The number of initial states of the RC4 cipher with the same cycle structure

Marina Pudovkina

Moscow Engineering Physics Institute (State University)

maricap@online.ru

Abstract. RC4 cipher is the most widely used stream cipher in software applications. It was designed by R. Rivest in 1987. In this paper we find the number of keys of the RC4 cipher generating initial permutations with the same cycle structure. We obtain that the distribution of initial permutations is not uniform.

1. Introduction

RC4 cipher is the most widely used stream cipher in software applications. It was designed by R. Rivest in 1987. There are several papers by analysis of RC4, where several attacks and vulnerabilities were described. Most of these attacks revolve around the concept of a distinguisher. The first distinguisher was Golic [1] that exploited correlation between z_i and z_{i+2}. The results due to [2, 3] and their generalization [8] are of most practical importance.

In [9] is introduced the idealized model of RC4 and considered the cipher as a random walk on a symmetric group. Mironov proved a necessary condition for existence of strong distinguishers in the idealized model.

In this paper we find the number of keys of the RC4 cipher generating initial permutations with the same cycle structure. We obtain that the distribution of initial permutations is not uniform.

Thus, we show that the probability of generating the identical permutation is in $\sqrt{\pi}\ \dfrac{m^{(m+1)/2}}{e^{3/2m-\sqrt{m}+1/4}}$

more what you would expect.

2. Description of the RC4 cipher

The RC4 stream cipher is modeled by a finite automaton $A_g = (F, f, Z_m \times Z_m \times S_m, Z_m)$, where F: $Z_m \times Z_m \times S_m \to Z_m \times Z_m \times S_m$ is a next-state function, f: $Z_m \times Z_m \times S_m \to Z_m$ is an output function. The RC4 stream cipher depends on $m=2^n$, $n \in N$.

The state of the RC4 cipher at time t is $(i_t, j_t, s_t) \in Z_m \times Z_m \times S_m$ and the initial state is $(0, 0, s_0,)$.

Key schedule algorithm ρ

s_0 is the identical permutation, $i_0 = j_0 = 0$.

For $t = \overline{1, m}$ do:

1. $i_t = t - 1$,
2. $j_t = j_{t-1} + s_{t-1}[i_t] + k_{t-1(mod\ L)}$ (mod m)*,
3. $s_t[i_t] = s_{t-1}[j_t]$, $s_t[j_t] = s_{t-1}[i_t]$;
4. $s_t[r] = s_{t-1}[r]$, $r = \overline{0, m-1} \setminus \{i_t, j_t\}$.

Consider the RC4 cipher at time t (t=1,2….).

The next-state function F

1. $i_t = i_{t-1} + 1 \pmod{m}$;
2. $j_t = j_{t-1} + s_{t-1}[i_t] \pmod{m}$;
3. $s_t[i_t] = s_{t-1}[j_t]$, $s_t[j_t] = s_{t-1}[i_t]$;
4. $s_t[r] = s_{t-1}[r]$, $r = \overline{0, m-1} \setminus \{i_t, j_t\}$.

The output function f
Output: $z_t = s_t[(s_t[j_t] + s_t[i_t])(\bmod\ m)]$.

Encryption x_t: $c_t = x_t \oplus z_t$. Decryption c_t: $x_t = c_t \oplus z_t$.

3. The number of states with the same cycle structure

In this section we prove our main result, i.e. we find the number of keys of RC4 generating initial permutations with the same cycle structure. We begin with definitions.

By $B(\alpha_1, \ldots, \alpha_m)$ denote the set of all permutations from S_m with the cycle structure $\{1^{\alpha_1} 2^{\alpha_2} \ldots m^{\alpha_m}\}$, where $1 \cdot \alpha_1 + 2 \cdot \alpha_2 + \ldots + m \cdot \alpha_m = m$. It is known that $|B(\alpha_1, \ldots, \alpha_m)| = \dfrac{m!}{1^{\alpha_1} 2^{\alpha_2} \ldots m^{\alpha_m} \cdot \alpha_1! \alpha_2! \ldots \cdot \alpha_m!}$. Let a permutation $s \in S_m$ are chosen randomly from S_m, then $P\{s \in B(\alpha_1, \ldots, \alpha_m)\} = \dfrac{1}{1^{\alpha_1} 2^{\alpha_2} \ldots m^{\alpha_m} \cdot \alpha_1! \alpha_2! \ldots \cdot \alpha_m!}$.

