sensors

MDPI

Article

A Generic Model of the Pseudo-Random Generator Based on Permutations Suitable for Security Solutions in Computationally-Constrained Environments

Tomislav Unkašević [1,*,†], Zoran Banjac [1,†] and Milan Milosavljević [2]

[1] VLATACOM Institute, 11070 Belgrade, Serbia; zoran.banjac@vlatacom.com
[2] Singidunum University, 11000 Belgrade, Serbia; mmilosavljevic@singidunum.ac.rs
* Correspondence: tomislav.unkasevic@vlatacom.com; Tel.: +381-11-377-115-0
† Current address: Milutina Milankovića Boulevard 5, Belgrade, Serbia.

Received: 31 October 2019; Accepted: 27 November 2019; Published: 3 December 2019

Abstract: Symmetric cryptography methods have an important role in security solutions design in data protection. In that context, symmetric cryptography algorithms and pseudo-random generators connected with them have strong influence on designed security solutions. In the computationally constrained environment, security efficiency is also important. In this paper we proposed the design of a new efficient pseudo-random generator parameterized by two pseudo-random sequences. By the probabilistic, information-theoretic and number theory methods we analyze characteristics of the generator. Analysis produced several results. We derived sufficient conditions, regarding parameterizing sequences, so that the output sequence has uniform distribution. Sufficient conditions under which there is no correlation between parameterizing sequences and output sequence are also derived. Moreover, it is shown that mutual information between the output sequence and parameterizing sequences tends to zero when the generated output sequence length tends to infinity. Regarding periodicity, it is shown that, with appropriately selected parameterizing sequences, the period of the generated sequence is significantly longer than the periods of the parameterizing sequences. All this characteristics are desirable regarding security applications. The efficiency of the proposed construction can be achieved by selection parameterizing sequences from the set of efficient pseudo-random number generators, for example, multiple linear feedback shift registers.

Keywords: pseudo-random generator; security; wireless sensor networks; IoT; probability distribution; correlation; information leakage

1. Introduction

The expansion of communication and network technologies, as well as technological advances in the design and implementation of microprocessor devices, have led to the ability to informational connecting different devices and creation of intelligent systems capable of monitoring and managing complex processes. Communication devices utilize the Internet infrastructure and protocols to create a world of connected devices, like Wireless Sensor Networks (WSN) and Internet of Things (IoT). This technological advancement enables the progress of many technological and life processes bringing to us smart cities, autonomous vehicles, robotization and intelligent robot behavior [1–4]. In that context, information security has a very important role in compromising the integrity and privacy of data in such an integrated world can cause serious damage, even to the level of a general disaster [5–9]. Therefore, in addition to security mechanisms incorporated into Internet protocols, additional security mechanisms incorporated into devices and systems are used to prevent unintended behavior. Moreover,

a huge number of that type of devices (sensors, cameras, surveillance systems) need to work in real time fashion so that the defined security mechanisms do not disrupt system behavior. They must be designed in such a way that it is easy to implement them both in hardware and software and their application should not disrupt system behavior i.e., they must be efficient [10,11].

There are various IoT applications, connected with different types of sensors, that have become an integral part of our lives, and most of them can be classified in common areas such as smart healthcare services, smart home, intelligent transportation, smart grid, etc. However, as a consequence of mass deployment, many IoT challenges have arisen, such as limited processing capability and memory resources, large amount of data to transmit, different operating characteristics of hardware, and heterogeneous data and networks types [12–16]. Moreover, personal privacy, data confidentiality and integrity are also a great challenge of IoT that must be overcome, particularly for devices with limited resources and heterogeneous technologies [13,14,17–19]. Cryptography can be used to protect confidentiality (or secrecy) of data and communication. It can also be used to ensure the integrity (or accuracy) of information as well as for authentication (and non-repudiation) services [20]. An important point in the IoT world is that most IoT solutions have a "closed design", so it is often very difficult or even impossible to incorporate additional security mechanisms after the production process is completed. On the other hand, as a consequence of the limited software and hardware resources of IoT devices, the suite of cryptographic algorithms that can be implemented is narrowing, so the right measure must be found between the desired level of security and implementation capabilities, which makes the security issue even more challenging [13,16]. Different cryptographic algorithms that offer roughly the same level of security may require different power and resource consumption, so you need to choose the right one, subject to the limitations of some specific IoT application and deployed hardware [19]. Given that public-key crypto algorithms, compared to symmetric crypto algorithms, have far greater power and resource consumption due to their high processing time [21], it is a natural choice to use a symmetric algorithm in IoT security solution design. Detailed analysis and comparison of symmetric block-type algorithms such as AES, RC6, Twofish, SPECK128, LEA, and ChaCha20-Poly1305 algorithms in IoT devices are given in [16]. On the other hand, stream or sequential symmetric key ciphers are typically faster than block-type. Block ciphers, in general, require more memory resources to encrypt/decrypt larger chunks (block) of data, while sequential ciphers usually take only one or a few bits at a time, they have relatively low memory requirements and therefore are suitable to implement in limited scenarios. Stream cryptography algorithms, as a subgroup of symmetric cryptography algorithms, are among the most common cryptography data protection techniques. The idea comes from Shannon's one-time pad system, where instead of a random sequence of encryption bits, a series of bits obtained from a pseudo-random generator is used [20]. The sequence generated by the pseudo-random generator is used for plain-text encryption and its properties determine the security of the protected data. Therefore, the basic cryptography goal in stream cryptography systems is to design pseudo-random bit/symbol generators with good cryptographic characteristics. Many ideas have been implemented in the last fifty years, with more or less success.

One of the most popular and widely used pseudo-random sequence generator is the RC4 generator defined in 1987 by Ron Rivest. The description of the algorithm was revealed by reverse engineering of the RSA INC software [22], and the correctness of the algorithm description obtained was confirmed by Rivest himself [23,24].

The RC4 algorithm owes its popularity to its simplicity and ease of implementation in both software and hardware. The high popularity and applicability has attracted the attention of the cryptanalytic community. The results of a deep and thorough analysis of this algorithm led to the detection of a number of weaknesses of the algorithm. A comprehensive review of the weaknesses identified is given in [25] where the empirically detected weaknesses are theoretically proved as well as the original results of the authors of the article. The compromitation of this algorithm was additionally

contributed by the implementation methods in security protocols, so that its use in security protocols has not been recommended since 2015 [26].

On the other hand, the beauty and elegance of the idea itself suggest the possibility of its exploitation.

Our work is aimed to define low complexity and efficient generic model of the pseudo-random generator that does not suffer from the weaknesses immanent to the RC4 and which is suitable for the implementation of the security solution in the computational constrained microprocessor environments, e.g., WSN and IoT. This paper defines a pseudo-random generator that can, in some way, be considered as a generalization of ideas related to RC4 because it uses the time varying permutations, sequences for permutation changing and addressing output element from the current generator state. In order to prove plausible cryptographic properties of the proposed pseudo-random generator different mathematical techniques are used to analyze probability distribution of the output sequence, correlation properties, information leakage between the state of the generator and output sequence and periodicity.

The paper is organized in four parts. After the first part containing motivation for this work and introduction in the second part we introduce necessary notation, describe proposed pseudo-random generator and his relationship to RC4 in brief. The third part contains the analysis of the generator, some comments and remarks. The fourth part contains a summary of the paper's results.

2. Notation and Generating Algorithm Description

Let $I_k = \{0, 1, \ldots, k-1\}$. Then with $\mathcal{P}_k$ we will denote a set of all bijections from I_k to I_k, and as usual, its elements will be named permutations. Set $\mathcal{P}_k$ is a totally ordered set by the, so called, lexicographic order and have exactly $k!$ elements, where ! denotes factorial operation. Elements of the set $\mathcal{P}_k$ we will denote by $\Pi_1, \Pi_2, \ldots, \Pi_{k!}$.

By $P\{X\}$ we will denote the probability of set X.

Let $S = \{f_0, f_1, \ldots, f_{m-1}\} \subseteq \mathcal{P}_k$ be a set of permutations on I_k with the following properties:

1. For any $\Pi_j \in \mathcal{P}_k$, $j = 1, 2, \ldots, k!$, exists a number $l > 0$ and set of integers $i_1, i_2, \ldots, i_l \in I_m$ such that $\Pi_j = f_{i_l} \circ f_{i_{l-1}} \circ \ldots \circ f_{i_1}$, where $\circ$ is composition of functions.
2. For all $p, q \in I_m$ if $p \neq q$ then $f_p \neq f_q$.
3. $\Pi_1 = I \in S$ where I is identical permutation.

