On Sufficient Randomness for Secure Public-Key Cryptosystems

Takeshi Koshiba

Secure Computing Lab., Fujitsu Laboratories Ltd.,
4-1-1 Kamikodanaka, Nakahara-ku, Kawasaki 211-8588, Japan
koshiba@acm.org

Abstract. In this paper, we consider what condition is sufficient for random inputs to secure probabilistic public-key encryption schemes. Although a framework given in [16] enables us to discuss uniformly and comprehensively security notions of public-key encryption schemes even for the case where cryptographically weak pseudorandom generator is used as random nonce generator to encrypt single plaintext messages, the results are rather theoretical. Here we naturally generalize the framework in order to handle security for the situation where we want to encrypt many messages with the same key. We extend some results w.r.t. single message security in [16] – separation results between security notions and a non-trivial sufficient condition for the equivalence between security notions – to multiple messages security. Besides the generalization, we show another separation between security notions for k-tuple messages and for $(k+1)$-tuple messages. The natural generalization, obtained here, rather improves to understand the security of public-key encryption schemes and eases the discussion of the security of practical public-key encryption schemes. In other words, the framework contributes to elucidating the role of randomness in public-key encryption scheme. As application of results in the generalized framework, we consider compatibility between the ElGamal encryption scheme and some sequence generators. Especially, we consider the applicability of the linear congruential generator (LCG) to the ElGamal encryption scheme.

1 Introduction

One of the important goals in computational cryptography is to provide a public-key encryption scheme that achieves a security level as strong as possible under various circumstances. For this purpose, several security notions have been introduced. In particular, we will discuss in this paper the notions of "semantic security" and "ciphertext indistinguishability" introduced in [14], which have been shown to be equivalent [14,20]. For another major security notion, we have "non-malleability" introduced in [8]. These notions are basically defined in terms of an adversary who is given only a challenge ciphertext. This attack model is called *ciphertext only attack* (abbreviated COA). Besides COA, three major attack models have been studied in the literature. One is called *chosen plaintext*

D. Naccache and P. Paillier (Eds.): PKC 2002, LNCS 2274, pp. 34–47, 2002.

attack (abbreviated CPA) model, in which the adversary can encrypt any plaintext messages of his choice. For more stronger attack models, chosen ciphertext attack and adaptive chosen ciphertext attack have been also considered in the literature [21,22].

Although these security notions have been studied quite well (see, e.g., [1,3]), we think that there are still some important issues that have not been addressed in the previous research. Security when used with a "pseudorandom" resource is one of such issues. Usually, security notions are defined assuming that ideal (i.e., true) random resource is available. Furthermore, it has been shown that one can safely use any "cryptographically strong polynomial-time pseudorandom" generator (see, e.g., [4,26]) for the substitute of the true random resource; that is, most security notions do not change by using the polynomial-time pseudorandomness for the true randomness. Although we have several "cryptographically strong" polynomial-time pseudorandom generators, they are unfortunately not fast enough for practical use, and much faster but less reliable pseudorandom generators have been used in many practical situations. Then the above security notions (and their relations) may be no longer valid with such weak pseudorandomness. In fact, it has been shown [2] that if DSS is used with the linear congruential generator, then its secret key can be easily detected after seeing a few signatures. Though this result indicates that the linear congruential generator is unsuitable for cryptographic purposes, it does not mean that the linear congruential generator is useless at all for *all* cryptographic systems. It is certainly important to study more carefully which aspect of the randomness is indeed important for discussing several security levels.

A framework introduced in [16] enables to discuss uniformly and comprehensively "semantic security" and "ciphertext indistinguishability" notions even for the case where some cryptographically weak pseudorandom generator is used as random nonce generator to encrypt plaintext messages. It has been shown that semantic security and ciphertext indistinguishability in the framework are not equivalent and a non-trivial sufficient condition for the equivalence has been given. Unfortunately, security notions only for the situation where we encrypt a single message per key generated can be handled in the framework. Clearly, in reality, we want to encrypt many messages with the same key. Nevertheless, security for multiple messages has not been intensively studied except results in [11]. In [11], security notions for multiple messages have been shown to coincide with their respective security notions for single messages. We note that such coincidence is proved only when cryptographically strong pseudorandom generators are used.

Here we naturally generalize the framework, proposed in [16], in order to handle security for multiple messages. We extend results w.r.t. single message security in [16] to multiple messages security. That is, we show that semantic security for k-tuple messages and ciphertext indistinguishability for k-tuple messages are not equivalent for any $k \geq 1$ and give a sufficient condition for the equivalence. Since these generalized results are easily derived from the original results, we stress that the generalization improves to understand the security of

public-key encryption schemes and eases the discussion of the security of practical public-key encryption schemes. Besides the generalization, we show another separation between security notions for k-tuple messages and for $(k+1)$-tuple messages. Moreover, the generalized framework enables us to discuss compatibility between public-key encryption schemes and practical pseudorandom generators. We stress that though the generalization is a natural extension, it may have an impact upon designing pseudorandom generators within practical public-key cryptosystems. In the single message security setting, it is hard to grasp the practical meaning of the results. On the other hand, generalized results with respect to the multiple message security help us to figure out involvement with practical systems. As application of results in the generalized framework, we consider compatibility between the ElGamal encryption scheme and some sequence generators. Especially, we show that linear congruential generator (LCG) is applicable to the ElGamal encryption scheme without losing security on some new and acceptable assumption.

