UDC 51-74

Babash A.V.,
Dr. of Mathematical
and Physical sciences, Prof.
Kudiyarov D.S.,
postgraduate student
Russian State Social
University,
Russia

Conference participants,
National championship
in scientific analytics

ARCFOUR FINITE STATE MACHINE MODEL

Pseudorandom number generator ARCFOUR is modeled by two consecutively connectedfinite state machines. Using this modelsome characteristics ofARCFOUR internal permutation sequence period are calculated andprovided.

Keywords: ARCFOUR, pseudorandom, generator, machine, model

ARCFOURis a pseudorandom number generator known as RC4. The name RC4 is a trademark, so RC4 is often called ARCFOUR or ARC4 (Alleged RC4) to avoid trademark problems. [1]

Pseudorandom number generator RC4 was designed by Ron Rivest of RSA Security in 1987. The RC acronym usually understood by Ron's Code. [1]

Initially RC4 was a RSA company's property, but In September 1994 a description of it was anonymously posted on the Cypherpunks web page and in sci.crypt newsgroup. Then this information spreaded in the Internet.The posted code providedthe same output sequences that licensed RC4 did, so posted code was considered to be genuine. [2]

RSA hasn't officially published the algorithm yet. [1]

RC4 use is very wide because of such features as speed and simplicity. It's a part of standards like TLS, WEP and WPA. [2]

ARCFOUR is a set of pseudorandom number generators. It's internal state depends on n parameter. ARCFOURwithn = 8 isusedinpracticalapplications.

Let $V_N = Z_N \times Z_N \times S_N$ be an internal state set, where Z_N is a ring of integers modulo N, S_N is a symmetric group and $N = 2^n$. So the internal state for $n = 8$ is $V_{256} = Z_{256} \times Z_{256} \times S_{256}$.

Atapointoftime tARCFOURinternal state is $v_t = (i_t, j_t, s_t)$, where $v_t \in V_N$.

ARCFOURinclude two stages, called Key Scheduling Algorithm (KSA) and Pseudorandom Generation Algorithm (PRGA). The first one is an algorithm, which derives the initial state of ARCFOURfrom a key. PRGA isanalgorithm, whichgeneratesoutputdataoninitialstatebasis.

Initial state of ARCFOUR is derived from internal state achieved after KSA and equals $v_0 = (i_0, j_0, s_0) = (0, 0, s)$, where s is a permutation in internal state achieved after KSA.

At every moment $t \in [1; \infty)$ of PRGA ARCFOUR turns into the new internal state $v_t = (i_t, j_t, s_t)$ and one output number γ_t is generated:

$$i_t = i_{t-1} \oplus 1;$$
$$j_t = j_{t-1} \oplus s_{t-1}(i_t);$$
$$s_t = s_{t-1} \circ (s_{t-1}(i_t), s_{t-1}(j_t));$$
$$\gamma_t = s_t(s(i_t) \oplus s_t(j_t)),$$

where $\oplus$ – is addition modulo N operation, $s_t(z)$ – is a z-th number in permutation s_t, $\circ$ – permutation composition operation in symmetric group S_N, and (x, y) – is a transposition of x-th and y-th elements in symmetric group S_N (if $x = y$, then (x, y) – is identity permutation).

Let Z_N be a set of possible internal states of Moore machine A, δ: $Z_N \rightarrow Z_N$ is its transition function, λ: $Z_N \rightarrow Z_N$ is its output function. Let B to be the Mealy machine, X is its input alphabet, Σ is its state set, $(h_x)_{x \in X}$ are its partial transition functions, $(f_x)_{x \in X}$, $f_x : \Sigma \rightarrow Z$ are its partial output functions, Y is its output alphabet. We will use σ, $h_{x(1)}\sigma$, $h_{x(2)}h_{x(1)}\sigma$, ..., $h_{x(k)}... h_{x(2)}h_{x(1)}$, ... to denote a B internal states sequence, derived from initial state σ and input sequence $x(1)$, $x(2)$, ..., $x(k)$. c|u means c divides u.

We would remind that sequence b_1, b_2, …, b_t, … of some alphabet elements is called periodic if there is a natural number d which fulfills the condition that $b_t = b_{t+kd}$ for any natural numbers t, k. Minimal d that fulfills this condition is a period of the sequence.

