Wuhan University
Journal of Natural Sciences

2010, Vol.15 No.5, 415-421

Article ID 1007-1202(2010)05-0415-07
DOI 10.1007/s11859-010-0676-9

An Efficient Conversion Scheme for Enhancing Security of Diffie-Hellman-Based Encryption

☐ **ZHANG Xi, HANG Huanhua**

College of Computer and Software, Shenzhen University, Shenzhen 518060, Guangdong, China

Abstract: Nowadays, indistinguishability against adaptive chosen-ciphertext attacks (IND-CCA2) has been widely accepted as a proper security criterion for encryption schemes. In this paper, an efficient conversion is proposed to satisfy the IND-CCA2 security. It uses the random oracle methodology and the idea of hybrid encryption, and can enhance any Diffie-Hellman based encryption scheme, which is only one-way under plaintext-checking attack. Compared with other existing conversions, this conversion has the advantages of short ciphertext and low computation overhead, especially when it is applied to the multi-recipient setting.

Key words: Diffie-Hellman-based encryption; adaptive chosen-ciphertext attack; multi-recipient setting; randomness-reusing
CLC number: TP 305

Received date: 2009-12-03
Foundation item: Supported by the National Natural Science Foundation of China (60903178)
Biography: ZHANG Xi, male, Associate professor, research direction: information security. E-mail:zxsay@126.com

0 Introduction

A fundamental task of cryptography is to protect the secrecy of messages transmitted over public communication lines[1]. For this purpose one can use encryption schemes to encode messages in a way that an eavesdropper cannot decode it. However, as networks become more and more open and accessible, an adversary may not only eavesdrop but also implement active attacks on the channel, e.g., she may try to interact with honest parties by sending ciphertexts to them and analyze their responses. Such active attacks are much more powerful and hard to combat than passive ones. To model this type of attacks, the notion of chosen-ciphertext security was introduced by Naor *et al*[2] and developed in Refs. [3-4]. Nowadays, indistinguishability against adaptive chosen-ciphertext attacks (IND-CCA2) has been widely accepted as a proper security criterion for encryption schemes.

Diffie-Hellman-based encryption schemes have been widely used in the past two decades, e.g., ElGamal encryption scheme[5], Cramer-Shoup encryption scheme[6], and most of the identity-based encryption schemes. However, many of them are just secure in a weaker sense. It is necessary to enhance them to meet the IND-CCA2 security. A promising way is to use a conversion to transform them into IND-CCA2 secure ones. Several conversions (e.g., Refs. [7-14]) have been proposed up to now. However, when these conversions are applied to Diffie-Hellman-based encryptions, they introduce either long ciphertexts or computational overhead, especially

when they are applied in the multi-recipient setting. In view of this, we propose a new conversion, which can enhance any Diffie-Hellman based encryption scheme, which is only one-way under plaintext-checking attack, to satisfy the IND-CCA2 security. Compared with other conversions, the proposed conversion has the advantages of short ciphertext and low computational cost. These advantages become more obvious in the multi-recipient settings.

1 Preliminaries

In this section, we shall review the definitions and security notions for asymmetric encryption and symmetric encryption. In addition, we shall formalize the definition of Diffie-Hellman-based encryption, and then take ElGamal encryption scheme as an example to explain it.

1.1 Asymmetric Encryption

Definition 1 An asymmetric scheme AE = (G, K, E, D) consists of four algorithms:

● $G(k)$: The common key generation algorithm takes as inputting a security parameter k and outputting a common key I, denoted by $I \leftarrow G(k)$;

● $K(I)$: The key generation algorithm takes as inputting the common key I and returning a public/private key pair (pk, sk), denoted by $(\mathrm{pk}, \mathrm{sk}) \leftarrow K(I)$;

● $E_{\mathrm{pk}}(m, r)$: The encryption algorithm takes as inputting a public key pk, a plaintext $m \in M$ and a random coin $r \in \Omega$, and returning a ciphertext C, denoted by $C \leftarrow E_{\mathrm{pk}}(m, r)$. When the random coins are useless in the discussion, denoted by $C \leftarrow E_{\mathrm{pk}}(m)$;

● $D_{\mathrm{sk}}(C)$: The decryption algorithm takes as inputting the secret key sk and a ciphertext C, and returning the corresponding plaintext m or a special symbol $\perp$ indicating that the ciphertext is invalid, denoted by $m \leftarrow D_{\mathrm{sk}}(C)$.