The main contribution of this paper is rather providing a framework in which we can easily discuss security notions of practical public-key encryption schemes under more various circumstances than theoretical results. In addition, we stress that the framework elucidates the role of randomness in public-key encryption scheme.

Notations and Conventions

We introduce some useful notations and conventions for discussing probabilistic algorithms. If A is a probabilistic algorithm, then for any input x, the notation $A(x)$ refers to the probability space which assigns to the string y the probability that A, on input x, outputs y. If S is a probability space, denote by $\Pr_{e \leftarrow S}[e]$ (or $\Pr_S[e]$) the probability that S associates with element e. When we consider finite sample sets, it is convenient to consider separately a sample set and probability distribution on the set. If S is a finite set and D is a probability distribution on S, denote by $\Pr_{e \in_D S}[e]$ the probability that element $e \in S$ is chosen according to D. If S is a finite set, denote by $\Pr_{e \in_U S}[e]$ the probability that element $e \in S$ is chosen uniformly.

By 1^n we denote the unary representation of the integer n. A function $f : \{0,1\}^* \to \{0,1\}^*$ is *polynomially-bounded* if there exists a polynomial $p(\cdot)$ such that $|f(x)| \leq p(|x|)$ for all $x \in \{0,1\}^*$.

2 New Framework

In this section, we prepare a framework in which we can uniformly and comprehensively discuss "semantic security" and "ciphertext indistinguishability" notions for multiple messages even for the case where some cryptographically weak pseudorandom generator is used as random nonce generator to encrypt messages. This framework is a slightly generalized version of the framework proposed in [16]. We stress that the generalization improves to understand the

security of public-key encryption schemes and eases the discussion of the security of practical public-key encryption schemes thought the generalization itself is slight.

2.1 *R*-sequence for Random Inputs to Encryption Algorithms

We begin with introducing the notion of "R-sequence" and some notations. An R-sequence is just a sequence of strings (of certain length ℓ) randomly and uniformly chosen from some (finite) subset of initial segments of sequences of strings (of length ℓ). More specifically, we consider the following set family of string sequences.

Definition 1. Let $q(\cdot)$ be a polynomial. A $q(n)$-*R-sequence set family* (abbreviated RSSF) $\{R_n\}_{n\in\mathbb{N}}$ is a set family of sequences of strings of length $q(n)$.

Below we usually use $\{R_n\}$ to denote some RSSF. On the other hand, we consider a special $q(n)$-RSSF, where $q(n)$-RSSF $\{T_n\}$ is just a collection of sets of all infinite sequences of strings of length $q(n)$, and denote the special RSSF by TSSF. We sometimes use TSSF instead of true randomness in the sequel. Note that, in order to regard sequences in R_n as infinite ones, we sometimes consider the concatenation of the finite sequence in R_n and some infinite sequences of constant dummy strings. Although each element in R_n is possibly infinite, we use its finite initial segments only. So, we prepare some operation $\mathit{Pref}(\cdot,\cdot)$ on R_n; $\mathit{Pref}(R_n, i)$ denotes a set $\{(r_1,\ldots,r_i) : (r_1,\ldots,r_i)$ is the initial segment of a sequence in $R_n\}$. This is because we avoid a tedious discussion of random variables of infinite domain.

Our ultimate purpose is to give a taxonomy of RSSF from a viewpoint of the security of public-key encryption schemes. We will enumerate some conditions over RSSF to begin with.

While the well-known fact can be restated in our framework as the polynomial-time pseudorandomness (see, e.g., [4,26]) is *sufficient* to have the equivalence between semantic security and ciphertext indistinguishability, we show that the polynomial-time pseudorandomness is not necessary to have the equivalence. This implies that there may be more usable sufficient conditions for the equivalence. It is easy to consider separately "efficient samplability" and "indistinguishability from true randomness" as some properties on R-sequences. We call the former property *samplability* simply and the latter *semi-randomness* to distinguish from pseudorandomness. "Samplability" is quite a natural property because generators without samplability is, in general, difficult to use algorithmically. Especially in Monte-Carlo simulation, far efficient samplability is required much. On the other hand, "semi-randomness" is also one of important properties. Semi-random sequences pass many feasible statistical tests. Some sequences that are obtained from physical sources such as electronic noise or the quantum effects in a semiconductor. When the sequences pass all known feasible statistical tests, it is often that such sequences may have the semi-randomness property. So, in this paper, we study these two properties on RSSF.

We begin with definition of "semi-randomness." Semi-random sequences are ones which are not distinguished from the true randomness by any polynomial-size circuit. More specifically, we consider the following definition.

