Wuhan University
Journal of Natural Sciences

2014, Vol.19 No.5, 426-432

Article ID 1007-1202(2014)05-0426-07
DOI 10.1007/s11859-014-1035-z

Multi-Party Identity-Based Symmetric Privacy-Preserving Matching with Cloud Storage

□ QIU Shuo, LIU Jiqiang†, SHI Yanfeng, HAN Zhen

School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China

Abstract: In this paper, we address the problem of multi-party privacy-preserving matching (PPM) over the encrypted data. We firstly construct an efficient identity-based re-encryption scheme like ElGmal (IBR-ElGmal) using combined public keys, which not only ensures the privacy of the information during the transmission process but also holds perfect multiplicative homomorphic property. Then we construct a multi-party identity-based symmetric privacy-preserving matching (M-IBSPM) protocol based on IBR-ElGmal scheme in cloud environments, which realizes the privacy-preserving matching among multiple different parties as well as getting the symmetric output. Furthermore, with our M-IBSPM protocol, most of the computation costs are taken over by cloud service provider without leaking any privacy, and our protocol achieves perfect security and privacy in the semi-honest model. Finally, we analyze the efficiency for our protocol.

Key words: cloud computing; symmetric privacy-preserving matching; identity-based re-encryption
CLC number: TP 305

Received date: 2014-04-26
Foundation item: Supported by the Fundamental Research Funds for the Central Universities (2014YJS038)
Biography: QIU Shuo, female, Ph.D. candidate, research direction: privacy-preserving and information security. E-mail: 12112079@bjtu.edu.cn
† To whom correspondence should be addressed. E-mail: jqliu@bjtu.edu.cn

0 Introduction

Cloud computing is an epitome of on-demand and scalable computing. It provides virtualized computing resources as services over the Internet. Individuals or enterprises outsource their data to the cloud service provider who provides an abstraction of unlimited processing power and storage facility. However, the outsourced data maybe include some sensitive information, such as financial transaction, medical information and so on. So the data users should encrypt their private data before outsourcing to ensure the confidentiality. Privacy-preserving matching (PPM) introduced by Freedman *et al* [1] solves the problem that one party finds other similar parties without leaking any private information. In this paper, we focus on privacy-preserving matching under cloud computing setting. The categories of privacy-preserving matching protocols mainly include the two-party protocols and the multi-party protocols.

For the two-party protocols, Freedman *et al* [1] presented the first private matching protocol using oblivious polynomial evaluation [2] in the semi-honest model. Li *et al* [3] presented a taxonomy of design criteria for private matching and summarized some open problems of private matching including how to design more complicated protocols of multiple parties. Sang *et al* [4] proposed an efficient protocol for privacy-preserving matching on distributed datasets, which greatly reduces the computation and communication time. After that, Ye *et al* [5] proposed a distributed private matching scheme based on FNP scheme [1]. Hazay *et al* [6] proposed a pattern matching protocol using oblivious pseudorandom functions, which achieves linear complexity. However, their protocol has a large number of exponentiations. Later, Jarecki *et al* [7] improved its efficiency. Zhang *et al* [8] proposed fine-grained private matching for a specific application (proximity-based mobile social network).

For the multi-party protocols, Vaidya [9] realized the multi-party setting protocol using commutative encryption with information secret sharing proposed by Agrawal [10]. However, their methods provided weak security due to the deterministic property of commutative encryption. Li *et al* [11] proposed an unconditionally secure protocol for multi-party set intersection using the secret sharing technology. All inputs are shared and computed by those parties. This increases each party's computation overhead. To avoid Li's costly computing, Narayanan *et al* [12] proposed a multi-party protocol based on NP scheme and achieved low computation complexities. Recently, Li *et al* [13, 14] proposed a multi-party private matching protocol for mobile social network and ensured that the minimal private information was exchanged among different parties. Later, there are another two private matching protocols [15, 16] proposed by Gao and Zhou.

But all of these private matching protocols are constructed in traditional public-key setting. The identity-based cryptosystem was introduced by Shamir[17] to simplify the certificate management. The first identity-based private matching protocol was proposed by Wu *et al* [18]. As the description of Zhong *et al* [19], Wu's protocol only realized two-party private matching and its security is not very clear. So Zhong *et al* [19] presented an efficient and secure multi-party identity-based private matching protocol. However, the parties in their protocol need to directly interact with each other, which leads that all the parties must take over all the heavy computation of private matching themselves. It is not suitable for cloud environments and cannot get the symmetric output of the matching result. Qiu *et al* [20] proposed a symmetric identity-based private set intersection protocol without pairing and adopted a semi-honest third party to perform much of the protocol computation. However, their protocol is in two-party setting.

