Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 7920495, 10 pages
http://dx.doi.org/10.1155/2016/7920495

Research Article

An Encryption Technique for Provably Secure Transmission from a High Performance Computing Entity to a Tiny One

Miodrag J. Mihaljević,[1] Aleksandar Kavčić,[2] and Kanta Matsuura[3]

[1]*Mathematical Institute, Serbian Academy of Sciences and Arts, Kneza Mihaila 36, 11000 Belgrade, Serbia*
[2]*Department of Electrical Engineering, University of Hawaii, 2540 Dole Street, Honolulu, HI 96822, USA*
[3]*Institute of Industrial Science, University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan*

Correspondence should be addressed to Miodrag J. Mihaljević; miodragm@turing.mi.sanu.ac.rs

Received 25 December 2015; Accepted 22 March 2016

Academic Editor: Veljko Milutinovic

An encryption/decryption approach is proposed dedicated to one-way communication between a transmitter which is a computationally powerful party and a receiver with limited computational capabilities. The proposed encryption technique combines traditional stream ciphering and simulation of a binary channel which degrades channel input by inserting random bits. A statistical model of the proposed encryption is analyzed from the information-theoretic point of view. In the addressed model an attacker faces the problem implied by observing the messages through a channel with random bits insertion. The paper points out a number of security related implications of the considered channel. These implications have been addressed by estimation of the mutual information between the channel input and output and estimation of the number of candidate channel inputs for a given channel output. It is shown that deliberate and secret key controlled insertion of random bits into the basic ciphertext provides security enhancement of the resulting encryption scheme.

1. Introduction

It is well recognized that communications should be secure and accordingly encrypted in order to avoid misuse of the transmitted information. Consequently, contemporary cryptographic algorithms for encryption play a very important role in data communication systems for various areas of applications. A particular challenge is related to addressing the resource constrained environments, where the requirements include lightweight algorithms and hardware designs. To select a suitable encryption algorithm for an application or an environment, the algorithmic requirements as well as the implementation constraints have to be taken into account. This is also in line with a discussion recently reported in [1].

On the other hand, in a number of scenarios the communication parties are with very different capabilities: one party could be with a tiny capability and the other with much higher ones. As an illustration, we point to a communication scenario over the Internet of Things (IoT) where a tiny machine (a tiny sensor, e.g.) should communicate with a more powerful one (sink of a sensor network or a gate, e.g.). According to the current state of the art, the following two problems appear as the still open ones: (i) developing encryption/decryption techniques which take into account asymmetric capabilities of the entities involved in encryption/decryption and (ii) enhancing cryptographic security of encryption in a lightweight and provable manner.

Consequently, in this paper we consider the problem of designing a dedicated encryption/decryption algorithm which fits into the communications scenarios which include the following: (i) a high performance computing party should deliver encrypted messagesin a one-way communication scenario to a number of parties which have tiny computational capabilities; (ii) implementation limitations at the tiny entity imply employment of a lightweight keystream generator (from certain reported lightweight stream ciphers); (iii) developed encryption scheme should have enhanced security in comparison with the one offered by the employed keystream generator.

A certain number of reported encryption approaches jointly employ elements of traditional stream ciphers and

elements of coding theory as well as features of certain communication channels (see, e.g., [2–8]), and this paper follows the same track. We consider an encryption approach which involves a communication channel with the synchronization errors which appear in the form of *inserted bits*. In this approach, the transmitting/encrypting side requires a source of random bits and capability to insert them between message bits. Under the assumption that the transmitter has a method to inform the intended receiver about the *locations* (and not necessarily the values) of the inserted random bits, the intended receiver can perform decimation (i.e., discard the inserted bits) of the obtained sequence so that it can be a subject of simple traditional decryption.

Summary of the Results. This paper focuses on the following two issues which have not been addressed in the literature: (i) developing of an encryption/decryption technique which has asymmetric implementation complexity and provides lightweight decryption and (ii) security enhancement of the involved keystream generator employing paradigm of the binary channels with random insertions. An encryption/decryption technique for data transfer between a computationally powerful party and a party with limited computational capabilities is proposed which provides a tradeoff between implementation complexities at the involved parties: the implementation overhead is reduced at the low-capability party at the expense of a higher (but still moderate) one at the party with high capabilities. In order to achieve security enhancement of the employed traditional keystream generator the proposed encryption technique at the transmitting side involves a simulator of the binary channel with synchronization errors. Security enhancement of encryption archived by the proposed scheme in comparison with the security of the employed keystream generator is based on the design paradigm and results on the mutual information between inputs and outputs of the channels with bit insertion.

Organization. The paper is organized as follows. In Section 2, we give the underlying ideas for the design and proposal of an encryption/decryption framework. In Section 3, we provide some information-theoretic results for the proposed scheme; that is, we mostly derive various mutual information rates of interest for the security evaluation. In Section 4, we provide the cryptographic security evaluation based on implications which link the information-theoretic quantities to computational complexity based ones. Accordingly, Sections 5 and 6 provide evaluation of the computational complexity security enhancement employing numerical estimation of the mutual information and enumeration of input candidates for the given output after a binary channel with insertion of random bits, respectively. (Also note that this paper is a significantly revised and expanded version of [8].)

2. A Proposal of a Dedicated Encryption Technique

This section proposes an encryption/decryption technique which provides asymmetric implementation complexity at the communicating parties and provably enhanced cryptographic security. Both asymmetric implementation complexity and enhanced security appear as a consequence of the design based on employment of a simulator for binary channels with insertion errors.

2.1. Underlying Ideas. Our main design goals/approaches could be summarized as follows:

(i) Enhance security based on information-theoretic and coding results over channels with synchronization errors.

(ii) Assuming that Party I is more powerful than Party II move the more complex operations to the side of Party I without implications on the cryptographic security.

This paper proposes a stream cipher developed based on the following two construction principles: (i) adjustment of the construction to the asymmetric capabilities of the involved parties; (ii) employment of the results regarding binary channels with insertion errors for enhancing security. The goals are that the party with more powerful resources performs more complex operations and that the entire scheme provides a highly and provably secure level of cryptographic security resulting from the employment of the insertion communications channel paradigm.

Our design is based on employment of the following building blocks:

(i) a lightweight binary keystream generator;

(ii) a block for insertion (embedding) t random bits into a given n-dimensional binary vector;

(iii) a block for decimation of a given $(n + t)$-dimensional binary vector which selects certain n-bits.

