sensors

Article

Privacy-Preserving Multi-Receiver Certificateless Broadcast Encryption Scheme with De-Duplication

Jianhong Zhang [1,2,3,*,†] **and Peirong Ou** [4,†]

[1] School of Information Sciences and Technology, North China University of Technology, Beijing 100144, China
[2] National Engineering Laboratory for Big Data Collaborative Security Technology, Beijing 100015, China
[3] Guangxi Key Laboratory of Cryptography and Information Security, Guilin 541004, China
[4] College of Sciences, North China University of Technology, Beijing 100144, China
* Correspondence: jhzhang@ncut.edu.cn
† These authors contributed equally to this work.

Received: 24 April 2019; Accepted: 23 July 2019; Published: 31 July 2019

Abstract: Nowadays, the widely deployed and high performance Internet of Things (IoT) facilitates the communication between its terminal nodes. To enhance data sharing among terminal devices and ensure the recipients' privacy protection, a few anonymous multi-recipient broadcast encryption (AMBE) proposals are recently given. Nevertheless, the majority of these AMBE proposals are only proven be securely against adaptively chosen plain-text attack (CPA) or selectively chosen ciphertext attack (CCA). Furthermore, all AMBE proposals are subjected to key escrow issue due to inherent characteristics of the ID-based public cryptography (ID-PKC), and cannot furnish secure de-duplication detection. However, for cloud storage, it is very important for expurgating duplicate copies of the identical message since de-duplication can save the bandwidth of network and storage space. To address the above problems, in the work, we present a privacy-preserving multi-receiver certificateless broadcast encryption scheme with de-duplication (PMCBED) in the cloud-computing setting based on certificateless cryptography and anonymous broadcast encryption. In comparison with the prior AMBE proposals, our scheme has the following three characteristics. First, it can fulfill semantic security notions of data-confidentiality and receiver identity anonymity, whereas the existing proposals only accomplish them by formalizing the weaker security models. Second, it achieves duplication detection of the ciphertext for the identical message encrypted with our broadcast encryption. Finally, it also avoids the key escrow problem of the AMBE schemes.

Keywords: data sharing; anonymous broadcast encryption; security proof; secure de-duplication

1. Introduction

With development of various Internet of Things (IoT) applications, the communication amongst smart IoT devices has become more and more frequent and convenient. As an important one-to-many communication model, broadcast encryption (BE, for short), which was first formally proposed by Amos Fiat and Moni Naor [1], allows for the broadcaster to deliver the encrypted data to the authorized subset S of the receivers that are monitoring the broadcast channel. In addition, only the receivers that belong to the subset S can recover the message by their private key, while the other receivers outside of S can obtain no information about the delivered data. In general, broadcast encryption is capable of saving more computational complexity and communication overhead than traditional encryption in the peer-to-peer model. Therefore, it has very important applications in communications field [2,3] and IoT [4], etc.

However, IoT devices have some non-negligible vulnerabilities during data sharing and anonymity protection [5–7]. At the same time, anonymity is also an important security property in

BE schemes, which indicates that any receiver is unable to gain any information of the other receivers' identity from the ciphertexts. Let us consider an example: a user wants to share some sensitive files with its friends in the cloud; for individual privacy, the user does not want its friends to learn about the others' identity because they might be the opponent. This problem is very similar to blind carbon copy (BCC) in the email system. To solve this problem, many cryptographers have given many solutions, for instance, Bellare et al.'s public key encryption with key-privacy [8], ciphertext-policy attribute-based encryption with hidden-policy [9], anonymous identity-based encryption [10], anonymous broadcast encryption [11–13], and anonymous Certificate-Based Encryption [14,15], where anonymous broadcast encryption is the most efficient method in the multi-user setting. In the cloud environment, anonymity is more important due to its openness. Thus, many applications in keyword search and data retrieval [16–18] have considered how to achieve strong anonymity in their schemes. The existing anonymous broadcast encryption schemes are classified into two types, one type is based on public key certificate, and the other type is based on ID-based cryptography. Attribute-based encryption provides scalable encryption while supporting anonymity for users in the same group, that is, with the same attributes [19,20]. They have also been applied widely in cloud computing to support access control for data sharing [21]. However, because of the open problem of revocation in attribute-based encryption, it still suffers from the user revocation in practical application [22–24]. Some of the corresponding data could easily be recovered from IoT devices by using forensic techniques [25,26]. Fortunately, Antonis Michalas et al. recently proposed two hybrid encryption schemes [27,28] which can solve the open problem of revocation in attribute-based encryption.

Although cloud storage servers have abundant storage space, the identical data's different encryption can result in multi-replica; this not only wastes space, but also brings a heavy burden on data maintenance. To save the storage space across multiple users in the cloud storage service, de-duplication is an important candidate technique. However, not all of the public encryption schemes can directly support the de-duplication of the ciphertext since random numbers are introduced in the encrypting process. The convergent encryption and the related security definition have been formalized for addressing the de-duplication of ciphertext [29]. Because random numbers are introduced in the encryption algorithm, it is very difficult that the existing anonymous multi-receiver ID-based broadcast encryption schemes (AMIBE) directly support the de-duplication of ciphertext. To overcome the above de-duplication problem, in this work, we propose a secure Privacy-preserving Multi-receiver Certificateless Broadcast Encryption Scheme with De-duplication. Our construction is characterised as follows: firstly, it is the first anonymous certificateless broadcast encryption scheme with de-duplication; secondly, it is capable of simultaneously achieving confidentiality and anonymity of the receivers' identities under adaptive CCA security. Thirdly, the key escrow problem does not exist.

2. Related Works

In 2006, Barth et al. presented the first public key cryptography-based anonymous BE scheme with chosen-ciphertext security [11]. However, the complexity of decryption linearly grows with the size of the set of the receivers. In 2012, Libert et al. put forth a fully anonymous BE scheme with adaptive chosen-ciphertext security with the random oracle model [30]. Subsequently, Fazio et al. proposed two sublinear ciphertext-size anonymous broadcast encryption schemes in [31] which are proven to be securely against adaptive CPA and adaptive CCA in the standard security model, respectively. In 2007, Delerablee constructed a constant-size ciphertext BE scheme [32]; however, the receivers' public keys need to be attached in the ciphertext. Until now, the PKC-based anonymous broadcast encryption scheme can achieve constant ciphertext and resist adaptive chosen-ciphertext attack (CCA) in the standard-security model.

ID-based BE (IBBE) is an extension of broadcast encryption in the ID-PKC system [33] in which the user's public key is replaced with the user's identity. It simplifies public key management and eliminates the public key certificate. To furnish anonymity protection of the receiver's identity, the first anonymous multi-receiver identity-based broadcast encryption (AMIBE) scheme [12] was introduced. Nevertheless, their scheme was shown to be insecure by Wang et al. [34] and Chien [35] since it can not achieve anonymity protection of the receiver's identity, whereafter, Wang et al. also presented a modified proposal to fulfill the anonymity of the receiver's identity in [34]. Very regretfully, Wang et al.'s modified proposal was pointed to be insecure by Zhang et al. in [36]. In 2018, Tseng et al. presented an improved vision of Fan et al.'s AMIBE by revising receiver anonymity's security definition in [37] and their scheme was shown to be secure in the random oracle model. In Asia-CCS16, based on the multilinear map, Xu et al. gave an AMIBE scheme which is against anonymity attacks and chosen-plaintext attacks in the standard model [38,39]. However, all multilinear map candidates are broken [40]; thus, their proposal is infeasible in reality. Recently, He et al. proposed an ID-based anonymous BE scheme that can concurrently achieve data indistinguishability and anonymity of the receiver identities under the adaptively chosen ciphertext attacks [41].

ID-based cryptographic protocols cut out complex maintenance of certificates; however, an inherent problem called "*key escrow*" exists. This problem can make the PKG be able to execute any cryptographic operation in the name of users since it knows all users' private keys. Thus, the problem might result in potential security threats for the ID-based crypto-system. To avoid the key escrow problem, Al-Riyami and Paterson gave a variant of ID-based PKC: certificateless cryptography in [42]. Not only do the advantages of ID-based cryptography remain, but they also prevent the key escrow problem of ID-based PKC. In 2004, Yum et al. presented a general construction construction of certificateless encryption (CLE) [43]. Unfortunately, Yum et al.'s scheme was shown to be insecure by Libert et al. in [44] since it does not satisfy CCA security of CLE. In addition, therewith, Libert et al. put forward a novel construction of CLE achieving CCA security.

Recently, lIslam et al. put forward a pairing-free anonymous multi-receiver certificateless encryption scheme (AMCLE, for short) by combining AMIBE with CLE in [45]. Their scheme can achieve receivers' anonymity and the ciphertext length is linear with the number of the authorized receivers. When more than one person sends the same data, it will bring a heavy burden to the receivers for data storage. Thus, de-duplication is a wise choice to address the growing demand for storage.

To reconcile de-duplication, Douceur et al. presented a method *convergent encryption* (CE) [46], which is a deterministic symmetric encryption with secret key $H(m)$. If two users Alice and Bob encrypt the same plaintext m, they can obtain the same ciphertext $C = E_{H(m)}(m)$. Its attractive merit makes it be applied in some commercial system. However, it lacks the detailed security analysis and it is not explicit what its basic security goal precisely is. To solve de-duplication of the identical message which is encrypted under the different secret keys, Bellare et al. put forth a novel notion *Message-Locked Encryption* (MLE) [47]. However, MLE is only capable of providing security of unpredictable data. Recently, Bellare et al. proposed an Interactive message-locked encryption and secure de-duplication [48] which can solve the correlated message's security problem. Until now, numerous secure de-duplication schemes have been presented for settling data de-duplication in cloud [49–51].