So, how to construct multi-party identity-based symmetric privacy-preserving matching (M-IBSPM) with cloud storage is still an open problem. To solve this problem, we construct an efficient identity based proxy re-encryption with multiplicative homomorphic property and then propose an M-IBSPM protocol based on it. Our contributions in this paper can be summarized as follows:

① We propose an identity-based re-encryption scheme using combined public keys based on the basic ElGmal encryption (IBR-ElGmal), which can simplify the certificate management. Furthermore, IBR-ElGmal scheme has perfect multiplicative homomorphic property.

② We propose a multi-party identity-based symmetric privacy-preserving matching (M-IBSPM) protocol based on IBR-ElGmal scheme under cloud environments, which not only realizes the privacy-preserving matching among multiple different parties, but also gets the symmetric output of the matching result.

③ Our M-IBSPM protocol reduces much computation overhead for all the parties by delegating most of the matching operations to the cloud service provider. We also analyze its security and privacy through rigorous simulation between a real world and an ideal world.

The rest of this paper is organized as follows. We introduce some preliminaries in Section 1. We present our identity-based proxy re-encryption scheme and privacy-preserving matching protocol in Section 2 and 3, respectively. We give a rigorous security analysis of our encryption scheme and matching protocol in Section 4. Then we analyze the efficiency of our protocol in Section 5. Finally, we conclude this paper in Section 6.

1 Preliminaries

1.1 Multiplicative Homomorphic Encryption

We use a semantic secure multiplicative homomorphic encryption scheme as a building block in our protocol. The homomorphic property can be stated as follows:

- Given two ciphertexts $\mathbf{Enc}(m_1)$ and $\mathbf{Enc}(m_2)$, we have $\mathbf{Enc}(m_1)\mathbf{Enc}(m_2)=\mathbf{Enc}(m_1m_2)$;
- Given a ciphertext $\mathbf{Enc}(m)$ and a constant c, we can efficiently compute $\mathbf{Enc}(m)^c=\mathbf{Enc}(m^c)$.

1.2 Divisible Decisional Diffie-Hellman Assumption

DDDH Assumption Let g be a generator of an Abelian group G, where the order of G is q. The challenger chooses $a,b,c\xleftarrow{\$}Z_q^*$ and a bit $\tau\xleftarrow{\$}\{0,1\}$, if $\tau=1$, he outputs the tuple $(g,g^a,g^b,g^{b/a})$; otherwise, he outputs the tuple (g,g^a,g^b,g^c). Then the adversary outputs a guess τ' of τ. The adversary can have ϵ

advantage to guess τ if $\left|\Pr[\tau=\tau']-\frac{1}{2}\right|=\epsilon$.

Note that, according to Ref. [21], we know that DDDH assumption is equivalent to the basic Decisional Diffie-Hellman (DDH) assumption.

Definition 1 The ϵ-DDDH assumption holds in G if no PPT adversary has at least ϵ advantage in solving above problem.

1.3 M-IBSPM Protocol

System Model We consider our multi-party identity-based symmetric privacy-preserving matching system in the following setting: the system consists of $n+1$ parties $A, B_1, B_2, \cdots, B_n$ and a semi-honest cloud service provider CSP. We denote the initiating party by A who launches the matching process and intends to find the parties that match with her, from the rest of the parties $B_1, B_2, \cdots, B_n$. Each party has an input data of their profile information (including interests or background), which will be mapped to a certain length (maybe through a one-way hash function).

We assume that CSP is semi-honest ("honest-but-curious") in this paper, meaning that CSP will honestly follow the protocol and not deviate from the protocol, but it is curious to find out as much as private information from the data that it receives and stores. In addition, it will keep a record of all its intermediate computations in order to learn additional information.

Definition 2 Similar to Ref. [19], we let D be all the possible inputs set. A has an input $a \in D$, each B_i has an input $b_i \in D$. Finally, the protocol outputs $F(a, b_1, b_2, \cdots, b_n) = (z_1, z_2, \cdots, z_n)$, where

$$z_i = \begin{cases} 1, & a = b_i \\ 0, & \text{otherwise} \end{cases}$$

Definition 3 (Computationally Indistinguishable) Let $\{X_n\}_{n\in\mathbf{N}}$ and $\{Y_n\}_{n\in\mathbf{N}}$ represent the probability distributions of X_n and Y_n. X_n, Y_n are computationally indistinguishable, noted as $X_n \overset{c}{\equiv} Y_n$, if for any PPT distinguishing algorithm D, there exists a negligible function $\mathrm{negl}(n)$, such that

$$\left|\Pr[D(1^n, X_n)=1] - \Pr[D(1^n, Y_n)=1]\right| \leqslant \mathrm{negl}(n)$$

The privacy definition of M-IBSPM includes two-folds as follows:

① The indistinguishability between the two models ;

② The privacy of any elements of A and B_i.

