Tatra Mt. Math. Publ. **29** (2004), 157–171

TATRA MOUNTAINS
Mathematical Publications

SOME MODIFICATIONS OF ONE-TIME PAD CIPHER

Oto Strauch

ABSTRACT. The main step in the presented modifications of one-time pad (also called Vernam cipher) is a key sequence x_n, $n = 1, 2, \dots$, computed from private vector sequence $\mathbf{b}_n$ and random sequence $\mathbf{x}_n$ by the relation

$$x_n = \begin{cases} 1 & \text{for } f(\mathbf{b}_n, \mathbf{x}_n) \in [A, B), \\ 0 & \text{for } f(\mathbf{b}_n, \mathbf{x}_n) \notin [A, B), \end{cases}$$

where f is a continuous function and the interval $[A, B)$ satisfies $g(B) - g(A) = \frac{1}{2}$, where g is an asymptotic distribution function of the sequence $f(\mathbf{b}_n, \mathbf{x}_n)$. We analyze the presented modification with respect to the disclosed message.

0. Introduction

This paper is devoted to security analysis of modified one-time pad cipher firstly described in the author's previous paper [S], which gives an application of uniformly distributed (shortly u.d.) sequences, for definitions and basic properties cf. [DT]. Here we use the following notations:

- $\mathbf{b}$ and $\mathbf{x}$ are s-dimensional vectors in a Jordan-measurable bounded subset K of $\mathbb{R}^s$ with positive Lebesgue measure $|K|$;
- $f(\mathbf{b}, \mathbf{x})$ is a continuous real function defined on $K \times K$, where its values form an interval $[A_0, B_0]$;
- $\mathbf{b}_n$ and $\mathbf{x}_n$, $n = 1, 2, \dots$, are u.d. and statistically independent sequences of s-dimensional vectors in K, i.e., $2s$-dimensional sequence $(\mathbf{b}_n, \mathbf{x}_n)$ is u.d. in $K \times K$;

2000 Mathematics Subject Classification: 11K31, 94A60.
Keywords: sequence, uniform distribution, statistically independent sequences, distribution function, scalar product, mixed product, one-time pad.

This research was supported by the Slovak Academy of Sciences Grant 2/1130/21 and by INNA, Ltd. Prague.

• Then the sequence $f(\mathbf{b}_n, \mathbf{x}_n)$, $n = 1, 2, \dots$, has the asymptotic distribution function (shortly a.d.f.) defined by

$$g(t) = \lim_{N\to\infty} \frac{\#\{n \le N : f(\mathbf{b}_n, \mathbf{x}_n) < t\}}{N}$$

for $t \in [A_0, B_0]$, and it can be computed by

$$g(t) = \frac{|\{(\mathbf{b}, \mathbf{x}) \in K^2 : f(\mathbf{b}, \mathbf{x}) < t\}|}{|K|^2}.$$

Starting with private $\mathbf{b}_n$ and random $\mathbf{x}_n$, the following key sequence x_n can be used in one-time pad:

$$x_n = \begin{cases} 1 & \text{if } f(\mathbf{b}_n, \mathbf{x}_n) \in [A, B), \\ 0 & \text{if } f(\mathbf{b}_n, \mathbf{x}_n) \notin [A, B), \end{cases}$$

where $[A, B)$ is an interval satisfying $g(B) - g(A) = \frac{1}{2}$.

• Similar, a.d.f. has the sequence $f\big(\mathbf{b}_n^{(1)}, \mathbf{b}_n^{(2)}, \dots, \mathbf{b}_n^{(s-1)}, \mathbf{x}_n\big)$, $n = 1, 2, \dots$, with statistically independent and u.d. sequences $\mathbf{b}_n^{(1)}, \mathbf{b}_n^{(2)}, \dots, \mathbf{b}_n^{(s-1)}, \mathbf{x}_n$ in K and it can be computed as

$$g(t) = \frac{|\{(\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(s-1)}, \mathbf{x}) \in K^s : f(\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(s-1)}, \mathbf{x}) < t\}|}{|K^s|}.$$

The u.d. and the statistical independence of $\mathbf{b}_n^{(1)}, \mathbf{b}_n^{(2)}, \dots, \mathbf{b}_n^{(s-1)}, \mathbf{x}_n$ in K is equivalent to u.d. of $(\mathbf{b}_n^{(1)}, \mathbf{b}_n^{(2)}, \dots, \mathbf{b}_n^{(s-1)}, \mathbf{x}_n)$ in K^s and the 0–1 sequence x_n defined by

$$x_n = \begin{cases} 1 & \text{for } f\big(\mathbf{b}_n^{(1)}, \dots, \mathbf{b}_n^{(s-1)}, \mathbf{x}_n\big) \in [A, B), \\ 0 & \text{for } f\big(\mathbf{b}_n^{(1)}, \dots, \mathbf{b}_n^{(s-1)}, \mathbf{x}_n\big) \notin [A, B), \end{cases}$$

where $g(B) - g(A) = \frac{1}{2}$, can be used as a key sequence, again.

• For two s-dimensional real vectors $\mathbf{b} = (b_1, \dots, b_s)$ and $\mathbf{x} = (x_1, \dots, x_s)$ the scalar (inner) product is

$$\mathbf{b} \cdot \mathbf{x} = b_1 x_1 + \dots + b_s x_s .$$

The scalar product of $\mathbf{b}$ and a set I of s-dimensional vectors is defined as usual $\mathbf{b} \cdot I = \{\mathbf{b} \cdot \mathbf{x} : \mathbf{x} \in I\}$. Similarly it can be defined $f(\mathbf{b}, I)$ and $I_1 + I_2 = \{\mathbf{x}_1 + \mathbf{x}_2 : \mathbf{x}_1 \in I_1, \mathbf{x}_2 \in I_2\}$.

