ISSN 1342-2812

Research Reports on Mathematical and Computing Sciences

Public-key Steganography with Authentication

Hirotoshi Takebe and Keisuke Tanaka

December 2007, C–251

Department of
Mathematical and
Computing Sciences
Tokyo Institute of Technology

SERIES C: Computer Science

Public-key Steganography with Authentication

Hirotoshi Takebe and Keisuke Tanaka

Dept. of Mathematical and Computing Sciences
Tokyo Institute of Technology
W8-55, 2-12-1 Ookayama Meguro-ku, Tokyo 152-8552, Japan
{takebe3, keisuke}@is.titech.ac.jp

January 28, 2008

Abstract

Steganography is the science of sending messages hidden in harmless communications over a public channel so that an adversary eavesdropping on the channel cannot even detect the presence of the hidden messages. Several models for steganography have been introduced. Some are private-key settings, and the others are public-key settings. In this paper, we propose a model of public-key steganography with authentication. We formalize its security condition. We also construct a concrete scheme of public-key steganography with authentication via a public-key encryption scheme and a digital signature scheme.

Keywords: public-key encryption, digital signature, steganography.

1 Introduction

Background. Steganography is the science of hiding information by embedding messages within other ones which are seemingly harmless. As the goal of steganography is to hide the presence of a message, it can be seen as the complement of cryptography, whose goal is to hide the content of a message.

We consider two parties linked by a public communication channel which is under watch by an adversary. The sender sends a message which is seemingly harmless because their conversation is watched at any time. A genuine communication message is called *covertext*. However, if he wants to send a message which he does not want to be known to an adversary, he may embed it as hidden information in another message, which is also seemingly harmless. Such a message is called *stegotext*. The adversary, who knows the distribution of the covertext, tries to detect whether a given message is covertext or stegotext.

Related Work. Various protocols have been proposed for steganography, which are surveyed by Anderson and Petitcolas [1]. Formal models for steganography were recently introduced.Foe example, there are several information-theoretic formalizations [4, 13, 9] and one complexity-theoretic model [7]. However, these models have addressed *private-key* steganography. In other words, only the parties who share a secret or a private-key in advance can use these protocols. In contrast, *public-key* steganography allows parties to communicate steganographically with no prior exchange of secrets.

Public-key steganography was formalized by von Ahn and Hopper [11]. They defined the security notion for public-key steganography, which was the analogue of the security against the chosen-plaintext attack of public-key cryptosystem. They constructed the stegosystem which satisfied this notion. This stegosystem consists of two conversion methods, which are a public-key

cryptosystem and a procedure Basic_Encode. In order to construct the secure stegosystem, they defined a security notion for public-key cryptosystem, which was the indistinguishability from random bits under the chosen-plaintext attack. They proposed some public-key encryption schemes satisfying this notion. They also formalized the notion of the steganographic key exchange, and constructed the secure steganographic key exchange protocol under the Decisional Diffie-Hellman assumption.

Backes and Cachin [3] defined a new security notion for public-key steganography which was stronger than that of von Ahn and Hopper [11]. A stegosystem which satisfies this notion is called *steganographically secure against the adaptive chosen-covertext attack* (SS-CCA). Analogously to the standard cryptographic notion of the chosen-ciphertext attack, this seems to be the most general type of possible attacks on a stegosystem. They also defined another security notion of the security. It is the steganographic security against the *replayable* adaptive chosen-covertext attack (SS-RCCA), which is relaxed notion of SS-CCA. They showed that an SS-RCCA stegosystem could be constructed from any RCCA-secure public-key cryptosystem [5] whose ciphertexts were pseudorandom. Hopper [6] constructed an SS-CCA stegosystem, which relied on the existence of public-key encryption schemes which satisfied the indistinguishability from random bits under the chosen-ciphertext attack (IND$-CCA [6]). They showed the existence of such an encryption scheme under the Decisional Diffie-Hellman assumption.

Contribution. In previous setting, the goal of adversary is detection. Indeed, their schemes are secure against the detecting attacks. Now, we consider the impersonation. In some previous schemes, the steganographic-encoding algorithm is so public that an eavesdropper Eve may pretend to be the sender. In particular, Eve may make stegotexts as which the receiver will accept from the valid sender. Our idea, to prevent such an attack, is that the steganographic-encoding algorithm runs with sender's secret information so that anyone except the valid sender cannot make stegotexts as which will be accepted from the valid sender. Namely, we consider the authenticity of stegotexts.

In this paper, we propose public-key steganography with authentication by employing the idea of von Ahn and Hopper [11] for constructing the public-key steganography. We define the security of public-key steganography with authentication, the steganographic security and the unforgeability.

We also construct a concrete scheme of public-key steganography with authentication by modifying a public-key steganography scheme proposed by Hopper [6]. We construct it via a public-key encryption scheme and a digital signature scheme. We show that our proposed scheme of public-key steganography with authentication is steganographically secure and unforgeable if the underlying public-key encryption scheme satisfies the indistinguishability from random bits under the chosen-ciphertext attack (IND$-CCA) and the underlying digital signature scheme satisfies the existential unforgeability from the chosen-message attack (EUF-CMA).

Organization. We give preliminaries in section 2. We propose definitions and the security properties for public-key steganography with authentication in section 3. We construct a concrete scheme of public-key steganography with authentication and give security proofs in section 4. We give the conclusion in section 5.

2 Preliminaries

We say that a function $\mu : \mathbb{N} \to [0,1]$ is *negligible* in n if for every $c > 0$, there exists n_0 such that $\mu(n) < \frac{1}{n^c}$ for all $n > n_0$.

We denote the uniform distribution on k bit strings by $\mathcal{U}_k$. Let $\mathcal{D}$ be a probability distribution. We denote $x \leftarrow \mathcal{D}$ as the action of drawing a sample x according to $\mathcal{D}$. We denote the *minimum entropy* of a probability distribution $\mathcal{D}$ with finite set X by $H_\infty(\mathcal{D}) = \min_{x \in X} \left\{ \log_2 \frac{1}{\Pr[x \leftarrow \mathcal{D}]} \right\}$.

Let $\mathcal{D}$ be a probability distribution on the finite set X. We say that a function $f : X \to \{0,1\}$ is *ϵ-biased* on $\mathcal{D}$ if $|\Pr[f(x) = 0 | x \leftarrow \mathcal{D}] - \frac{1}{2}| < \epsilon$. We also say that f is *perfectly unbiased* on $\mathcal{D}$ if $\Pr[f(x) = 0 | x \leftarrow \mathcal{D}] = \frac{1}{2}$.

Let X and Y be finite sets and F a family of functions $f : X \to Y$. We say that a family F is *strongly universal* [12] if for all distinct x_1, $x_2 \in X$ and all y_1, $y_2 \in Y$ which are not necessarily distinct, the number of functions $f \in F$ such that $f(x_1) = y_1$ and $f(x_2) = y_2$ is exactly $\frac{|F|}{|Y|^2}$.

2.1 Public-Key Encryption

First, we define a public-key encryption scheme.

Definition 1 (public-key encryption). *A public-key encryption scheme $\mathcal{PE}$ is a tuple of four algorithms denoted by* $(\mathsf{Enc\_CGen}, \mathsf{Enc\_KGen}, \mathsf{Enc}, \mathsf{Dec})$.

- $\mathsf{Enc\_CGen}$: *The common parameter generation algorithm* $\mathsf{Enc\_CGen}$ *is a probabilistic algorithm. On input a security parameter k,* $\mathsf{Enc\_CGen}$ *returns a sequence of common parameters cp_{enc} containing the security parameter k and other system-wide parameters such as the description of computational groups and hash functions. We write this as $cp_{enc} \leftarrow \mathsf{Enc\_CGen}(1^k)$.*
- $\mathsf{Enc\_KGen}$: *The key generation algorithm* $\mathsf{Enc\_KGen}$ *is a probabilistic algorithm. On input a common parameter cp_{enc},* $\mathsf{Enc\_KGen}$ *returns a pair of (pk, sk). pk and sk are public and secret keys, respectively. We write this as $(pk, sk) \leftarrow \mathsf{Enc\_KGen}(cp_{enc})$.*
- Enc: *The encryption algorithm* Enc *is a probabilistic algorithm. On input a common parameter cp_{enc}, a public key pk, and a message m,* Enc *returns a ciphertext c. We write this as $c \leftarrow \mathsf{Enc}(cp_{enc}, pk, m)$.*
- Dec: *The decryption algorithm* Dec *is a deterministic algorithm. On input a common parameter cp_{enc}, a secret key sk, and a ciphertext c,* Dec *returns either a message m or a symbol $\perp$ which indicates that the ciphertext c is invalid. We write this as $m/\perp \leftarrow \mathsf{Dec}(cp_{enc}, sk, c)$.*

We denote $\mathcal{M}_{\mathcal{PE}}$ as the message space of $\mathcal{PE}$. We require that for all cp_{enc} which can be output by $\mathsf{Enc\_CGen}(1^k)$, for all (pk, sk) which can be output by $\mathsf{Enc\_KGen}(cp_{enc})$, for all $m \in \mathcal{M}_{\mathcal{PE}}$, and for all c which can be output by $\mathsf{Enc}(cp_{enc}, pk, m)$, we have that $\mathsf{Dec}(cp_{enc}, sk, c) = m$.

