A Secure and Efficient On-line/Off-line Group Key Distribution Protocol

Keju Meng $\,\cdot\,$ Fuyou Miao $\,\cdot\,$ Yue Yu

Received: date / Accepted: date

Abstract Nowadays, group communications are getting more and more popular. In order to secure the communication, all participating users need to share a common group key in advance. The paper proposes a secure and efficient group key distribution protocol based on Shamir's secret sharing scheme. In the protocol, 1) each user only needs to send registration message in privacy, while all the other messages can be transported in public. Meanwhile, 2) the scheme supports authentication for group keys without any assumption of hard mathematics problem. Moreover, 3) the protocol introduces the notion of on-line/off-line into group key distribution and thus the speeds of group key response and recovery are greatly improved. Analyses show that our scheme is resistant to passive attack, impersonation attack and reply attack.

 $\begin{array}{l} \textbf{Keywords} \ \mbox{Group key distribution} \cdot \mbox{Group key authentication} \cdot \mbox{Secret sharing} \cdot \mbox{Lagrange interpolation polynomial} \cdot \mbox{On-line}/\mbox{Off-line} \\ \textbf{MSC 2010 Codes: 94A60;94A62} \end{array}$

1 Introduction

With the rapid development of network, communication patterns are not limited to 1-to-1 or 1-to-m (i.e. sever/clients). Group communications with m-to-m pat-

Keju Meng

Fuyou Miao (corresponding author) No.96, JinZhai Road Baohe District,Hefei,Anhui, 230026 E-mail: mfy@ustc.edu.cn

Yue Yu

No.96, JinZhai Road Baohe District,Hefei,Anhui, 230026 E-mail: yuyue204@mail.ustc.edu.cn University of Science and Technology of China

No.96, JinZhai Road Baohe District,Hefei,Anhui, 230026 E-mail: mkj@mail.ustc.edu.cn University of Science and Technology of China

University of Science and Technology of China

tern become more and more popular. Before a secure group communication, all participating users are required to share a group key. Group key establishment protocols have been studied in many literatures. According to whether a trusted key generation center (KGC) exists or not, these protocols are classified into two classes: group key agreement (or exchange) protocols and group key distribution (or transfer) protocols.

In group key agreement protocols, there is not a KGC trusted by all users. Initially, the group key agreement protocol [3] based on Diffie-Hellman algorithm [6] was proposed by Ingemarsson et al. in 1982. The protocol requires all n participants to communicates synchronously and needs n-1 rounds of interaction to establish a group key. Steiner et al. [28] extends basic Diffie-Hellman key exchange to a n-party Diffie-Hellman key exchange. In 2004, Kim et al. [16] proposed a fault-tolerant protocol named Tree-based Group Diffie-Hellman. Protocols [4,14, 36] are also based on Diffie-Hellman algorithm. Moreover, there are some group key agreement protocols based on other methods. For example, Chen et al. [5] proposed a key agreement protocol based on bilinear pairing to improve the key establishment efficiency. Irshad et al. [15] proposed the protocol based on Chebyshev chaotic map without a registration center. Because there is no KGC and all the users are equal in these protocols, more communication cost is required during key establishment.

Distinct from group key agreement protocols, there exists a KGC trusted by all users in group key distribution protocols [31, 17, 24, 11]. The group key is generated and distributed by the KGC, so that group key establishment can be accomplished more efficiently. However, it is difficult to distribute a group key for the following reasons:

- 1. Since a group key is transferred to multiple users, it is easier to be intercepted during distribution.
- 2. Group users may change from time to time. When a user leaves or joins a group, the group key needs to be updated such that users out of the group have no information about the new group key.
- 3. Even if a group does not change for a long time, a session key still needs to be updated after a period of time. Otherwise, it could be cracked by adversaries.

As a group-oriented cryptographic tool, secret sharing has the potential to address the above problems. Therefore, based on Shamir's (t, n) threshold secret sharing (SS) scheme, we propose an on-line/off-line group key distribution protocol.

(t, n) threshold SS was first proposed by Shamir [26] and Blakley [2] separately in 1979. In a (t, n) SS scheme, a secret is divided into n shares and each share is allocated to a shareholder secretly such that any t or more than t shareholders can reconstruct the secret while less than t shareholders cannot obtain the secret. Shamir's (t, n) threshold SS and Blakley's (t, n) threshold SS are based on polynomial interpolation and hyperplane geometry respectively, while Mignotte [22] and Asmuth [1] et al.'s schemes are both based on Chinese Remainder Theorem. All these schemes are not based on any hard mathematical problems.

There are some literatures using (t, n) threshold SS to design group key distribution protocols [21,8,29,35]. Guo et al. [9] proposed a group key distribution protocol based on Asmuth-Bloom's (t, n) threshold SS. Harn and Lin proposed an authenticated group key transfer protocol [12] based on Shamir's (t, n) threshold SS. However, the protocol is not information theoretically secure since it empolys large integer factorization problem to thwart Insider attack. Liu et al. [20] tried to improve the security of Harn-Lin's protocol by utilizing two hash functions, but their protocol is also based on large integer factorization problem.

In these protocols [12, 20, 25], one-way hash functions are used to support group key authentication. The KGC announces the hash value of group key in advance, so that users can verify whether the group key is correctly distributed by the KGC. In our proposed protocol, hash function is not only used for authentication but also utilized to compute an offset to protect the group key.

In the above distribution protocols, the KGC starts to compute and distribute group key after it has received all users' requests. However, in the scenes where quick response is required, more efficient distribution protocol is needed. Therefore, we introduce on-line/off-line method to speed the response of group key distribution protocol for the first time.

The concept of on-line/off-line was first proposed by Even et al. [7] to design an efficient digital signature scheme. The idea is that the signature generating procedure is divided into two phases. Part of signature can be computed during off-line phase (before the signed message is given), so that on-line phase (after the signed message is received) works very fast. Subsequently, Tanaka [30], Liu [19] and some others [27, 13, 18, 34, 10], extended the notion to design some other schemes. In our group key distribution protocol, the period before group key requests are sent to the KGC is the off-line phase. KGC can carry out most calculation of group key generation during this phase. When KGC receives group key request messages from users, the protocol goes into on-line phase. Then, KGC generates group key and distributes it to users. In this way, KGC has a rapid response in on-line phase.