Given the definition of ①, firstly we discuss two models similar to Ref. [19] as follows:

Ideal Model In ideal model, there exists a trusted third party. Each of the $n+1$ parties $A, B_1, B_2, \cdots, B_n$ is supposed to send her/his input to the trusted party. One of these $n+1$ parties is honest and she sends her/his input faithfully; the other n parties are controlled by a polynomial-time adversary, so each of them sends an arbitrary message, or an arbitrary element of D or $\perp$ (Here a polynomial-time adversary is an adversary who tries to launch an attack by running an algorithm in polynomial time).

For each i, with Definition 2, the trusted party computes

$$z_i = \begin{cases} 1, & \text{if } a' = b' \in D \\ 0, & \text{if } a' \neq b' \wedge a', b' \in D \\ \perp, & \text{if } a' = \perp \vee b' = \perp \end{cases}$$

Notice that, a' is the message from A to the trusted party and b' is the message from B_i to the trusted party. The trusted party sends $(z_1, z_2, \cdots, z_n)$ to the $n+1$ parties.

Real Model The real model represents what happens in the real world. In this model, only one honest party follows the protocol and all the other parties controlled by a polynomial adversary have arbitrary behaviors such as sending arbitrary messages or skipping messages which need to be sent.

The privacy of our protocol requires that the real model is indistinguishable from the ideal model in adversary's view.

Specifically, we denote $\mathrm{ideal}_{A^{(i)}, B_1{}^{(i)}, \cdots, B_n{}^{(i)}}(a, b_1, \cdots, b_n)$ and $\mathrm{real}_{A^{(r)}, B_1{}^{(r)}, \cdots, B_n{}^{(r)}}(a, b_1, \cdots, b_n)$ as the output of the malicious parties in the ideal model and the real model, respectively, when the inputs of the $n+1$ parties are $(a, b_1, \cdots, b_n)$. We show the protocol is private if for all $(A^{(r)}, B_1{}^{(r)}, \cdots, B_n{}^{(r)})$ in the real model that includes an honest party and n malicious parties controlled by a polynomial-time adversary, there should exist $(A^{(i)}, B_1{}^{(i)}, \cdots, B_n{}^{(i)})$ in the ideal model that includes an honest party and n malicious parties controlled by a polynomial-time adversary, such that

$$\{\mathrm{ideal}_{A^{(i)}, B_1{}^{(i)}, \cdots, B_n{}^{(i)}}(a, b_1, \cdots, b_n)\} \overset{c}{\equiv} \{\mathrm{real}_{A^{(r)}, B_1{}^{(r)}, \cdots, B_n{}^{(r)}}(a, b_1, \cdots, b_n)\}$$

2 IBR-ElGmal Scheme

2.1 Scheme Construction

Similar with Qiu's definition of IBE-CPK[20], we construct our IBR-ElGmal scheme using combined public keys[22]. Our scheme has perfect multiplicative homomorphism.

Setup: Let G be an Abelian group such that the order of G is q, where q is a prime. Let g be a generator of G. The plaintext space is G. we choose k secret elements $x_i \overset{\$}{\leftarrow} Z_q^* \ (1 \leqslant i \leqslant k)$, and let

$$X = (x_1, x_2, \cdots, x_k),\ Y = (y_1, y_2, \cdots, y_k)$$

where $y_i = g^{x_i} \ (1 \leqslant i \leqslant k)$. Additionally, we choose a secure hash function $H : \{0,1\}^* \to \{0,1\}^k$ and set the master secret key msk as X and the public parameters pp as (g, G, q, k, Y, H).