The security proof for our proposed scheme will involve two security notions for asymmetric encryption, i.e., one-way under plaintext-checking attack (OW-PCA) and indistinguishable against adaptive chosen-ciphertext attack (IND-CCA2). Thus, we review these two security notions as follows:

Definition 2[11] Let O_{PCA} be a plaintext-checking oracle, which can show whether a ciphertext/plaintext pair (C, m) is valid, i.e., whether $D_{\mathrm{sk}}(C) = m$ holds. An asymmetric encryption scheme AE = (G, K, E, D) is said to be (t, q, ε) OW-PCA secure, if for any adversary A which runs within time t and makes at most q queries to O_{PCA}, its advantage defined as below satisfies

$$\mathrm{Adv}_{\mathrm{AE},A}^{\text{OW-PCA}} = \Pr\left[\begin{array}{l} I \leftarrow G(k); (\mathrm{pk}, \mathrm{sk}) \leftarrow K(I); m \in_{\mathrm{R}} M; r \in_{\mathrm{R}} \Omega; \\ C^* \leftarrow E_{\mathrm{pk}}(m, r) : A^{O_{\mathrm{PCA}}}(\mathrm{pk}, C^*) = m \end{array}\right] < \varepsilon,$$

where C^* denotes the challenge ciphertext, and the above probability is also taken over the random choice of A.

Definition 3[15] Let $O_{D_{\mathrm{sk}}}$ be a decryption oracle, which on input a ciphertext C returns the result of $D_{\mathrm{sk}}(C)$. An asymmetric encryption scheme AE = (G, K, E, D) is said to be (t, q, ε) IND-CCA2 secure, if for any adversary A which runs within time t and makes at most q queries to the decryption oracle $O_{D_{\mathrm{sk}}}$, its advantage defined as below satisfies

$$\mathrm{Adv}_{AE,A}^{\text{IND-CCA2}}(k) = \left| 2 \times \Pr\left[\begin{array}{l} I \leftarrow G(k); (\mathrm{pk}, \mathrm{sk}) \leftarrow K(I); \\ (m_0, m_1) \leftarrow A^{O_{D_{\mathrm{sk}}}}(\mathrm{find}, \mathrm{pk}); \\ b \in_{\mathrm{R}} \{0,1\}; r \in_{\mathrm{R}} \Omega; C^* \leftarrow E_{\mathrm{pk}} \\ (m_b, r) : A^{O_{D_{\mathrm{sk}}}}(\mathrm{guest}, C^*) = b \end{array}\right] - 1 \right| < \varepsilon$$

where A is not allowed to query C^* to the decryption oracle in the guess stage, and the probability is also taken over the random choice of A.

1.2 Symmetric Encryption

Definition 4 A symmetric encryption scheme SE=(K^{sym}, E^{sym}, D^{sym}) consists of three algorithms:

● $K^{\mathrm{sym}}(k)$: The key generation algorithm takes as inputting the security parameter k and returning a symmetric key K, denoted by $K \leftarrow K^{\mathrm{sym}}(k)$;

● $E_K^{\mathrm{sym}}(m)$: The encryption algorithm takes as inputting the symmetric key K and a plaintext $m \in M$, and outputting a ciphertext C, denoted by $C \leftarrow \mathrm{E}_K^{\mathrm{sym}}(m)$;

● $D_K^{\mathrm{sym}}(C)$: The decryption algorithm takes as inputting the symmetric key K and a ciphertext C, and outputting the plaintext m, denoted by $m \leftarrow D_K^{\mathrm{sym}}(C)$.

The widely accepted security notion for symmetric encryption is indistinguishable against chosen-ciphertext attack (IND-CCA), which is reviewed as below:

Definition 5[16] Let O_{E} be an encryption oracle which on input a plaintext returns the ciphertext, and O_{D} be a decryption oracle which on input a ciphertext returns the plaintext. A symmetric encryption scheme SE=(K^{sym}, E^{sym}, D^{sym}) is said to be (t, ε) IND-CCA secure, if for any adversary A with running time bounded by t, its advantage defined as below satisfies

$$\mathrm{Adv}_{\mathrm{SE},A}^{\text{IND-CCA}}(k) =$$

$$\left|2\times\Pr\left[\begin{array}{l}K\leftarrow K(k);(m_0,m_1)\leftarrow A^{O_{\mathrm{E}},O_{\mathrm{D}}}(\mathrm{find});\\ b\in_{\mathrm{R}}\{0,1\};C^*\leftarrow \mathrm{E}_K^{\mathrm{sym}}(m_b):\\ A^{O_{\mathrm{E}},O_{\mathrm{D}}}(\mathrm{guess},C^*)=b\end{array}\right]-1\right|<\varepsilon$$

where A is disallowed to query m_0 and m_1 to the encryption oracle, A is also disallowed to query C^* to the decryption oracle in the guess stage, and the probability is also taken over the random choice of A.

1.3 Diffie-Hellman-Based Encryption

Although Diffie-Hellman-based encryption schemes have been widely used, no formal definition of them has been given so far. In this section, inspired by the definition of ElGamal encryption scheme[5], we formalize the definition of Diffie-Hellman-based encryption schemes, and give an example for this type of schemes.