Finite-state machine PRGA model is a consecutively connected machine A and B. Machine A input parameters are $\delta(i) = i \oplus 1$, $\lambda(i) = i \oplus 1$. It's clear, that $\lambda = \delta$. We should note, that if the initial state of machine A is $i = 0$, then its output sequence is $Q = 1, 2, \ldots N - 1, 0, \ldots$ with a period of N. Machine B parameters are

$$X = Z_N, \Sigma = Z_N \times S_N, Y = Z_N,$$
$$h_x : Z_N \times S_N \rightarrow Z_N \times S_N,$$
$$f_x : Z_N \times S_N \rightarrow Z_N$$

where $x \in Z_N$ and

$$h_x(j,s) = (j \oplus s(x)),$$
$$s \circ (s(x), s(j \oplus s(x))),$$
$$f_x(j,s) = s(s(x) \oplus s(j))$$

Further machine B with such parameters will be denoted as B*.

The initial state of PRGA is $v_0 = (i_0, j_0, s_0) = (0, 0, s_0)$. So the output sequence is $\gamma_t = s_t(s(i_t) \oplus s_t(j_t))$ where $t \in \{1, 2, ...\}$ and equals output sequence of machine B^*, generated from the initial state $(0, s_0)$ and input sequence $Q = 1, 2, \ldots N - 1, 0, 1, 2...$.

We are interested in possible and impossible periods of B^* output sequence $\tau(\gamma)$. The information about B^* internal state permutation $s_t \in \{1, 2, ...\}$ sequence $\tau(s)$ is also important. It's clear that $\tau(\gamma)|\tau(s)$ and $\tau(s)$ divides internal state sequence period (j_t, s_t), $t \in \{1, 2, ...\}$ of machine B^* with internal state s_0 and input sequence Q.

Statement 1. Sequences $\gamma_t \in \{1, 2, ...\}$, s_t, $t \in \{1, 2, ...\}$ are periodic.

Statement 1 proving. If we prove that B^* internal state sequence (j_t, s_t),

$t \in \{1, 2, ...\}$ corresponded to initial state s_0 is periodic we will prove the Statement 1. The proof that B^* is a permutative machine (means that $(hx)_{x\in X}$ are bijective) is enough to prove it.

Let's prove that machine B^* is permutative. Let's assume that B^* is not permutative machine. So for some input symbol $x \in Z_N$ and some internal states (j, s), (j^*, s^*) from $Z_N \times S_N$ the equation $h_x(j, s) = h_x(j^*, s^*)$ is true. So

$$\begin{aligned}&(j \oplus s(x)), s \circ (s(x), s(j \oplus (s(x)))) = \\ &= (j^* \oplus s^*(x)), s^* \circ (s^*(x), s^*(x)))\end{aligned}$$

and consequently

$$\begin{aligned}&j \oplus s(x) = j^* \oplus s^*(x), \\ &s \circ (s(x), s(j \oplus (s(x))) = \\ &= s^* \circ (s^*(x), s^*(j^* \oplus s^*(x)).\end{aligned}$$

Let

$$j'' = j \oplus s(x) = j^* \oplus s^*(x).$$

So

$$s \circ (s(x), s(j'') = s^* \circ (s^*(x), s^*(j'')).$$

It means that $s = s^*$ and consequently $j = j^*$. Itisacollision, whichproves Statement 1.

Most valuable result of current article is contained in further theorem.

Theorem. If input sequence is $Q = 1, 2, ... N - 1, 0, 1, 2...$ then the permutation s_t, $t \in \{1, 2, ...\}$ sequence period of machine B^* is multiple of 2^{n-1} for any initial permutation s_0. If inequality $s_0(1) \neq 1+2$ is satisfied for substitution s_0 then the permutation s_t, $t \in \{1, 2, ...\}$ sequence period of machine B^* is a multiple of $N = 2^n$.

We provide two lemmas in order to prove the theorem.

Lemma 1. There are two natural numbers $n(1)$, $n(2)$, which satisfy $a_n(1) - b_n(2) = d$, for any two natural numbers a, b satisfied equality $(a, b) = d$.

Lemma 1 may be proven with the well-known statement about existence of two integers n(1), n(2) fulfilled condition given above.