Definition 2. A $q(n)$-RSSF $\{R_n\}$ is said to be $t(n)$-*semi-random* if for any polynomial-size circuit family $\{C_n\}_{n\in\mathbb{N}}$, any polynomial $p(\cdot)$, all sufficiently large n,

$$\left| \Pr_{\substack{(r_1,\ldots,r_{t(n)})\in U \\ Pref(R_n,t(n))}} \big[C_n(r_1,\ldots,r_{t(n)})=1\big] - \Pr_{\substack{(r'_1,\ldots,r'_{t(n)})\in U \\ Pref(T_n,t(n))}} \big[C_n(r'_1,\ldots,r'_{t(n)})=1\big] \right| < \frac{1}{p(n)},$$

where $\{T_n\}$ is TSSF.

We note that semi-random sequences are different from output sequences by polynomial-time pseudorandom generators. Semi-random sequences need not to be recursive nor generated efficiently.

Next, we give a definition of "samplability." For any samplable sequence, there exists a (polynomial-size) generator $\{S_n\}_{n\in\mathbb{N}}$ whose output is statistically close to the samplable sequence. More specifically, we consider the following definition.

Definition 3. A $q(n)$-RSSF $\{R_n\}$ is said to be $t(n)$-*samplable* if there exists a polynomial-size circuit family $\{S_n\}_{n\in\mathbb{N}}$ so that for every polynomial $p(\cdot)$ and all sufficiently large n,

$$\max_A \left\{ \left| \Pr_{r\in_U\{0,1\}^{q(n)}} \big[S_n(r)\in A\big] - \Pr_{\substack{(r_1,\ldots,r_{t(n)})\in U \\ Pref(R_n,t(n))}} \big[(r_1,\ldots,r_{t(n)})\in A\big] \right| \right\} < \frac{1}{p(n)},$$

where the maximum is taken all over the subsets A of $Pref(T_n,t(n))$.

We note that the maximum value in the above definition is so called "statistical difference" between two probability distributions: $\{S_n(r)\}_{r\in_U\{0,1\}^{q(n)}}$ and the uniform distribution on $Pref(R_n,t(n))$.

We extend the notion of public-key encryption scheme in order to cope with RSSF instead of true randomness. The following is our treatment for public-key encryption schemes in the new framework. The following definition seems to be cumbersome. Since nonces in encryption are not necessarily independent of each other, the definition below seems to be more complex than the original (simplified) definition.

Definition 4 (public-key encryption scheme, revisited). A *public-key encryption scheme* is a quadruple (G,M,E,D), where the following conditions hold.

1. G, called the *key generator*, is a probabilistic polynomial-time algorithm which, on input 1^n, outputs a pair of binary strings. (Although the key generator also uses randomness, we disregard it here in order to cast light on roles of randomness in encrypting. So, we assume that randomness in key generator is always ideal.)

2. $M = \{M_n\}_{n\in\mathbb{N}}$ is a family of message spaces from which all plaintext messages will be drawn. In order to make our notation simpler (but without loss of generality), we will assume that $M_n = \{0,1\}^n$.
3. For every polynomial $q(\cdot)$, every $q(n)$-RSSF $\{R_n\}$, every n, every pair (e,d) in the support of $G(1^n)$, for any integer $k \geq 1$ and for any $\alpha_1, \ldots, \alpha_k \in M_n$, (*encryption*) "deterministic" polynomial-time algorithm E and (*decryption*) deterministic polynomial-time algorithm D satisfy

$$\Pr_{\substack{(r_1,\ldots,r_k)\in_U \\ Pref(R_n,k)}}\left[\bigwedge_{i=1}^{k} D(d, E(e,\alpha_i;r_i)) = \alpha_i\right] = 1,$$

where the probability is over the uniform distribution on $Pref(R_n, k)$.

Hereafter, we write $E_e(\alpha; r)$ instead of $E(e, \alpha; r)$ and $D_d(\beta)$ instead of $D(d, \beta)$. We note that the argument r in the term $E_e(\alpha; r)$ denotes the random input to the encryption algorithm E. Also, we let $G_1(1^n)$ denote the first element (i.e., encryption key) in the pair $G(1^n)$. Without loss of generality, we treat the encryption algorithm as deterministic one fed with a plaintext message and a (random) supplementary input of length $q(n)$.

2.2 Security Notions in the New Framework

In this subsection, we reformulate the notions of semantic security and indistinguishability to suit the new framework.

Since Goldwasser and Micali defined semantic security and ciphertext indistinguishability (a.k.a. polynomial security), several ways to define such notions are shown. In this paper, we adopt a non-uniform formulation as in [11] in order to simplify the exposition. We note that employing such a non-uniform formulation (rather than a uniform one) may strengthen the definitions; yet, it does weaken the implications proven between the definitions, since proofs make free usage of non-uniformity.

A transformation is a uniform algorithm which, on inputs $\overline{C_n}$, outputs $\overline{C'_n}$, where $\overline{C_n}$ (resp., $\overline{C'_n}$) is the representation of a circuit C_n (resp., C'_n) in some standard encoding. Without loss of generality, we identify a circuit with its representation (in the standard encoding).