3. Preliminaries

3.1. Bilinear Groups

Throughout the paper, we only consider a Type 2 pairing since our scheme is based on such construction. In the following, we review some concepts of such bilinear group pair.

1. $\mathbb{G}_1$ and $\mathbb{G}_2$ denote two additional groups of the same prime p; $\mathbb{G}_T$ denotes a multiplicative group. In addition, it is deemed to be hard for solving the discrete logarithm problem in group $\mathbb{G}_i, i \in \{1, 2, T\}$.
2. P_i denotes the generator of group $\mathbb{G}_i$,for $i \in \{1, 2\}$.
3. Let $\varphi : \mathbb{G}_2 \to \mathbb{G}_1$ be a computable isomorphism map which satisfies $\varphi(P_2) = P_1$; and
4. Let $\hat{e} : \mathbb{G}_1 \times \mathbb{G}_2 \to \mathbb{G}_T$ denote a computable bilinear map, which meets the following criteria:

- Bilinearity: For arbitrary $a, b \in Z_p$ and all $Q \in \mathbb{G}_1$, $F \in \mathbb{G}_2$, we have $\hat{e}(aQ, bF) = \hat{e}(Q, F)^{ab}$;
- Non-degeneracy: $\hat{e}(P_1, P_2) \neq 1$.

3.2. Security Assumptions

In this subsection, we give several security assumptions [33,52] which are the security foundation to construct the proposed scheme.

ε-BDH-2 problem [33] **in** $(\mathbb{G}_1, \mathbb{G}_2)$. Given group elements $a_1P_2, b_1P_2 \in \mathbb{G}_2$ and $c_1P_1 \in \mathbb{G}_1$, where $P_2 \in \mathbb{G}_2$, $P_1 \in \mathbb{G}_1$, and $a_1, b_1, c_1 \in Z_p^*$; if there exists a PPT-algorithm $\mathcal{A}$ which takes $(P_1, P_2, a_1P_2, b_1P_2, c_1P_1)$ as inputs and outputs, the Type 2 pairing $X = e(P_1, P_2)^{a_1b_1c_1} \in \mathbb{G}_T$. $\mathcal{A}$'s advantage is defined as

$$\varepsilon = Pr[e(P_1, P_2)^{a_1b_1c_1} \leftarrow A(P_1, P_2, aP_2, bP_2, c_1P_1)].$$

We think that ε-bilinear Diffie–Hellman problem in $\mathbb{G}_2$ and $\mathbb{G}_1$ holds against $\mathcal{A}$ if the algorithm $\mathcal{A}$ is not capable of obtaining $\hat{e}(P_1, P_2)^{a_1b_1c_1}$ with a non-negligible probability greater than ε.

ε-BDDH-2 problem in $(\mathbb{G}_1, \mathbb{G}_2)$ [33]. It is hard to distinguish the distributions $D_1 = (P_1, P_2, a_1P_1, b_1P_1, c_1P_2, e(P_1, P_2)^{a_1b_1c_1})$ and $D_2 = (P_1, P_2, a_1P_1, b_1P_1, c_1P_2, Z)$, where $Z \in \mathbb{G}_T$ and $a_1, b_1, c_1 \in_R Z_p$. In general, D_1 is denoted as the BDDH tuple, and D_2 is called "random tuple". For a PPT algorithm $\mathcal{B}$, $\mathcal{B}$'s advantage of breaking the BDDH-2 problem in $(\mathbb{G}_2, \mathbb{G}_1)$ is defined as

$$\begin{aligned} \varepsilon = &|Pr[B(P_1, P_2, a_1P_1, b_1P_1, c_1P_2, e(P_1, P_2)^{a_1b_1c_1}) = 1] \\ &-Pr[B(P_1, P_2, a_1P_1, b_1P_1, c_1P_2, Z) = 1]|. \end{aligned}$$

We think that ε-decisional Bilinear Diffie–Hellman problem in $(\mathbb{G}_2, \mathbb{G}_1)$ holds against $\mathcal{B}$ if the algorithm $\mathcal{B}$ is capable of distinguishing the difference of the above two distributions in a non-negligible probability $\varepsilon > 1/2$.

The Computational Diffie–Hellman problem (CDH) in $\mathbb{G}_1$. Let $(P_1, a_1P_1, b_1P_1) \in \mathbb{G}_1^3$ be a random 3-tuple where $a_1, b_1 \in Z_p$; there does not exist an efficient algorithm $\mathcal{A}$ that can calculate abP_1. A's advantage of breaking the Computational Diffie–Hellman problem in $\mathbb{G}_1$ is defined as

$$\varepsilon = Pr[A(P_1, aP_1, bP_2) = abP_1].$$

We think that the CDH problem holds against $\mathcal{A}$ if the algorithm $\mathcal{A}$ is capable of outputting $a_1b_1P_1$ in a non-negligible probability ε.

The Decisional Diffie–Hellman problem (DDH) in $\mathbb{G}_1$. Given a 4-tuple $(P_1, a_1P_1, b_1P_1, W) \in \mathbb{G}_1$ where $a_1, b_1 \in Z_p$ and $W \in \mathbb{G}_1$, there does not exist an efficient algorithm $\mathcal{A}$ that determines $a_1b_1P_1 = W$. A's advantage of breaking the Decisional Diffie–Hellman problem in $\mathbb{G}_1$ is defined as

$$\begin{aligned} \varepsilon \quad = \quad &|Pr[A(P_1, a_1P_1, b_1P_1, a_1b_1P_1) = 1] - \\ &Pr[A(P_1, a_1P_1, b_1P_1, W) = 1]|. \end{aligned}$$

We think that the DDH problem holds against $\mathcal{A}$ if the algorithm $\mathcal{A}$ is capable of distinguishing the difference of $a_1b_1P_1$ and W in a non-negligible probability $\varepsilon > 1/2$.

4. Basic System Model and Security Model

4.1. System Model

According to the definitions of certificateless encryption and broadcast encryption, we give the basic system model of privacy-preserving multireceiver certificateless broadcast encryption with de-duplication (PMCBED) schemes. The PMCBED scheme mainly borrows the idea in [12,37,38] to achieve privacy protection of receiver identities in the certificateless broadcast encryption scheme and offer the ciphertext de-duplication function. Its framework is showed in Figure 1. It includes four entities: key generation center (KGC), the receivers, the broadcaster and the de-duplicator. Their detailed roles are shown as follows:

1. KGC: it is a trustworthy entity that is responsible for producing a partial private key of the receiver.
2. the Broadcaster: It is a sender of the message. It first selects a subset of the receivers and calculates the ciphertext of the transmitted message. Afterwards, it sends these ciphertexts to the de-duplicator.
3. The de-duplicator: It is an honest-but-curious entity. It can be acted on by the cloud server. Its goal is to check whether the received ciphertext has its replica existing in the cloud.
4. The Receiver: It is the receiver of the ciphertext, its goal is to decrypt the ciphertext. If and only if it is an authenticated receiver, then it can decrypt the ciphertext.

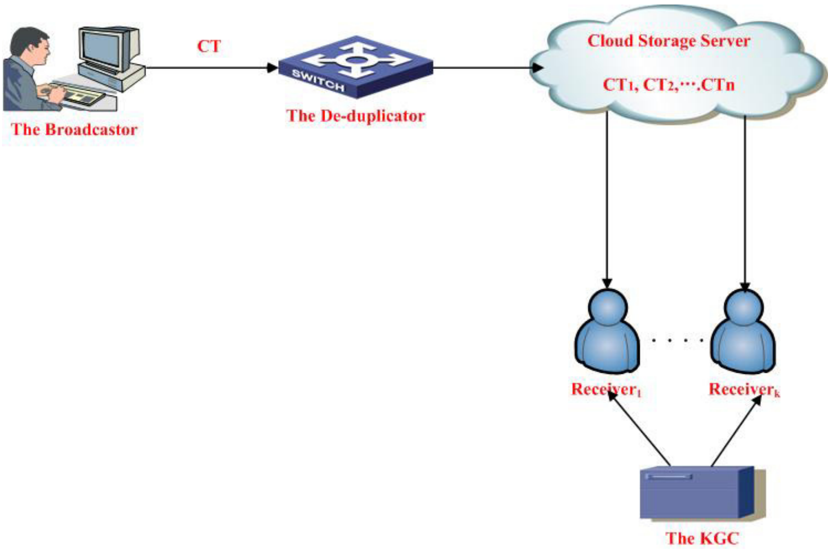

Figure 1. The system model of the PMCBED scheme.