Using the above key sequences, in the previous paper [S] we have described three modifications of the classical one-time pad. We shall repeat this in Parts 1–3 (similar modifications we did not find in books [K], [L] and [MvV]). The main Part 4 of this paper is devoted to the study of security of such ciphers with respect to disclosed messages. In the classical one-time pad cipher (i.e., Vernam cipher, cf. [MvV, Chap. 6]), two users X and Y have a common private 0–1 key sequence x_n, $n = 1, 2, \ldots, N$, and X sends the 0–1 plaintext sequence u_n, $n = 1, 2, \ldots, N$, to Y as the ciphertext sequence $y_n = u_n + x_n \pmod 2$. Since the y_n are sent through public line, if u_n is disclosed, then the key sequence x_n is also disclosed. Thus, for security reasons, x_n should be applied only once. In Part 4 we prove that in our modifications the private sequence $\mathbf{b}_n$ can be used many times.

1. First modification

(I) Two users X and Y agree on

- a private vector sequence
$$\mathbf{b}_n, \qquad n = 1, 2, \ldots, N,$$
in $K \subset \mathbb{R}^s$, and
- a private sequence of regular $(s+1) \times (s+1)$ real matrices
$$\mathbf{A}_n, \qquad n = 1, 2, \ldots, N,$$
where N is sufficiently large, and
- a continuous function
$$f(\mathbf{b}, \mathbf{x}) \qquad \text{for } \mathbf{b}, \mathbf{x} \in K,$$
where the values $f(\mathbf{b}, \mathbf{x})$ form an interval $[A_0, B_0]$. The users compute the d.f.
$$g(t) = \frac{|\{(\mathbf{b}, \mathbf{x}) \in K^2; (\mathbf{b}, \mathbf{x}) < t\}|}{|K|^2} \qquad \text{for } t \in [A_0, B_0]$$
and select an interval $[A, B) \subset [A_0, B_0]$ such that
$$g(B) - g(A) = \frac{1}{2}.$$
The interval $[A, B)$ is also private.

(II) If X wants to send a message written as a 0–1 sequence (i.e., the plaintext digits)
$$u_n, \qquad n = 1, 2, \ldots, N,$$
to Y, he makes the following steps:

- X selects a random (or pseudo-random) sequence of vectors $\mathbf{x}_n$, $n = 1, 2, \ldots, N$, in the region K.

- X computes enciphering sequence (the key digits)

$$x_n = \begin{cases} 1 & \text{if } f(\mathbf{b}_n, \mathbf{x}_n) \in [A, B)\,, \\ 0 & \text{if } f(\mathbf{b}_n, \mathbf{x}_n) \notin [A, B)\,. \end{cases}$$

- X enciphers u_n to the sequence (ciphertext digits)

$$y_n = u_n + x_n \pmod 2\,, \qquad n = 1, 2, \dots, N\,.$$

- X mixes the sequence $(\mathbf{x}_n, y_n)$ of $s+1$-dimensional vectors using matrices $\mathbf{A}_n$ to

$$\mathbf{z}_n^T = \mathbf{A}_n \cdot (\mathbf{x}_n, y_n)^T.$$

and then sends $\mathbf{z}_n$ to Y through a public line.

(III) Y decrypts the received ciphertext as follows:

- Y decomposes $\mathbf{z}_n$ applying formula

$$(\mathbf{x}_n, y_n) = \mathbf{A}_n^{-1} \cdot \mathbf{z}_n^T\,.$$

- Using $\mathbf{x}_n$ he computes x_n in the same manner as X.
- He finds the plaintext digits

$$u_n = y_n + x_n \pmod 2\,, \qquad n = 1, 2, \dots, N\,.$$

In [S] we have proposed to take for f

- $f(\mathbf{b}, \mathbf{x}) = \mathbf{b} \cdot \mathbf{x}$—the scalar (inner) product and $K = [0,1]^s$, and we found the d.f.

$$g_s(t) = \left|\left\{(\mathbf{b}, \mathbf{x}) \in [0,1]^{2s} : \mathbf{b} \cdot \mathbf{x} < t\right\}\right|, \qquad t \in [0, s]\,,$$

explicitly only for $s = 1, 2, 3$ and $t \in [0, 1]$:

$$g_1(t) = t - t \log t\,;$$

$$g_2(t) = \tfrac{t^2}{2}\big((\log t)^2 - 3 \log t + \tfrac{7}{2} - \tfrac{1}{6}\pi^2\big)\,;$$

$$g_3(t) = \tfrac{t^3}{27}\Big(-\tfrac{9}{2}(\log t)^3 + \tfrac{99}{4}(\log t)^2 + \big(-\tfrac{255}{4} + \tfrac{9}{4}\pi^2\big)\log t + \tfrac{575}{8} - \tfrac{33}{8}\pi^2 - 9\zeta(3)\Big)\,,$$

where $\zeta(s)$ is the classical Riemann's zeta function. For $g_3(t)$ we have computed an interval $[0, 0.710837116\ldots)$, where $g_3(0.710837116\ldots) = 0.5$.