Let $\{A_n\}_{n=1}^{\infty}$ and $\{C_n\}_{n=1}^{\infty}$ be a two sequences of independent identically distributed random variables with $P\{A_n = l\} = a_l, \quad l = 0, 1, ..., k-1$ and $P\{C_n = l\} = c_l, \quad l = 0, 1, ..., (m-1)$.

Then we will define a pseudo-random sequence $\{Z_n\}_{n=1}^{\infty}$ with equations

$$\begin{array}{ll} g_0 = p & p \in \mathcal{P}_k \\ g_{n+1} = f_{C_{n+1}} \circ g_n & n \geq 0 \\ Z_{n+1} = g_{n+1}(A_{n+1}) & n \geq 0. \end{array} \tag{1}$$

From (1) generating algorithm for the sequence $\{Z_n\}_{n=1}^{\infty}$ is obvious. As a first step we will construct the sequence of permutations $\{g_n\}_{n=0}^{\infty}$, $g_n \in \mathcal{P}_k$ using sequence $\{C_n\}_{n=1}^{\infty}$ and element Z_n of the sequence $\{Z_n\}_{n=1}^{\infty}$ is computed as a value of the function g_n at the point A_n. Graphical presentation of the sequence generating process is given on the Figure 1.

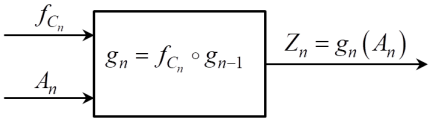

Figure 1. The generator graphic presentation.

Defined generator algorithm apply time-varying permutations as well as the RC4 but in RC4 fixed set of permutations, set of transpositions, is used. Graphical presentation of the RC4 algorithm is given on the Figure 2. where the summatios and numbers with denote reduction modulo 256.

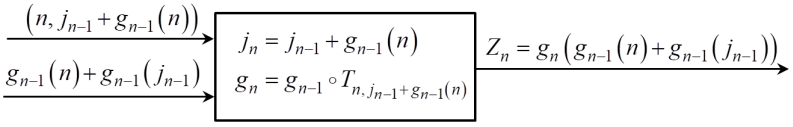

Figure 2. The RC4 graphic presentation, all numbers and summations assumes reduction modulo 256.

In the defined generator case any set S which is generator of $\mathcal{P}_k$ can be used. Sequences that are used in RC4 applied transposition determination and address of permutation table position corresponds to sequences $\{C_n\}_{n=1}^{\infty}$ and $\{A_n\}_{n=1}^{\infty}$ in our algorithm respectively. While the mentioned sequences from RC4 are precisely defined, in our case sequences $\{C_n\}_{n=1}^{\infty}$ and $\{A_n\}_{n=1}^{\infty}$ are arbitrarily chosen and the generator is parameterized by those two sequences. Later in the paper we define sufficient conditions for sequences $\{C_n\}_{n=1}^{\infty}$ and $\{A_n\}_{n=1}^{\infty}$ to achieve good pseudo-random and security properties.

3. Analysis of the Generator

For every pseudo-random generator it is necessary to analyze its properties regarding the possibilities of output sequence prediction or reconstruction of the generator initial state. In that sense desirable properties are uniform distribution of the output sequence, nonexistence of the correlation between the output sequence and elements of the generator, nonexistence of the output sequence auto-correlation and long period of the output sequence. These features are especially important for generators used in security solutions and lack any of them usually have serious consequences on the security of the system. Different examples can be found in [20,27].

3.1. Distribution of the Generated Sequence

Intuitively one can expect that $\{Z_n\}_{n=1}^{\infty}$ has a uniform distribution but relatively weak constraints demanded for the sequences $\{A_n\}_{n=1}^{\infty}$ and $\{C_n\}_{n=1}^{\infty}$ require formal proof for the expectations. By the next theorem we will show that $\{Z_n\}_{n=1}^{\infty}$ has an asymptotically uniform distribution.

Theorem 1. *If*

1. $\sum_{i=1}^{k} a_i = 1$ *and* $a_i > 0$ *for all* $i \in I_k$,
2. $\sum_{i=0}^{m-1} c_i = 1$ *and* $c_i > 0$ *for all* $i \in I_m$,

then pseudo-random sequence $\{Z_n\}_{n=1}^{\infty}$ *has an asymptotically uniform distribution i.e.,*

$$(\forall l \in I_k) \quad \left(\lim_{n\to\infty} P\{Z_n = l\} = \frac{1}{k}\right)$$

for $l \in I_k$.

The proof of the Theorem 1 will be derived in two steps. First, using Markov chains theory, we will show that sequence $\{g_n\}_{n=0}^{\infty}$ has asymptotically uniform distribution and after that, in the second step, using that result we will show the statement of the theorem.

Proof. First we will analyze the sequence of functions $\{g_n\}_{n=0}^{\infty}$. It is obvious that $g_n \in \mathcal{P}_k$, as a consequence of $(\mathcal{P}_k, \circ)$ being group. Next we will show that

$$(\forall i \in \{1, 2, \ldots, k!\}) \left(\lim_{n\to\infty} P\{g_n = \Pi_i\} = \frac{1}{k!} \right) \tag{2}$$

To prove (2) we will observe the sequence $\{g_n\}_{n=0}^{\infty}$ as a stationary Markov chain over the set of states $\mathcal{P}_k$. Indeed, according to the definition of $\{g_n\}_{n=0}^{\infty}$, transition from the g_n to g_{n+1} doesn't depend on the history of g_n, but only on the current state g_n and the value of C_n.

Denote by $G_n = [P\{g_n = \Pi_i\}]_{1\times k!}$ a row matrix whose elements are the probabilities that after n steps the chain is in the state Π_i. Let t_{ij} be the probability that the chain changes state from Π_i to Π_j in one step and $T = \left[t_{ij}\right]_{k!\times k!}$ be one step transition matrix for the Markov chain. Denote by $T_n = \left[t_{ij}^n\right]_{k!\times k!}$ n-step probability transition matrix of that system starting at the state Π_i changes to state Π_j after exactly n steps. It is well known, (see [28,29]), that $T_n = T^n$ and that,

$$\begin{aligned} G_n &= G_0 T_n \\ \lim_{n\to\infty} G_n &= \lim_{n\to\infty} G_0 T_n = G_0 \lim_{n\to\infty} T_n \end{aligned} \tag{3}$$

When the limit values in (3) exists. To show that $\lim_{n\to\infty} T_n$ exists it is sufficient to show that such $n_0 \in \mathbb{N}$ exists for which $t_{ij}^{n_0} > 0$ for all $i, j \in \{1, 2, ..., k!\}$ (see [28,29]).

Let us define numbers n_{ij} as

$$n_{ij} = \min_r \left\{ r \mid (\exists (i_1, i_2, \ldots, i_r) \in I_m^r) \left(f_{i_r} \circ \ldots \circ f_{i_2} \circ f_{i_1} = \Pi_j \circ \Pi_i^{-1} \right) \right\}. \tag{4}$$

Due to the properties of the set S it is clear that $n_{ij} > 0$. Let n_0 be $\max_{i,j\in\{1,2,\ldots,k!\}} n_{ij}$ and show that $t_{ij}^{n_0}$ is greater than zero. Because $(\mathcal{P}_k, \circ)$ is group then the equation $x \circ \Pi_i = \Pi_j$ has exactly one solution, $\Pi_j \circ \Pi_i^{-1}$, so we can write

$$\begin{aligned} t_{ij}^{n_0} &= \sum_{(i_1,\ldots,i_{n_0})} P\left\{ f_{i_{n_0}} \circ \ldots \circ f_{i_2} \circ f_{i_1} \circ \Pi_i = \Pi_j \right\} \\ &= \sum_{(i_1,\ldots,i_{n_0})} P\left\{ f_{i_{n_0}} \circ \ldots \circ f_{i_2} \circ f_{i_1} = \Pi_j \circ \Pi_i^{-1} \right\} \\ &= \sum_{\substack{(i_1,\ldots,i_{n_0}) \\ f_{i_{n_0}} \circ \ldots \circ f_{i_2} \circ f_{i_1} = \Pi_i^{-1} \circ \Pi_j}} \prod_{l=1}^{n_0} P\{C_l = i_l\} \end{aligned} \tag{5}$$

