This is a revised version of the extended abstract which appeared in
The Cryptographers' Track at RSA Conference '2001
(8 – 12 april 2001, San Francisco, California, USA)
D. Naccache Ed., LNCS 2020, pages 159–175, Springer-Verlag, 2001.

REACT: Rapid Enhanced-security Asymmetric Cryptosystem Transform

Tatsuaki Okamoto[1] and David Pointcheval[2]

[1] NTT Labs, 1-1 Hikarinooka, Yokosuka-shi 239-0847 Japan.
E-mail: okamoto@isl.ntt.co.jp.
[2] Dépt d'Informatique, ENS – CNRS, 45 rue d'Ulm, 75230 Paris Cedex 05, France.
E-mail: David.Pointcheval@ens.fr – URL: http://www.di.ens.fr/~pointche.

Abstract. Seven years after the optimal asymmetric encryption padding (OAEP) which makes chosen-ciphertext secure encryption scheme from any trapdoor one-way permutation (but whose unique application is RSA), this paper presents REACT, a new conversion which applies to any weakly secure cryptosystem, in the random oracle model: it is optimal from both the computational and the security points of view. Indeed, the overload is negligible, since it just consists of two more hashings for both encryption and decryption, and the reduction is very tight. Furthermore, advantages of REACT beyond OAEP are numerous:

1. it is more general since it applies to any partially trapdoor one-way function (a.k.a. weakly secure public-key encryption scheme) and therefore provides security relative to RSA but also to the Diffie-Hellman problem or the factorization;
2. it is possible to integrate symmetric encryption (block and stream ciphers) to reach very high speed rates;
3. it provides a key distribution with session key encryption, whose overall scheme achieves chosen-ciphertext security even with weakly secure symmetric scheme.

Therefore, REACT could become a new alternative to OAEP, and even reach security relative to factorization, while allowing symmetric integration.

Keywords: Public-Key Encryption, Semantic Security, Chosen-Ciphertext Attacks, Gap Problems

1 Introduction

For a long time many conversions from a weakly secure encryption scheme into a chosen-ciphertext secure cryptosystem have been attempted, with variable success. Such a goal is of greatest interest since many one-way encryption schemes are known, with variable efficiency and various properties, whereas chosen-ciphertext secure schemes are very rare.

1.1 Chosen-Ciphertext Secure Cryptosystems

Until few years ago, the description of a cryptosystem, together with some heuristic arguments for security, were enough to convince and to make a scheme to be widely adopted. Formal semantic security [18] and further non-malleability [13] were just seen as theoretical properties. However, after multiple cryptanalyses of international standards [7, 10, 9], provable security has been realized to be important and even became a basic requirement for any new cryptographic protocol. Therefore, for the last few years, many cryptosystems have been proposed. Some furthermore introduced new algebraic problems, and assumptions [25, 1, 2, 19, 26, 29, 31, 34], other are intricate constructions, over old schemes, to reach

chosen-ciphertext security (from El Gamal [20, 41, 40, 11], D-RSA [33] or Paillier [32]), with specific security proofs.

Indeed, it is easy to describe a one-way cryptosystem from any trapdoor problem. Furthermore, such a trapdoor problems is not so rare (Diffie-Hellman [12], factorization, RSA [37], elliptic curves [22], McEliece [24], NTRU [19], etc). A very nice result would be a generic and *efficient* conversion from any such a trapdoor problem into a chosen-ciphertext secure encryption scheme.

1.2 Related Work

In 1994, Bellare and Rogaway [5] suggested such a conversion, the so-called OAEP (Optimal Asymmetric Encryption Padding). However, its application domain was restricted to trapdoor one-way *permutations*, which is a very rare object (RSA, with a few variants, is the only one application). Nevertheless, it provided the most efficient RSA-based cryptosystem, the so-called OAEP-RSA, provably chosen-ciphertext secure, and thus became the new RSA standard – PKCS #1 [38], and has been introduced in many world wide used applications.

At PKC '99, Fujisaki and Okamoto [15, 17] proposed another conversion with further important improvements [16, 35]. Therefore it looked like the expected goal was reached: a generic conversion from any one-way cryptosystem into a chosen-ciphertext secure encryption scheme. However, the resulting scheme is not optimal, from the computational point of view. Namely, the decryption phase is more heavy than one could expect, since it requires a re-encryption.

As a consequence, with those conversions, one cannot expect to obtain a scheme with a fast decryption phase (unless both encryption and decryption are very fast, which is very unlikely). Nevertheless, decryption is usually implemented on a smart card. Therefore, cryptosystem with efficient decryption process is a challenge with a quite practical impact.

1.3 Achievement: a New and Efficient Conversion

The present work provides a new conversion in the random oracle model [4] which is optimal from the computational point of view in both the encryption and decryption phases. Indeed, the encryption needs an evaluation of the one-way function, and the decryption just makes one call to the inverting function. Further light computations are to be done, but just an XOR and two hashings. Moreover, many interesting features appear with integration of symmetric encryption schemes.

The way the new conversion works is very natural: it roughly first encrypts a session key using the asymmetric scheme, and then encrypts the plaintext with any symmetric encryption scheme, which is *semantically-secure* under simple passive attacks (possibly the one-time pad), using the session key as secret key. Of course this simple and actually used scheme does not reach chosen-ciphertext security. However, just making the session key more unpredictable and adding a checksum, it can be made so:

$$C = \mathcal{E}^{\mathsf{asym}}_{\mathsf{pk}}(R) \text{ and } c = \mathcal{E}^{\mathsf{sym}}_{K}(m), \text{ where } K = G(R)$$
$$\mathcal{E}_{\mathsf{pk}}(m) = C||c||H(R, m, C, c),$$

where G and H are any hash functions. Therefore, this conversion is not totally new. Moreover, in [4], a similar construction has been suggested, but in the particular setting where $\mathcal{E}^{\mathsf{asym}}$ is a trapdoor permutation (as in OAEP) and the one-time pad for $\mathcal{E}^{\mathsf{sym}}$. Thus, our construction is much more general, and we provide a new security analysis. Moreover, if one uses a semantically secure symmetric encryption scheme against basic passive attacks (no known-plaintext attacks), the last two parts of the ciphertext, which are very fast since they only make calls to a hash function and to a symmetric encryption, can be used more than once, with many messages. This makes a highly secure use of a session key, with symmetric encryption $\mathcal{E}^{\mathsf{sym}}$ which initially just meets a very weak security property:

$$C = \mathcal{E}^{\mathsf{asym}}_{\mathsf{pk}}(R) \text{ and } K = G(R)$$
$$\mathcal{E}_{\mathsf{pk}}(m_i) = C||c_i = \mathcal{E}^{\mathsf{sym}}_{K}(m_i)||H(R, m_i, C, c_i) \text{ for } i = 1, \ldots$$