2. Second modification

The cryptosystem in Part 1 can be modified taking a function $f\big(\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(s-1)}, \mathbf{x}\big)$ and u.d. vectors sequence $\big(\mathbf{b}_n^{(1)}, \dots, \mathbf{b}_n^{(s-1)}, \mathbf{x}_n\big)$ in K^s, where

- $(\mathbf{b}_n^{(1)}, \dots, \mathbf{b}_n^{(s-1)})$ is a common private sequence for users X and Y;

- $\mathbf{x}_n$ is a sequence in K chosen randomly by X;
- The enciphering sequence is computed by

$$x_n = \begin{cases} 1 & \text{if } f\big(\mathbf{b}_n^{(1)}, \dots, \mathbf{b}_n^{(s-1)}, \mathbf{x}_n\big) \in [A, B)\,, \\ 0 & \text{if } f\big(\mathbf{b}_n^{(1)}, \dots, \mathbf{b}_n^{(s-1)}, \mathbf{x}_n\big) \notin [A, B)\,, \end{cases}$$

where $g(B) - g(A) = \frac{1}{2}$ and

$$g(t) = \frac{\left|\{(\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(s-1)}, \mathbf{x}) \in K^s \colon f(\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(s-1)}, \mathbf{x}) < t\}\right|}{|K^s|}\,.$$

In [S] we proposed to use the determinant function

- $f\big(\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(s-1)}, \mathbf{x}\big) = \big|\det\big(\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(s-1)}, \mathbf{x}\big)\big|$,

i.e., the absolute value of the mixed product, where the vectors are lying in the s-dimensional ball $K = B(r)$ in $\mathbb{R}^s$ with center $(0, \dots, 0)$ and radius r. The a.d.f.

$$g_s(r, t) = \frac{\left|\{(\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(s-1)}, \mathbf{x}) \in B^s(r) \colon |\det(\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(s-1)}, \mathbf{x})| < t\}\right|}{|B(r)|^s}$$

for all $t \in [0, r^s]$ can be written as $g_s(r, t) = \tilde{g}_s(\lambda)$, $\lambda \in [0, 1]$, where $\lambda = \frac{t}{r^s}$. In [S] we find the following explicit forms only for $s = 1, 2, 3$:

$\tilde{g}_1(\lambda) = \lambda$;

$\tilde{g}_2(\lambda) = \frac{2}{\pi}(1 + 2\lambda^2) \arcsin \lambda + \frac{6}{\pi}\lambda\sqrt{1 - \lambda^2} - 2\lambda^2$;

$\tilde{g}_3(\lambda) = 1 + \frac{9}{4}\lambda \int\limits_{\lambda}^{1} \frac{\arccos x}{x}\,dx - \frac{3}{4}\lambda^3 \arccos \lambda - \sqrt{1 - \lambda^2} + \frac{7}{4}\lambda^2\sqrt{1 - \lambda^2}$,

where $\lambda \in [0, 1]$.

3. Third modification

The further modification of the one-time pad cryptosystem is as follows:

- In a given s-dimensional region K the users X and Y have a common private sequence

$$\big(\mathbf{b}_n^{(1)}, \dots, \mathbf{b}_n^{(s-2)}\big)\,, \qquad n = 1, 2, \dots$$

Prior to the encryption of a message, users X and Y exchange through a public line two sequences $\mathbf{u}_n$ and $\mathbf{w}_n$ computed as follows:

- X selects $\mathbf{x}_n = (x_{n,1}, \dots, x_{n,s})$, $n = 1, 2, \dots, N$, randomly;
- Y selects $\mathbf{y}_n = (y_{n,1}, \dots, y_{n,s})$, $n = 1, 2, \dots, N$, randomly;

- X sends to Y through a public line the sequence of vectors

$$\mathbf{u}_n = \begin{vmatrix} \mathbf{i}_1 & \dots & \mathbf{i}_s \\ b_{n,1}^{(1)} & \dots & b_{n,s}^{(1)} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ b_{n,1}^{(s-2)} & \dots & b_{n,s}^{(s-2)} \\ x_{n,1} & \dots & x_{n,s} \end{vmatrix},$$

  where $\mathbf{i}_1, \mathbf{i}_2, \dots, \mathbf{i}_s$ are s-dimensional unit vectors and $\mathbf{b}_n^{(i)} = \left(b_{n,1}^{(i)}, b_{n,2}^{(i)}, \dots, b_{n,s}^{(i)}\right)$;
- Y sends to X through a public line the sequence

$$\mathbf{w}_n = \begin{vmatrix} \mathbf{i}_1 & \dots & \mathbf{i}_s \\ b_{n,1}^{(1)} & \dots & b_{n,s}^{(1)} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ b_{n,1}^{(s-2)} & \dots & b_{n,s}^{(s-2)} \\ y_{n,1} & \dots & y_{n,s} \end{vmatrix};$$

- X and Y compute the common sequence of absolute values of scalar products

$$v_n = |\mathbf{u}_n \cdot \mathbf{y}_n| = |\mathbf{w}_n \cdot \mathbf{x}_n|;$$

- X and Y compute the key sequence x_n, $n = 1, 2, \dots, N$, by

$$x_n = \begin{cases} 1 & \text{if } v_n \in [A, B), \\ 0 & \text{if } v_n \notin [A, B), \end{cases}$$

  where $g(B) - g(A) = \frac{1}{2}$ and $g(t)$ is a d.f. of u.d. sequence

$$\left(\mathbf{x}_n, \mathbf{b}_n^{(1)}, \dots, \mathbf{b}_n^{(s-2)}, \mathbf{y}_n\right)$$

  in K. If $K = B(r)$, we can take $g(t) = g_s(r, t)$ as in Part 2.
- Note that users X and Y can test $\mathbf{u}_n$ and $\mathbf{w}_n$ by

$$\mathbf{u}_n \cdot \mathbf{b}_n^{(i)} = \mathbf{w}_n \cdot \mathbf{b}_n^{(i)} = 0,$$

  $i = 1, 2, \dots, s-2$.