Note that if the distribution of initial permutations $\rho(k)$ of the RC4 cipher is uniform, then $P\{\rho(k) \in B(\alpha_1, \ldots, \alpha_m)\} = \Omega(\alpha_1, \ldots, \alpha_m) \dfrac{1}{m^m} = \dfrac{1}{1^{\alpha_1} 2^{\alpha_2} \ldots m^{\alpha_m} \cdot \alpha_1! \alpha_2! \ldots \cdot \alpha_m!}$, where $\Omega(\alpha_1, \ldots, \alpha_m) = |\{k \in Z_m^m : \rho(k) \in B(\alpha_1, \ldots, \alpha_m)\}|$. The average of keys generating initial permutations with the same cyclic structure is $N_m(\alpha_1, \ldots, \alpha_m) = \dfrac{m^m \cdot m!}{m! 1^{\alpha_1} 2^{\alpha_2} \ldots m^{\alpha_m} \cdot \alpha_1! \alpha_2! \ldots \cdot \alpha_m!} = \dfrac{m^m}{1^{\alpha_1} 2^{\alpha_2} \ldots m^{\alpha_m} \cdot \alpha_1! \alpha_2! \ldots \cdot \alpha_m!}$.

Note that multiplication to the left transposition (i_t, j_t) by s_t is swapped elements $s_t[i_t]$ and $s_t[j_t]$, i.e. $s_{t+1} = (i_t, j_t) s_t$. Therefore, the initial permutation of RC4 can be represented as $s_m = (m-1, j_m) \ldots (1, j_2)(0, j_1) \cdot E$, where E is the identical permutation.

Thus, we'll consider the multiset $\vartheta_m = \{(m-1, j_m) \ldots (1, j_2)(0, j_1) \mid (j_m, \ldots, j_1) \in Z_m^m\}$, which elements are permutations that can be represented as $(m-1, j_m) \ldots (1, j_2)(0, j_1)$, where $k = \overline{1, m}$, $j_k \in \overline{0, m-1}$. Note also that the permutation s can be represented as $(m-1, j_m) \ldots (1, j_2)(0, j_1)$ different ways.

Let $A_\eta = \{i_1, \ldots, i_\eta\}$, $|A_\eta| = \eta$, be a directed set and $i_1 < i_2 < \ldots < i_\eta$. Let the set $J_\eta = \{j_\eta, \ldots, j_1\}$, where $|J_\eta| \le \eta$ and either $J_\eta \cap A_\eta = \varnothing$ or $J_\eta \cap A_\eta \ne \varnothing$. Denote $\varpi = |A_\eta \cup J_\eta|$.

Let us consider the product of transpositions $(i_\eta, j_\eta) \ldots (i_2, j_2)(i_1, j_1)$ that is generated by the key-schedule algorithm of RC4. The product of transpositions $(i_\eta, j_\eta) \ldots (i_2, j_2)(i_1, j_1)$ corresponds to the ϖ vertices graph Ξ_ϖ that vertices are the elements of the set $A_\eta \cup J_\eta$ and that edges are the pairs (i_r, j_r) labeled by i_r, $r = \overline{1, \eta}$. Note that the transposition (α, α) is a loop of Ξ_ϖ; the labels' set of Ξ_η is A_η and the number of edges is η.

The order relation on the set A_η is induced on the set of labels, i.e. $i_1 < i_2 < \ldots < i_\eta$ on the labels' set. For $A_\eta = \{0, 1, \ldots, m-1\}$, $J_\eta \subseteq A_\eta$, $\eta = m$ and $i_k = k-1$, the product of transpositions $(0, j_1), (1, j_2), \ldots, (m-1, j_m)$ corresponding to some key of RC4 uniquely defines the graph Ξ_m.

The Ξ_m graph may have several components. Let Λ_k on k vertices labeled by elements of $X_k = \{x_1, x_2, \ldots, x_k\}$, where $X_k \subseteq \{0, 1, \ldots, m-1\}$, be a component of Ξ_m.

We say that the connected graph Λ_n disintegrates into two cycles of length i and j, if the product of transpositions corresponding to Λ_n can be represented as products of two independent cycles of length i and j. By $\Gamma_{i,j}$ denote a connected graph on $i+j$ vertices labeled by elements of $X=\{x_1, x_2,\ldots,x_{i+j}\}$ that disintegrates into two cycles of length i and j.

To prove the main theorem we need the following proposition.