1.4 Outline of the Paper

We first review, in Section 2, the security notions about encryption schemes (both symmetric and asymmetric) required in the rest of the paper, with namely the semantic security. Then, in the next section (Section 3), we describe a new attack scenario, we call the Plaintext-Checking Attack. It then leads to the introduction of a new class of problems, the so-called Gap-Problems [28]. Then in Section 4, we describe our new conversion together with the security proofs. The next section (Section 5) presents some interesting applications of this conversion. Then comes the conclusion.

2 Security Notions for Encryption Schemes

2.1 Asymmetric Encryption Schemes

In this part, we formally define public-key encryption schemes, together with the security notions.

Definition 1 (Asymmetric Encryption Scheme). An asymmetric encryption scheme on a message-space $\mathcal{M}$ consists of 3 algorithms $(\mathcal{K}^{\mathsf{asym}}, \mathcal{E}^{\mathsf{asym}}, \mathcal{D}^{\mathsf{asym}})$:

- the key generation algorithm $\mathcal{K}^{\mathsf{asym}}(1^k)$ outputs a random pair of secret-public keys $(\mathsf{sk}, \mathsf{pk})$, relatively to the security parameter k;
- the encryption algorithm $\mathcal{E}^{\mathsf{asym}}_{\mathsf{pk}}(m; r)$ outputs a ciphertext c corresponding to the plaintext $m \in \mathcal{M}$ (using the random coins $r \in \Omega$);
- the decryption algorithm $\mathcal{D}^{\mathsf{asym}}_{\mathsf{sk}}(c)$ outputs the plaintext m associated to the ciphertext c.

Remark 2. As written above, $\mathcal{E}^{\mathsf{asym}}_{\mathsf{pk}}(m; r)$ denotes the encryption of a message $m \in \mathcal{M}$ using the random coins $r \in \Omega$. When the random coins are useless in the discussion, we simply note $\mathcal{E}^{\mathsf{asym}}_{\mathsf{pk}}(m)$, as done above in the introduction.

The basic security notion required from an encryption scheme is the *one-wayness*, which roughly means that, from the ciphertext, one cannot recover the whole plaintext.

Definition 3 (One-Way). An asymmetric encryption scheme is said to be *one-way* if no polynomial-time attacker can recover the whole plaintext from a given ciphertext with non-negligible probability. More formally, an asymmetric encryption scheme is said (t, ε)-OW if for any adversary $\mathcal{A}$ with running time bounded by t, its inverting probability is less than ε:

$$\mathsf{Succ}^{\mathsf{ow}}(\mathcal{A}) = \Pr_{\substack{m \overset{R}{\leftarrow} \mathcal{M} \\ r \overset{R}{\leftarrow} \Omega}} [(\mathsf{sk}, \mathsf{pk}) \leftarrow \mathcal{K}^{\mathsf{asym}}(1^k) : \mathcal{A}(\mathcal{E}^{\mathsf{asym}}_{\mathsf{pk}}(m; r)) \overset{?}{=} m] < \varepsilon,$$

where the probability is also taken over the random coins of the adversary.

A by now more and more required property is the *semantic security* [18] also known as *indistinguishability of encryptions* or *polynomial security* since it is the computational version of perfect security [39].

Definition 4 (Semantic Security). An asymmetric encryption scheme is said to be *semantically secure* if no polynomial-time attacker can learn any bit of information about the plaintext from the ciphertext, excepted the length. More formally, an asymmetric encryption scheme is said (t, ε)-IND if for any adversary $\mathcal{A} = (A_1, A_2)$ with running time bounded by t,

$$\mathsf{Adv}^{\mathsf{ind}}(\mathcal{A}) = 2 \times \Pr_{\substack{b \overset{R}{\leftarrow} \{0,1\} \\ r \overset{R}{\leftarrow} \Omega}} \left[\begin{array}{l} (\mathsf{sk}, \mathsf{pk}) \leftarrow \mathcal{K}^{\mathsf{asym}}(1^k), (m_0, m_1, s) \leftarrow A_1(\mathsf{pk}) \\ c \leftarrow \mathcal{E}^{\mathsf{asym}}_{\mathsf{pk}}(m_b; r) : A_2(c, s) \overset{?}{=} b \end{array} \right] - 1 < \varepsilon,$$

where the probability is also taken over the random coins of the adversary, and m_0, m_1 are two identical-length plaintexts chosen by the adversary in the message-space $\mathcal{M}$.

Both notions are denoted OW and IND respectively in the following.

Another security notion has been defined, called *non-malleability* [13]. It roughly means that it is impossible to derive, from a given ciphertext, a new ciphertext such that the plaintexts are meaningfully related. But we won't detail it since this notion has been proven equivalent to semantic security against parallel attacks [6].

Indeed, the adversary considered above may obtain, in some situations, more informations than just the public key. With just the public key, we say that she plays a *chosen–plaintext attack* since she can encrypt any plaintext of her choice, thanks to the public key. It is denoted CPA. But she may have, for some time, access to a decryption oracle. She then plays a *chosen–ciphertext attack*, which is either *non-adaptive* [27] if this access is limited in time, or *adaptive* [36] if this access is unlimited, and the adversary can therefore ask any query of her choice to the decryption oracle, but of course she is restricted not to use it on the challenge ciphertext. It has already been proven [3] that under this latter attack, the adaptive chosen-ciphertext attacks, denoted CCA, the semantic security and the non-malleability notions are equivalent, and this is the strongest security notion that one could expect, in the standard model of communication. We therefore call this security level in this scenario the *chosen–ciphertext security*.

2.2 Symmetric Encryption Schemes

In this part, we briefly focus on symmetric encryption schemes.