Definition 5. A public-key encryption scheme (G, M, E, D) is *semantically secure for $t(n)$-tuple messages w.r.t. $q(n)$-RSSF $\{R_n\}$* if there exists a probabilistic polynomial-time transformation T so that every polynomial-size circuit family $\{C_n\}_{n\in\mathbb{N}}$, for every probability ensemble $\{\bar{X}_n\}_{n\in\mathbb{N}}$ satisfying that $\bar{X}_n$ is a probability distribution on $M_n^{t(n)}$, every pair of polynomially-bounded functions $f, h : \{0,1\}^* \to \{0,1\}^*$, every polynomial $p(\cdot)$ and all sufficiently large n,

$$\Pr_{\substack{G,\bar{X};(r_1,\ldots,r_{t(n)})\in_U \\ Pref(R_n,t(n))}}\left[C_n(G_1(1^n), \bar{E}_{G_1(1^n)}(\bar{X}_n;\bar{r}), 1^n, h(\bar{X}_n)) = f(\bar{X}_n)\right]$$
$$< \Pr_{T,G,\bar{X}_n}\left[C'_n(G_1(1^n), 1^n, h(\bar{X}_n)) = f(\bar{X}_n)\right] + \frac{1}{p(n)}$$

where $C'_n = T(C_n)$, $\bar{X}_n = (X_n^{(1)}, \ldots, X_n^{(t(n))})$, and $\bar{E}_e(\bar{X}_n; \bar{r}) = E_e(X_n^{(1)}; r_1), \ldots, E_e(X_n^{(t(n))}; r_{t(n)})$.

Some explanation on the attack model is needed here. In the above definition, an adversary C_n is given only an encryption key $G_1(1^n)$ and a ciphertext message $E_{G_1(1^n)}(X_n; r)$ (and some supplementary information $h(X_n)$). Thus, it is considered as ciphertext only attack (COA) model. But note here that we may consider any polynomial-size circuit C_n for the adversary; hence, we may assume that the encryption algorithm is also included in C_n. In the situation where the true randomness is available, this immediately includes the chosen plaintext attack (CPA) model where the adversary can encrypt any plaintext messages of his choice. This is not true any more in the new framework because there is no guarantee that some (randomized) polynomial-size circuit can generate R-sequences in R_n uniformly at random. Moreover, we consider our revised COA model. For our COA model, we consider the situation where an adversary cannot directly access to R-sequence generators. The situation means that those who use public-key encryption scheme have their own *private* R-sequence generators. In general, they do not have to publicize their R-sequence generators which are used in public-key encryption scheme. In addition, the case where R-sequence generators are *privately* used is more secure than the case where R-sequence generators are *publicly* used. Thus, we can say that our COA model makes sense.

Definition 6. A public-key encryption scheme (G, M, E, D) is *ciphertext indistinguishable for $t(n)$-tuple messages w.r.t. $q(n)$-RSSF $\{R_n\}$* if for every polynomial-size circuit family $\{C_n\}_{n\in\mathbb{N}}$, every polynomial $p(\cdot)$, all sufficiently large n and every $x_1, \ldots, x_{t(n)}, y_1, \ldots, y_{t(n)} \in M_n$,

$$\left| \Pr_{\substack{G;(r_1,\ldots,r_{t(n)})\in_U \\ Pref(R_n,t(n))}} \left[C_n(G_1(1^n), \bar{E}_{G_1(1^n)}(\bar{x}; \bar{r})) = 1\right] - \Pr_{\substack{G;(r'_1,\ldots,r'_{t(n)})\in_U \\ Pref(R_n,t(n))}} \left[C_n(G_1(1^n), \bar{E}_{G_1(1^n)}(\bar{y}; \bar{r}')) = 1\right] \right| < \frac{1}{p(n)}$$

where $\bar{x} = (x_1, \ldots, x_{t(n)})$, $\bar{y} = (y_1, \ldots, y_{t(n)})$, $\bar{E}_e(\bar{x}; \bar{r}) = E_e(x_1; r_1), \ldots, E_e(x_{t(n)}; r_{t(n)})$, and $\bar{E}_e(\bar{y}; \bar{r}') = E_e(y_1; r'_1), \ldots, E_e(y_{t(n)}; r'_{t(n)})$.

The following notion is somewhat artificial. However, it is useful to characterize the notions of semantic security and ciphertext indistinguishability.

Definition 7. A public-key encryption scheme (G, M, E, D) is *ciphertext skew-indistinguishable for $t(n)$-tuple messages w.r.t. $q(n)$-RSSF $\{R_n\}$* if for every polynomial-size circuit family $\{C_n\}_{n\in\mathbb{N}}$, every polynomial $p(\cdot)$, all sufficiently

large n and every $x_1, \ldots, x_{t(n)}$, $y_1, \ldots, y_{t(n)} \in M_n$,

$$\left| \Pr_{\substack{G;(r_1,\ldots,r_{t(n)})\in U \\ Pref(R_n,t(n))}} \left[C_n(G_1(1^n), \bar{E}_{G_1(1^n)}(\bar{x};\bar{r})) = 1\right] - \Pr_{\substack{G;(r'_1,\ldots,r'_{t(n)})\in U \\ Pref(T_n,t(n))}} \left[C_n(G_1(1^n), \bar{E}_{G_1(1^n)}(\bar{y};\tilde{r})) = 1\right] \right| < \frac{1}{p(n)}$$

where $\bar{x} = (x_1, \ldots, x_{t(n)})$, $\bar{y} = (y_1, \ldots, y_{t(n)})$, $\bar{E}_e(\bar{x};\bar{r}) = E_e(x_1;r_1), \ldots, E_e(x_{t(n)};r_{t(n)})$, and $\bar{E}_e(\bar{y};\tilde{r}) = E_e(y_1;r'_1), \ldots, E_e(y_{t(n)};r'_{t(n)})$.

