

Chosen-Ciphertext Security for any One-Way Cryptosystem

David Pointcheval

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. For two years, public key encryption has become an essential topic in cryptography, namely with security against chosen-ciphertext attacks. This paper presents a generic technique to make a highly secure cryptosystem from any partially trapdoor one-way function, in the random oracle model. More concretely, any suitable problem providing a one-way cryptosystem can be efficiently derived into a chosen-ciphertext secure encryption scheme. Indeed, the overhead only consists of two hashing and a XOR. As application, we provide the most efficient El Gamal encryption variant, therefore secure relative to the computational Diffie-Hellman problem. Furthermore, we present the first scheme whose security is relative to the factorization of large integers, with a perfect reduction (factorization is performed within the same time and with identical probability of success as the security break).

Keywords: Public Key Encryption Schemes, Semantic Security, Chosen-Ciphertext Attacks, Partially Trapdoor One-Way Functions

1 Introduction

1.1 Background

In 1976, Diffie and Hellman [13] proposed the concept of public-key cryptography, and namely of public-key encryption using trapdoor one-way functions. However, despite research efforts of many cryptographers, very few practical cryptosystems have been proposed.

In the mean time, a lot of security notions have been proposed, from one-wayness, where the attacker tries to recover the whole plaintext given the ciphertext and only public data, to non-malleability under chosen-ciphertext attacks [14, 32, 6], where the attacker tries to derive a new ciphertext, whose plaintext is meaningfully related to the original one, with access to the decryption algorithm.

More recently, a general study of security notions [2] has been driven, leading to the conclusion that all those notions are equivalent under chosen-ciphertext attacks, which we regroup under the name of “chosen-ciphertext security”. It therefore represents the most powerful of the practical security notions, just under the theoretical notion of plaintext-awareness, only defined in the random oracle model (see [4, 2] for more details). For a long time, one-wayness or even just heuristic security of an encryption scheme have been considered enough. Semantic security [18]/non-malleability [14] were just seen as theoretical properties, for people from the complexity theory. But after the numerous practical attacks, namely the recent Bleichenbacher's [9], Coron-Naccache-Stern's and Coppersmith-Halevi-Jutla's [11, 10] ones, provable security has been realized to be of important interest for many applications. Therefore, chosen-ciphertext security has become a requirement for encryption schemes.

1.2 Related Work

The reason of this small number of candidates as secure cryptosystems relies on the fact that hardly satisfied properties are required. And even with a well-suited problem, a cryptosystem is not easily derived and does not necessarily lead to an efficient resulting scheme.

Before 1994, only theoretical and impractical schemes were proposed [22, 32]. Then Bellare and Rogaway [4] designed a generic padding, called Optimal Asymmetric Encryption Padding, to make a chosen-ciphertext secure cryptosystem from any trapdoor one-way permutation. However, such permutations are very rare: indeed, RSA [33] is the only one (with the recently proposed, at last PKC '99, by Paillier [26], but also based on the RSA assumption). Therefore OAEP–RSA has been a long time the only practical cryptosystem secure against chosen-ciphertext attacks. Then it has been incorporated in SET, the Secure Electronic Transaction system [19] proposed by VISA and Master Card, and has become the new RSA encryption standard PKCS #1 v2.0 [34].

The last few years, many new schemes has been proposed with proven security relative to decisional problems:

- the decisional Diffie-Hellman problem [13]: at PKC '98, Tsionis–Yung [38] proposed the first El Gamal [15] based cryptosystem, but using an unproven assumption about the unforgeability of Schnorr's Signatures [35]. The same year, Shoup–Gennaro [37] and Cramer–Shoup [12] proposed other variants, the latter is even secure in the standard model.
- the decisional dependent–RSA problem: the author of this work [29, 30] proposed schemes based on a new problem, called the dependent–RSA problem. Some variants have also been proven chosen-ciphertext secure relative to RSA, which became the first alternative to the OAEP–RSA.
- the higher residuosity [7, 8]: Paillier–Pointcheval [28] proposed a variant of the Paillier's scheme [27], whose main interest is the efficiency of the decryption process.

However, those schemes came one by one, with new and specific intricate proofs of security for each. Last year, at PKC '99, Fujisaki–Okamoto [16] proposed a second generic transformation. This very simple transformation enhances the security of any public-key encryption scheme in the random oracle model: from a semantically secure scheme (against just chosen-plaintext attacks), it makes a chosen-ciphertext secure scheme. It may be applied to many, more or less recently proposed, schemes also based on above decisional problems [18, 15, 21, 24]. In other words, any trapdoor decisional problem can be derived into chosen-ciphertext secure schemes.

However, decisional problems generally rely on strong assumptions, whereas one-way schemes are based on weaker ones, such as the classical computational problems: RSA [33], Diffie-Hellman [15], factorization [24], classes of residuosity/partial discrete logarithm [27]), etc. Indeed, computational problems are clearly more difficult to solve and are moreover more numerous than decisional ones.

At last crypto, Fujisaki–Okamoto [17] improved their generic transformation, to make security related to the computational problem instead of just the decisional one.

1.3 Outline of the Paper

The present research, independently driven from the Fujisaki–Okamoto [17] work, reaches a similar result but with more efficient reductions: it presents the most efficient and general transformation. Indeed, it allows to make chosen-ciphertext secure schemes from any one-way encryption scheme: for any one-way function, if a trapdoor allows to get back a part of the preimage, one can base a chosen-ciphertext secure encryption scheme on the relying computational problem.

More concretely, from any one-way encryption scheme (which is the weakest requirement one can make about an encryption scheme) and just two more hashing, one can make a highly secure cryptosystem relying only on the same assumption as the one-wayness of the original scheme, which is generally a really difficult computational problem (at least more difficult than just decisional ones).

We then apply this generic transformation to many well-known one-way functions to provide the best schemes of their families: the most efficient scheme based on the computational Diffie–Hellman problem and the first scheme as secure as factorization.

