Efficient and Provable Security Amplifications

Ronald Cramer (cramer@cwi.nl)
Ivan Damgård (ivan@daimi.aau.dk)
Torben Pedersen (torbenp@daimi.aau.dk

CWI, P.O. Box 94079, NL-1090 GB Amsterdam, The Netherlands
Aarhus University, Ny Munkegade, DK-8000 Aarhus C, Denmark
CryptoMathic, Gustav Wiedsvej 10, 8000 Aarhus C, Denmark

Abstract. Even, Goldreich and Micali showed at Crypto'89 that the existence of signature schemes secure against known message attacks implies the existence of schemes secure against adaptively chosen message attacks. Unfortunately, this transformation leads to a rather impractical scheme. We exhibit a similar security amplification, which takes the given scheme to a new signature scheme that is not even existentially forgeable under adaptively chosen message attacks. Additionally, however, our transformation will be practical: The complexity of the resulting scheme is twice that of the original scheme.

The principles of both transformations carry over to block encryption systems. It is shown how they can be used to convert a block encryption system secure against known plaintext attacks to a system secure against chosen plaintext attacks. For both schemes it is shown that if the transformed scheme can be broken given a number, T, of encryptions of adaptively chosen plaintexts, then the original scheme can be broken given encryptions of T uniformly chosen plaintexts. In this case, however, the application of the technique of Even, Goldreich and Micali leads to the more efficient scheme. The transformed scheme has the same key length as the original, and ciphertexts are doubled in length. As an example, when applied to DES the transformed scheme is secure against differential cryptanalysis, which relies on the ability to get encryptions of plaintext pairs with proper differences.

1 Introduction

Goldwasser, Micali and Rivest [5] distinguish between several levels of security for digital signature schemes. Any such level, is defined by the extent to which an attacker has access to a true signer and the goal of the attacker. A digital signature scheme is said to have a given security level, if the scheme is secure against the corresponding attacker.

In the strongest level of security described in [5], the attacker is allowed to first use a true signer as an oracle, i.e., he can obtain a signature on any message of his choice. The attacker's goal is to generate a signature on some *new* message, i.e., a message he hasn't requested the oracle to sign. A digital signature scheme that is secure against such an attacker is called *not existentially forgeable under adaptively chosen message attacks.* This attacker has the weakest possible goal while having the strongest possible access to the signer. Therefore, this level of security is thought to be the most desirable. In [5], a scheme is exhibited with this level of security, under the assumption that a family of claw-free one-way trapdoor permutations exists. After [1], [8] and [10], the matter of secure signature schemes is, at least theoretically, settled: the existence of one-way functions is a sufficient and necessary condition for the existence of signature schemes with this security level. It must be noted, however, that these theoretical schemes have, by their impracticality, little value in real life.

We will consider schemes that are *not existentially forgeable under known plaintext attacks.* Here, an attacker has only passive access to the signature oracle. More precisely, the attacker receives signatures on messages that are chosen uniformly at random, after which he has to produce a forgery on a new message. Our goal will be to transform a signature scheme that is secure against such an attacker into a signature scheme that is not existentially forgeable under adaptively chosen message attacks. Furthermore, we will demand that the complexity of the resulting scheme is almost the same as that of the original scheme. The resulting transformation is presented in Section 2.

In the context of, so-called, on-line/off-line signatures, Even, Goldreich and Micali gave (in [3]) a transformation whose effect is the same security amplification, but with a loss of efficiency that would make the resulting scheme impractical for most applications where the signer is not able to perform off-line computations.

As in [5], and in many other cryptographic schemes, their approach works with two independently generated instances of the signature scheme Σ that is given as input. The resulting keys constitute the keys for the instance of $\overline{\Sigma}$.

Given a message m of length n that is to be signed. The first instance is used to authenticate the concatenation of $2n$ bit-strings, chosen uniformly at random by the signer. Bit-wise, the message m is used to select n of these strings which are finally authenticated, one-by-one, using the second instance of Σ. For each new message, this procedure is repeated. As a result of this bit-wise signing technique, the complexity of the transformed scheme becomes, roughly, the complexity of Σ times the number of bits that are signed. Therefore the transformation from [3] is not suitable to serve as a basis for security amplifications of practical signature schemes.