We note that, in three definitions above, any adversary does not directly access to RSSF but gets ciphertext messages encrypted using the RSSF as inputs.

We have seen some security notions for public-key encryption schemes. Here we mention known results w.r.t. multiple messages security by our terminology.

Theorem 1 ([11,14,20]). *Let (G, M, E, D) be a public-key encryption scheme. The following statements are equivalent.*

1. *(G,M,E,D) is semantically secure for single message w.r.t. TSSF.*
2. *(G,M,E,D) is ciphertext indistinguishable for single message w.r.t. TSSF.*
3. *(G,M,E,D) is semantically secure for polynomial-tuple messages w.r.t. TSSF.*
4. *(G,M,E,D) is ciphertext indistinguishable for polynomial-tuple messages w.r.t. TSSF.*

Recall that TSSF is a special case of RSSF. So, the equivalence is satisfied if the true randomness (say, TSSF) is used as random inputs. In what follows, we discuss general cases.

3 Results

3.1 Separation Results

In this subsection, we consider classes of pairs of RSSF and public-key encryptions scheme w.r.t. the RSSF. We especially show that semantic security and ciphertext indistinguishability for multiple messages are separable from each other.

We denote by $\mathcal{SS}_r^{t(n)}$ the class of pairs of encryption scheme (G, M, E, D) and RSSF $\{R_n\}$ satisfying that (G, M, E, D) w.r.t. $\{R_n\}$ is semantically secure for $t(n)$-tuple messages. We also denote $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{SS}_r^{t(n)}$ if an encryption scheme (G, M, E, D) which is semantically secure for $t(n)$-tuple messages w.r.t. a RSSF $\{R_n\}$. We denote by $\mathcal{INC}_{rr}^{t(n)}$ the class of pairs of encryption schemes (G, M, E, D) and RSSF $\{R_n\}$ satisfying that (G, M, E, D) w.r.t. RSSF $\{R_n\}$ is ciphertext indistinguishable for $t(n)$-tuple messages. We denote by $\mathcal{INC}_{rt}^{t(n)}$ the class of pairs of encryption scheme (G, M, E, D) and

RSSF $\{R_n\}$ satisfying that (G, M, E, D) w.r.t. RSSF $\{R_n\}$ is ciphertext skew-indistinguishable for $t(n)$-tuple messages.

In [16], some relations among security notions for single message have been already shown. In case of multiple messages, we can obtain similar results to the case of single message.

Theorem 2. *Suppose that there exists a public-key encryption scheme w.r.t. TSSF. Then, for any polynomial $t(n)$, $\mathcal{INC}\mathcal{D}_{rt}^{t(n)} \subsetneq \mathcal{SS}_{r}^{t(n)} \subsetneq \mathcal{INC}\mathcal{D}_{rr}^{t(n)}$.*

We omit the proof on account of space constraints. We note that the theorem can be similarly shown as a proof in [16].

3.2 Sufficient Condition for the Equivalence

In this subsection, we consider how properties of RSSF affect on the security of encryption schemes. We especially give a sufficient condition that semantic security and ciphertext indistinguishability for multiple messages become equivalent.

Theorem 3. *Let $t(n)$ be a polynomial. Suppose that $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{INC}\mathcal{D}_{rr}^{t(n)}$. If $\{R_n\}$ is $t(n)$-semi-random, then $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{INC}\mathcal{D}_{rt}^{t(n)}$.*

Theorem 4. *Let $t(n)$ be a polynomial. Suppose that $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{INC}\mathcal{D}_{rr}^{t(n)}$. If $\{R_n\}$ is $t(n)$-samplable, then $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{SS}_{r}^{t(n)}$.*

We omit the proofs for the above two theorems on account of space constraints. We note that the theorem can be similarly shown as a proof in [16].

Corollary 1. *Let $t(n)$ be a polynomial. Suppose that $\{R_n\}$ is $t(n)$-semi-random or $t(n)$-samplable. Then $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{INC}\mathcal{D}_{rr}^{t(n)}$ if and only if $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{SS}_{r}^{t(n)}$.*

Although we have a better sufficient condition for the equivalence between semantic security and ciphertext indistinguishability, the condition is not necessary for the equivalence.