2 Security Notions for Public Key Encryption

2.1 Definitions

The first common security notion that one would like an encryption scheme to satisfy is the *one-wayness*: with just public data, an attacker can not get back the whole plaintext of a given ciphertext. This notion was satisfied by the RSA cryptosystem [33], relative to the RSA assumption, and by the El Gamal encryption scheme [15], relative to the computational Diffie–Hellman problem. Many applications require more from an encryption scheme, namely

- semantic security (a.k.a. *polynomial security/indistinguishability of encryptions* [18]): if the attacker has some information about the plaintext, for example that it is either “yes” or “no” to a crucial query, she should not learn more with the view of the ciphertext. It is computationally impossible to distinguish between two messages which one has been encrypted. This implies encryption schemes to be probabilistic, such as the El Gamal scheme [15], but not like RSA [33].
- non-malleability [14]: for the problem of encrypted bids, an attacker may just want to under-bid a ciphertext of an unknown amount, without learning anything about this amount or her own proposition. This property has been formally defined under the notion of *non-malleability* [14, 6]: from a given ciphertext any attacker can not derive a new ciphertext in such a way that the plaintexts underlying the two ciphertexts are meaningfully related.

On the other hand, an attacker can play many kinds of attacks: she may just have access to public data, and then encrypt any plaintext of her choice (*chosen-plaintext attacks*) or moreover query the decryption algorithm (*adaptively/non-adaptively chosen-ciphertext attacks* [22, 32]).

A general study of these security notions has been recently driven [2], we therefore refer the reader to this paper for more details, concluding with a complete hierarchy. More precisely, semantic security and non-malleability are equivalent in the adaptively chosen-ciphertext scenario, which defines the strongest practical security notion. In what follows, we call it “*chosen-ciphertext security*”.

2.2 Model of Security

For the last few years, after the numerous attacks against unprovable schemes, provable security has been realized to be of greatest interest. Such a proof is led in the complexity theory setting: one tries to polynomially reduce a well-established difficult problem to an attack. Therefore, an efficient attacker would help to solve the difficult problem: this leads to a contradiction.

Very few schemes have been proven using only such polynomial reductions, without any other assumption. Furthermore, they hardly reach efficiency. The last years, the so-called “random oracle model” [3] has boosted researches, providing an interesting tool in proving the security of very efficient schemes. Indeed, this model, where some concrete cryptographic objects are idealized, namely the hash functions which are assumed to be really random ones, helped to provide security proofs for many encryption schemes [3, 4, 37, 29, 25, 16, 17, 28] and digital signature schemes [5, 23, 31].

3 The New Construction

Our new construction can be applied from any partially trapdoor one-way injective function, and not just from fully trapdoor one-way permutations [4] or already semantically secure encryption schemes [16]. This long sentence “partially trapdoor one-way injective function” is nothing else than a classical encryption scheme which is secure in the weakest sense, just one-way against chosen-plaintext attacks.

Then, in order to formalize notations, we first define this notion of partially trapdoor one-way functions, which informally characterizes one-way functions for which a trapdoor allows a partial recovery of a preimage.

3.1 Partially Trapdoor One-Way Function

Let us consider any function f , from the product space $\mathcal{X} \times \mathcal{Y}$ into \mathcal{Z} .

Definition 1 (One-Way). The function f is said to be *one-way* if, for any given $z = f(x, y)$, it is computationally impossible to get back the couple (x, y) . More formally, for any polynomial time adversary \mathcal{A} , its success $\text{Succ}_{\mathcal{A}}$, defined by $\text{Succ}_{\mathcal{A}} = \Pr_{x,y}[f(\mathcal{A}(f(x, y))) = f(x, y)]$, is negligible.

Whereas we will use the above denomination in the rest of the paper, our result will deal with an even weaker notion of one-wayness, where an adversary should not only have to invert with non-negligible probability of success, but she should also have to rarely output wrong solutions. Of course, she is allowed to output “Reject” when she can not solve the problem.

Definition 2 (Weakly One-Way). The function f is said to be *weakly one-way* if, it is either *one-way* or, for any polynomial time adversary \mathcal{A} , its error probability, defined by

$$\text{Err}_{\mathcal{A}} = \Pr_{x,y}[f(u, v) \neq f(x, y) \mid \mathcal{A}(f(x, y)) \neq \text{“Reject”} \wedge (u, v) = \mathcal{A}(f(x, y))],$$

is non-negligible.

It is a weaker assumption about a problem. Indeed, an adversary may be able to return a correct answer half the time, and therefore break “one-wayness”. But the other half of answers can be junk, which can not be detected if the decisional problem is also difficult (such as the Diffie-Hellman problem [13], the Residuosity problem [7, 8, 24, 21, 27], etc).

Definition 3 (Trapdoor). A (*weakly*) *one-way* function is said to be *trapdoor*, if for some extra information (the trapdoor), for any given $z \in f(\mathcal{X} \times \mathcal{Y})$, it is easily possible to get back a couple (x, y) such that $f(x, y) = z$. Whereas it was computationally impossible without the trapdoor.

As already remarked, such functions which also need to be injective for cryptographic use are very rare (just RSA, and some few related ones), but they are required to apply OAEP [4]. This relativizes the practical impact of the OAEP-transformation.

However, for encryption purpose, it is not required to get back the whole preimage of z , it is the reason why we define the new *partially trapdoor one-way* notion which is more common (see El Gamal [15], or more recently Okamoto–Uchiyama [24], Naccache–Stern [21] and Paillier [27]).

Definition 4 (Partially Trapdoor One-Way). The function f is said to be *partially trapdoor one-way* if,

- for any given $z = f(x, y)$, it is computationally impossible to get back an available x . Such an x is called a *partial preimage* of z .