Now, to prove $t_{ij}^{n_0} > 0$ it is sufficient to show that at least one summand in (5) is greater then zero, see [28,29]. Because $n_{ij} > 0$ we can find a set of indices $\left\{ i_1, i_2, \ldots, i_{n_{ij}} \right\}$, $n_{ij} \leq n_0$ such that $f_{i_{n_{ij}}} \circ \ldots \circ f_{i_2} \circ f_{i_1} = \Pi_i^{-1} \circ \Pi_j$. Because the index of identical permutation is 1, then the summand which corresponds to the set of indices $\left\{ i_1, i_2, \ldots, i_{n_{ij}}, \underbrace{1, 1, \ldots, 1}_{n_0 - n_{ij}} \right\}$ is evidently greater then zero and we showed that $\lim_{n\to\infty} T_n$ exists. Because convergence is component wise, $\lim_{n\to\infty} t_{ij}^n = t_j^*$

exists. Limit value $\lim_{n\to\infty} T_n$ can be determined as a solution of the system of a equations known as Champman-Kolmogorov equations.

$$\begin{aligned} \sum_{j=1}^{k!} t_j^* t_{jl} &= t_l^* \\ \sum_{j=1}^{k!} t_j^* &= 1 \end{aligned} \tag{6}$$

It is easy to check, by substitution, that the $t_1^* = t_2^* = \ldots = t_{k!}^* = \frac{1}{k!}$ is unique solution of the system (6).

From now on the proof is straightforward. Let $l \in I_k$ be arbitrary, then

$$\begin{aligned} P\{Z_n = l\} &= P\{g_n(A_n) = l\} \\ &= \sum_{i=1}^{k!} P\{g_n(A_n) = l \mid g_n = \Pi_i\}\ P\{g_n = \Pi_i\} \\ &= \sum_{i=1}^{k!} P\{\Pi_i(A_n) = l \mid g_n = \Pi_i\}\ P\{g_n = \Pi_i\} \\ &= \sum_{i=1}^{k!}\sum_{j=0}^{k-1} P\{\Pi_i(j) = l \mid A_n = j\}\ P\{A_n = j\}\, P\{g_n = \Pi_i\}. \end{aligned} \tag{7}$$

Because $\{\Pi_i(j) = l \mid\}$ and $\{A_n = j\}$ are independent random variables it follows from (7) that

$$\begin{aligned} P\{Z_n = l\} &= \sum_{i=1}^{k!}\sum_{j=0}^{k-1} P\{\Pi_i(j) = l \mid A_n = j\}\ P\{A_n = j\}\, P\{g_n = \Pi_i\} \\ &= \sum_{i=1}^{k!}\sum_{j=0}^{k-1} P\{\Pi_i(j) = l\}\ P\{A_n = j\}\, P\{g_n = \Pi_i\} \\ &= \sum_{j=0}^{k-1} P\{A_n = j\} \sum_{i=1}^{k!} P\{g_n = \Pi_i\}\, P\{\Pi_i(j) = l\} \end{aligned} \tag{8}$$

Finding limit of the both sides of the (8) it follows

$$\begin{aligned} \lim_{n\to\infty} P\{Z_n = l\} &= \lim_{n\to\infty} \sum_{j=0}^{k-1} P\{A_n = j\} \sum_{i=1}^{k!} P\{g_n = \Pi_i\}\, P\{\Pi_i(j) = l\} \\ &= \sum_{j=0}^{k-1} P\{A_n = j\} \sum_{i=1}^{k!} P\{\Pi_i(j) = l\} \lim_{n\to\infty} P\{g_n = \Pi_i\} \end{aligned} \tag{9}$$

because $P\{A_n = j\}$, $P\{\Pi_i(j) = l\}$ do not depend on n. Using that $\lim_{n\to\infty} \mathrm{P}\{g_n = \Pi_i\} = \frac{1}{k!}$ from (9) it follows that

$$\begin{aligned} \lim_{n\to\infty} P\{Z_n = l\} &= \sum_{j=0}^{k-1} P\{A_n = j\} \sum_{i=1}^{k!} P\{\Pi_i(j) = l\} \lim_{n\to\infty} P\{g_n = \Pi_i\} \\ &= \frac{1}{k!} \sum_{j=0}^{k-1} P\{A_n = j\} \sum_{i=1}^{k!} P\{\Pi_i(j) = l\} \\ &= \frac{1}{k!} \sum_{j=0}^{k-1} P\{A_n = j\}\,(k-1)! \\ &= \frac{(k-1)!}{k!} \sum_{j=0}^{k-1} P\{A_n = j\} \end{aligned}$$

$$= \frac{1}{k}$$

which proves the theorem. □

Remark 1. *Asymptotically uniform distribution of the sequence* Z_n, $n = 1, 2, \ldots\ldots$, *as we have shown, is a consequence of the asymptotically uniform distribution of* $\{g_n\}_{n=0}^{\infty}$. *If* $G_0 = \left[\frac{1}{k!}, \frac{1}{k!}, \ldots, \frac{1}{k!}_i\right]_{1\times k!}$ *it is easy to verify that* $\{g_n\}_{n=0}^{\infty}$ *has a uniform distribution and as a consequence* Z_n, $n = 1, 2, \ldots$ *has a uniform distribution too.*

Theorem 2. *If the random variables* A_n, $n = 1, 2, \ldots\ldots$ *are uniformly distributed then for all* $z \in I_k$

$$P\{Z_n = z\} = \frac{1}{k}$$

Proof. By the generator definition we have

$$\begin{aligned} P\{Z_n = z\} &= \sum_{i=1}^{k!}\sum_{j=0}^{k-1} P\{Z_n = z \,|\, g_n = \Pi_i \wedge A_n = j\} \cdot P\{g_n = \Pi_i \wedge A_n = j\} \\ &= \sum_{i=1}^{k!}\sum_{j=0}^{k-1} P\{\Pi_i(j) = z\} \cdot P\{g_n = \Pi_i\} \cdot P\{A_n = j\} \end{aligned} \tag{10}$$

because A_n is uniformly distributed from (10) it follows that

$$\begin{aligned} P\{Z_n = z\} &= \sum_{i=1}^{k!}\sum_{j=0}^{k-1} P\{\Pi_i(j) = z\} \cdot P\{g_n = \Pi_i\} \cdot \frac{1}{k} \\ &= \frac{1}{k}\sum_{i=1}^{k!}\sum_{j=0}^{k-1} P\{\Pi_i(j) = z\} \cdot P\{g_n = \Pi_i\} \\ &= \frac{1}{k}\sum_{i=1}^{k!} P\{g_n = \Pi_i\} \sum_{j=0}^{k-1} P\{\Pi_i(j) = z\} \\ &= \frac{1}{k} \end{aligned}$$

because $\sum_{i=1}^{k!} P\{g_n = \Pi_i\} = 1$ and $\sum_{j=0}^{k-1} P\{\Pi_i(j) = z\} = 1$ and theorem is proved. □

By these theorems we showed that generator has at least asymptotically uniform distribution of the values in the generator output sequence. This is important security feature because it indicates impossibility of the prediction of the generator output sequence based on the probability distribution of the output sequence values.

Correlation Properties

Theorem 3. *If the random variables* A_n, $n = 1, 2, \ldots\ldots$ *are uniformly distributed then for all* $a, b \in I_k$,

1. $P\{Z_{n+l} = b \wedge Z_n = a\} = \frac{1}{k^2}$
2. $P\{Z_{n+l} = b | Z_n = a\} = \frac{1}{k}$

Proof.