Theorem 5. *Let $t(n)$ be a polynomial. Suppose that there exists a public-key encryption scheme that is semantically secure w.r.t. TSSF. There exists an encryption scheme (G, M, E, D) such that $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{INC}\mathcal{D}_{rt}^{t(n)}$ and $\{R_n\}$ is not $t(n)$-semi-random.*

Proof. Suppose that $\langle (G, M, E, D), \{T_n\}\rangle \in \mathcal{INC}\mathcal{D}_{rt}^{t(n)}$, where $\{T_n\}$ is $q(n)$-TSSF. Then there exists an encryption scheme (G, M, E', D') such that $\langle (G, M, E', D'), \{T'_n\}\rangle \in \mathcal{INC}\mathcal{D}_{rt}^{t(n)}$, where $\{T'_n\}$ is $(q(n)+1)$-TSSF, $E'_e(\alpha; r) = E_e(\alpha; r_1)$, $D'_d(\beta) = D_d(\beta)$, $r = r_1 r_2$ and $|r_2| = 1$. We consider a RSSF $\{R_n\} = (\{0,1\}^{q(n)}1)^{t(n)}$. It is easy to see that $\langle (G, M, E', D'), \{R_n\}\rangle \in \mathcal{INC}\mathcal{D}_{rt}^{t(n)}$, because the last bit of random supplementary bit is not used in encrypting.

On the other hand, it is easy to see that $\{R_n\}$ and $\{T'_n\}$ are distinguishable. In other words, $\{R_n\}$ is not $t(n)$-semi-random. □

3.3 Multiplicity

It is easy to see that the parameter of RSSF is available as a measure of compatibility between RSSFs and encryption schemes. The following theorem says that, for any pair of RSSF and encryption scheme, there may exist limitation on the numbers of messages which are encrypted with the same key and without losing security.

Theorem 6. *Let $t(\cdot)$ and $t'(\cdot)$ be polynomials. Suppose that there exists a public-key encryption scheme that is semantically secure w.r.t. TSSF. If $t(n) < t'(n)$ then $\mathcal{IND}_{rr}^{t'(n)} \subsetneq \mathcal{IND}_{rr}^{t(n)}$, $\mathcal{IND}_{rt}^{t'(n)} \subsetneq \mathcal{IND}_{rt}^{t(n)}$, and $\mathcal{SS}_{r}^{t'(n)} \subsetneq \mathcal{SS}_{r}^{t(n)}$.*

Proof. Let g be a pseudorandom generator which, given a seed s of length $q(n)$, outputs a string of length $t(n)q(n)$. We consider two RSSFs $R_n = \{g(s) : s \in \{0,1\}^{q(n)}\}$ and $R'_n = \{(g(s), g_l(s)) : s \in \{0,1\}^{q(n)}\}$, where $g_l(s)$ denotes the suffix of $g(s)$ of length $q(n)$. Let $\{T_n\}$ be $q(n)$-TSSF. Suppose that $\langle (G, M, E, D), \{T_n\}\rangle \in \mathcal{IND}_{rt}^{t(n)}$. It is easy to see that $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{IND}_{rt}^{t(n)}$, which implies that $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{SS}_{r}^{t(n)}$ and $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{IND}_{rr}^{t(n)}$. It is also easy to see that $\langle (G, M, E, D), \{R'_n\}\rangle \notin \mathcal{IND}_{rt}^{t(n)+1}$ and $\langle (G,M,E,D), \{R'_n\}\rangle \notin \mathcal{IND}_{rr}^{t(n)+1}$, because encryptions of $x_1, \ldots, x_{t(n)}, x_{t(n)}$ and $y_1, \ldots, y_{t(n)}, y_{t(n)+1}$ are distinguishable. We consider a function f such that $f(x_1, \ldots, x_{t(n)}, x_{t(n)+1})=1$ if and only if $x_{t(n)} = x_{t(n)+1}$. Then $\langle (G, M, E, D), \{R'_n\}\rangle \notin \mathcal{SS}_{r}^{t(n)+1}$.

It is easy to see that $\mathcal{IND}_{rr}^{t'(n)} \subseteq \mathcal{IND}_{rr}^{t(n)}$, $\mathcal{IND}_{rt}^{t'(n)} \subseteq \mathcal{IND}_{rt}^{t(n)}$, and $\mathcal{SS}_{r}^{t'(n)} \subseteq \mathcal{SS}_{r}^{t(n)}$. This completes the proof. □

4 Application

In [25], it is shown that the ElGamal encryption scheme [10] is semantically secure on condition that the decision Diffie-Hellman (DDH) problem is intractable. Let $\mathcal{G}$ be a group of some odd prime order q. Roughly speaking, the DDH problem is one to distinguish the uniform distribution on $\{(g, g^a, g^b, g^{ab}) : g \in \mathcal{G}, a, b \in \mathbb{Z}_q\} \subset \mathcal{G}^4$ from the uniform distribution on $\mathcal{G}^4$ (see, e.g., [5,7,19]). In this section, we consider the compatibility between the ElGamal encryption scheme and linear congruential sequences.

Let us give a simple description of the ElGamal encryption scheme $\mathbf{EG} = (G_{eg}, M_{eg}, E_{eg}, D_{eg})$. Key generation algorithm G_{eg} chooses an n-bit prime number p such that $p = 2q + 1$ and q is a prime number. Let $\mathcal{G}_p$ be the unique non-trivial subgroup of $\mathbb{Z}_p^*$. G_{eg} also chooses uniformly and randomly a generator $g \in \mathcal{G}_p$ and $x \in \mathbb{Z}_q$. G_{eg} finally outputs $((p, g, g^x), x)$. Message space M_{eg} is set to be $\mathcal{G}_p$. (Although, in the definition of encryption scheme, message space depends only on the security parameter, we use prime-dependent message space without loss of generality.) Encryption algorithm E_{eg}, given an encryption key (p, g, y), a message m and a random input r, outputs (g^r, my^r). We note that group operation is carried using the value p. Decryption algorithm D_{eg}, given a decryption key x and ciphertext (c_1, c_2), outputs $c_2/(c_1)^x$.