Accordingly, we assume that the employed keystream generator outputs certain pseudo-random sequences denoted as C^n and G'^n. Also, we assume that a deterministic mapping exists which maps a given G'^n into G^n. We assume that the message M^n is additively combined (i.e., encrypted) with the shared pseudo-randomness C^n to obtain X^n, that is,

$$X^n = M^n \oplus C^n, \tag{1}$$

and X^n is subject of further mapping by a simulated binary channel with random insertions where positions of random bits embedding are specified by G^n so that the channel outputs $Y^{(n)}$. The intended receiver (Bob), knowing both C^n and G^n, can easily decimate $Y^{(n)}$ to obtain X^n and further perform $M^n = X^n \oplus C^n$, to obtain the message M^n.

Since Bob can easily recover the transmitted message using a simple decimation technique, the system requires no special hardware overhead for decryption. This is especially useful if the intended receiver is a low-power device. On the transmitter's side encryption requires simulation of a binary channel with insertion errors and the transmitter needs to send $(1 - i)^{-1}$ times more symbols than it otherwise would, which means that the power consumption of

the transmitter goes up by a factor of $(1-i)^{-1}$. Hence, it may be reasonable to use this scheme when the transmitter is a high computational/power device and the receiver is a low computation/power device. In essence, a properly adjusted synchronization error scheme (an insertion scheme) seems to be well suited for a resources-asymmetric communication scenario in which a base station has ample resources while each of the numerous distributed nodes has severely constrained resources.

2.2. Framework for Encryption and Decryption. This section proposes an encryption/decryption technique for one-way communication from a transmitting party with high computational and other resources towards a receiving party with limited computational capabilities. Accordingly, the design follows the asymmetric implementation and execution constraints and the requirement regarding provable security.

As usual, it is assumed that encryption and decryption parties share a secret key and that before a transmission session, based on the common secret key and the public data, both parties (encryption and decryption ones) establish a session key to be used for the transmission session.

The encryption/decryption technique is designed employing the following components:

(a) Encryption side:

(i) a lightweight stream cipher (keystream generator);

(ii) a block which provides deterministic mapping (see Figure 1) of a given keystream segment of dimension $n+t$ into a vector with predetermined weight equal to t, that is, with a number of ones equal to t which determines positions of the embedded bits;

(iii) a simulator of a binary channel with random bits insertions controlled by keystream generator which performs mapping $\{0,1\}^n \to \{0,1\}^{n+t}$.

(b) Decryption side:

(i) a lightweight stream cipher (keystream generator);

(ii) a block for deterministic mapping of a given keystream segment into a vector with predetermined weight, that is, the number of ones, the same as that at the encryption side;

(iii) a block for decimation controlled by keystream generator which performs mapping $\{0,1\}^{n+t} \to \{0,1\}^n$.

We assume that implementation and execution complexity of a keystream controlled simulator of a binary channel with random insertions is highly dominant in the considered encryption/decryption scheme.

Assuming that n and t are the parameters, for specification of the proposed encryption/decryption, the following notation is employed:

(i) $\mathbf{M}$ is n-dimensional binary vector of data which should be encrypted;

(ii) $\mathbf{C}$ is n-dimensional binary vector of keystream for stream ciphering;

(iii) $\mathbf{G}'$ is $(n+t)$-dimensional binary vector of keystream nonoverlapping with $\mathbf{C}$;

(iv) $\mathbf{G}$ is $(n+t)$-dimensional binary vector of the weight exactly t obtained by a deterministic mapping of $\mathbf{G}'$;

(v) $\mathbf{X}$ is n-dimensional binary vector defined as $\mathbf{X} = \mathbf{M} \oplus \mathbf{C}$;

(vi) $\mathbf{Y}$ is $(n+t)$-dimensional binary vector which is equal to $\mathbf{X}$ with t inserted random bits.

The proposed encryption/decryption is displayed in Figure 1.

3. Information-Theoretic Analysis

This section yields an information-theoretic analysis of a (statistical) model of the considered encryption displayed in Figure 1.

A random variable is denoted by an uppercase letter (e.g., X) and its realization is denoted by a lowercase letter (e.g., x). An index (subscript) denotes discrete time. A discrete-time sequence of n random variables, for example, $X_1, X_2, \ldots, X_n$, is shortly denoted by $X^n = (X_1, X_2, \ldots, X_n)$. Since our channel has synchronization errors, we have a need to distinguish *strings* from *sequences*. We denote a random string (indexed by discrete-time k) as $Y_{(k)}$. The string $Y_{(k)}$ may not have a fixed length, and we denote its length (which is a random variable if the string itself is a random variable) as $\mathscr{L}(Y_{(k)})$. A concatenation of two strings a and b is denoted by $a \parallel b$. As short notation, we denote the concatenation of n strings $Y_{(1)}$ through $Y_{(n)}$ as $Y^{(n)} = Y_{(1)} \parallel Y_{(2)} \parallel \cdots \parallel Y_{(n)}$. The entropy of a random object X is denoted by $H(X)$, and the mutual information between two random objects X and Y is denoted by $I(X;Y)$. The binary entropy function is denoted by $h(p) = -p\log_2 p - (1-p)\log_2(1-p)$.

Let the channel input X_k be a binary random variable drawn from the alphabet $\mathscr{X} = \{0,1\}$. The vector of all channel inputs up to time n is denoted by $X^n \triangleq (X_1, X_2, \ldots, X_n)$. The transmitter (Alice) observes the pseudo-random sequence $G^n \triangleq (G_1, G_2, \ldots, G_n)$ provided by a shared source of randomness (shared with Bob) and uses it to create a channel output (ciphertext) $Y^{(n)}$. Even though G^n is a pseudo-random sequence, we assume that the variables G_k are statistically indistinguishable from independent and identically distributed (iid) *geometric random variables* with parameter i; that is, for any integer $\ell \geq 0$, we have

$$\Pr\{G_k = \ell\} = (1-i)\, i^{\ell}. \tag{2}$$

Here, the parameter i denotes the *insertion probability*. Namely, between any two symbols X_k and X_{k+1}, Alice inserts a string $B_{(k)}$ that consists of Bernoulli-1/2 random variables, such that the length of $B_{(k)}$ equals $\mathscr{L}(B_{(k)}) = G_k$. Since G^n is a sequence of iid geometric random variables with parameter i, it is clear that Alice's transmission scheme is equivalent to randomly inserting a Bernoulli-1/2 random variable at any point of time during the communication. Formally, we state