1. By the generator definition we have

$$P\{Z_{n+l} = b \wedge Z_n = a\} = P\{g_{n+k}(A_{n+k}) = b \wedge g_n(A_n) = a\} =$$

$$= \sum_{i=1}^{k!}\sum_{j=1}^{k!} P\left\{g_{n+k}\left(A_{n+k}\right) = b \wedge g_n\left(A_n\right) = a \,|\, g_{n+k} = \Pi_i \wedge g_n = \Pi_j\right\}$$
$$\cdot P\left\{g_{n+k} = \Pi_i \wedge g_n = \Pi_j\right\} \quad (11)$$
$$= \sum_{i=1}^{k!}\sum_{j=1}^{k!} P\left\{\Pi_i\left(A_{n+k}\right) = b \wedge \Pi_j\left(A_n\right) = a\right\} \cdot P\left\{g_{n+k} = \Pi_i \wedge g_n = \Pi_j\right\}$$
$$= \sum_{i=1}^{k!}\sum_{j=1}^{k!} P\left\{\Pi_i\left(A_{n+k}\right) = b \wedge \Pi_j\left(A_n\right) = a\right\} \cdot P\left\{g_{n+k} = \Pi_i \,|\, g_n = \Pi_j\right\}$$
$$\cdot P\left\{g_n = \Pi_j\right\}.$$

Now, using notation from the Theorem 1 $P\left\{g_{n+k} = \Pi_i \,|\, g_n = \Pi_j\right\}$ is equal to $t_{i,j}^k$ and putting it in (11) it follows that

$$P\left\{Z_{n+l} = b \wedge Z_n = a\right\} =$$
$$= \sum_{i=1}^{k!}\sum_{j=1}^{k!} P\left\{\Pi_i\left(A_{n+k}\right) = b \wedge \Pi_j\left(A_n\right) = a\right\} \cdot P\left\{g_{n+k} = \Pi_i \,|\, g_n = \Pi_j\right\} \quad (12)$$
$$\cdot P\left\{g_n = \Pi_j\right\}$$
$$= \sum_{i=1}^{k!}\sum_{j=1}^{k!} P\left\{\Pi_i\left(A_{n+k}\right) = b \wedge \Pi_j\left(A_n\right) = a\right\} \cdot t_{i,j}^k \cdot P\left\{g_n = \Pi_j\right\}.$$

Because Π_i, Π_j are permutations $P\left\{\Pi_i\left(A_{n+k}\right) = b \wedge \Pi_j\left(A_n\right) = a\right\}$ is equal to $P\left\{A_{n+k} = \Pi_i^{-1}(b) \wedge A_n = \Pi_j^{-1}(a)\right\}$ and using that A_{n+k}, A_n are independent random variables from (12) it follows that

$$P\left\{Z_{n+l} = b \wedge Z_n = a\right\} =$$
$$= \sum_{i=1}^{k!}\sum_{j=1}^{k!} P\left\{\Pi_i\left(A_{n+k}\right) = b \wedge \Pi_j\left(A_n\right) = a\right\} \cdot t_{i,j}^k \cdot P\left\{g_n = \Pi_j\right\} \quad (13)$$
$$= \sum_{i=1}^{k!}\sum_{j=1}^{k!} P\left\{A_{n+k} = \Pi_i^{-1}(b) \wedge A_n = \Pi_j^{-1}(a)\right\} \cdot t_{i,j}^k \cdot P\left\{g_n = \Pi_j\right\}$$
$$= \sum_{i=1}^{k!}\sum_{j=1}^{k!} P\left\{A_{n+k} = \Pi_i^{-1}(b)\right\} \cdot P\left\{A_n = \Pi_j^{-1}(a)\right\} \cdot t_{i,j}^k \cdot P\left\{g_n = \Pi_j\right\}$$

Now, using that A_{n+k} and A_n are uniformly distributed independent random variables it follows from (13) that

$$P\left\{Z_{n+l} = b \wedge Z_n = a\right\} =$$
$$= \sum_{i=1}^{k!}\sum_{j=1}^{k!} P\left\{A_{n+k} = \Pi_i^{-1}(b)\right\} \cdot P\left\{A_n = \Pi_j^{-1}(a)\right\} \cdot t_{i,j}^k \cdot P\left\{g_n = \Pi_j\right\}$$
$$= \sum_{i=1}^{k!}\sum_{j=1}^{k!} \frac{1}{k} \cdot \frac{1}{k} \cdot t_{i,j}^k \cdot P\left\{g_n = \Pi_j\right\} \quad (14)$$
$$= \frac{1}{k^2}\sum_{i=1}^{k!} P\left\{g_n = \Pi_i\right\} \sum_{j=1}^{k!} t_{i,j}^k$$
$$= \frac{1}{k^2}$$

because $\sum_{j=1}^{k!} t_{i,j}^{k} = 1$ and $\sum_{i=1}^{k!} P\{g_n = \Pi_i\} = 1$, and statement is proved.

2. Using statement of the Theorem 2 that $P\{Z_n = a\} = \frac{1}{k}$ by the definition of conditional probability it follows that

$$P\{Z_{n+l} = b | Z_n = a\} = \frac{P\{Z_{n+k} = b \wedge Z_n = a\}}{P\{Z_n = a\}} = \frac{\frac{1}{k^2}}{\frac{1}{k}} = \frac{1}{k}$$

which proves the statement.

□

3.2. Information Leakage

Information leakage means existence of the correlation between generator output sequence and elements of its inner state. Such type correlation may be base for the process of reconstruction of the some generator state during his generation history. Knowledge of the one state during the generator work and knowledge of the generator algorithm allows prediction of the future elements of the output sequence which is undesirable in security applications. In this part, correlation with the state element sequences $\{A_n\}_{n=1}^{\infty}$ and $\{C_n\}_{n=1}^{\infty}$ with $\{Z_n\}_{n=1}^{\infty}$ is considered.

Theorem 4. *Under the conditions of the Theorem 1 we have*

1. $\lim_{n\to\infty} P(Z_n = z | C_n = c) = \frac{1}{k}$
2. $\lim_{n\to\infty} I(Z_n, C_n) = 0$

where $z \in I_k$ and $c \in I_m$

Proof.

1. By the definition

$$\begin{aligned} &P(Z_n = z | C_n = c) = \\ &= \frac{P(Z_n = z \wedge C_n = c)}{P(C_n = c)} = \frac{P(g_n(A_n) = z \wedge C_n = c)}{P(C_n = c)} \\ &= \frac{P\left(\left(\bigcup_{i=1}^{k!} g_n = \Pi_i \wedge \Pi_i(A_n) = z\right) \wedge C_n = c\right)}{P(C_n = c)} \\ &= \frac{P\left(\left(\bigcup_{i=1}^{k!} f_{C_n} \circ g_{n-1} = \Pi_i \wedge \Pi_i(A_n) = z\right) \wedge C_n = c\right)}{P(C_n = c)} \\ &= \frac{P\left(\bigcup_{i=1}^{k!} ((f_{C_n} \circ g_{n-1} = \Pi_i \wedge \Pi_i(A_n) = z) \wedge C_n = c)\right)}{P(C_n = c)}. \end{aligned} \tag{15}$$

Because $\{f_{C_n} \circ g_{n-1} = \Pi_i\}$ and $\{f_{C_n} \circ g_{n-1} = \Pi_j\}$ are disjoint events for $i, j \in I_k$ and $i \neq j$, from (15) it follows that

$$\begin{aligned} P(Z_n = z | C_n = c) &= \frac{P\left(\bigcup_{i=1}^{k!} ((f_{C_n} \circ g_{n-1} = \Pi_i \wedge \Pi_i(A_n) = z) \wedge C_n = c)\right)}{P(C_n = c)} \\ &= \frac{\sum_{i=1}^{k!} P(((f_{C_n} \circ g_{n-1} = \Pi_i \wedge \Pi_i(A_n) = z) \wedge C_n = c))}{P(C_n = c)} \end{aligned} \tag{16}$$