Definition 6 An asymmetric encryption scheme AE = (G, K, E, D) working as follows is called a Diffie-Hellman-based encryption scheme:

- $G(k)$: $I\leftarrow G(k)$;
- $K(I)$: (pk, sk) $\leftarrow K(I)$;
- $E_{\mathrm{pk}}(m, r)$: The encryption algorithm can be partitioned into two phases: taken as inputting pk and r, it generates one ciphertext part C_{DH} and a special value AK; it then encrypts m into another ciphertext part C_M using AK. Formally, we write $C_{\mathrm{DH}}\leftarrow f_{\mathrm{pk}}(r)$, $\mathrm{AK}\leftarrow \hat{f}^1_{\mathrm{pk}}(r)$ and $C_M\leftarrow h_{\mathrm{AK}}(m)$, where $f, \hat{f}^1$, and h denote functions and can be determined in concrete schemes. The final ciphertext consists of $C=(C_{\mathrm{DH}}, C_M)$.
- $D_{\mathrm{sk}}(C)$: The decryption algorithm can also be partitioned into two phases: it first recovers AK from C_{DH} using sk; then it recovers the plaintext m from C_M using AK. Formally, we write $\mathrm{AK}\leftarrow \hat{f}^2_{\mathrm{sk}}(C_{\mathrm{DH}})$, $m\leftarrow h^{-1}_{\mathrm{AK}}(C_M)$, where $\hat{f}^2$ is a function and can be determined in concrete schemes, h^{-1} is the inverse function of h, i.e., for any $m\in M$, $h^{-1}_{\mathrm{AK}}(h_{\mathrm{AK}}(m))=m$ always holds.

Note that AK plays an important role in this definition. On one hand, AK can be computed by the encrypter using r. On the other hand, it can be recovered from C_{DH} by the decrypter using sk. In this sense, AK is similar to an agreed-key in a Diffile-Hellman key agreement protocol. That is why we name this type of schemes Diffie-Hellman-based encryption schemes. We call AK the agreed-key of C_{DH}. We also call (C_{DH}, AK) a valid ciphertext/agreed-key pair if $\hat{f}^2_{\mathrm{sk}}(C_{\mathrm{DH}})=\mathrm{AK}$ holds. Note that (C_{DH}, AK) is a valid ciphertext/agreed-key pair if ((C_{DH}, AK), m) is a valid ciphertext / plaintext pair.

For an easier understanding of Definition 6, we take ElGamal encryption scheme[5] as an example:

Example 1 ElGamal encryption scheme AE = (G, K, E, D) works as follows.

- $G(k)$: Choose a large prime q with $2^{k-1}<q<2^k$ and $2q+1$ is a prime. A q-order group G_q and a generator g of G_q are also chosen. The common key is $I=(q, G_q, g)$;
- $K(I)$: Choose $x\in_{\mathrm{R}} Z_q^*$ and compute $y=g^x$. The secret key is (q, g, x) and the public key is (q, g, y);
- $E_{\mathrm{pk}}(m, r)$: Choose $r\in_{\mathrm{R}} Z_q^*$ and compute $C_{\mathrm{DH}}=f_{\mathrm{pk}}(r)=g^r$, $\mathrm{AK}=\hat{f}^1_{\mathrm{pk}}(r)=y^r=g^{xr}$ and $C_M=h_{\mathrm{AK}}(m)=\mathrm{AK}\cdot m$. The ciphertext consists of $C=(C_{\mathrm{DH}}, C_M)$.
- $D_{\mathrm{sk}}(C)$: Parse C as (C_{DH}, C_M), compute AK= $\hat{f}^2_{\mathrm{sk}}(C_{\mathrm{DH}})=C_{\mathrm{DH}}^x=g^{xr}$, and recover the plaintext as $m=h^{-1}_{\mathrm{AK}}(C_M)=C_M/\mathrm{AK}$.

The encrypter computes the agreed-key AK with $y^r=g^{xr}$, while the decrypter computes AK with $C_{\mathrm{DH}}^x=g^{rx}$. It has been proven in Ref. [11] that the OW-PCA security of ElGamal encryption scheme is based on the gap Diffie-Hellman (GDH) assumption.

2 Proposed Conversion Scheme

2.1 Scheme Description

We consider two encryption schemes. One is an OW-PCA secure Diffie-Hellman-based encryption scheme AE = (G, K, E, D), the other is an IND-CCA secure length preserving symmetric encryption scheme SE = (K^{sym}, E^{sym}, D^{sym}). Supposing the symmetric key length of SE is l, we also consider a hash function H which outputs l-bit strings. Then, we construct a hybrid scheme HE = (G^{hyb}, K^{hyb}, E^{hyb}, D^{hyb}) working as follows:

- G^{hyb}: run $G(k)$ and output a common key I;
- $K^{\mathrm{hyb}}(I)$: run $K(I)$ and output a public/secret key pair (pk, sk);
- $\mathrm{E}^{\mathrm{hyb}}_{\mathrm{pk}}(m, r)$: For any message $m\in M$ and random coin $r\in\Omega$, it first computes $C_{\mathrm{DH}}=f_{\mathrm{pk}}(r)$ and $\mathrm{AK}=\hat{f}^1_{\mathrm{pk}}(r)$, then computes $K=(C_{\mathrm{DH}},\mathrm{AK})$ as the symmetric key. Finally, it computes another ciphertext part as $C_{\mathrm{sym}}=E^{\mathrm{sym}}_K(m)$. The final ciphertext is $C=(C_{\mathrm{DH}}, C_{\mathrm{sym}})$;
- $D^{\mathrm{hyb}}_{\mathrm{sk}}(C)$: parse C=(C_{DH}, C_{sym}), extract AK = $\hat{f}^2_{\mathrm{sk}}(C_{\mathrm{DH}})$ and compute K=(C_{DH}, AK). Finally, the plaintext can be recovered as $m=D^{\mathrm{sym}}_K(C_{\mathrm{sym}})$.