More formally, for any polynomial time adversary \mathcal{A} , its success, defined by $\text{Succ}_{\mathcal{A}} = \Pr_{x,y}[\exists y', f(x', y') = f(x, y) \mid x' = \mathcal{A}(f(x, y))]$, is negligible. It is *one-way* even for just finding partial-preimage, thus *partial one-wayness*.

- for some extra information, for any given $z \in f(\mathcal{X} \times \mathcal{Y})$, it is easily possible to get back an x , such that there exists a y which satisfies $f(x, y) = z$. The trapdoor does not allow a total inversion, but just a partial one, it is thus called a *partial trapdoor*.

Definition 5 (Partially Trapdoor Weakly One-Way). The function f is said to be *partially trapdoor weakly one-way* if it is *partially trapdoor one-way*

but furthermore, for any polynomial time adversary \mathcal{A} , either its success $\text{Succ}_{\mathcal{A}}$ is negligible or its error probability, defined by

$$\text{Err}_{\mathcal{A}} = \Pr_{x,y}[\forall y', f(x', y') \neq f(x, y) \mid \mathcal{A}(f(x, y)) \neq \text{“Reject”} \wedge (x', y') = \mathcal{A}(f(x, y))],$$

is non-negligible. It is *weakly one-way* even for just finding partial-preimage, thus *partial one-wayness*.

Such partially trapdoor (weakly) one-way functions, moreover *injective*, are of common use in many cryptosystems, however they usually only provide the “one-wayness” of the encryption scheme, which is a very weak property for a cryptosystem. Whereas even semantic security against chosen-plaintext attacks relies on much stronger assumptions or is simply not provable/achieved. Let us briefly recall some of them.

3.2 Some Partially Trapdoor One-Way Injective Functions

The Diffie-Hellman Problem. The most popular encryption scheme based on a partially trapdoor one-way function is the El Gamal cryptosystem [15] based on the Diffie-Hellman distribution key problem [13].

- The Computational Diffie-Hellman Problem: given g , g^a and g^b in a group G , compute g^{ab} .
- The Decisional Diffie-Hellman Problem: given g , g^a , g^b and g^c in a group G , decide whether $g^c = g^{ab}$ or not.
- The El Gamal encryption scheme: given a message m (theoretically in G), a ciphertext is a pair $(g^a, y^a \cdot m)$, where $y = g^b$ is the public key, while b is kept secret.

Computing g^{ab} just given g^a and g^b is assumed impossible (Computational Diffie-Hellman Problem), whereas given b , it only consists of an exponentiation. Then the partially trapdoor one-way function is the following, where $y = g^b$, and q the order of g :

$$\begin{array}{ll} f : G \times \mathbb{Z}_q \longrightarrow G \times G & g : G \times G \longrightarrow G \\ (m, a) \longmapsto (g^a, y^a \cdot m) & (x, z) \longmapsto m = z/x^b \end{array}$$

The one-wayness of the El Gamal encryption scheme clearly relies on the *partial* one-wayness of this function, without the knowledge of b : the Computational Diffie-Hellman Problem, which is almost as difficult as the discrete logarithm problem [20]. However, semantic security requires a much stronger assumption, the Decisional Diffie-Hellman one.

The Partial Discrete Logarithm Problem. More recently, at Crypto '98, Okamoto–Uchiyama [24], at ACM CCS '98, Naccache–Stern [21] and, at Eurocrypt '99, Paillier [27] proposed new encryption schemes based on trapdoor discrete logarithms. More precisely, a trapdoor (the factorization of the composite

modulus) allows to partially compute discrete logarithms. The encryption process puts the message to be encrypted in this recoverable part. The one-wayness of those schemes relies on the factorization, the higher residues and the partial discrete logarithms respectively. However, the semantic security (even against chosen-plaintext attacks) relies on higher residues [7, 8], a weaker problem than factorization and even RSA [27].

The aim of this work is to provide a generic transformation to make any encryption scheme, whose one-wayness is provable, semantically secure even against adaptively chosen-ciphertext attacks, adding just the random oracle assumption.

3.3 Generic Construction

Let us consider such a partially trapdoor one-way injective function f , from the product space $\mathcal{X} \times \mathcal{Y}$ into \mathcal{Z} , and we denote by g its partial invert:

$$\begin{aligned} f : \mathcal{X} \times \mathcal{Y} &\longrightarrow \mathcal{Z} & g : \mathcal{Z} &\longrightarrow \mathcal{X} \\ (x, y) &\longmapsto z & z &\longmapsto x \quad \text{s.t. } \exists y \in \mathcal{Y}, z = f(x, y) \end{aligned}$$

We furthermore need two functions, a hash function H and a generator function G , both assumed to be ideal random functions [3], where k is a security parameter:

$$H : \{0, 1\}^k \rightarrow \mathcal{Y} \quad G : \mathcal{X} \rightarrow \{0, 1\}^k.$$

The names, “hash” and “generator” functions, come from the fact that, in practice, \mathcal{X} and \mathcal{Y} will be of similar size, but maybe smaller than $\{0, 1\}^k$. The cryptosystem is designed in Figure 1, with $k = k_0 + k_1$, where k_0 and k_1 denote the lengths of the messages to be encrypted and the error-parameter respectively. Moreover, $[M]_{k_0}$ denotes the truncation of the bit-string M to its k_0 left bits.

Encryption of $m \in \mathcal{M} = \{0, 1\}^{k_0} \rightarrow (a, b)$
$r \in \mathcal{X}$ and $s \in \{0, 1\}^{k_1}$ are randomly chosen $a = f(r, H(m\ s))$ $b = (m\ s) \oplus G(r)$ $\longrightarrow (a, b)$ is the ciphertext
Decryption of (a, b)
Given $a \in \mathcal{Z}$ and $b \in \{0, 1\}^k$ $r = g(a)$ $M = b \oplus G(r)$ if $a = f(r, H(M))$ $\longrightarrow m = [M]_{k_0}$ is the plaintext otherwise, “Reject: invalid ciphertext”

Fig. 1. Our Generic Construction \mathcal{Enc}_f

Concerning this scheme, called \mathcal{Enc}_f , we show that, under some assumptions about the function f , an attacker against semantic security under an adaptively

chosen-ciphertext attack can be used to efficiently simulate g , and thus partially invert the one-way function f without the trapdoor information, which is computationally impossible under the *partially trapdoor one-way assumption* for the function f .