Proposition 1

1. *The number $N_{i,j}(n)$ of graphs $\Gamma_{i,j}$ on $i+j$ vertices, which are labeled by elements of the set $\{x_1,x_2,\ldots,x_{i+j}\}$, and having a length n cycle is*

$$N_{i,j}(n)= \binom{i+j}{n}\left(\sum_{k=1}^{n-1} \binom{i+j-n}{i-k} \cdot k \cdot (n-k) \cdot i^{i-k-1} \cdot j^{j-n+k-1} A_{n-1,k-1} \right), \quad (1)$$

2. *The number $N_{i,j}(n)$ of graphs $\Gamma_{i,j}$ on $i+j$ vertices, which are labeled by elements of the set $\{x_1,x_2,\ldots,x_{i+j}\}$*

a) for $i>0$ and $j>0$ is

$$N_{i,j}=\sum_{n=2}^{i+j} \binom{i+j}{n}\left(\sum_{k=1}^{n-1} \binom{i+j-n}{i-k} \cdot k \cdot (n-k) \cdot i^{i-k-1} \cdot j^{j-n+k-1} A_{n-1,k-1} \right),$$

b) for $i>0$, $j=0$ or $j>0$, $i=0$ is

$$N_{i,0}=j^{j-1},$$

where $A_{n,k} = \sum_{t=0}^{k}(-1)^t \binom{n+1}{t}(k-t+1)^n$, $k=\overline{0,n-1}$.

The proof is by direct calculations.

By $\Lambda(k_{i,j})$ denote a graph with $k_{i,j}$ components that disintegrate into cycles of length i and j. Let the graph $\Lambda=\bigcup_{k_{ij}}\Lambda(k_{ij})$. The m-dimensional vector $\vec{k}_j =(k_{0j}, k_{1j},\ldots, k_{mj})$ that the i-th coordinate is the number of components $\Lambda(k_{i,j})$ of Λ.

Our main result is the following.

Theorem 2. *The number of different products of transpositions $(m-1,j_m)\ldots(1,j_2)(0,j_1)$, where $j_k\in\{\overline{0,m-1}\}$, $k=\overline{1,m}$, generating permutations with the same cyclic structure $\{1^{\alpha_1}2^{\alpha_2}\ldots m^{\alpha_m}\}$, $1\cdot\alpha_1+2\cdot\alpha_2+\ldots+m\cdot\alpha_m=m$, is*

$$\Omega(\alpha_1,\ldots,\alpha_m)= m!\cdot\prod_{r=1}^{m} r^{(r-1)\alpha_r} \cdot \sum_{\substack{(\vec{k}_1,\ldots,\vec{k}_m):0\leq k_{0j}\leq\alpha_j, \\ 0<i\leq j, k_{ij}\leq\min(\alpha_j,\alpha_i), \\ \sum_{t=0}^{m}k_{tj}+k_{jj}=\alpha_j,k_{ij}=k_{ji}}} \prod_{i\leq j} \frac{\left(C_{ij}\right)^{k_{ij}}}{k_{ij}!\cdot(i+j)!^{k_{ij}}}, \quad (2)$$

where summation is carried out by the first $j+1$ coordinates $(k_{0j}, k_{1j},\ldots, k_{jj})$ of the vector $\vec{k}_j =(k_{0j}, k_{1j},\ldots, k_{mj})\in Z_m^{m+1}$, $j=\overline{1,m}$, and

$$C_{ij}= \sum_{n=2}^{i+j} \binom{i+j}{n}\left(\sum_{k=1}^{i} \binom{i+j-n}{i-k} \cdot k \cdot (n-k) \cdot i^{-k} \cdot j^{k-n} A_{n-1,k-1} \right) \text{ for } 0<i\leq j,$$

$$C_{0j}=1,$$

$$A_{n,k} = \sum_{t=0}^{k}(-1)^t \binom{n+1}{t}(k-t+1)^n, \quad k=\overline{0,n-1}.$$

Proof. We stress that $\Lambda(k_{i,j})$ consists of $k_{i,j}$ graphs $\Gamma_{i,j}$, which vertices are labeled by elements of the set $\{x_1, x_2,\ldots, x_{k_{ij}(i+j)}\}$. The number of ways to distribute of labels between $k_{i,j}$ components is

$$\frac{\binom{k_{ij}(i+j)}{i+j}\binom{(k_{ij}-1)(i+j)}{i+j}\cdots\binom{i+j}{i+j}}{k_{ij}!}=\frac{(k_{ij}(i+j))!}{k_{ij}!\cdot(i+j)!^{k_{ij}}}.$$

Therefore, by proposition 1 the number of various graphs $\Lambda(k_{i,j})$ having $k_{i,j}$ components is

$$M(k_{ij})=\frac{(k_{ij}(i+j))!}{k_{ij}!\cdot(i+j)^{k_{ij}}}\cdot\left(N_{ij}\right)^{k_{ij}}. \qquad (3)$$

Let the graph $\Lambda=\bigcup_{k_{ij}}\Lambda(k_{ij})$ which vertices labeled by elements of the set $\{x_1, x_2,\ldots, x_m\}$ have $\nu=\sum_{i\leq j}k_{ij}$ components.