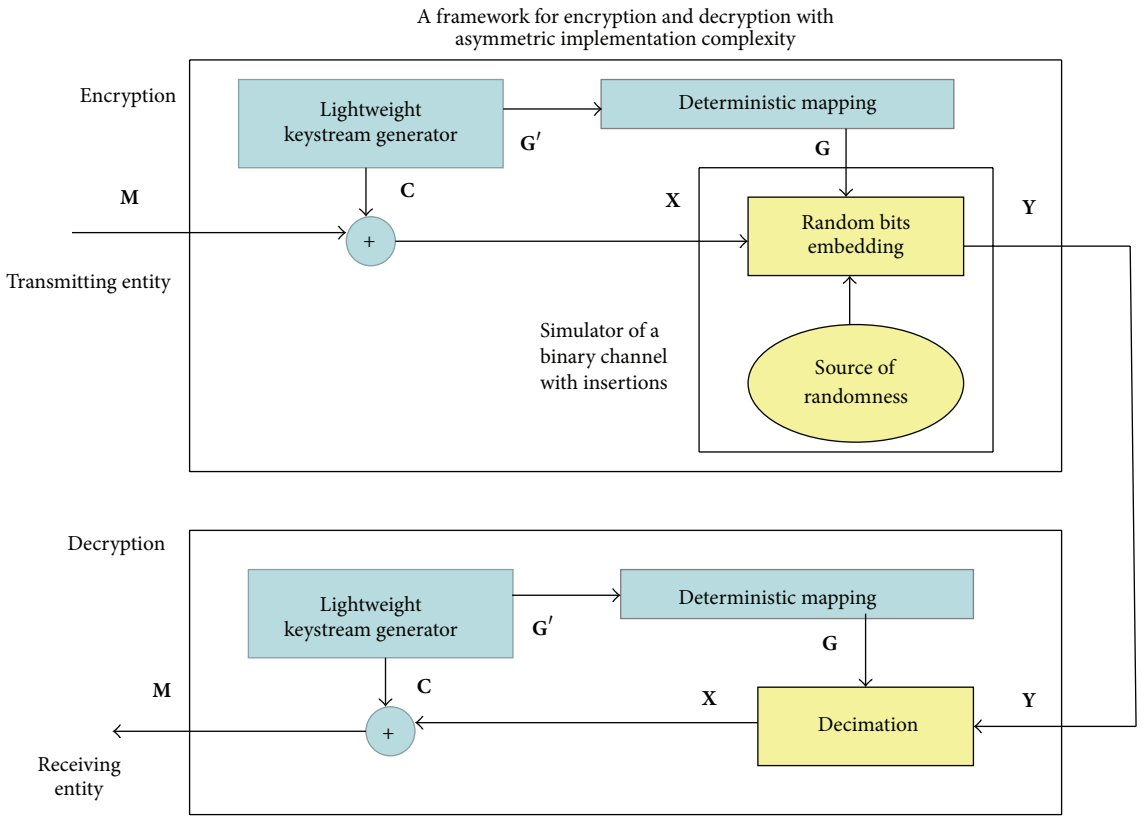

Figure 1: Encryption/decryption technique for scenarios with one-way communications between the entities with high performance computing capabilities and the very tiny ones.

that Alice creates a *string* $Y^{(n)}$ obtained as a *concatenation* of individual strings $Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$, that is,

$$Y^{(n)} = Y_{(1)} \parallel Y_{(2)} \parallel \cdots \parallel Y_{(n)}, \tag{3}$$

where each individual string $Y_{(k)}$ is obtained as

$$Y_{(k)} = X_k \parallel B_{(k)}. \tag{4}$$

The length of the string $Y^{(n)}$ equals

$$\begin{aligned} \mathscr{L}\left(Y^{(n)}\right) &= n + \sum_{k=1}^{n} G_k, \\ E\left[\mathscr{L}\left(Y^{(n)}\right)\right] &= \frac{n}{1-i}; \end{aligned} \tag{5}$$

that is, on average, Alice inserts $i/(1-i)$ Bernoulli-1/2 random variables between any two symbols X_k and X_{k+1}.

Eve (the eavesdropper) and Bob (the intended receiver) both receive the string $Y^{(n)}$ containing the randomly inserted symbols. The eavesdropper, not having access to the shared source of randomness G^n, cannot easily parse the string $Y^{(n)}$ to recover X^n. The intended receiver, on the other hand, has access to G^n, and since G_k represents the length of the inserted string between any two symbols X_k and X_{k+1}, the intended receiver (Bob) can easily remove the inserted symbols $\underline{B}_k$ from $Y^{(n)}$ (i.e., decimate $Y^{(n)}$) to recover X^n. In other words, by sharing the source of randomness G^n, Bob can resynchronize himself with Alice; see Figure 1.

The sequence C^n is a pseudo-random sequence, but for the purpose of computing information-theoretic quantities, we assume that C^n is modeled to be statistically indistinguishable from a sequence of iid Bernoulli-1/2 random variables. (It should not be understood that C^n implements a one-time pad. The variables C_k are only *statistically* modeled as Bernoulli-1/2 for the purposes of deriving (and computing) some information-theoretic quantities that we later use to derive a cryptographic security measure.)

Here, no assumptions are made on the statistical properties of the message M^n, but because C^n is iid Bernoulli-1/2, we have that X^n is also iid Bernoulli-1/2. Hence, the information-theoretic quantity of interest is the *iud information rate* defined as the information rate between X^n and $Y^{(n)}$ when the symbols X_k are independent and uniformly distributed (iud):

$$\mathscr{I}_{\text{iud}}(X;Y) \triangleq \lim_{n\to\infty} \frac{1}{n} I\left(X^n; Y^{(n)}\right)\Big|_{p(x^n)=2^{-n}}. \tag{6}$$

The information rate $\mathscr{I}_{\text{iud}}(X;Y)$ represents the amount of information that the eavesdropper can "*learn*," on average, about X after observing Y. The information rate $\mathscr{I}_{\text{iud}}(X;Y)$ is

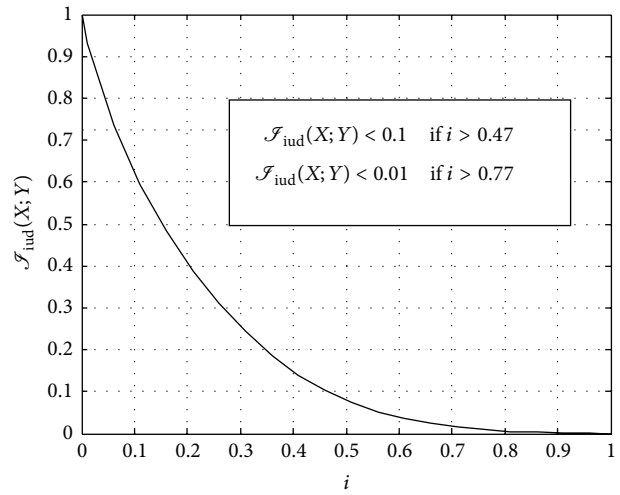

FIGURE 2: Information rate $\mathscr{I}_{\text{iud}}(X;Y)$ as a function of insertion probability i.