Definition 5 (Symmetric Encryption Scheme). A symmetric encryption scheme with a key-length k, on messages of length ℓ, consists of 2 algorithms $(\mathcal{E}^{\mathsf{sym}}, \mathcal{D}^{\mathsf{sym}})$ which depends on the k-bit string k, the secret key:

- the encryption algorithm $\mathcal{E}^{\mathsf{sym}}_{\mathsf{k}}(m)$ outputs a ciphertext c corresponding to the plaintext $m \in \{0,1\}^{\ell}$, in a deterministic way;
- the decryption algorithm $\mathcal{D}^{\mathsf{sym}}_{\mathsf{k}}(c)$ gives back the plaintext m associated to the ciphertext c.

As for asymmetric encryption, impossibility for any adversary to get back the whole plaintext just given the ciphertext is the basic requirement. However, we directly consider *semantic security*.

Definition 6 (Semantic Security). A symmetric encryption scheme is said to be *semantically secure* if no polynomial-time attacker can learn any bit of information about the plaintext from the ciphertext, excepted the length. More formally, a symmetric encryption scheme is said (t,ε)-IND if for any adversary $\mathcal{A} = (A_1, A_2)$ with running time bounded by t, $\mathsf{Adv}^{\mathsf{ind}}(\mathcal{A}) < \varepsilon$, where

$$\mathsf{Adv}^{\mathsf{ind}}(\mathcal{A}) = 2 \times \Pr_{\substack{\mathsf{k} \xleftarrow{R} \{0,1\}^k \\ b \xleftarrow{R} \{0,1\}}} [(m_0, m_1, s) \leftarrow A_1(k), c \leftarrow \mathcal{E}^{\mathsf{sym}}_{\mathsf{k}}(m_b) : A_2(c, s) \stackrel{?}{=} b] - 1,$$

in which the probability is also taken over the random coins of the adversary, and m_0, m_1 are two identical-length plaintexts chosen by the adversary in the message-space $\{0,1\}^{\ell}$.

In the basic scenario, the adversary just sees some ciphertexts, but nothing else. However, many stronger scenarios can also be considered. The first which seemed natural for public-key cryptosystems are the known/chosen-plaintext attacks, where the adversary sees some plaintext-ciphertext pairs with the plaintext possibly chosen by herself. These attacks are not trivial in the symmetric encryption setting, since the adversary is unable to encrypt by herself.

The strongest scenario considers the adaptive chosen-plaintext/ciphertext attacks, where the adversary has access to both an encryption and a decryption oracle, such as in the so-called boomerang attack [42].

However, just the security against the basic no-plaintext/ciphertext attacks (a.k.a. passive attacks) is enough in our application. Therefore, one can remark that it is a very weak requirement. Indeed, if one considers AES candidates, cryptanalysts even fail in breaking efficiently semantic security using adaptive chosen plaintext/ciphertext attacks: with respect to pseudo-random permutations, semantic security is equivalent to say that the family $(\mathcal{E}^{\mathsf{sym}}_{\mathsf{k}})_{\mathsf{k}}$ is (t,ε)-indistinguishable from the uniform distribution on all the possible permutations over the message-space, after just one query to the oracle which is either $\mathcal{E}^{\mathsf{sym}}_{\mathsf{k}}$ for some random k or a random permutation (*cf.* universal hash functions [8])!

Remark 7. One should remark that the one-time pad provides a perfect semantically secure symmetric encryption: for any t it is $(t,0)$-semantically secure, for $\ell = k$.

3 The Plaintext-Checking Attacks

3.1 Definitions

We have recalled above all the classical security notions together with the classical scenarios of attacks in the asymmetric setting. A new kind of attacks (parallel attacks) has been recently defined [6], which have no real practical meaning, but the goal was just to deal with non-malleability. In this paper, we define a new one, where the adversary can check whether a message-ciphertext pair (m, c) is valid: the *Plaintext-Checking Attack*.

Definition 8 (Plaintext-Checking Attack). The attacker has access to a *Plaintext-Checking Oracle* which takes as input a plaintext m and a ciphertext c and outputs 1 or 0 whether c encrypts m or not.

It is clear that such an oracle is less powerful than a decryption oracle. This scenario will be denoted by PCA, and will be always assumed to be fully adaptive: the attacker has always access to this oracle without any restriction (we even allows her to include the challenge ciphertext in the query.) It is a very weak security notion.

Remark 9. One can remark that semantic security under this attack cannot be reached. Thus, we will just consider the *one-wayness* in this scenario. Moreover, for any deterministic asymmetric encryption scheme, the PCA-scenario is equivalent to the CPA-one. Indeed, the Plaintext-Checking oracle does just give an information that one can easily obtain by oneself. Namely, any trapdoor one-way permutation provides a OW-PCA-secure encryption scheme (*eg.* RSA [37]).

3.2 Examples

Let us consider some famous public-key encryption schemes in order to study their OW-PCA-security.

The RSA Cryptosystem. In 1978, Rivest–Shamir–Adleman [37] defined the first asymmetric encryption scheme based on the RSA–assumption. It works as follows:

- The user chooses two large primes p and q and publishes the product $n = pq$ together with any exponent e, relatively prime to $\varphi(n)$. He keeps p and q secret, or the invert exponent $d = e^{-1} \bmod \varphi(n)$.
- To encrypt a message $m \in \mathbb{Z}_n^\star$, one just has to compute $c = m^e \bmod n$.
- The recipient can recover the message thanks to d, $m = c^d \bmod n$.

The *one-wayness* (against CPA) of this scheme relies on the RSA problem. Since this scheme is deterministic, it is still one-way, even against PCA, relative to the RSA problem: the RSA-cryptosystem is OW-PCA relative to the RSA problem.

The El Gamal Cryptosystem. In 1985, El Gamal [14] defined an asymmetric encryption scheme based on the Diffie-Hellman key distribution problem [12]. It works as follows:

- An authority chooses and publishes an Abelian group $\mathcal{G}$ of order q, denoted multiplicatively but it could be an elliptic curve or any Abelian variety, together with a generator g. Each user chooses a secret key x in $\mathbb{Z}_q^\star$ and publishes $y = g^x$.
- To encrypt a message m, one has to choose a random element k in $\mathbb{Z}_q^\star$ and sends the pair $(r = g^k, s = m \times y^k)$ as the ciphertext.
- The recipient can recover the message from a pair (r, s) since $m = s/r^x$, where x is his secret key.