In what follows, X and Y denote the sizes of \mathcal{X} and \mathcal{Y} respectively, whereas q_G , q_H and q_D denote the numbers of queries asked to the random oracles G and H and to the decryption oracle D , respectively.

Lemma 6. *Let us consider an attacker \mathcal{A} against the semantic security of Enc_f in a chosen-plaintext scenario. If we denote by ε the advantage of this attacker, one can design an algorithm \mathcal{B} that outputs, for any given z , a set \mathcal{S} of values such that a partial preimage of z is in \mathcal{S} with probability greater than $\varepsilon - q_H/2^{k_1}$.*

Proof. Let us consider an adversary $\mathcal{A} = (A_1, A_2)$ against the semantic security of this scheme, where A_1 denotes the “find”-stage and A_2 the “guess”-stage. We then use this adversary to construct a machine \mathcal{B} able to output candidates as partial preimages of f . Let z be a given value in \mathcal{Z} for which we want to find the partial preimage in \mathcal{X} . Our machine \mathcal{B} works as follows:

- It first runs the attacker A_1 where any query to the oracles G and H are intercepted. For any new query q asked to the oracle H , \mathcal{B} chooses a random H_q in \mathcal{Y} and outputs it as the value $H(q)$. The same way, for any new query q asked to the oracle G , \mathcal{B} chooses a random G_q in $\{0, 1\}^k$ and outputs it as the value $G(q)$. The attacker A_1 finally outputs two messages m_0 and m_1 . Our machine \mathcal{B} chooses a random bit c and a random string $b \in_R \{0, 1\}^k$, then it defines $a = z$ and outputs (a, b) as an encryption of m_c .
- The attacker A_2 is fed with this ciphertext (a, b) , and queries to oracles are again intercepted and answered as above.
- When the attacker returns its answer d , our machine \mathcal{B} returns the set \mathcal{S} of all the queries asked to the oracle G during the whole attack.

Now, let us assume that $z = f(x, y)$ for some (x, y) . Because of injectivity, if such a pair exists, it is unique. We consider the game presented in Figure 2, where some values of the oracle are defined, if they have not already been. We define the following events:

- AskG, the query x is asked to G ;
- AskH, a query $m||s$ is asked to H , for some message $m \in \mathcal{M}$, but the specific value s chosen at the beginning of the game.

We say that the attacker wins the game if some of both events occur or if, at the end, the value d returned by A_2 is equal to c . Then the advantage of the attacker is defined by $\text{Adv} = 2 \Pr[\text{wins}] - 1$.

With a random simulation of G and H , as described above, it is clear that this game perfectly simulates the real life excepted the unlikely case where x or $m_c||s$ have already been asked to G or H respectively during the find stage (before their assignment). But this case makes the attacker to win in our game, then $\text{Adv} \geq \text{Adv}_{\mathcal{A}} = \varepsilon$. However, since no advantage can be gained by the adversary without AskG nor AskH, by splitting the game in two cases, depending on both

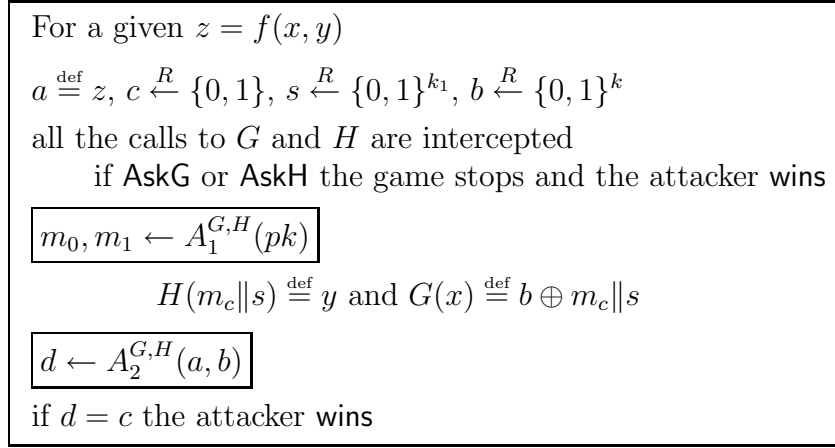


Fig. 2. The Game

events AskG and AskH, one obtains that $\Pr[\text{wins}] \leq \frac{1}{2} \times (1 + \Pr[\text{AskG} \vee \text{AskH}])$. Finally, this leads to $\Pr[\text{AskG} \vee \text{AskH}] \geq \varepsilon$.

Another remark that one can do is that AskH is very unlikely without AskG since it is the only way to gain information about s . More precisely,

$$\Pr[\text{AskH} \mid \neg \text{AskG}] \leq \frac{q_H}{2^{k_1}}.$$

Finally, one can conclude that $\varepsilon \leq \Pr[\text{AskG}] + \Pr[\text{AskH} \mid \neg \text{AskG}]$, and therefore

$$\Pr[\text{AskG}] \geq \varepsilon - \frac{q_H}{2^{k_1}}.$$

This means that with probability greater than $\varepsilon - q_H/2^{k_1}$, x lies in the set \mathcal{S} of queries asked to the oracle G . \square

Thanks to an easy simulation (a plaintext-extractor [4]), one can state the following result.