Let us consider the prime-indexed RSSF $\{R_p\}$ which corresponds to linear congruential sequences, where $R_p = \{(r, f_p(r), \ldots, f_p^{t(n)-1}(r)) : r \in \mathbb{Z}_q\}$ and f_p is a function of the form $f_p(r) = ar + b \bmod q$.

Now we are ready to consider the security of the ElGamal encryption scheme w.r.t. linear congruential sequence for random inputs. First, we restate some trivial statements using our terminology.

Proposition 1. *Suppose that the DDH problem is intractable. If the parameter a for linear congruential sequence is public, then $\langle \mathbf{EG}, \{R_p\}\rangle \in \mathcal{SS}_r^1$ and $\langle \mathbf{EG}, \{R_p\}\rangle \notin \mathcal{SS}_r^k$ for any $k \geq 2$.*

The above proposition seems to say that the linear congruential sequence is useless at all for the ElGamal encryption scheme. However, we do not have to publicize the parameter of the linear congruential sequence.

Proposition 2. *Suppose that the DDH problem is intractable. If the parameter of the linear congruential sequence is not public but randomly and uniformly distributed, then $\langle \mathbf{EG}, \{R_p\}\rangle \in \mathcal{SS}_r^2$.*

We do not know whether or not $\langle \mathbf{EG}, \{R_p\}\rangle \in \mathcal{SS}_r^3$ on the same assumption. So, we consider a bit stronger assumption. Let $\mathcal{L}_k = \{(gh, g^a h, \ldots, g^{a^k} h) : g, h \in \mathcal{G}_p, a \in \mathbb{Z}_q\} \subset (\mathcal{G}_p)^{k+1}$. We call the problem to distinguish the uniform distribution on $\mathcal{L}_k$ from the uniform distribution on $(\mathcal{G}_p)^{k+1}$ *decision k-skew-power series* (k-DSPS) problem. If $h = 1$ then the k-DSPS problem is reducible to the DDH problem. It seems that the k-DSPS problem is somewhat artificial. However, it is just a subproblem of a natural problem. We note that $\mathcal{G}_p$ is a commutative ring w.r.t. two operators $\oplus_g$ and $\otimes_g$, where $g^a \oplus_g g^b = g^{a+b}$ and $g^a \otimes_g g^b = g^{ab}$. The DDH problem is considered as the equivalence problem between $\alpha \otimes_g \beta$ and γ, where $\alpha, \beta, \gamma \in \mathcal{G}_p$. Similarly, the (computational) Diffie-Hellman problem is considered as the evaluating problem for $\alpha \otimes_g \beta$, where $\alpha, \beta \in \mathcal{G}_p$. Naturally, we can define *expression* on $\mathcal{G}_p$ using the additive operator $\oplus_g$ and the multiplicative operator $\otimes_g$. So, the equivalence problem for two expressions on $\mathcal{G}_p$ is more general than the DDH problem. It is easy to see that the k-DSPS problem is also a subproblem of the equivalence problem for two expressions on $\mathcal{G}_p$. We note that if both of two expressions on $\mathcal{G}_p$ do not include any multiplicative operator, the subproblem is easily solved.

Theorem 7. *Suppose that the k-DSPS problem is intractable, where k is a constant. If the parameter of the linear congruential sequence is not public but randomly and uniformly distributed, then $\langle \mathbf{EG}, \{R_p\}\rangle \in \mathcal{SS}_r^{k+1}$.*

Proof. The ciphertext skew-indistinguishability w.r.t. LCG follows directly from the assumption. Using Theorem 2, we get the assertion. □

We consider a bit stronger assumption that $z(n)$-DSPS problem is intractable for any polynomial $z(\cdot)$ and name it *DSPS assumption*. Then we get the following.

Corollary 2. *Under the DSPS assumption,* $\langle \mathbf{EG}, \{R_p\}\rangle \in \mathcal{SS}_r^{v(n)}$ *for any polynomial* $v(\cdot)$.

We note that in the case of the DSS in [2] the secret key can be detected by solving some simultaneous linear equations. However, in the case of the ElGamal encryption scheme w.r.t. LCG, such equations do not appear in ciphertext. Thus, the techniques in [2] do not seem to be applicable to the case of the ElGamal encryption scheme w.r.t. LCG.

5 Concluding Remarks

We have extended the framework proposed in [16] where we can uniformly and comprehensively discuss security notions of public-key encryption schemes even for the case where some cryptographically weak pseudorandom generator is used as random nonce generator to encrypt plaintext messages. We have also shown some separation results between security notions for multiple messages and given a sufficient condition for the equivalence between the security notions. Obtained results give us a clear sight for designing sequence generators for random inputs to public-key encryption schemes. We have shown that the LCG is available to random inputs to the ElGamal encryption schemes on some similar assumption with the DDH assumption, although the LCG itself is cryptographically weak [6,17,24]. However, reliability of the assumption may be controversial, thought it is weaker than a natural assumption where the equivalence problem for two expressions on $\mathcal{G}_p$ is intractable.

