The Stream Ciphers and the Keystream Generator Based on Inverse Permutation

V. Scholtz [*] **J. Scholtzová** [**]

[*] *Department of Physics and Measurement, Faculty of Chemical Engineering, Institute of Chemical Technology in Prague, Czech Republic (e-mail: scholtzv@vscht.cz).*
[**] *Department of Mathematics, Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic (e-mail: scholtzo@math.feld.cvut.cz)*

Abstract: This paper describes a keystream generator of the stream ciphers based on inverse permutation and discuss its cryptographic criteria. The good cryptographic criteria as linear and spherical complexity are satisfied. These criteria are in this case of binary generator equivalent with other criteria. Consequently it was shown, that the generated binary sequence has good pseudorandom quality.

Keywords: Stream ciphers, keystream generator, cryptographic criteria, inverse permutation, empirical tests.

1. INTRODUCTION

Stream ciphers are suitable for the time-critical applications or processing-constrained devices to meet requirements of performance extremes. Some of its application is the audio/video encryption in communications, e.g. VoIP (Voice over IP), digital video broadcasting system like pay-TV or wireless communication protocols (WEP, WPA, SSL) and others, see Lu (2006). Good review and the state of the art can be find in Biryukov (2004).

In this work, the stream ciphers are introduced and the basic principles and requirements for its construction are described. Consequently, generator of keystream based on the inverse permutation is constructed and its cryptographic criteria are discussed.

2. MATHEMATICAL BACKGROUND

The private key ciphering systems can be generally classified into two main parts: block and stream ciphers. The principal distinction between block and stream ciphers is in memory, Cusick et al. (1998). The block cipher divides the message into blocks and enciphers each block by a key. Otherwise, the stream cipher specifies a device with internal memory that enciphers each digit of the message stream to other digit which depends on both the secret key and the internal state of the stream cipher at time. The sequence which controls the enciphering is called the key stream and the automaton to produce this key stream is called keystream generator. For more information or detailed study see the monography Cusick et al. (1998).

2.1 Additive synchronous stream ciphers

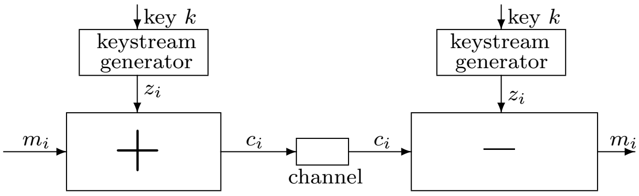

Fig. 1. Additive synchronous stream cipher.

The keystream in synchronous stream ciphers systems are independent of the message stream. The keystream is deterministic so that the stream can be reproduced for the decipherment. One type of the synchronous stream ciphers are the additive synchronous stream ciphers. The characters of the keystream comes from an Abelian group $(G, +)$ $\mathbb{Z}_N$. The ciphertext c_i is calculated by the addition operation "+" of the keystream character z_i and the message stream character m_i. On the receiver side, the original message stream is reconstructed by the inverse operation "−" of the ciphertext stream character c_i and the some keystream character z_i as on the sender side. This process is schematically depicted in fig. 1 and the adequate equation of the ciphering and deciphering are following:

$$\begin{aligned} c_i &= m_i + z_i, \\ m_i &= c_i - z_i. \end{aligned} \qquad (1)$$

2.2 Cryptographic aspects of sequences

The keystream sequences are required to have some properties to make it impossible (computational difficult) to find the key generator and decipher the ciphered message. Due to results published in works of Cusick et al. (1998) and Birkhoff et al. (1981), there are some common cryptographic measures of their strength such as the linear complexity, sphere complexity, nonlinearity and randomness.

Linear complexity. Linear homogeneous recurrence relations code of order m over $\mathbb{Z}_N$ is an equation of the form

$$\begin{aligned} &c_0 s_i + c_1 s_{i-1} + \ldots + c_m s_{i-m} = 0, \\ &i \geq m; \quad c_m \neq 0. \end{aligned} \tag{2}$$

This sequence produces an infinity sequence $s^\infty = (s_0, s_1, ...)$ for each initial block of $a = (s_0, s_1, ..., s_{m-1})$. If some block of $2m+1$ characters of this sequence is known, it is possible to calculate all the coefficients $c_0, \ldots, c_m$ and reconstruct the whole sequence s^∞. From the cryptographic point of view, it is desirable that the sequence s^∞ should be generated by an utmost order.

The polynomial $p(x) = c_0 + c_1 x + \cdots + c_{k-1} x^{k-1}$ is called the **characteristic polynomial** of the sequence s^∞.