Theorem 7. *Let us consider an attacker \mathcal{A} against the semantic security of Enc_f in a chosen-ciphertext scenario. If we denote by ε the advantage of this attacker, one can design an algorithm \mathcal{B} that outputs, for any given z , a set \mathcal{S} of values such that a partial preimage of z is in \mathcal{S} with probability greater than*

$$\varepsilon - \frac{(q_D + 1)q_H}{2^{k_1}} - \frac{q_D}{Y}.$$

Proof. Since we have the semantic security against chosen-plaintext attacks, we just have to provide a plaintext-extractor [4], to prove the plaintext-awareness of this scheme which implies security against chosen-ciphertext attacks [2]. The plaintext-extractor is a simulator of the decryption oracle. For a given ciphertext (a, b) , it works as follows: the simulator \mathcal{S} considers all the query-answer pairs (r, G_r) obtained from G , and computes for each $M = b \oplus G_r$. If for some M , $f(r, H(M))$ is equal to a , which can hold for one pair at most, because of the injectivity of f , $m = [M]_{k_0}$ is returned. Otherwise, the ciphertext is considered

as an invalid one, and therefore rejected. Remark that to obtain the above value $H(M)$, one uses the previously described simulation of H .

With this decryption simulation, it is clear that only valid ciphertexts will be decrypted. But will all valid ciphertexts be decrypted? Definitely not, since a valid ciphertext can be produced without asking $G(r)$. But since f is an injection, at most one value for $H(m||s)$ can be accepted: y , if $a = f(r, y)$.

$$\begin{aligned} \Pr[\text{valid} \mid \neg\text{AskG}] &= \Pr[\text{valid} \wedge \text{AskH} \mid \neg\text{AskG}] + \Pr[\text{valid} \wedge \neg\text{AskH} \mid \neg\text{AskG}] \\ &\leq \Pr[\text{AskH} \mid \neg\text{AskG}] + \Pr[\text{valid} \mid \neg\text{AskH} \wedge \neg\text{AskG}] \\ &\leq \frac{q_H}{2^{k_1}} + \frac{1}{Y}. \end{aligned}$$

Finally, the probability of wrong decryption (rejection of valid ciphertext) is upper-bounded by $1/Y + q_H/2^{k_1}$. Therefore, the probability to get no wrong decryption during the attack is lower-bounded by

$$\left(1 - \frac{1}{Y} - \frac{q_H}{2^{k_1}}\right)^{q_D} \geq 1 - \frac{q_D}{Y} - \frac{q_H q_D}{2^{k_1}},$$

which concludes the proof. \square

Remark 8. As one can remark, if \mathcal{Y} is too small, which can even be empty in the case of a fully trapdoor function, and therefore Y not exponentially large, above result is meaningless. However, one can easily extend \mathcal{Y} : let f be a partially trapdoor one-way function $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}$, for any ℓ one defines $\mathcal{Y}_\ell = \mathcal{Y} \times \{0, 1\}^\ell$ and $\mathcal{Z}_\ell = \mathcal{Z} \times \{0, 1\}^\ell$ as well as

$$\begin{aligned} f_\ell : \mathcal{X} \times \mathcal{Y}_\ell &\longrightarrow \mathcal{Z}_\ell \\ (x, (y||r)) &\longmapsto f(x, y)||r \end{aligned}$$

Then, with no computational extra cost, one exponentially increases the size of the set \mathcal{Y} : $Y_\ell = Y \cdot 2^\ell$. However, in many cases, such extensions are not required.

Depending on the kind of problem, even the **weak one-wayness** can be really broken with various efficiency. However, with no particular extra property, randomly choosing x in the set \mathcal{S} , one can just break **one-wayness**, and then state the ‘‘General Case Theorem’’.

Theorem 9 (The General Case). *Let us consider an attacker \mathcal{A} against the semantic security of Enc_f in a chosen-ciphertext scenario. If we denote by ε the advantage of this attacker, one can design an algorithm \mathcal{B} that returns, for any given z , a candidate as partial preimage of z by f which is correct with probability greater than*

$$\frac{1}{q_G} \times \left(\varepsilon - \frac{q_H(q_D + 1)}{2^{k_1}} - \frac{q_D}{Y} \right).$$

However, mathematical problems used in cryptography very often also satisfy a *strong* random self-reducible property (RSA, Diffie-Hellman, etc): from any instance can be derived a random instance whose solution easily provides a solution to the initial instance, using a *strong ring*-homomorphic reduction, where

a *strong ring* is a ring whose cardinality only possesses large prime factors. This is usually the case, since the ring is generally, either a large prime field (Diffie-Hellman problem) or a \mathbb{Z}_n -ring, where n is an RSA-modulus ($n = pq$, RSA, residuosity, partial discrete logarithm [27]) or at least difficult to factor ($n = p^2q$ [24]).

Such problems, as any random self-reducible problem, have the particularity to be uniformly difficult (or easy), there is no worst case nor best case but just average ones. For a *strong* random self-reducible problem, one can then state an improved result. It is derived from a generalization of the Shoup's theorem [36] about the faulty Diffie-Hellman oracles.

Lemma 10. *Let us consider a (k, δ) -oracle \mathcal{A} for a strong random self-reducible problem, which returns a list of k candidates that actually contains the solution with probability greater than $\delta > 7/8$. We can construct a probabilistic algorithm that breaks the **weak one-wayness** of the problem with the following properties. For given $\ell \in \mathbb{N}$, the algorithm makes 24ℓ queries to the oracle \mathcal{A} and performs $\mathcal{O}(\ell \cdot k)$ self-reductions. For all inputs, the output is correct with probability at least $1 - 2^{-\ell}$.*

Proof. With a strong random self-reducible partially one-way problem denoted by $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}$, the structure $(\mathcal{X}, +, *)$ is a strong ring, of size X , and there exists a function $R : \mathcal{Z} \times \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{Z}$, which is invertible for any fixed second and third parameters. The function R maps any instance in \mathcal{Z} into a random one, using random elements r_+, r_* in \mathcal{X} , in such a way that the solution of the resulting instance x' is related with the solution of the initial instance x by $x' = x * r_* + r_+$. We assume that p , the smallest prime factor of X , is large enough and namely that $k^2 < p/8$.