References

1. M. Bellare, A. Desai, D. Pointcheval, and P. Rogaway. Relations among notions of security for public-key encryption schemes. In H. Krawczyk, editor, *Advances in Cryptology — CRYPTO'98*, volume 1462 of *Lecture Notes in Computer Science*, pages 26–45. Springer-Verlag, 1998.
2. M. Bellare, S. Goldwasser, and D. Micciancio. Pseudo-random number generation within cryptographic algorithms: The DSS case. In B. S. Kaliski Jr., editor, *Advances in Cryptology — CRYPTO'97*, volume 1294 of *Lecture Notes in Computer Science*, pages 277–291. Springer-Verlag, 1997.
3. M. Bellare and A. Sahai. Non-malleable encryption: Equivalence between two notions, and an indistinguishability-based characterization. In M. Wiener, editor, *Advances in Cryptology — CRYPTO'99*, volume 1666 of *Lecture Notes in Computer Science*, pages 519–536. Springer-Verlag, 1999.
4. M. Blum and S. Micali. How to generate cryptographically strong sequences of pseudo-random bits. *SIAM Journal on Computing*, 13(4):850–864, 1984.
5. D. Boneh. The decision Diffie-Hellman problem. In J. P. Buhler, editor, *Proceedings of the 3rd International Symposium on Algorithmic Number Theory (ANTS-3)*, volume 1423 of *Lecture Notes in Computer Science*, pages 48–63. Springer-Verlag, 1998.
6. J. Boyar. Inferring sequences produced by pseudo-random number generators. *Journal of the Association for Computing Machinery*, 36(1):129–141, 1989.

7. W. Diffie and M. E. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, IT-22(6):644–654, 1976.
8. D. Dolev, C. Dwork, and M. Naor. Non-malleable cryptography. In *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing*, pages 542–552. ACM Press, 1991.
9. D. Dolev, C. Dwork, and M. Naor. Non-malleable cryptography. *SIAM Journal on Computing*, 30(2):391–437, 2000.
10. T. ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Transactions on Information Theory*, IT-31(4):469–472, 1985.
11. O. Goldreich. *Foundation of Cryptography (Fragment of a Book – Version 2.03)*, 1998.
12. O. Goldreich. *Modern Cryptography, Probabilistic Proofs and Pseudorandomness*, volume 17 of *Algorithms and Combinatorics*. Springer-Verlag, 1999.
13. O. Goldreich. *Foundations of Cryptography: Basic Tools*. Cambridge University Press, 2001.
14. S. Goldwasser and S. Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28(2):270–299, 1984.
15. T. Koshiba. A theory of randomness for public key cryptosystems: The ElGamal cryptosystem case. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E83-A(4):614–619, 2000.
16. T. Koshiba. A new aspect for security notions: Secure randomness in public-key encryption schemes. In K. Kim, editor, *Proceeding of the 4th International Workshop on Practice and Theory in Public Key Cryptography (PKC2001)*, volume 1992 of *Lecture Notes in Computer Science*, pages 87–103. Springer-Verlag, 2001.
17. H. Krawczyk. How to predict congruential generators. *Journal of Algorithms*, 13(4):527–545, 1992.
18. M. Luby. *Pseudorandomness and Cryptographic Applications*. Princeton Univ. Press, 1996.
19. U. M. Maurer and S. Wolf. Diffie-Hellman protocol. *Designs, Codes and Cryptography*, 19(2-3):147–171, 2000.
20. S. Micali, C. Rackoff, and B. Sloan. The notion of security for probabilistic cryptosystems. *SIAM Journal on Computing*, 17(2):412–426, 1988.
21. M. Naor and M. Yung. Public-key cryptosystems provably secure against chosen ciphertext attacks. In *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing*, pages 427–437. ACM Press, 1990.
22. C. Rackoff and D. R. Simon. Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack. In J. Feigenbaum, editor, *Advances in Cryptology — CRYPTO'91*, volume 576 of *Lecture Notes in Computer Science*, pages 433–444. Springer-Verlag, 1992.
23. T. Saito, T. Koshiba, and A. Yamamura. The decision Diffie-Hellman assumption and the quadratic residuosity assumption. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E84-A(1):165–171, 2001.
24. J. Stern. Secret linear congruential generators are not cryptographically secure. In *Proceedings of the 28th Annual IEEE Symposium on Foundations of Computer Science*, pages 421–426. IEEE Computer Society Press, 1987.
25. Y. Tsiounis and M. Yung. On the security of ElGamal based encryption. In H. Imai and Y. Zheng, editors, *Proceedings of the 1st International Workshop on Practice and Theory in Public Key Cryptography (PKC'98)*, volume 1431 of *Lecture Notes in Computer Science*, pages 117–134. Springer-Verlag, 1998.

26. A. C. Yao. Theory and applications of trapdoor functions. In *Proceedings of the 23rd Annual IEEE Symposium on Foundations of Computer Science*, pages 80–91. IEEE Computer Society Press, 1982.