Second, we review the notion of public-key encryption proposed by Hopper [6]. It is the indistinguishability from random bits under the chosen-ciphertext attack. Let $\mathcal{PE} = (\mathsf{Enc\_CGen}, \mathsf{Enc\_KGen}, \mathsf{Enc}, \mathsf{Dec})$ be a public-key encryption scheme and k a security parameter. Let ℓ be the function which implies the length of the ciphertext. We define a distinguishing game under the chosen-ciphertext attack against $\mathcal{PE}$ by an adversary A_d and a challenger. We consider the experiments $\mathbf{Exp}^i_{\mathbf{CCA}}$ for $i \in \{0,1\}$ as follows:

$\mathbf{Exp}^i_{\mathbf{CCA}}(1^k)$

1. $cp_{enc} \leftarrow \mathsf{Enc\_CGen}(1^k)$ and $(pk, sk) \leftarrow \mathsf{Enc\_KGen}(cp_{enc})$.
2. A_d is given cp_{enc} and pk.
3. A_d can make access to the decoding oracle $\mathtt{DEC}_{sk}$. A_d queries the ciphertext c to $\mathtt{DEC}_{sk}$, and receives either the corresponding plaintext m or a symbol $\perp$.
4. A_d produces a message m^* and passes it to the challenger. The challenger passes s^*_i to A_d.
5. A_d continues to query the decoding oracle $\mathtt{DEC}_{sk}$ with the restriction that A_d may not query s^*_i.
6. A_d outputs a bit γ.

7. Return γ.

We define s_i^* for $i \in \{0,1\}$ as follows:

- s_0: The challenger computes $s_0 \leftarrow \mathsf{Enc}(cp_{enc}, pk, m)$.
- s_1: The challenger samples $s_1 \leftarrow \mathcal{U}_{l^*}$.

We define A_d's advantage against $\mathcal{PE}$ by

$$\mathbf{Adv}^{\mathsf{ind\$-cca}}_{\mathcal{PE},A_d}(k) = |\Pr[\mathbf{Exp}^0_{\mathbf{CCA},A_d}(1^k) = 1] - \Pr[\mathbf{Exp}^1_{\mathbf{CCA},A_d}(1^k) = 1]| \ .$$

We denote the set of A_d who make q_D decoding queries in running time t by $\mathcal{A}_d(t, q_D)$.

Definition 2 (IND$-CCA). *We say that $\mathcal{PE}$ is indistinguishable from random bits under the chosen-ciphertext attack* (IND$-CCA) *if for every probabilistic polynomial adversary $A_d \in \mathcal{A}_d(t, q_D)$, $\mathbf{Adv}^{\mathsf{ind\$-cca}}_{\mathcal{PE},A_d}(k)$ is negligible in k.*

2.2 Digital Signature

First, we define a digital signature scheme.

Definition 3 (digital signature). *A digital signature scheme $\mathcal{DS}$ is a tuple of four algorithms denoted by* (Sig_CGen, Sig_KGen, Sig, Ver).

- Sig_CGen: *The common parameter generation algorithm* Sig_CGen *is a probabilistic algorithm. On input a security parameter k,* Sig_CGen *returns a sequence of common parameters cp_{sig} containing the security parameter k and other system-wide parameters such as the description of computational groups and hash functions. We write this as $cp_{sig} \leftarrow \mathsf{Sig\_CGen}(1^k)$.*
- Sig_KGen: *The key generation algorithm* Sig_KGen *is a probabilistic algorithm. On input a common parameter cp_{sig},* Sig_KGen *returns a pair of (pk, sk). pk and sk are public and secret keys, respectively. We write this as $(pk, sk) \leftarrow \mathsf{Sig\_KGen}(cp_{sig})$.*
- Sig: *The signing algorithm* Sig *is a probabilistic algorithm. On input a common parameter cp_{sig}, a secret key sk, and a message m,* Sig *returns a signature σ for m. We write this as $\sigma \leftarrow \mathsf{Sig}(cp_{sig}, sk, m)$.*
- Ver: *The verification algorithm* Ver *is a deterministic algorithm. On input a common parameter cp_{sig}, a pubic key pk, a message m, and a candidate signature σ for m,* Ver *returns 1 if σ is the valid signature for m. Otherwise,* Ver *returns 0. We write this as $0/1 \leftarrow \mathsf{Ver}(cp_{sig}, pk, m, \sigma)$.*

We denote $\mathcal{M}_{\mathcal{DS}}$ as the message space of $\mathcal{DS}$. We require that for all cp_{sig} which can be output by $\mathsf{Sig\_CGen}(1^k)$, for all (pk, sk) which can be output by $\mathsf{Sig\_KGen}(cp_{sig})$, for all $m \in \mathcal{M}_{\mathcal{DS}}$, and for all σ which can be output by $\mathsf{Sig}(cp_{sig}, sk_A, m)$, we have that $\mathsf{Ver}(cp_{sig}, pk_A, m, \sigma) = 1$.

We next define the security of digital signature. Let $\mathcal{DS} = (\mathsf{Sig\_CGen}, \mathsf{Sig\_KGen}, \mathsf{Sig}, \mathsf{Ver})$ be a digital signature scheme and k a security parameter. Let A_f be an adversary who forges a signature. We consider the experiments $\mathbf{Exp}_{\mathbf{CMA}}$ as follows:

$\mathbf{Exp}_{\mathbf{CMA}}(1^k)$

1. $cp_{sig} \leftarrow \mathsf{Sig\_CGen}(1^k)$, $(pk, sk) \leftarrow \mathsf{Sig\_KGen}(cp_{sig})$.
2. A_f is given cp_{sig} and pk.
3. A_f queries messages to the signing oracle $\mathtt{SIG}_{sk}$, and receives the corresponding signatures, adaptively.

4. In the end, A_f outputs (m^*, σ^*). Let $d \leftarrow \mathsf{Ver}(cp_{sig}, pk_A, m^*, \sigma^*)$. If $d = 1$ and A_f has not queried m^*, then return 1. Otherwise, return 0.

We define A_f's advantage against $\mathcal{DS}$ by

$$\mathbf{Adv}^{\mathsf{euf-cma}}_{\mathcal{DS},A_f}(k) = \Pr[\mathbf{Exp}^{}_{\mathbf{CMA},A_f}(1^k) = 1].$$

We denote the set of A_f who make q_S signing queries in running time t by $\mathcal{A}_f(t, q_S)$.

Definition 4 (EUF-CMA). *We say that $\mathcal{DS}$ is existentially unforgeable under the chosen-message attack* (EUF-CMA) *if for any probabilistic polynomial adversary $A_f \in \mathcal{A}_f(t, q_S)$, $\mathbf{Adv}^{\mathsf{euf-cma}}_{\mathcal{DS},A_f}(k)$ is negligible in k.*

2.3 Pseudorandom Generators

We define a pseudorandom generator. Let $G : \{0,1\}^k \to \{0,1\}^{l(k)}$ be a function which is computable in polynomial time and $k < l(k)$. We define a distinguishing game by an adversary A_{pr} and a challenger. We consider the experiments $\mathbf{Exp}^{i}_{\mathbf{PRG}}(1^k)$ for $i \in \{0,1\}$ as follows:

$\mathbf{Exp}^{i}_{\mathbf{PRG}}(1^k)$

1. The challenger passes z_i to A_{pr}.
2. A_{pr} outputs a bit γ.
3. Return γ.

We define r_i for $i \in \{0,1\}$ as follows:

- z_0: The challenger chooses $x \leftarrow \mathcal{U}_k$ and computes $z_0 = G(x)$.
- z_1: The challenger chooses $z_1 \leftarrow \mathcal{U}_{l(k)}$.

We define A_{pr}'s advantage against G by

$$\mathbf{Adv}^{\mathsf{prg}}_{G,A_{pr}}(k) = |\Pr[\mathbf{Exp}^{0}_{\mathbf{PRG},G,A_{pr}}(1^k) = 1] - \Pr[\mathbf{Exp}^{1}_{\mathbf{PRG},G,A_{pr}}(1^k) = 1]|.$$

We denote the set of A_{pr} in running time t by $\mathcal{A}_{pr}(t)$.