In Part 5 we prove that this Modification 3 is equivalent to the Modification 1, where f is given by the absolute value of the scalar product of two s^2-dimensional vectors and without using the scrambling matrices.

4. Security analysis of Modification 1

The security problem can be divided into three parts:

(i) Security of $\mathbf{b}_n$,

(ii) security of key sequence x_n, and

(iii) quality of x_n.

We shall be interested in the security of Modification 1 with respect to disclosed messages. For simplification we consider Modification 1 without the mixed part $\mathbf{A}_n \cdot (\mathbf{x}_n, y_n)^T$, i.e., we shall assume that the user X sends

- the random vectors $\mathbf{x}_n$, $n = 1, 2, \ldots, N$, by a public line, and
- the ciphertext y_n, $n = 1, 2, \ldots, N$, by a public line.

Moreover, we assume that

- the message u_n, $n = 1, 2, \ldots, N$, is disclosed, and
- the interval $[A, B)$ is not secrete.

Since the ciphertext y_n through a public line the above assumptions implies

- the enciphering sequence x_n, $n = 1, 2, \ldots, N$, is disclosed,

and this implies that, for secrete $\mathbf{b}_n$, $n = 1, 2, \ldots, N$,

$$\mathbf{b}_n \in f^{-1}\big([A, B), \mathbf{x}_n\big) \quad \text{or} \quad \mathbf{b}_n \notin f^{-1}\big([A, B), \mathbf{x}\big),$$

is known, where $f^{-1}(I, \mathbf{x}) = \big\{\mathbf{b} \in K \colon f(\mathbf{b}, \mathbf{x}) \in I\big\}$.

Assume that the disclosed messages are repeated k times and the involved sequences are:

$\mathbf{x}_n^{(i)}$, $n = 1, 2, \ldots, N_i$, is the random sequence for the ith application,

$x_n^{(i)}$, $n = 1, 2, \ldots, N_i$, is the enciphering sequence in the ith application,

and

$v_n^{(i)} = f\big(\mathbf{b}_n, \mathbf{x}_n^{(i)}\big)$, $n = 1, 2, \ldots, N_i$, is the auxiliary sequence for computing enciphering $x_n^{(i)}$ by

$x_n^{(i)} = 1 \Longleftrightarrow v_n^{(i)} \in [A, B)$, where $i = 1, 2, \ldots, k$.

For a further application[1] of Modification 1 we assume that a message u_n, $n = 1, 2, \ldots, N$, is secrete and the related sequences are

$\mathbf{x}_n$, $n = 1, 2, \ldots, N$,

x_n, $n = 1, 2, \ldots, N$,

$v_n = f(\mathbf{b}_n, \mathbf{x}_n)$, $n = 1, 2, \ldots, N$.

[1] $(k+1)$–application

(i) Security of $\mathbf{b}_n$

Fixing n and assuming that $k = k_1 + k_2$, where $x_n^{(i_l)} = 1$ for $l = 1, 2, \dots, k_1$ and $x_n^{(i_l)} = 0$ for $l = k_1 + 1, k_1 + 2, \dots, k_2$, we have

$$\mathbf{b}_n \in J_n = \left(\bigcap_{l=1}^{k_1} f^{-1}\big([A, B), \mathbf{x}_n^{(i_l)}\big) \right) \cap \left(\bigcap_{l=k_1+1}^{k_2} f^{-1}\big([A_0, B_0] - [A, B), \mathbf{x}_n^{(i_l)}\big) \right).$$

This $J_n \subset [0,1]^s$ provides only some possibilities to approximate of $\mathbf{b}_n$. For the auxiliary sequence $v_n = f(\mathbf{b}_n, \mathbf{x}_n)$ we have

$$v_n \in Y_n = f(J_n, \mathbf{x}_n).$$

Here the set J_n and $Y_n \subset [A_0, B_0]$ are publicly known[2] (see also security of x_n, part (b)). E.g., if $Y_n \subset [A, B]$, then $x_n = 1$.

(**a**) For example, if $\mathbf{b}, \mathbf{x} \in K = [0,1]^s$, $f(\mathbf{b}, \mathbf{x}) = \mathbf{b} \cdot \mathbf{x}$, $[A, B) = [0, B)$, then the set J_n is convex. This follows from the fact that the vector $\mathbf{b}_n$ lies under ($x_n^{(i)} = 1$) or over ($x_n^{(i)} = 0$) the hyperplane

$$\mathbf{y} \cdot \frac{\mathbf{x}_n^{(i)}}{\left|\mathbf{x}_n^{(i)}\right|} = \frac{B}{\left|\mathbf{x}_n^{(i)}\right|},$$

for every $i = 1, 2, \dots, k$, where $\mathbf{y} = (y_1, \dots, y_s)$ is a variable vector. Thus Y_n is a subinterval of $[0, s)$.

(**b**) Furthermore, assume that $k = k_1$ (i.e., $x_n^{(i)} = 1$ for $i = 1, 2, \dots, k$). Since every set $f^{-1}\big([0, B), \mathbf{x}_n^{(i)}\big)$ contains the intersection $B(r) \cap [0,1]^s$, where $r = 1/\sqrt{s}$, we have

$$B(r) \cap [0,1]^s \subset J_n.$$

(**c**) If for given n and $i = 1, 2, \dots, k$, the Euclidean norms satisfy $\left|\mathbf{x}_n^{(i)}\right| < B/\sqrt{s}$, then we have $J_n = [0,1]^s$, but in this case (see below (a)), $x_n = 1$ automatically.

