UDC 519.248; 003.26

M.T. Jenaliyev[1], A.T. Nurtazin[2], Z.G. Khisamiev[2]

[1]*Institute of Mathematics and Mathematical Modeling CS MES RK, Almaty, Kazakhstan;*
[2]*Institute of Information and Computing Technologies CS MES RK, Almaty, Kazakhstan*
(E-mail: muvasharkhan@gmail.com)

Estimation of randomity of generators of keys of the crypto system

The aim of the study is to develop an algorithm for constructing an integer random variable representing random components of the keys of the RSA cryptosystem, and determining the table for the distribution of its probabilities. The constructed random variable can be used to estimate the hypothesis that a given generator generates its values with the same probability with which a random variable takes these values. A hypothesis testing algorithm is provided, which is accompanied by a calculation table for a uniformly distributed random variable that takes primes from a given half-open interval.

Keywords: RSA-cipher, keys, random variable, prime number, probability distribution.

Introduction

To ensure that the practical stability of the cryptosystem is consistent with its theoretical stability, it is necessary to fulfill all the requirements for using this cryptosystem. The most important requirement is the requirement of generating by generator keys with a frequency corresponding to the calculated probability. If the random elements and groups of interrelated random elements can be represented as the corresponding discrete random variables and their probability distribution can be determined, then it is possible to evaluate the hypothesis on the feasibility of requirements for the correctness of operation to this generator. In this paper, we propose an algorithm for constructing a random variable representing interrelated elements p, q, a of key of cipher RSA, where p, q are prime numbers, a is mutually prime with the numbers $(p-1)(q-1)$. The second part provides an algorithm for testing the hypothesis that this generator generates its values with a frequency that tends in probability to the probabilities of the corresponding values of a given random variable with unlimited increase in the number of output values (generation of keys by generator with a frequency which corresponds to the calculated probability). The algorithm is accompanied by a calculation table for verifying the above hypothesis for several generators, generating prime numbers from a given half-interval and for a uniformly distributed random variable, taking the prime values from this half-interval. Note that the uniform distribution of a random variable has not values when evaluating a hypothesis. Note also that the random variable constructed in this work for the RSA cipher is not uniformly distributed, while this requirement is the main one in the works [1, 2].

1 Algorithm for constructing a random variable

Let p, q be random prime numbers from the integer half-interval $\mathfrak{G} = (0, G]$, a is a random number, $1 \leqslant a < \varphi(pq)$, mutually prime number with $\varphi(pq) = (p-1)(q-1)$, i.e. $(a, \varphi(pq)) = 1$. Three random numbers p, q, a are random elements that generate the key $k(p, q, a)$ of the cipher RSA. These random numbers are dependent: the choice of p implies the elimination of p from the set of values for q; the choice of p, q determines the possible values and the corresponding probabilities for a. We construct a discrete random variable $\xi(p, q, a)$ such that the probability $\mathcal{P}(p, q, a) = \mathcal{P}(\xi(p, q, a))(*)$. Since in each key $k(p, q, a)$ of the cipher RSA the following inequality holds $p \neq q$, and also the following equality holds $k(p, q, a) = k(q, p, a)$, it suffices to define $\xi(p, q, a)$ with the condition (*) for $p < q$. Further $\xi(p, q, a)$ must be one-to-one because a partition on a probability space of triples (p, q, a) must correspond to a partition on the number axis and vice versa.

Proposition. Let $\xi(p, q, a) = pqa^2$ be prime numbers , where p, q and $p < q < G, a < \varphi(pq), (a, \varphi(p, q)) = 1$. Then the function ξ is one-to-one.

Proof. Let $(p, q, a) \neq (p_1, q_1, a_1)$. Case 1. Let $(p, q) = (p_1, q_1)$. Then $a \neq a_1$ and hence $pqa^2 \neq p_1 q_1 a_1^2$. Case 2.1. Let $(p, q) \neq (p_1, q_1), p \neq p_1, q \neq q_1$. Then $pq \neq p_1 q_1$ and let $pq < p_1 q_1$. We assume $pqa^2 = p_1 q_1 a_1^2$,

then $p_1q_1|a^2$, from here $p_1q_1|a$. But $a < pq < p_1q_1$ and hence $a = 0$, contradiction. Case 2.2. Let $(p,q) \neq (p_1,q_1)$, $p = p_1, q \neq q_1$. We assume $pqa^2 = p_1q_1a_1^2$, then $qa^2 = q_1a_1^2$, hence $q|a_1^2$ and $q^2|a_1^2$. Let k be such that $q^{2k}|a_1^2$, and $q^{2(k+1)} \nmid a_1^2$, then $q^{2k+1}|a^2$ и $q^{2(k+1)}|a^2$. From here $q^{2(k+1)}|a_1^2$. Contradiction. As the case 2.2, the case 2.3 is considered $(p,q) \neq (p_1,q_1), p \neq p_1, q = q_1$. Proposition is proved.