According to the above, we summarize the characteristics of our on-line/off-line group key distribution protocol as below.

- Each user just needs to share a coordinate with the KGC in privacy during registration. And then, all messages can be transported publicly.
- All users share a common group key, and they can verify whether or not the key is sent from KGC.
- When some users drop out of a group, they will not get any information about new group keys.
- The process of group key distribution is divided into on-line and off-line phases.
 KGC carries out most computation during off-line phase, so that it distributes the key to users rapidly in on-line phase.
- Our protocol is resistant to passive attack, impersonation attack, and reply attack. Moreover, it is not based on any hard mathematics problems.

The rest of the paper is organized as follows: In next section, we present some preliminaries about Shamir's (t, n) threshold SS and Harn-Lin's protocol. The entities and attack models are given in section 3. In section 4, we propose our group key distribution protocol in detail. Correctness analysis and security analysis are given in section 5 and section 6 respectively. In section 7, we compare our protocol to related ones. Section 8 presents some experimental data and section 9 concludes the paper.

2 Preliminaries

2.1 Shamir's (t, n) threshold SS

Shamir's (t, n) threshold SS is based on polynomial interpolation. It consists of the following two steps:

Share Generation: At first, the dealer D randomly chooses a polynomial f(x) of degree t - 1: $f(x) = a_0 + a_1x + \ldots + a_{t-1}x^{t-1} \mod p$. The constant term a_0 is regarded as the secret s, i.e., $s = f(0) = a_0$. All the coefficients $a_0, a_1, \ldots, a_{t-1}$ are in the finite field GF(p), where p is a prime number. The dealer D selects n abscissas $\mathcal{X} = 1, 2, \ldots, n$ to generate a share set $\delta = \{s_i | f(x_i), i \in \mathcal{X}\}$. Then, the dealer D distributes each share s_i to the corresponding shareholder U_i secretly while makes all the abscissas x_i known to the public.

Secret Reconstruction: When m $(m \geq t)$ legal shareholders attempt to recover the secret s, they pool shares together privately. As a result, each shareholder get a share set $\delta_m = \{s_i | f(x_i), i \in \mathcal{X}_m\}$ associated to the abscissas set $\mathcal{X}_m = \{1, 2, ..., m\}$. Then, the secret s can be computed as

$$s = f(0) = \sum_{i \in \mathcal{X}_m} s_i \prod_{j \in \mathcal{X}_m, j \neq i} \frac{-x_j}{x_i - x_j} \mod p$$

where $s_i \prod_{j \in \mathcal{X}_{I_m}, j \neq i} \frac{-x_j}{x_i - x_j}$ is a Lagrange component.

Remark 2.1. In Shamir's (t, n) threshold SS, if a shareholder gets any t or more than t shares, it can compute the secret s easily by Lagrange Interpolation Polynomial. However, any less than t shareholders cannot obtain any information about the secret. Likewise, less than t-1 Lagrange components also cannot recover generation polynomial f(x) or the secret s. It plays an important role in our proposed protocol. Due to the problem without any computational assumption, Shamir's (t, n) threshold SS is an information theoretically secure scheme.

2.2 Harn-Lin's protocol

In this protocol, the KGC first picks a RSA-number n, where n = p * q. Both p and q are large safe primes. Then, KGC shares a coordinate (x_i, y_i) with each user U_i during registration, where i = 1, 2, ..., m. The group key generation and distribution process is described as follow:

Step 1: The initiator sends a key initiation message to KGC.

Step 2: KGC broadcasts a response message to all users.

Step 3: Each user sends a random challenge R_i to KGC, where $R_i \in Z_n$.

Step 4: KGC randomly selects a group key k and constructs an m-th degree polynomial f(x) to pass through (m + 1) coordinates, (0, k) and $(x_i, y_i \oplus R_i)$, for i = 1, 2, ..., m. KGC computes m additional coordinates P_i on f(x) and authentication information $auth = h(k, U_1, ..., U_m, R_1, ..., R_m)$, where h(.) is a public

one-way hash function. KGC distributes $\{auth, P_1, ..., P_m\}$ to all users publicly. Step 5: Each user U_i can recover the group key k by one private coordinate (x_i, y_i) and m public coordinates P_i , for i = 1, 2, ..., m. Then, it computes $auth = h(k, U_1, ..., U_m, R_1, ..., R_m)$ to check whether or not the key is sent from KGC.

Remark 2.2. In Harn-Lin's protocol, KGC should choose a larger parameter n as the module of polynomials to replace the module p in Shamir's (t, n) threshold SS, where n is a RSA-number. In this way, the coordinate (x_i, y_i) of U_i can be reused many times and kept private. Harn and Lin use the big integer factorization problem to thwart Insider attack skillfully. Because this protocol is based on not only Shamir's (t, n) threshold SS but also hard math problem, it is not information theoretically secure. Readers can read the original paper [21] to acquire more detailed information about the problem. However, all computations have to be over \mathbb{Z}_n , and KGC cannot carry out any calculations of group key before it receives all users' random challenges. Consequently, Harn-Lin's protocol is inefficient.

3 Entities and attack models

In our group key distribution protocol, entities are classified into three types: 1) KGC, 2) User and 3) adversary. Furthermore, there are three attack models for adversaries: 1) passive attack, 2) impersonation attack, and 3) reply attack.

3.1 Entities

3.1.1 KGC

KGC is the entity trusted by all users and it is in charge of distributing group keys to the users. Firstly, it attests to users' identities and confers with each legal user on a private coordinate secretly. In fact, the process is similar to the shares generation in a (t, n) threshold SS scheme. Then, before users initiate a group key request, KGC goes into off-line phase to carry out part of computations about group key generation. Next, when KGC receives a group key initialization message, it goes into on-line phase to complete key generation. Finally, it sends group key distribution messages to users quickly.

$3.1.2 \ User$

A user is an entity who shares a private coordinate with KGC during registration. Afterwards, users are expected to maintain their own coordinates secure all the time, because the coordinates will be reused time and again. When users need a group key, each of them submits a request to KGC. After a user receives its corresponding distribution message, it recovers a session key and verifies whether or not the key is transferred from KGC.

$3.1.3 \ Adversary$

An adversary is an entity who wants to attack the protocol. Suppose that all users form a set \mathcal{U} . The adversaries are classified as Outsiders and Insiders according to whether or not they are in \mathcal{U} .