The *one-wayness* of this scheme is well-known to rely on the Computational Diffie-Hellman problem. However, to reach semantic security, this scheme requires m to be encoded into an element in the group $\mathcal{G}$. And then, it is equivalent to the Decision Diffie-Hellman problem, where the Diffie-Hellman problems are defined as follows:

- *The Computational Diffie-Hellman Problem (CDH)*: given a pair (g^a, g^b), find the element $C = g^{ab}$.
- *The Decision Diffie-Hellman Problem (DDH)*: given a triple (g^a, g^b, g^c), decide whether $c = ab \bmod q$ or not.
- *The Gap–Diffie-Hellman Problem (GDH)*: solve the CDH problem with the help of a DDH Oracle (which answers whether a given triple is a Diffie-Hellman triple or not).

Proposition 10. *The El Gamal encryption scheme is OW-PCA relative to the GDH problem.*

Proof. The proof directly comes from the fact that a Plaintext-Checking Oracle, for a given public key $y = g^x$ and a ciphertext $(r = g^k, s = m \times y^k)$, simply checks whether the triple $(y = g^x, r = g^k, s/m)$ is a DH-triple. It is exactly a DDH Oracle. □

Since no polynomial time reduction (even a probabilistic one) is known from the CDH problem to the DDH problem [23], the GDH assumption seems as reasonable as the DDH assumption (the reader is referred to [28] for more details).

4 Description of REACT

4.1 The Basic Conversion

Let us consider $(\mathcal{K}^{\mathsf{asym}}, \mathcal{E}^{\mathsf{asym}}, \mathcal{D}^{\mathsf{asym}})$, any OW-PCA–secure asymmetric encryption scheme, as well as two hash functions G and H which output k_1-bit strings and k_2-bit strings respectively. Then, the new scheme $(\mathcal{K}, \mathcal{E}, \mathcal{D})$ works as follows:

- $\mathcal{K}(1^k)$: it simply runs $\mathcal{K}^{\mathsf{asym}}(1^k)$ to get a pair of keys $(\mathsf{sk}, \mathsf{pk})$, and outputs it.

- $\mathcal{E}_{\mathsf{pk}}(m; R, r)$: for any k_1-bit message m and random values $R \in \mathcal{M}$ and $r \in \Omega$, it gets $c_1 = \mathcal{E}^{\mathsf{asym}}_{\mathsf{pk}}(R; r)$, then it computes the session key $K = G(R)$, $c_2 = K \oplus m$ as well as $c_3 = H(R, m, c_1, c_2)$. The ciphertext consists of the triple $C = (c_1, c_2, c_3)$.
- $\mathcal{D}_{\mathsf{sk}}(c_1, c_2, c_3)$: it first extracts R from c_1 by decrypting it, $R = \mathcal{D}^{\mathsf{asym}}_{\mathsf{sk}}(c_1)$. It verifies whether $R \in \mathcal{M}$. It can therefore recover the session key $K = G(R)$ and $m = K \oplus c_2$ which is returned if and only if $c_3 = H(R, m, c_1, c_2)$ and $R \in \mathcal{M}$. Otherwise, it outputs "Reject".

The overload is minimal. Actually, if we consider the encryption phase, it just adds the computation of two hash values and an XOR. Concerning the decryption phase, which had been made heavy in previous conversions [15, 16, 35] with a re-encryption to check the validity, we also just add the computation of two hash values and an XOR, as in the encryption process. Indeed, to compare with previous conversions, the validity of the ciphertext was checked by a full re-encryption. In our conversion, this validity is simply checked by a hash value.

4.2 The Hybrid Conversion

As it has already been done with some previous encryption schemes [15, 16, 30, 33, 35], the "one-time pad" encryption can be generalized to any symmetric encryption scheme which is not perfectly secure, but semantically secure against passive attacks.

Let us consider two encryption schemes, $(\mathcal{K}^{\mathsf{asym}}, \mathcal{E}^{\mathsf{asym}}, \mathcal{D}^{\mathsf{asym}})$ is a OW-PCA–secure asymmetric scheme and $(\mathcal{E}^{\mathsf{sym}}, \mathcal{D}^{\mathsf{sym}})$ is a IND–secure symmetric scheme on ℓ-bit long messages, which uses k_1-bit long keys, as well as two hash functions G and H which output k_1-bit strings and k_2-bit strings respectively. Then, the hybrid scheme $(\mathcal{K}^{\mathsf{hyb}}, \mathcal{E}^{\mathsf{hyb}}, \mathcal{D}^{\mathsf{hyb}})$ works as follows:

- $\mathcal{K}^{\mathsf{hyb}}(1^k)$: exactly has above, for $\mathcal{K}(1^k)$.
- $\mathcal{E}^{\mathsf{hyb}}_{\mathsf{pk}}(m; R, r)$: for any ℓ-bit message m and random values $R \in \mathcal{M}$ and $r \in \Omega$, it gets $c_1 = \mathcal{E}_{\mathsf{pk}}(R; r)$ and a random session key $K = G(R)$. It computes $c_2 = \mathcal{E}^{\mathsf{sym}}_K(m)$ as well as the checking part $c_3 = H(R, m, c_1, c_2)$. The ciphertext consists of $C = (c_1, c_2, c_3)$.
- $\mathcal{D}^{\mathsf{hyb}}_{\mathsf{sk}}(c_1, c_2, c_3)$: it first extracts R from c_1 by decrypting it, $R = \mathcal{D}^{\mathsf{asym}}_{\mathsf{sk}}(c_1)$. It verifies whether $R \in \mathcal{M}$ or not. It can therefore recover the session key $K = G(R)$ as well as the plaintext $m = \mathcal{D}^{\mathsf{sym}}_K(c_2)$ which is returned if and only if $c_3 = H(R, m, c_1, c_2)$ and $R \in \mathcal{M}$. Otherwise, it outputs "Reject".

The overload is similar to the previous conversion one, but then, the plaintext can be longer. Furthermore, the required property for the symmetric encryption is very weak. Indeed, as it will be seen in the security analysis (see the next section), it is just required for the symmetric encryption scheme to be semantically secure in the basic scenario (no plaintext/ciphertext attacks).