Now, using Bayes theorem from (16) it follows that

$$\begin{aligned} P\left(Z_n = z | C_n = c\right) &= \frac{\sum_{i=1}^{k!} P\left(\left(\left(f_{C_n} \circ g_{n-1} = \Pi_i \wedge \Pi_i\left(A_n\right) = z\right) \wedge C_n = c\right)\right)}{P\left(C_n = c\right)} \\ &= \frac{\sum_{i=1}^{k!} \sum_{j=0}^{k-1} P\left(\left(\left(f_{C_n} \circ g_{n-1} = \Pi_i \wedge \Pi_i\left(A_n\right) = z\right) \wedge C_n = c\right) \mid \left(A_n = j\right)\right) \cdot P\left(A_n = j\right)}{P\left(C_n = c\right)} \\ &= \frac{\sum_{i=1}^{k!} \sum_{j=0}^{k-1} P\left(\left(\left(f_{C_n} \circ g_{n-1} = \Pi_i \wedge \Pi_i\left(j\right) = z\right) \wedge C_n = c\right)\right) \cdot P\left(A_n = j\right)}{P\left(C_n = c\right)} \\ &= \frac{\sum_{i=1}^{k!} \sum_{j=0}^{k-1} P\left(\left(g_{n-1} = f_{C_n}^{-1} \circ \Pi_i \wedge C_n = c\right) \wedge \Pi_i\left(j\right) = z\right) \cdot P\left(A_n = j\right)}{P\left(C_n = c\right)}. \end{aligned} \tag{17}$$

Now, because events $\left\{g_{n-1} = f_{C_n}^{-1} \circ \Pi_i \wedge C_n = c\right\}$ and $\left\{\Pi_i\left(j\right) = z\right\}$ are independent from (17) it follows that

$$\begin{aligned} &P\left(Z_n = z | C_n = c\right) = \\ &= \frac{\sum_{i=1}^{k!} \sum_{j=0}^{k-1} P\left(\left(g_{n-1} = f_{C_n}^{-1} \circ \Pi_i \wedge C_n = c\right) \wedge \Pi_i\left(j\right) = z\right) \cdot P\left(A_n = j\right)}{P\left(C_n = c\right)} \\ &= \frac{\sum_{i=1}^{k!} \sum_{j=0}^{k-1} P\left(g_{n-1} = f_{C_n}^{-1} \circ \Pi_i \wedge C_n = c\right) \cdot P\left(\Pi_i\left(j\right) = z\right) \cdot P\left(A_n = j\right)}{P\left(C_n = c\right)} \end{aligned} \tag{18}$$

Grouping summands which depends on j from (18) it follows that

$$\begin{aligned} &P\left(Z_n = z | C_n = c\right) = \\ &= \frac{\sum_{i=1}^{k!} \sum_{j=0}^{k-1} P\left(g_{n-1} = f_{C_n}^{-1} \circ \Pi_i \wedge C_n = c\right) \cdot P\left(\Pi_i\left(j\right) = z\right) \cdot P\left(A_n = j\right)}{P\left(C_n = c\right)} \\ &= \sum_{i=1}^{k!} \frac{P\left(g_{n-1} = f_{C_n}^{-1} \circ \Pi_i \wedge C_n = c\right) \sum_{j=0}^{k-1} P\left(A_n = j\right) \cdot P\left(\Pi_i\left(j\right) = z\right)}{P\left(C_n = c\right)} \\ &= \sum_{i=1}^{k!} \frac{P\left(g_{n-1} = f_{C_n}^{-1} \circ \Pi_i \wedge C_n = c\right)}{P\left(C_n = c\right)} \sum_{j=0}^{k-1} P\left(A_n = j\right) \cdot P\left(\Pi_i\left(j\right) = z\right) \\ &= \sum_{i=1}^{k!} P\left(g_{n-1} = f_{C_n}^{-1} \circ \Pi_i | C_n = c\right) \sum_{j=0}^{k-1} P\left(A_n = j\right) \cdot P\left(\Pi_i\left(j\right) = z\right) \\ &= \sum_{i=1}^{k!} P\left(g_{n-1} = f_c^{-1} \circ \Pi_i\right) \sum_{j=0}^{k-1} P\left(A_n = j\right) \cdot P\left(\Pi_i\left(j\right) = z\right) \end{aligned} \tag{19}$$

Taking the limit from the both sides in (19) it follows

$$\lim_{n \to \infty} P\left(Z_n = z | C_n = c\right) =$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \sum_{j=0}^{k-1} P\left(A_n = j\right) \sum_{i=1}^{k!} P\left(g_{n-1} = f_c^{-1} \circ \Pi_i\right) \cdot P\left(\Pi_i\left(j\right) = z\right) \\
&= \sum_{j=0}^{k-1} P\left(A_n = j\right) \sum_{i=1}^{k!} P\left(\Pi_i\left(j\right) = z\right) \cdot \lim_{n\to\infty} P\left(g_{n-1} = f_c^{-1} \circ \Pi_i\right) = \\
&= \sum_{j=0}^{k-1} P\left(A_n = j\right) \sum_{i=1}^{k!} P\left(\Pi_i\left(j\right) = z\right) \cdot \frac{1}{k!} \\
&= \frac{1}{k!} \sum_{j=0}^{k-1} P\left(A_n = j\right) \sum_{i=1}^{k!} P\left(\Pi_i\left(j\right) = z\right) \\
&= \frac{1}{k!} \sum_{j=0}^{k-1} P\left(A_n = j\right) \cdot (k-1)! = \\
&= \frac{(k-1)!}{k!} \sum_{j=0}^{k-1} P\left(A_n = j\right) = \frac{1}{k}
\end{aligned}
$$

which proves the statement.

2. By the definition of mutual information we have that

$$I\left(Z_n, C_n\right) = H\left(Z_n\right) - H\left(Z_n|C_n\right) \tag{20}$$

We will start with computing $H\left(Z_n\right)$.

$$H\left(Z_n\right) = \sum_{i=0}^{k-1} P\left(Z_n = i\right) \log_2 \frac{1}{P\left(Z_n = i\right)} \tag{21}$$

Taking the limit from the both sides in (21) and using Theorem 1 we have

$$\lim_{n\to\infty} H\left(Z_n\right) = \log_2 k. \tag{22}$$

In the same way it follows that

$$
\begin{aligned}
H\left(Z_n|C_n\right) &= \sum_{i=0}^{m-1} P\left(C_n = i\right) \cdot H\left(Z_n|C_n = i\right) = \\
&= \sum_{i=0}^{m-1} P\left(C_n = i\right) \cdot \sum_{j=0}^{k-1} P\left(Z_n = j|C_n = c_i\right) \log_2 \frac{1}{P\left(Z_n = j|C_n = i\right)}
\end{aligned} \tag{23}
$$

Taking the limit from both sides in (23) and using part 1 of the Theorem 4 we obtain

$$
\begin{aligned}
&\lim_{n\to\infty} H\left(Z_n|C_n\right) = \\
&= \sum_{i=0}^{m-1} P\left(C_n = i\right) \cdot \sum_{j=0}^{k-1} \lim_{n\to\infty} P\left(Z_n = j|C_n = i\right) \cdot \lim_{n\to\infty} \log_2 \frac{1}{P\left(Z_n = j|C_n = i\right)} \\
&= \sum_{i=0}^{m-1} P\left(C_n = i\right) \cdot \sum_{j=0}^{k-1} \frac{1}{k} \log_2 k = \log_2 k \cdot \sum_{i=0}^{m-1} P\left(C_n = i\right) = \log_2 k.
\end{aligned} \tag{24}
$$

Using (22) and(24) in (20) it follows that

$$\lim_{n\to\infty} I\left(Z_n, C_n\right) = \lim_{n\to\infty} H\left(Z_n\right) - \lim_{n\to\infty} H\left(Z_n|C_n\right) = \log_2 k - \log_2 k = 0 \tag{25}$$

which proves the statement. □

Theorem 5. *Under the conditions of the Theorem 1 we have*

1. $\lim_{n\to\infty} P\left(Z_n = z | A_n = a\right) = \frac{1}{k}$
2. $\lim_{n\to\infty} I\left(Z_n, A_n\right) = 0$

where $z, a \in I_k$.

Proof.