1) Insider: If a user not only wants to obtain the group key, but also attempts to derive coordinates kept by other users in \mathcal{U} , it is named Insider.

2) Outsider: If any others not in \mathcal{U} want to attack the protocol, they are called Outsiders. Outsiders always aim at capturing group key of \mathcal{U} , or preventing users in \mathcal{U} from obtaining a valid group key.

3.2 Attack models

3.2.1 Passive attack

Passive attack means that adversaries break confidentiality of a protocol by capturing messages among legal members. In our protocol, all messages during distribution phase are transported publicly, so that adversaries are able to easily capture the messages.

3.2.2 Impersonation attack

Impersonation attack refers to the fact that an entity pretends to be another to attack a protocol. In our protocol, it means that an Outsider sends group key request to KGC in the name of a legal user, or an adversary pretends to be KGC to distribute group key.

3.2.3 Reply attack

Reply attack is that an entity resends outdated messages to others. In our protocol, it is subdivided into two types:

1) An Outsider resends an outdated group key request to KGC.

2) An adversary redistributes an outdated group key to users.

4 Our proposed protocol

In this section, we show our group key distribution protocol in detail. It consists of three phases: 1) Preparatory phase, 2) Distribution phase, 3) Group key recovery and authentication.

4.1 Symbol definition

Before describing the protocol, we first define some notations as listed in Table 1.

Symbol	Descriptions
m	number of legal users
h(.)	hash function
k	group key selected by KGC
f(x)	group key generation function
U_i	the i^{th} user
U	group of legal users $\mathcal{U} = \{U_1, U_2,, U_m\}$
σ	group communication identifier
I	group key initialization message
R_{σ}	response message
$M_{\sigma,i}$	request message of U_i
(x_i, y_i)	private and permanent coordinate of user U_i
(x_i^*, y_i^*)	public coordinate selected by user U_i
$g_i(x)$	linear function constructed by (x_i, y_i) and (x_i^*, y_i^*)
$g_i^{-1}(y)$	inverse of $g_i(x)$
(x_i', y_i')	ephemeral coordinate select by KGC for user U_i
$\mathbf{x}_1,, \mathbf{x}_m$	public abscissas included in R_{σ}
d'_i	original group key distribution information
d_i	protected group key distribution information
\mathcal{K}_i	distribution message for U_i
Δ_i	Lagrange component of U_i
k_i	session key computed by U_i

Table 1: Notations

4.2 Preparatory phase

4.2.1 Initialization of KGC

The KGC selects a one-way hash function h(.) and a random prime p. Then, it makes them publicly known.

4.2.2 Users registration

As described above, each user should register with KGC before it joins in the group. In this process, each user U_i is required to share a private coordinate (x_i, y_i) with KGC, where both x_i and y_i are in the finite field GF(p). KGC should guarantee each $x_i \neq 0$ and $x_i \neq x_j$ for $i \neq j$. Each user just needs to register only once. Then, it makes its identity U_i public while keeps its coordinate (x_i, y_i) private.

4.3 Distribution phase

4.3.1 Off-line phase

step 1: Suppose that there are total m legal users who have registered with KGC. They form a user group $\mathcal{U} = \{U_1, U_2, ..., U_m\}$ and their private coordinates constitute a set $\Omega = \{(x_1, y_1), (x_2, y_2), ..., (x_m, y_m)\}$.

step 2: KGC randomly generates a polynomial $f(x) = a_0 + a_1x + \dots + a_mx^m$



Fig. 1: Sketch of users registration

of degree m and chooses a_0 as the group key k, i.e., $k = f(0) = a_0$.

step 3: KGC picks 2m different coordinates on f(x) to form other two sets $\Omega_1 = \{(x'_1, y'_1), (x'_2, y'_2), ..., (x'_m, y'_m)\}$ and $\Omega_2 = \{(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), ..., (\mathbf{x}_m, \mathbf{y}_m)\}$ such that $\Omega \cap \Omega_1 = \Omega \cap \Omega_2 = \emptyset$.

step 4: KGC uses each x'_i (i = 1, 2, ..., m) in Ω_1 and all elements in Ω_2 to compute original group key distribution information such as

$$d'_{i} = \sum_{t=1}^{m} \mathbf{y}_{t} \frac{-x'_{i}}{\mathbf{x}_{t} - x'_{i}} \prod_{j=1, j \neq t}^{m} \frac{-\mathbf{x}_{j}}{\mathbf{x}_{t} - \mathbf{x}_{j}} \bmod p$$

step 5: KGC computes an offset $h(x'_i, y'_i)$ to generate a value $d_i = d'_i + h(x'_i, y'_i) \mod p$ as protected group key distribution information.

4.3.2 On-line phase

step 1: The initiator sends a group key initialization message I to KGC.

step 2: When KGC receives the initialization message, it goes into on-line phase and broadcasts a message $R_{\sigma} = \{\sigma, \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m\}$ as response, where σ is a group communication identifier selected by KGC.

step 3: Each user U_i randomly picks a coordinate (x_i^*, y_i^*) , where $x_i^* \neq x_i$. Then it sends to KGC a request message $M_{\sigma,i}$ including the group communication identifier σ , its identity U_i and the selected coordinate (x_i^*, y_i^*) , such as

$$M_{\sigma,i} = \{\sigma, U_i, (x_i^*, y_i^*)\}$$

step 4: For each user U_i , KGC keeps a private coordinate (x_i, y_i) , and receives (x_i^*, y_i^*) in message $M_{\sigma,i}$. It uses the two coordinates to construct a linear function

$$g_i(x) = y_i \frac{x - x_i^*}{x_i - x_i^*} + y_i^* \frac{x - x_i}{x_i^* - x_i} \mod p$$

step 5: KGC uses (x'_i, y'_i) in Ω_1 to compute two values $g_i(x'_i)$ and $g_i^{-1}(y'_i)$, where $g_i^{-1}(y)$ is the inverse of $g_i(x)$.

step 6: KGC generates and sends distribution message $\mathcal{K}_i = \{U_i, g_i(x'_i), g_i^{-1}(y'_i), d_i, h(k, \sigma)\}$ to the corresponding user U_i , where $h(k, \sigma)$ is authentication information about the group key.

step 7: KGC completes the group key distribution and goes into off-line phase again to wait for next group key request.