4.3 Chosen-Ciphertext Security

Let us turn to the security analysis. Indeed, if the asymmetric encryption scheme $(\mathcal{K}^{\mathsf{asym}}, \mathcal{E}^{\mathsf{asym}}, \mathcal{D}^{\mathsf{asym}})$ is OW-PCA–secure and the symmetric encryption scheme

$(\mathcal{E}^{\mathsf{sym}}, \mathcal{D}^{\mathsf{sym}})$ is IND-secure, then the conversion $(\mathcal{K}^{\mathsf{hyb}}, \mathcal{E}^{\mathsf{hyb}}, \mathcal{D}^{\mathsf{hyb}})$ is IND-CCA in the random oracle model. More precisely, one can claim the following exact security result.

Theorem 11. *Let us consider a CCA–adversary $\mathcal{A}^{\mathsf{cca}}$ against the "semantic security" of the conversion $(\mathcal{K}^{\mathsf{hyb}}, \mathcal{E}^{\mathsf{hyb}}, \mathcal{D}^{\mathsf{hyb}})$, on ℓ-bit long messages, within a time bounded by t, with advantage ε, after q_D, q_G and q_H queries to the decryption oracle, and the hash functions G and H respectively. Then for any $0 < \nu < \varepsilon$, and*

$$t' \leq t + q_G \Phi + (q_H + q_G) O(1)$$

(Φ is the time complexity of $\mathcal{E}_K^{\mathsf{sym}}$), there either exists

- *an adversary $\mathcal{B}^{\mathsf{pca}}$ against the (t', φ)-OW-PCA-security of the asymmetric encryption scheme $(\mathcal{K}^{\mathsf{asym}}, \mathcal{E}^{\mathsf{asym}}, \mathcal{D}^{\mathsf{asym}})$, after less than $q_G + q_H$ queries to the Plaintext-Checking Oracle, where*

$$\varphi = \varepsilon - \nu - \frac{q_D}{2^{k_2}}.$$

- *or an adversary $\mathcal{B}$ against the (t', ν)-IND–security of the symmetric encryption scheme $(\mathcal{E}^{\mathsf{sym}}, \mathcal{D}^{\mathsf{sym}})$.*

Proof. More than semantically secure against chosen-ciphertext attacks, this converted scheme can be proven "plaintext–aware" [5, 3], which implies chosen-ciphertext security. To prove above Theorem, we first assume that the symmetric encryption scheme $(\mathcal{E}^{\mathsf{sym}}, \mathcal{D}^{\mathsf{sym}})$ is (t', ν)-IND–secure, for some probability $0 < \nu < \varepsilon$.

Semantic Security. The semantic security of this scheme intuitively comes from the fact that for any adversary, in order to have any information about the encrypted message m, she at least has to have asked $(R, \star, c_1, c_2)$ to H (which is called "event 1" and denoted by E_1) or R to G (which is called "event 2" and denoted by E_2). Therefore, for a given $c_1 = \mathcal{E}_{\mathsf{pk}}^{\mathsf{asym}}(R; r)$, R is in the list of the queries asked to G or H. Then, for any candidate R', one asks to the Plaintext Checking Oracle whether c_1 encrypts R' or not. The accepted one is returned as the inversion of $\mathcal{E}_{\mathsf{pk}}^{\mathsf{asym}}$ on the ciphertext c_1, which breaks the OW-PCA.

More precisely, let us consider $\mathcal{A} = (A_1, A_2)$, an adversary against the semantic security of the converted scheme, using an adaptive chosen-ciphertext attack. Within a time bound t, she asks q_D queries to the decryption oracle and q_G and q_H queries to the hash functions G and H respectively, and distinguishes the right plaintext with an advantage greater than ε. Actually, in the random oracle model, because of the randomness of G and H, if neither event 1 nor event 2 happen, she gets $c_2 = \mathcal{E}_K^{\mathsf{sym}}(m_b)$, for a totally random key K. Indeed, to the output (m_0, m_1, s) from A_1, A_2 is given c_1, the challenge ciphertext one wants to completely decrypt under $\mathcal{D}_{\mathsf{sk}}^{\mathsf{asym}}$, $c_2 \leftarrow \mathcal{E}_K^{\mathsf{sym}}(m_b)$ where K is a random k_1-bit string and b a random bit, and c_3 is a random k_2-bit string. During this simulation, the random oracles are furthermore simulated as follows:

- for any new query R' to the oracle G, one first checks whether this R' is the searched R (which should lead to the above random K). For that, one asks to the Plaintext-Checking Oracle to know whether c_1 actually encrypts R'. In this case, above K value is returned. Otherwise, a new random value is sent.
- for any new query (R', m', c_1', c_2') to the oracle H, if $(c_1', c_2', m') = (c_1, c_2, m_b)$, and R' is the searched R, which can be detected thanks to the Plaintext-Checking Oracle, above c_3 is returned. Otherwise, a random value is sent.

Then, she cannot gain any advantage greater than ν, when the running time is bounded by t': $\Pr_b[A_2(\mathcal{E}^{\mathsf{hyb}}_{\mathsf{pk}}(m_b; r), s) = b \,|\, \neg(\mathsf{E}_1 \vee \mathsf{E}_2)] \leq 1/2 + \nu/2$. However, splitting the success probability, according to $(\mathsf{E}_1 \vee \mathsf{E}_2)$, one gets the following

$$\frac{1}{2} + \frac{\varepsilon}{2} \leq \left(\frac{1}{2} + \frac{\nu}{2}\right)(1 - \Pr[\mathsf{E}_1 \vee \mathsf{E}_2]) + 1 \times \Pr[\mathsf{E}_1 \vee \mathsf{E}_2],$$

which leads to

$$\frac{\varepsilon}{2} \leq \frac{\nu}{2} + \left(\frac{1}{2} - \frac{\nu}{2}\right)\Pr[\mathsf{E}_1 \vee \mathsf{E}_2] \leq \frac{\nu}{2} + \frac{1}{2} \times \Pr[\mathsf{E}_1 \vee \mathsf{E}_2].$$

This is equivalent to $\Pr[\mathsf{E}_1 \vee \mathsf{E}_2] \geq \varepsilon - \nu$. If E_1 or E_2 occurred, an R' will be accepted and returned after at most $(q_G + q_H)$ queries to the Plaintext Checking Oracle.