For a given instance $z \in \mathcal{Z}$, one runs twice the oracle \mathcal{A} : once to find candidates, and a second time to check which one is correct. First, with input z , whose solution is x , it gets (x_1, \dots, x_k) . Then, with input $z' = R(z, r_1, r_2)$, for some random r_1, r_2 , whose solution is $x' = x * r_1 + r_2$, it gets (x'_1, \dots, x'_k) . If a right solution is found in both lists, then $x = x_i$ and $x * r_1 + r_2 = x'_j$ for some (i, j) . Therefore, for some pair (i, j) ,

$$x'_j = x_i * r_1 + r_2. \quad (1)$$

Now, let us assume that, $x_i \neq x$, then the probability that above equality (1) holds is at most the conditional probability that for random elements r_1 and r_2 , $x'_j = x_i * r_1 + r_2$, given $x' = x * r_1 + r_2$. This is equal to the probability that for fixed x' and random r , $x'_j = (x_i - x) * r + x'$. Thanks to Shoup's Lemma 1, in [36], one knows that this latter is at most $1/p$.

Then, our algorithm either outputs x_i , if exactly one pair (i, j) satisfies equality (1), or reports failure. And therefore, three exclusive events may happen: (F) failure, (I) incorrect output, (C) correct answer. $\Pr[F] + \Pr[I]$ is upper bounded by the probability that one of the lists does not contain the correct output or that an extraneous relation (1) holds. This occurs with probability bounded by $1/8 + 1/8 + k^2/p \leq 3/8$. However, incorrect output can just occur

when at least one of lists does not contain the correct output: $\Pr[I] \leq 1/8 + 1/8 = 1/4$. Then it follows that $\Pr[C] \geq 1 - 3/8 > 5/8$. Therefore

$$(\Pr[C] + \Pr[I]) / \Pr[I] = 1 + \Pr[C] / \Pr[I] \geq 1 + (5/8) / (1/4) = 7/2.$$

Finally, we obtain that $\Pr[F] \leq 3/8$ and $\Pr[C \mid \neg F] \geq 5/7$.

Now, let us run this algorithm 12ℓ times, on randomly self-reduced instances, and output the majority of the non-failure answers. On average, we get more than $u = 7.5\ell$ answers. The probability of error is upper-bounded by the probability to get more than $v = u/2$ incorrect answers among the u ones:

$$\begin{aligned} \Pr[\text{error}] &\leq \sum_{r=v+1}^u \binom{u}{r} \left(\frac{2}{7}\right)^r \left(\frac{5}{7}\right)^{u-r} = \left(\frac{5}{7}\right)^u \times \sum_{r=v+1}^u \binom{u}{r} \left(\frac{2}{5}\right)^r \\ &\leq \left(\frac{25}{49}\right)^v \left(\frac{2}{5}\right)^{v+1} \times \sum_{r=0}^v \binom{u}{r+v+1} \left(\frac{2}{5}\right)^r \\ &\leq \frac{2}{5} \times \left(\frac{10}{49}\right)^v \times \sum_{r=0}^v \binom{u}{r+v+1} \leq \frac{2}{5} \times \left(\frac{10}{49}\right)^v \times \frac{2^v}{2} \leq \frac{2}{5} \times \left(\frac{40}{49}\right)^v \end{aligned}$$

Finally, this probability is upper-bounded by $(1/2)^\ell$, since $v = 15\ell/4$. \square

Theorem 11 (The Strong Random Self-Reducible Problem Case). *Let us consider an attacker \mathcal{A} against the semantic security of Enc_f in a chosen-ciphertext scenario running within a time bound T . If we denote by ε the advantage of this attacker, one can design an algorithm \mathcal{B} that returns, for any given z , a partial preimage of z by f in an expected time bounded by $21\ell T/\delta$, with a negligible probability of error upper-bounded by $2^{-\ell}$, for any parameter ℓ , where*

$$\delta = \left(\varepsilon - \frac{q_H(q_D + 1)}{2^{k_1}} - \frac{q_D}{Y} \right) \approx \varepsilon.$$

Proof. It is an easy corollary of above lemma. Indeed, if one runs $7/8\delta$ times the general reduction (with randomly self-reduced instances), collecting all the output sets, the global set contains the correct solution with probability greater than $7/8$. \square

Another situation may exist where the verification of the rightness of the candidate is easy. In this case, the efficiency of the reduction is much better. Indeed, from the list of candidates, one has just to check if one of them is the solution.

Theorem 12 (The Easy Verifiable Case). *Let us consider an attacker \mathcal{A} against the semantic security of Enc_f in a chosen-ciphertext scenario. If we denote by ε the advantage of this attacker, one can design an algorithm \mathcal{B} which runs within almost the same time and outputs a partial preimage by f of any given z with probability greater than $\varepsilon - (q_H(q_D + 1))/2^{k_1} - q_D/Y$.*

4 Applications

Let us apply this result to some encryption schemes to make provide semantic security, even against adaptively chosen-ciphertext attacks in the random oracle model, without any more assumption than the one-wayness of the original encryption scheme.