4.4 Group key recovery and authentication

step 1: Each user U_i also constructs $g_i(x)$ and $g_i^{-1}(y)$ by (x_i, y_i) and (x_i^*, y_i^*) . After U_i receives message \mathcal{K}_i from KGC, it recovers $x'_i = g_i^{-1}(g_i(x'_i))$ and $y'_i = g_i(g_i^{-1}(y'_i))$.

step 2: Each user U_i uses the coordinate (x'_i, y'_i) , *m* public abscissas $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$ in message R_{σ} , to compute a Lagrange component, such as

$$\Delta_i = y'_i \prod_{j=1}^m \frac{-\mathbf{x}_j}{x'_i - \mathbf{x}_j} \bmod p$$

step 3: The group key can be computed as

$$k_i = d_i + \Delta_i - h(x'_i, y'_i) \mod p$$

step 4: Each user U_i also uses the same one-way hash function to compute a value

$$h_i = h(k_i, \sigma)$$

If $h_i = h\{k, \sigma\}$ holds, it means that the group key is the correct key sent from KGC, i.e. $k_i = k$. Then, U_i is allowed to use the certified key k to communicate with others in \mathcal{U} . Otherwise, users should initiate a new group key require to ensure their communication security.





5 Correctness analysis

In this section, we give two theorems to prove the correctness of our proposed protocol: 1) all users in \mathcal{U} can obtain a common group key; 2) the group key can be validated by the authentication information $h\{k, \sigma\}$.

Theorem 5.1. All the legal users can compute a common group key k, i.e. $k = d_i + \Delta_i - h(x'_i, y'_i) \mod p = d_j + \Delta_j - h(x'_j, y'_j) \mod p$, where $i \neq j$.

Proof. Like secret reconstruction in Shamir's (t, n) threshold SS, each user U_i requires at least m + 1 Lagrange components to recover the group key, due to the generation function f(x) with degree m. However, KGC uses only m coordinates $(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), ..., (\mathbf{x}_m, \mathbf{y}_m)$ and the abscissa x'_i to compute d'_i , which is the sum of m Lagrange components, and it sends the value $d_i = d'_i + h(x'_i, y'_i) \mod p$ to U_i . Therefore, the user U_i need just compute its own Lagrange component Δ_i and the offset $h(x'_i, y'_i)$ to recover the group key k. That is

$$\begin{aligned} & d_i + \Delta_i - h(x'_i, y'_i) \mod p \\ &= d'_i + h(x'_i, y'_i) + \Delta_i - h(x'_i, y'_i) \mod p \\ &= d'_i + \Delta_i \mod p \\ &= y'_i \prod_{j=1}^m \frac{-\mathbf{x}_j}{x'_i - \mathbf{x}_j} + \sum_{t=1}^m \mathbf{y}_t \frac{-x'_i}{\mathbf{x}_t - x'_i} \prod_{j=1, j \neq t}^m \frac{-\mathbf{x}_j}{\mathbf{x}_t - \mathbf{x}_j} \mod p \\ &= k \qquad \Box \end{aligned}$$

Theorem 5.2. Each user U_i is assured that its calculated group key is sent from KGC, if $h_i = h(k, \sigma)$ holds.

Proof. During distribution phase, only KGC knows the group key k. Thus, the valid authentication information of k can be computed by none but KGC. After distribution phase, each user U_i computes a key k_i . It also uses the hash function h(.) with k_i and σ to compute $h_i = h(k_i, \sigma)$. Obviously, if and only if $h_i = h(k, \sigma)$ (without regard for hash collision), its calculated key k_i is equal to k which is selected by KGC.

6 Security analysis

Detailed security analyses are presented in this section. We first give a lemma as the security foundation. Then, five theorems are presented to prove that our protocol is resistant to passive attack, impersonal attack and reply attack.

Note that all the calculations are over finite field GF(p), thus an event is deemed to be impossible if the probability of its occurrence is equal to or less than 1/p.

Lemma 6.1. Suppose that there are t public coordinates $(x_1, y_1), (x_2, y_2), ..., (x_t, y_t),$ (each $x_i \neq 0$), on a generation function g(x) with degree t - 1, such as $g(x) = a_0 + a_1x + a_2x^2 + ... + a_tx^t \mod p$, where all the parameters are in the finite field GF(p). The probability of reconstructing generation function g(x) by the t known coordinates is only 1/p.

Proof. Obviously, there are t+1 unknown parameters $a_0, a_1, ..., a_t$ for the function g(x) with degree t. Now, assume that the constant term a_0 is a known quantity. It means we get an extra coordinate (x_0, y_0) , where $x_0 = 0$ and $y_0 = a_0$. In this

way, we can use Lagrange Interpolation Polynomial to construct a function, such as

$$G(x) = \sum_{i=0}^{t} y_i \prod_{j=i, j \neq i}^{l} \frac{x - x_j}{x_i - x_j} \mod p$$

Due to $a_0 \in \mathbb{Z}_p$, there are totally p candidates for the value of a_0 . Therefore, different p functions can be computed corresponding to the different p possible values of a_0 , but only one function is the correct generation function f(x). In other words, the probability of reconstructing t-th generation function g(x) by t coordinates is only 1/p.

6.1 Resistance to passive attack

Model of passive attack: During distribution phase, all the messages are transported publicly. Thus, adversaries can easily capture response message R_{σ} , request messages $M_{\sigma,i}$, and distribution messages \mathcal{K}_i .

Theorem 6.1. Outsiders cannot obtain the group key and no adversary can derive the private coordinates of legal users by passive attack.

Proof. As for an Outsider, if it wants to obtain the group key, there are two methods:

1) The Outsider attempts to recover the generation function f(x). In order to reconstruct f(x) of degree m, an Outsider needs at least m+1 coordinates on f(x) or m+1 valid Lagrange components. However, no complete coordinate on f(x) is included in these public messages. KGC broadcasts only m protected sums of Lagrange components. Hence, the generation function f(x) cannot be recovered.