KeyGen: Let ID be a user's identity. Suppose that $h_i \ (i = 1, \cdots, k)$ is the ith bit of $H(\mathrm{ID})$. The private key and public key of the user are, respectively

$$x_{\mathrm{ID}} = \sum_{i=1}^{k} h_i x_i \bmod q,$$

$$y_{\mathrm{ID}} = \prod_{i=1}^{k} (y_i)^{h_i} = \prod_{i=1}^{k} g^{h_i x_i} = g^{x_{\mathrm{ID}}}$$

RekeyGen: Given two users' private keys x_{ID_1} and x_{ID_2}, it outputs the re-encryption key

$$\mathrm{RK}_{\mathrm{ID}_1 \to \mathrm{ID}_2} = x_{\mathrm{ID}_2} / x_{\mathrm{ID}_1}$$

The proxy can use the re-encryption key to convert the ciphertext under ID_1 into the ciphertext under ID_2.

Enc: Given a message $m \in G$, it encrypts the message as $C = (y_{\mathrm{ID}_1}{}^{r}, mg^r)$, where $r \overset{\$}{\leftarrow} Z_q^*$.

ReEnc: Given the ciphertext $C = (C_1, C_2)$ under ID_1, it computes

$$C_1' = C_1^{\mathrm{RK}_{\mathrm{ID}_1 \to \mathrm{ID}_2}} = y_{\mathrm{ID}_1}^{\ r\mathrm{RK}_{\mathrm{ID}_1 \to \mathrm{ID}_2}} = g^{x_{\mathrm{ID}_1} r (x_{\mathrm{ID}_2}/x_{\mathrm{ID}_1})} = y_{\mathrm{ID}_2}^{\ r}$$

Thus it converts $C = (C_1, C_2)$ into the ciphertext $C' = (C_1', C_2)$ under ID_2.

Dec: Let $C = (C_1, C_2)$ be a valid ciphertext of message m under a user's identity ID, then the user can decrypt C using his private key x_{ID} as:

$$C_2 / C_1^{(x_{\mathrm{ID}})^{-1}} = mg^r / g^{rx_{\mathrm{ID}}(x_{\mathrm{ID}})^{-1}} = mg^r / g^r = m$$

2.2 Homomorphism Verification

Obviously, our IBR-ElGmal scheme holds perfect multiplicative homomorphic property:

$$\begin{aligned} &\mathbf{Enc}(m_1, r_1) \cdot \mathbf{Enc}(m_2, r_2) \\ &= (y_{\mathrm{ID}}{}^{r_1}, m_1 g^{r_1}) \cdot (y_{\mathrm{ID}}{}^{r_2}, m_2 g^{r_2}) \\ &= (y_{\mathrm{ID}}{}^{r_1} \cdot y_{\mathrm{ID}}{}^{r_2}, (m_1 g^{r_1}) \cdot (m_2 g^{r_2})) \\ &= (y_{\mathrm{ID}}{}^{r_1 + r_2}, m_1 m_2 g^{r_1 + r_2}) \\ &= \mathbf{Enc}(m_1 m_2, r_1 + r_2) \end{aligned}$$

$$\begin{aligned} &\mathbf{Enc}(m, r)^c = (y_{\mathrm{ID}}{}^{r}, mg^r)^c \\ &= (y_{\mathrm{ID}}{}^{rc}, m^c g^{rc}) \\ &= (y_{\mathrm{ID}}{}^{r'}, m^c g^{r'}) \\ &= \mathbf{Enc}(m^c, r') \end{aligned}$$

3 M-IBSPM Protocol

Next, we construct an M-IBSPM protocol based on IBR-ElGmal scheme with a semi-honest cloud service provider CSP and $n+1$ parties $A, B_1, B_2, \cdots, B_n$.

We firstly present a specific **Re-Key Gen** algorithm for CSP to generate the re-encryption key of A and B_i.

Re-Key Gen CSP gets the re-encryption key of A and $B_i (i = 1, 2, \cdots, n)$ as follows:

- A encrypts her private key x_{ID_A} under the CSP 's identity ID_C, then she encrypts $\mathbf{Enc}_{\mathrm{ID}_C}(x_{\mathrm{ID}_A})$ under B_i's identity to get $\mathbf{Enc}_{\mathrm{ID}_{B_i}}(\mathbf{Enc}_{\mathrm{D}_C}(x_{\mathrm{ID}_A}))$, and sends it to B_i.
- Each B_i receives $\mathbf{Enc}_{\mathrm{ID}_{B_i}}(\mathbf{Enc}_{\mathrm{ID}_C}(x_{\mathrm{ID}_A}))$ and decrypts it to get $\mathbf{Enc}_{\mathrm{ID}_C}(x_{\mathrm{ID}_A})$, then he encrypts his private key under CSP 's identity to get $\mathbf{Enc}_{\mathrm{ID}_C}(x_{\mathrm{ID}_{B_i}})$ and computes
$$\mathbf{Enc}_{\mathrm{ID}_C}(x_{\mathrm{ID}_{B_i}}) / \mathbf{Enc}_{\mathrm{ID}_C}(x_{\mathrm{ID}_A}) = \mathbf{Enc}_{\mathrm{ID}_C}(x_{\mathrm{ID}_{B_i}} / x_{\mathrm{ID}_A}),$$
and he sends $\mathbf{Enc}_{\mathrm{ID}_C}(x_{\mathrm{ID}_{B_i}} / x_{\mathrm{ID}_A})$ to CSP .
- CSP decrypts the ciphertext and gets the re-encryption key $\mathrm{RK}_{A \to B_i} = x_{\mathrm{ID}_{B_i}} / x_{\mathrm{ID}_A}$ and computes $\mathrm{RK}_{B_i \to A} = 1 / \mathrm{RK}_{A \to B_i} = x_{\mathrm{ID}_A} / x_{\mathrm{ID}_{B_i}}$. Finally, it gets the bidirectional re-encryption keys between A and $B_i (i = 1, 2, \cdots, n)$.

M-IBSPM We assume $D \subseteq G$, the M-IBSPM protocol performs as follows:

Phase 1 A and $B_i \ (i = 1, 2, \cdots, n)$ perform as follows:

- A and B_i run the **KeyGen** to generate their public/private key pairs, respectively. Then they run **Re-Key Gen** with CSP to generate re-encryption keys.
- A and B_i run the **Enc** to encrypt their private data and send their ciphertexts $\mathbf{Enc}_{\mathrm{ID}_A}(a, r_A)$, $\mathbf{Enc}_{\mathrm{ID}_{B_i}}(b_i, r_{B_i})$ (where $r_A, r_{B_i} \in Z_q^*$) to CSP .

Phase 2 When A wants to request matching query, the protocol performs as follows:

- A sends a matching request Req to CSP .
- CSP receives A's request, then performs as follows:

- It uses $\mathrm{RK}_{A \to B_i}$ to re-encryption $C_a = \mathbf{Enc}_{\mathrm{ID}_A}(a, r_A)$ to $C_a' = \mathbf{Enc}_{\mathrm{ID}_{B_i}}(a, r_A)$.

- It randomly chooses $\alpha_i \in Z_q^*$, then computes

$$\begin{aligned}T_i &= (C'_a / C_{b_i})^{\alpha_i}\\ &=(\mathbf{Enc}_{\mathrm{ID}_{B_i}}(a, r_A)/\mathbf{Enc}_{\mathrm{ID}_{B_i}}(b_i, r_{B_i}))^{\alpha_i}\\ &=\mathbf{Enc}_{\mathrm{ID}_{B_i}}(a/b_i, r_{A,B_i})^{\alpha_i}\\ &=\mathbf{Enc}_{\mathrm{ID}_{B_i}}((a/b_i)^{\alpha_i}, r'_{A,B_i})\end{aligned}$$

- Then it sends T_i to B_i and re-encryption T_i to get $T'_i = \mathbf{Enc}_{\mathrm{ID}_A}((a/b_i)^{\alpha_i}, r'_{A,B_i})$ at the same time, then sends T'_i to A.

- A decrypts to get $t'_i = \mathbf{Dec}_{x_{\mathrm{ID}_A}}(T'_i)$ and B_i decrypts to get $t_i = \mathbf{Dec}_{x_{\mathrm{ID}_{B_i}}}(T_i)$. If $t'_i = t_i = 1$, then output 1; If $t'_i \neq 1$ or $t_i \neq 1$, then output 0.

4 Security Analysis

Theorem 1 Our IBR-ElGmal scheme is (t,ϵ)-semantically secure under $\frac{\epsilon}{2}\left(1-\frac{1}{e}-2^{t-k}\right)$-EDDH assumption in the random oracle model.

The detailed proof of this theorem refers to Qiu [20]. In the following part, we prove the correctness and privacy of our M-IBSPM protocol.