1. By the definition of conditional probability it follows that

$$\begin{aligned} & P\left(Z_n = z | A_n = a\right) = \\ & = \frac{P\left(Z_n = z \wedge A_n = a\right)}{P\left(A_n = a\right)} = \frac{P\left(g_n\left(A_n\right) = z \wedge A_n = a\right)}{P\left(A_n = a\right)} \\ & = \frac{P\left(\left(\bigcup_{i=1}^{k!} g_n = \Pi_i \wedge \Pi_i\left(A_n\right) = z\right) \wedge A_n = a\right)}{P\left(A_n = a\right)} \\ & = \frac{P\left(\left(\bigcup_{i=1}^{k!} g_n = \Pi_i \wedge \Pi_i\left(A_n\right) = z\right) \wedge A_n = a\right)}{P\left(A_n = a\right)} \\ & = \frac{P\left(\bigcup_{i=1}^{k!} \left(\left(\left(g_n = \Pi_i \wedge \Pi_i\left(A_n\right) = z\right) \wedge A_n = a\right)\right)\right)}{P\left(A_n = a\right)}. \end{aligned} \tag{26}$$

Because $\left(\left(\left(g_n = \Pi_i \wedge \Pi_i\left(A_n\right) = z\right) \wedge A_n = a\right)\right)$ and $\left(\left(\left(g_n = \Pi_j \wedge \Pi_j\left(A_n\right) = z\right) \wedge A_n = a\right)\right)$ are disjoint events when $i \neq j$ from (26) it follows that

$$\begin{aligned} & P\left(Z_n = z | A_n = a\right) = \\ & = \frac{P\left(\bigcup_{i=1}^{k!} \left(\left(\left(g_n = \Pi_i \wedge \Pi_i\left(A_n\right) = z\right) \wedge A_n = a\right)\right)\right)}{P\left(A_n = a\right)} \\ & = \frac{\sum_{i=1}^{k!} P\left(\left(g_n = \Pi_i \wedge \left(\Pi_i\left(A_n\right) = z \wedge A_n = a\right)\right)\right)}{P\left(A_n = a\right)}. \end{aligned} \tag{27}$$

Using that $P\left(A \wedge B\right) = P\left(A|B\right) \cdot P\left(B\right)$ from (27) it follows that

$$\begin{aligned} & P\left(Z_n = z | A_n = a\right) = \\ & = \frac{\sum_{i=1}^{k!} P\left(\left(g_n = \Pi_i \wedge \left(\Pi_i\left(A_n\right) = z \wedge A_n = a\right)\right)\right)}{P\left(A_n = a\right)} \\ & = \frac{\sum_{i=1}^{k!} P\left(g_n = \Pi_i | \left(\Pi_i\left(A_n\right) = z \wedge A_n = a\right)\right) \cdot P\left(\Pi_i\left(a\right) = z \wedge A_n = a\right)}{P\left(A_n = a\right)} \\ & = \sum_{i=1}^{k!} P\left(g_n = \Pi_i | \left(\Pi_i\left(A_n\right) = z \wedge A_n = a\right)\right) \cdot \frac{P\left(\Pi_i\left(a\right) = z \wedge A_n = a\right)}{P\left(A_n = a\right)} \\ & = \sum_{i=1}^{k!} P\left(g_n = \Pi_i \,|\, \left(\Pi_i\left(A_n\right) = z \wedge A_n = a\right)\right) \cdot P\left(\Pi_i\left(A_n\right) = z \,|\, A_n = a\right). \end{aligned} \tag{28}$$

Because $\{g_n = \Pi_i\}$ and $\{\Pi_i(A_n) = z \wedge A_n = a\}$ are independent random variables from (28) it follows that

$$\begin{aligned} & P(Z_n = z | A_n = a) = \\ & = \frac{\sum_{i=1}^{k!} P((g_n = \Pi_i \wedge (\Pi_i(A_n) = z \wedge A_n = a)))}{P(A_n = a)} \\ & = \frac{\sum_{i=1}^{k!} P(g_n = \Pi_i | (\Pi_i(A_n) = z \wedge A_n = a)) \cdot P(\Pi_i(a) = z \wedge A_n = a)}{P(A_n = a)} = \\ & = \sum_{i=1}^{k!} P(g_n = \Pi_i \mid (\Pi_i(A_n) = z \wedge A_n = a)) \cdot P(\Pi_i(A_n) = z \mid A_n = a) \\ & = \sum_{i=1}^{k!} P(g_n = \Pi_i) \cdot P(\Pi_i(a) = z) \end{aligned} \tag{29}$$

Taking the limits from the both sides in (29) it follows

$$\begin{aligned} & \lim_{n\to\infty} P(Z_n = z | A_n = a) = \\ & = \lim_{n\to\infty} \sum_{i=1}^{k!} P(g_n = \Pi_i) \cdot P(\Pi_i(a) = z) \\ & = \sum_{i=1}^{k!} \lim_{n\to\infty} P(g_n = \Pi_i) \cdot P(\Pi_i(a) = z) \\ & = \sum_{i=1}^{k!} \frac{1}{k!} \cdot P(\Pi_i(a) = z) \\ & = \frac{1}{k!} \sum_{i=1}^{k!} P(\Pi_i(a) = z) \\ & = \frac{1}{k!} \cdot (k-1)! = \frac{1}{k} \end{aligned}$$

2. In the same way as in the Theorem 4 it follows that

$$\lim_{n\to\infty} H(Z_n) = \log_2 k$$

By the definition of conditional entropy it follows that

$$\begin{aligned} H(Z_n | A_n) &= \sum_{i=0}^{k-1} P(A_n = i) \cdot H(Z_n | A_n = i) \\ &= \sum_{i=0}^{k-1} P(A_n = i) \cdot \sum_{j=0}^{k-1} P(Z_n = j | A_n = i) \log_2 \frac{1}{P(Z_n = j | A_n = i)} \end{aligned} \tag{30}$$

Taking the limit of both sides in (30) and statement of the part 1 we obtain

$$\begin{aligned} & \lim_{n\to\infty} H(Z_n | A_n) = \\ & = \sum_{i=0}^{k-1} P(A_n = i) \cdot \sum_{j=0}^{k-1} \lim_{n\to\infty} P(Z_n = j | A_n = i) \cdot \lim_{n\to\infty} \log_2 \frac{1}{P(Z_n = j | A_n = i)} \end{aligned} \tag{31}$$

$$= \sum_{i=0}^{k-1} P\left(A_n = i\right) \cdot \sum_{j=0}^{k-1} \frac{1}{k} \log_2 k = \log_2 k \cdot \sum_{i=0}^{k-1} P\left(A_n = i\right) = \log_2 k$$

And finally, using the (22) and (31) in the definition for the $I\left(Z_n, A_n\right)$, it follows that

$$\lim_{n\to\infty} I\left(Z_n, A_n\right) = \lim_{n\to\infty} H\left(Z_n\right) - \lim_{n\to\infty} H\left(Z_n|A_n\right) = \log_2 k - \log_2 k = 0$$

which proves the statement. □

3.3. Periodicity

Every pseudo-random generator can be viewed as a finite automaton with output over the finite set of states and symbols. Because the automaton transition function is deterministic it follows that the output sequence must be periodic. So, $\{A_n\}_{n=1}^{\infty}$, and $\{C_n\}_{n=1}^{\infty}$ are periodic and denote their periods by A and B respectively. It is easy to verify that $\{g_n\}_{n=0}^{\infty}$, and $\{Z_n\}_{n=1}^{\infty}$ are periodic too and denote their periods G, Z respectively. In this part relations between A, C, G and Z are considered and some sufficient conditions under A, C, GM are defined which improves the value of Z. For that we need a few Lemmas.

To find out period of $\{Z_n\}_{n=1}^{\infty}$ we will first determine the period of the $\{g_n\}_{n=0}^{\infty}$.

Lemma 1. *Denote by* $l \in \{1, 2, ..., k!\}$ *the order of the permutation* $\prod_{i=1}^{C} f_{C_i}$. *Then the period of* $\{g_n\}_{n=0}^{\infty}$ *is* lC.

Proof. First we have to prove that lC is a period, but it is straightforward.

$$g_{k+\lambda lC} = f_{C_1} \circ f_{C_2} \circ \ldots \circ f_{C_{\lambda lC}} \circ f_{C_{1+\lambda lC}} \circ \ldots \circ f_{C_{k+\lambda lC}}$$

and because C is period of the $\{C_n\}_{n=1}^{\infty}$ we have

$$\begin{aligned} & f_{C_1} \circ f_{C_2} \circ \ldots \circ f_{C_{\lambda lC}} \circ f_{C_{1+\lambda lC}} \circ \ldots \circ f_{C_{k+\lambda lC}} = \\ & = \left(f_{C_1} \circ f_{C_2} \circ \ldots \circ f_{C_C}\right)^{\lambda l} \circ f_{C_1} \circ \ldots \circ f_{C_k} \\ & = f_{C_1} \circ \ldots \circ f_{C_k} \\ & = g_k \end{aligned}$$

Next step is to prove that lC is the fundamental period i.e., that every other period is divisible by lC.