2) The Outsider tries to obtain a valid Lagrange component Δ_i and the corresponding offset $h(x'_i, y'_i)$ to compute $k = d_i + \Delta_i - h(x'_i, y'_i) \mod p$ just like a legal user U_i . If so, it must know the exact coordinate (x'_i, y'_i) , because of the following two reasons:

I: In equation $\Delta_i = y'_i \prod_{j=1}^m \frac{-\mathbf{x}_j}{x'_i - \mathbf{x}_j} \mod p, \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$ are public while x'_i and

 y'_i are unknown. So, the Outsider needs to know x'_i and y'_i to compute Δ_i .

II: Because of unidirectional characteristic of one-way hash function, the Outsider must know the values of x'_i and y'_i to compute the offset $h(x'_i, y'_i)$.

However, KGC only releases $g_i(x'_i)$ and $g_i^{-1}(y'_i)$ included in \mathcal{K}_i . The Outsider still needs the function $g_i(x)$ and $g_i^{-1}(y)$ to compute (x'_i, y'_i) . During *on-line phase*, the Outsider can capture only one coordinate (x^*_i, y^*_i) on $g_i(x)$ from $M_{\sigma,i}$. On the basis of **Lemma 6.1.**, the probability of reconstructing the linear polynomial $g_i(x)$ from only one coordinate is 1/p.

If an Outsider wants to obtain a private coordinate (x_i, y_i) , the probability is much less than 1/p. Because even if an Outsider gets one function $g_i(x)$, it still does not know which coordinate on $g_i(x)$ is (x_i, y_i) . It must capture at least two functions $g_i(x)$ s from different group key distribution processes, and the intersection of these function images is (x_i, y_i) . But actually, the probability of an Outsider obtaining any one function $g_i(x)$ is only 1/p, not to mention more functions. As for an Insider, on the one hand, it is a legal user in \mathcal{U} , so that it is able to compute the group key easily from its corresponding distribution message. On the other hand, it still wants to obtain some private coordinates (x_i, y_i) of other users. For this purpose, an Insider also needs to know (x'_i, y'_i) as foundation. There are two methods that might work for an Insider:

I: If the Insider attempts to recover the function $g_i(x)$ by only one coordinate (x_i^*, y_i^*) like an Outsider, the probability has been given above and it is much less than 1/p.

II: The Insider uses group key k to compute $\tau_i = (k - d_i) \mod p$ and get an equation

$$\begin{aligned} \tau_i &= \Delta_i - h(x'_i, y'_i) \mod p \\ \tau_i &= y'_i \prod_{j=1}^m \frac{-\mathbf{x}_j}{x'_i - \mathbf{x}_j} - h(x'_i, y'_i) \mod p \end{aligned}$$

In this equation, only x'_i and y'_i are unknown. However, $h(x'_i, y'_i)$ is a one-way hash value. It is impossible to compute exact (x'_i, y'_i) from the equation. Otherwise, it runs counter to the unidirectional characteristic of h(.).

As a result, an Outsider could capture neither the group key nor any valid coordinates by passive attack. For an Insider, it can only obtain the group key but no other users' coordinates can be derived. $\hfill \Box$

6.2 Resistance to impersonation attack

Model of impersonation attack 1: An Outsider can pretend to be a legal user U_i . And then, it also sends a request message $M_{\sigma,i} = \{\sigma, U_i, (x_i^*, y_i^*)\}$ to KGC and receives a corresponding distribution message $\mathcal{K}_i = \{g_i(x'_i), g_i^{-1}(y'_i), d_i, h(k, \sigma)\}$.

Theorem 6.2. Outsiders can neither obtain the group key nor prevent other legal users from obtaining the group key by impersonation attack 1.

Proof. In order to compute the group key according to \mathcal{K}_i , the Outsider must reconstruct the function $g_i(x) = y_i \frac{x - x_i^*}{x_i - x_i^*} + y_i^* \frac{x - x_i}{x_i^* - x_i} \mod p$ to recover the coordinate (x'_i, y'_i) . Although the Outsider can select (x_i^*, y_i^*) by itself, it still does not know the exact coordinate (x_i, y_i) . On account of **Lemma 6.1.**, the Outsider cannot reconstruct the function $g_i(x)$. Therefore, it is not able to obtain the group key by impersonation attack 1.

Meanwhile, for a legal user U_j , it can still recover the coordinate (x'_j, y'_j) to compute the group key as usual. That is to say, legal users are not influenced by the Outsider's mock request message.

Model of impersonation attack 2: An adversary imitates KGC to distribute group key. It first picks an *m*-th polynomial f(x). After the initiator sends a group key initialization message, it broadcasts a response message $R_{\sigma} = \{\sigma, \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m\}$. Then after users send request messages, it sends distribution message \mathcal{K}_i to U_i .

Theorem 6.3. Any adversary cannot distribute a fake group key to the users by impersonation attack 2.

Proof. Although the adversary can also select f(x) and compute distribution information d_i like KGC does in off-line phase, it still cannot generate the function $g_i(x)$ to compute accurate values of $g_i(x'_i)$ and $g_i^{-1}(y'_i)$ in \mathcal{K}_i because it does not know (x_i, y_i) of U_i . If the adversary uses false values $fg_i(x'_i)$ and $fg_i^{-1}(y'_i)$ to construct $\mathcal{K}_i = \{U_i, fg_i(x'_i), fg_i^{-1}(y'_i), d_i, h(k, \sigma)\}$, users will compute a different group key from the constant term in f(x) picked by the adversary. Then, the user will be aware of receiving a fake group key when it uses the authentication information in \mathcal{K}_i to check its calculated key. In brief, an adversary is completely powerless to distribute a fake group key by impersonation attack 2.

6.3 Resistance to reply attack

Model of reply attack 1: Suppose that an Outsider obtains an outdated group key ok and a request message $oM_{o\sigma,i} = \{o\sigma, U_i, (ox_i^*, oy_i^*)\}$. When the group requests a new key, the Outsider uses new group communication identifier σ to revise the outdated request message as $M_{\sigma,i} = \{\sigma, U_i, (ox_i^*, oy_i^*)\}$. Then, it sends $M_{\sigma,i}$ to KGC.

Theorem 6.4. An Outsider cannot obtain an outdated group key by reply attack 1.

Proof. As described in section 4, each group key is selected by KGC during the off-line phase. In other words, the process of key generation is completely independent of users' request messages. Even if all users utilize old coordinates (ox_i^*, oy_i^*) to send request messages, they still obtains a new group key k foreign to previous key ok. Therefore, an Outsider cannot obtain any information about the new group key by reply attack 1.