4.1 Correctness

Theorem 2 If A, B_i and CSP exactly follow the protocol M-IBSPM, then M-IBSPM can be with high probability to output $F(a,b_1,b_2,\cdots,b_n)=(z_1,z_2,\cdots,z_n)$, where

$$z_i = \begin{cases}1, & a=b_i\\ 0, & \text{otherwise}\end{cases}$$

Proof Since B_i has the same decryption results with A, here we give A's decryption process as follows:

$$\begin{aligned}t'_i &= \mathbf{Dec}_{x_{\mathrm{ID}_A}}(T'_i)\\ &= \mathbf{Dec}_{x_{\mathrm{ID}_A}}(\mathbf{Enc}_{\mathrm{ID}_A}((a/b_i)^{\alpha_i}, r'_{A,B_i}))\\ &= (a/b_i)^{\alpha_i}\end{aligned}$$

Obviously, if $a=b_i$, $t'_i=1$, then $z_i=1$, otherwise $z_i=0$. We complete the proof of Theorem 2.

4.2 Privacy

Theorem 3 The M-IBSPM protocol is computationally indistinguishable between the ideal model and the real model under the semantically secure identity based re-encryption with multiplicative homomorphic property.

Proof Intuitively, we analyze the protocol as follows:

- A: For the ciphertext $\mathbf{Enc}_{\mathrm{ID}_A}((a/b_i)^{\alpha_i}, r'_{A,B_i})$, A can decrypt it to $(a/b_i)^{\alpha_i}$. It is randomized by α_1, so A can not learn any information about b_i unless $a=b_i$.

- B_i: For the ciphertext $\mathbf{Enc}_{\mathrm{ID}_{B_i}}(\mathbf{Enc}_{\mathrm{ID}_C}(x_{\mathrm{ID}_A}))$, B_i decrypts it to $\mathbf{Enc}_{\mathrm{ID}_C}(x_{\mathrm{ID}_A})$, since our IBR-ElGmal scheme is IND-ID-CPA secure, it is indistinguishable between $\mathbf{Enc}_{\mathrm{ID}_{B_i}}((a/b_i)^{\alpha_i}, r'_{A,B_i})$ and a random element. Obviously, B_i can decrypt it to $(a/b_i)^{\alpha_i}$ and it is randomized by α_1. So B_i can't learn any information about $a=b_i$ unless a.

- CSP: For CSP, all the ciphertexts $\{\mathbf{Enc}_{\mathrm{ID}_A}(a, r_A), \mathbf{Enc}_{\mathrm{ID}_{B_1}}(b_1, r_{B_i}), \cdots, \mathbf{Enc}_{\mathrm{ID}_{B_n}}(b_n, r_{B_i})\}$ are indistinguishable from random elements. And for $\mathbf{Enc}_{\mathrm{ID}_C}(x_{\mathrm{ID}_{B_i}}/x_{\mathrm{ID}_A})$, CSP can decrypt it to $x_{\mathrm{ID}_{B_i}}/x_{\mathrm{ID}_A}$. But it can't learn anything about $x_{\mathrm{ID}_{B_i}}$ or x_{ID_A}.

In the following part, we formally prove Theorem 3.

First, we consider the case $A^{(r)}$ and CSP are honest, and in the ideal model $A^{(i)}$ is honest. In the ideal model, the adversary chooses $x_{\mathrm{ID}_A} \xleftarrow{\$} Z_q^*$ and an element $a \xleftarrow{\$} D$, and encrypts them to $\mathbf{Enc}_{\mathrm{ID}_\mathrm{C}}(x_{\mathrm{ID}_A})$ and $\mathbf{Enc}_{\mathrm{ID}_A}(a)$ respectively. Then $A^{(i)}$ feeds $\mathbf{Enc}_{\mathrm{ID}_C}(\mathrm{x}_{\mathrm{ID}_A})$ and $\mathbf{Enc}_{\mathrm{ID}_A}(a)$ to $B_i^{(r)}$. Suppose the trusted party returns $(z_1,\cdots,z_n)$ to each $B_i^{(r)}$, if $z_i=1$, the adversary feeds $B_i^{(r)}$ with encryption of 1. If $z_i=0$, the adversary feeds $B_i^{(r)}$ with encryption of a random element, and each $B_i^{(i)}$ outputs what $B_i^{(r)}$ outputs. Since our encryption scheme is IND-ID-CPA secure, the outputs are computationally indistinguishable with that of $B_i^{(r)}$. So the outputs of $B_i^{(i)}$ in the ideal model are computationally indistinguishable from the outputs of $B_i^{(r)}$ in the real model.