Suppose contrary, that lC isn't the fundamental period i.e., that the fundamental period is d, $d \mid lC$ and $d < lC$. From $g_{k+\lambda d} = g_k$ we have

$$\prod_{i=k+1}^{k+\lambda d} f_{C_i} = I,\ k \in \mathbb{N}, k \geq 1 \tag{32}$$

where I is the identical permutation. Multiplying (32) with f_{C_k} from the left we have

$$\prod_{i=k}^{k+\lambda d-1} f_{C_i} \circ f_{C_{k+\lambda d}} = f_{C_k}$$

and applying (32) we have

$$f_{C_{k+\lambda d}} = f_{C_k},\ k \geq 1.$$

From this we have

$$C_{k+\lambda d} = C_k, \quad k \geq 1$$

which means that d is a period of $\{C_n\}_{n=1}^{\infty}$ and that $C \,|d$ i.e., $d = rC$, $r < l$. Now, look at $g_{1+\lambda d}$

$$g_{1+\lambda d} = \prod_{i=1}^{1+\lambda d} f_{C_i} = \prod_{i=1}^{\lambda rC} f_{C_i} \circ f_{C_{1+\lambda rC}} =$$
$$= \left(\prod_{i=1}^{C} f_{C_i}\right)^{\lambda r} \circ f_{C_1} = f_{C_1} = g_1$$

and we conclude that

$$\left(\prod_{i=1}^{C} f_{C_i}\right)^{\lambda r} = I$$

for every $\lambda \geq 1$. Finally we have

$$\left(\prod_{i=1}^{C} f_{C_i}\right)^{r} = I$$

and conclude that $l|r$ which is in contradiction with $r < l$ so we proved that lC is the fundamental period of $\{g_n\}_{n=0}^{\infty}$. □

Lemma 2. *Let G be the fundamental period of $\{g_n\}_{n=0}^{\infty}$. If $(G, A) = 1$ and $\{A_1, A_2, \ldots, A_n, \ldots\} = I_k$ then period of $\{Z_n\}_{n=1}^{\infty}$ is GA.*

Proof. By straightforward computation we can easily check that GA is a period of the $\{Z_n\}_{n=1}^{\infty}$. We need only to prove that GA is fundamental period of $\{Z_n\}_{n=1}^{\infty}$.

Suppose that the fundamental period isn't GA but it is d. Then $d \mid GA$ and because $(G, A) = 1$ we have $d = d_1 d_2, (d_1, d_2) = 1$ and $d_1|G,\ d_2|A$.

Further,

$$g_{k+\lambda d}\left(A_{k+\lambda d}\right) = g_k\left(A_k\right) \tag{33}$$

for every $\lambda \geq 1, \lambda \in \mathbb{N}$. We can set $\lambda = \lambda_1 \cdot \frac{G}{d_1}, \lambda_1 \geq 1$ in the equation above which transform it to

$$g_{k+\lambda_1 G d_2}\left(A_{k+\lambda_1 G d_2}\right) = g_k\left(A_k\right)$$

and having in mind that G is a period of $\{g_n\}_{n=0}^{\infty}$ we obtain

$$g_k\left(A_{k+\lambda_1 G d_2}\right) = g_k\left(A_k\right).$$

From the fact that g_k is bijection it follows that

$$A_{k+\lambda_1 G d_2} = A_k$$

which means that Gd_2 is a period of $\{A_n\}_{n=1}^{\infty}$ and consequently that $A|Gd_2$. Using that $(G, A) = 1$ we have that $A|d_2$ and with $d_2|A$ we conclude that $A = d_2$. According to former observations we have that d must be of the form $d_1 A$ and rewriting of (33) yields

$$g_{k+\lambda d_1 A}\left(A_{k+\lambda d_1 A}\right) = g_k\left(A_k\right).$$

This equation we can simplify to

$$g_{k+\lambda d_1 A}\left(A_k\right) = g_k\left(A_k\right)$$

because A is a period of $\{A_n\}_{n=1}^{\infty}$. Now we put our attention to $g_{k+nG+\lambda d_1 A}\left(A_{k+nG+\lambda d_1 A}\right)$.

$$\begin{aligned} g_{k+\lambda d_1 A}\left(A_{k+nG}\right) &= g_{k+nG+\lambda d_1 A}\left(A_{k+nG+\lambda d_1 A}\right) = \\ &= g_{k+nG}\left(A_{k+nG}\right) = \\ &= g_k\left(A_{k+nG}\right) \end{aligned}$$

so we have that functions $g_{k+\lambda d_1 A}, g_k$ are equal on the set $\{A_{k+nG} \mid n \in \mathbb{N}\}$. Set $\{x \mid k+nG \equiv_A x,\ n \in \mathbb{N}\}$ is equal I_A, because $(G, A) = 1$, and from the periodicity of $\{A_n\}_{n=1}^{\infty}$ follows $\{A_{k+nG} \mid n \in \mathbb{N}\} = I_k$. Functions $g_{k+\lambda d_1 A}, g_k$ are equal on their domain so we have that

$$g_{k+\lambda d_1 A} = g_k\ , \lambda \geq 1, \lambda \in \mathbb{N}$$

which means that $d_1 A$ is a period of the $\{g_n\}_{n=0}^{\infty}$. In the same way as above we have that $d_1 = G$ and we have that the fundamental period of the $\{Z_n\}_{n=1}^{\infty}$ is GA. □

Following Corollary is a trivial consequence of the Lemma 2.

Corollary 1. *If* $(G, A) = 1$ *then the period of the* $\{Z_n\}_{n=1}^{\infty}$ *is greater or equal to* A.

This corollary shows that with suitably chosen pseudo-random sequence $\{A_n\}_{n=1}^{\infty}$ output sequence will have the period greater or equal to period of the $\{A_n\}_{n=1}^{\infty}$.

The next theorem, the main result of this paragraph, is a straightforward application of the former Lemmas.

Theorem 6. *Let* $l \in \{1, 2, ..., k!\}$ *denote the order of the permutation* $\prod_{i=1}^{C} f_{C_i}$. *Then if* $(lC, A) = 1$ *and* $\{A_1, A_2, \ldots, A_n, \ldots\} = I_k$, *the period of* $\{Z_n\}_{n=1}^{\infty}$ *is* lCA.

A statement of this theorem is a stronger variant of the Corollary 1 and it shows that with appropriately selected pseudo-random sequences $\{A_n\}_{n=1}^{\infty}$ and $\{B_n\}_{n=1}^{\infty}$ period of the generator output sequence is significantly longer then period sequences $\{A_n\}_{n=1}^{\infty}$ and $\{B_n\}_{n=1}^{\infty}$.

4. Conclusions

Confidentiality of different sensor networks is a very serious requirement since such networks cannot fully achieve their purpose without having the necessary security. Namely, for various IoT applications, which typically have limited processing capabilities, restricted memory capacity and power constraints, one of the key challenges is to design an efficient and reliable cryptographic generator that meets the desired security requirements. In this paper, we defined a pseudo-random generator that can, in some way, be considered as a generalization of ideas related to RC4 because it uses time varying permutations and sequences for permutation changing and addressing output element from the current permutation are considered in a general fashion. In the paper, we analyze properties of the purposed generator. The proposed pseudo-random generator can be implemented efficiently in software and hardware, for example by using output of the multiple linear shift registers as input sequences $\{A_n\}_{n=1}^{\infty}$, and $\{C_n\}_{n=1}^{\infty}$ of the generator. The security characteristics considered in the paper potentiate application of the generator in the computational constrained environments security solutions.

In the first part of the proposed generator properties analysis, the generator output sequence probability distribution is considered. Theorem 1 establishes sufficient conditions for the generator output sequence have an asymptotically uniform probability distribution. Moreover, sufficient conditions are established for the distribution of the output sequence to have the exact uniform distribution, Remark 1 and Theorem 2 . The generator output sequence uniform distribution indicates resistance of the generator to attacks based on output sequence elements prediction.