Model of reply attack 2: Suppose that an adversary obtains an outdated group key ok. Define an outdated request message as $oM_{o\sigma,i} = \{o\sigma, U_i, (ox_i^*, oy_i^*)\}$ and outdated distribution message as $o\mathcal{K}_i = \{U_i, og_i(ox_i'), og_i^{-1}(oy_i'), od_i, h(ok, o\sigma)\}$. When users utilize new coordinate (x_i^*, y_i^*) to request a new group key, the adversary resends the outdated response and distribution messages to them.

Theorem 6.5. An adversary cannot distribute an outdated group key to users by reply attack 2.

Proof. In order to let legal user recover group key, KGC is required to construct g(x) to transport (x'_i, y'_i) , where g(x) is generated by private coordinate (x_i, y_i) and public coordinate (x^*_i, y^*_i) . The private coordinate (x_i, y_i) kept by U_i is fixed while (x^*_i, y^*_i) is always changed for different request messages $M_{\sigma,i}$. As described attack in *Model of reply attack* 2, if (x_i, y_i) , (x^*_i, y^*_i) and (ox^*_i, oy^*_i) are not in a straight line, the outdated function og(x) is different from new function g(x). As a result, the user U_i is not able use g(x) to recover the previous coordinate (ox'_i, oy'_i) from $\mathcal{K}_i = \{U_i, og_i(ox'_i), og_i^{-1}(oy'_i), od_i, h(ok, \sigma)\}$, so that it cannot obtain the outdated group key. In brief, an adversary cannot distribute an outdated group key to users by reply attack 2.

Corollary 6.1. If a legal user drops out of the group, it will not get any information about the new group key. **Proof.** If a user drops out of the group but still wants to obtain new group keys, the KGC will not compute and send its corresponding distribution messages. So, it is deemed as an Outsider. In virtue of **Theorem 6.1. - 6.5.**, an Outsider can not obtain any information of group keys distributed by KGC. Therefore, the corollary holds.

7 Properties

In this section, we compare our proposed protocol with Harn-Lin's protocol [12] and Liu et al.'s protocol [20], because all the three protocols are group key distribution (transfer) protocols based on Shamir's (t, n) threshold SS. The nine aspects are considered in comparisons: 1) hard problems, 2) security, 3) on-line/off-line mechanism, 4) computation complexity for KGC, 5) computation complexity for a user, 6) communication overhead, 7) storage complexity, 8) response speed of the group key request and 9) calculation speed of group key recovery and authentication.

7.1 Hard problem

In all the three protocols, each user is just required to register at the KGC only once, and then it should keep the private coordinate (x_i, y_i) as a long-term secret. For keeping the private coordinate unknown to the others, both Harn-Lin's and Liu et al.'s protocols are based on integer factorization problem. But in our protocol, the KGC distributes sums of *Lagrange components* protected by hash values instead of coordinates on the generation function f(x), so that no hard problems are needed in our protocol.

7.2 Security

On account of cask principles, the security of a protocol depends on the weakest point. Because Shamir's SS is unconditional secure, the security of Harn-Lin's and Liu et al.'s protocols is based on large integer factorization problem, while our protocol depends on one-way hash function.

All the three protocols are resistant to passive attack and impersonation attack. Liu et al.'s protocol and our protocol can also resist reply attack. However, in the light of Nam et al.'s paper [23], Harn-Lin's protocol is vulnerable to reply attack. For detailed attack method, reader can refer to the original paper.

7.3 On-line/off-line mechanism

In both Harn-Lin's and Liu et al.'s protocols, the KGC has to wait users to send group key request messages. Only when KGC receives all messages, it starts to pick group key k, construct generation function f(x), select coordinates on f(x) and compute authentication value. In other words, KGC cannot do anything before it receives request messages from users. Instead, the KGC in our protocol is able to carry out most computations in off-line phase (before users send request messages). As a result, KGC has a very short response time in on-line phase (after users send request messages).

The above comparisons are shown in the Table 2.

protocol	Harn-Lin's	Liu et al.'s	Our protocol		
	protocol	protocol			
hash function	1	2	1		
hard problem	yes	yes	no		
resistance to passive attack	yes	yes	yes		
resistance to impersonation attack	yes	yes	yes		
resistance to reply attack	no	yes	yes		
on-line/off-line	no	no	yes		

Table 2: Comparisons table

7.4 Computation complexity for KGC

In Harn-Lin's protocol, KGC performs about m exclusive - OR operations, (m-1)(m+1) additive operations, $m^2 + (m-1)^2$ multiplicative operations and 1 hash operation to generate group key distribution messages, where m is the total number of legal users. Liu et al.'s protocol needs m more hash operations than Harn-Lin's protocol. Hence, the computational complexity is $O(m^2)$ of the two protocols. In our proposed protocol, the KGC performs m(m+3) additive operations, m(m+1) multiplicative operations and m + 1 hash operations in off-line phase. Thus, the computational complexity is also $O(m^2)$. However, the computation is over \mathbb{Z}_p in our protocol while it is over \mathbb{Z}_n in the other protocols, where n is much greater than p. Moreover, because KGC carries out most computation in off-line phase, it just needs 2m + 3 additive operations, 2m + 2 multiplicative operations and 1 hash operation in on-line and the computational complexity is just O(m).

7.5 Computation complexity for a user

After a user receives group key distribution message from the KGC, it starts to recover and verify the group key. In Harn-Lin's protocol, a user needs 1 exclusive -OR operation, about m^2 additive operations, m^2 multiplicative operations, and 1 hash operation. Liu et al.'s protocol needs 1 more hash operation than Harn-Lin's protocol. Hence, the computational complexity are $O(m^2)$ for the two protocols. In our protocol, a user just needs m+10 additive operations, m+10 multiplicative operations and 2 hash operations. Hence, the computational complexity is O(m).

7.6 Communication overhead

In both Harn-Lin and Liu et al. protocols, users send to KGC about 2m numbers and KGC transports 3m numbers to users. Thus, the communication overhead is

5m numbers and each number is in \mathbb{Z}_n . In our protocol, users send to KGC about 3m numbers and KGC transports 6m numbers to users. Hence, the communication overhead is 9m numbers and each number is in \mathbb{Z}_p . However, KGC in our protocol has to sends each group key distribution message to the corresponding user one by one. KGC in the other protocols can broadcast distribution messages to all users.