Let's look at machine B defined earlier. Let $B(\sigma)$ be a connected component of machine B which contains state $\sigma \in \Sigma$ and $B(\sigma, P)$ is an output sequence of permutative machine B with initial state σ and input sequence P. Denote $P^j = x(j), x(j + 1), \ldots$ for $P = x(1), x(2), ...$.

Lemma 2. If $P = x(1), x(2), ..., x(L), ...$ is periodic input sequence of permutative machine $B = (X, \Sigma, Y, (h_x)_{x\in X}, (f_x)_{x\in X})$ with period $\omega(P)$ and if $W(\sigma, P)$ is period of machine B output sequence $B(\sigma, P) = y_1, y_1, ..., y_L, ...$ then for any multigram with length L system $P^{kd+1} = x_{kd+1}, x_{kd+2}, ...$ where $k\in\{0, 1, ...\}$, $P^1 = P$, $d = (\omega(P), W(\sigma(P)))$ always exist states $s^*_{1+d}, s^*_{1+2d}, ..., s^*_{1+kd}, ...$ of connected component $B(\sigma_1)$, which satisfies

$$\begin{aligned}&B(\sigma, x_1, x_2, ..., x_L) = y_1, y_2, ..., y_L = \\ &= B(\sigma^*_{1+d}, P^{1+d}) = B(\sigma^*_{1+2d}, P^{1+2d}) = \\ &= B(\sigma^*_{1+kd}, P^{1+kd})... .\end{aligned}$$

Lemma 2 Proof. Machine B is permutative so for any machine B state σ and any periodic sequence P the output sequence $B(\sigma, P) = y_1, y_2, ..., y_L, ...$ is periodic.

Denote $\omega(P) = \omega$, $W(\sigma, P) = W$ further in this proof. It's evident that if $(w, W) = $ w then lemma 2 is true.

Assume that $(w, W) = d^1w$. Let k is non-negative integer. In according to lemma 1 there are integers $n(1)$, $n(2)$ satisfied $kn(1)w - kn(2)W = kd$. So

$$\begin{aligned}B(\sigma, P) &= B(\sigma, P^1) = \\ &= B(\sigma, P^{kd-kd+1}) = \\ &= B(\sigma, P^{d\ -kn(1)w+\ kn(2)W+1}) = \\ &= B(\sigma, P^{kd\ +\ kn(2)W+1}). \qquad (1)\end{aligned}$$

Let's choosenaturalnumbern (3) which fulfisthe condition $w|n(2) + n(3)$ andrewrite (1) (notethat $w(P^{kd+1}) = w$):

$$\begin{aligned}&B(\sigma, P^{kd\ +\ kn(2)W+1}) = \\ &= B(h_{x(kd+kn(2)W+n(3)W)}\cdots \rightarrow \\ &\rightarrow ...\ h_{x(kd+kn(2)W+1)}\sigma, P^{kd+n(2)W+n(3)W+1}) = \\ &= B(h_{x(kd+n(2)W+n(3)W)}\cdots \rightarrow \\ &\rightarrow ...\ h_{x(kd+n(2)W+1)}\sigma, P^{kd+1}).\end{aligned}$$

Evidently, the state

$$\sigma^*_{kd+1} = h_{x(kd+n(2)W+n(3)W)}...h_{x(kd+n(2)W+1)}\sigma$$

belongs to machine B connected component $B(\sigma)$. So existence of the states s^*_{d+1}, ..., s^*_{kd+1} and machine B connected component $B(\sigma)$ satisfied

$$\begin{aligned}&B(\sigma,P) = B(\sigma^*_{d+1}, P^{(d+1)}) = ... \rightarrow \\ &\rightarrow ... = B(\sigma^*_{kd+1}, P^{((k)d+1)})\end{aligned}$$

is proven. Solemma 2 isproventoo.