not computable in closed-form but is attainable using Monde Carlo techniques. For example, known bounds are [10]

$$\mathscr{I}_{\text{iud}}(X;Y) \geq \frac{1}{n} I\left(X^n; Y^{(n)}\right)\Big|_{p(x^n)=2^{-n}} - \frac{1}{n} H\left(\mathscr{L}\left(Y^{(n)}\right)\right), \tag{7}$$

$$\mathscr{I}_{\text{iud}}(X;Y) \leq \frac{1}{n} I\left(X^n; Y^{(n)}\right)\Big|_{p(x^n)=2^{-n}}. \tag{8}$$

For large n, the correction term $(1/n)H(\mathscr{L}(Y^{(n)}))$ in (7) equals

$$\frac{1}{n} H\left(\mathscr{L}\left(Y^{(n)}\right)\right) = \frac{1}{2n}\log_2\left(\frac{2\pi e \cdot i \cdot n}{(1-i)^2}\right) + O\left(n^{-2}\right). \tag{9}$$

If our desired accuracy of computing (bounding) $\mathscr{I}_{\text{iud}}(X;Y)$ is 10^{-4} and if $i = 0.95$, considerations of (7)–(9) dictate that $n \geq 1.5 \cdot 10^5$. For details on how to compute $\mathscr{I}_{\text{iud}}(X;Y)$ using "rhomboidal" trellis techniques such that both the desired correction term (9) and the confidence interval are kept under a predetermined accuracy (e.g., 10^{-4}), see [10]. Here, we only give numerical results in Figure 2, which reveal that the information rate $\mathscr{I}_{\text{iud}}(X;Y)$ is only a small fraction of the entropy rate $H(X_k) = 1$, especially when $i > 0.5$. These results are very favorable for secret communication because only a small fraction of the uncertainty in X^n can be learned from observing $Y^{(n)}$, as the next section demonstrates.

We already established that learning X after observing Y is extremely unfavorable for the eavesdropper because the information rate $\mathscr{I}_{\text{iud}}(X;Y)$ is low for large insertion probabilities i. However, the eavesdropper may adopt a strategy in which she first attempts to *learn* the sequence G^n and then attempt to crack X^n. To study the effects of this strategy, let us define the following quantities:

$$\mathscr{I}_{\text{iud}}(G;Y) \triangleq \lim_{n\to\infty} \frac{1}{n} I\left(G^n; Y^{(n)}\right)\Big|_{p(x^n)=2^{-n}},$$

$$\mathscr{I}_{\text{iud}}(X,G;Y) \triangleq \lim_{n\to\infty} \frac{1}{n} I\left(X^n, G^n; Y^{(n)}\right)\Big|_{p(x^n)=2^{-n}},$$

$$\mathscr{I}_{\text{iud}}(X;Y \mid G) \triangleq \lim_{n\to\infty} \frac{1}{n} I\left(X^n; Y^{(n)} \mid G^n\right)\Big|_{p(x^n)=2^{-n}},$$

$$\mathscr{I}_{\text{iud}}(G;Y \mid X) \triangleq \lim_{n\to\infty} \frac{1}{n} I\left(G^n; Y^{(n)} \mid X^n\right)\Big|_{p(x^n)=2^{-n}}. \tag{10}$$

Proposition 1. *Consider*

$$\mathscr{I}_{iud}(G;Y) = 0, \tag{11}$$

$$\mathscr{I}_{iud}(X;Y \mid G) = 1, \tag{12}$$

$$\mathscr{I}_{iud}(X,G;Y) = 1, \tag{13}$$

$$\mathscr{I}_{iud}(G;Y \mid X) = 1 - \mathscr{I}_{iud}(X;Y). \tag{14}$$

Proof. First, notice that

$$\lim_{n\to\infty} \frac{H\left(Y^{(n)}\right)}{n} = \frac{1}{1-i} \tag{15}$$

because $Y^{(n)}$ is a string of Bernoulli-1/2 random variables whose length is $\mathscr{L}(Y^{(n)})$, and as $n \to \infty$, we have

$$\lim_{n\to\infty} \frac{\mathscr{L}\left(Y^{(n)}\right)}{n} \overset{\text{wp 1}}{=} \frac{\mathrm{E}\left[\mathscr{L}\left(Y^{(n)}\right)\right]}{n} = \frac{1}{1-i}. \tag{16}$$

Next, we also have

$$\lim_{n\to\infty} \frac{H\left(Y^{(n)} \mid G^n\right)}{n} = \frac{n + \mathrm{E}\left(\sum_{k=1}^{n} G_k\right)}{n} = \frac{1}{1-i}, \tag{17}$$

and (11) is now a direct consequence of (15) and (17). Equality (12) follows from the fact that X^n is uniquely determined (by decimation) if G^n and $Y^{(n)}$ are known; that is, $H(X^n \mid G^n, Y^{(n)}) = 0$. Finally, (13) follows by adding (11) to (12) and applying the chain rule for mutual information, and (14) follows from (13) also using the chain rule. □

By equality (11) of Proposition 1, it is clear that the eavesdropper cannot learn G^n simply by observing $Y^{(n)}$. Also, from Figure 2, it is clear that, from the eavesdropper's perspective, learning X^n from $Y^{(n)}$ is extremely unfavorable because she can only learn a small fraction $\mathscr{I}_{\text{iud}}(X;Y)$ of $H(X) \triangleq H(X_k) = 1$ by observing $Y^{(n)}$. However, equality (12) of Proposition 1 reveals a potential vulnerability in that if the eavesdropper were to somehow learn G^n, then secrecy would be lost because $\mathscr{I}_{\text{iud}}(X;Y \mid G) = H(X) = 1$. Since learning either G^n or X^n *individually* is not favorable to the eavesdropper, the eavesdropper's strategy could be to go after the pair (X,G). Indeed, equality (13) of Proposition 1 reveals that, theoretically, the eavesdropper could gain substantial knowledge of the pair (X,G) by observing $Y^{(n)}$. Even for large i, this posterior knowledge of the pair (X,G), quantified as $\mathscr{I}_{\text{iud}}(X,G;Y)$, is not a negligible fraction of the entropy

$$H(X,G) \triangleq H\left(X_k\right) + H\left(G_k\right) = 1 + \frac{h(i)}{1-i}. \tag{18}$$

In the next section, we further explore the cryptographic implications by studying the connection between computational complexity and the information-theoretic quantities.

4. Generic Framework for the Security Evaluation

Note that the above information-theoretic analysis is based on modeling the pseudo-random sequence C^n as a random sequence. In this section, we now take into account the fact that the sequence is indeed pseudo-random. We show that the considered encryption (see Figure 1) based on employing the binary insertion channel $[X^n \to Y^{(n)}]$ provides enhanced security compared to the basic scheme that outputs only X^n.