(ii) Security of x_n

Some information about key sequence x_n can be found without any information about secrete $\mathbf{b}_n$. Here we again assume that $f(\mathbf{b}, \mathbf{x}) = \mathbf{b} \cdot \mathbf{x} \in [0, s]$ for $\mathbf{b}, \mathbf{x} \in K = [0,1]^s$.

(**a**) If $[A, B) = [0, B)$, then we have

$$|\mathbf{x}_n| < \frac{B}{\sqrt{s}} \implies \mathbf{b}_n \cdot \mathbf{x}_n < \sqrt{s}\,\frac{B}{\sqrt{s}} = B$$

[2] i.e., J_n and Y_n can be computed from the publicly known f, $[A, B)$, $\mathbf{x}_n^{(i)}$, $i = 1, 2, \dots, k$, and $\mathbf{x}_n$.

and thus $x_n = 1$ for such $\mathbf{x}_n$, automatically. Since $\mathbf{x}_n$, $n = 1, 2, \dots$, is u.d. in $K = [0,1]^s$, then the relative density of the publicly known part of the key sequence x_n, $n = 1, 2, \dots$, is $\left|B(r) \cap [0,1]^s\right|$ for $r = \frac{B}{\sqrt{s}}$ (i.e., $< \left(\frac{B}{\sqrt{s}}\right)^s$).

(a1) To eliminate (a), we must use only interval $[A, B)$ with $A > 0$ or more generally, m non-overlapping intervals $[A_i, B_i)$, $i = 1, 2, \dots, m$, satisfying

$$\sum_{i=1}^{m} g(B_i) - g(A_i) = \frac{1}{2}\,.$$

(a2) The (a) can also be eliminated by using a big s. For the smallest s we need to apply the mixed part of Modification 1.

(**b**) Assume that $k = s$ and that the random vectors $\mathbf{x}_n^{(1)}, \dots, \mathbf{x}_n^{(s)}$ are linearly independent over $\mathbb{R}$, for given fixed n. Then, for the random vector $\mathbf{x}_n$ in the next application of Modification 1, can be expressed as a liner combination

$$\mathbf{x}_n = t_{n,1}\mathbf{x}_n^{(1)} + \dots + t_{n,s}\mathbf{x}_n^{(s)}\,,$$

which implies that

$$v_n = t_{n,1}v_n^{(1)} + \dots + t_{n,s}v_n^{(s)}\,.$$

Let

$$\begin{aligned}
&\mathbf{t}_n = (t_{n,1}, \dots, t_{n,s}) \in K = [0,1]^s\,,\\
&\mathbf{v}_n = \left(v_n^{(1)}, \dots, v_n^{(s)}\right),\\
&I_n^{(i)} = [A, B) \text{ if } x_n^{(i)} = 1 \quad \text{or} \quad I_n^{(i)} = [0, s] - [A, B) \text{ if } x_n^{(i)} = 0\,,\\
&I_n = I_n^{(1)} \times \dots \times I_n^{(s)}\,,\\
&X_n = \mathbf{t}_n \cdot I_n\,.
\end{aligned}$$

Then $\mathbf{v}_n \in I_n$, where $\mathbf{v}_n$ is private, I_n is publicly known, and $\mathbf{t}_n \cdot \mathbf{v}_n = v_n \in X_n$, where X_n can be computed, i.e., X_n is a publicly known set from $[0, s]$. Let p_n be the probability that $x_n = 1$ (the values n, t_n are fixed and $\mathbf{v}_n$ is u.d. in I_n). Assume that $[A, B) = [0, B)$. Then I_n is a box of the volume $|I_n| = B^{k_1}(s-B)^{k_2}$ ($k_1 + k_2 = s$) and X_n is an interval in $[0, s]$ and let $X_n = [A_n, B_n]$. Then $X_n \cap [0, B) = \left[\min(B, A_n), \min(B, B_n)\right)$ and we have

$$p_n = g\left(\min(B, B_n)\right) - g\left(\min(B, A_n)\right),$$

where g is a related d.f. Let $P_N = \prod_{n=1}^{N} \max(p_n, 1-p_n)$. For a satisfactory attack[3] to the private key sequence x_n, $n = 1, 2, \dots, N$, determined by Modification 1, we need more than $M = [P_N^{-1}]$ random 0–1 sequences x'_n, $n = 1, 2, \dots, N$ (realizing $x'_n = 1$ with probability p_n). Moreover, eliminating those n for which $|\mathbf{x}_n| < B/\sqrt{s}$, we can decrease N to $N\left(1 - (B/\sqrt{s})^s\right)$. Since in this possible attack we have $k \geq s$, we can write:

[3] We approximate the plaintext digits u_n by $u'_n = y_n + x'_n \pmod 2$ and test its meaning.