Next, we consider the case $B_i^{(r)}$ and CSP are honest, and in the ideal model $B_i^{(i)}$ is honest. In the ideal model, the adversary chooses $x_{\mathrm{ID}_{B_i}} \xleftarrow{\$} Z_q^*$ and an element $b_i \xleftarrow{\$} D$, and gets $\mathbf{Enc}_{\mathrm{ID}_\mathrm{C}}(x_{\mathrm{ID}_{B_i}}/x_{\mathrm{ID}_A})$ and $\mathbf{Enc}_{\mathrm{ID}_{B_i}}(b_i)$. Then $B_i^{(i)}$ feeds $\mathbf{Enc}_{\mathrm{ID}_\mathrm{C}}(x_{\mathrm{ID}_{B_i}}/x_{\mathrm{ID}_A})$ and $\mathbf{Enc}_{\mathrm{ID}_{B_i}}(b_i)$ to $A^{(r)}$ and $B_j^{(r)}$, where $j=1,\cdots,i-1, i+1, \cdots, n$. Suppose the trusted third party returns $(z_1,\cdots,z_n)$ to $A^{(r)}$, and z_j to $B_j^{(r)}$, respectively. If $z_j=1$, the adversary feeds $A^{(r)}$ and $B_j^{(r)}$ with encryption of 1. If

$z_j = 0$, the adversary feeds $A^{(r)}$ and $B_j^{(r)}$ with encryption of a random element, and each $A^{(i)}, B_j^{(i)}$ outputs what $A^{(r)}, B_j^{(r)}$ outputs, where $j = 1, \cdots, i-1, i+1, \cdots, n$. Since our encryption scheme is IND-ID-CPA secure, the outputs are computationally indistinguishable with that of $A^{(r)}$ and $B_i^{(r)}$. So the outputs of $A^{(i)}$ and $B_i^{(i)}$ in the ideal model are computationally indistinguishable from the outputs of $A^{(r)}$ and $B_i^{(r)}$ in the real model.

For CSP, all the messages of A and each B_i are in encrypted form, since our encryption scheme is IND-ID-CPA secure, it can't learn any information about A and each B_i. We complete the proof of Theorem 3.

5 Complexity Analysis

Table 1 describes the asymptotic complexities of three main algorithms in our IBR-ElGmal scheme. Note that Exp denotes the exponentiation operation. We show that our re-encryption scheme is as efficient as the basic ElGmal scheme.

Table 1 Asymptotic complexities of IBR-ElGmal scheme

Algorithm	Computational complexity	Output size
Enc	2 Exp	2|*G*|
ReEnc	1 Exp	2|*G*|
Dec	1 Exp	—

With the above complexities analysis of re-encryption scheme, we describe the computational complexities and communication overhead for three phases (RE-Key, Phase 1, Phase 2) in our M-IBSPM protocol in Table 2 and 3, where |Enc| denotes the output size of Enc in Table 3. We can know that CSP takes over most of the matching computation from Table 2 and all the communication overhead are just in a polynomial size of n based on Table 3.

Table 2 Computational complexities of M-IBSPM protocol

Phase	A	B_i	CSP
Re-Key Gen	$(n+1)$ Enc	1 Dec + 1 Enc	1 Dec
Phase 1	1 Enc	1 Enc	—
Phase 2	n Dec	1 Dec	$2n$ ReEnc+ n Exp

Table 3 Communication overhead of M-IBSPM protocol

Phase	Communication overhead
Re-Key Gen	$2n$ |Enc|
Phase 1	(n+1) |Enc|
Phase 2	$2n$|Enc|

6 Conclusion

In this paper, we propose an identity-based re-encryption scheme using combined public keys based on the basic ElGmal encryption (IBR-ElGmal). It not only simplifies the certificate management but also has perfect multiplicative homomorphic property. We also propose a multi-party identity-based symmetric privacy-preserving matching (M-IBSPM) protocol based on IBR-ElGmal scheme in cloud environments, which leads to a symmetric output of matching result. Our M-IBSPM scheme reduces much computation overhead of all the parties by delegating most of the matching operations to *CSP* and we prove the security and privacy through rigorous analysis in the semi-honest model. Finally, we give a detailed complexity analysis of our protocol. However, in our M-IBPSM protocol, A needs to decrypt all the ciphertexts to get the last matching result (sometimes maybe only several parties are matched), so it will waste much of *A's* computation overhead. Therefore, one of our future work is let the parties A and B_i directly get the matching result without decryption.

References

[1] Freedman M J, Nissim K, Pinkas B. Efficient private matching and set intersection[C]//*Advances in Cryptology-EUROCRYPT* 2004. Berlin, Heidelberg: Springer-Verlag, 2004: 1-19.

[2] Naor M, Pinkas B. Oblivious transfer and polynomial evaluation[C]//*Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*. New York：ACM Press, 1999: 245-254.