For a PMCBED scheme, it has eight algorithms: **System-setup**, **Extract partial-private key**, **Set secret-value**, **Set-public-key**, **Set-private-key**, **Encryption**, **Decryption** and **Equality-test**. For each algorithm, its detailed definition is given as follows:

- System-setup (1^{λ}). λ is a security parameter, and this algorithm is run by a Key Generation Center (KGC) which takes as input λ, return the public parameters PP and the master secret key msk of KGC. The public parameters PP should be published publicly.
- Extract partial-private key (msk, ID). In general, this algorithm is run by KGC. It takes as inputs public parameters PP, master key msk and a receiver's identity ID, and outputs the partial-private key d_{ID} of the receiver.
- Set secret-value (ID). The algorithm is run by the receiver. It takes as inputs public parameters PP and the identity ID of the receiver, and returns x_{ID} as the receiver's secret value.
- Set private-key (x_{ID}, d_{ID}): This algorithm is run by the receiver, it takes as inputs the partial-private key d_{ID} and secret-value x_{ID} of the receiver, and outputs private key $SK_{ID} = (d_{ID}, x_{ID})$ of the receiver.
- Set public-key (ID): The algorithm is used to produce the public key of the receiver. It takes as inputs secret value x_{ID} of the receiver and public parameters PP, and outputs the corresponding public key Y_{ID}.
- Encrypt $(m, (ID_1, Y_{ID1}), \cdots, (ID_t, Y_{IDt}))$. The broadcaster runs this algorithm by inputting a plaintext m, public parameters PP, a set $S = (ID_1, Y_{ID1}), \cdots, (ID_t, Y_{IDt})$ of receivers' identities/public keys, and outputs a ciphertext $C = Encrypt(m, params, S)$.

- Decrypt (C): The algorithm is run by the receiver. It takes as inputs a ciphertext C, public parameters PP and the private key SK_{ID} of the receiver, returns a recovered message m or a symbol $\perp$ that indicates decryption error.
- Equality-test (sk_{TTP}, CT, CT'): It is a deterministic algorithm, run by a de-duplicator which is an honest-but-curious entity, it takes public parameter PP, the de-duplicator's secret key sk_{TTP} and two ciphertexts CT and CT' as inputs, and returns 1 if CT and CT' are from the identical plaintext, otherwise, returns 0.

4.2. Security Models

For a secure public key encryption scheme, it should ensure the confidentiality of the encrypted message, this property is referred to as ciphertext-indistinguishability which can be defined in two security models of chosen-plaintext-attack (CPA) and chosen-ciphertext-attack (CCA) [53]. However, for IND-CPA and IND-CCA, indistinguishability does not hold in a secure de-duplication public key encryption in that it is easily breached by an IND-CPA adversary or an IND-CCA adversary in the game [53]. In the **Challenge phase** of the IND-CPA/CCA security game, the adversary is allowed to select two plaintexts mt_0 and mt_1, and then a challenge C^* for a plaintext mt_b with $b \in \{0,1\}$ is returned. By invoking the **Equality-test** algorithm, the adversary is able to output the corresponding b by computing a ciphertext $\hat{C}$ for plaintext mt_b and checking whether $\hat{C}$ matches the challenge ciphertext C^*. The reason to produce such problem is that, given two ciphertexts, any one can run an **Equality-test** algorithm to check their matching-ability.

To provide IND-CCA security in the public key encryption with de-duplication, a trusted-third party (TTP) is introduced to execute an **Equality-test** algorithm by inputting its private key. Meanwhile, the adversary is not allowed to have access to TTP in the security game. Thus, the **Equality-test** query is not involved in the following security games. In the context of the rest of this paper, we let the de-duplicator act as the TTP.

Inspired by security models of certificateless encryption (CLE) and anonymous BE, the security model of our PMCBED schemes defines two security notations "confidentiality" and " anonymity of the receivers' identities". For confidentiality, it indicates that an adversary is not capable of obtaining any information of the encrypted message from ciphertext. For anonymity of the receivers' identities, it indicates that an adversary is not capable of obtaining any identity information of the other receivers from ciphertext.

In the following, we first define the IND-CCA security game for PMCBED . Let Adv_I, Adv_{II} be Type I and Type II probabilistic polynomial time (PPT) adversaries, respectively. In the following, Adv_I/Adv_{II} will make an interactive game with the challenger C.

Definition 1. *A PMCBED scheme is defined to be secure against adaptive-chosen-ciphertext attack ("IND-CCA security") if there does not exist a Type I/II of adversaries having a non-ignorable superiority in the following game:*

- *Setup: Let λ be a security parameter, C be a Challenger. C invokes a $Setup\ (1^{\lambda})$ algorithm to return public parameters PP and master secret key msk; afterwards, C transmits PP to Adv. If Adv is the Type II adversary Adv_{II}, then msk is also sent to Adv. Otherwise, msk is secretly kept by the Challenger and then sends system public parameters PP to adversary Adv who also receives the master secret key msk if it is of Type II. Otherwise, the master secret key msk is kept secret.*
- *Phase 1: In this phase, Adv can adaptively make a series of queries:*
  - *Public key query oracle: Upon receiving public key query of the receiver ID, if it is the first query of the receiver, then C invokes **Set public-key** algorithm to produce public key PK_{ID} and return PK_{ID} to Adv. Otherwise, it returns the matching public key in the list.*
  - *Extract partial-private key oracle: On receiving a partial private key query of the receiver ID, C inputs msk to invoke the **Extract partial-private key** algorithm and return d_{ID} if Adv is the Type I Adv_I; otherwise, the oracle is not required if A is of Type II.*

- *Extract secret-key oracle: Upon receiving the secret key query of the receiver ID from the adversary Adv, C invokes the* ***Set secret-value*** *algorithm to produce secret value x_{ID} and return it to Adv.*
- *Decrypt oracle: On receiving the decrypting query of (CT, ID) from Adv, C invokes the* ***Set secret-value*** *algorithm and* ***Extract partial-private key*** *algorithm to obtain private key SK_{ID} of the receiver ID; then, it runs a Decryption(CT, SK_{ID}) algorithm to recover the corresponding plaintext.*

Note that when Adv is the Type I Adv_I, it also needs to query ***Public-key-replace oracle*** *in which the receiver's public key Y_{ID} is replaced with a new public key Y'_{ID} when inputting a receiver's identity ID and its corresponding public key Y_{ID}.*

- *Challenge: The adversary Adv submits two distinct equivalent-length messages m_0 and m_1 as well as a set of the receivers' identities/public-keys $S^* = (ID_1/Y_1, \cdots, ID_k/Y_k)$. It is required that Adv cannot query Extract partial-private-key oracle with the identity $ID_i \in S^*$. The challenger C randomly samples a bit $b \in \{0,1\}$ to compute the challenge ciphertext $C^* = Encrypt(m_b, PP, S^*)$ and returns it to adversary A.*
- *Phase 2: Adversary Adv can continue to adaptively issue a new sequence of queries as in Phase 1. In addition, $(ID^*/Y^*, C^*)$ is not permitted to issue Decryption query, where $ID^*/Y^* \in S^*$.*

  Meanwhile, in a Type I attack, Adv is not allowed to issue Extract partial-private-key query and Public-key-replace query on identity ID^*, where $ID^* \in S^*$.
- *Guess: At last, a guess bit $b' \in \{0,1\}$ is returned by the adversary Adv. Adv wins this game if $b' = b$.*

Definition 2. *A PMCBED scheme is defined as ANO-CCA security if there does not exist a Type I or II of adversary Adv which has a non-ignorable superiority in the following games:*

- ***Setup*** *and* ***Phase 1****: In the two phases, they are the same as those in the above IND-CCA Game.*
- *Challenge: In this phase, Adv produces two challenge sets $\hat{S}_0$ and $\hat{S}_1$, where $|\hat{S}_1| = |\hat{S}_0|$. In addition, it then submits a message m^* and $(\hat{S}_0, \hat{S}_1)$ to C. In addition, the constraint conditions are as follows: (1) Adv is not permitted to issue Extract partial-private-key queries on ID^* when Adv is the Type I adversary Adv_I,(2) a Adv is not permitted to issue Extract secret-key queries on ID^* when Adv is the Type II adversary Adv_{II}, where $ID^* \in \hat{S}_1 \oplus \hat{S}_0$ and $\hat{S}_1 \oplus \hat{S}_0 = \hat{S}_1 \cup \hat{S}_0 - \hat{S}_0 \cap \hat{S}_1$. Then, C uniformly samples a bit $\alpha \in \{0,1\}$ to calculate the ciphertext $C^* = Encrypt(PP, \hat{S}_\alpha, m^*)$ and returns it to Adv.*
- *Phase 2. In this phase, Adv adaptively issues a new series of queries as in Phase 1 with the following constraint conditions :(1) Adv is not permitted to issue Extract partial-private-key queries on ID^*, (2) Public-key-replace queries on ID^* are not allowed when Adv is the Type I adversary Adv_I, (3) Extract secret-key queries on ID^* are not allowed when Adv is the Type II adversary Adv_{II}, and (4) Adv is not allowed to issue Decryption Query on $(ID^*/Y^*; C^*)$, where $ID^* \in \hat{S}_1 \oplus \hat{S}_0$.*
- *Guess: At last, a guess bit $\alpha' \in \{0,1\}$ is outputted by Adv. Adv wins this game if $\alpha = \alpha'$.*

5. Our Scheme

Setup: Let λ be a security parameter, **Setup** (λ) algorithm takes as input λ, and outputs a bilinear map $e : \mathbb{G}_1 \times \mathbb{G}_2 \to \mathbb{G}_T$, where $\mathbb{G}_1$ and $\mathbb{G}_2$ are two groups satisfying $\mathbb{G}_1 = \ < P_1 >$ and $\mathbb{G}_2 = \ < P_2 >$. In addition, they has the same order p. **Note that** $P_1 = \varphi(P_2)$ and $\varphi : \mathbb{G}_2 \to \mathbb{G}_1$ is an isomorphism. Let $H : \{0,1\}^* \to \mathbb{G}_1$,$H_1 : \mathbb{G}_T \leftarrow \mathbb{G}_1$, $H_3 : \mathbb{G}_T \leftarrow Z_p$ be three cryptographical hash function. $H_2()$ and $f()$ are two one-way functions. H_0 is a random generator of group $\mathbb{G}_2$. For the KGC, it picks a number $s \in Z_p$ at random to calculate its public key $PK_{pub} = sP_2$. Let $TPK = x_T P_1$ denote the public key of de-duplicator, $x_T \in Z_p$ be its private key. $(E(\cdot), D(\cdot))$ denotes the encryption/decryption algorithm of AES. Finally, the public parameters are $Param = (P_1, P_2, \mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, p, \varphi, TPK, PK_{pub}, e, H(), H_1, H_2, H_3, f, H_0, (E, D))$. $msk = s$ acts as a master secret key and is kept secretly.