Plaintext–Extractor. Since we are in an adaptive chosen-ciphertext scenario, we have to simulate the decryption oracle, or to provide a plaintext-extractor. When the adversary asks a query (c_1, c_2, c_3), the simulator looks for all the pairs (m, R) in the table of the query/answer's previously got from the hash function H. More precisely, it looks for all the pairs (m, R) such that $R \in \mathcal{M}$ and the query (R, m, c_1, c_2) has been asked to H with answer c_3. For any of theses pairs, it computes $K = G(R)$, using above simulation, and checks whether $c_2 = \mathcal{E}^{\mathsf{sym}}_K(m)$ and asks to the Plaintext-Checking Oracle whether c_1 encrypts the given R (therefore globally at most q_H queries to this oracle, whatever the number of queries to the decryption oracle, since R and c_1 are both included in the H-query). In the positive case, it has found a pair (m, R) such that, $R \in \mathcal{M}$, $K = G(R)$ and for some r', $c_1 = \mathcal{E}^{\mathsf{asym}}_{\mathsf{pk}}(R; r')$, $c_2 = \mathcal{E}^{\mathsf{sym}}_K(m)$ and $c_3 = H(R, m, c_1, c_2)$. The corresponding plaintext is therefore m, exactly as would have done the decryption oracle. Otherwise, it rejects the ciphertext.

Some decryptions may be incorrect, but only rejecting a valid ciphertext: a ciphertext is refused if the query (R, m, c_1, c_2) has not been asked to H. This may just leads to two situations:

- either the c_3 has been obtained from the encryption oracle, which means that it is a part of the challenge ciphertext. Because of R, m, c_1 and c_2 in the quadruple H-input, the decryption oracle query is exactly the challenge ciphertext.
- or the attacker has guessed the right value for $H(R, m, c_1, c_2)$ without having asked for it, but only with probability $1/2^{k_2}$;

Conclusion:

Finally, a (c_1, c_2, c_3) decryption-oracle query is not correctly answered with probability limited by $1/2^{k_2}$. Therefore, using this plaintext-extractor, we obtain,

$$\Pr[(\mathsf{E}_1 \vee \mathsf{E}_2) \wedge \text{ no incorrect decryption}] \geq \varepsilon - \nu - \frac{q_D}{2^{k_2}}$$

in which cases one solves the *one-wayness*, simply using the Plaintext-Checking Oracle to check which element, in the list of queries asked to G and H, is the solution. The decryption simulation will just also require Plaintext-Checking on some (R, c_1) which appeared in the H queries. If one memorizes all the obtained answers from the Plaintext-Checking Oracle, putting a tag to each H-input/output values, less than $q_G + q_H$ queries are asked. The running time of adversary, $\mathcal{B}$ or $\mathcal{B}^{\mathsf{pca}}$, is bounded by the running time of $\mathcal{A}$, q_G executions of $\mathcal{E}_K^{\mathsf{sym}}$, and $(q_G + q_H)O(1)$ queries to (G, H and Plaintext-Checking) oracles. That is, $t' \leq t + q_G\Phi + (q_H + q_G)O(1)$. □

5 Some Examples

We now apply this conversion to some classical encryption schemes which are clearly OW-PCA under well defined assumptions.

5.1 With the RSA Encryption Scheme: REACT–RSA

We refer the reader to the section 3.2 for the description and the notations used for the RSA cryptosystem. Let us consider two hash functions G and H which output k_1-bit strings and k_2-bit strings respectively, and any semantically secure symmetric encryption scheme $(\mathcal{E}^{\mathsf{sym}}, \mathcal{D}^{\mathsf{sym}})$.

- $\mathcal{K}(1^k)$: it chooses two large primes p and q greater than 2^k, computes the product $n = pq$. A key pair is composed by a random exponent e, relatively prime to $\varphi(n)$ and its inverse $d = e^{-1} \bmod \varphi(n)$.
- $\mathcal{E}_{e,n}(m; R)$: with $R \in \mathbb{Z}_n^\star$, it gets $c_1 = R^e \bmod n$, then it computes $K = G(R)$ and $c_2 = \mathcal{E}_K^{\mathsf{sym}}(m)$ as well as $c_3 = H(R, m, c_1, c_2)$. The ciphertext consists of the triple $C = (c_1, c_2, c_3)$.
- $\mathcal{D}_{d,n}(c_1, c_2, c_3)$: it first extracts $R = c_1^d \bmod n$. Then it recovers $K = G(R)$ and $m = \mathcal{D}_K^{\mathsf{sym}}(c_2)$ which is returned if and only if $c_3 = H(R, m, c_1, c_2)$. Otherwise, it outputs "Reject".

Theorem 12. *The REACT–RSA encryption scheme is IND-CCA in the random oracle model, relative to the RSA problem (and the semantic security of the symmetric encryption scheme under the basic passive attack).*

Proof. We have just seen before that the plain-RSA encryption is OW-PCA, relative to the RSA problem, which completes the proof. □

This becomes the *best* alternative to OAEP–RSA [5, 38]. Indeed, if one simply uses the "one-time pad", the ciphertext is a bit longer than in the OAEP situation, but one can also use any semantically secure encryption scheme to provide high-speed rates, which is not possible with OAEP.

5.2 With the El Gamal Encryption Scheme: REACT–El Gamal

We also refer the reader to the section 3.2 for the description and the notations used for the El Gamal cryptosystem. Let us consider two hash functions G and H which output k_1-bit strings and k_2-bit strings respectively, and any semantically secure symmetric encryption scheme $(\mathcal{E}^{\mathsf{sym}}, \mathcal{D}^{\mathsf{sym}})$.

- $\mathcal{K}(1^k)$: it chooses a large prime q, greater than 2^k, a group $\mathcal{G}$ of order q and a generator g of $\mathcal{G}$. A key pair is composed by a random element x in $\mathbb{Z}_q^\star$ and $y = g^x$.
- $\mathcal{E}_y(m; R, r)$: with R a random string, of the same length as the encoding of the $\mathcal{G}$-elements, and $r \in \mathbb{Z}_q$, it gets $c_1 = g^r$ and $c_1' = R \oplus y^r$, then it computes $K = G(R)$ and $c_2 = \mathcal{E}_K^{\mathsf{sym}}(m)$ as well as $c_3 = H(R, m, c_1, c_1', c_2)$. The ciphertext therefore consists of the tuple $C = (c_1, c_1', c_2, c_3)$.
- $\mathcal{D}_x(c_1, c_1', c_2, c_3)$: it first extracts $R = c_1' \oplus c_1^x$. Then it recovers $K = G(R)$ and $m = \mathcal{D}_K^{\mathsf{sym}}(c_2)$ which is returned if and only if $c_3 = H(R, m, c_1, c_1', c_2)$. Otherwise, it outputs "Reject".