2.2 Security and Comparison in the Single Recipient Setting

About the security analysis, we have the following theorems:

Theorem 1 Suppose the hash function H acts as a random oracle. Then, the resulting scheme HE obtained from our conversion is IND-CCA2 secure in the random oracle model, assuming the asymmetric encryption scheme AE is OW-PCA secure and the symmetric en-

cryption scheme SE is IND-CCA secure. Concretely, if there exists a (t, q_H, q_D, ε) adversary A against the IND-CCA2 security of HE, then for any $0<v<\varepsilon$, there exists either

● a (t', q_O, φ) adversary B against the OW-PCA security of AE, where $q_O \leqslant q_H$, $\varphi \geqslant \varepsilon - v - \frac{q_D}{2^l}$, and $t' \leqslant t + q_H(T_{O_{PCA}} + T_{O_E}) + q_D T_{O_D}$;

or

● a (t',v) adversary C against the IND-CCA security of SE, where q_H, q_D and q_O are the number of oracle queries to H, O_D and O_{PCA} respectively. T_{PCA}, T_{O_E}, and T_{O_D} are the running time of the PCA oracle, symmetric encryption oracle and symmetric decryption oracle, respectively.

Proof Without loss of generality, we assume that SE is (t', v) -IND-CCA secure for some probability $0<v<\varepsilon$. Now, we construct an adversary B against the (t', q_O, φ)-OW-PCA security of AE by interacting with A.

Suppose B is given a public key pk and a challenge ciphertext $\hat{C}=(C^*_{DH}, \hat{C}_M)$. B's goal is to output the corresponding plaintext $\hat{m}$ such that $D_{sk}(\hat{C})=\hat{m}$ holds. B interacts with A as follows:

B forwards pk to A. In the find stage, B answers the H-hash queries and the decryption queries for A as follows:

● H-queries: B maintains a hash list H-list, which is initially empty. When A submits a pair ($C_{DH,i}$, AK_i) to the H oracle, B checks whether ($C_{DH,i}$, AK_i) exists in H-list for some $K_i \in_R \{0,1\}^l$ and $b_i \in_R \{0,1\}$. If it does, K_i is returned to A; Otherwise, B chooses $m_i \in_R M$, computes $C_{M,i}=h_{AK_i}(m_i)$ and queries $((C_{DH,i}, C_{M,i}), m_i)$ to its plaintext-checking oracle.

➢ If $\left((C_{DH,i}, C_{M,i}), m_i\right)$ is a valid ciphertext/plaintext pair and $C_{DH,i}=C^*_{DH}$ holds (denote this event by **E**), B runs $h^{-1}_{AK_i}(C^*_M)$, outputs the associated plaintext as the OW-PCA solution and halts.

➢ If $((C_{DH,i}, C_{M,i}), m_i)$ is a valid ciphertext/plaintext pair and $C_{DH,i} \neq C^*_{DH}$, B checks whether $(C_{DH,i}, -)$ exists in H-list for some $K_i \in \{0,1\}^l$ and some b_i=1. If yes, B adds $(C_{DH,i}, AK_i, K_i, b_i)$ in H-list and deletes tuple $(C_{DH,i}, -, K_i, b_i)$ from H-list; otherwise, B chooses $K_i \in_R \{0,1\}^l$, sets $b_i=1$ and adds the tuple $(C_{DH,i}, AK_i, K_i, b_i)$ in H-list;

➢ If $((C_{DH,i}, C_{M,i}), m_i)$ is not a valid ciphertext/plaintext pair, then B chooses $K_i \in_R \{0,1\}^l$, set $b_i=0$ and adds $(C_{DH,i}, AK_i, K_i, b_i)$ in H-list;

● Decryption queries: when a decryption query $C_i=(C_{DH,i}, C_{sym,i})$ is requested, if H-list contains a tuple $(C_{DH,i}, AK_i, K_i, 1)$ or $(C_{DH,i}, -, K_i, 1)$ for some $K_i \in \{0,1\}^l$, B uses K_i as the symmetric key to recover the plaintext as $m_i = D^{sym}_{K_i}(C_{sym,i})$ and returns it to A; Otherwise, B chooses $K_i \in_R \{0,1\}^l$, set $b_i=0$ and adds $(C_{DH,i}, -, K_i, b_i)$ on H-list. Then, B uses K_i as the symmetric key to recover the plaintext as $m_i = D^{sym}_{K_i}(C_{sym,i})$ and returns it to A.