Interestingly, both the technique of [3] and the one used in this paper can also be applied to conventional encryption schemes. As shown in Section 3 this will yield, given an encryption scheme secure against known plaintext attacks, a new scheme secure against chosen plaintext attacks. More precisely, for both transformations it is shown that if the new scheme can be broken (i.e., a given ciphertext is decrypted or the key is recovered) given the encryptions of T chosen plaintexts, then the original scheme can be broken given the encryptions of T randomly chosen plaintexts. Furthermore, if the transformation is based on the principles of [3], then the transformed scheme has the same key space as the original scheme and encryption of one message block requires just one application of the given scheme (plus access to a number of random bits). In particular this implies that attacks on the transformed scheme based on differential cryptanalysis, which require the ability to get encryptions of pairs of plaintexts with proper differences (see [2]), will not be more efficient than attacks, where the plaintexts are chosen uniformly at random. Furthermore, it is shown that the new schemes are not more vulnerable against chosen ciphertext attack that the given scheme.

2 Security Amplification of Signatures

A signature scheme, Σ, is defined by a tuple (k, M, G, σ, V), where k is the security parameter, M the message space, G a key generation algorithm, σ a signature algorithm and V a verification algorithm. All algorithms are polynomial time in k, and k determines the length of messages that can be signed (see [5] for further details).

Let Σ be any such signature scheme and let $M(k)$ denote the message space and $|M(k)|$ its size corresponding to k. Moreover we assume that $M(k) = \{0,1\}^{t(k)}$, where $t(k)$ is some non-constant polynomial in k.

Let $\overline{M}(k)$ denote a subset of $M(k)$ that consists of a negligible[1] large fraction $\rho(k)$ of $M(k)$. For instance, $\overline{M}(k)$ could consist of all bit-strings of length t, with the last $t/2$ bits set to zero. Using Σ as a building block, a new scheme $\overline{\Sigma} = (k, \overline{M}, \overline{G}, \overline{\sigma}, \overline{V})$ is constructed as follows:

Message space

The message space for security parameter, k, is $\overline{M}(k)$ as defined above.

Initialisation

Let the security parameter k be given. To generate an instance of $\overline{\Sigma}$, the signer runs G twice, yielding two key-pairs (pk_1, sk_1) and (pk_2, sk_2). The public-key pk for the instance of $\overline{\Sigma}$ will be (pk_1, pk_2), and the secret key sk will be (sk_1, sk_2).

Signing

Let $m \in \overline{M}(k)$ be the message to be signed. The signer chooses a random

[1] A non-negative function $f : \mathbb{N} \to \mathbb{R}$ is negligible iff $\forall c > 0\, \exists n_0 \in \mathbb{N}\, \forall n \in \mathbb{N} : n > n_0 \Rightarrow f(n) \leq n^{-c}$.

pair (m_1, m_2), with $m_1, m_2 \in M(k)$, such that $m_1 \oplus m_2 = m$, and computes $\sigma_i(m_i)$ for $i = 1, 2$, where $\sigma_i(m_i)$ denotes a signature in Σ, with respect to the key-pair (pk_i, sk_i). The signature, $\overline{\sigma}(m)$, in $\overline{\Sigma}$ is $(m_1, m_2, \sigma_1(m_1), \sigma_2(m_2))$.

Verification

To verify a signature $\overline{\sigma}(m)$ on $m \in \overline{M}(k)$ with respect to pk, the receiver checks whether $m_1 \oplus m_2 \stackrel{?}{=} m$, and whether $\sigma_i(m_i)$ is a valid signature in Σ with respect to pk_i for $i = 1, 2$.

It is now shown that if Σ is secure against a known message attack, then $\overline{\Sigma}$ is secure against a chosen message attack.

We first need a lemma which says that it is very unlikely that $\sigma_1(m_1)$ and $\sigma_2(m_2')$ corresponding to signatures on two different messages, $m, m' \in \overline{M}(k)$, can be combined to a valid signature in $\overline{\Sigma}$.