Definition 5 (pseudorandom generator). *We say that G is a pseudorandom generator if for any probabilistic polynomial adversary $A_{pr} \in \mathcal{A}_{pr}(t)$, $\mathbf{Adv}^{\mathsf{prg}}_{G,A_{pr}}(k)$ is negligible in k.*

2.4 Channels

Intuitively, the communication between the parties follows the distribution relied on the previous communications. In order to define this notion, we follow previous works [7, 8, 11, 3, 6] on steganography.

We model the communication between two parties by a *channel*. Let D be a finite set of documents, we define that $D^* = D \times D \times \cdots$. We define a channel $\mathcal{C} = \{\mathcal{C}_h | h \in D^*\}$, which is a family of probability distributions on a set of documents D, indexed by sequences $h \in D^*$. We call the index h the *history*. For an integer ℓ, we define the distribution $\mathcal{C}^{\ell}_h = \mathcal{C}_h \times \mathcal{C}_{(h||d_1)} \times \mathcal{C}_{(h||d_1||d_2)} \times \cdots \times \mathcal{C}_{(h||d_1||\ldots||d_{\ell-1})}$, where $d_1 \leftarrow \mathcal{C}_h, d_2 \leftarrow \mathcal{C}_{(h||d_1)}, \ldots, d_\ell \leftarrow \mathcal{C}_{(h||d_1||\ldots||d_{\ell-1})}$. A history $h = (d_1||d_2||\ldots||d_\ell)$ is called *legal* with respect to $\mathcal{C}$ if for all i, $\Pr[d_i \leftarrow \mathcal{C}_{(d_1||\ldots||d_{i-1})} | d_1 \leftarrow \mathcal{C}_\nu, d_2 \leftarrow \mathcal{C}_{d_1}, \cdots, d_{i-1} \leftarrow \mathcal{C}_{(d_1||d_2||\ldots||d_{i-2})}] > 0$ where ν is an empty string.

In this setting, we allow all parties to access to the channel oracle for any h about $\mathcal{C}$. In other words, we allow the adversary to learn the covertext distribution on all communications by an oracle. The adversary queries the history h to the channel oracle, and receives the document d

where $d \leftarrow \mathcal{C}_h$. If the query h is not legal, then the channel oracle returns $\perp$. While attacking, the adversary can access to the channel oracle at any time.

We require the following property with respect to the channel proposed by Hopper [6]: the channel has the sampleability with efficiency. In particular, there is an efficiently computable algorithm $\mathsf{channel}$ such that the output distribution of $\mathsf{channel}(h, r)$ where $r \leftarrow \mathcal{U}_k$ and $\mathcal{C}_h$ are computationally indistinguishable.

For a function $f : D \to \{0, 1\}$, we define the following property: if $|\Pr[f(x) = 0 \mid r \leftarrow \mathcal{U}_k, x \leftarrow \mathsf{channel}(h, r)] - \frac{1}{2}| < \epsilon$ for any legal h, we say that f is ϵ-*biased* with respect to $\mathsf{channel}$.

3 Steganography with Authentication

In this section, we propose the definition and the security properties for public-key steganography with authentication.

Formalization. We first define the scheme of public-key steganography with authentication.

Definition 6 (public-key steganography with authentication). *A scheme of public-key steganography with authentication $\mathcal{ASS}$ is a tuple of five algorithms* $(\mathsf{Stg\_CGen}, \mathsf{Stg\_KGen}_A, \mathsf{Stg\_KGen}_B, \mathsf{Stg\_Enc}, \mathsf{Stg\_Dec})$ *as follows:*

- $\mathsf{Stg\_CGen}$: *The common parameter generation algorithm* $\mathsf{Stg\_CGen}$ *is a probabilistic algorithm. On input a security parameter* 1^k, $\mathsf{Stg\_CGen}$ *returns a sequence of common parameters cp containing the security parameter k and other system-wide parameters such as the description of computational groups and hash functions. We write this as* $cp \leftarrow \mathsf{Stg\_CGen}(1^k)$.

- $\mathsf{Stg\_KGen}_A$: *The key generation algorithm for the sender* $\mathsf{Stg\_KGen}_A$ *is a probabilistic algorithm. On input a common parameter sequence cp,* $\mathsf{Stg\_KGen}_A$ *returns a pair of* (pk_A, sk_A). *pk_A and sk_A are sender's public and secret keys, respectively. We write this as* $(pk_A, sk_A) \leftarrow \mathsf{Stg\_KGen}_A(cp)$.

- $\mathsf{Stg\_KGen}_B$: *The key generation algorithm for the receiver* $\mathsf{Stg\_KGen}_B$ *is a probabilistic algorithm. On input a common parameter sequence cp,* $\mathsf{Stg\_KGen}_B$ *returns a pair of* (pk_B, sk_B). *pk_B and sk_B are receiver's public and secret keys, respectively. We write this as* $(pk_B, sk_B) \leftarrow \mathsf{Stg\_KGen}_B(cp)$.

- $\mathsf{Stg\_Enc}$: *The steganographic encoding algorithm* $\mathsf{Stg\_Enc}$ *is a probabilistic algorithm.* $\mathsf{Stg\_Enc}$ *takes a common parameter cp, sender's secret key sk_A, receiver's public key pk_B, a message m, and a history h as inputs.* $\mathsf{Stg\_Enc}$ *has access to a channel oracle for some channel $\mathcal{C}$, which can sample from $\mathcal{C}_h$ for any h.* $\mathsf{Stg\_Enc}$ *returns a sequence of documents* $(s_1, s_2, \ldots, s_l)$ *from the support of $\mathcal{C}_h^l$. We write this as* $(s_1, s_2, \ldots, s_l) \leftarrow \mathsf{Stg\_Enc}(cp, sk_A, pk_B, m, h)$. *We call* $(s_1, s_2, \ldots, s_l)$ *a stegotext, and often simply write s.*

- $\mathsf{Stg\_Dec}$: $\mathsf{Stg\_Dec}$ *is a steganographic decoding algorithm. On input a common parameter cp, receiver's secret key sk_B, sender's public key pk_A, a stegotext* $s = (s_1, s_2, \ldots, s_l)$, *and a history h,* $\mathsf{Stg\_Dec}$ *returns either a message m or a symbol $\perp$ which indicates that the stegotext is invalid. We write this as* $m/\perp \leftarrow \mathsf{Stg\_Dec}(cp, sk_B, pk_A, s, h)$.

We denote $\mathcal{M}_{\mathcal{ASS}}$ as the message space of $\mathcal{ASS}$, and $\mathcal{H}$ as the set of legal histories with respect to $\mathcal{C}$. We require that for all cp which can be output by $\mathsf{Stg\_CGen}(1^k)$, for all (pk_A, sk_A) which can be output by $\mathsf{Stg\_KGen}_A(cp)$, for all (pk_B, sk_B) which can be output by $\mathsf{Stg\_KGen}_B(cp)$, for all $(m, h) \in \mathcal{M}_{\mathcal{ASS}} \times \mathcal{H}$, and for all stegotext s which can be output by $\mathsf{Stg\_Enc}(cp, sk_A, pk_B, m, h)$, there is a negligible function $\nu(k)$ such that $\Pr[m \leftarrow \mathsf{Stg\_Dec}(cp, sk_B, pk_A, s, h)] \geq 1 - \nu(k)$.

Steganographic Security. We define the security property of public-key steganography with authentication. First, we define the steganographic security, which is the property that any adversary eavesdropping the channel between the sender and the receiver cannot detect whether the target text conceals a hiddentext or not. Let $\mathcal{ASS} = (\mathsf{Stg\_CGen}, \mathsf{Stg\_KGen}_A, \mathsf{Stg\_KGen}_B, \mathsf{Stg\_Enc}, \mathsf{Stg\_Dec})$ be a scheme of public-key steganography with authentication and k a security parameter, and $\mathcal{C}$ a channel. Let ℓ^* be the function which implies the length of the stegotext. We define a distinguishing game under the chosen message-and-history attack against $\mathcal{ASS}$ by an adversary W_d and a challenger. We consider the experiments $\mathbf{Exp}^i_{\mathbf{CMSHA}}$ for $i \in \{0, 1\}$ as follows:

$\mathbf{Exp}^i_{\mathbf{CMSHA}}(1^k)$

1. $cp \leftarrow \mathsf{Stg\_CGen}(1^k), (pk_A, sk_A) \leftarrow \mathsf{Stg\_KGen}_A(cp)$, and $(pk_B, sk_B) \leftarrow \mathsf{Stg\_KGen}_B(cp)$.
2. W_d is given cp, pk_A, and pk_B.
3. W_d can make access to the steganographic-encoding oracle $\mathtt{STG\_ENC}_{sk_A,pk_B}$. W_d queries a message-history pair (m, h) to $\mathtt{STG\_ENC}_{sk_A,pk_B}$, and receives the corresponding stegotext s. W_d can also make access to the steganographic-decoding oracle $\mathtt{STG\_DEC}_{sk_B,pk_A}$. W_d queries a stegotext-history pair (s, h) to $\mathtt{STG\_DEC}_{sk_B,pk_A}$, and receives either the corresponding message m or a symbol $\perp$.
4. W_d produces a message m^* and a history h^* and passes it to the challenger. The challenger passes s^*_i to W_d.
5. W_d continues to query the encoding oracle $\mathtt{STG\_ENC}_{sk_A,pk_B}$ and the decoding oracle $\mathtt{STG\_DEC}_{sk_B,pk_A}$ with the restriction that W_d may not query s^*_i to $\mathtt{STG\_DEC}_{sk_B,pk_A}$.
6. W_d outputs a bit γ.
7. Return γ.