After the find stage, A outputs two equal-length messages $m_0, m_1 \in M$. B flips a bit coin $b_i \in_R \{0,1\}$, chooses $K_i \in_R \{0,1\}^l$ and computes $C^*_{sym}=E^{sym}_{K^*}(m_b)$. B gives $C^*=(C^*_{DH}, C^*_{sym})$ as the challenge ciphertext to A.

In the guess stage, B processes the H-queries and the decryption queries in the same way as in the find stage.

Finally, A outputs a bit b' as the guess for b.

This completes the description of the simulation. Now, we begin to analyze the simulation. From the above specification of B, it can be easily seen that the running time of B is bounded by $t' \leqslant t + q_H(T_{O_{PCA}} + T_{O_E}) + q_D T_{O_D}$, and the number of PCA oracle queries is at most q_H. Next, we proceed to examine the advantage of B. Note that the responses to A's H-queries are indistinguishable from the real environment, since each response is uniformly random and independently distributed in $\{0,1\}^l$. C^* is a valid challenge ciphertext for A since H is a random hash function. The only failure of the simulation provided for A is the event that some decryptions may be incorrect (denote this event by **FD**).

Now, we proceed to bound the probability of event **FD**. Note that **FD** happens only when A asks some ciphertexts, which are valid and produced without asking the hash function H, to the decryption oracle. However, without asking the H oracle, the only way for A to produce a valid ciphertext is to guess the l-bit long symmetric key. Since H is a random hash function mapping to l-bit long string, the probability of extracting the right symmetric key is at most $1/2^l$. This implies that A can produce a valid ciphertext without asking H with probability limited by $1/2^l$. Thus, we have $\Pr[\mathbf{FD}] \leqslant \frac{q_D}{2^l}$.

If event **FD** does not happen, then the simulation provided for A is indistinguishable from the real environment. Since $\Pr[b'=b]=\frac{\varepsilon+1}{2}$ holds in the real environment, $\Pr[b'=b \mid \neg\mathbf{FD}]=\frac{\varepsilon+1}{2}$ also holds.

Next, we shall bound the probability of event **E**. From the specification of B, it is obvious that if event **E** does not happen, the only way for A to produce a correct guess for b is to succeed in a chosen-ciphertext attack against the

symmetric encryption scheme SE. Since we assume that SE is (t',v) IND-CCA secure, $\Pr[b'=b|\neg\mathbf{E}]\leqslant\frac{v+1}{2}$ holds in the real environment. Thus $\Pr[(b'=b|\neg\mathbf{E})|\neg\mathbf{FD}]\leqslant\frac{v+1}{2}$ also holds. Then, we have

$$\begin{aligned}
&\frac{(\varepsilon+1)}{2}=\Pr[b'=b|\neg\mathbf{FD}]\\
&=\Pr[(b'=b|\mathbf{E})|\neg\mathbf{FD}]\Pr[\mathbf{E}|\neg\mathbf{FD}]\\
&\quad+\Pr[(b'=b|\neg\mathbf{E})|\neg\mathbf{FD}]\Pr[\neg\mathbf{E}|\neg\mathbf{FD}]\\
&\leqslant\Pr[\mathbf{E}|\neg\mathbf{FD}]+\Pr[(b'=b|\neg\mathbf{E})|\neg\mathbf{FD}](1-\Pr[\mathbf{E}|\neg\mathbf{FD}])\\
&\leqslant\Pr[\mathbf{E}|\neg\mathbf{FD}]+\frac{(1+v)}{2}(1-\Pr[\mathbf{E}|\neg\mathbf{FD}])\\
&=\frac{(1+v)}{2}+\frac{(1-v)}{2}\Pr[\mathbf{E}|\neg\mathbf{FD}]\\
&\leqslant\frac{(1+v)}{2}+\frac{1}{2}\Pr[\mathbf{E}|\neg\mathbf{FD}]
\end{aligned}$$

which implies $\Pr[\mathbf{E}|\neg\mathbf{FD}]\geqslant\varepsilon-v$. Therefore,

$$\begin{aligned}
\Pr[\mathbf{E}]&\geqslant\Pr[\mathbf{E}|\neg\mathbf{FD}]\Pr[\neg\mathbf{FD}]\\
&\geqslant(\varepsilon-v)\left(1-\frac{q_{\mathrm{D}}}{2^l}\right)>\varepsilon-v-\frac{q_{\mathrm{D}}}{2^l}
\end{aligned}$$

According to the specification of B, if event $\mathbf{E}$ does happen, then B can recover $\hat{m}$ and break the OW-PCA security of AE successfully. This implies that B can break the OW-PCA security of AE with probability at least $\varepsilon-v-\frac{q_{\mathrm{D}}}{2^l}$.

Thus, we successfully construct a (t',q_O,φ) adversary B against the OW-PCA security of AE, where $\varphi\geqslant\varepsilon-v-\frac{q_{\mathrm{D}}}{2^l}$, $q_O\leqslant q_{\mathrm{H}}$, $t'\leqslant t+q_{\mathrm{H}}(T_{O_{\mathrm{PCA}}}+T_{O_{\mathrm{E}}})+q_{\mathrm{D}}T_{O_{\mathrm{D}}}$. Therefore, the proof of Theorem 1 is completed.