[3] Li Y, Tygar J D, Hellerstein J M. Private matching[C]//*Computer Security in the* 21*st Century*. Berlin, Heidelberg: Springer-Verlag, 2005: 25-50.

[4] Sang Y, Shen H, Tan Y, *et al*. Efficient protocols for privacy preserving matching against distributed datasets[C]// *Information and Communications Security*. Berlin, Heidelberg: Springer-Verlag, 2006: 210-227.

[5] Ye Q, Wang H, Pieprzyk J. Distributed private matching and set operations [C]//*Information Security Practice and Experience*. Berlin, Heidelberg: Springer-Verlag, 2008: 347-360.

[6] Hazay C, Lindell Y. Efficient protocols for set intersection and pattern matching with security against malicious and covert adversaries [C]//*Theory of Cryptography.* Berlin, Heidelberg: Springer-Verlag, 2008: 155-175.

[7] Jarecki S, Liu X. Efficient oblivious pseudorandom function with applications to adaptive ot and secure computation of set intersection[C]//*Theory of Cryptography.* Berlin, Heidelberg: Springer-Verlag, 2009: 577-594.

[8] Zhang R, Zhang Y, Sun J, *et al*. Fine-grained private matching for proximity-based mobile social networking [C]// *INFOCOM*, 2012 *Proceedings IEEE*. Piscataway N J: IEEE Press, 2012: 1969-1977.

[9] Vaidya J, Clifton C. Secure set intersection cardinality with application to association rule mining[J]. *Journal of Computer Security*, 2005, **13**(4): 593-622.

[10] Agrawal R, Evfimievski A, Srikant R. Information sharing across private databases[C]//*Proceedings of the* 2003 *ACM SIGMOD International Conference on Management of Data*. New York: ACM Press, 2003: 86-97.

[11] Li R, Wu C. An unconditionally secure protocol for multi-party set intersection[C]//*Applied Cryptography and Network Security*. Berlin, Heidelberg: Springer-Verlag, 2007: 226-236.

[12] Narayanan G S, Aishwarya T, Agrawal A, *et al*. Multi party distributed private matching, set disjointness and cardinality of set intersection with information theoretic security [C] //*Cryptology and Network Security*. Berlin, Heidelberg: Springer-Verlag, 2009: 21-40.

[13] Li M, Cao N, Yu S, *et al*. Findu: Privacy-preserving personal profile matching in mobile social networks[C]//*INFOCOM*, 2011 *Proceedings IEEE.* Piscataway N J: IEEE Press, 2011: 2435-2443.

[14] Li M, Yu S, Cao N, *et al*. Privacy-preserving distributed profile matching in proximity-based mobile social networks[J]. *IEEE Transactions on Wireless Communications*, 2013, **12**(5): 2024-2033.

[15] Gao Z, Du S, Li M, *et al*. Fairness-aware and privacy-preserving friend matching protocol in mobile social networks[J]. *IEEE Transactions on Emerging Topics in Computing*, 2013, **1**(1): 192-200.

[16] Zhou B, Pei J. Preserving privacy in social networks against neighborhood attacks[C]// *IEEE* 24*th International Conference on Data Engineering.* Piscataway N J: IEEE Press, 2008: 506-515.

[17] Shamir A. Identity-based cryptosystem and signature schemes[C] //*Advances in Cryptology—EUROCRYPT*'98. Berlin Heidelberg: Springer-Verlag, 1984:47-53.

[18] Wu Z, Chen Z, Guo F, *et al*. Identity based private matching[C]// *Third International Workshop on Security*, *Privacy and Trust in Pervasive and Ubiquitous Computing*, 2007. Piscataway N J: IEEE Press, 2007: 85-90.

[19] Zhong S, Chen T. An efficient identity-based protocol for private matching[J]. *International Journal of Communication Systems*, 2011, **24**(4): 543-552.

[20] Qiu S, Liu J, Shi Y. Identity-based symmetric private set intersection[C]// 2013 *International Conference on Social Computing* (*Social Com*). Piscataway N J: IEEE Press, 2013: 653-658.

[21] Liu J, Zhong S. Fast Identity-based encryption using combined public keys[EB/OL]. [2014-03-20]. *http://www.paper.edu.cn/releasepaper/content/*200903-756.

[22] Bao F, Deng R H, Zhu H. Variations of Diffie-Hellman problem[C]//*Information and Communications Security*. Berlin, Heidelberg: Springer-Verlag, 2003: 301-312.

□