Consider the following game involving two players A and B. Player B submits any member $m^1 \in \overline{M}(k)$ to A, and A returns a random pair (m_1^1, m_2^1), with $m_1^1, m_2^1 \in M(k)$, such that $m_1^1 \oplus m_2^1 = m^1$. They repeat this procedure, say, r times. This results in a sequence

$$(m^1, m_1^1, m_2^1), \ldots, (m^r, m_1^r, m_2^r),$$

such that $m_1^j \oplus m_2^j = m^j$ for $j = 1 \ldots r$. B is allowed to choose the values of m^j adaptively. B wins if he can find a pair (m_1^u, m_2^v) such that $m_1^u \oplus m_2^v \in \overline{M}(k)$ and $u \neq v$ and $1 \leq u, v \leq r$.

Lemma 2.1 *In the game described above, B's probability of losing the game is at least $1 - r(r-1)\rho(k)$.*

Proof Define for $1 \leq u, v \leq r$ and $u \neq v$, the stochastic variable $X_{u,v} = m_1^u \oplus m_2^v$. The probability that $X_{u,v} \in \overline{M}(k)$ is clearly fully determined by A's uniform coin flips, and therefore equal to $\rho(k)$. As there are $r(r-1)$ pairs (u, v), B will win with probability at most $r(r-1)\rho(k)$ and hence lose the game with the claimed probability. □

Now consider the signature scheme $\overline{\Sigma}$ described above. Let $\mathcal{A}$ be any probabilistic polynomial time algorithm that executes an adaptively chosen message attack on $\overline{\Sigma}$, and let $\mathcal{A}$'s signature requests be on messages

$$m^1, \ldots, m^{r(k)} \in \overline{M}(k),$$

with $r(k)$ polynomially bounded. The signer then returns

$$(\sigma_1(m_1^1), \sigma_2(m_2^1)), \ldots, (\sigma_1(m_1^{r(k)})), \sigma_2(m_2^{r(k)}),$$

as required.

Proposition 2.2 *If the signature scheme Σ is not existentially forgeable under known message attacks, the attacker $\mathcal{A}$ has only negligible probability of outputting a signature $\tilde{\sigma}(\tilde{m})$ in Σ, where $\tilde{m} \neq m_i^j$ for $j = 1, \ldots, r(k)$, and $\tilde{\sigma}(\tilde{m})$ is a valid signature with respect to pk_i, with $i = 1$ or $i = 2$.*

Proof By standard simulation techniques. Suppose $\mathcal{A}$'s probability of success is non-negligible (in k). Let a signer S in Σ, with public key pk, be given. We will use the attacker $\mathcal{A}$ to conduct a successful known message attack on signer S, thus contradicting the assumption on Σ.

Generate an instance (pk', sk') in Σ. Choose i at random in $\{1, 2\}$, and put $pk = pk_i$ and $pk' = pk_{3-i}$. Now present the resulting key (pk_1, pk_2) for $\overline{\Sigma}$ to the attacker. The signer S with public key pk_i, used as a subroutine in the simulation, will output signatures on randomly chosen messages. More specifically, the simulation works as follows.

1. Receive message $m \in \overline{M}(k)$ from the attacker.
2. Receive a signature $\sigma_i(m_i)$ from S, where S chooses $m_i \in M(k)$ uniformly at random.
3. Compute $m_{3-i} = m_i \oplus m$ and $\sigma_{3-i}(m_{3-i})$. Forward $\overline{\sigma}(m)$ to the attacker.

As the attacker cannot distinguish this simulation from a true signer in $\overline{\Sigma}$, the probability that $\tilde{\sigma}(\tilde{m})$ is a forgery of S's signature is half $\mathcal{A}$'s success probability. This is still non-negligible. □

Theorem 2.3 *Let Σ be any signature scheme that is not existentially forgeable under known message attacks. Then the signature scheme $\overline{\Sigma}$ is not existentially forgeable under adaptively chosen message attacks*