Next, we compare our conversion with other conversions in the standard setting, i.e., single recipient setting. To proceed, we also take the ElGamal encryption scheme as an example. We here first explain the notations used in Table 1. Let SM, M and I denote the computation cost of exponentiation, multiplication and inversion in G_q respectively. E_{sym} and D_{sym} denote the computation cost of symmetric encryption and decryption algorithm, respectively. H denotes the computation cost of a hash function. Let G denote the element from G_q. L_{sym} and L_{H} denote the bit length of the output of symmetric encryption scheme and hash function, respectively.

Table 1 indicates that our conversion enjoys short ciphertext and low computation overhead compared with those conversions in Refs. [8, 9, 11, 12]. Even compared with the OAEP 3-round conversion[14] without redundancy, our conversion has competitive performance. As to the computation cost, our conversion saves two Hs and substitutes one M (resp., I) with one E_{sym} (resp., D_{sym}) in the encryption (resp., decryption) phase. Note that symmetric encryption and decryption can be performed very efficiently. As to the bandwidth, our conversion substitutes one G with one L_{sym}. Note that the scheme obtained from our conversion is a hybrid one, which supports the encryption of long plaintext. On the contrary, to encrypt long plaintext, the scheme obtained from OAEP 3-round conversion has to be carried out several times. Thus, our conversion can yield savings in the long plaintext scenario. The next section will further show that the advantages of our conversion over other existing conversions become significant in the multi-recipient setting.

2.3 Security and Comparison in the Multi-Recipient Setting

We first review some definitions and results in the multi-recipient encryption setting. For more details, we refer to Refs. [17, 18].

The multi-recipient encryption setting considers the following scenarios. There are n recipients, numbered by 1, ···, n. Each recipient i has a public/secret key pair (pk_i, sk_i). Suppose a sender want to send messages m_1, ···, m_n for recipients 1, ···, n, respectively. The sender encrypts these messages into a ciphertext C and sends it to all the recipients. After receiving the ciphertext C, each recipient i uses his secret key sk_i to recover message m_i from C. For this purpose, there is of course a naive way: the

Table 1 Comparison of our conversion and other conversion schemes in the single recipient setting

Conversion scheme	Computation cost		Bandwidth
	Encryption	Decryption	
Fujisaki-Okamoto[9]	2SM+1M+1E_{sym}+2H	3SM+1M+1I+1D_{sym}+2H	2G+1L_{sym}
Pointcheval[12]	2SM+1M+2H	3SM+1M+1I+2H	2G+1L_{H}
REACT[11]	2SM+1M+1E_{sym}+2H	1SM+1I+1D_{sym}+2H	2G+1L_{sym}+1L_{H}
GEM[8]	2SM+1M+1E_{sym}+3H	1SM+1I+1D_{sym}+3H	2G+1L_{sym}
OAEP 3-round[14]	2SM+1M+3H	1SM+1I+3H	2G
Our conversion	2SM+1E_{sym}+1H	1SM+1D_{sym}+1H	1G+1L_{sym}

ciphertext C is just the concatenation of independently encrypted messages for n recipients, i.e., $C = C_1||\cdots||C_n$ where $C_i = E_{\mathrm{pk}_i}(m_i, r_i)$ and $r_i \in_{\mathrm{R}} \Omega$. Kurosawa[18] for the first time noticed that randomness-reusing in some schemes can yield savings in computation cost and bandwidth. Randomness-reusing means the encryption algorithm consumes the same random coin r_i for all the recipients. However, as pointed out by Bellare *et al*[17], randomness-reusing in some schemes will affect the security of these schemes in the multi-recipient setting. To study the security of encryption schemes in randomness-reusing multi-recipient setting, they introduced a definition named reproducibility and a theorem called reproducible theorem.

Let us briefly review the definition of reproducibility. Let pk_1, pk_2 be public keys, and let $C_1 = E_{\mathrm{pk}_1}(m_1, r_1)$. Roughly speaking, an encryption scheme is called reproducible, if given I, pk_1, pk_2, C_1, any message m_2, and the secret key sk_2 corresponding to pk_2, there exists a polynomial time reproduction algorithm that returns the ciphertext $C_2 = E_{\mathrm{pk}_2}(m_2, r)$.

Formally, reproducibility is defined as follows:

Definition 7[17] An asymmetric encryption scheme AE = (G, K, E, D) is called reproducible, if for any k there exists a polynomial time algorithm R called reproduction algorithm such that the following experiment outputs 1 with probability 1:

Experiment $\mathrm{Exp}^{\mathrm{repr}}_{\mathrm{AE},R}(k)$:

$I \leftarrow G(k); (\mathrm{pk}_1, \mathrm{sk}_1) \leftarrow K(I); m_1 \in_{\mathrm{R}} M; r \in_{\mathrm{R}} \Omega;$

$C_1 \leftarrow E_{\mathrm{pk}_1}(m_1, r); (\mathrm{pk}_2, \mathrm{sk}_2) \leftarrow K(I); m_2 \in_{\mathrm{R}} M;$

$C_2 \leftarrow R(I, \mathrm{pk}_1, C_1, m_2, \mathrm{pk}_2, \mathrm{sk}_2);$

If $E_{\mathrm{pk}_2}(m_2, r) = C_2$ holds, then return 1; else return 0

End.