7.7 Storage complexity

In both Harn-Lin and Liu et al. protocols, KGC just needs to store m private coordinates of users. In other words, it keeps 2m numbers which are in \mathbb{Z}_n . But in our protocol, KGC has to keep 6m numbers including m private coordinates (x_i, y_i) , m ephemeral coordinates (x'_i, y'_i) , m abscissas \mathbf{x}_i and m protected group key distribution values d_i . Moreover, each user need 2 numbers to store its private coordinate in all the three protocols.

7.8 Response speed of group key request

In Harn-Lin's protocol, before KGC receives users' request messages, it cannot carry out any computation. After that, the computation complexity for KGC is $O(m^2)$. Liu et al.'s protocol needs m more hash operations than Harn-Lin's protocol. Hence, the computational complexity is also $O(m^2)$. But in our proposed protocol, KGC carries out most computation in off-line phase and the computational complexity for KGC is just O(m) in of-line phase. Therefore, the response speed of group key request is faster than it in the other protocols.

7.9 Speed of group key recovery and authentication by users

After a user receives group key distribution message from the KGC, it starts to recover and verify the group key. In both Harn-Lin's and Liu et al.'s protocols, the computational complexity for a user is $O(m^2)$ while it is just O(m). Therefore, the speed of group key recovery and authentication by users in our protocol is also faster than it in the other protocols.

8 Experiment

In this section, we verify the theories in **7.5** and **7.6** through experiments. Because the computational complexity of Liu et al's protocol approximately equals Harn-Lin's, we just compare our protocol with Harn-Lin's. We chose a RSA-number with 1024 bits as the modulus in Harn-Lin's protocol, where n = 13506641086599522334 96032162788059699388814756056670275244851438515265106048595338339402871 50571909441798207282164471551373680419703964191743046496589274256239341 02086438320211037295872576235850964311056407350150818751067659462920556 36855294752135008528794163773285339061097505443349998111500569772368909 27563. And SHA-256 is used in both protocols as the hash function. Because the hash value of SHA-256 bits, we choose a 260 bits number as the modulus for

our protocol, where p = 15476004057485807498789328096422564586042495822915855764974651495000499083147583.

From the perspective of cryptography, the computational complexity of factoring n is about $e^{\sqrt{\ln(n)\ln(\ln n)}} \approx e^{70}$. Although SHA-1 and MD5 have been proved insecurity by Wang et al. [33,32], there is still no more effective method than birthday attack to break SHA-256 currently. The computational complexity of breaking SHA-256 by birthday attack is 2^{128} , which is greater than e^{70} . Therefore, our protocol is more secure than the other in this experiment.

Remark 8.1. In our protocol, the hash value should be in \mathbb{Z}_p for security reason. Otherwise, it increases the possibility of hash function collision when different hash values modulo p. For efficiency reason, the smaller modulus p is, the faster KGC and users calculate. Therefore, the modulus p should be set as a number which is a bit greater than the hash value.

Table 3 shows the simulation environment and conditions of the experiment.

CPU	Core i7-3630QM 2.40GHz			
RAM	8G DDR3 1333MHz			
Operation system	Windows 10 Enterprise 64-bit			
Programming language	Python			
Programming software	Pycharm			

Table 3: Simulation environment and conditions

Protocol User number	Harn-Lin's protocol	Our protocol
50	4.571370	0.104552
100	20.673966	0.134091
150	52.657993	0.259405
200	103.841224	0.328321
250	211.968105	0.420336
300	419.852612	0.633199
350	634.937633	0.805593
400	935.913335	1.047877
450	1219.035164	1.196793
500	1593.049735	1.295323

Table 4: Response time of group key request (s)

Analysis 1: From Table 4 and Table 5, it can be seen that, for the same user number, both response time of group key request and time of key recovery and authentication in Harn-Lin's protocol are far longer than those in our protocol. There are two reasons:

1: The computational complexity of Harn-Lin's protocol is $O(m^2)$ while our protocol is O(m) in on-line phase.

2: More importantly, the computation of Harn-Lin's protocol is over \mathbb{Z}_n , where n is 1024 bits. But it of our protocol is over \mathbb{Z}_p , where p is 260 bits.

Table 5. This of key recovery and authentication (5)					
Protocol User number	Harn-Lin's protocol	Our protocol			
50	3.947241	0.019082			
100	18.507578	0.037025			
150	40.831115	0.054113			
200	67.566742	0.091200			
250	110.791448	0.128122			
300	202.860842	0.160992			
350	229.557842	0.210194			
400	312.193951	0.245661			
450	483.446120	0.285706			
500	616.605906	0.315709			

Table 5: Time of key recovery and authentication (s)





Fig. 4: Time of our protocol

Analysis 2: Fig 3 and Fig 4 show that, with the rise of user number, the speed of time increase in Harn-Lin's protocol is faster than our protocol. This means that the computational complexity of Harn-Lin's protocol is higher than our protocol.



Fig. 5: Time of our protocol with massive users

Moreover, in order to show the high-efficiency of our protocol, we use massive users to collect more data about time of group key request and key recovery in our protocol. The results are shown in Fig 5.

9 Conclusion

In this paper, we propose a secure on-line/off-line group key distribution protocol. In the proposed protocol, only users' registration messages need to be transported in privacy, while all the other messages can be transported in public. In terms of safety, the protocol is resistant to passive attack, impersonation attack and reply attack. And it supports authentication function. More importantly, KGC can carry out most computations of group key generation in off-line phase. Consequently, the response speed of group key request and calculation speed of group key recovery and authentication can be greatly improved.

Acknowledgement

This work is supposed by National Natural Science Foundation of China under 61572454, 61572453, 61472382, 61520106007.