Proof Let $\tilde{m} \in \overline{M}(k)$ and let $\tilde{\sigma}(\tilde{m}) = (m_1, m_2, \sigma_1(m_1), \sigma_2(m_2))$ be a forgery in $\overline{\Sigma}$ on a new message, obtained after an adaptively chosen message attack. By Proposition 2.2, except with negligible probability $\sigma_1(m_1) = \sigma_1(m_1^u)$, and $\sigma_2(m_2) = \sigma_2(m_2^v)$, for some u, v with $1 \leq u, v \leq r(k)$ and $u \neq v$ (notation as in Proposition 2.2). So we must have that $m_1^u \oplus m_2^v = \tilde{m}$. However, by Lemma 2.1, this has only negligible probability. □

3 Symmetric Encryption Schemes

Let an encryption scheme with key space, K, plaintext space, M, and ciphertext space, C, be given. Encryption of $m \in M$ under key $k \in K$ is denoted by $E_k(m)$ and decryption of $c \in C$ under k is denoted by $D_k(c)$. Such a scheme will be denoted by (E, D, K, M, C).

As for signature schemes it is possible to classify attacks against an encryption scheme in terms of the goal of the attack and the amount of information, which is available to the attacker. Usually two goals are distinguished:

- Decrypting a given ciphertext, $c \in C$.
- Finding the key.

Other goals could be to find some information about the key or to find an alternative algorithm for decrypting (corresponding to "universal break" of signatures; see [5]). Of these two goals it is clearly harder to find the secret key than to just decrypt a given ciphertext. Four types of attacks can be distinguished:

- Given just the knowledge of the encryption scheme;
- Given a number of pairs $(m, E_k(m))$, where m is chosen by the owner of k according to some distribution;
- Chosen plaintext attack: The attacker may choose adaptively a number of plaintexts, $m_1, m_2, \ldots$ and get $E_k(m_1), E_k(m_2), \ldots$.
- Chosen ciphertext attack: The attacker may choose adaptively a number of ciphertexts $c_1, c_2, \ldots$ and get $D_k(m_1), D_k(m_2), \ldots$.

The last two attack are similar in many situations. Based on these definitions the highest security level of an encryption scheme is security against getting the key under a chosen plaintext or ciphertext attack.

In the following it is shown how the principle used in the previous section can be used to turn a scheme, secure against known plaintext attack where the messages are chosen uniformly at random, into a scheme which is secure against a chosen plaintext attack. This transformation requires twice as long key and the ciphertext is doubled. This may make the transformed system inadequate for many applications. Therefore, the construction is improved in Section 3.2, where a similar transformation is given which does not double the key size.

3.1 Applying the Basic Method

In the following it will be assumed that $M = \{0,1\}^n$ for some parameter n, but the construction works for any message space for which

- There is a binary operator, $\odot$, on M and a neutral element $m_0 \in M$ for this operator such that $(M, \odot, m_0)$ is a group.
- Efficient algorithms for selecting elements in M and computing both $\odot$ and its inverse exist.

Given an encryption scheme (E, D, K, M, C) ,
a new scheme $(E^{(1)}, D^{(1)}, K^{(1)}, M^{(1)}, C^{(1)})$ is defined as follows:

- $K^{(1)} = K \times K$.
- $M^{(1)} = M$.
- $C^{(1)} = C \times C$.
- $E^{(1)}_{k_1,k_2}(m) = (E_{k_1}(m_1), E_{k_2}(m_2))$ where m_1 is chosen uniformly at random in M and $m_2 = m \oplus m_1$.
- $D^{(1)}_{k_1,k_2}(c_1, c_2) = D_{k_1}(c_1) \oplus D_{k_2}(c_2)$.

Theorem 3.1 *If (E, D, K, M, C) is secure against known plaintext attacks after getting $T \in \mathbb{N}$ encryptions of uniformly chosen plaintexts.*
Then $(E^{(1)}, D^{(1)}, K^{(1)}, M^{(1)}, C^{(1)})$ is secure against chosen plaintext attacks where the attacker is allowed to choose T plaintexts adaptively.

Furthermore, if $(E^{(1)}, D^{(1)}, K^{(1)}, M^{(1)}, C^{(1)})$ can be broken in a chosen ciphertext attack given decryptions of T ciphertexts, then (E, D, K, M, C) can be broken in a similar attack also requiring T decryptions.