One can confirm that many Diffie-Hellman-based encryption schemes are reproducible, e.g., ElGamal scheme[5], Boneh-Franklin identity-based encryption scheme[19], and Cramer-Shoup scheme[20], etc.

Given the security of a reproducible encryption scheme, the following reproducible theorem[17] can determine its security in randomness-reusing multi-recipient setting.

Lemma 1 If an asymmetric encryption scheme AE = (G, K, E, D) is reproducible and IND-CPA (resp., IND-CCA2) secure, then it is also IND-CPA (resp., IND-CCA2) secure in randomness reusing multi-recipient setting.

Interestingly, our conversion has the following result.

Lemma 2 Assuming that a Diffie-Hellman-based encryption scheme AE=(G, K, E, D) is reproducible, then the encryption scheme HE obtained from our conversion is reproducible.

Proof Let HE = (G^{hyb}, K^{hyb}, E^{hyb}, D^{hyb}). Since AE is reproducible, there exists a polynomial time reproduction algorithm R for AE. Now, we show how to construct a polynomial time reproduction algorithm R' for HE using R.

Given $I \leftarrow G^{\mathrm{hyb}}(k), (\mathrm{pk}_1, \mathrm{sk}_1) \leftarrow K^{\mathrm{hyb}}(I), m_1 \in_{\mathrm{R}} M$, $r \in_{\mathrm{R}} \Omega, (C_{\mathrm{DH},1}, C_{\mathrm{sym},1}) \leftarrow \mathrm{E}^{\mathrm{hyb}}_{\mathrm{pk}_1}(m_1, r), (\mathrm{pk}_2, \mathrm{sk}_2) \leftarrow K^{\mathrm{hyb}}(I)$ and $m_2 \in_{\mathrm{R}} M$, the reproduction algorithm R' for HE works as follows:

Algorithm $R'(I, \mathrm{pk}_1, (C_{\mathrm{DH},1}, C_{\mathrm{sym},1}), m_2, \mathrm{pk}_2, \mathrm{sk}_2)$

Choose a random ciphertext part $C^*_{M,1}$;

Compute $(C_{\mathrm{DH},2}, C^*_{\mathrm{M},2}) \leftarrow R(I, \mathrm{pk}_1, (C_{\mathrm{DH},1}, C^*_{M,1}), m_2,$ $\mathrm{pk}_2, \mathrm{sk}_2)$;

Compute $\mathrm{AK}_2 \leftarrow \hat{f}^2_{\mathrm{sk}_2}(C_{\mathrm{DH},2})$;

Compute $K_2 = H(C_{\mathrm{DH},2}, \mathrm{AK}_2)$ and $C_{\mathrm{sym},2} = \mathrm{E}^{\mathrm{sym}}_{K_2}(m_2)$;

Return $(C_{\mathrm{DH},2}, C_{\mathrm{sym},2})$

End.

Note that in the above algorithm, $\mathrm{E}^{\mathrm{hyb}}_{\mathrm{pk}_2}(m_2, r) = (C_{\mathrm{DH},2}, C_{\mathrm{sym},2})$ holds. Thus, the proof of Lemma 2 is completed.

According to Theorem 1 and Lemmas 1 and 2, the following theorem can be drawn.

Theorem 2 Assuming that a Diffie-Hellman-based encryption scheme AE is reproducible and OW-PCA secure, then the resulting encryption scheme HE obtained from our conversion is IND-CCA2 secure in randomness-reusing multi-recipient encryption setting.

Theorem 2 shows that our conversion can be applied to those reproducible and OW-PCA secure schemes and the resulting schemes are IND-CCA2 secure in the randomness-reusing multi-recipient setting. Furthermore, we shall show that randomness-reusing enables these resulting schemes to enjoy short ciphertext and low computation cost. However, those schemes obtained from other conversions do not enjoy these advantages.

In Table 2, we compare the scheme obtained from our conversion with those obtained from other conversions in the multi-recipient setting. Table 2 indicates that the schemes obtained from other conversions have the same bandwidth and computation overhead as the naive one in the multi-recipient setting. On the contrary, ours can save $(n-1)$SM in the encryption phase, and $(n-1)G$ in bandwidth. According to Tables 1 and 2, it can be found that the advantages of our conversion over other conversions become significant in the multi-recipient setting.