We define s_i for $i \in \{0, 1\}$ as follows:

- s_0: The challenger computes $s_0 \leftarrow \mathsf{Stg\_Enc}(cp, sk_A, pk_B, m^*, h^*)$.
- s_1: The challenger samples $s_1 \leftarrow \mathcal{C}^{\ell^*}_{h^*}$.

We define W_d's advantage against $\mathcal{ASS}$ with respect to $\mathcal{C}$ by

$$\mathbf{Adv}^{\mathsf{ss-cmsha}}_{\mathcal{ASS},\mathcal{C},W_d}(k) = |\Pr[\mathbf{Exp}^0_{\mathbf{CMSHA},\mathcal{C},W_d}(1^k)) = 1] - \Pr[\mathbf{Exp}^1_{\mathbf{CMSHA},\mathcal{C},W_d}(1^k)) = 1]|.$$

We denote the set of W_d who make q_{SE} steganographic encoding queries and q_{SD} steganographic decoding queries in running time t by $\mathcal{W}_d(t, q_{SE}, q_{SD})$.

Definition 7 (SS-CMSHA). *We say that $\mathcal{ASS}$ is steganographically secure under the chosen message/stegotext-and-history attack with respect to $\mathcal{C}$* (SS-CMSHA) *if for any probabilistic polynomial adversary $W_d \in \mathcal{W}_d(t, q_{SE}, q_{SD})$, $\mathbf{Adv}^{\mathsf{ss-cmsha}}_{\mathcal{ASS},\mathcal{C},W_d}(k)$ is negligible in k.*

Unforgeability. Second, we define the unforgeability, which is the property that anyone except the sender cannot forge the stegotext which will be accepted by the receiver so that it contains some hiddentext from the sender. Let $\mathcal{ASS} = (\mathsf{Stg\_CGen}, \mathsf{Stg\_KGen}_A, \mathsf{Stg\_KGen}_B, \mathsf{Stg\_Enc}, \mathsf{Stg\_Dec})$ be a scheme of public-key steganography with authentication and k a security parameter, and $\mathcal{C}$ a channel. Let W_f be an adversary who forges the stegotext. We consider the experiments $\mathbf{Exp}_{\mathbf{CMHA}}$ as follows:

$\mathbf{Exp}_{\mathbf{CMHA}}(1^k)$

1. $cp \leftarrow \mathsf{Stg\_CGen}(1^k)$ and $(pk_A, sk_A) \leftarrow \mathsf{Stg\_KGen}_A(cp)$.
2. W_f is given cp and pk_A.
3. W_f queries (m, h, pk_R), where m is a message, h is a history, and pk_R is receiver's public key which W_f chooses arbitrarily, to the steganographic-encoding oracle $\mathtt{STG\_ENC}_{sk_A,pk_R}$, and receives the stegotext, adaptively.
4. In the end, W_f outputs $(s^*, h^*, pk_R^*, sk_R^*)$. Let $m^* \leftarrow \mathsf{Stg\_Dec}(cp, sk_R^*, pk_A, s^*, h^*)$. If $m^* \neq \perp$ and W_f has not queried $(m^*, \cdot, \cdot)$, then return 1. Otherwise, return 0.

We define W_f's advantage against $\mathcal{ASS}$ with respect to $\mathcal{C}$ by

$$\mathbf{Adv}^{\mathsf{euf-cmha}}_{\mathcal{ASS},\mathcal{C},W_f}(k) = \Pr[\mathbf{Exp}^{\mathbf{CMHA}}_{\mathcal{C},W_f}(1^k) = 1].$$

We denote the set of W_f who make q_{SE} steganographic encoding queries in running time t by $\mathcal{W}_f(t, q_{SE})$.

Definition 8 (EUF-CMHA). *We say that $\mathcal{ASS}$ is existentially unforgeable under the chosen message-and-history attack with respect to $\mathcal{C}$* (EUF-CMHA) *if for any probabilistic polynomial adversary $W_f \in \mathcal{W}_f(t, q_{SE})$, $\mathbf{Adv}^{\mathsf{euf-cmha}}_{\mathcal{ASS},\mathcal{C},W_f}(k)$ is negligible in k.*

4 The Construction

In this section, We construct a concrete scheme of public-key steganography with authentication via a public-key encryption scheme and a digital signature scheme. We employ the idea for constructing the public-key steganography by von Ahn and Hopper [11], Backes and Cachin [3], and Hopper [6]. We show that our scheme is steganographically secure if the underlying public-key encryption scheme satisfies IND$-CCA. We also show that our scheme is unforgeable if the underlying digital signature scheme satisfies EUF-CMA.

4.1 Preparations

In this section, we prepare to construct a scheme with public-key steganography with authentication by using a public-key encryption and a digital signature scheme. We employ the idea for constructing the public-key steganography by von Ahn and Hopper [11], and Backes and Cachin [3]. Let $f : D \to \{0,1\}$ be a hash function. Then, the following procedures Basic_Encode and Basic_Decode have been proposed [11, 3].

Procedure Basic_Encode:
Input: target $c \in \{0,1\}^a$, history h, bound k
Parse c as $c_1||c_2||\cdots||c_a$, where $c_i \in \{0,1\}$ for $1 \leq i \leq a$
for $i = 1, 2, \ldots, a$ do
  Let $j = 0$
  repeat:
    sample $s_i \leftarrow \mathcal{C}_h$, increment j
  until $f(s_i) = c_i$ or $j > k$
  set $h = h||s_i$.
Output: $s_1, s_2, \ldots, s_a$

Procedure Basic_Decode:
Input: documents $s_1, s_2, \ldots, s_a$ where $s_i \in D$ for $1 \leq i \leq a$
for $i = 1, 2, \ldots, a$ do
  compute $c_i = f(s_i)$
set $c = c_1||c_2||\cdots||c_a$.
Output: c

In this construction, $\mathsf{Basic\_Decode}(\mathsf{Basic\_Encode}(c,h,k))$ might not equal c. We call this the *encoding error* and denote EE_B as the event that the encoding error happens through the procedure.

Proposition 9. *If f is ϵ-biased on $\mathcal{C}_h$ for all h, then for all c and h, there exists a negligible function $\hat{\epsilon}$ in k such that* $\Pr[\mathsf{EE}_\mathsf{B}] \leq \hat{\epsilon}$.

Proof. We assume that $|c| = 1$. Since f is ϵ-biased on $\mathcal{C}_h$ for all h, we have that $\Pr[\mathsf{Basic\_Decode}(\mathsf{Basic\_Encode}(c,h,k)) \neq c] < (\frac{1}{2}+\epsilon)^k$ where $c \in \{0,1\}$. Considering the case that $|c|$ is polynomial, $\Pr[\mathsf{EE}_\mathsf{B}] < |c|(\frac{1}{2}+\epsilon)^k$ by the union bound. We denote $|c|(\frac{1}{2}+\epsilon)^k$ by $\hat{\epsilon}$, which is negligible in k. □

We denote $\mathtt{A}(s,a,h,k)$ as $\Pr[\mathsf{Basic\_Encode}(c,h,k) = s_1,s_2,\ldots,s_a | c \leftarrow \mathcal{U}_a]$, and $\mathtt{B}(s,a,h,k)$ as $\Pr[s = s_1,s_2,\ldots,s_a | s \leftarrow \mathcal{C}_h^a]$. If f is perfectly unbiased on $\mathcal{C}_h$ for all h, then $\mathtt{A}(s,a,h,k) = \mathtt{B}(s,a,h,k)$. In other words, the distribution of output of $\mathsf{Basic\_Encode}(c,h,k)$ where c is chosen randomly from $\{0,1\}^a$ is identical with that of documents according to $\mathcal{C}_h^a$. To keep more general, we consider the case that f is ϵ-biased on $\mathcal{C}_h$ for all h.

Proposition 10. *If f is ϵ-biased on $\mathcal{C}_h$ for all h, then for any k and $s = s_1,s_2,\ldots,s_a$,* $|\mathtt{A}(s,a,h,k) - \mathtt{B}(s,a,h,k)| \leq a\epsilon$.