4.1. Preliminaries: Security Notation. A definition of security consists of two distinct components: a specification of the assumed power of the adversary and a description of what constitutes a "break" of the scheme. Generally speaking, a cryptographic scheme is secure in a computational sense, if, for every probabilistic polynomial-time adversary $\mathscr{A}$ carrying out an attack of some specified type and for every polynomial $p(n)$, there exists an integer N such that the probability that $\mathscr{A}$ succeeds in this attack (where success is also well defined) is less than $1/p(n)$ for every $n > N$. Accordingly, the following two definitions specify a security evaluation scenario and a security statement.

Definition 2. The adversarial indistinguishability experiment consists of the following steps:

(1) The adversary $\mathscr{A}$ chooses a pair of messages $(\mathbf{m}_0; \mathbf{m}_1)$ of the same length n and passes them onto the encryption system for encrypting.

(2) A bit $b \in \{0, 1\}$ is chosen uniformly at random, and only one of the two messages $(\mathbf{m}_0; \mathbf{m}_1)$, precisely $\mathbf{m}_b$, is encrypted into ciphertext $\mathrm{Enc}(\mathbf{m}_b)$ and returned to $\mathscr{A}$.

(3) Upon observing $\mathrm{Enc}(\mathbf{m}_b)$, and without knowledge of b, the adversary $\mathscr{A}$ outputs a bit b_0.

(4) The experiment output is defined to be 1 if $b_0 = b$, and 0 otherwise; if the experiment output is 1, denoted shortly as the event $(\mathscr{A} \to 1)$, one says that $\mathscr{A}$ has succeeded.

Definition 3. An encryption scheme provides indistinguishable encryptions in the presence of an eavesdropper, if for all probabilistic polynomial-time adversaries $\mathscr{A}$

$$\Pr\left[\mathscr{A} \longrightarrow 1 \mid \mathrm{Enc}\left(\mathbf{m}_b\right)\right] \leq \frac{1}{2} + \epsilon, \tag{19}$$

where $\epsilon = \mathrm{negl}(n)$ is a negligibly small function.

Definitions 2 and 3 are more precisely discussed in [11].

4.2. Evaluation of the Security Gain Based on the Mutual Information. We consider the encryption system displayed in Figure 1 taking into account the fact that the legitimate parties share pseudo-random secret sequences instead of random ones. Our goal is to estimate the advantage of $\mathscr{A}$ in the indistinguishability game specified by Definition 2 when $\mathbf{y} \leftarrow \mathrm{Enc}(\mathbf{m}_b)$, where $\mathbf{y}$ is a particular realization of $Y^{(n)}$, assuming that the advantage of $\mathscr{A}$ is known when $\mathbf{m}_0$ and $\mathbf{m}_1$ are two chosen realizations of M^n and the corresponding realization of X^n is known.

Proposition 4. *Let the encrypted mapping of M^n into X^n be such that $1/2 + \epsilon$ equals the advantage of the adversary $\mathscr{A}$ (specified by Definition 3) to win the indistinguishability game (specified by Definition 2), and let the mutual information $\mathscr{I}_{iud}(X; Y)$ be known. Under these assumptions, for large n,*

$$\Pr\left[\mathscr{A} \longrightarrow 1 \mid Y^{(n)} = \mathbf{y}\right] = \frac{1}{2} + \epsilon \cdot \delta, \quad \textit{where } \delta \triangleq \Pr\left(X^n = \mathbf{x}_b \mid Y^{(n)} = \mathbf{y}\right) < \frac{1}{n} + \frac{1}{n} I\left(X^n, Y^{(n)}\right)\Big|_{p(x^n)=2^{-n}}. \tag{20}$$

Proof. Note that, for simplicity of the proof, Proposition 4 addresses a restricted case where it is assumed that $1/2 + \epsilon$ equals the advantage of the adversary $\mathscr{A}$ (specified by Definition 3) to win the indistinguishability game. Let the index b of the selected message be realization of the random variable B whose distribution reflects that of the output of adversary $\mathscr{A}$. The probability $\Pr(B = b \mid Y^{(n)} = \mathbf{y})$ that $\mathscr{A}$ wins the game is determined by the following:

$$\begin{aligned}
\Pr\left(B = b \mid Y^{(n)} = \mathbf{y}\right) &= \frac{\Pr\left(B = b, Y^{(n)} = \mathbf{y}\right)}{\Pr\left(Y^{(n)} = \mathbf{y}\right)} \\
&= \frac{\sum_{\mathbf{x}} \Pr\left(B = b, Y^{(n)} = \mathbf{y}, X^n = \mathbf{x}\right)}{\Pr\left(Y^{(n)} = \mathbf{y}\right)} \\
&= \frac{\sum_{\mathbf{x}} \Pr\left(B = b \mid Y^{(n)} = \mathbf{y}, X^n = \mathbf{x}\right) \Pr\left(Y^{(n)} = \mathbf{y}, X^n = \mathbf{x}\right)}{\Pr\left(Y^{(n)} = \mathbf{y}\right)} \\
&= \frac{\sum_{\mathbf{x}} \Pr\left(B = b \mid X^n = \mathbf{x}\right) \Pr\left(Y^{(n)} = \mathbf{y}, X^n = \mathbf{x}\right)}{\Pr\left(Y^{(n)} = \mathbf{y}\right)}.
\end{aligned} \tag{21}$$

According to the proposition assumption we have

$$\Pr\left(B = b \mid X^n = \mathbf{x}_b\right) = \frac{1}{2} + \epsilon, \tag{22}$$

where $\mathbf{x}_b$ corresponds to the selected $\mathbf{m}_b$, and

$$\Pr\left(B = b \mid X^n = \mathbf{x}\right) = \frac{1}{2} \quad \text{for any } \mathbf{x} \neq \mathbf{x}_b. \tag{23}$$

Consequently,

$$\begin{aligned}\Pr\left(B=b \mid Y^{(n)}=\mathbf{y}\right) &= \frac{\Pr\left(B=b \mid X^n=\mathbf{x}_b\right)\Pr\left(Y^{(n)}=\mathbf{y}, X^n=\mathbf{x}_b\right)}{\Pr\left(Y^{(n)}=\mathbf{y}\right)} \\ &\quad + \frac{\sum_{\mathbf{x}:\mathbf{x}\neq\mathbf{x}_b}\Pr\left(B=b \mid X^n=\mathbf{x}\right)\Pr\left(Y^{(n)}=\mathbf{y}, X^n=\mathbf{x}\right)}{\Pr\left(Y^{(n)}=\mathbf{y}\right)}, \\ \Pr\left(B=b \mid Y^{(n)}=\mathbf{y}\right) &= \frac{(1/2+\epsilon)\Pr\left(Y^{(n)}=\mathbf{y}, X^n=\mathbf{x}_b\right)-(1/2)\Pr\left(Y^{(n)}=\mathbf{y}, X^n=\mathbf{x}_b\right)}{\Pr\left(Y^{(n)}=\mathbf{y}\right)} \\ &\quad + \frac{(1/2)\sum_{\mathbf{x}}\Pr\left(Y^{(n)}=\mathbf{y}, X^n=\mathbf{x}\right)}{\Pr\left(Y^{(n)}=\mathbf{y}\right)} = \frac{1}{2}+\epsilon\cdot\Pr\left(X^n=\mathbf{x}_b \mid Y^{(n)}=\mathbf{y}\right).\end{aligned} \tag{24}$$