Table 2 Comparison of our conversion and other conversion schemes in the multi-recipient setting

Conversion scheme	Computation cost		Bandwidth
	Encryption	Decryption	
Fujisaki-Okamoto[9]	$2n$SM+nM+nE_{sym}+$2nH$	$3n$SM+nM+nI+nD_{sym}+$2nH$	$2nG$+nL_{sym}
Pointcheval[12]	$2n$SM+nM+$2nH$	3SM+1M+1I+$2H$	$2nG$+nL_H
REACT[11]	$2n$SM+nM+nE_{sym}+$2nH$	nSM+nI+nD_{sym}+$2nH$	$2nG$+nL_{sym}+nL_H
GEM[8]	$2n$SM+nM+nE_{sym}+$3nH$	nSM+nI+nD_{sym}+$3nH$	$2nG$+nL_{sym}
OAEP 3-round[14]	$2n$SM+nM+$3nH$	nSM+nI+$3nH$	$2nG$
Our conversion	(n+1)SM+ nE_{sym}+ nH	nSM+nD_{sym}+nH	1G+nL_{sym}

3 Conclusion

In this paper, a new conversion for Diffie-Hellman-based encryption schemes is presented. Compared with other existing conversions, ours enjoys short ciphertext and low computation overhead. It is interesting to note that these advantages become significant in the multi-recipient setting.

References

[1] Abe M, Gennaro R, Kurosawa K. Tag-KEM/DEM: A new framework for hybrid encryption[C]//*Proceedings of EUROCRYPT*'05. Berlin: Springer-Verlag, 2005: 128-146.

[2] Naor M, Yung M. Public-key cryptosystems provably secure against chosen ciphertext attacks [C]//*Proceedings of STOC*'90. New York: ACM Press, 1990: 427-437.

[3] Dolev D, Dwork C, Naor M. Non-malleable cryptography [C]//*Proceedings of STOC*'91. New York: ACM Press, 1991: 542-552.

[4] Rackoff C, Simon D. Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack[C]//*Proceedings of CRYPTO*'91. Berlin: Springer-Verlag, 1992: 433-444.

[5] ElGamal T. A Public key cryptosystem and a signature scheme based on discrete logarithms [J]. *IEEE Transactions on Information Theory*, 1985, **IT-31**(4): 469-472.

[6] Cramer R, Shoup V. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack[C]// *Proceedings of Crypto*'98. Berlin: Springer-Verlag, 1998: 13-25.

[7] Bellare M, Rogaway, P. Optimal asymmetric encryption—how to encrypt with RSA[C]//*Proceedings of Eurocrypto*'94. Berlin: Springer-Verlag, 1995: 92-111.

[8] Coron J S, Handschuh H, Joye M, *et al*. GEM: A generic chosen-ciphertext secure encryption method[C]//*Proceedings of CT-RSA*'02. Berlin: Springer- Verlag, 2002: 263-276.

[9] Fujisaki E, Okamoto T. Secure integration of asymmetric and symmetric encryption schemes[C]// *Proceedings of Crypto*'99. Berlin: Springer-Verlag, 1999: 537-554.

[10] Fujisaki E, Okamoto T. How to Enhance the security of public-key encryption at minimum cost[C]// *Proceedings of PKC*'99. Berlin: Springer-Verlag, 1999: 53-68.

[11] Okamoto T, Pointcheval D. REACT: Rapid enhanced-security asymmetric cryptosystem transform[C]//*Proceedings of CT-RSA*'01. Berlin: Springer-Verlag, 2001: 159-175.

[12] Pointcheval D. Chosen-ciphertext security for any one-way cryptosystem[C]//*Proceedings of PKC*'00. Berlin: Springer-Verlag, 2000: 129-146.

[13] Phan D H, Pointcheval D. Chosen-ciphertext security without redundancy[C]//*Proceedings of Asiacrypto*'03. Berlin: Springer-Verlag, 2003: 1-18.

[14] Phan D H, Pointcheval D. OAEP 3-Round: a generic and secure asymmetric encryption padding[C]// *Proceedings of Asiacrypto*'04 (LNCS 3329). Berlin: Springer-Verlag, 2004: 63-77.

[15] Goldwasser S, Micali S. Probabilistic encryption[J]. *Journal of Computer and System Sciences*, 1984, **28**: 270-299.

[16] Kurosawa K, Matsuo T. How to remove MAC from DHIES [C]//*Proceedings of ACISP*'04. Berlin: Springer-Verlag, 2004: 236-247.

[17] Bellare M, Boldyreva A, Staddon J. Multi-recipient encryption schemes: Security notions and randomness re-use[C]// *Proceedings of PKC*'03. Berlin: Springer-Verlag, 2003: 85-99.

[18] Kurosawa K. Multi-recipient public-key encryption with shortened ciphertext[C]//*Proceedings of PKC*'02. Berlin: Springer-Verlag, 2002: 48-63.

[19] Boneh D, Franklin M. Identity-based encryption from the Weil pairing[C]//*Proceedings of CRYPTO*'01, Berlin: Springer-Verlag, 2001: 213-229.

[20] Cramer R, Shoup V. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack [C]// *Proceedings of CRYPTO*'98. Berlin: Springer-Verlag, 1998: 13-25.

□