Theorem 13. *The REACT–El Gamal encryption scheme is IND-CCA in the random oracle model, relative to the GDH problem (and the semantic security of the symmetric encryption scheme under the basic passive attack).*

Proof. We have seen above that the plain-El Gamal encryption scheme is OW-PCA, relative to the GDH problem [28], which completes the proof. □

5.3 With the Okamoto-Uchiyama Encryption Scheme

Description of the Original Scheme. In 1998, Okamoto–Uchiyama [29] defined an asymmetric encryption scheme based on a trapdoor discrete logarithm. It works as follows:

- Each user chooses two large primes p and q and computes $n = p^2 q$. He also chooses an element $g \in \mathbb{Z}_n^\star$ such that $g_p = g^{p-1} \bmod p^2$ is of order p and computes $h = g^n \bmod n$. The modulus n and the elements g and h are made public while p and q are kept secret.
- To encrypt a message m, smaller than p, one has to choose a random element $r \in \mathbb{Z}_n$ and sends $c = g^m h^r \bmod n$ as the ciphertext.
- From a ciphertext c, the recipient can easily recover the message m since

$$m = L(c_p)/L(g_p) \bmod p,$$

  where $L(x) = (x-1)/p \bmod p$ for any $x = 1 \bmod p$, and $c_p = c^{p-1} \bmod p^2$.

The *semantic security* of this scheme relies on the p-subgroup assumption (a.k.a. p-residuosity or more generally high-residuosity), while the *one-wayness* relies on the factorization of the modulus n. The OW-PCA relies on the gap problem, the Gap–High-Residuosity problem, which consists in factoring an RSA modulus with access to a p-residuosity oracle.

Remark 14. Since the encryption process is public, the bound p is unknown. A public bound has to be defined, for example $n^{1/4}$ which is clearly smaller than p, or 2^k where $2^k < p, q < 2^{k+1}$ (see some remarks in [21] about the EPOC application of this scheme [30].)

The Converted Scheme: REACT–Okamoto-Uchiyama. Let us consider two hash functions G and H which output k_1-bit strings and k_2-bit strings respectively, and any semantically secure symmetric encryption scheme $(\mathcal{E}^{\mathsf{sym}}, \mathcal{D}^{\mathsf{sym}})$.

- $\mathcal{K}(1^k)$: it chooses two large primes p and q greater than 2^k, as well as g as described above. It then computes $n = p^2q$ and $h = g^n \bmod n$.
- $\mathcal{E}_{n,g,h}(m; R, r)$: with $R < 2^k$ and $r \in \mathbb{Z}_n$, it computes $c_1 = g^R h^r \bmod n$, then it gets $K = G(R)$ and $c_2 = \mathcal{E}_K^{\mathsf{sym}}(m)$ as well as $c_3 = H(R, m, c_1, c_2)$. The ciphertext consists of the triple $C = (c_1, c_2, c_3)$.
- $\mathcal{D}_p(c_1, c_2, c_3)$: it first extracts $R = L(c_{1p})/L(g_p)$. Then it recovers $K = G(R)$ and $m = \mathcal{D}_K^{\mathsf{sym}}(c_2)$ which is returned if and only if $R < 2^k$ and $c_3 = H(R, m, c_1, c_2)$. Otherwise, it outputs "Reject".

Theorem 15. *The REACT–Okamoto-Uchiyama cryptosystem is IND-CCA in the random oracle model, relative to the Gap–High-Residuosity problem (and the semantic security of the symmetric encryption scheme under the basic passive attack).*

Proof. We have just seen that the plain-Okamoto-Uchiyama encryption scheme is OW-PCA, relative to the Gap–High-Residuosity problem. □

6 Conclusion

This paper presents REACT, a new conversion which applies to any weakly secure cryptosystem: the overload is as negligible as for OAEP [5], but its application domain is more general. Therefore, REACT provides a very efficient solution to realize a provably secure (in the strongest security sense) asymmetric or hybrid encryption scheme based on any practical asymmetric encryption primitive, in the random oracle model.

Acknowledgements

We thank Markus Jakobsson and Moti Yung for helpful discussions. Thanks also to the anonymous reviewers for their comments.

References

1. M. Abdalla, M. Bellare, and P. Rogaway. DHAES: An Encryption Scheme Based on the Diffie-Hellman Problem. Submission to IEEE P1363a. September 1998.
2. M. Abdalla, M. Bellare, and P. Rogaway. The Oracle Diffie-Hellman Assumptions and an Analysis of DHIES. In *RSA '2001*, LNCS. Springer-Verlag, Berlin, 2001.
3. M. Bellare, A. Desai, D. Pointcheval, and P. Rogaway. Relations among Notions of Security for Public-Key Encryption Schemes. In *Crypto '98*, LNCS 1462, pages 26–45. Springer-Verlag, Berlin, 1998.
4. M. Bellare and P. Rogaway. Random Oracles Are Practical: a Paradigm for Designing Efficient Protocols. In *Proc. of the 1st CCS*, pages 62–73. ACM Press, New York, 1993.
5. M. Bellare and P. Rogaway. Optimal Asymmetric Encryption – How to Encrypt with RSA. In *Eurocrypt '94*, LNCS 950, pages 92–111. Springer-Verlag, Berlin, 1995.
6. M. Bellare and A. Sahai. Non-Malleable Encryption: Equivalence between Two Notions, and an Indistinguishability-Based Characterization. In *Crypto '99*, LNCS 1666, pages 519–536. Springer-Verlag, Berlin, 1999.