The number ways of distributing labels between ν components of Λ is

$$\binom{m}{k_{01}}\binom{m-k_{01}}{2k_{02}}\binom{m-k_{01}-2k_{02}}{2k_{11}}\binom{m-k_{01}-2k_{02}-2k_{11}}{3k_{03}}\cdots\binom{m-k_{01}\ldots-\beta\cdot k_{i,\beta-i}}{\beta\cdot k_{i,\beta-i}}\cdots\binom{m-k_{01}\ldots-mk_{0,m}}{mk_{1,m-1}}=$$

$$\frac{m!}{\prod_{i\leq j}\left(k_{ij}(i+j)\right)!}. \qquad (4)$$

From (3), (4) we obtain that the number of products transpositions $(m-1,j_m)\ldots(1,j_2)(0,j_1)$ generating permutations with the same cycle structure $\{1^{\alpha_1}2^{\alpha_2}\ldots m^{\alpha_m}\}$ and with the $\{\vec{k}_1,\ldots,\vec{k}_m\}$ vectors is

$$\Omega(\alpha_1,\ldots,\alpha_m;\ \vec{k}_1,\ldots,\vec{k}_m)=\frac{m!}{\prod_{i\leq j}\left(k_{ij}(i+j)\right)!}\cdot\prod_{i\leq j}\frac{(k_{ij}(i+j))!}{k_{ij}!\cdot(i+j)!^{k_{ij}}}\cdot\left(N_{ij}\right)^{k_{ij}},$$

where $\sum_{i=0}^{m}k_{ij}+k_{jj}=\alpha_j$, $j=\overline{1,m}$.

Therefore,

$$\Omega(\alpha_1,\ldots,\alpha_m)=\sum_{\substack{(\vec{k}_1,\ldots,\vec{k}_m):0\leq k_{0j}\leq\alpha_j,\\ 0<i\leq j,k_{ij}\leq\min(\alpha_j,\alpha_i),\\ \sum_{t=0}^{m}k_{tj}+k_{jj}=\alpha_j,k_{ij}=k_{ji}}}\Omega(\alpha_1,\ldots,\alpha_m;\vec{k}_1,\ldots,\vec{k}_m)=m!\cdot\sum_{\substack{(\vec{k}_1,\ldots,\vec{k}_m):0\leq k_{0j}\leq\alpha_j,\\ 0<i\leq j,k_{ij}\leq\min(\alpha_j,\alpha_i),\\ \sum_{t=0}^{m}k_{tj}+k_{jj}=\alpha_j,k_{ij}=k_{ji}}}\prod_{i\leq j}\frac{\left(N_{ij}\right)^{k_{ij}}}{k_{ij}!\cdot(i+j)!^{k_{ij}}} \qquad (5)$$

To compute (5) we note that

$$N_{i,j}=C_{ij}\cdot i^{i-1}\cdot j^{j-1},\ N_{i,i}=C_{ii}\cdot i^{2i-2},$$

where

$$C_{ij}=\sum_{n=2}^{i+j}\binom{i+j}{n}\left(\sum_{k=1}^{i}\binom{i+j-n}{i-k}\cdot k\cdot(n-k)\cdot i^{-k}\cdot j^{k-n}A_{n-1,k-1}\right)\text{ for } 0<i\leq j,$$

$$C_{0j}=1.$$

We also note that

$$\prod_{i\leq j}\left(N_{ij}\right)^{k_{ij}}=\prod_{r=1}^{m}r^{(r-1)\left(\sum_{j=1}^{m}k_{rj}+k_{rr}\right)}\cdot\prod_{i\leq j}\left(C_{ij}\right)^{k_{ij}}=\prod_{r=1}^{m}r^{(r-1)\alpha_r}\cdot\prod_{i\leq j}\left(C_{ij}\right)^{k_{ij}}.$$

Thus, we rewrite (5) as

$$\Omega(\alpha_1,\ldots,\alpha_m)= m!\cdot\prod_{r=1}^{m} r^{(r-1)\alpha_r} \cdot \sum_{\substack{(\bar{k}_1,\ldots,\bar{k}_m):0\le k_{0j}\le\alpha_j, \\ 0<i\le j, k_{ij}\le\min(\alpha_j,\alpha_i), \\ \sum_{t=0}^{m} k_{tj}+k_{jj}=\alpha_j, k_{ij}=k_{ji}}} \prod_{i\le j} \frac{(C_{ij})^{k_{ij}}}{k_{ij}!\cdot(i+j)!^{k_{ij}}} \cdot$$

The theorem is proved.

4. The asymptotic distribution of the number initial states with some cycle structure

In this section we describe distributions of the number keys of RC4 generating permutations with the following cycle structure $\{1^0\,2^0\ldots m^1\}$, $\{1^m\,2^0\ldots m^0\}$, $\{1^{m-d}\,2^0\ldots d^1\ldots m^0\}$.