Proof. We assume that $a = 1$. Since f is ϵ-biased on $\mathcal{C}_h$ for all h, we also assume without loss of generality that $\Pr[f(x) = 0 | x \leftarrow \mathcal{C}_h] = \frac{1}{2} + \delta$ where $0 \leq \delta < \epsilon$. Furthermore, we assume that $f(s_1) = 0$. We note that $0 \leq \Pr[s_1 \leftarrow \mathcal{C}_h] \leq \frac{1}{2} + \delta$. Since c is chosen randomly from $\{0,1\}$, we have that

$$\mathtt{A}(s,1,h,k) = \frac{1}{2}\Big\{\sum_{i=0}^{k-1}\Big(\frac{1}{2}-\delta\Big)^i \Pr[s_1 \leftarrow \mathcal{C}_h]\Big\} + \frac{1}{2}\Big(\frac{1}{2}+\delta\Big)^{k-1}\Pr[s_1 \leftarrow \mathcal{C}_h].$$

On the other hand, $\mathtt{B}(s,1,h,k) = \Pr[s_1 \leftarrow \mathcal{C}_h]$. Thus,

$$\begin{aligned}
|\mathtt{A}(s,1,h,k) - \mathtt{B}(s,1,h,k)| &= \Pr[s_1 \leftarrow \mathcal{C}_h]\Big|\frac{1}{2}\Big\{\sum_{i=0}^{k-1}\Big(\frac{1}{2}-\delta\Big)^i\Big\} + \frac{1}{2}\Big(\frac{1}{2}+\delta\Big)^{k-1} - 1\Big| \\
&\leq \Big(\frac{1}{2}+\delta\Big)\Big|\frac{1}{2}\Big\{\sum_{i=0}^{k-1}\Big(\frac{1}{2}-\delta\Big)^i\Big\} + \frac{1}{2}\Big(\frac{1}{2}+\delta\Big)^{k-1} - 1\Big| \\
&\leq \Big|\frac{1}{2}\Big\{\Big(\frac{1}{2}+\delta\Big)^k - \Big(\frac{1}{2}-\delta\Big)^k\Big\} - \delta\Big| \leq \delta < \epsilon.
\end{aligned}$$

In the case $f(s_1) = 1$, we can analyze the same as above. Furthermore, we can consider the case $a > 1$ by applying the hybrid arguments to the analysis, and get the claimed result. □

In order to choose the hash function f, we can apply the notion of the universal family of hash function.

Proposition 11. *Let $\mathcal{F}$ be a strongly universal family of hash functions $D \to \{0,1\}$ and F a random variable with the uniform distribution on $\mathcal{F}$. Let $h_1,\ldots,h_m$ be any sequence of legal histories. For $1 \leq i \leq m$, let A_i be a random variable with the distribution $\mathcal{C}_{h_i}$ and B_i a random variable with the uniform distribution on $\{0,1\}$. Let $L(k) = \min_i H_\infty(\mathcal{C}_{h_i})$. Then,*

$$\Delta[F, F(A_1),\ldots,F(A_m); F, B_1,\ldots,B_m] \leq m2^{-L(k)/2}.$$

The proposition is a direct consequence of the leftover hash lemma [10]. We assume that for every legal history h, $H_\infty(\mathcal{C}_h) \in \omega(\log k)$. Then from this proposition, for every polynomial p and any h, $\mathsf{Basic\_Encode}$ can operate on samples from $\mathcal{C}_h^p$ and induce only a negligible statistical distance in its output distribution.

Hopper [6] proposed the deterministic-encoding procedure $\mathsf{D\_Encode}$ for constructing the scheme of public-key steganography satisfying stronger security. Let $f : D \to \{0,1\}$ be a hash function. Then, the procedure $\mathsf{D\_Encode}$ is as follows:

Procedure $\mathsf{D\_Encode}$:
Input: target $c \in \{0,1\}^a$, history h, bound k, randomness $r_1, \ldots, r_{ak}$, where $r_i \in \{0,1\}^k$ for $1 \le i \le a$
Parse c as $c_1||c_2||\cdots||c_a$, where $c_i \in \{0,1\}$ for $1 \le i \le a$
Let $b = 0$; for $i = 1, 2, \ldots, a$ do
  Let $j = 0$
  repeat:
    compute $s_i = \mathsf{channel}(h, r_b)$, increment j, b
  until $f(s_i) = c_i$ or $j > k$
  set $h = h||s_i$.
Output: $s_1, s_2, \ldots, s_a$

In this construction, $\mathsf{Basic\_Decode}(\mathsf{D\_Encode}(c, h, k, R))$ where $R \leftarrow \mathcal{U}_{ak}$ might not equal c. We call this the *encoding error* and denote EE_D as the event that the encoding error happens through the procedure.

Proposition 12. *If f is ϵ-biased with respect to* $\mathsf{channel}$ *for all h, then for all c and h, there exists a negligible function $\hat{\epsilon}$ in k such that* $\Pr[\mathsf{EE}_\mathsf{D}] \le \hat{\epsilon}$.

Proof. The proof is similar to that for Proposition 9. □

4.2 The Algorithm

We now propose a scheme of public-key steganography with authentication via a public-key encryption scheme, a digital signature scheme, and a procedure $\mathsf{D\_Encode}$.

Definition 13. *Our proposed scheme of public-key steganography with authentication* $\mathcal{ASS} = (\mathsf{Stg\_CGen}, \mathsf{Stg\_KGen}_A, \mathsf{Stg\_KGen}_B, \mathsf{Stg\_Enc}, \mathsf{Stg\_Dec})$ *is as follows. Let* $\mathcal{PE} = (\mathsf{Enc\_CGen}, \mathsf{Enc\_KGen}, \mathsf{Enc}, \mathsf{Dec})$ *be a public-key encryption scheme whose message space is $\mathcal{M}_{\mathcal{PE}}$ and the length of the ciphertext generated by $\mathcal{PE}$ is $l_{\mathcal{PE}}$, and* $\mathcal{DS} = (\mathsf{Sig\_CGen}, \mathsf{Sig\_KGen}, \mathsf{Sig}, \mathsf{Ver})$ *a digital signature scheme whose message space and signature space are $\mathcal{M}_{\mathcal{DS}}$ and $\mathcal{S}_{\mathcal{DS}}$, respectively, and the length of the ciphertext generated by $\mathcal{DS}$ is $l_{\mathcal{DS}}$. We assume that both the security parameter of $\mathcal{PE}$ and that of $\mathcal{DS}$ are k, then the security parameter of $\mathcal{ASS}$ is the same k. We also assume that $\mathcal{M}_{\mathcal{PE}} = \mathcal{M}_{\mathcal{DS}} \times \{0,1\}^k \times \mathcal{S}_{\mathcal{DS}}$. Let $f : D \to \{0,1\}$ be a hash function and $G : \{0,1\}^k \to \{0,1\}^{k \times l_{\mathcal{PE}} k}$ a pseudorandom generator.*

Algorithm $\mathsf{Stg\_CGen}$:
Input: 1^k
Compute $cp_{enc} \leftarrow \mathsf{Enc\_CGen}(1^k)$
Compute $cp_{sig} \leftarrow \mathsf{Sig\_CGen}(1^k)$
Output: (cp_{enc}, cp_{sig})

Algorithm $\mathsf{Stg\_KGen}_A$:

Input: cp_{enc}, cp_{sig}
Compute $(pk_A, sk_A) \leftarrow \mathsf{Sig\_KGen}(cp_{sig})$
Output: (pk_A, sk_A)

Algorithm $\mathsf{Stg\_KGen}_B$:
Input: cp_{enc}, cp_{sig}
Compute $(pk_B, sk_B) \leftarrow \mathsf{Enc\_KGen}(cp_{enc})$
Output: (pk_B, sk_B)

Algorithm Stg_Enc:
Input: $cp_{enc}, cp_{sig}, sk_A, pk_B, m, h$
Choose $r \leftarrow \mathcal{U}_k$
Compute $\sigma \leftarrow \mathsf{Sig}(cp_{sig}, sk_A, m||r)$
Compute $c \leftarrow \mathsf{Enc}(cp_{enc}, pk_B, m||r||\sigma)$
Compute $s = \mathsf{D\_Encode}(c, h, k, G(r))$
Output: s

Algorithm Stg_Dec:
Input: $cp_{enc}, cp_{sig}, sk_B, pk_A, s$ *(where* $s = (s_1, s_2, \ldots, s_{l_{\mathcal{PE}}}))$, h
Compute $c \leftarrow \mathsf{Basic\_Decode}(s)$
Compute $M \leftarrow \mathsf{Dec}(cp_{enc}, sk_B, c_1||c_2|| \cdots ||c_{l_{\mathcal{PE}}})$
if $M = \perp$ *then return* $\perp$
Parse $M = m||r||\sigma$ *where* $|r| = k$ *and* $|\sigma| = l_{\mathcal{DS}}$
Compute $d \leftarrow \mathsf{Ver}(cp_{sig}, pk_A, m||r, \sigma)$
if $d \neq 1$ *then return* $\perp$
if $s \neq \mathsf{DEncode}(c, h, k, G(r))$ *then return* $\perp$
Output: m

4.3 Security Proofs

In this section, we give the security proofs for our scheme. First, We show that our scheme is steganographically secure if $\mathcal{PE}$ satisfies IND$-CCA.