Proof The proof goes along the lines of the proof of Proposition 2.2, but is somewhat simpler.

First consider a chosen plaintext attack aiming at decrypting a given ciphertext (c_1, c_2). Then a given ciphertext c can be decrypted in the original system in a known message attack as follows. If k is the (unknown) key in this system we select a key-pair for the new system by choosing $i \in \{1, 2, \}$ at random and letting $k_i = k$ and choosing $k_{3-i} \in K$ at random. Construct a ciphertext as $c_i = c$ and $c_{3-i} = E_{k_{3-i}}(m')$ where $m' \in M$ is chosen at random. The chosen plaintext attack can now be simulated in the same way as the adaptively chosen message attack on $\overline{\Sigma}$ in the proof of Proposition 2.2. In the end we will get a plaintext m'' corresponding to (c_1, c_2) and finally output $m = m' \oplus m''$ as the plaintext corresponding to c.

The simulated attack will construct ciphertexts with the same distribution as the real attack. Thus, it will output the plaintext corresponding to (c_1, c_2) with the same probability, and by the definition of the new scheme the decryption of c is derived correctly.

Next consider attack aiming at recovering the secret key. This situation is as above, except that we don't have to generate (c_1, c_2). If the attack outputs the entire key, then we can get k as k_i. If the attack only outputs one of the two keys we will get k with probability $\frac{1}{2}$.[2]

Finally, a chosen ciphertext attack against $(E^{(1)}, D^{(1)}, K^{(1)}, M^{(1)}, C^{(1)})$ can be simulated given a chosen ciphertext attack against (E, D, K, M, C) in such a way that each decryption in the new scheme requires one decryption in the old scheme. □

An advantage of $(E^{(1)}, D^{(1)}, K^{(1)}, M^{(1)}, C^{(1)})$ is that if known plaintext attacks against (E, D, K, M, C) require time linear in $|K|$ (the size of K), no chosen plaintext attack can be better. A disadvantage is that, although the key length is doubled, exhaustive key search in $(E^{(1)}, D^{(1)}, K^{(1)}, M^{(1)}, C^{(1)})$ can still be done in time linear in $|K|$ ([6]). Thus we get a lower bound of at most $|K| = \sqrt{|K^{(1)}|}$ for chosen plaintext attacks. This is not a sufficient security bound.

3.2 Improving the Security

As mentioned above, the security bound of $(E^{(1)}, D^{(1)}, K^{(1)}, M^{(1)}, C^{(1)})$ is not satisfactory, and furthermore, it might be too inefficient to use two encryptions in (E, D, K, M, C) in order to encrypt a single message block. In the following another construction is given, which improves on both of these deficiencies, while keeping the security amplification. This construction can be seen as an application of the principles of [3] to encryption schemes: use the given system (signature or encryption) to instantiate a one-time system (by signing the public key/encrypting the key) and use the one-time system on the input

[2] The proof of this case is the reason for choosing $k = k_1$ with probability $\frac{1}{2}$ — in the proofs of the other claims it would have been sufficient to let $k = k_1$ with probability 1.

(message/plaintext). The new scheme $(E^{(2)}, D^{(2)}, K^{(2)}, M^{(2)}, C^{(2)})$ is defined as follows:

- $K^{(2)} = K$.
- $M^{(2)} = M$.
- $C^{(2)} = C \times C$.
- $E_k^{(2)}(m) = (E_k(m_1), m_2)$, where m_1 is chosen uniformly at random in M and $m_2 = m \oplus m_1$.
- $D_k^{(2)}(c_1, c_2) = D_k(c_1) \oplus c_2$.

Theorem 3.2 *If (E, D, K, M, C) is secure against known plaintext attacks after getting $T \in \mathbb{N}$ encryptions of uniformly chosen plaintexts.*
Then $(E^{(2)}, D^{(2)}, K^{(2)}, M^{(2)}, C^{(2)})$ is secure against chosen plaintext attacks where the attacker is allowed to choose T plaintexts adaptively.