Description of the algorithm

Let π be the number of primes in the half-interval $\mathfrak{G}$. The algorithm consists of $\pi - 1$ stages, there are exactly so many primes $p \in \mathfrak{G}$, for which there exists $q \in \mathfrak{G}$ such that $p < q$. At each stage i the probability $\mathcal{P}(p_i)$ of choice for i-th prime number p_i is determined, it equals to $\frac{1}{\pi-1}$, and $\pi - i$ is number of sub-stages. That is on the i-th stage $\pi - i$ of sub-stages are fulfilled to determine the probabilities of the choice of numbers of the form p_iq, where $p_i < q$. For i-th prime number p_i there are exactly $\pi - i$ prime numbers q, for which $p_i < q$. Then, at each sub-stage of the stage i we define q and probabilities $\mathcal{P}(p_iq) = \frac{1}{(\pi-1)(\pi-i)}$ (see Fig. 1). Further, at each sub-stage of the stage i with a certain number p_iq, $\varphi^2(p_iq)$ steps are taken, that complete the calculation of probabilities $\mathcal{P}(\xi = p_iqa^2) = \frac{1}{(\pi-1)(\pi-i)\varphi^2(p_iq)}$, where $a \in M_{p_iq}$, $M_{p_iq} = \{a : 1 \leq a < p_iq, (a, \varphi(p_iq)) = 1\}$.

p	$\mathcal{P}(p)$	q	$\mathcal{P}(q)$	pq	$\mathcal{P}(pq)$	$\varphi(pq)$	$\varphi^2(pq)$	$\mathcal{P}(a^2),\ a \in M_{pq}$	$\mathcal{P}(pqa^2),\ a \in M_{pq}$
2	$\frac{1}{(\pi-1)}$	3	$\frac{1}{(\pi-1)}$	6	$\frac{1}{(\pi-1)^2}$	2	1	1	$\frac{1}{(\pi-1)^2}$
		...	...	...	...	...	...	...	...
		p_π	$\frac{1}{(\pi-1)}$	$2p_\pi$	$\frac{1}{(\pi-1)^2}$	$p_\pi - 1$	$\varphi(p_\pi - 1)$	$\frac{1}{\varphi(p_\pi-1)}$	$\frac{1}{\varphi(p_\pi-1)(\pi-1)^2}$
...	...	...	...	...	...	...	...	...	...
$p_{\pi-2}$	$\frac{1}{(\pi-1)}$	$p_{\pi-1}$	$\frac{1}{2}$	$p_{\pi-2}\cdot$ $\cdot p_{\pi-1}$	$\frac{1}{2(\pi-1)}$	$\varphi(p_{\pi-2}p_{\pi-1})$	$\varphi^2(p_{\pi-2}p_{\pi-1})$	$\frac{1}{\varphi^2(p_{\pi-2}p_{\pi-1})}$	$\frac{1}{\varphi^2(p_{\pi-2}p_{\pi-1})2(\pi-1)}$
		p_π	$\frac{1}{2}$	$p_{\pi-2}\cdot$ $\cdot p_\pi$	$\frac{1}{2(\pi-1)}$	$\varphi(p_{\pi-2}p_\pi)$	$\varphi^2(p_{\pi-2}p_\pi)$	$\frac{1}{\varphi^2(p_{\pi-2}p_\pi)}$	$\frac{1}{\varphi^2(p_{\pi-2}p_\pi)2(\pi-1)}$
$p_{\pi-1}$	$\frac{1}{(\pi-1)}$	p_π	1	$p_{\pi-1}\cdot$ $\cdot p_\pi$	$\frac{1}{(\pi-1)}$	$\varphi(p_{\pi-1}p_\pi)$	$\varphi^2(p_{\pi-1}p_\pi)$	$\frac{1}{\varphi^2(p_{\pi-1}p_\pi)}$	$\frac{1}{\varphi^2(p_{\pi-1}p_\pi)(\pi-1)}$

Figure 1. Calculation of the probability for distribution ξ

Thus, a discrete probability distribution for a random variable $\xi(p,q,a) = pqa^2$ will be determined. On the Figure 2 application of the algorithm for $G = 10$ is considered.