7. D. Bleichenbacher. A Chosen Ciphertext Attack against Protocols based on the RSA Encryption Standard PKCS #1. In *Crypto '98*, LNCS 1462, pages 1–12. Springer-Verlag, Berlin, 1998.
8. L. Carter and M. Wegman. Universal Hash Functions. *Journal of Computer and System Sciences*, 18:143–154, 1979.
9. D. Coppersmith, S. Halevi, and C. S. Jutla. ISO 9796 and the New Forgery Strategy. Working Draft presented at the Rump Session of Crypto '99, 1999.
10. J.-S. Coron, D. Naccache, and J. P. Stern. On the Security of RSA Padding. In *Crypto '99*, LNCS 1666, pages 1–18. Springer-Verlag, Berlin, 1999.
11. R. Cramer and V. Shoup. A Practical Public Key Cryptosystem Provably Secure against Adaptive Chosen Ciphertext Attack. In *Crypto '98*, LNCS 1462, pages 13–25. Springer-Verlag, Berlin, 1998.
12. W. Diffie and M. E. Hellman. New Directions in Cryptography. *IEEE Transactions on Information Theory*, IT–22(6):644–654, November 1976.
13. D. Dolev, C. Dwork, and M. Naor. Non-Malleable Cryptography. In *Proc. of the 23rd STOC*. ACM Press, New York, 1991.
14. T. El Gamal. A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms. *IEEE Transactions on Information Theory*, IT–31(4):469–472, July 1985.
15. E. Fujisaki and T. Okamoto. How to Enhance the Security of Public-Key Encryption at Minimum Cost. In *PKC '99*, LNCS 1560, pages 53–68. Springer-Verlag, Berlin, 1999.
16. E. Fujisaki and T. Okamoto. Secure Integration of Asymmetric and Symmetric Encryption Schemes. In *Crypto '99*, LNCS 1666, pages 537–554. Springer-Verlag, Berlin, 1999.
17. E. Fujisaki and T. Okamoto. How to Enhance the Security of Public-Key Encryption at Minimum Cost. *IEICE Transaction of Fundamentals of Electronic Communications and Computer Science*, E83-A(1):24–32, January 2000.
18. S. Goldwasser and S. Micali. Probabilistic Encryption. *Journal of Computer and System Sciences*, 28:270–299, 1984.
19. J. Hoffstein, J. Pipher, and J. H. Silverman. NTRU: A Ring Based Public Key Cryptosystem. In *Algorithmic Number Theory Symposium (ANTS III)*, LNCS 1423, pages 267–288. Springer-Verlag, Berlin, 1998.
20. M. Jakobsson. A Practical Mix. In *Eurocrypt '98*, LNCS 1403, pages 448–461. Springer-Verlag, Berlin, 1998.
21. M. Joye, J. J. Quisquater, and M. Yung. On the Power of Misbehaving Adversaries and Cryptanalysis of EPOC. In *RSA '2001*, LNCS. Springer-Verlag, Berlin, 2001.
22. N. Koblitz. Elliptic Curve Cryptosystems. *Mathematics of Computation*, 48(177):203–209, January 1987.
23. U. M. Maurer and S. Wolf. The Diffie-Hellman Protocol. *Designs, Codes, and Cryptography*, 19:147–171, 2000.
24. R. J. McEliece. A Public-Key Cryptosystem Based on Algebraic Coding Theory. *DSN progress report*, 42-44:114–116, 1978. Jet Propulsion Laboratories, CALTECH.
25. D. Naccache and J. Stern. A New Public-Key Cryptosystem. In *Eurocrypt '97*, LNCS 1233, pages 27–36. Springer-Verlag, Berlin, 1997.
26. D. Naccache and J. Stern. A New Cryptosystem based on Higher Residues. In *Proc. of the 5th CCS*, pages 59–66. ACM Press, New York, 1998.
27. M. Naor and M. Yung. Public-Key Cryptosystems Provably Secure against Chosen Ciphertext Attacks. In *Proc. of the 22nd STOC*, pages 427–437. ACM Press, New York, 1990.
28. T. Okamoto and D. Pointcheval. The Gap-Problems: a New Class of Problems for the Security of Cryptographic Schemes. In *PKC '2001*, LNCS. Springer-Verlag, Berlin, 2001.
29. T. Okamoto and S. Uchiyama. A New Public Key Cryptosystem as Secure as Factoring. In *Eurocrypt '98*, LNCS 1403, pages 308–318. Springer-Verlag, Berlin, 1998.
30. T. Okamoto, S. Uchiyama, and E. Fujisaki. EPOC: Efficient Probabilistic Public-Key Encryption. Submission to IEEE P1363a. November 1998.
31. P. Paillier. Public-Key Cryptosystems Based on Discrete Logarithms Residues. In *Eurocrypt '99*, LNCS 1592, pages 223–238. Springer-Verlag, Berlin, 1999.
32. P. Paillier and D. Pointcheval. Efficient Public-Key Cryptosystems Provably Secure against Active Adversaries. In *Asiacrypt '99*, LNCS 1716, pages 165–179. Springer-Verlag, Berlin, 1999.
33. D. Pointcheval. HD–RSA: Hybrid Dependent RSA – a New Public-Key Encryption Scheme. Submission to IEEE P1363a. October 1999.
34. D. Pointcheval. New Public Key Cryptosystems based on the Dependent-RSA Problems. In *Eurocrypt '99*, LNCS 1592, pages 239–254. Springer-Verlag, Berlin, 1999.
35. D. Pointcheval. Chosen-Ciphertext Security for any One-Way Cryptosystem. In *PKC '2000*, LNCS 1751, pages 129–146. Springer-Verlag, Berlin, 2000.

36. C. Rackoff and D. R. Simon. Non-Interactive Zero-Knowledge Proof of Knowledge and Chosen Ciphertext Attack. In *Crypto '91*, LNCS 576, pages 433–444. Springer-Verlag, Berlin, 1992.
37. R. Rivest, A. Shamir, and L. Adleman. A Method for Obtaining Digital Signatures and Public Key Cryptosystems. *Communications of the ACM*, 21(2):120–126, February 1978.
38. RSA Data Security, Inc. Public Key Cryptography Standards – PKCS.
39. C. E. Shannon. Communication Theory of Secrecy Systems. *Bell System Technical Journal*, 28(4):656–715, 1949.
40. V. Shoup and R. Gennaro. Securing Threshold Cryptosystems against Chosen Ciphertext Attack. In *Eurocrypt '98*, LNCS 1403, pages 1–16. Springer-Verlag, Berlin, 1998.
41. Y. Tsiounis and M. Yung. On the Security of El Gamal based Encryption. In *PKC '98*, LNCS. Springer-Verlag, Berlin, 1998.
42. D. Wagner. The Boomerang Attack. In *Proc. of the 6th FSE*, LNCS 1636. Springer-Verlag, Berlin, 1999.