First we compute the number of RC4 keys generating the cycle structure $\{1^0\,2^0\ldots m^1\}$. By (2) we have $\Omega(0,\ldots,0,1)=m^{m-1}$. Therefore, $P\{\rho(k)\in B(0,\ldots,0,1)\}=\dfrac{1}{m}$. Note that if s is randomly chosen from S_m, then $P\{s\in B(0,\ldots,0,1)=\dfrac{1}{m}$.

Let us now compute the number of keys of RC4 generating identical permutations. By (2) we get

$$\Omega(m,0,\ldots,0)= \sum_{k_{01}+2k_{11}=m} \frac{m!}{k_{01}!k_{11}!\cdot2^{k_{11}}} =\sum_{k=0}^{m/2} \frac{m!}{k!\cdot(m-2\cdot k)!\cdot2^{k}} \cdot$$

It is known [10] that $\Omega(m,0,\ldots,0) = \dfrac{1}{e^{1/4}\sqrt{2}}\left(\dfrac{m}{e}\right)^{m/2}\cdot e^{\sqrt{m}}\left(1+O\left(\dfrac{1}{\sqrt{m}}\right)\right)$ as $m\to\infty$.

Let $\quad P\{s\in B(\alpha_1,\ldots,\alpha_m)\}=\dfrac{1}{1^{\alpha_1}2^{\alpha_2}\ldots m^{\alpha_m}\cdot\alpha_1!\alpha_2!\ldots\cdot\alpha_m!}$, $\quad N_m(\alpha_1,\ldots,\alpha_m)=\dfrac{m^m}{1^{\alpha_1}2^{\alpha_2}\ldots m^{\alpha_m}\cdot\alpha_1!\alpha_2!\ldots\cdot\alpha_m!}$

$P\{\rho(k)\in B(\alpha_1,\ldots,\alpha_m)\}=\Omega(\alpha_1,\ldots,\alpha_m)\dfrac{1}{m^m}$, i.e. if suppose that s is uniformly random.

Proposition 3. *Let* $\Delta_m(\alpha_1,\ldots,\alpha_m)=|\ \Omega(\alpha_1,\ldots,\alpha_m)\ -\ N_m(\alpha_1,\ldots,\alpha_m)|$, $\delta_m(\alpha_1,\ldots,\alpha_m)=\Omega(\alpha_1,\ldots,\alpha_m)/N_m(\alpha_1,\ldots,\alpha_m)$, $m\ge4$. *Then for* $m\to\infty$

$$\Delta_m(m,0,\ldots,0) \sim \frac{1}{e^{1/4}\sqrt{2}}\left(\frac{m}{e}\right)^{m/2}\cdot e^{\sqrt{m}}-\frac{e^m}{\sqrt{2\pi m}}, \qquad (6)$$

$$\delta_m(m,0,\ldots,0) \sim \sqrt{\pi}\,\frac{m^{(m+1)/2}}{e^{3/2m-\sqrt{m}+1/4}} \cdot$$

Proof. The following are true.

$$\Delta_m(m,0,\ldots,0)= \sum_{k=0}^{m/2} \frac{m!}{k!\cdot(m-2\cdot k)!\cdot2^{k}} - \frac{m^m}{m!} \sim \frac{1}{e^{1/4}\sqrt{2}}\left(\frac{m}{e}\right)^{m/2}\cdot e^{\sqrt{m}}-\frac{e^m}{\sqrt{2\pi m}} \cdot$$

$$\delta_m(\alpha_1,\ldots,\alpha_m) \sim \frac{1}{e^{1/4}\sqrt{2}}\left(\frac{m}{e}\right)^{m/2}\cdot e^{\sqrt{m}}\cdot\frac{\sqrt{2\pi m}}{e^m}=\sqrt{\pi}\,\frac{m^{(m+1)/2}}{e^{3/2m-\sqrt{m}+1/4}}$$

This completes the proof.

For some values of m numeric dates are resulted in tab. 1.

Table 1
The number of keys of RC4 generating identical permutations for different values of m