Extract partial-private key: First, in all, a receiver submits its identity ID to the KGC; then, the KGC utilizes its master secret key msk to produce partial-private key d_{ID} of the receiver, where $d_{ID} = sH(ID)$.

Set secret value: For a receiver with identifier ID_i, it uniformly samples a number $xk_i \in Z_p$ and returns xk_i to act as its secret value.

Set private-key: For a receiver with identifier ID_i, let d_{ID_i} be its partial-private key, and xk_i be its secret value. Its private key SK_{ID_i} is set to be $SK_{ID_i} = (xk_i, d_{ID_i})$.

Set public-key: In this algorithm, a receiver with identifier ID_i takes an input secret value xk_i, and outputs its public key $Y_i = xk_i P_1$.

Encrypt: Given a transmitted message M and a group of the receivers with public keys and identifiers $\{ID_i, Y_i\}_{i=1,2,\cdots,n}$, a broadcaster computes as follows:

1. For $i = 1$ to n, it calculates $x_i = H_2(ID_i)$, and then it produces the polynomial

$$C_i(x) = \prod_{j=1, j\neq i}^{n} \frac{x - x_j}{x_i - x_j} = \sum_{j=0}^{n-1} b_{i,j} x^j \mod p.$$

   Obviously, we find $C_i(x_i) = 1$ and $C_i(x_j) = 0$ for $i \neq j$.
2. It randomly chooses $k \in Z_p$ to compute $C_1 = kP_2$.
3. Then, it selects $Q \in \mathbb{G}_T$ and $\tau \in Z_p$ to compute $K = H_3(Q)$ and $C_3 = E(K, M||\tau)$.
4. Next, choose a random number $r_1 \in Z_p$, and then for $j \in \{1, 2, \cdots, n\}$, it calculates

$$R_j = H_1(e(H(ID_j), k \cdot PK_{pub})) + r_1 Y_j.$$

5. In addition, it computes $C_2 = e(P_1, r_1 P_2)^k \cdot Q$.
6. In addition, for each $t \in \{1, 2, \cdots, n\}$, it computes

$$Q_t = \sum_{j=1}^{n} b_{j,t-1} R_j.$$

7. Compute $C_0 = (f(M) + f(\tau)) \cdot TPK$ and $C_{-1} = e(P_1, P_2)^{f(\tau)}$.
8. Finally, the resultant ciphertext is as below:

$$CT = (C_{-1}, C_0, C_1, C_2, C_3, Q_1, \cdots, Q_n).$$

Decrypt: For a given broadcast-ciphertext $CT = (C_{-1}, C_0, C_1, C_2, C_3, Q_1, \cdots, Q_n)$, an authorized receiver with identity ID_i inputs public parameters $Param$, system public key PK_{pub} and its private key SK_{ID_i} to decrypt broadcast–ciphertext CT by the following steps:

1. First, it computes $x_i = H(ID_i)$.
2. Then, it calculates

$$\hat{R}_i = Q_1 + \sum_{j=2}^{n} (x_i^{j-1} Q_j).$$

3. It computes $W = xk_i^{-1} \cdot (\hat{R}_i - H_1(e(sH(ID_i), C_1)))$;
4. In addition, it obtains the decryption key $K' = H_1(C_2 / \ e(W, C_2\ ))$.
5. Finally, it obtains the plaintext $M = D(K', C_3)$ and checks $C_0 \overset{?}{=} (f(M) + f(\tau))TPK$. If it holds, output TRUE.

Equality-test: Given two ciphertexts CT and CT', where $CT' = (C'_{-1}, C'_0, C'_1, C'_2, C'_3, Q'_1, \cdots, Q'_n)$ and $CT = (C_{-1}, C_0, C_1, C_2, C_3, Q_1, \cdots, Q_n)$, the de-duplicator makes use of its private key x_T to execute as follows:

$$e(C_0 - C'_0, P_2)^{x_T^{-1}} \overset{?}{=} C_{-1} / C'_{-1}. \quad (1)$$

Finally, it returns 1 if the above-mentioned Equation (1) holds; otherwise, output $\perp$.

Discussion

For the above construction, we can know that, if the receiver's identity ID is involved in the set of the designated receivers, then this receiver can decrypt the corresponding ciphertext CT since, when this receiver's identifier satisfies $ID_i \in S$, where $S = \{ID_1/PK_1, \cdots, ID_n/PK_n\}$, let $x_i = H(ID_{t_i})$, we have $C_j(x_i) = 0$ for $j \neq i$ and

$$\begin{aligned}\hat{R}_i &= Q_1 + x_iQ_2 + x_i^2Q_3 + \cdots + x_i^{n-1}Q_n \\ &= (b_{1,0}R_1 + b_{2,0}R_2 + \cdots + b_{n,0}R_n) + \\ &\quad x_i(b_{1,1}R_1 + b_{2,1}R_2 + \cdots + b_{n,1}R_n) + \cdots + \\ &\quad x_i^n(b_{1,n-1}R_1 + b_{2,n-1}R_2 + \cdots + b_{n,n-1}Q_n) \\ &= (b_{1,0} + b_{1,1}x_i + \cdots + b_{1,n-1}x_i^{n-1})R_1 + \\ &\quad (b_{2,0} + b_{2,1}x_i + \cdots + b_{2,n-1}x_i^{n-1})R_2 + \\ &\quad \cdots + (b_{n,0} + b_{n,1}x_i + \cdots + b_{n,n-1}x_i^{n-1})R_n \\ &\quad = C_i(x_i)R_i = R_i.\end{aligned}$$

Thus, the receiver with identifier ID_i is capable of obtaining r_1P_1 by utilizing its partial-private key d_{ID_i}, namely,

$$r_1P_1 = \hat{R}_i - H_1(e(d_{ID_i}, C_1)).$$

It means that the receiver with identifier ID_i is able to decrypt the message by the key $K = H_1(C_2/e(r_1P_1, C_1))$.

6. Security Analysis

In the following theorems, we will show that our aforementioned construction can achieve two security properties: anonymity of the receiver's identity and confidentiality.

Theorem 1. *Let H, H_1 and H_2 denote random oracles. If the BDH-2 problem and the DDH problem in $(\mathbb{G}_1, \mathbb{G}_2)$ are hard, then our proposed construction can be proven to be secure against the IND-PMCBED-CCA attack of the Type I adversary.*

Proof. Suppose there exists a Type I of adversary A_I in an IND-PMCBED-CCA game. If it can break our construction in a non-negligible probability ϵ, then we are capable of building an algorithm B which solves the BDH-2 problem and the DDH problem in $(\mathbb{G}_1, \mathbb{G}_2)$. □

Let (P_2, aP_2, bP_2, cP_1) be a random instance of the BDH-2 problem, where a, b and c are unknown random numbers from Z_p; the target is to compute $e(P_1, P_2)^{abc}$. In addition, let $(P_1, \beta_1P_1, \beta_2P_1, V)$ be a random instance of the DDH problem, its target is to determine $V \overset{?}{=} \beta_1\beta_2P_1$. Therefore, B simulates the following security game with the adversary A_I.

Setup. Let $PP = \{P_1, P_2, \mathbb{G}_1, \mathbb{G}_2, e, p, H, H_1, H_2, H_3, (E, D)\}$ be system parameters; they are built by B. In addition, B sets $PK = aP_1 = \varphi(aP_2)$ and $TPK = \beta_1P_1$. Then, B sends public parameters PP to the adversary A_I. In the following proof,H_2 acts as a one-way function. H, H_1 and H_3 are random oracles.

Phase 1. In this phase, $\mathcal{A}_I$ is capable of adaptively issuing a series of queries.