The second part of the generator analysis deala with the correlation properties of the generator output sequence in which it was shown, by Theorem 3 , that the output sequence elements are asymptotically independent and accordingly no immanent remote correlations are detected, unlike to the RC4 generator (see [25]). This property indicates resistance to autocorrelation type atacks.

The third part of the analysis relates to the possibility of information leaking about the internal state of the generator, sequence $\{A_n\}_{n=1}^{\infty}$ and $\{C_n\}_{n=1}^{\infty}$, through the output sequence $\{Z_n\}_{n=1}^{\infty}$. Theorems 4 and 5 show that the amount of information that flows through an output sequence tends to zero when the length of the sequence tends to infinity. In practical terms, this means that, by rejecting the initial segment of a generated sequence of a given length, the amount of information about the state of the generator flowing into the output sequence is arbitrarily small.

In the last part of the generator analysis, the period length of the output sequence is analyzed. It has been shown that if sequences $\{A_n\}_{n=1}^{\infty}$ and $\{C_n\}_{n=1}^{\infty}$ are chosen in a suitable manner and their periods satisfy the conditions of Theorem 6 , the output sequence has a significantly longer period than the sequences $\{A_n\}_{n=1}^{\infty}$ and $\{C_n\}_{n=1}^{\infty}$.

According to the performed analysis results proposed generator has provably good security characteristics.

Complexity and implementation considerations about this proposal are determined by the generation and complexity of the sequences $\{C_n\}_{n=1}^{\infty}$ and $\{A_n\}_{n=1}^{\infty}$. Relatively weak constraints demanded for the probability distribution for the sequences $\{C_n\}_{n=1}^{\infty}$ and $\{A_n\}_{n=1}^{\infty}$ in the Theorem 1 allow implementation of the efficiently generated sequences, for example sequences generated by the multiple linear feedback shift registers .

The method described in this paper makes it possible to obtain a pseudo-random sequence with asymptotically uniform distribution and longer period using two pseudo-random sequences with irregular (non-uniform) probability distributions. Required initial conditions for two pseudo-random sequences are not serious limitations for this method because they describe natural requirements for the pseudo-random sequences, i.e., that values of their elements exhaust the set on which they are defined. An interesting question arising in this context is the speed of convergence in the Theorem 1, i.e., the number of steps after which we can use the sequence $\{Z_n\}_{n=1}^{\infty}$ as uniformly distributed. It is not possible to answer this question generally because the matrix T is defined by the chosen set S and probability distribution of the random variable C. Consequently, for each set S and random variable C it has to be analyzed separately. In practice, this does not make any restrictions on the application of the proposed generator because in every concrete case it is possible to compute number of transition steps to achieve representation of the limit values by the desired accuracy.

Author Contributions: Conceptualization, T.U.; Formal analysis, T.U.; Supervision, Z.B. and M.M.; Writing—original draft, T.U.; Writing—review & editing, Z.B. and M.M.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rehman, R.A.; Khan, B. IoT Elements, Layered Architectures and Security Issues: A Comprehensive Survey. *Sensors* **2018**, *18*, 2796.
2. Salah, K. *The Era of Internet of Things*, 2nd ed.; Springer: Cham, Switzerland, 2019.
3. Rayes, A.; Samer, S. *Internet of Things from Hype to Reality*, 2nd ed.; Springer: Cham, Switzerland, 2019.
4. Atlam, H.F.; Walters, R.J.; Wills, G.B. Internet of Things: State-of-the-art, Challenges, Applications, and Open Issues. *IJICR* **2018**, *9*, 928–938.
5. Costa, D.G.; Figuerêdo, S.; Oliveira, G. Cryptography in Wireless Multimedia Sensor Networks: A Survey and Research Directions. *Cryptography* **2017**, *1*, 4. [CrossRef]
6. Kambourakis, G.; Marmol, F.G.; Wang, G. Security and Privacy in Wireless and Mobile Networks. *Future Internet* **2018**, *10*, 18. [CrossRef]
7. Ziegler, S. *Internet of Things Security and Data Protection*, 2nd ed.; Springer: Cham, Switzerland, 2019.

8. Cheruvu, S.; Kumar, A.; Smith, N.; Wheeler, D.M. *Demystifying Internet of Things Security: Successful IoT Device/Edge and Platform Security Deployment*; Apress: Berkeley, CA, USA, 2019.
9. Mahmood, Z. (Ed.) *Security, Privacy and Trust in the IoT Environment*; Springer: Cham, Switzerland, 2019.
10. Banday, M.T. *Cryptographic Security Solutions for the Internet of Things*; IGI Global: Hershey, PA, USA, 2019.
11. Biryukov, A.; Perrin, L. State of the Art in Lightweight Symmetric Cryptography. IACR Cryptology ePrint Archive, 2017. Available online: https://eprint.iacr.org/2017/511 (accessed on 28 October 2019).
12. Jing, Q.; Vasilakos, A.V.; Wan, J.; Lu, J.; Qiu, D. Security of the Internet of Things: Perspectives and challenges. *Wirel. Netw.* **2014**, *20*, 2481–2501. [CrossRef]
13. Frustaci, M.; Pace, P.; Aloi, G.; Fortino, G. Evaluating Critical Security Issues of the IoT World: Present and Future Challenges. *IEEE Internet Things J.* **2017**, *5*, 2483–2495. [CrossRef]
14. Xiao, Y.; Shen, X.; Sun, B.; Cai, L. Security and Privacy in RFID and Applications in Telemedicine. *IEEE Commun. Mag.* **2006** *44*, 64–72. [CrossRef]
15. Kumar, P.; Lee, H.J. Security Issues in Healthcare Applications Using Wireless Medical Sensor Networks: A Survey. *Sensors* **2012**, *12*, 55–91. [CrossRef] [PubMed]
16. Saraiva, D.A.F.; Leithardt, V.R.Q.; de Paula, D.; Sales Mendes, A.; González, G.V.; Crocker, P. PRISEC: Comparison of Symmetric Key Algorithms for IoT Devices. *Sensors* **2019**, *19*, 4312. [CrossRef] [PubMed]
17. Krishna, B.V.S.; Gnanasekaran, T. A systematic study of security issues in Internet-of-Things (IoT). In Proceedings of the International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 10–11 Februaty 2017; pp. 107–111. [CrossRef]
18. Carsten, M. Security and privacy in the internet of things. *J. Cyber Policy* **2017**, *2*, 155–184. [CrossRef]
19. Hamad, F.; Smalov, L.; James, A. Energy-aware Security in M-Commerce and the Internet of Things. *IETE Tech. Rev.* **2009**, *26*, 357–362. [CrossRef]
20. von zur Gathen, J. *CryptoSchool*; Springer: Berlin/Heidelberg, Germany, 2015.
21. Bilal, M.; Kang, S.G. An Authentication Protocol for Future Sensor Networks. *Sensors* **2017**, *17*, 979. [CrossRef] [PubMed]
22. Anderson, B. Thank You Bob Anderson. Cypherpunks (Mailing List). Available online: http://cypherpunks.venona.com/date/1994/09/msg00304.html (accessed on 28 October 2019).
23. Rrivest, R. *6.857 Computer and Network Security Lectures and Handouts*; MIT: Cambridge, MA, USA, 2008.
24. Rivest, R.; Schuldt, J. Spritz—A Spongy RC4-Like Stream Cipher and Hash Function. Available online: https://en.wikipedia.org/wiki/RC4#cite_note-Rivest2014-14 (accessed on 28 October 2019).
25. Sen Gupta, S.; Maitra, S.; Paul, G.; Sarkar, S. (Non-)Random Sequences from (Non-)Random Permutations—Analysis of RC4 Stream Cipher. Available online: https://eprint.iacr.org/2011/448.pdf (accessed on 28 October 2019).
26. Popov, A. *RFC 7465 Prohibiting RC4 Cipher Suites*; IETF: Fremont, CA, USA, 2015.
27. Stamp, M.; Low, R. *Applied Cryptanalysis : Breaking Ciphers in the Real World*; Wiley: Hoboken, NJ, USA, 2007.
28. Privoult, N. *Understanding Markov Chains*; Springer: Singapore, 2018.
29. Sericola, B. *Markov Chains—Theory, Algorithms and Applications*; Wiley: Hoboken, NJ, USA, 2013.