Theorem 14. *Let* $\mathcal{PE}$ *be a public-key encryption scheme,* $\mathcal{DS}$ *a digital signature scheme, and* $\mathcal{ASS}$ *our proposed scheme of public-key steganography with authentication via* $\mathcal{PE}$ *and* $\mathcal{DS}$*. We assume that a hash function* $f : D \to \{0,1\}$ *is* ϵ*-biased on* $\mathcal{C}_h$ *for all* h *where* ϵ *is negligible in* k*. We also assume that* f *is* $\acute{\epsilon}$*-biased with respect to* channel *for all* h *where* $\acute{\epsilon}$ *is negligible in* k*. If* $\mathcal{PE}$ *satisfies* IND$-CCA *and* G *is a pseudorandom generator, then* $\mathcal{ASS}$ *satisfies the steganographic security for any* $\mathcal{C}$*.*

Proof. Let W_d be an adversary in $\mathcal{W}_d(t, q_{SE}, q_{SD})$ who breaks the steganographic security of $\mathcal{ASS}$ with respect to $\mathcal{C}$. We consider the experiments $\mathbf{Exp}^i_{\mathbf{PS}}$ for $i \in \{0, 1, 2, 3, 4\}$ as follows:

$\mathbf{Exp}^i_{\mathbf{PS}}(1^k)$

1. $cp_{enc} \leftarrow \mathsf{Enc\_CGen}(1^k)$, $cp_{sig} \leftarrow \mathsf{Sig\_CGen}(1^k)$, $(pk_A, sk_A) \leftarrow \mathsf{Sig\_KGen}(cp_{sig})$, and $(pk_B, sk_B) \leftarrow \mathsf{Enc\_KGen}(cp_{enc})$.
2. W_d is given cp_{enc}, cp_{sig}, pk_A, and pk_B.
3. W_d can make access to the steganographic-encoding oracle $\mathtt{STG\_ENC}_{sk_A, pk_B}$. W_d queries a message-history pair (m, h) to $\mathtt{STG\_ENC}_{sk_A, pk_B}$, and receives the corresponding stegotext s. W_d can also make access to the steganographic-decoding or-

acle $\mathtt{STG\_DEC}_{sk_B,pk_A}$. W_d queries a stegotext-history pair (s,h) to $\mathtt{STG\_DEC}_{sk_B,pk_A}$, and receives either the corresponding message m or a symbol $\perp$.

4. W_d produces a message m^* and a history h^* and passes it to the challenger. The challenger passes s_i^* to W_d.
5. W_d continues to query the encoding oracle $\mathtt{STG\_ENC}_{sk_A,pk_B}$ and the decoding oracle $\mathtt{STG\_DEC}_{sk_B,pk_A}$ with the restriction that W_d may not query s_i^* to $\mathtt{STG\_DEC}_{sk_B,pk_A}$.
6. W_d outputs a bit γ.
7. Return γ.

We define s_i for $i \in \{0,1,2,3,4\}$ as follows:

- s_0^*: The challenger chooses $r^* \leftarrow \mathcal{U}_K$. The challenger computes $\sigma^* \leftarrow \mathsf{Sig}(cp_{sig}, sk_A, m^*||r^*)$, $c^* \leftarrow \mathsf{Enc}(cp_{enc}, pk_A, m^*||r^*||\sigma^*)$, and $s_0^* = \mathsf{DEncode}(c^*, h^*, k, G(r^*))$.
- s_1^*: The challenger chooses $C^* \leftarrow \mathcal{U}_{l_{\mathcal{PE}}}$. The challenger chooses $r^* \leftarrow \mathcal{U}_k$ and computes $s_1^* = \mathsf{DEncode}(C^*, h^*, k, G(r^*))$.
- s_2^*: The challenger chooses $C^* \leftarrow \mathcal{U}_{l_{\mathcal{PE}}}$. The challenger chooses $R^* \leftarrow \mathcal{U}_{k \times l_{\mathcal{PE}} k}$ and computes $s_2^* = \mathsf{DEncode}(C^*, h^*, k, R^*)$.
- s_3^*: The challenger chooses $C^* \leftarrow \mathcal{U}_{l_{\mathcal{PE}}}$ and computes $s_3^* = \mathsf{Basic\_Encode}(C^*, h^*, k)$.
- s_4^*: The challenger samples $s_4^* \leftarrow \mathcal{C}_{h^*}^{l_{\mathcal{PE}}}$.

We denote the set of W_d who make q_{SE} steganographic-encoding queries and q_{SD} steganographic-decoding queries in running time t by $\mathcal{W}_d(t, q_{SE}, q_{SD})$. Let $\mathbf{Adv}^i_{\mathcal{C},W_d}(k) = |\Pr[\mathbf{Exp}^i_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1] - \Pr[\mathbf{Exp}^{i+1}_{\mathbf{PS},\mathcal{C},W_d}(1^k) = 1]|$. Then we have that

$$\begin{aligned}
\mathbf{Adv}^{\mathsf{ss-cmsha}}_{\mathcal{ASS},\mathcal{C},W_d}(k) &= |\Pr[\mathbf{Exp}^0_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1] - \Pr[\mathbf{Exp}^4_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1]| \\
&\le |\Pr[\mathbf{Exp}^0_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1] - \Pr[\mathbf{Exp}^1_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1]| \\
&\quad + |\Pr[\mathbf{Exp}^1_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1] - \Pr[\mathbf{Exp}^2_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1]| \\
&\quad + |\Pr[\mathbf{Exp}^2_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1] - \Pr[\mathbf{Exp}^3_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1]| \\
&\quad + |\Pr[\mathbf{Exp}^3_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1] - \Pr[\mathbf{Exp}^4_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1]| \\
&= \mathbf{Adv}^0_{\mathcal{C},W_d}(k) + \mathbf{Adv}^1_{\mathcal{C},W_d}(k) + \mathbf{Adv}^2_{\mathcal{C},W_d}(k) + \mathbf{Adv}^3_{\mathcal{C},W_d}(k).
\end{aligned}$$

From Proposition 10, $\mathbf{Adv}^3_{\mathcal{C},W_d}(k) \le l_{\mathcal{PE}}\epsilon$. From the definition of $\mathsf{channel}$, there exists a negligible function $\tilde{\epsilon}$ in k such that $\mathbf{Adv}^2_{\mathcal{C},W_d}(k) \le \tilde{\epsilon}$. Applying these and the following lemmas, we get the claimed result.

Lemma 15. *For some $\mathcal{C}$, if there exists $W_d \in \mathcal{W}_d(t, q_{SE}, q_{SD})$, then there exists $A_d \in \mathcal{A}_d(t + O(q_{SE}) + O(q_{SD}), q_{SD})$ such that*

$$\mathbf{Adv}^{\mathsf{ind\$-cca}}_{\mathcal{PE},A_d}(k) = (1-\hat{\epsilon})\mathbf{Adv}^0_{\mathcal{C},W_d}(k),$$

where $\hat{\epsilon}$ is negligible in k and t_f is the time of computing the hash function f.

Proof. We construct an adversary A_d attacking the indistinguishability from random bits of $\mathcal{PE}$ by using W_d.

A_d takes a common parameter cp_{enc} and a public key pk_B where $cp_{enc} \leftarrow \mathsf{Enc\_CGen}(1^k)$ and $(pk_B, sk_B) \leftarrow \mathsf{Enc\_KGen}(cp_{enc})$. A_d runs $cp_{sig} \leftarrow \mathsf{Sig\_CGen}(1^k)$ and $(pk_A, sk_A) \leftarrow \mathsf{Sig\_KGen}(cp_{sig})$, and gets cp_{sig} and (pk_A, sk_A). Then A_d passes W_d (cp_{enc}, cp_{sig}), pk_A, and pk_B as a common parameter, sender's public key, and receiver's public key, respectively.

If W_d makes a steganographic-encoding query (m, h), A_d chooses $r \leftarrow \mathcal{U}_k$ and computes $\sigma \leftarrow \mathsf{Sig}(cp_{sig}, sk_A, m||r)$, $c \leftarrow \mathsf{Enc}(cp_{enc}, pk_B, m||r||\sigma)$, and $s \leftarrow \mathsf{D\_Encode}(c, h, k, G(r))$. Then A_d passes s to W_d.