- H-Hash Query: When receiving the H-hash query on ID_i from $\mathcal{A}_I$, B answers as below. If a record ID_i have appeared in a tuple (ID_i, Q_i, η_i, q_i) in the H-list which is originally empty, it sends back Q_i; otherwise, it generates $\eta_i \in \{0,1\}$, and randomly chooses $q_i \in Z_p$. If $\eta_i = 0$, it sets $Q_i = q_iP_1$, else it sets $Q_i = q_ibP_1 = q_i \cdot \varphi(bP_2)$ and adds (ID_i, Q_i, η_i, q_i) in the H-list. It returns Q_i to A_I

- H_1-Query: On input, an identity X_i, if (X_i, T_i) exists in the H_1-list, then it returns T_i to A_I; otherwise, it picks $T_i \in \mathbb{G}_1$ to return A_I and adds (ID_i, T_i) into the H_1-list. Note that H_1-list is originally empty.
- H_3-Query: On input, D_i, if (D_i, k_i) is in the H_3-list which being originally empty, it sends back k_i to A_I; otherwise, it picks $k_i \in Z_p$ to return A_I and adds (D_i, k_i) into the H_3-list.
- Public-key query: When A_I makes a public key query with ID_i, if the 3-tuple (ID_i, Y_i, xk_i) appears in the PK-list which is initially empty. Y_i is returned to A_I; otherwise, B picks $xk_i \in Z_p$ to set $Y_i = xk_i P_1$, and adds (ID_i, Y_i, xk_i) in the PK-list. Finally, it returns Y_i to A_I.
- Extract partial-private key Query: Upon receiving a Partial-private key query of the identity ID_i, if the record (ID_i, Q_i, η_i, q_i) had appeared in the H-list and $\eta_i = 0$, then B computes $d_{ID_i} = q_i \cdot aP_1 = q_i \cdot \varphi(aP_2)$. Otherwise, abort it and output $\perp$.
- Extract secret-value Query: When A_i issues a query on an identity ID_i, if 3-tuple (ID_i, Y_i, xk_i) exists on the PK-list, then xk_i is returned to A_I, otherwise, B randomly selects $xk_i \in Z_p$ to compute $Y_i = xk_i P_1$ and adds (ID_i, Y_i, xk_i) in the PK-list.
- Public-key-replace Query: When A_I makes a public key replace query with (ID_i, Y_i'), the corresponding tuple (ID_i, Y_i, xk_i) is replaced into a new tuple $(ID_i, Y_i', \perp)$ in the PK-list.
- Decryption Queries: On input, a ciphertext CT and an identity ID_i, where $CT = (C_1, C_2, C_3, Q_1', \cdots, Q_n')$, B first issues a H-query with ID_i to obtain the tuple (ID_i, Q_i, η_i, q_i), if $\eta_i = 0$, it sets $d_{ID_i} = q_i \cdot P_1$ and make a Extract-secret-value query with ID_i, if $xk_i \neq \perp$ is returned, B can make use of (d_{ID_i}, xk_i) to decrypt CT and respond the Decryption Query. Otherwise, B does the following steps:

  1. For $j = 1$ to q_{H_3} {
     it retrieves k_i from H_3-list and decrypts CT to recover $M||\tau = D(k_i, CT)$ with k_i to parse it into M and τ which can recover τTPK. (**Note that** we assume that the H_3-query had been made before the adversary issues the decryption-query with CT).
     if $C_0 = f(M) \cdot TPK + f(\tau)TPK$
         break;
     }
  2. If $j \leq q_{H_3}$, B sends back M to A_I. Otherwise, it aborts it.

Challenge. In this phase [13], A_I submits two equivalent-size plaintext M_0 and M_1, as well as a challenge set of identities/public-keys $S^* = (ID_1/Y_1, ID_2/Y_2, \cdots, ID_l/\ PK_l)$ with the restriction conditions which A_I have not issued partial-private-key Oracle with $ID_i \in S^*$ in phase 1 and each $\eta_i = 1$ in the tuple (ID_i, Q_i, η_i, q_i) of H_1-list, where Y_i is a public key which corresponds to the identity ID_i.

Then, B computes as follows:

1. iIt sets $C_1^* = cP_2$.
2. For $j = 1$ to l, it computes $x_j^* = H_2(ID_j)$.
3. Next, for $j = 1$ to l, it constructs the polynomial

$$f_j(x) = \prod_{i=1, j\neq i}^{l} \frac{x - x_i^*}{x_j^* - x_i^*} = \sum_{i=0}^{l} a_{ji} x^i.$$

4. B randomly chooses $r_1 \in Z_p$.
5. For $j = 1$ to l, it randomly chooses $T_i \in \mathbb{G}_1$ to compute $R_j = T_j + r_1 Y_j$.
6. For $j = 1, 2, \cdots, l$, B computes $Q_j = \sum_{i=0}^{l} a_{i,j-1} R_i$
7. B randomly chooses $Q \in \mathbb{G}_T$ and $\tau \in \{0,1\}^t$ to compute $C_2^* = e(P_1, C_1^*)^{r_1} \cdot Q$ and $C_3^* = E(K, M_\beta||\tau)$, where $K = H_3(Q)$,$\beta \in \{0,1\}$.
8. It computes $C_0^* = f(M_\beta)TPK + V$ and $C_{-1}^* = e(\alpha_2 P_1, P_2)$. Note that in fact $(TPK = \alpha_1 P_1, P_1, V, \alpha_2 P_1)$ is also an instance of DDH problem when $(P_1, \alpha_1 P_1, \alpha_2 P, V)$ is an instance of DDH problem, since $P_1 = \alpha_1^{-1} \cdot TPK, V = \alpha_2 \cdot TPK$ and $\alpha_2 P_1 = \alpha_1^{-1}\alpha_2 \cdot TPK$.

9. The resultant ciphertext $CT^* = (C^*_{-1}, C^*_0, C^*_1, C^*_2, C^*_3, Q_1, \cdots, Q_l)$ is returned to A_I.

Phase 2. A_I can adaptively make a new series of queries as in Phase 1 with the constraints:

1. CT^* can not be made into Decryption queries.
2. All $ID_i \in S^*$ is not allowed to issue Extract partial-private-key queries.

Guess. Eventually, A_I outputs its guess $\beta' \in \{0, 1\}$.

When $V = \beta_1\beta_2P_1$, the challenge ciphertext CT^* is a valid one. For the perspective of A_I, the challenger's simulation is indistinguishable from the real game. When V is a random element of $\mathbb{G}_1$, the challenge ciphertext has the same distribution as the real ciphertext. Furthermore, we assume that A_I must have previously issued H_1 query with $X_i = e(H(ID_i), PK)^c$ Because $C^*_1 = cP_1$,$H(ID_i) = q_ibP_1$ and $PK = aP_1$, it means that B can compute $e(P_1, P_2)^{abc} = (X_i)^{q_i^{-1}}$.

Therefore, it is impossible to have an IND-PMCBED-CCA adversary A_I which breaks our PMCBED scheme. □

Theorem 2. *Under the DDH problem in $\mathbb{G}_1$, our proposed PMCBED scheme is provably secure against the IND-PMCBED-CCA attack of Type II adversary A_{II}.*

Proof. Assume that there is a Type II of adversary A_{II} in the IND-PMCBED-CCA game. If it breaks our construction, then we are capable of constructing an algorithm B to solve the DDH problem. Let (P_1, aP_1, bP_1, Z) be an instance of DDH problem in group $\mathbb{G}_1$, where $a, b \in Z_p$ are unknown, its goal is to determine $Z = abP_1$. □

Setup. Algorithm B randomly chooses $\alpha \in Z_p$ to compute $PK = \alpha P_1$ and let $TPK = aP_1$. Let PP be public parameters, where $PP = (P_1, PK, TPK, e, \mathbb{G}_1, \mathbb{G}_2, P_2, H, H_1, H_2, H_3, E, D, f)$. Then, it delivers PP and α to the adversary A_{II}. H, H_1, H_3 are three random oracles which are controlled by B.
Phase 1. A_{II} can adaptively issue a series of queries.
***H*-Hash Queries.** Upon receiving a receiver's identifier ID_j, B first checks that (ID_j, Q_j) has appeared in the H-list which is initially empty; if it is, then Q_j is returned. Otherwise, B picks $q_j \in Z_p$ at random to calculate $Q_j = q_jP_1$ and adds (ID_j, Q_j, q_j) in the H_1-list. Finally, Q_j is returned.
***H*₁-Hash Queries. It is the same as that of Theorem 1.
***H*₃-Hash Queries. It is the same as that of Theorem 1.
Public-Key Queries. Upon receiving an identity ID_i, if the 3-tuple (ID_i, Y_i, xk_i) has existed in the PK-list that was originally empty, then Y_i is returned. Otherwise, it produces $\eta_i \in \{0, 1\}$ and randomly chooses $a_i \in Z_p$. If $\eta_i = 0$, it sets $Y_i = a_iP_1$, else it sets $Q_i = a_ibP_1$ and adds (ID_i, Y_i, η_i, a_i) in the PK-list. It returns Y_i to A_{II}.
Decryption Query. Upon receiving (CT, ID_i), if ID_i had existed in the PK-list and the corresponding $\eta_i = 0$ holds, then B decrypts the ciphertext CT by $(\alpha \cdot H(ID_i), a_i)$ and returns the decrypted message M to the adversary A_{II}. Otherwise, B does the following steps:

```
1.  For $j = 1$ to $q_{H_3}$ {
    it retrieves $k_i$ from $H_3$-list and decrypts $CT$ to recover $M = D(k_i, CT)$ with $k_i$;
    if $C_0 = f(M) \cdot TPK + f(\tau)TPK$
        break;
    }
2.  If $j \leq q_{H_3}$, $B$ sends back $M$ to $A_{II}$. If not, it aborts it.
```

Challenge Phase. Let $S^* = (ID_1/Y_1, ID_2/Y_2, \cdots, ID_l\ /\ Y_l)$. In this phase, the adversary A_{II} outputs two equivalent length messages M_0 and M_1, and a set of identites/public-keys S^* with the restriction conditions with each η_i in the tuple (ID_i, Y_i, η_i, a_i), where $ID_i \in S^*$ satisfies $\eta_i = 1$.