Next, we have the following general upper bound on the entropy (see [12] or [13], e.g.):

$$H\left(X^n \mid Y^{(n)}\right) \leq h\left(P_{\text{err}}\right) + P_{\text{err}}\log_2\left(2^n-1\right), \tag{25}$$

where $h(\cdot)\leq 1$ is the binary entropy function and $P_{\text{err}} = 1 - \Pr(\mathbf{x}_b \mid \mathbf{y})$, implying

$$\begin{aligned}\delta &\triangleq \Pr\left(X^n=\mathbf{x}_b \mid Y^{(n)}=\mathbf{y}\right) \\ &< \frac{1}{n}+1-\frac{1}{n}H\left(X^n \mid Y^{(n)}\right) \\ &= \frac{1}{n}+\frac{1}{n}I\left(X^n, Y^{(n)}\right)\Big|_{p(x^n)=2^{-n}}.\end{aligned} \tag{26}$$

□

5. Evaluation of the Security Gain Based on Numerical Estimation of the Mutual Information

Theorem 5. *Let the encrypted mapping of M^n into X^n be such that $1/2+\epsilon$ equals the advantage of the adversary $\mathscr{A}$ (specified by Definition 3) to win the indistinguishability game (specified by Definition 2), and let the mutual information $\mathscr{I}_{iud}(X;Y)$ be known (see Figure 2, e.g.). Under these assumptions, for large n,*

$$\begin{aligned}&\Pr\left[\mathscr{A}\longrightarrow 1 \mid Y^{(n)}=\mathbf{y}\right] = \frac{1}{2}+\epsilon\cdot\delta, \\ &\text{where } \delta < \mathscr{I}_{iud}(X;Y) + \frac{\log_2\left[(8\pi e\cdot i\cdot n)/(1-i)^2\right]}{2n} + O\left(n^{-2}\right).\end{aligned} \tag{27}$$

Proof. Consider

$$\begin{aligned}\delta &\triangleq \Pr\left(\mathbf{x}_b \mid \mathbf{y}\right) < \frac{1}{n}+1-\frac{1}{n}H\left(X^n \mid Y^{(n)}\right) \\ &= \frac{1}{n}+\frac{1}{n}I\left(X^n, Y^{(n)}\right)\Big|_{p(x^n)=2^{-n}}.\end{aligned} \tag{28}$$

Substitution of (7) and (9) into (28) finalizes the proof. □

Accordingly, the encryption mapping $M^n \rightarrow Y^{(n)}$ enhances security by a factor δ in comparison to the encryption mapping $M^n \rightarrow X^n$ because the probability that $\mathscr{A}$ wins the game becomes closer to 1/2, which corresponds to random guessing.

6. Evaluation of the Security Enhancement Employing Enumeration of Channel Input Candidates for the Given Output

6.1. Preliminaries. Let $\mathbf{Z} \in \{0,1\}^{\ell}$ be a binary string of length ℓ, and let $t \leq \ell$ be a parameter. Recently, in [9], improved bounds on the number of subsequences obtained from a binary string $\mathbf{Z}$ of length ℓ under t deletions have been reported. It is known that the number of subsequences in this setting strongly depends on the number of runs in the string $\mathbf{Z}$, where a run is a maximal substring of the same character. The improved bounds are obtained by a structural analysis of the family of r-run strings $\mathbf{Z}$, an analysis in which the extremal strings with respect to the number of subsequences have been identified. Specifically, for every r, r-run strings with the minimum (resp., maximum) number of subsequences under any t deletions have been considered, an exact analysis of the number of subsequences of these extremal strings has been presented, and it has been shown that this number can be calculated in polynomial time.

Let $D_t(\mathbf{Z})$ be a set of subsequences of $\mathbf{Z}$ that can be obtained from $\mathbf{Z}$ after t deletions. The analysis of $D_t(\mathbf{Z})$ and its size are challenging as the number of subsequences of a string $\mathbf{Z}$ obtained by deletions not only depends on its length ℓ and the number t of deletions, but also strongly depends on its structure. For example, $D_t(0^{\ell})$ is of size 1 and equals the single string $0^{\ell-t}$. Clearly, $|D_t(\mathbf{Z})|$ is at most $2^{\ell-t}$ (as after t deletions we remain with a binary string of length $\ell-t$). It has been shown that the number of subsequences $|D_t(\mathbf{Z})|$ strongly depends on the *number of runs* r in the string. Here, a run is a maximal substring of the same character, and

the number of runs $r = \rho(\cdot)$ in a given string $\mathbf{Z}$ is denoted by $\rho(\mathbf{Z})$. It has been proven that

$$\binom{\rho(\mathbf{Z}) - t + 1}{t} \leq \left|D_t(\mathbf{Z})\right| \leq \binom{\rho(\mathbf{Z}) + t - 1}{t}. \tag{29}$$

Also, it has been shown that the maximal number of subsequences is obtained from certain strings $\mathbf{Z}$, known as cyclic strings $\mathbf{Z}_\ell^C$, in which $|\mathbf{Z}| = \rho(\mathbf{Z})$, and it has been shown that

$$\binom{\rho(\mathbf{Z}) - t + 1}{t} \leq \left|D_t(\mathbf{Z})\right| \leq \left|D_t\left(\mathbf{Z}_\ell^C\right)\right|, \tag{30}$$

which has been further improved so that the following has been shown:

$$\begin{aligned} \sum_{i=1}^{t} \binom{\rho(\mathbf{Z}) - t}{i} &= \left|D_t\left(\mathbf{Z}_r^C\right)\right| \leq \left|D_t(\mathbf{Z})\right|, \\ \left|D_t(\mathbf{Z})\right| &\leq \left|D_t\left(\mathbf{Z}_\ell^C\right)\right| = \sum_{i=1}^{t} \binom{\ell - t}{i}, \end{aligned} \tag{31}$$

where $\mathbf{Z}_r^C$ is a string of length r with r runs.

