verlag, 2009, pp. 196–214.

[10]. R. Canetti, S. Halevi, and J. Katz, "Chosen-ciphertext security from identity-based encryption," in Advances in Cryptology–EUROCRYPT (Lecture Notes in Computer Science), vol. 3027. Berlin, Germany: Springer-Verlag, 2004, pp. 207–222.