4.1 The El Gamal Encryption Scheme

If one applies our transformation to the famous El Gamal encryption scheme [15], which means to the Diffie-Hellman problem, one gets the scheme presented in Figure 3, together with the following security properties. As previously seen, the

$\mathcal{G} = \langle g \rangle$ of order q $H : \{0, 1\}^k \rightarrow \mathbb{Z}_q$ and $G : \mathcal{G} \rightarrow \{0, 1\}^k$
Secret Key: $x \in \mathbb{Z}_q$ Public Key: $y = g^x$
Encryption
$r \in_R \mathcal{G}$ and $s \in_R \{0, 1\}^{k_1}$ $d = H(m s)$ $\mathcal{Enc}(m, r s) = \begin{cases} a = g^d \\ b = y^d \cdot r \\ c = (m s) \oplus G(r) \end{cases}$
Decryption
$\mathcal{Dec}(a, b, c) = \begin{cases} r = b/a^x & t = c \oplus G(r) \\ \text{if } a = g^{H(t)} & \text{then } m = [t]_{k_0} \end{cases}$

Fig. 3. The DH-based Encryption Scheme

partially trapdoor one-way injection is known as relying on the Computational Diffie-Hellman Problem:

$$\begin{array}{ll}
 f : \mathcal{G} \times \mathbb{Z}_q \longrightarrow \mathcal{G} \times \mathcal{G} & g : \mathcal{G} \times \mathcal{G} \longrightarrow \mathcal{G} \\
 (m, a) \longmapsto (g^a, y^a \cdot m) & (x, z) \longmapsto m = z/x^b
 \end{array}$$

Furthermore, it is well-known to be random self-reducible, even in our strong sense: for a given $(\alpha, \beta) = f(m, a)$, for random $u, v, w \in \mathbb{Z}_q$, $(\alpha^u g^v, \beta^u y^v g^w) = f(m^u g^w, au + v)$, with $(m^u g^w, au + v)$ uniformly distributed in $\mathcal{G} \times \mathbb{Z}_q$.

One has just to remark that, during the decryption phase, if $a = g^d \bmod p$, then $b = a^x r = y^d r \bmod p$ holds.

Theorem 13 (The DH-based Encryption Scheme). *Any algorithm \mathcal{A} able to break the semantic security of the DH-based Encryption Scheme under adaptively chosen-ciphertext attacks within time T can be used as a subroutine to an algorithm \mathcal{B} that breaks the Computational Diffie-Hellman problem in an expected time bounded by $30T\ell/\varepsilon$, with a negligible probability of error bounded by $2^{-\ell}$, for any parameter ℓ , where*

$$\varepsilon = \left(\text{Adv}_{\mathcal{A}} - \frac{(q_D + 1)q_H}{2^{k_1}} - \frac{q_D}{q} \right) \approx \text{Adv}_{\mathcal{A}}.$$

Advantages of the DH-based Encryption Scheme considered as an El Gamal variant. At PKC '98, Tsionis and Yung [38] studied El Gamal based encryption schemes. They were the first to propose a variant secure against chosen-ciphertext attacks, in the random oracle model. However, it was also based on

both the Decisional Diffie-Hellman problem and an unproven assumption about the unforgeability of Schnorr signatures [35]. Furthermore, for weaker schemes, it required more computations: three exponentiations instead of only two for both encryption and decryption in ours.

Later in the same year, Shoup and Gennaro [37] proposed a new variant provably secure against chosen-ciphertext attacks in the random oracle model, under the sole assumption of the Decisional Diffie-Hellman problem. Once again, efficiency was a serious backward: encryption required five exponentiations instead of two for ours, and decryption required seven exponentiations instead of two! However, it was the first convincing El Gamal variant.

Finally, one could consider the Fujisaki-Okamoto variant [16], with a similar efficiency in the random oracle model, or the Cramer-Shoup variant [12], twice slower but, for the first time, proven in the standard model. However, in both cases, security is relative to the Decisional Diffie-Hellman problem, a much stronger assumption than the Computational one.

Consequently, this El Gamal variant is the most efficient from our knowledge (only two exponentiations per encryption or decryption). Furthermore, it is semantically secure against adaptively chosen-ciphertext attacks under the sole assumption of the Computational Diffie-Hellman problem (and not the Decisional one), in the random oracle model.

4.2 The Okamoto-Uchiyama Encryption Scheme

Let us turn to the Okamoto-Uchiyama encryption scheme [24], which is one-way related to the factorization. Our transformation leads to the scheme presented in Figure 4, together with the following security properties. The partially trapdoor

p, q large prime integers of same length, and $n = p^2q$ $H : \{0, 1\}^k \rightarrow \mathbb{Z}_n$ and $G : \mathbb{Z}_n \rightarrow \{0, 1\}^k$ $g \in \mathbb{Z}_n^*$ such that the order of $g_p = g^{p-1} \bmod p^2$ is p $h = g^n \bmod n$
Encryption
$r \in_R \mathbb{Z}_n, s \in_R \{0, 1\}^{k_1}$ $\mathcal{Enc}(m, r \ s) = \begin{cases} a = g^r h^{H(m \ s)} \\ b = (m \ s) \oplus G(r) \end{cases}$
Decryption
$\mathcal{Dec}(a, b, c) = \begin{cases} r = L(y_p) / L(g_p) \bmod p \\ m \ s = b \oplus G(r) \\ a \stackrel{?}{=} g^r h^{H(m \ s)} \end{cases}$ where $y_p = y^{p-1} \bmod p^2$ and $L(x) = (x - 1) / p$.

Fig. 4. The OU-based Encryption Scheme

one-way injection is known as relying on the factorization of the large integer

$$n = p^2q:$$

$$f : \mathbb{Z}_p \times \mathbb{Z}_{(p-1)(q-1)} \longrightarrow \mathbb{Z}_n^* \quad g : \mathbb{Z}_n^* \longrightarrow \mathbb{Z}_p$$

$$(x, r) \longmapsto g^x \times h^r \bmod n \quad y \longmapsto \frac{L(y_p)}{L(g_p)} \bmod p$$

This problem is well-known to be random self-reducible, however, one can furthermore use the “easy verifiable” property. Indeed, we can use the attacker to suggest candidates as partial-preimage of a known $y = g^x h^n \bmod n$, with a rather large x . With all the candidates a , one computes $\gcd(x - a, n)$ which should provide p for the right solution.