The polynomial $m(x)$ is called a **minimal characteristic polynomial** of the sequence s^∞ if for any other characteristic polynomial $p(x)$ is satisfied

$$\deg(m(x)) \leq \deg(p(x)), \tag{3}$$

where deg is an order of the polynomial.

The **linear complexity** $L(s^\infty)$ is defined as the order of minimal polynomial of the sequence s^∞, such that

$$L(s^\infty) = \deg(m(x)). \tag{4}$$

In other words, the linear complexity represents a length of the minimal continuous sequence part necessary to calculate the whole sequence.

Sphere complexity. On principle, the potentional disturber does not need to find the key stream exactly. In the major part of situations some key stream sequence approximation may be sufficient.

The **sphere complexity** $SC_u(s)$ of sequence s represents the minimum of linear complexities of all sequences with the Hamming distance not greater that u:

$$SC_u(s) \;=\; min\{L(s+y),\; WH(y) \leq u\}, \tag{5}$$

where L is the linear complexity and $WH(y)$ is the Hamming weight of vector y.

Recall that the Hamming weight of vector y is the number of components of y that are different from zero and the Hamming distance of two sequences is the Hamming weight of their difference.

The secret keystream sequence must satisfy not only the good linear complexity but the sphere complexity also.

Nonlinearity. Let $g(x)$ be a function $g : (G, +) \rightarrow (H, +)$. The **local nonlinearity** of the function $g(x)$ for $a \neq 0$ is measured by

$$P_g(a) = \max_{b \in H} Pr(f(x+a) - f(x) = b) \tag{6}$$

and the **global nonlinearity** by

$$P_g = \max_{0 \neq a \in G} \max_{b \in H} Pr(f(x+a) - f(x) = b), \tag{7}$$

where $Pr(A)$ denotes the probability that event A occurs.

In other words, the nonlinearity represents the probability, that in the sequence occurs some linear part.

2.3 Empirical tests

Empirical tests are methods of statistical testing of some selected quality of sequences. The idea is to predicate some property of pseudorandom sequence $(s_0, s_1, \ldots)$ and construct a test to divide the characters into several categories $A = \{a_1, a_2, \ldots, a_k\}$. From the predicated attributes the theoretical probability of the categories are calculated and compared with the measured ones. For the comparison, the statistical χ^2 test with $k-1$ degrees of freedom may be used.

χ^2 test Let have a test with k categories. Each digit of the sequence may be assign into one of the categories. Value $v = k-1$ is the degree of freedom of the test. Let $(Y_1, Y_2, \ldots, Y_k)$ be the theoretical probabilities of each category, and $(P_1, P_2, \ldots, P_k)$ the measured probabilities. The number

$$V = \sum_{i=1}^{k} \frac{(Y_i - P_1)^2}{P_i} \tag{8}$$

represents a statistical value. This value V is matched to some probability P, which represents a probability that the each digit of the pseudorandom sequence acquire statistical value due to the theoretical predication. Some values of this χ^2 test are shown in table 1, for more details see the book Knuth (1969). Due to the same book, good pseudorandom generators should have this probability

$$P \in \langle 10\%, 90\% \rangle. \tag{9}$$

Table 1. Some values of the statistical number V for adequate probability P in χ^2 test with v degrees of freedom.

	$P = 96\%$	$P = 75\%$	$P = 50\%$	$P = 25\%$	$P = 5\%$
$v = 1$	0.00393	0.1015	0.4549	1.323	3.841
$v = 2$	0.1026	0.5753	1.386	2.773	5.991
$v = 3$	0.3518	1.213	2.366	4.108	7.815
$v = 4$	0.7107	1.923	3.357	5.385	9.488
$v = 5$	1.1455	2.675	4.351	6.626	11.07

3. BINARY SEQUENCES

The digits of binary sequences are of the $\mathbb{Z}_2$. In this section, some properties of binary sequences are introduced.

The **period** of the sequence $s^\infty = (s_0, s_1, \ldots)$ is minimal P, for which $s_{i+p} = s_i,\ i \in N$.

If p and $2p+1$ are primes, then p is **Sophie Germain prime**.

The **primitive element** (or generator) of the group is such an element which powers generates all the elements of this group.

Following rules are postulated e.g. in the book of Cusick et al. (1998):

Theorem 1. Let r be an odd prime, $N = r^k$ for $k \geq 1$ and the q be primitive element from $\mathbb{Z}_N$, then for each nonconstant sequence s^∞ with the period N over $\mathbb{Z}_q$,