m	$\Omega(m,0,\dots,0)$	$N_m(m,0,\dots,0)$	$P\{\rho(k)\in B(m,0,\dots,0))\}$	$P\{s\in B(m,0,\dots,0)\}$
1024	$3{,}6{\cdot}10^{1332}$	$6{,}4{\cdot}10^{442}$	$1{,}0{\cdot}10^{-1749}$	$1{,}8{\cdot}10^{-2639}$
512	$9{,}2{\cdot}10^{591}$	$4{,}0{\cdot}10^{220}$	$6{,}6{\cdot}10^{-796}$	$2{,}9{\cdot}10^{-1167}$
256	$2{,}30{\cdot}10^{260}$	$3{,}76{\cdot}10^{110}$	$7{,}12{\cdot}10^{-358}$	$1{,}16{\cdot}10^{-507}$
128	$5{,}3{\cdot}10^{111}$	$1{,}3{\cdot}10^{54}$	$1{,}0{\cdot}10^{-158}$	$2{,}6{\cdot}10^{-216}$
64	$1{,}4{\cdot}10^{47}$	$1{,}3{\cdot}10^{26}$	$3{,}4{\cdot}10^{-69}$	$7{,}9{\cdot}10^{-90}$
32	$2{,}2{\cdot}10^{19}$	$5{,}6{\cdot}10^{12}$	$1{,}5{\cdot}10^{-29}$	$3{,}8{\cdot}10^{-36}$
16	$4{,}6{\cdot}10^{7}$	$8{,}8{\cdot}10^{5}$	$2{,}5{\cdot}10^{-12}$	$4{,}8{\cdot}10^{-14}$

From table 1 we see that the distribution of initial permutations is not uniform.

Remark. It is known [10] that the number of $T_m^{(p)}$ solutions of the equation $s^p=E$, where p is a prime number, in the symmetric group S_m is

$$T_m^{(p)}= \sum_{k=0}^{m/p} \frac{m!}{k!\cdot(m-p\cdot k)!\cdot p^k}\,,$$

and as $m\to\infty$

$$T_m^{(2)}=\frac{1}{e^{1/4}\sqrt{2}}\left(\frac{m}{e}\right)^{m/2}\cdot e^{\sqrt{m}}\left(1+O\left(\frac{1}{\sqrt{m}}\right)\right),$$

$$T_m^{(p)}=\left(\frac{m}{e}\right)^{m(1-1/p)} p^{-1/2}\cdot e^{\sqrt[p]{m}}\left(1+o(1)\right) \text{ for } p>2.$$

It is easy to see that $\Omega(m,0,\dots,0)=T_m^{(2)}$.

Proposition 4. *Let $\Omega(m{-}d,0,\dots,1,0\dots0)$ be the number of keys of RC4 generating the set of permutations with the cycle structure $\{1^{m-d}\,2^0\dots d^l\dots m^0\}$. Then*

1. $$\Omega(m{-}d,0,\dots,1,0\dots0)=m!\cdot d^{d-l}\cdot\left(\frac{1}{d!(m-d)!}\sum_{k}^{\left[\frac{m-d}{2}\right]}\frac{(m-d)!}{k!\cdot(m-d-2k)!\cdot2^k}+\frac{d}{(d+1)!(m-d-1)!}\right.$$

$$\left.\sum_{k}^{\left[\frac{m-d-1}{2}\right]}\frac{(m-d-1)!}{k!\cdot(m-d-1-2k)!\cdot2^k}\right);$$

2. *for $(m{-}d)\to\infty$, $d\le c$, where c is a constant*

$$\Omega(m{-}d,0,\dots,1,0\dots0)\sim d^{d-l}\cdot\frac{\sqrt{2}\cdot e^{\sqrt{(m-d)}}}{d!\cdot e^{1/4}}\left(\frac{m}{e}\right)^{(m+d)/2},$$

3. *for $(m{-}d)\to\infty$, $d\to\infty$, $d=O(m)$*

$$\Omega(m{-}d,0,\dots,1,0\dots0)\sim\frac{e^{\sqrt{(m-d)}+d}}{\sqrt{\pi d}\,d\cdot e^{1/4}}\left(\frac{m}{e}\right)^{(m+d)/2},$$

4. *for $(m{-}d)\to\infty$, $d=o(m)$, $(m{-}d)/m\to c<1$,*

$$\Omega(m{-}d,0,\dots,1,0\dots0)\sim\left(\frac{1}{c}\right)^{(c\cdot m+1)/2}\frac{m^{m(2-c)/2-3/2}}{e^{m/2+1/4-\sqrt{m\cdot c}}}\,\frac{1}{\sqrt{\pi}(1-c)^{3/2}},$$

5. *for $(m{-}d)\to0$,*

$$\Omega(m{-}d,0,\dots,1,0\dots0)\sim m^{m-l}.$$

The proof is by direct calculation.