Then, B is computed as below:

1. It uniformly samples $k \in Z_p$ to compute $C_1^* = kP_2$ and $C_0^* = f(M_\beta)TPK + Z$ as well as $C_{-1}^* = e(bP_1, P_2)$. **Note that** we have the relation $(D_0 = aP_1, D_1 = P_1 = D_0^{a^{-1}}, D_2 = Z = D_0^b, D_4 = bP_1 = D_0^{a^{-1}b})$ which is the instance of the CDH problem if $Z = abP_1$.
2. For $j = 1$ to l, it calculates $x_i^* = H_2(ID_i)$;
3. Then, for $j = 1$ to l, it builds the polynomial

$$f_j(x) = \prod_{j \neq i}^{l} \frac{x - x_i^*}{x_j^* - x_i^*} = \sum_{i=0}^{l} a_{ji} x^i.$$

4. For $j = 1$ to l, B computes

$$R_j = H_1(e(\alpha \cdot H(ID_i), C_1^*)) + a_i \cdot Z.$$

Note that r_1 in the original encryption is set as $r_1 = a$ but is unknown.

5. For $i \in \{1, 2 \cdots, l\}$, it calculates

$$Q_i = \sum_{j=1}^{l} a_{j,i-1} R_j.$$

6. It randomly selects $Q \in \mathbb{G}_T$ to compute $K = H_3(Q)$ and $C_3^* = E(K, M_\beta || x_Z)$, x_Z denotes the x-coordination of point Z.
7. It computes $C_2^* = e(aP_1, P_2)^k \cdot Q$.
8. The ciphertext is $CT^* = (C_{-1}^*, C_0^*, C_1^*, C_2^*, C_3^*, Q_1, \cdots, Q_l)$.

Phase 2. A_{II} may issue a new series of queries which is the same as what it did in Phase 1 with the restriction that CT^* is not made in the Decryption query.

Guess. Finally, A_{II} gives its guess β'. If $\beta = \beta'$, A_{II} wins this game with non-ignorable advantage ε. When $Z = abP_1$, the ciphertext $CT^* = (C_0^*, C_1^*, C_2^*, C_3^*, Q_1, \cdots, Q_l)$ is a valid one since

$$\begin{aligned}
R_j &= H_1(e(\alpha \cdot H(ID_i), C_1^*)) + a_i \cdot Z \\
&= H_1(e(\alpha \cdot H(ID_i), C_1^*)) + a(a_i \cdot b)P_1 \\
&= H_1(e(\alpha \cdot H(ID_i), C_1^*)) + a \cdot Y_i, \\
C_2^* &= e(aP_1, P_2)^k \cdot Q = e(P_1, aP_2)^k \cdot Q, \\
C_0^* &= f(M_\beta)TPK + abP_1 = f(M_\beta)TPK + bTPK, \\
C_{-1}^* &= e(P_1, P_2)^b = e(P_1, P_2)^\tau.
\end{aligned}$$

This means that $r_1 = a$ and $\tau = b$ in the encryption. Thus, if A_{II} breaks our scheme , then B is able to solve the DDH problem. □

Theorem 3. *Let hash functions H, H_1, H_3 be random oracles. If the decisional bilinear Diffie–Hellman problem (DBDH) is hard, then our construction is able to be proved to be secure against the Type I adversary in the ANON-ID-CCA attack game.*

Proof. Let A_I be an ANON-ID-CCA adversary. If it breaks the proposed AMCLE scheme in a non-ignorable advantage, then we are capable of building a new algorithm B to solve the DBDH problem. □

Setup. Firstly, let $PK = aP_1$ act as a master public key, and B builds the following parameters $PP = (\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, H, H_1, H_2, H_3, e, p, f, E, D)$, and delivers PP to the adversary A_I. H, H_1, H_3 are three hash functions that act as random oracles.

Phase 1. A_I is capable of adaptively making a sequence of security queries which are the same as those in Theorem 1.

Challenge. After terminating Phase 1, A_I submits a challenge message M and two disparate sets of identities/public-keys $S_0^* = (ID_0^*/Y_0^*, ID_2/Y_2, \cdots, ID_l \ / \ Y_l)$ and $S_1^* = (ID_1^*/Y_1^*, ID_2/\ Y_2, \cdots, ID_l \ / \ Y_l)$ with the constraint in which A_I can not issue **Extract Partial-private-key queries** with ID_i for $ID_i \in \{S_0^*, S_1^*\}$. B randomly selects $\beta \in \{0,1\}$ to compute as follows:

1. It sets $C_1^* = cP_2$.
2. B retrieves $(ID_\beta^*, Q_\beta^*, \eta_\beta^*, q_\beta^*)$ by issuing a H-query on ID_β^*, if $\eta_\beta^* = 0$ holds, then it aborts it and outputs $\perp$; if $\eta_\beta^* = 1$, then let $Q_\beta^* = q_\beta^* \cdot bP_1$ and $X_\beta^* = Z^{q_\beta^*}$. Next, it issues H_2-queries with X_β^* to obtain T_β^*.
3. Compute $x_\beta^* = H_2(ID_\beta^*)$, and for $j = 2$ to l, it computes $x_i^* = H_2(ID_i)$.
4. Next, for $j \in \{2, 3, \cdots, l\}$, it constructs the polynomial

$$f_j(x) = \prod_{i \neq j, i = 1}^{l} \frac{1}{x_j^* - x_i^*} \cdot (x - x_i^*) = \sum_{i = 0}^{l} a_{ji} x^i.$$

5. B randomly chooses $r_1 \in Z_p$ and for $j \in \{2, 3, \cdots, l\}$, it randomly chooses $T_i \in \mathbb{G}_1$ to compute $R_j = T_j + r_1 Y_j$; and then it computes $R_\beta = T_\beta + r_1 Y_\beta^*$.
6. For $j \in \{\beta, 2, 3, \cdots, l\}$, B computes $Q_j = \sum_{i = 0}^{l} a_{i,j-1} R_i$.
7. B randomly chooses $Q \in \mathbb{G}_T$ and $\tau \in Z_p$ to compute $C_2^* = e(P_1, C_1^*)^{r_1}$ and $C_3^* = E(K, M_\beta || x_\tau)$, where $K = H_3(Q)$ and x_τ is the x-coordination of point $\tau \cdot TPK$.
8. B computes $C_0^* = (f(M_\beta) + \tau)TPK$ and $C_{-1}^* = e(P_1, P_2)^\tau$.
9. The ciphertext is $CT^* = (C_{-1}^*, C_0^*, C_1^*, C_2^*, C_3^*, Q_1, \cdots, Q_l)$ to the adversary A_I.

Phase 2. A_I sequentially issues a new series of queries with the following restrictions:

1. A_I can not issue **Extract Partial-Private-Key Queries** with ID, where $ID \in \{ID_0^*, ID_1^*\}$.
2. A_I can not issue **Public-Key Replace** with ID, where $ID \in \{ID_0^*, ID_1^*\}$.
3. A_I can not issue **Decryption Queries** with (ID, CT^*), where $ID \in \{ID_0^*, ID_1^*\}$.

Guess. Finally, A_I outputs its guess β'. B outputs 1 when $\beta = \beta'$, it means that $Z = e(P_1, P_2)^{abc}$; if $\beta \neq \beta'$, outputs 0, it means $Z \neq e(P_1, P_2)^{abc}$.
Analysis: In the above game, the simulation is indistinguishable from the scheme. If $Z = e(P_1, P_2)^{abc}$, then we let $k^* = c$. All this time, CT^* has the same distribution as the ciphertext in the real game; If Z is a random element in $\mathbb{G}_T$, then the ciphtertext has the uniform distribution in the ciphertext space since $C_3^* = E(K, x_\tau || M_\beta)$, where $K = H_3(Q)$ is a random element. Thus, in the adversary A_I's view, M_β is independent, and it cannot provide any information to A_I. □

Theorem 4. *Let hash functions H, H_1 and H_3 be a random oracle. If the DDH assumption in groups $(\mathbb{G}_1, \mathbb{G}_2)$ is difficult, then our construction is proven to be secure against the Type II of adversary A_{II} in the ANON-ID-CCA attack game.*

Proof. Let A_{II} be an adversary. If it breaks our construction, then we are capable of constructing a novel algorithm B which solves the DDH problem. Let (P_1, aP_1, bP_1, Z) be a random instance of DDH problem in groups $(\mathbb{G}_1, \mathbb{G}_2)$, where $a, b \in Z_p$ are unknown, its goal is to determine $Z = abP_1$. □

Setup. Algorithm B randomly chooses $\alpha \in Z_p$ to set $PK = \alpha P_1$. Let $PP = (P_1, P_2, e, p, PK, f, \mathbb{G}_1, \mathbb{G}_2, H, H_1, H_2, H_3, (E, D))$ denote public parameters that are built by B. Then, it delivers PP and α to the adversary A_{II}. Here H, H_1, H_3 are three random oracles that are controlled by B.
Phase 1. A_{II} is capable of issuing a series of the same queries as those of Theorem 2.
Challenge. A_{II} outputs a challenge plaintext M^* and two different sets S_0^* and S_1^* of identities/public-keys, where $S_0^* = (ID_0^*/Y_0^*, ID_2/Y_2, \cdots, ID_l \ / \ Y_l)$ and $S_1^* = (ID_1^*/\ Y_1^*, ID_2/\ Y_2, \cdots, ID_l \ / \ Y_l)$. In addition, the following constraints need to be satisfied: A_{II} cannot issue **Extract partial-private-key queries** on ID_i in **Phase 1**, where $ID_i \in \{ID_0^*, ID_1^*\}$. In addition, then B randomly selects $\beta \in \{0,1\}$ to compute as below:

1. First, it makes a **Public-key Query** on ID^*_β to obtain $(ID^*_\beta, Y^*_\beta, \eta^*_\beta, a^*_\beta)$. If $\eta^*_\beta = 0$, output $\perp$ and abort it. If $\eta^*_\beta = 1$, it means that $Y^*_\beta = a^*_\beta \cdot bP_1$.
2. For $j \in \{\eta, 2, 3, \cdots, l\}$, it calculates $x^*_i = H_2(ID_i)$;
3. Then, for $j \in \{\beta, 2, 3, \cdots, l\}$, it builds the polynomial

$$f_j(x) = \prod_{i \neq i}^{l} \frac{x - x^*_i}{x^*_j - x^*_i} = \sum_{i=0}^{l} a_{ji} x^i.$$

4. For $j \in \{\beta, 2, 3, \cdots, l\}$, B issues **Public-key Queries** with ID_j to obtain (ID_j, Y_j, η_j, a_j). If $\eta^*_\beta = 0$, B computes

$$R_j = H_1(e(\alpha \cdot H(ID_j), C^*_1)) + a_j \cdot aP_1.$$

If $\eta^*_\beta = 1$, it computes $R_j = H_1(e(\alpha \cdot H(ID_j), C^*_1)) + a_j \cdot Z$.
5. For $j \in \{\beta, 2, 3, \cdots, l\}$, it computes $Q_i = \sum_{j=1}^{l} a_{j,i-1} R_j$.
6. It randomly selects $Q \in \mathbb{G}_T$ and $\tau \in Z_p$ to compute $K = H_3(Q)$ and $C^*_3 = E(K, x_\tau || M_\beta)$.
7. It randomly chooses $k \in Z_p$ to compute $C^*_1 = kP_2$ and $C^*_0 = f(M_\beta) + \tau \cdot TPK$ as well as $C^*_{-1} = e(P_1, P_2)^\tau$.
8. It computes $C^*_2 = e(aP_1, P_2)^k \cdot Q$.
9. The resultant ciphertext is $CT^* = (C^*_{-1}, C^*_0, C^*_1, C^*_2, C^*_3, Q_1, \cdots, Q_l)$.

Phase 2. A_{II} can still adaptively issue the queries with the following constraints.

1. A_{II} is not capable of issuing **Public-key Query** with ID, where $ID \in \{ID^*_0, ID^*_1\}$.
2. A_{II} is not capable of issuing **Decryption Query** with (CT^*, ID), where $ID \in \{ID^*_0, ID^*_1\}$.

Guess. Finally, A_{II} returns its guess bit β'. B outputs 1 if $\beta = \beta'$; it means that $Z = abP_1$; otherwise, outputs 0 meaning $Z \neq abP_1$.
Analysis: In the above game, the simulation is indistinguishable from the scheme. When $Z = abP_1$, assume $r_1 = a$. The challenge ciphertext has the same distribution as that in the real game, in addition to when Z is a random element of $\mathbb{G}_1$, C^*_2 and C^*_3 in the ciphtertext has the form $C^*_2 = e(aP_1, P_2)^k \cdot Q$ and $C^*_3 = E(K, x_\tau || M_\beta)$, where $K = H_3(Q)$ and Q are uniform and random. Thus, from the adversary A_{II}'s view, M_β is independent; it provides no information to A_{II}. □

7. Performance Analysis

To evaluate the efficiency of the proposed scheme, we give the corresponding computational cost of the main algorithm by comparing with the Hung et al. scheme [37] and Islam et al. scheme [45]. For convenience, we define the following notations. Let T_p, T_m, T_e and T_h denote the time of executing a pairing operation, a scalar multiplication operation and an exponentiation operation as well as a map-to-point hash function, respectively. The computation cost of the main algorithms for the three schemes are shown in Table 1.

Table 1. Comparison of computation costs in the three schemes.

	Islam et al. Scheme [45]	**Hung et al. Scheme [37]**	**Our Scheme**
Computational cost of encryption for n receivers	$(2n+1)T_p + (n^2+n)T_m$	$nT_p + nT_e + (n+1)T_m + nT_h$	$(n+1)T_p + (n+2)T_m + 2T_e + nT_h$
Complexity of encryption	$O(n^2)$	$O(n)$	$O(n^2)$
Computational cost of decryption for each receiver	$T_m + nT_h$	$T_p + 1T_M$	$3T_p + (n+3)T_M + T_e$
Complexity of decryption	$O(n)$	$O(1)$	$O(n)$
De-duplication	No	No	Yes
Security	selective-CCA security	selective-CCA security	CCA-security

From Table 1, we find that our proposed scheme has more computational costs than the other two schemes. However, our proposed scheme has better security and functionality.

8. Conclusions

The users are increasingly concerned about anonymity. To protect the identity anonymity of the receiver, we construct a privacy-preserving Multi-receiver Certificateless Broadcast Encryption Scheme with De-duplication scheme in this work. It can not only simultaneously achieve confidentiality and the receiver's identity anonymity, but also achieve duplicate detection to determine whether two different ciphertexts are from the identical message. Thus, our proposal can efficiently reduce the cloud server's storage burden. It is very significant for cloud storage. Nevertheless, the ciphertext size is linear to the number of the receivers. A very important challenge will be how to construct a PMCBED scheme with constant-size ciphertext.

Author Contributions: All the authors contributed to the research and wrote the article. J.Z. proposed the idea, designed, security analysis and performed the evaluation. P.O. suggested directions for the detailed designs and evaluation, as well as coordinating the research.

Funding: This research was supported by the Open Fund of National Engineering Laboratory for Big Data Collaborative Security Technology, the Guangxi Key Laboratory of Cryptography and Information Security (No. GCIS 201710), Beijing Natural Science Foundation—Haidian Original Innovation Joint Fund Project Task Book (Key Research Topic) (Nos. L182039), and Beijing Municipal Natural Science Foundation (Nos. 4162020).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fiat, A.; Naor, M. Broadcast encryption. In Proceedings of the Annual International Cryptology Conference, Santa Barbara, CA, USA, 22–26 August 1993; pp. 480–491.
2. Fan, L.; Lei, X.; Yang, N.; Duong, T.Q.; Karagiannidis, G.K. Secure multiple amplify-and-forward relaying with cochannel interference. *IEEE J. Sel. Top. Signal Process.* **2016**, *10*, 1494–1505. [CrossRef]
3. Fan, L.; Lei, X.; Yang, N.; Duong, T.Q.; Karagiannidis, G.K. Secrecy cooperative networks with outdated relay selection over correlated fading channels. *IEEE Trans. Veh. Technol.* **2017**, *66*, 7599–7603. [CrossRef]
4. Shen, J.; Zhou, T.; Chen, X.; Li, J.; Susilo, W. Anonymous and traceable group data sharing in cloud computing. *IEEE Trans. Inf. Forensics Secur.* **2017**, *13*, 912–925. [CrossRef]
5. D'Orazio, C.J.; Choo, K.K.R.; Yang, L.T. Data exfiltration from Internet of Things devices: iOS devices as case studies. *IEEE Internet Things J.* **2016**, *4*, 524–535. [CrossRef]
6. Do, Q.; Martini, B.; Choo, K.K.R. Cyber-physical systems information gathering: A smart home case study. *Comput. Netw.* **2018**, *138*, 1–12. [CrossRef]
7. Zhang, J. Improvement of ID-based proxy re-signature scheme with pairing-free. *Wireless Netw.* **2019**. [CrossRef]
8. Bellare, M.; Boldyreva, A.; Desai, A.; Pointcheval, D. Key-privacy in public-key encryption. In *International Conference on the Theory and Application of Cryptology and Information Security*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 566–582.
9. Li, X.; Gu, D.; Ren, Y.; Ding, N.; Yuan, K. Efficient ciphertext-policy attribute based encryption with hidden policy. In Proceedings of the International Conference on Internet and Distributed Computing Systems, Wu Yi Shan, China, 21–23 November 2012; pp. 146–159.
10. Camenisch, J.; Kohlweiss, M.; Rial, A.; Sheedy, C. Blind and anonymous identity-based encryption and authorised private searches on public key encrypted data. In Proceedings of the International Conference on Practice and Theory in Public Key Cryptography—PKC, Irvine, CA, USA, 18–20 March 2009; pp. 96–214.
11. Barth, A.; Boneh, D.; Waters, B. Privacy in encrypted content distribution using private broadcast encryption. In *International Conference on Financial Cryptography and Data Security*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 52–64.
12. Fan, C.I.; Tseng, Y.F. Anonymous multi-receiver identity-based authenticated encryption with CCA security. *Symmetry* **2015**, *7*, 1856–1881. [CrossRef]
13. Lai, J.; Mu, Y.; Guo, F.; Chen, R. Fully privacy-preserving ID-based broadcast encryption with authorization. *Comput. J.* **2017**, *60*, 1809–1821. [CrossRef]