References

- 1. Asmuth C, Bloom J. A modular approach to key safeguarding[J]. IEEE transactions on information theory, 29(2): 208-210 (1983).
- Blakley G R. Safeguarding cryptographic keys[C]//Proceedings of the national computer conference. 48: 313-317 (1979).
- Boyd C. On key agreement and conference key agreement[C]//Australasian Conference on Information Security and Privacy. Springer, Berlin, Heidelberg, 294-302 (1997).
- Bresson E, Chevassut O, Pointcheval D. Provably secure authenticated group Diffie-Hellman key exchange[J]. ACM Transactions on Information and System Security (TIS-SEC), 10(3): 10 (2007).
- Chen L Q, Sun C F, Xu C J. An Authenticated Group Key Agreement Scheme for Wireless Sensor Networks Based on Bilinear Pairings[C]//Advanced Materials Research. Trans Tech Publications, 846: 876-882 (2014).
- Diffie W, Hellman M. New directions in cryptography[J]. IEEE transactions on Information Theory, 22(6): 644-654 (1976).
- Even S, Goldreich O, Micali S. On-line/off-line digital signatures[C]//Conference on the Theory and Application of Cryptology. Springer, New York, NY, 263-275 (1989).
- Fiat A, Naor M. Broadcast encryption[C]//Annual International Cryptology Conference. Springer, Berlin, Heidelberg, 480-491 (1993).
- 9. Guo C, Chang C C. An authenticated group key distribution protocol based on the generalized Chinese remainder theorem[J]. International Journal of Communication Systems, 27(1): 126-134 (2014).
- Guo F, Mu Y, Chen Z. Identity-based online/offline encryption[C]//International Conference on Financial Cryptography and Data Security. Springer, Berlin, Heidelberg, 247-261 (2008).
- 11. Harn L, Hsu C F. A Practical Hybrid Group Key Establishment for Secure Group Communications[J]. The Computer Journal, 60(11): 1582-1589 (2017).
- Harn L, Lin C. Authenticated group key transfer protocol based on secret sharing[J]. IEEE transactions on computers, 59(6): 842-846 (2010).
- Hohenberger S, Waters B. Online/offline attribute-based encryption[C]//International Workshop on Public Key Cryptography. Springer, Berlin, Heidelberg, 293-310 (2014).
- Hsu C, Zeng B, Zhang M. A novel group key transfer for big data security[J]. Applied Mathematics and Computation, 249: 436-443 (2014).
- 15. Irshad A, Sher M, Chaudhary S A, et al. An efficient and anonymous multi-server authenticated key agreement based on chaotic map without engaging Registration Centre[J]. The Journal of Supercomputing, 72(4): 1623-1644 (2016).
- Kim Y, Perrig A, Tsudik G. Tree-based group key agreement[J]. ACM Transactions on Information and System Security (TISSEC), 7(1): 60-96 (2004).

- Kumar V, Kumar R, Pandey S K. A computationally efficient centralized group key distribution protocol for secure multicast communications based upon RSA public key cryptosystem[J]. Journal of King Saud University-Computer and Information Sciences (2018).
- Lai J, Mu Y, Guo F. Efficient identity-based online/offline encryption and signcryption with short ciphertext[J]. International Journal of Information Security, 16(3): 299-311 (2017).
- Liu J K, Baek J, Zhou J, et al. Efficient online/offline identity-based signature for wireless sensor network[J]. International Journal of Information Security, 9(4): 287-296 (2010).
- 20. Liu Y, Cheng C, Cao J, et al. An improved authenticated group key transfer protocol based on secret sharing[J]. IEEE Transactions on computers, 62(11): 2335-2336 (2013).
- Mayer A, Yung M. Generalized secret sharing and group-key distribution using short keys[C]//Compression and Complexity of Sequences 1997. Proceedings. IEEE, 30-44 (1997).
- Mignotte M. How to share a secret[C]//Workshop on Cryptography. Springer, Berlin, Heidelberg, 371-375 (1982).
- Nam J, Kim M, Paik J, et al. Cryptanalysis of a group key transfer protocol based on secret sharing[C]//International Conference on Future Generation Information Technology. Springer, Berlin, Heidelberg, 309-315 (2011).
- Naoui S, Elhdhili M E, Saidane L A. Lightweight Enhanced Collaborative Key Management Scheme for Smart Home Application[C]//High Performance Computing & Simulation (HPCS), 2017 International Conference on. IEEE, 777-784 (2017).
- 25. Olimid R F. On the security of an authenticated group key transfer protocol based on secret sharing[C]//Information and Communication Technology-EurAsia Conference. Springer, Berlin, Heidelberg, 399-408 (2013).
- 26. Shamir A. How to share a secret[J]. Communications of the ACM, 22(11): 612-613 (1979).
- Shamir A, Tauman Y. Improved online/offline signature schemes[C]//Annual International Cryptology Conference. Springer, Berlin, Heidelberg, 355-367 (2001).
- Steiner M, Tsudik G, Waidner M. Diffie-Hellman key distribution extended to group communication[C]//Proceedings of the 3rd ACM conference on Computer and communications security. ACM, 31-37 (1996).
- 29. Sun Y, Wen Q, Sun H, et al. An authenticated group key transfer protocol based on secret sharing[J]. Procedia Engineering, 29: 403-408 (2012).
- Tanaka H, Nakajima K, Ishigaki K, et al. Hybrid pen-input character recognition system based on integration of online-offline recognition[C]//Document Analysis and Recognition, 1999. ICDAR'99. Proceedings of the Fifth International Conference on. IEEE, 209-212 (1999).
- Ustaoglu B. Obtaining a secure and efficient key agreement protocol from (H) MQV and NAXOS[J]. Designs, Codes and Cryptography, 46(3): 329-342 (2008).
- Wang X, Yin Y L, Yu H. Finding collisions in the full SHA-1[C]//Annual international cryptology conference. Springer, Berlin, Heidelberg, 17-36 (2005).
- 33. Wang X, Yu H. How to break MD5 and other hash functions[C]//Annual international conference on the theory and applications of cryptographic techniques. Springer, Berlin, Heidelberg, 19-35 (2005).
- 34. Xu S, Mu Y, Susilo W. Online/Offline signatures and multisignatures for AODV and DSR routing security[C]//Australasian conference on information security and privacy. Springer, Berlin, Heidelberg, 99-110 (2006).
- 35. Yuan W, Hu L, Li H, et al. Security and improvement of an authenticated group key transfer protocol based on secret sharing[J]. Applied Mathematics & Information Sciences, 7(5): 1943 (2013).
- 36. Zhang L, Wu Q, Domingo-Ferrer J, et al. Round-efficient and sender-unrestricted dynamic group key agreement protocol for secure group communications[J]. IEEE Transactions on Information Forensics and Security, 10(11): 2352-2364 (2015).