Proposition 5. *Let* $\Delta_m(\alpha_1,...,\alpha_m)=|\ \Omega(\alpha_1,...,\alpha_m)\ -\ N_m(\alpha_1,...,\alpha_m)|,\ \delta_m(\alpha_1,...,\alpha_m)=\ \Omega(\alpha_1,...,\alpha_m)/$
$N_m(\alpha_1,...,\alpha_m),\ m\geq4.$ *Then*

1. *for $(m{-}d)\to\infty$, $d\leq c$, where c is a constant*

$$\Delta_m(m{-}d,0,...,1,0...0) \sim d^{d-1}\cdot\frac{\sqrt{2}\cdot e^{\sqrt{(m-d)}}}{d!\cdot e^{1/4}}\left(\frac{m}{e}\right)^{(m+d)/2}-\frac{m^d\cdot e^{m-d}}{\sqrt{2\pi md}},$$

$$\delta_m(m{-}d,0,...,1,0...0)\sim\sqrt{\pi m}\ \left(\frac{m}{e}\right)^{(m-d)/2}\frac{2d^d}{d!e^{m-\sqrt{m-d}+1/4}},$$

2. *for $(m{-}d)\to\infty$, $d\to\infty$, $d=O(m)$*

$$\Delta_m(m{-}d,0,...,1,0...0) \sim \frac{e^{\sqrt{(m-d)}+d}}{\sqrt{\pi d}d\cdot e^{1/4}}\left(\frac{m}{e}\right)^{(m+d)/2}-\frac{m^d\cdot e^{m-d}}{\sqrt{2\pi md}},\qquad(7)$$

$$\delta_m(m{-}d,0,...,1,0...0)\sim\sqrt{\frac{2}{d}}\ \frac{m^{(m-d+1)/2}}{e^{(m-5)/2-\sqrt{m-d}+1/4}}.$$

3. *for $(m{-}d)\to\infty$, $d=o(m)$, $(m{-}d)/m\to c<1$,*

$$\Delta_m(m{-}d,0,...,1,0...0)\sim\left(\frac{1}{c}\right)^{(c\cdot m+1)/2}\frac{m^{m(2-c)/2-3/2}}{e^{m/2+1/4-\sqrt{m\cdot c}}}\ \frac{1}{\sqrt{\pi}(1-c)^{3/2}}-\frac{m^{m(1-c)}\cdot e^{c\cdot m}}{\sqrt{2\pi m}m(1-c)},$$

$$\delta_m(m{-}d,0,...,1,0...0)\sim\left(\frac{1}{c}\right)^{(c\cdot m+1)/2}\ \sqrt{\frac{2}{1-c}}\ \frac{m^{m/2-m(1-c)/2}}{e^{m/2+c\cdot m+1/4-\sqrt{m\cdot c}}}.$$

The proof is by direct calculation.

Let $\Omega_{m,d}=\Omega(m{-}d,0,...,1,0...0)$, $N_{m,d}=N_m(m{-}d,0,...,1,0...0)$, $P_{m,d}^{(t)}=P\{\rho(k)\in B(m{-}d,0,...,1,$
$0...0)\}$, $P_{m,d}^{(p)}=P\{s\in B(m{-}d,0,...,1,0...0)\}$.

Table 2
The number of keys of RC4 generating the set of permutations with the cyclic structure $\{1^{m-d}$
$2^0...d^1...m^0\}$ for different m

m	d	$\Omega_{m,d}$	$N_{m,d}$	$P_{m,d}^{(t)}$	$P_{m,d}^{(p)}$
2048	1024	$2,3\cdot10^{5028}$	$7,2\cdot10^{4138}$	$5,7\cdot10^{-1754}$	$1,8\cdot10^{-2643}$
1024	512	$2,6\cdot10^{2284}$	$2,0\cdot10^{1915}$	$7,4\cdot10^{-799}$	$5,6\cdot10^{-1169}$
512	256	$2,3\cdot10^{1027}$	$6,4\cdot10^{877}$	$1,6\cdot10^{-360}$	$4,6\cdot10^{-510}$
256	128	$1,5\cdot10^{456}$	$6,5\cdot10^{398}$	$4,7\cdot10^{-160}$	$2,0\cdot10^{-217}$
128	64	$1,8\cdot10^{201}$	$6,5\cdot10^{178}$	$3,1\cdot10^{-71}$	$1,2\cdot10^{-90}$
2048	64	$1,8\cdot10^{3097}$	$1,2\cdot10^{1097}$	$4,5\cdot10^{-3685}$	$2,9\cdot10^{-5685}$
1024	32	$1,3\cdot10^{1393}$	$2,7\cdot10^{537}$	$3,8\cdot10^{-1690}$	$7,6\cdot10^{-2545}$
512	16	$3,3\cdot10^{619}$	$4,4\cdot10^{262}$	$2,3\cdot10^{-768}$	$3,2\cdot10^{-1124}$
256	8	$5,7\cdot10^{271}$	$7,8\cdot10^{127}$	$1,7\cdot10^{-345}$	$2,4\cdot10^{-489}$
128	4	$1,9\cdot10^{117}$	$8,8\cdot10^{61}$	$3,5\cdot10^{-153}$	$1,7\cdot10^{-208}$
2048	2	$1,9\cdot10^{2970}$	$5,0\cdot10^{893}$	$4,7\cdot10^{-3812}$	$1,3\cdot10^{-5888}$

Obtained results are presented in Fig. 1.