14. Fan, C.I.; Tsai, P.J.; Huang, J.J.; Chen, W.T. Anonymous multi-receiver certificate-based encryption. In Proceedings of the 2013 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, CyberC, Beijing, China, 10–12 October 2013; pp. 19–26.
15. Katz, J.; Sahai, A.; Waters, B. Predicate encryption supporting disjunctions, polynomial equations, and inner products. In Proceedings of the Theory and Applications of Cryptographic Techniques International Conference on Advances in Cryptology, Istanbul, Turkey, 2008; pp. 146–162.
16. Liu, Q.; Guo, Y.; Wu, J.; Wang, G. Effective query grouping strategy in clouds. *J. Comput. Sci. Technol.* **2017**, *32*, 1231–1249. [CrossRef]
17. Liu, Z.; Huang, Y.; Li, J.; Cheng, X.; Shen, C. Divoram: Towards a practical oblivious ram with variable block size. *Inf. Sci.* **2018**, *447*, 1–11. [CrossRef]
18. Jhaveri, R.H.; Patel, N.M.; Zhong, Y.; Sangaiah, A.K. Sensitivity analysis of an attack-pattern discovery based trusted routing scheme for mobile ad-hoc networks in industrial iot. *IEEE Access* **2018**, *6*, 23344–23355. [CrossRef]
19. Cai, Z.; Yan, H.; Li, P.; Huang, Z.A.; Gao, C. Towards secure and flexible ehr sharing in mobile health cloud under static assumptions. *Clust. Comput.* **2017**, *20*, 2415–2422. [CrossRef]
20. Li, J.; Chen, X.; Chow, S.S.; Huang, Q.; Wong, D.S.; Liu, Z. Multi-authority fine-grained access control with accountability and its application in cloud. *J. Netw. Comput. Appl.* **2018**, *112*, 89–96. [CrossRef]
21. Wang, H.; Zheng, Z.; Wu, L.; Li, P. New directly revocable attribute-based encryption scheme and its application in cloud storage environment. *Clust. Comput.* **2017**, *20*, 2385–2392. [CrossRef]
22. Li, J.; Li, J.; Chen, X.; Jia, C.; Lou, W. Identitybased encryption with outsourced revocation in cloud computing. *IEEE Trans. Comput.* **2015**, *64*, 425–437. [CrossRef]
23. Yang, L.; Han, Z.; Huang, Z.; Ma, J. A remotely keyed file encryption scheme under mobile cloud computing. *J. Netw. Comput. Appl.* **2018**, *106*, 90–99. [CrossRef]
24. Wu, Z.; Tian, L.; Li, P.; Wu, T.; Jiang, M.; Wu, C. Generating stable biometric keys for flexible cloud computing authentication using finger vein. *Inf. Sci.* **2016**, *433*, 431–447. [CrossRef]
25. Zhang, J.;Bai, W.; Wang, Y. Non-Interactive ID-Based Proxy Re-Signature Scheme for IoT Based on Mobile Edge Computing. *IEEE Access* **2019**, *7*, 37865–37875. [CrossRef]
26. Quick, D.; Kwang Raymond Choo, K. Pervasive social networking forensics: Intelligence and evidence from mobile device extracts. *J. Netw. Comput. Appl.* **2016**, *86*, 24–33. [CrossRef]
27. Bakas, A.; Michalas, A. Modern family: A revocable hybrid encryption scheme based on attribute-based encryption, symmetric searchable encryption and SGX. In Proceedings of the 15th EAI International Conference on Security and Privacy in Communication Networks (SecureComm '19), Orlando, FL, USA, 23–25 October 2019.
28. Michalas, A. The lord of the shares: Combining attribute-based encryption and searchable encryption for flexible data sharing. In Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing (SAC), Limassol, Cyprus, 8–12 April 2019.
29. Li, J.; Chen, X.; Li, M.; Li, J.; Lee, P.P.C.; Lou, W. Secure de-duplication with efficient and reliable convergent key management. *IEEE Trans. Parallel Distrib. Syst.* **2014**, *25*, 1615–1625. [CrossRef]
30. Libert, B.; Paterson, K.G.; Quaglia, E.A. Anonymous broadcast encryption: Adaptive security and efficient constructions in the standard model. In *International Workshop on Public Key Cryptography LNCS*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7293, pp. 206–224.
31. Fazio, N.; Perera, I.M. Outsider-anonymous broadcast encryption with sublinear ciphertexts. In *International Workshop on Public Key Cryptography*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 225–242.
32. Delerablee, C. Identity-based broadcast encryption with constant size ciphertexts and private keys. In *International Conference on the Theory and Application of Cryptology and Information Security. ASIACRYPT 2007, Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2007; Volume 4833, pp. 200–215.
33. Galindo, D. Boneh-Franklin identity based encryption revisited. In *International Colloquium on Automata, Languages, and Programming*; ICALP 2005, LNCS 3580; Springer: Berlin/Heidelberg, Germany, 2005; pp. 791–802.
34. Wang, H.; Zhang, Y.; Xiong, H.; Qin, B. Cryptanalysis and improvements of an anonymous multi-receiver identity-based encryption scheme. *Inf. Secur. IET* **2012**, *6*, 20–27. [CrossRef]
35. Chien, H.Y. *Improved Anonymous Multi-Receiver Identity-Based Encryption*; Oxford University Press: Oxford, UK, 2012; Volume 55.

36. Zhang, J.; Xu, Y.; Zou, J. Comment on Wang et al.'s anonymous multi-receiver id-based encryption scheme and its improved schemes. *Int. J. Intell. Inf. Database Syst.* **2013**, *7*, 400–413. [CrossRef]
37. Hung, Y.H.; Huang, S.S.; Tseng, Y.M.; Tsai, T.T. Efficient anonymous multireceiver certificateless encryption. *IEEE Syst. J.* **2017**, *11*, 1–12. [CrossRef]
38. Xu, P.; Li, J.; Wang, W.; Jin, H. Anonymous identity-based broadcast encryption with constant decryption complexity and strong security. In Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security, Xi'an, China, 30 May–3 June 2016; pp. 223–233.
39. Liang, K.; Chu, C.K.; Tan, X.; Wong, D.S.; Tang, C.; Zhou, J. Chosen-ciphertext secure multi-hop identity-based conditional proxy re-encryption with constant-size ciphertexts. *Theor. Comput. Sci.* **2014**, *539*, 87–105. [CrossRef]
40. Boyen, X.; Haines, T. Forward-secure linkable ring signatures. In *Australasian Conference on Information Security and Privacy*; Springer: Cham, Switzerland, 2018; pp. 245–264.
41. He, K.; Weng, J.; Liu, J.N.; Liu, J.K.; Liu, W.; Deng, R.H. Anonymous identity-based broadcast encryption with chosen-ciphertext security. In Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security, Xi'an, China, 30 May–3 June 2016; pp. 247–255.
42. Al-Riyami, S.S.; Paterson, K.G. Certificateless Public Key Cryptography. In *Advances in Cryptology—ASIACRYPT 2003. Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2003; Volume 2894, pp. 452–473.
43. Yum, D.H.; Lee, P.J. Generic construction of certificateless encryption. In *International Conference on Computational Science and Its Applications*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 802–811.
44. Libert, B.; Quisquater, J.J. On constructing certificateless cryptosystems from identity based encryption. In *International Workshop on Public Key Cryptography*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 474–490.
45. Islam, S.H.; Khan, M.K.; Al-Khouri, A.M. Anonymous and provably secure certificateless multireceiver encryption without bilinear pairing. *Secur. Commun. Netw.* **2015**, *8*, 2214–2231. [CrossRef]
46. Douceur, J.R.; Adya, A.; Bolosky, W.J.; Simon, P.; Theimer, M. Reclaiming space from duplicate files in a serverless distributed file system. In Proceedings of the 22nd International Conference on Distributed Computing Systems, Vienna, Austria, 2–5 July 2002; pp. 617–624.
47. Bellare, M.; Keelveedhi, S.; Ristenpart, T. Message-locked encryption and secure de-duplication. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 296–312.
48. Bellare, M.; Keelveedhi, S. Interactive message-locked encryption and secure de-duplication. In *IACR International Workshop on Public Key Cryptography*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 516–538.
49. Li, J.; Li, Y.; Chen, X.; Lee, P.; Lou, W. A hybrid cloud approach for secure authorized de-duplication. *IEEE Trans. Parallel Distrib. Syst.* **2015**, *26*, 1206–1216. [CrossRef]
50. Li, J.; Chen, X.; Huang, X.; Tang, S.; Xiang, Y.; Hassan, M.M.; Alelaiwi, A. Secure distributed de-duplication systems with improved reliability. *IEEE Trans. Comput.* **2015**, *64*, 3569–3579. [CrossRef]
51. Li, X.; Li, J.; Huang, F. A secure cloud storage system supporting privacy-preserving fuzzy de-duplication. *Soft Comput.* **2016**, *20*, 1437–1448. [CrossRef]
52. Boneh, D.; Di Crescenzo, G.; Ostrovsky, R.; Persiano, G. Public key encryption with keyword search. In Proceedings of the International Conference on the Theoryand Applications of Cryptographic Techniques, EUROCRYPT 2004, Interlaken, Switzerland, 2–6 May 2004; pp. 506–522.
53. Cui, H.; Deng, R.H.; Li, Y.; Wu, G. Attribute-based storage supporting secure de-duplication of encrypted data in cloud. *IEEE Trans. Big Data* **2017**. [CrossRef]