If W_d makes a steganographic-decoding query (s, h), A_d computes $\hat{c} = \mathsf{Basic\_Decode}(s)$ and queries $\hat{c}$ to A_d's decoding oracle $\mathtt{DEC}_{sk_A}$ and receives $\hat{M}$. If $\hat{M} = \perp$, then A_d returns $\perp$ to W_d. Otherwise, A_d parses $\hat{M} = m||r||\sigma$ where $|r| = k$ and $|\sigma| = l_{\mathcal{DS}}$ and computes $d \leftarrow \mathsf{Ver}(cp_{sig}, pk_A, m||r, \sigma)$. If $d \neq 1$, then A_d returns $\perp$ to W_d. Otherwise, A_d computes $\hat{s} = \mathsf{DEncode}(\hat{c}, h, k, G(r))$. If $\hat{s} \neq s$, then A_d returns $\perp$ to W_d. Otherwise, A_d passes m to W_d.

If W_d outputs (m^*, h^*) as W_d's challenge, A_d chooses $r \leftarrow \mathcal{U}_k$ and computes $\sigma^* = \mathsf{Sig}(cp_{sig}, sk_A, m^*||r^*)$ Then A_d outputs $m^*||r^*||\sigma^*$ to A_d's challenge oracle and receives c^*. A_d computes $s^* = \mathsf{D\_Encode}(c^*, h^*, k, G(r^*))$ and passes s^* to W_d.

After the challenge phase, A_d continues to respond steganographic-encoding and steganographic-decoding queries of W_d as before with the following restriction. In the steganographic-decoding query phase, if $\hat{c}$ is identical with c^*, then A_d returns $\perp$ to W_d.

However, if the encoding error happens in this simulation, A_d stops this attack.

Finally, if W_d outputs a bit γ, A_d outputs the same bit γ.

We note that when A_d is given a ciphertext of $m^*||r^*||\sigma^*$, A_d perfectly simulates $\mathbf{Exp}^0_{\mathbf{PS},\mathcal{C},W_d}(1^k)$ for W_d. Therefore $\Pr[\mathbf{Exp}^0_{\mathbf{CCA},A_d}(1^k) = 1] = \Pr[\overline{\mathsf{EE_D}}] \Pr[\mathbf{Exp}^0_{\mathbf{PS},\mathcal{C},W_d}(1^k) = 1]$. On the other hand, when A_d is given a random string, A_d perfectly simulates $\mathbf{Exp}^1_{\mathbf{PS},\mathcal{C},W_d}(1^k)$ for W_d. Therefore $\Pr[\mathbf{Exp}^1_{\mathbf{CCA},A_d}(1^k) = 1] = \Pr[\overline{\mathsf{EE_D}}] \Pr[\mathbf{Exp}^1_{\mathbf{PS},\mathcal{C},W_d}(1^k) = 1]$. From Proposition 12, there exists a negligible $\hat{\epsilon}$ such that $\Pr[\overline{\mathsf{EE_D}}] = 1 - \hat{\epsilon}$. Hence,

$$\begin{aligned}
\mathbf{Adv}^{\mathsf{ind\$-cca}}_{\mathcal{PE},A_d}(k) &= |\Pr[\mathbf{Exp}^0_{\mathbf{CCA},A_d}(1^k) = 1] - \Pr[\mathbf{Exp}^1_{\mathbf{CCA},A_d}(1^k) = 1]| \\
&= |\Pr[\overline{\mathsf{EE_D}}] \Pr[\mathbf{Exp}^0_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1] - \Pr[\overline{\mathsf{EE_D}}] \Pr[\mathbf{Exp}^1_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1]| \\
&= (1 - \hat{\epsilon})\mathbf{Adv}^0_{\mathcal{C},W_d}(k).
\end{aligned}$$

□

Lemma 16. *For some $\mathcal{C}$, if there exists $W_d \in \mathcal{W}_d(t, q_{SE}, q_{SD})$, then there exists $A_{pr} \in \mathcal{A}_{pr}(t + O(q_{SE}) + O(q_{SD}))$ such that*

$$\mathbf{Adv}^{\mathsf{prg}}_{G,A_{pr}}(k) = (1 - \hat{\epsilon})\mathbf{Adv}^1_{\mathcal{C},W_d}(k),$$

where $\hat{\epsilon}$ is negligible in k.

Proof. We construct an adversary A_{pr} attacking the randomness of G by using W_d.

A_{pr} takes as input $z^* \in \{0,1\}^{k \times l_{\mathcal{PE}} k}$ from A_{pr}'s challenger. A_{pr} runs $cp_{enc} \leftarrow \mathsf{Enc\_CGen}(1^k)$, $(pk_B, sk_B) \leftarrow \mathsf{Enc\_KGen}(cp_{enc})$, $cp_{sig} \leftarrow \mathsf{Sig\_CGen}(1^k)$, and $(pk_A, sk_A) \leftarrow \mathsf{Sig\_KGen}(cp_{sig})$. Then A_{pr} passes W_d (cp_{enc}, cp_{sig}), pk_A, and pk_B as a common parameter, sender's public key, and receiver's public key, respectively.

If W_d makes a steganographic-encoding query (m, h), A_{pr} chooses $r \leftarrow \mathcal{U}_k$ and computes $\sigma \leftarrow \mathsf{Sig}(cp_{sig}, sk_A, m||r)$, $c \leftarrow \mathsf{Enc}(cp_{enc}, pk_B, m||r||\sigma)$, and $s \leftarrow \mathsf{D\_Encode}(c, h, k, G(r))$. Then A_{pr} passes s to W_d.

If W_d makes a steganographic-decoding query (s, h), A_{pr} computes $\hat{c} = \mathsf{Basic\_Decode}(s)$ and $\hat{M} \leftarrow \mathsf{Dec}(cp_{enc}, sk_B, s)$. If $\hat{M} = \perp$, then A_{pr} returns $\perp$ to W_d. Otherwise, A_{pr} parses $\hat{M} = m||r||\sigma$ where $|r| = k$ and $|\sigma| = l_{\mathcal{DS}}$ and computes $d \leftarrow \mathsf{Ver}(cp_{sig}, pk_A, m||r, \sigma)$. If $d \neq 1$, then A_{pr} returns $\perp$ to W_d. Otherwise, A_{pr} computes $\hat{s} = \mathsf{DEncode}(\hat{c}, h, k, G(r))$. If $\hat{s} \neq s$, then A_{pr} returns $\perp$ to W_d. Otherwise, A_{pr} passes m to W_d.

If W_d outputs (m^*, h^*) as W_d's challenge, A_{pr} chooses $C^* \leftarrow \mathcal{U}_{l_{\mathcal{PE}}}$ and computes $s^* = \mathsf{D\_Encode}(C^*, h^*, k, z^*)$, and passes s^* to W_d.

After the challenge phase, A_{pr} continues to respond steganographic-encoding and steganographic-decoding queries of W_d as before with the following restriction. In the steganographic-decoding query phase, if $\hat{c}$ is identical with c^*, then A_{pr} returns $\perp$ to W_d.

However, if the encoding error happens in this simulation, A_{pr} stops this attack.

Finally, if W_d outputs a bit γ, A_{pr} outputs the same bit γ.

We note that if z^* is the value $G(r^*)$ where $r^* \leftarrow \mathcal{U}_k$, A_{pr} perfectly simulates $\mathbf{Exp}^0_{\mathbf{PS},\mathcal{C},W_d}(1^k)$ for W_d. Therefore $\Pr[\mathbf{Exp}^0_{\mathbf{PRG},G,A_{pr}}(1^k) = 1] = \Pr[\overline{\mathsf{EE}_\mathsf{D}}]\Pr[\mathbf{Exp}^1_{\mathbf{PS},\mathcal{C},W_d}(1^k) = 1]$. On the other hand, if z^* is chosen randomly from $\{0,1\}^{k \times l_{\mathcal{PE}} k}$, A_d perfectly simulates $\mathbf{Exp}^2_{\mathbf{PS},\mathcal{C},W_d}(1^k)$ for W_d. Therefore $\Pr[\mathbf{Exp}^1_{\mathbf{PRG},G,A_{pr}}(1^k) = 1] = \Pr[\overline{\mathsf{EE}_\mathsf{D}}]\Pr[\mathbf{Exp}^2_{\mathbf{PS},\mathcal{C},W_d}(1^k) = 1]$. From Proposition 12, there exists a negligible $\hat{\epsilon}$ such that $\Pr[\overline{\mathsf{EE}_\mathsf{D}}] = 1 - \hat{\epsilon}$. Hence,

$$\begin{aligned}
\mathbf{Adv}^{\mathsf{prg}}_{G,A_{pr}}(k) &= |\Pr[\mathbf{Exp}^0_{\mathbf{PRG},G,A_{pr}}(1^k) = 1] - \Pr[\mathbf{Exp}^1_{\mathbf{PRG},G,A_{pr}}(1^k) = 1]| \\
&= |\Pr[\overline{\mathsf{EE}_\mathsf{D}}]\Pr[\mathbf{Exp}^1_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1] - \Pr[\overline{\mathsf{EE}_\mathsf{D}}]\Pr[\mathbf{Exp}^2_{\mathbf{PS},\mathcal{C},W_d}(1^k)) = 1]| \\
&= (1 - \hat{\epsilon})\mathbf{Adv}^1_{\mathcal{C},W_d}(k).
\end{aligned}$$

□

□

Second, we show that our scheme is unforgeable if $\mathcal{DS}$ satisfies $\mathsf{EUF\text{-}CMA}$.