Claim. *For sufficiently large dimension s, the simplified Modification 1 (with $\mathbf{b}, \mathbf{x} \in [0,1]^s$, $f(\mathbf{b}, \mathbf{x}) = \mathbf{b} \cdot \mathbf{x}$, and without the mixed part) can be securely applied at least s-times.*

(iii) Quality of x_n

It is known that the quality of the enciphering sequence x_n is given by the degree of the u.d. of x_n, i.e., by the maximal number j for which the j-dimensional sequence $(x_n, \dots, x_{n+j-1})$, $n = 1, 2, \dots$ is u.d. Precisely, since we use only finite 0–1 sequence x_n, $n = 1, 2, \dots, N$, a measure of quality of x_n is the discrepancy

$$D_N\big((x_n, \dots, x_{n+j-1})\big) = \max_{\mathbf{e} \in \{0,1\}^j} \left| \frac{\#\{n \le N : (x_n, \dots, x_{n+j-1}) = \mathbf{e}\}}{N} - \frac{1}{2^j} \right| .$$

Assuming that the finite sequence $(\mathbf{b}_n, \mathbf{x}_n)$, $n = 1, 2, \dots, N$, in K^2 is an initial segment of an infinite sequence $(\mathbf{b}_n, \mathbf{x}_n)$, $n = 1, 2, \dots$, with is u.d. in K^2, with a.d.f. $g(t)$ of the sequence $v_n = f(\mathbf{b}_n, \mathbf{x}_n)$ given by

$$g(t) = \lim_{N \to \infty} \frac{\#\{n \le N : f(\mathbf{b}_n, \mathbf{x}_n) < t\}}{N} ,$$

then we see that the key 0–1 sequence x_n, $n = 1, 2, \dots, N$, is an initial segment of the u.d. sequence x_n, $n = 1, 2, \dots$ with $x_n = 1 \iff v_n \in [A, B)$, where $g(B) - g(A) = 1/2$. Similarly, the u.d. of the $2sj$-dimensional sequence

$$(\mathbf{b}_n, \mathbf{b}_{n+1}, \dots, \mathbf{b}_{n+j-1}, \mathbf{x}_n, \mathbf{x}_{n+1}, \dots, \mathbf{x}_{n+j-1}), \qquad n = 1, 2, \dots ,$$

implies the u.d. of j-dimensional sequence

$$(x_n, x_{n+1}, \dots, x_{n+j-1}), \qquad n = 1, 2, \dots$$

Thus the quality of input $(\mathbf{b}_n, \mathbf{x}_n)$ has an impact on the quality of x_n. Precisely, recapitulate and extend the notations in the Method 1

v_n is a sequence in $[A_0, B_0]$ with a.d.f. $g(t)$;
$[A_0, B) \subset [A_0, B_0]$ such that $g(B) - g(A_0) = \frac{1}{2}$;
$x_n = 1 \iff v_n \in [A_0, B)$;
$f : [0,1]^{2s} \to [A_0, B_0]$ is a continuous function having convex $f^{-1}([A_0, B])$ and $f^{-1}([B, B_0])$;
$\mathbf{z}_n = (\mathbf{b}_n, \mathbf{x}_n)$ is an u.d. sequence in $[0,1]^{2s}$;
$v_n = f(\mathbf{z}_n)$ for $n = 1, 2, \dots$;
$I = I_1 \times \dots \times I_j$, where either $I_i = [A_0, B)$ or $I_i = [B, B_0]$, $i = 1, 2, \dots, j$;
$\mu(I)$ is the measure I with respect to d.f. $g(t_1, \dots, t_j) = g(t_1) \dots g(t_j)$ (clearly $\mu(I) = 1/2^j$);

$U = f^{-1}(I_1) \times \cdots \times f^{-1}(I_j)$;
$|U|$ is the Lebesgue measure of U;
$\Delta_N((v_n, \dots, v_{n+j-1})) = \sup_I \left| \frac{\#\{n \le N:\ (v_n,\dots,v_{n+j-1}) \in I\}}{N} - \mu(I) \right|$;
$\Delta'_N((\mathbf{z}_n, \dots, \mathbf{z}_{n+j-1})) = \sup_U \left| \frac{\#\{n \le N:\ (\mathbf{z}_n,\dots,\mathbf{z}_{n+j-1}) \in U\}}{N} - |U| \right|$.

For finite $2sj$-dimensional sequence $(\mathbf{z}_n, \dots, \mathbf{z}_{n+j-1})$, $n = 1, 2 \dots, N$, defines two classical discrepancies:

(1) *The extremal discrepancy* [DT, p. 4]

$$D_N((\mathbf{z}_n, \dots, \mathbf{z}_{n+j-1})) = \sup_X \left| \frac{\#\{n \le N:\ (\mathbf{z}_n,\dots,\mathbf{z}_{n+j-1}) \in X\}}{N} - |X| \right| ,$$

where the supremum is taken over all boxes $X \subset [0,1]^{2sj}$, and

(2) *the isotropic discrepancy* [DT, p. 9] $J_N((\mathbf{z}_n, \dots, \mathbf{z}_{n+j-1}))$, where the supremum is taken over all convex subsets $X \subset [0,1]^{2sj}$.

The definitions imply

$$\begin{aligned} D_N((x_n, \dots, x_{n+j-1})) &= \Delta_N((v_n, \dots, v_{n+j-1})) \\ &= \Delta'_N((\mathbf{z}_n, \dots, \mathbf{z}_{n+j-1})) \le J_N((\mathbf{z}_n, \dots, \mathbf{z}_{n+j-1})) . \end{aligned}$$

By [DT, p. 9, Th.1.12]

$$J_N((\mathbf{z}_n, \dots, \mathbf{z}_{n+j-1})) \le (8sj\sqrt{2sj} + 1)\left(D_N((\mathbf{z}_n, \dots, \mathbf{z}_{n+j-1}))\right)^{1/(2sj)}$$

and the extremal discrepancy D_N can be estimated by Erdős–Turán–Koksma inequality [DT, p. 23, Th. 1.28]. It is conjectured [DT, p. 39] that for any finite sequence in $[0,1]^r$ we have $D_N \ge c_r (\log N)^{r-1}/N$. Thus, for a large s, the extremal discrepancy $D_N((\mathbf{z}_n, \dots, \mathbf{z}_{n+j-1}))$ cannot give a good estimate of the discrepancy $D_N((x_n, \dots, x_{n+s-1}))$ of the enciphering x_n, because $r = 2sj$.