Furthermore, if $(E^{(2)}, D^{(2)}, K^{(2)}, M^{(2)}, C^{(2)})$ can be broken in a chosen ciphertext attack given decryptions of T ciphertexts, then (E, D, K, M, C) can be broken in a similar attack also requiring T decryptions.

Proof The proof follows the previous quite closely. Whenever, the attacker against $(E^{(2)}, D^{(2)}, K^{(2)}, M^{(2)}, C^{(2)})$ asks for an encryption of m, the simulator gets a pair $(m_1, c_1) = (m_1, E_k(m_1))$ for a randomly chosen $m_1 \in M$ from the owner of k and returns $(c_1, m_1 \oplus m)$. This simulator constructs ciphertexts with the same distribution as if they were obtained by proper encryptions in $(E^{(2)}, D^{(2)}, K^{(2)}, M^{(2)}, C^{(2)})$.

If the attack aims at decrypting a ciphertext (c_1, c_2) in $(E^{(2)}, D^{(2)}, K^{(2)}, M^{(2)}, C^{(2)})$ then a ciphertext, c in (E, D, K, M, C) can be decrypted by choosing $c_1 = c$ and c_2 at random and then using the method sketched above to simulate the chosen plaintext attack. If m is the plaintext corresponding to (c_1, c_2) then $m \oplus c_2$ is the plaintext corresponding to c. If the attack outputs information about the key in $(E^{(2)}, D^{(2)}, K^{(2)}, M^{(2)}, C^{(2)})$, then the same information is obtained about the key in (E, D, K, M, C) as they are equal. Again, a chosen ciphertext attack against $(E^{(2)}, D^{(2)}, K^{(2)}, M^{(2)}, C^{(2)})$ can easily be simulated given a chosen ciphertext attack against (E, D, K, M, C) □

4 Other Applications

The techniques desribed in this paper immediately carry over to message authenticity codes (MAC's) based on pseudo-random functions. We describe this application only briefly, and leave it to the reader to complete the details.

Given a family of pseudo-random functions that is secure against known plaintext attacks, we construct a MAC with two independent functions f_1 and f_2 from the family. The message space for the MAC is defined as in our application to signature schemes, i.e., superpolynomial in the size of the security parameter, but a negligible fraction of the domain of the functions. Splitting a message is also done in the same way. Given a message m and a random splitting m_1, m_2,

the MAC is now computed as $(f_1(m_1), f_2(m_2))$. This MAC is then secure against adaptively chosen plaintext attacks.

In the context of a special class of proofs of knowledge, "split-and-divert" techniques similar to those employed in this paper, have been used for efficient transformations from honest verifier zero knowledge protocols to witness hiding protocols in [4].

References

1. M. Bellare, S. Micali: *How to Sign Given Any Trapdoor Function.* Proceedings of STOC '88, pp.32–42.
2. E. Biham, A. Shamir: *Differential Cryptanalysis of DES-like Cryptosystems.* Proceedings of Crypto'90, pp. 2–21.
3. S. Even, O. Goldreich and S. Micali: *On-Line/Off-Line Digital Signatures.* Proceedings of Crypto '89, pp.263–275.
4. R. Cramer, I. Damgård, B. Schoenmakers, : "Proofs of Partial Knowledge and Simplified Design of Witness Hiding Protocols", Proceedings of Crypto '94, pp.174–187.
5. S. Goldwasser, S. Micali and R. Rivest: *A Digital Signature Scheme Secure Against Chosen Message Attacks.* SIAM Journal on Computing, 17(2): 281–308, 1988.
6. L.R. Knudsen. Personal communication.
7. M. Naor, C. Dwork: *An Efficient Existentially Unforgeable Signature Scheme and its Applications.* Proceedings of Crypto '94, pp.234–246.
8. M. Naor, M. Yung: *Universal One Way Hash functions and their Cryptographic Applications.* Proceedings of STOC '89, pp.33–43.
9. National Bureau of Standards. *Data encryption standard.* Federal Information Processing Standard (FIPS), Publication 46, National Bureau of Standards, U.S. Department of Commerce, Washington DC, January 1977.
10. J. Rompel: *One Way Functions are Necessary and Sufficient for signatures.* Proceedings of STOC '90, pp.387–394.