Theorem 17. *Let $\mathcal{PS}$ be a public-key encryption scheme, $\mathcal{DS}$ a digital signature scheme, and $\mathcal{ASS}$ our proposed scheme of public-key steganography with authentication via $\mathcal{PE}$ and $\mathcal{DS}$. We assume that a hash function $f : D \to \{0,1\}$ is ϵ-biased on $\mathcal{C}$ for all h where ϵ is negligible in k. We also assume that f is $\hat{\epsilon}$-biased with respect to* channel *for all h where $\hat{\epsilon}$ is negligible in k. If $\mathcal{DS}$ satisfies* EUF-CMA*, then $\mathcal{ASS}$ satisfies the unforgeability for any $\mathcal{C}$. In particular, for some $\mathcal{C}$, if there exists $W_f \in \mathcal{W}_f(t, q_{SE})$, then there exists $A_f \in \mathcal{A}_f(t + O(q_{SE}), q_{SE})$ such that*

$$\mathbf{Adv}^{\mathsf{euf-cma}}_{\mathcal{DS},A_f}(k) = (1 - \hat{\epsilon})\mathbf{Adv}^{\mathsf{euf-cmha}}_{\mathcal{ASS},\mathcal{C},W_f}(k),$$

where $\hat{\epsilon}$ is negligible in k.

Proof. We construct an adversary A_f who forges a signature of $\mathcal{DS}$ by using W_f.

A_f takes a common parameter cp_{sig} and public key pk_A where $cp_{sig} \leftarrow \mathsf{Sig\_CGen}(1^k)$ and $(pk_A, sk_A) \leftarrow \mathsf{Sig\_KGen}(cp_{sig})$, A_f runs $cp_{enc} \leftarrow \mathsf{Enc\_CGen}(1^k)$. Then A_f passes W_f (cp_{enc}, cp_{sig}) and pk_A as a common parameter and sender's public key, respectively.

If W_f makes a steganographic-encoding query (m, h, pk_R), A_f chooses $r \leftarrow \mathcal{U}_k$ and queries $m||r$ to A_f's signing oracle and receives σ which is a signature for $m||r$. A_f computes $c \leftarrow \mathsf{Enc}(cp_{enc}, pk_R, m||r||\sigma)$ and $s = \mathsf{D\_Encode}(c, h, k, G(r))$, and passes s to W_f.

However, if the encoding error happens in this simulation, A_f stops this attack.

W_f outputs $(s^*, h^*, pk^*_R, sk^*_R)$. A_f computes $c^* = \mathsf{Basic\_Decode}(s^*)$, $M^* \leftarrow \mathsf{Dec}(cp_{enc}, sk^*_R, s^*)$, and phases $M^* = m^*||r^*||\sigma^*$ where $|r^*| = k$ and $|\sigma^*| = l_{\mathcal{DS}}$. Then A_f outputs $(m^*||r^*, \sigma^*)$. By construction, if s^* is a valid stegotext for (m^*, h^*), then σ^* is a valid signature for $m^*||r^*$. From Proposition 12, there exists a negligible $\hat{\epsilon}$ such that $\Pr[\overline{\mathsf{EE}_\mathsf{D}}] = 1 - \hat{\epsilon}$. Thus, $\mathbf{Adv}^{\mathsf{euf-cma}}_{\mathcal{DS},A_f}(k) = \Pr[\overline{\mathsf{EE}_\mathsf{D}}]\mathbf{Adv}^{\mathsf{euf-cmha}}_{\mathcal{ASS},\mathcal{C},W_f}(k) = (1 - \hat{\epsilon})\mathbf{Adv}^{\mathsf{euf-cmha}}_{\mathcal{ASS},\mathcal{C},W_f}(k)$. □

As a result, we can show the following theorem.

Theorem 18. *Let $\mathcal{PS}$ be a public-key encryption scheme, $\mathcal{DS}$ a digital signature scheme, and $\mathcal{ASS}$ our proposed scheme of public-key steganography with authentication via $\mathcal{PE}$ and $\mathcal{DS}$. We assume that a hash function $f : D \to \{0,1\}$ is ϵ-biased on $\mathcal{C}$ for all h where ϵ is negligible in k. We also assume that f is $\hat{\epsilon}$-biased with respect to* channel *for all h where $\hat{\epsilon}$ is negligible in k. If $\mathcal{PE}$ satisfies* IND$-CCA, *$\mathcal{DS}$ satisfies* EUF-CMA*, and G is a pseudorandom generator, then $\mathcal{ASS}$ satisfies the steganographic security and the unforgeability for any $\mathcal{C}$.*

Proof. If $\mathcal{ASS}$ does not satisfy SS-CMSHA, then we have that either $\mathcal{PE}$ does not satisfy IND$-CCA or G is not a pseudorandom generator because of Theorem 14. If $\mathcal{ASS}$ does not satisfy EUF-CMHA, then we have that $\mathcal{DS}$ does not satisfy EUF-CMA because of Theorem 17. Therefore, we get the claimed result. □

5 Conclusion

In this paper, we have proposed public-key steganography with authentication. We have defined the security notion of public-key steganography with authentication, which were the steganographic security and the unforgeability. We have constructed a concrete scheme of public-key steganography with authentication via a public-key encryption scheme and a digital signature scheme. We have shown that our proposed scheme of public-key steganography with authentication is steganographically secure and unforgeable if the underlying public-key encryption scheme satisfies IND$-CCA and the underlying digital signature scheme satisfies EUF-CMA.

References

[1] Anderson, R. J., and Petitcolas, F. A. P. On the limits of steganography. *IEEE Journal on Selected Areas in Communications 16*, 4 (May 1998), 463–473.

[2] Aucsmith, D., Ed. *Information Hiding, Second International Workshop, Portland, Oregon, USA, April 14-17, 1998, Proceedings* (1998), vol. 1525 of *Lecture Notes in Computer Science*, Springer.

[3] Backes, M., and Cachin, C. Public-key steganography with active attacks. In *TCC* (2005), J. Kilian, Ed., vol. 3378 of *Lecture Notes in Computer Science*, Springer, pp. 210–226.

[4] Cachin, C. An information-theoretic model for steganography. In Aucsmith [2], pp. 306–318.

[5] Canetti, R., Krawczyk, H., and Nielsen, J. B. Relaxing chosen-ciphertext security. In *Advances in Cryptology – CRYPTO 2003* (Santa Barbara, California, USA, August 2003), D. Boneh, Ed., vol. 2729 of *Lecture Notes in Computer Science*, Springer-Verlag, pp. 565–582.

[6] Hopper, N. On steganographic chosen covertext security. In *ICALP* (2005), L. Caires, G. F. Italiano, L. Monteiro, C. Palamidessi, and M. Yung, Eds., vol. 3580 of *Lecture Notes in Computer Science*, Springer, pp. 311–323.

[7] Hopper, N. J., Langford, J., and von Ahn, L. Provably Secure Steganography. In *Advances in Cryptology – CRYPTO 2002* (Santa Barbara, California, USA, August 2002), M. Yung, Ed., vol. 2442 of *Lecture Notes in Computer Science*, Springer-Verlag, pp. 77–92.

[8] Le, T. V., and Kurosawa, K. Efficient public key steganography secure against adaptively chosen stegotext attacks. Cryptology ePrint Archive, Report 2003/244.

[9] Mittelholzer, T. An information-theoretic approach to steganography and watermarking. In *Information Hiding* (1999), A. Pfitzmann, Ed., vol. 1768 of *Lecture Notes in Computer Science*, Springer, pp. 1–16.

[10] Shoup, V. *A computational introduction to number theory and algebra*. Cambridge University Press, pub-CAMBRIDGE:adr, 2005.

[11] von Ahn, L., and Hopper, N. J. Public-Key Steganography. In *Advances in Cryptology – EUROCRYPT 2004* (Interlaken, Switzerland, May 2004), C. Cachin and J. Camenisch, Eds., vol. 3027 of *Lecture Notes in Computer Science*, Springer-Verlag, pp. 323–341.

[12] Wegman, M. N., and Carter, L. New hash functions and their use in authentication and set equality. *J. Comput. Syst. Sci. 22*, 3 (1981), 265–279.

[13] Zöllner, J., Federrath, H., Klimant, H., Pfitzmann, A., Piotraschke, R., Westfeld, A., Wicke, G., and Wolf, G. Modeling the security of steganographic systems. In Aucsmith [2], pp. 344–354.