In [9], also a family of strings, named unbalanced strings, has been defined. A string is called unbalanced, if all of the runs of symbols in the string are of length 1, except for one run. Let $U_{\ell,r}^{(i)}$ be a binary string of length ℓ with r runs, in which all runs are of length 1, except for the ith run which is of length $\ell - r + 1$. Due to symmetry $|D_t(U_{\ell,r}^{(1)})| = |D_t(U_{\ell,r}^{(r)})|$, and consequently define

$$u(\ell, r, t) = \left|D_t\left(U_{\ell,r}^{(1)}\right)\right| = \left|D_t\left(U_{\ell,r}^{(r)}\right)\right|. \tag{32}$$

It has been shown in [9] that these extreme cases have the least number of subsequences among the unbalanced strings and also that they have the least number of subsequences among all strings. The following theorem has been proven in [9].

Theorem 6 (Theorem 3 [9]: closed-form formula for $u(\ell, r, t)$). *For all $t < \ell$, $2 < r \leq \ell$,*

(i) *when $r > t$,*

$$u(\ell, r, t) = d(r, t) + \sum_{i=t+r-\ell-1}^{t-2} d(r-2, i), \tag{33}$$

(ii) *when $r \leq t$,*

$$u(\ell, r, t) = 2 + \sum_{i=t+r-\ell-1}^{r-3} d(r-2, i), \tag{34}$$

where

$$d(r, i) = \left|D_i\left(\mathbf{Z}_r^C\right)\right| = \sum_{j=0}^{i} \binom{r-i}{j} \tag{35}$$

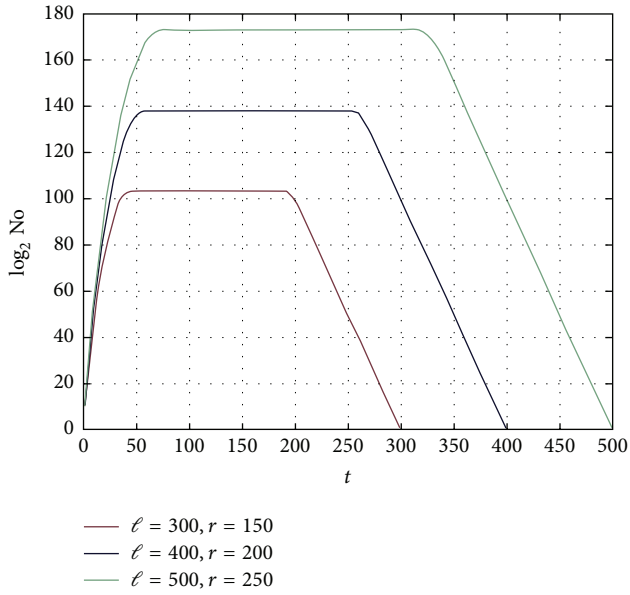

Figure 3: Number (No) of different subsequence of length ℓ which can be obtained from a binary sequence of length $\ell + t$: a numerical illustration of the statement of Theorem 3 [9].

assuming that $d(r, 0) = 1$ and, for $i < 0$, $d(r, i) = 0$ and that the following conventions are employed:

$$\begin{aligned} \sum_{i=j}^{k} a_i &= 0 \quad \textit{when } j > k, \\ \binom{\ell}{i} &= 0 \quad \textit{when } i < 0 \textit{ or } i > \ell. \end{aligned} \tag{36}$$

A numerical illustration of Theorem 6 is displayed in Figure 3.

6.2. Estimation of the Security Enhancement. Traditionally, as introduced in [14], the main information-theoretic security metric is the average information leaked, that is, the mutual information $I(\mathbf{M}; \mathbf{Y})$ between the message $\mathbf{M}$ and the related sample $\mathbf{Y}$, or, equivalently, the uncertainty, that is, the equivocation $H(\mathbf{M} \mid \mathbf{Y})$. Recently, certain information-theoretic security measures have been considered in [15] implying that, in our case, as a strong security metric the average mutual information $\overline{I}(\mathbf{M}, \mathbf{Y})$ should be addressed and $(1/n)\overline{I}(\mathbf{M}, \mathbf{Y})$ as a corresponding weak one.

Theorem 7. *Assuming that the employed keystream generator is such that the following is valid,*

$$\begin{aligned} I(\mathbf{M}; \mathbf{C}) &= 0, \\ I(\mathbf{M}; \mathbf{G}) &= 0, \\ I(\mathbf{C}; \mathbf{G}) &= 0, \\ I(\mathbf{M}; \mathbf{X}) &\leq \epsilon, \end{aligned} \tag{37}$$

the simulator of binary channel with random insertions provides

$$\frac{1}{n} I(\mathbf{M};\mathbf{Y}) \leq \frac{\alpha \cdot \epsilon}{n}, \qquad \alpha = 1 - \frac{1}{n}\log_2 (u(n+t,r,t)), \tag{38}$$

where $u(n+t,r,t)$ is the number of certain equally likely subsequences.

Sketch of the Proof. The uncertainty about the input (the argument) into a binary channel with random insertions given its output (the image) depends on the number of equally likely candidate arguments which can generate the given image. A lower bound on the number of these candidates can be obtained based on the lower bound on the number of the subsequences which can be obtained from the given one employing Theorem 6 (i.e., Theorem 3 from [9]). By adapting this result to the considered particular case we have the following. A lower bound on the number of the argument candidates $u(n+t,r,t)$, where r is a parameter, is given by (39) and (40):

(i) when $r > t$,

$$u(n+t,r,t) = d(r,t) + \sum_{i=r-n-1}^{m-2} d(r-2,i), \tag{39}$$

(ii) when $r \leq t$:

$$u(n+t,r,t) = 2 + \sum_{i=r-n-1}^{r-3} d(r-2,i), \tag{40}$$

where

$$d(r,i) = \sum_{j=0}^{i} \binom{r-i}{j} \tag{41}$$

assuming that $d(r,0) = 1$ and, for $i < 0$, $d(r,i) = 0$. Particularly note that the above enumerated subsequences are obtained from a sequence where all of the runs of symbols are of length 1, except for one run, and that the assumed decimation is a random one, and in addition, for simplicity of the evaluation we assume that the subsequences appear equally likely.