But, on the other hand, the resulting 0–1 sequence x_n can have a good discrepancy, since for any given 0–1 u.d. sequence x_n there exists u.d. sequence $\mathbf{z}_n$ in $[0,1]^{2s}$ such that x_n satisfies $x_n = 1 \iff f(\mathbf{z}_n) \in [A_0, B)$ for all n. To see this, use the following construction: Let $\mathbf{u}_n$ be u.d. in $f^{-1}([A_0, B])$ and $\mathbf{w}_n$ be u.d. in $f^{-1}([B, B_0])$, then the sequence

$$\mathbf{z}_n = x_n \mathbf{u}_n + (1 - x_n)\mathbf{w}_n$$

is u.d. in $[0,1]^{2s}$.

5. Transformation of Modification 3 to 1

Assume that we have used Modification 3 s-times, and let n what follows be fixed. Denote the vectors $\mathbf{x}_n$, $\mathbf{y}_n$, $\mathbf{u}_n$ and $\mathbf{w}_n$ in the ith application of

Modification 3 ($1 \le i \le s$), $\mathbf{x}_n^{(i)}$, $\mathbf{y}_n^{(i)}$, $\mathbf{u}_n^{(i)}$ and $\mathbf{w}_n^{(i)}$, respectively. We can assume that $\mathbf{x}_n^{(1)}, \dots, \mathbf{x}_n^{(s)}$ and $\mathbf{y}_n^{(1)}, \dots, \mathbf{y}_n^{(s)}$ are linearly independent in $\mathbb{R}^s$. Thus the vectors $\mathbf{x}_n$ and $\mathbf{y}_n$ in a further application of Modification 3 can be expressed as

$$\mathbf{x}_n = t_1 \mathbf{x}_n^{(1)} + \dots + t_s \mathbf{x}_n^{(s)}, \quad \mathbf{y}_n = t'_1 \mathbf{y}_n^{(1)} + \dots + t'_s \mathbf{y}_n^{(s)}.$$

The same holds for the computed $\mathbf{u}_n$ and $\mathbf{w}_n$

$$\mathbf{u}_n = t_1 \mathbf{u}_n^{(1)} + \dots + t_s \mathbf{u}_n^{(s)}, \quad \mathbf{w}_n = t'_1 \mathbf{w}_n^{(1)} + \dots + t'_s \mathbf{w}_n^{(s)},$$

and the sequence $v_n = |\mathbf{u}_n \cdot \mathbf{y}_n| = |\mathbf{w}_n \cdot \mathbf{x}_n|$ can be computed as

$$v_n = \left| \sum_{i=1}^{s} t_i \mathbf{u}_n^{(i)} \cdot \sum_{i=1}^{s} t'_i \mathbf{y}_n^{(i)} \right| = \left| \sum_{i,j=1}^{s} t_i t'_j \mathbf{u}_n^{(i)} \cdot \mathbf{y}_n^{(j)} \right|.$$

Denote the s^2-dimensional vectors (in the matrix form) by

$$\begin{aligned} \mathbf{X}_n &= \left(t_i t'_j\right), \qquad i,j = 1,2,\dots,s, \\ \mathbf{B}_n &= \left(\mathbf{u}_n^{(i)} \cdot \mathbf{y}_n^{(j)}\right), \qquad i,j = 1,2,\dots,s. \end{aligned}$$

Because the vectors $\mathbf{u}_n$ and $\mathbf{w}_n$ are transmitted through a public line, the vectors of coefficients $(t_1, \dots, t_s)$ and $(t'_1, \dots, t'_s)$ are not private and thus $\mathbf{X}_n$ is public. Contrary to this, $\mathbf{B}_n$ is private, because $\mathbf{y}_n^{(1)}, \dots, \mathbf{y}_n^{(s)}$ are private. Thus the enciphering sequence x_n in the Modification 3 can be computed as in Modification 1

$$f(\mathbf{X}_n, \mathbf{B}_n) = |\mathbf{X}_n \cdot \mathbf{B}_n| = v_n, \; x_n = 1 \Longleftrightarrow v_n \in [A,B), \; g(B) - g(A) = 1/2.$$

Here the d.f. $g(t)$ is the same as in Modification 3, but the region formed by $\mathbf{X}_n$ and $\mathbf{B}_n$ in $\mathbb{R}^{2s^2}$ is more complicated.

6. Pseudorandom $\mathbf{x}_n$ and $\mathbf{b}_n$

If the sequence $\mathbf{x}_n$ is randomly generated in K and $\mathbf{b}_n$ is u.d. in K, then they are statistically independent. If we use a pseudo-random generation of $\mathbf{x}_n$ we must have $\mathbf{z}_n = (\mathbf{b}_n, \mathbf{x}_n)$ u.d. in K^2 (or in Modification 2–3 u.d. of $\mathbf{z}_n = \left(\mathbf{b}_n^{(1)}, \mathbf{b}_n^{(2)}, \dots, \mathbf{x}_n\right)$). The following candidates for such sequences $\mathbf{z}_n$ are to be taken into account:

Kronecker sequences

The k-dimensional Kronecker sequence

$$\mathbf{z}_n = (n\theta_1, \dots, n\theta_k) \mod 1, \qquad n = 1,2,\dots,$$

is u.d. in $[0,1]^k$ if and only if $1,\theta_1,\dots,\theta_k$ are linearly independent over rationals $\mathbb{Q}$. For different primes $p_1,\dots,p_k$, the square-roots $1,\sqrt{p_1},\dots,\sqrt{p_j}$ are linearly independent and thus

$$\mathbf{z}_n = \left(n\sqrt{p_1},\dots,n\sqrt{p_k}\right) \mod 1\,, \qquad n=1,2,\dots,$$

is u.d. in $[0,1]^k$. This Kronecker type sequence was first used by R. D. Richtmyer (1951) [R].