Let's prove theorem 1. In order to apply lemma 2 to B^* permutation s_t, $t \in \{1, 2, ...\}$ (generated from input sequence $Q = 1, 2, ..., N-1, 0, 1, 2, ...)$ sequence period research we will review permutation s_t, $t \in \{1, 2, ...\}$ sequence of machine B^* generated from input sequence $Q = 1, 2, ..., N-1, 0, 1, 2, ...$ as output sequence of machine B^*_F differed from machine B^* only by partial output functions $(F_x)_{x\in X}$, $F_x(j, s) = = s \circ (s(x), s(j \oplus (s(x)))$ for any permutation $s \in S_N$. Let's review internal states sequence

$$\begin{aligned}&B^{**}_F((0,s_0),Q) = \\ &= (j_1,s_1),(j_1,s_2),...,(j_\gamma,s_\gamma),...\end{aligned}$$

of machine B^*_F generated from initial state $(0, s_0)$ and input sequence $Q = 1, 2, ..., N - 1, 0, 1, 2, ...$ with a period of N. Note that there is a shift

$$\begin{aligned}(B^*_F((0,s_0),Q))^V &= s_V, s_{V+1},... = \\ &= s_0,s_1,s_2,...,s_\gamma,...\end{aligned}$$

of B^*_F output sequence

$$B^*_F((0,s_0),Q) = s_1,s_2,...,s_\gamma,...$$

which generated from initial state (j, s) and input sequence $Q^{N-1} = 0, 1, 2$ ... because

$$\begin{aligned}&(B^{**}_F((0,s_0),Q))^v = \\ &= (j_v,s_v),(j_{v+1},s_{V+1}),... = \\ &= B^*_F(j_v,s_v),Q^{N-1}) = \\ &= (0,s_0),(j_1,s_1),(j_2,s_2),...,(j_\gamma,s_\gamma),...\end{aligned}$$

Evidently sequence $B^*_F((0, s_0), Q)$ period W equals its shift $B^*_F((j, s), Q^{N-1}) = = s_0, s_1, s_2, ..., s_\gamma, ...$ period. Denote $\omega = N = 2^n$, $d = 2^m$, $L = 2$. Assume that $(2^n, W) = 2^m$, $m < n$. Applying lemma 2 to sequence

$$B^*_F((j,s),Q^{N-1}) = s_0,s_1,s_2,...,s_\gamma,...,$$

we can see, that for any bigram from sequence $Q^{N-1} = 0, 1, 2, ...$ like (0,1); $(d, 1+d)$; $(2d, 1+2d)$; …; $(kd, 1+kd)$; … where $d = 2^m$ will be found states (j^*_d, s^*_d), (j^*_{2d}, s^*_{2d}), ..., (j^*_{kd}, s^*_{kd}), ... of connected component $B^*_F(0, s_0)$ satisfied

$$\begin{aligned}&B^*_F((j,s),0,1) = s_0,s_1, \\ &B^*_F((j^*_d,s^*_d),d,1+d) = s_0,s_1, \\ &\text{.......................................} \qquad (2) \\ &B^*_F((j^*_{kd},s^*_{kd}),kd,1+kd) = s_0,s_1, \\ &k \in \{1,2,...\}.\end{aligned}$$

From (2) weget:

$$\begin{aligned}&F_{kd}(j^*_{kd},s_0) = \\ &= s_0 \circ (s_0(1+kd),(s_0(j^*_{kd} \oplus s_0(1+kd))) = s_1;\end{aligned}$$

$$s_0 \circ (s_0(1), s_0(s_0(1)) =$$
$$= s_0 \circ (s_0(1+kd), (s_0(j^*_{kd} \oplus s_0(1+kd)));$$
$$(s_0(1), s_0(s_0(1)) =$$
$$= (s_0(1+kd), s_0(j^*_{kd} \oplus s_0(1+kd))).$$

If *kd* is not a multiple of *N* then from last equality we get:

$$s_0(1) = s_0(j^*_{kd} \oplus s_0(1+kd));$$
$$s_0(s_0(1)) = s_0(1+kd).$$

So:

$$\begin{aligned} 1 &= j^*_{kd} \oplus s_0(1+kd); \\ s_0(1) &= 1+kd. \end{aligned} \tag{3}$$

Equality (3) is possible just if $d = 2^{n-1}$. So the assumption $(2^n, W) = 2^m$, $m < n$ is possible just if $s_0(1) = 1+2^{n-1}$.

The theorem is proven.

References:

1. Schneier, Bruce / «Applied cryptography Second Edition : protocols, algorithms, and source codes in C» / 1996.

2. Rivest,Ron L. «Ronald L. Rivest : FAQ» / MIT Computer Science and Artificial Intelligence Laboratory. http://people.csail.mit.edu/rivest/faq.html#RonM. (30.08.2012).