5. Conclusion

In this paper we computed the number of keys of RC4 generating initial permutations with the same cyclic structure. We find that the distribution is not uniformly random and prove that the probability of generating the identical permutation is in $\sqrt{\pi}\ \dfrac{m^{(m+1)/2}}{e^{3/2m-\sqrt{m}+1/4}}$ more what you would expect. Therefore, to determine an initial permutation of RC4 we can do the following.

Let $\vec{\alpha}=(\alpha_1,\ldots,\alpha_m)$ be a m-dimensional vector that coordinates are a solution of the equation $1\alpha_1+2\alpha_2+\ldots+m\alpha_m=m$.

I. Preliminary step.

a) Compute $\Omega(\vec{\alpha})$, $N_m(\vec{\alpha})$, $\delta_m(\vec{\alpha})=\Omega(\vec{\alpha})/N_m(\vec{\alpha})$ for all solutions of the equation. $1\alpha_1+2\alpha_2+\ldots+m\alpha_m=m$.

b) On the set of vectors $\vec{\alpha}$ we define the order relation. We'll suppose that $\vec{\alpha}\geq\vec{\alpha}'$ if $\delta_m(\vec{\alpha})\geq\delta_m(\vec{\alpha}')$ and $\vec{\alpha}_1\geq\vec{\alpha}_2\geq\vec{\alpha}_3\geq\ldots$ if. $\vec{\alpha}_r\geq\vec{\alpha}_t$ for $r\leq t$.

II.

Suppose $\Lambda_1^{(1)}=\varnothing$, $r=1$.

j-step of testing.

1. Randomly from the set $B(\vec{\alpha}_r)\backslash \Lambda_r^{(j)}$ choose a permutation $s_r^{(j)}$;
2. Suppose $s_0=s_r^{(j)}$;
3. $\Lambda_r^{(j+1)}=\Lambda_r^{(j)}\cup s_r^{(j)}$;
4. For the state $(0, 0, s_0)$ compute $L=3/2m$ keystreams $z_1^*,\ldots,z_L^*$;
5. If $z_k=z_k^*$ for all $k=\overline{1,L}$, then we find true the initial state. If there exist $k\in\{\overline{1,L}\}$ such that $z_k\neq z_k^*$, then for $\Lambda_r^{(j+1)}=B(\vec{\alpha}_r)$ suppose $r=r+1$, $\Lambda_r^{(j+1)}=\varnothing$ and go to step II; for $B(\vec{\alpha}_r)\backslash \Lambda_r^{(j+1)}\neq\varnothing$ go to step 1.

6. Acknowledgment

References

[1] Golic, J. D, Linear Statistical Weakness of Alleged RC4 Keystream Generator. Advances in Cryptology -- EUROCRYPT '97.

[2] Fluhrer S.R., McGrew D. A. "Statistical analysis of the alleged RC4 keystream generator", Proceeding of FSE'2000, Springer-Verlag.

[3] Mantin I. Shamir A. "A practical attack on broadcast RC4", Proceeding of FSE'2001, Springer-Verlag.

[4] Mister S., Tavares S., "Cryptanalysis of RC4-like ciphers", Proceeding of SAC'98, Springer-Verlag.

[5] Knudsen L., Meier W., Preneel B., Rijmen V., Verdoolaege S, "Analysis method for (alleged) RC4", Proceeding of ASIACRYPT'98, Springer-Verlag.

[6] Grosul A.L., Wallach D.S. "A related key cryptanalysis of RC4", 2000, to appear.

[7] Pudovkina M. "Short cycles of the alleged RC4 keystream generator ", 3nd International Workshop on Computer Science and Information Technologies, CSIT'2001, YFA, 2001.

[8] M. Pudovkina "Statistical weakness in the alleged RC4 keystream generator", 4 International Workshop on Computer Science and Information Technologies, CSIT'2002, 2002.

[9] Mironov I., (Not so) Random shuffles of RC4, Advances in Cryptology --CRYPTO'2002, Springer-Verlag.

[10] Колчин В.Ф., Случайные графы, М: ФИЗМАТЛИТ, 2000. (in Russian)

Fig.1. $P\{\rho(k)\in B(m-d,0,\ldots0,1,0,\ldots0)\}= \Omega(m-d,0,\ldots0,1,0,\ldots0)\frac{1}{m^m}$ (the solid line) and $P\{s\in B(m-d,0,\ldots0, 1,0, \ldots0)\}$ (the dashed line) if s is uniformly random for m=64.