p	$\mathcal{P}(p)$	q	$\mathcal{P}(q)$	pq	$\mathcal{P}(pq)$	$\phi(pq)$	$a : a \in M_{pq}$	$a^2 : a \in M_{pq}$	$\mathcal{P}(a^2)$	$pqa^2 : a \in M_{pq}$	$\mathcal{P}(pqa^2)$
2	$\frac{1}{3}$	3	$\frac{1}{3}$	6	$\frac{1}{9}$	2	1	1	1	6	$\frac{1}{9}$
		5	$\frac{1}{3}$	10	$\frac{1}{9}$	4	1,3	1,9	$\frac{1}{2}$	$10 \cdot 1, 10 \cdot 9$ $10, 90$ (1.4)	$\frac{1}{18}$
		7	$\frac{1}{3}$	14	$\frac{1}{9}$	6	1,5	1,25	$\frac{1}{2}$	$14 \cdot 1, 14 \cdot 25$ $14, 350$ (1.5)	$\frac{1}{18}$
3	$\frac{1}{3}$	5	$\frac{1}{2}$	15	$\frac{1}{6}$	8	1,3,5,7	1,9,25,49	$\frac{1}{4}$	$15 \cdot 1, 15 \cdot 9, 15 \cdot 25, 15 \cdot 49$ 15,135,375,735 (1.6)	$\frac{1}{24}$
		7	$\frac{1}{2}$	21	$\frac{1}{6}$	12	1,5,7,11	1,25,49,121	$\frac{1}{4}$	$21 \cdot 1, 21 \cdot 25, 21 \cdot 49, 21 \cdot 121$ 21,525,1029,2541 (1.7)	$\frac{1}{24}$
5	$\frac{1}{3}$	7	1	35	$\frac{1}{3}$	24	1,5,7,11,13,17, 19,23	1,25,49,121, 169,289 361,529	$\frac{1}{8}$	$35 \cdot 1, 35 \cdot 25, 35 \cdot 49, 35 \cdot 121,$ 35·169,35·289,35·361,35·529 35,875,1715,4235,5915, (1.8) 10115,12635,18515	$\frac{1}{24}$

Figure 2. Calculation of the probability distribution $\xi, n = 10$

On the Figure 3 a discrete probability distribution function is given, the calculation of which is given on the Figure 2.

$$\begin{bmatrix} \frac{1}{9} & \frac{1}{18} & \frac{1}{18} & \frac{1}{24} & \frac{1}{24} & \frac{1}{24} & \frac{1}{18} & \frac{1}{24} & \frac{1}{18} & \frac{1}{24} & \frac{1}{24} & \frac{1}{24} & \frac{1}{24} & \frac{1}{24} & \frac{1}{24} & \frac{1}{24} & \frac{1}{24} & \frac{1}{24} & \frac{1}{24} & \frac{1}{24} \\ 6 & 10 & 14 & 15 & 21 & 35 & 90 & 135 & 350 & 375 & 525 & 735 & 875 & 1029 & 1715 & 2541 & 4235 & 5915 & 10115 & 12635 & 18515 \end{bmatrix}$$

Figure 3. Probability distributions $\xi, n = 10$

The first line in Figure 3 is the probabilities of values of a random variable $\xi(p,q,a) = pqa^2$, located in the second line.

2 Evaluation of the generator

Suppose that for a random discrete quantity ξ, whose values lie in the half-interval $\mathfrak{G} = (0, G]$, the probability of acceptance of each of its values is known. Let Ω be sample of size $\mathrm{Dim}\{\Omega\}$ of generator $\mathfrak{F}$ for numbers from domain of values for ξ.

Hypothesis $\mathbb{H}$: $\mathfrak{F}$ generates values with frequencies tending in probability to the probabilities of the corresponding values of the random variable ξ, with an unlimited increase in size $\mathrm{Dim}\{\Omega\}$ of sample Ω.

Checking algorithm $\mathbb{H}$.

1. Run random choice of size for sample $\mathrm{Dim}\{\Omega\}$, $100 \le \mathrm{Dim}\{\Omega\}$ and number of partitions r, $2 \le r \le 100$.

2. Run random partition $\mathfrak{G}_r$ of the half-interval $\mathfrak{G}$ by a random choice $r-2$ of points in $\mathfrak{G} : \mathfrak{G}_r = (\mathfrak{g}(1) = \mathfrak{o}, \mathfrak{g}(1), ..., \mathfrak{g}(r-1), \mathfrak{g}(r) = \mathfrak{G}$, where $g(1) < g(2) < ... < g(r-1) < g(r)$.