Theorem 14 (The OU-based Encryption Scheme). *Any algorithm \mathcal{A} able to break the semantic security of the OU-based Encryption Scheme under adaptively chosen-ciphertext attacks within time T can be used to factor n with probability greater than $\text{Adv}_{\mathcal{A}} - (q_H(q_D + 1))/2^{k_1} - q_D/Y$, within the same time T .*

Advantages of the OU-based Encryption Scheme. The main advantage of this scheme is clear: the original one [24] is totally breakable under a (non-adaptive) chosen-ciphertext attack, which is a serious drawback. However its main interest was the factorization-based security. But just the one-wayness was related to the factorization. Indeed, even semantic security against chosen-plaintext attacks requires the higher residues assumption.

The presented scheme does not increase so much the computational load, but considerably enhances the security: chosen-ciphertext security is related to factorization, by a *perfect reduction* (the underlying problem can be broken within the same time and identical probability as the security property).

5 Conclusion

In this paper, we have presented the most interesting generic transformation which provides chosen-ciphertext secure schemes from the weakest possible assumption: the existence of partially trapdoor one-way functions. Furthermore, the exact security provides very practical results in the most common cases, random self-reducible or easy verifiable problems. Indeed, in this latter case, the reduction is optimal: the underlying problem can be broken within the same time and with the same probability than the resulting encryption scheme.

Finally, applications to well-known problems lead to very useful schemes: the most efficient based on the Computational Diffie–Hellman problem and the first one as secure as factorization.

Furthermore, to improve efficiency, one can integrate symmetric encryption, with $G(r)$ as secret key, instead of using the one-time pad, as it has already been done in recent works [1, 25, 17, 30].

Acknowledgements

I would like to thank Dan Boneh for his help and many fruitful comments.

References

1. M. Abdalla, M. Bellare, and P. Rogaway. DHAES: An Encryption Scheme Based on the Diffie-Hellman Problem. IEEE P1363a Submission. September 1998.
Available from <http://grouper.ieee.org/groups/1363/addendum.html>.
2. 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.
3. 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.
4. 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.
5. M. Bellare and P. Rogaway. The Exact Security of Digital Signatures – How to Sign with RSA and Rabin. In *Eurocrypt '96*, LNCS 1070, pages 399–416. Springer-Verlag, Berlin, 1996.
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. J. D. Cohen (Benaloh). Improving Privacy in Cryptographic Elections. Technical Report TR-454, Yale University, February 1986.
8. J. D. Cohen (Benaloh). *Improving Privacy in Cryptographic Elections*. PhD thesis, Yale University, September 1987. Also available as technical report TR-561.
9. 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.
10. 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.
11. 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.
12. 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.
13. W. Diffie and M. E. Hellman. New Directions in Cryptography. *IEEE Transactions on Information Theory*, IT-22(6):644–654, November 1976.
14. D. Dolev, C. Dwork, and M. Naor. Non-Malleable Cryptography. In *Proc. of the 23rd STOC*. ACM Press, New York, 1991.
15. 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.
16. 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.
17. 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.
18. S. Goldwasser and S. Micali. Probabilistic Encryption. *Journal of Computer and System Sciences*, 28:270–299, 1984.
19. SET Secure Electronic Transaction LLC. SET Secure Electronic Transaction Specification – Book 3: Formal Protocol Definition, may 1997.
Available from <http://www.setco.org/>.
20. U. M. Maurer. Diffie Hellman Oracles. In *Crypto '96*, LNCS 1109, pages 268–282. Springer-Verlag, Berlin, 1996.
21. 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.
22. 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.
23. K. Ohta and T. Okamoto. On Concrete Security Treatment of Signatures Derived from Identification. In *Crypto '98*, LNCS 1462, pages 354–369. Springer-Verlag, Berlin, 1998.
24. 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.
25. T. Okamoto, S. Uchiyama, and E. Fujisaki. EPOC: Efficient Probabilistic Public-Key Encryption. Submission to IEEE P1363a. November 1998.
Available from <http://grouper.ieee.org/groups/1363/addendum.html>.
26. P. Paillier. A Trapdoor Permutation Equivalent to Factoring. In *PKC '99*, LNCS 1560, pages 219–222. Springer-Verlag, Berlin, 1999.
27. P. Paillier. Public-Key Cryptosystems Based on Discrete Logarithms Residues. In *Eurocrypt '99*, LNCS 1592, pages 223–238. Springer-Verlag, Berlin, 1999.

28. P. Paillier and D. Pointcheval. Efficient Public-Key Cryptosystems Provably Secure against Active Adversaries. In *Asiacrypt '99*, LNCS. Springer-Verlag, Berlin, 1999.
29. D. Pointcheval. New Public Key Cryptosystems based on the Dependent-RSA Problems. In *Eurocrypt '99*, LNCS 1592, pages 239–254. Springer-Verlag, Berlin, 1999.
30. D. Pointcheval. HD-RSA: Hybrid Dependent RSA - a New Public Key Encryption Scheme. Submission to IEEE P1363a. October 1999.
Available from <http://grouper.ieee.org/groups/1363/addendum.html>.
31. D. Pointcheval and J. Stern. Security Arguments for Digital Signatures and Blind Signatures. *Journal of Cryptology*, 1999.
Available from <http://www.di.ens.fr/~pointche>.
32. 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.
33. 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.
34. RSA Data Security, Inc. Public Key Cryptography Standards – PKCS.
Available from <http://www.rsa.com/rsalabs/pubs/PKCS/>.
35. C. P. Schnorr. Efficient Signature Generation by Smart Cards. *Journal of Cryptology*, 4(3):161–174, 1991.
36. V. Shoup. Lower Bounds for Discrete Logarithms and Related Problems. In *Eurocrypt '97*, LNCS 1233, pages 256–266. Springer-Verlag, Berlin, 1997.
37. V. Shoup and R. Gennaro. Securing Threshold Cryptosystems against Chosen Ciphertext Attack. In *Eurocrypt '98*, LNCS 1403, pages 1–16. Springer-Verlag, Berlin, 1998.
38. Y. Tsiounis and M. Yung. On the Security of El Gamal based Encryption. In *PKC '98*, LNCS. Springer-Verlag, Berlin, 1998.