A survey of IND-CCA constructions

Ganyu (Bruce) Xu (g66xu)

CO 789, Winter 2024

1 The Fujisaki-Okamoto Transformation

The Fujisaki-Okamoto transformation [1] takes a IND-CPA public-key encryption scheme (PKE) and a few other cryptographic primitives of weaker security as inputs, and constructs a hybrid encryption scheme that achieves IND-CCA2 security.

In the conference paper, the security reduction is performed under the random oracle model, and the security of the hybrid scheme degrades linearly with the number of hash queries and the number of decryption queries. On the other hand, this scheme is desirable for its simplicity, and the low requirements for the input primitives: the PKE only needs to be one-time one-way secure, and the symetric cipher only needs to be indistinguishable under one-time attack.

1.1 The hybrid scheme and security result

The hybrid encryption scheme contains three routines: key generation, encryption, and decryption. The input for the hybrid scheme includes a public-key encryption scheme $\text{KeyGen}^{\text{asym}}, E^{\text{asym}}, D^{\text{asym}}$, a symmetric encryption scheme $E^{\text{sym}}, D^{\text{sym}}$, and two hash functions $G : \mathcal{M}^{\text{asym}} \mapsto \mathcal{K}^{\text{sym}}, H : \mathcal{M}^{\text{asym}} \times \mathcal{C}^{\text{sym}} \mapsto \text{Coin}^{\text{asym}}$.

Algorithm 1 FO Key generation

$(\text{PK}^{\text{asym}}, \text{SK}^{\text{asym}}) \leftarrow \text{PKE.KeyGen}()$
$\text{PK}^{\text{hy}} \leftarrow \text{PK}^{\text{asym}}, \text{SK}^{\text{hy}} \leftarrow \text{SK}^{\text{asym}},$
return $(\text{PK}^{\text{hy}}, \text{SK}^{\text{hy}})$

Algorithm 2 FO Key encryption

Require: $m \in \mathcal{M}^{\text{sym}}$

Sample from the PKE's message space $\sigma \xleftarrow{\$} \mathcal{M}^{\text{asym}}$
$a \leftarrow G(\sigma)$, $c \leftarrow E_a^{\text{sym}}(m)$
$h \leftarrow H(\sigma, c)$
$e \leftarrow E^{\text{asym}}(\text{PK}^{\text{hy}}, \sigma, h)$
return (e, c)

The decryption routine:

Algorithm 3 FO Key decryption

Require: The ciphertext (e, c)

1: $\hat{\sigma} \leftarrow D^{\text{asym}}(\text{SK}^{\text{hy}}, e)$
2: $\hat{h} \leftarrow H(\hat{\sigma}, c)$
3: $\hat{e} \leftarrow E^{\text{asym}}(\text{PK}^{\text{hy}})$
4: **if** $\hat{e} \neq e$ **then**
5: **return** $\perp$
6: **end if**
7: $\hat{a} \leftarrow G(\hat{\sigma})$
8: $\hat{m} \leftarrow D^{\text{sym}}_{\hat{a}}(c)$
9: **return** $\hat{m}$

Theorem 1.1. *For all IND-CCA adversary $\mathcal{A}^{hy}$ against the hybrid encryption scheme with advantage ϵ^{hy}, there exists a one-way one-time-encryption adversary against the public-key encryption scheme with advantage ϵ^{asym} and an one-time indistinguishability adversary against the symmetric encryption scheme with advantage ϵ^{sym} such that*

$$\epsilon^{hy} \leq q_H \epsilon^{asym} + \epsilon^{sym} + q_D 2^{-\gamma}$$

where γ is the spread of the public-key encryption scheme, q_H is the total number of hash queries, and q_D is the total number of decryption queries

1.2 Proof of security result

Theorem 1.1 is proved using a sequence of games that involves $\mathcal{A}^{\text{hy}}$ as the main routine, and two games that involves $\mathcal{A}^{\text{hy}}$ as a sub-routine. The sequence of games is as follows:

1. Game 0 is the standard IND-CCA2 game
2. Game 1 is identical to game 0, except that the decryption oracle $\mathcal{O}_D$ is modified. Instead of using the true secret key SK^{hy} to decrypt the query (e_q, c_q), the decryption oracle checks the tape of hash function H for the existence of hash query (σ_H, c_H, h_H) such that $c_q = c_H$ and $e_q = E^{\text{asym}}(\text{PK}^{\text{asym}}, \sigma_H, h_H)$. If such a query exists, then $\mathcal{O}_D$ uses σ_H to derive the symmetric key $a_q \leftarrow G(\sigma_H)$ and decrypt the queried ciphertext c_q. If no such query exists, then $\mathcal{O}_D$ will reject the queried ciphertext and output a decryption error. It is worth noting that this modified decryption oracle does not require the hybrid secret key SK^{hy} to process decryption queries.
3. Game 2 is identical to game 1, except the routine of encrypting the challenge ciphertext is modified: $a^* \xleftarrow{\$} \mathcal{K}^{\text{sym}}$ is randomly sampled from the symmetric key space instead of being queried from G, and $h^* \xleftarrow{\$} \text{COIN}^{\text{asym}}$ is randomly sampled from the asymmetric coin space instead of being queried from H.

Let S_0, S_1, S_2 denote the event that $\mathcal{A}^{\text{hy}}$ wins game 0, game 1, and game 2, respectively.

Lemma 1.1.1. *Let q_D denote the number of decryption queries, and γ denote the spread of the PKE, then*

$$P[S_0] - P[S_1] \leq q_D 2^{-\gamma}$$

Proof. For each decryption query (e_q, c_q), there are three mutually exclusive possibilities:

1. The queried ciphertext is **honest**, meaning that there is a matching record on the tape of the hash function H
2. There is no matching record on the tape of the hash function H, and the check in step 4 in algorithm 3 will fail, outputing decryption error. Such a query is called **invalid**

3. There is no matching record on the tape of the hash function H, but the check in step 4 of algorithm 3 will succeed. Such a query is called **almost valid**

Observe that for both S_0, S_1:

$$\begin{aligned} P[S] = {} & P[S \cap \text{ all decryption queries are honest }] \\ & + P[S \cap \text{ some decryption queries are dishonest, but none is almost valid }] \\ & + P[S \cap \text{ at least one almost valid decryption queries }] \end{aligned}$$

When all decryption queries are honest, the decryption oracles will correctly decrypt the query in both game 0 and game 1. When all dishonest decryption queries are invalid, the decryption oracles will reject the query in both games. The only difference between the two games lies in how the decryption oracle processes almost valid decryption queries. Therefore:

$$\begin{aligned} & P[S_0] - P[S_1] \\ & = P[S_0 \cap \text{ at least one almost valid decryption queries }] \\ & - P[S_1 \cap \text{ at least one almost valid decryption queries }] \\ & \leq P[\text{ at least one almost valid decryption queries }] \end{aligned}$$

Let (e, c) be some decryption query made without querying H, then in the true decryption routine, $\hat{h} \leftarrow H(\hat{\sigma}, c)$ will be a truly random coin, and $\hat{e} \leftarrow E^{\text{asym}}(\hat{\sigma}, \hat{h})$ will be a truly random ciphertext for the given public key and $\hat{\sigma}$. Since the the PKE has γ spread, we know that $P[e = \hat{e}] = P[(e, c) \text{ is almost valid}] \leq 2^{-\gamma}$. Among q_D decryption query, the probability of having at least one almost valid query is bounded by sum of probability of each decryption query being almost valid: $P[\text{ at least one almost valid query }] \leq q_D 2^{-\gamma}$. $\square$

References

[1] Eiichiro Fujisaki and Tatsuaki Okamoto. Secure integration of asymmetric and symmetric encryption schemes. In *Annual international cryptology conference*, pages 537–554. Springer, 1999.