Consequently, the uncertainty $H(\mathbf{X} \mid \mathbf{Y})$ is lower-bounded as follows:

$$H(\mathbf{X} \mid \mathbf{Y}) \geq \log_2 (u(n+t,r,t)) \tag{42}$$

noting that $u(n+t,r,t)$ is at most $2^n = H(\mathbf{X})$ as after t deletions we remain with a binary string of length n. Taking into account that

$$\frac{1}{n} I(\mathbf{X};\mathbf{Y}) = \frac{1}{n}(H(\mathbf{X}) - H(\mathbf{X} \mid \mathbf{Y})) \tag{43}$$

we obtain

$$\frac{1}{n} I(\mathbf{M};\mathbf{Y}) \leq \frac{1}{n} I(\mathbf{M};\mathbf{X}) \left[1 - \frac{1}{n}\log_2 (u(n+t,r,t))\right] \tag{44}$$

and accordingly the theorem statement. □

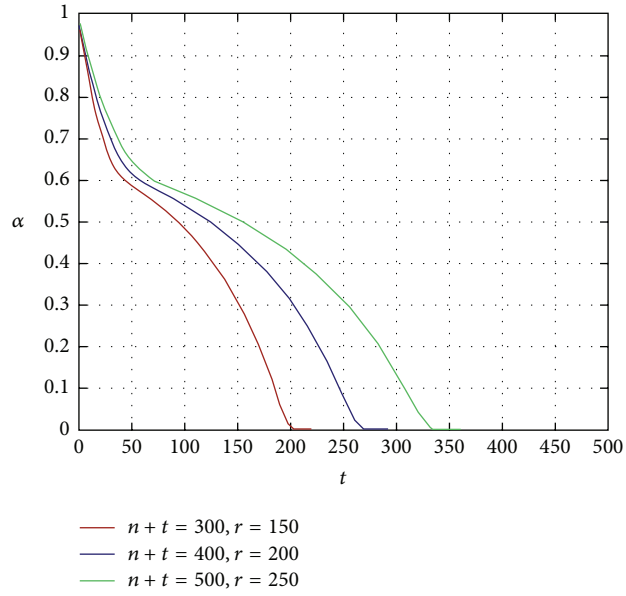

Figure 4: Numerical examples related to Theorem 7: illustration of the security gain implied by a binary channel with embedding of random bits noting that smaller α means higher security enhancement.

Figure 4 yields numerical illustrations of coefficient α which determines the security gain.

Note that, in order to achieve a desired high enhancement of the security, the insertion rate should be high enough as illustrated in Figure 4. When the insertion rate is low, the security enhancement is low as well, and this is analytically shown in the next corollary.

Corollary 8. *Consider*

$$\frac{1}{n} I(\mathbf{M};\mathbf{Y}) \leq \frac{1}{n} I(\mathbf{M};\mathbf{X}) \cdot \left(1 - \left(\log_2 \frac{1+\sqrt{5}}{2}\right)\frac{r}{n}\right) \tag{45}$$

when the parameters of the considered encryption fulfil the following constraints:

$$n > \frac{1+\sqrt{5}}{2} r, \qquad t \in [p^* r, n + t - r(1-p^*)] \quad \text{for } p^* \in [0.276, 0.278]. \tag{46}$$

Sketch of the Proof. For large values of t and r, the following approximation can be employed:

$$u(n+t,r,t) \approx \sum_{i=0}^{\min(r,t)} d(r,i), \tag{47}$$

where $x \approx y$ means that x is approximately y if x/y is a polynomial function of r and t. Accordingly,

$$d(r, pr) = \sum_{i=0}^{pr} \binom{r - pr}{i} \approx \begin{cases} 2^{r-pr} & \text{for } p \geq \frac{1}{3}, \\ \binom{r - pr}{pr} & \text{for } p < \frac{1}{3}. \end{cases} \quad (48)$$

Using the fact reported in [9] we have the following. Let $p^* = \arg\max_p d(r, pr)$. Numerical calculations reported in [9] show that $p^* \in [0.276, 0.278]$. Consequently, it is shown in [9] that for even r

$$d(r, p^* r) \approx \left(\frac{1 + \sqrt{5}}{2} \right)^r. \quad (49)$$

The above imply the corollary statement. □

Disclosure

This work was has been partially presented at IEEE Workshop on Information Theory, Korea, October 2015.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

The Ministry of Education, Science and Technological Development, Serbia, has partially funded this work.

References

[1] I. Ratković, N. Bežanić, O. S. Ünsal, A. Cristal, and V. Milutinović, "An overview of architecture-level power- and energy-efficient design techniques," *Advances in Computers*, vol. 98, pp. 1–57, 2015.

[2] M. J. Mihaljević, "A framework for stream ciphers based on pseudorandomness, randomness and error-correcting coding," in *Enhancing Cryptographic Primitives with Techniques from Error Correcting Codes*, B. Preneel, S. Dodunekov, V. Rijmen, and S. Nikova, Eds., vol. 23 of *NATO Science for Peace and Security Series D: Information and Communication Security*, pp. 117–139, IOS Press, Amsterdam, The Netherlands, 2009.

[3] M. J. Mihaljević and H. Imai, "An approach for stream ciphers design based on joint computing over random and secret data," *Computing*, vol. 85, no. 1-2, pp. 153–168, 2009.

[4] M. J. Mihaljević, "An approach for light-weight encryption employing dedicated coding," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM '12)*, pp. 892–898, Anaheim, Calif, USA, December 2012.

[5] M. J. Mihaljević, "On certain coding approaches for security evaluation and design of stream ciphers," *Transaction on Advanced Research*, vol. 8, no. 2, pp. 28–34, 2012.

[6] F. Oggier and M. J. Mihaljevic, "An information-theoretic security evaluation of a class of randomized encryption schemes," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 2, pp. 158–168, 2014.

[7] M. J. Mihaljević and K. Matsuura, "Evaluation of an approach for security enhancement of certain lightweight stream ciphers," in *Proceedings of the 32nd IEEE Symposium on Cryptography and Information Security (SCIS '15)*, Kokura, Japan, January 2015.

[8] A. Kavcic, M. J. Mihaljevic, and K. Matsuura, "Light-weight secrecy system using channels with insertion errors: cryptographic implications," in *Proceedings of the IEEE Information Theory Workshop (ITW '15)*, pp. 257–261, Jeju Island, South Korea, October 2015.

[9] Y. Liron and M. Langberg, "A characterization of the number of subsequences obtained via the deletion channel," *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2300–2312, 2015.

[10] J. Castiglione and A. Kavcic, "Trellis-based lower bounds on capacities of channels with synchronization errors," in *Proceedings of the IEEE Information Theory Workshop (ITW '15)*, pp. 11–15, Jeju Island, South Korea, October 2015.

[11] J. Katz and Y. Lindell, *Introduction to Modern Cryptography*, CRC Press, Boca Raton, Fla, USA, 2007.

[12] D. L. Tebbe and S. J. Dwyer III, "Uncertainty and the probability of error," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 516–518, 1968.

[13] M. Feder and N. Merhav, "Relations between entropy and error probability," *IEEE Transactions on Information Theory*, vol. 40, no. 1, pp. 259–266, 1994.

[14] C. E. Shannon, "Communication theory of secrecy systems," *The Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.

[15] M. R. Bloch and J. N. Laneman, "Strong secrecy from channel resolvability," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8077–8098, 2013.