Halton sequences

Let $n=\sum_{i=0}^{k(n)} a_i(n)q^i$, $a_i\in\{0,1,\dots,q-1\}$, be the expansion of integer n in the base $q\ge 2$ (q an integer), and define (so-called Van der Corput radical inverse function) $\gamma_q(n)=\sum_{i=0}^{k(n)} a_i(n)q^{-i-1}$. For the relative prime bases $q_1,\dots,q_k$, the sequence

$$\mathbf{z}_n = \left(\gamma_{q_1}(n),\dots,\gamma_{q_k}(n)\right), \qquad n=0,1,\dots,$$

is u.d. in $[0,1]^k$. This is the so called Halton sequence, cf. [DT, p. 369].

Faure sequences

The sequence $\mathbf{z}_n$ in $[0,1)^k$ is called (t,k)-sequence in base q, if for every $m>t$, all j, and all $0\le a_i<d^{d_i}$, $d_i\ge 0$ ($i=1,2,\dots,k$), satisfying $\sum_{i=1}^{k} d_i = m-t$, we have

$$\frac{\#\{\mathbf{z}_n : jq^m < n \le (j+1)q^m : \mathbf{z}_n\in I\}}{N} = |I|\,, \quad N=q^m\,,$$

where

$$I=\prod_{i=1}^{k}\left[\frac{a_i}{q^{d_i}},\frac{a_i+1}{q^{d_i}}\right).$$

Here the volume of k-dimensional interval I is $|I|=q^{t-m}$. These (t,k)-sequences are u.d. in $[0,1)^k$ and have the discrepancy

$$D_N = O\left(\frac{(\log N)^k}{N}\right).$$

For $t=0$, they are presently the "best possible" u.d. For $q=2$ this definition was found by I. M. Sobol' (1966) [So]. H. Faure (1982) [F] gives the following construction of the $(0,k)$-sequence $\mathbf{z}_n$:

Let $n = \sum_{j=0}^{m(n)} a_j(n)q^j$ be the q-adic expansion of n, $n = 0, 1, \dots$, and for any $\gamma \in [0,1)$ with q-adic expansion $\gamma = 0.c_1c_2\dots$ and for any infinite integer matrix C define $C \cdot \gamma = 0.d_1d_2\dots$, where $(d_1, d_2, \dots) = C \cdot (c_1, c_2, \dots)^T \pmod{q}$. Assume that

q is the smallest prime number with $q \geq s$;

C is the Pascal triangle matrix, for its

ith power $(C)^i = \left(C_{j,k}^{(i)}\right)$ we have

$$C_{j,k}^{(i)} = \begin{cases} \binom{j}{k} i^{j-k} & \text{for } 0 \leq k \leq j\,, \\ 0 & \text{for } k > j\,. \end{cases}$$

Let

$$z_{n,1} = \gamma_q(n) = \sum_{j=0}^{m(n)} a_j(n)q^{-(j+1)}\,, \qquad n = 0, 1, 2, \dots\,,$$

be the van der Corput sequence and and define

$$z_{n,i} = (C)^{i-1} \cdot z_{n,1}\,, \qquad i = 1, 2, \dots\,.$$

Then

$$\mathbf{z}_n = (z_{n,1}, \dots, z_{n,s})\,, \qquad n = 0, 1, 2, \dots,$$

is a $(0,s)$-sequence in base q. A general theory of (t,k)-sequences was developed by H. Niederreiter (see, e.g., [N]).

REFERENCES

[DT] DRMOTA, M.—TICHY, R. F.: in: Lecture Notes in Math., Vol. 1651, Springer-Verlag, Berlin, 1997.

[F] FAURE, H.: *Discrépance de suites associées à un système de numération (en dimension s)*, Acta Arith. **41** (1982), 337–351.

[K] KOBLITZ, N.: *A Course in Number Theory and Cryptography*, Grad. Text in Math., Vol. 114, Springer-Verlag, Berlin, 1987.

[L] LUBY, M.: *Pseudorandomness and Cryptographic Applications*, Princeton University Press, Princeton, 1996.

[MvV] MENEZES, A.—VAN OORSCHOT, P.—VANSTONE, S.: *Handbook of Applied Cryptography*, CRC Press, Singapore, 1996.

[N] NIEDERREITER, H.: *Random Number Generators and Quasi–Monte Carlo Methods*, (CBMS–NSF Regional Conf. Ser. in Appl. Math.), Vol. 63, Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[So] SOBOL', I. M.: *Distribution of points in a cube and integration nets*, Uspekhi Mat. Nauk **21** (1966), 271–272. (In Russian)

[S] STRAUCH, O.: *On the distribution functions of sequences generated by scalar and mixed product*, Math. Slovaca **53** (2003), 467–478.

Received November 17, 2003

Mathematical Institute
Slovak Academy of Sciences
Štefánikova 49
SK–814 73 Bratislava
SLOVAKIA
E-mail: strauch@ mat.savba.sk