3. Create an array of frequencies $\mathrm{Freq}\{\Omega_r\} = (\mathrm{Freq}\{\Omega_r\}[1], ..., \mathrm{Freq}\{\Omega_r\}[r])$, where $\mathrm{Freq}\{\Omega_r\}[i]$ is the number of prime numbers in the sample Ω that fall into the half-interval $(g(i), g(i+1)]$, $i = 1, ..., r-1$.

4. Create an array of probabilities $\mathcal{P}\xi_r = (\mathcal{P}\xi_r[1], ..., \mathcal{P}\xi_r[r-1])$, where $\mathcal{P}\xi_r[i]$ is the probability of falling ξ into the half-interval $(g(i), g(i+1)]$, $i = 1, ..., r-1$.

5. Compute value χ^2 for the array $\mathrm{Freq}\{\Omega_r\} : \chi^2(\mathrm{Freq}\{\Omega_r\}) = \sum_{i=1}^{r-1} \dfrac{\mathrm{Freq}\{\Omega_r^2\}[i]}{\mathrm{Dim}\{\Omega\} \cdot \mathcal{P}\xi_r[i]}$.

6. Test the hypothesis $\mathbb{H}$. If for a random variable χ^2 with $r-2$ degrees of freedom the probability $\mathcal{P}_{\mathfrak{F}}(r) = \mathcal{P}(\chi^2 \ge \chi^2(\mathrm{Freq}\{\Omega_r\})$, then there is reason to reject the hypothesis $\mathbb{H}$. Since it is considered unlikely in one test to obtain an event whose probability is less than 5%. End of algorithm.

Note that the number r is recommended to be taken in such a way that one of the conditions: i) $\chi^2 \ge \chi^2(\mathrm{Freq}\{\Omega_r\} \ge 10$; ii) $\mathrm{Freq}\{\Omega[i]\} \ge 10, i = 1, ..., r-1$ is satisfied. The conditions (i), (ii) are necessary for the practical application of the χ^2-Pearson criterion [3]. Repeated application of this algorithm to evaluate this generator with similar data may strengthen or weaken the grounds for rejecting the hypothesis.

3 Examples of estimates of prime numbers generators

Let the uniformly distributed discrete random variable ξ takes as values all prime numbers in the half-interval $\mathfrak{G} = (0, G]$ with probability $\mathcal{P}(\xi = p) = \int\limits_2^G \frac{1}{x} dx$, (asymptotic formula for the number of primes in the interval $\mathfrak{G}$) [4]. Let Ω be sample of size $\mathrm{Dim}\{\Omega\}$ for generator $\mathfrak{F}$ of primes in the interval $\mathfrak{G}$. Let us check the following hypothesis $\mathbb{H}$. $\mathbb{H}$: $\mathfrak{F}$ generates every prime number $p \in \mathfrak{G}$ with probability $\mathcal{P}(\xi = p)$. Figure 4 shows the results of the implementation of the algorithm for estimating 5 generators of prime numbers:

G	r	Dim	A Type	δ	$\mathcal{P}_\xi(r)$	H
10^3	6	100	Generator of Maple 13: Random selection of prime numbers – Ω	< 2.343 > 1.649	$>70\%$ < 80	Do not reject
			$\Delta[i] = \begin{cases} \text{random prime numbers, } \textit{if i odd} \\ p_{i2}, \textit{ if i even} \end{cases}$	< 11.668 > 9.488	$< 5\%$ $> 2\%$	Rather Reject
10^4	6	100	Generator of Maple 13: Random selection of prime numbers – Ω	> 2.195 < 3.357	$>50\%$ < 70	Do not reject
			$\Delta[i] = \begin{cases} \text{random prime numbers, } \textit{if i odd} \\ p_{10 \cdot t/2}, i \textit{ even} \end{cases}$	> 18.465	$< 0.1\%$	Rather Reject
10^5	6	100	Generator of Maple 13: Random selection of prime numbers – Ω	> 2.195 < 3.357	$>50\%$ < 70	Do not reject
			$\Delta[i] = \begin{cases} \text{random prime numbers, } \textit{if i odd} \\ p_{10^2 \cdot t/2}, \textit{if i even} \end{cases}$	> 18.465	$< 0.1\%$	Rather Reject
10^6	7	100	Generator of Maple 13: Random selection of prime numbers – Ω	> 2.343 > 1.610	$>80\%$ < 90	Do not reject
			$\Delta[i] = \begin{cases} \text{random prime numbers, } \textit{if i odd} \\ p_{10^3 \cdot t/2}, \textit{if i even} \end{cases}$	> 20.517	$< 0.1\%$	Rather Reject

Figure 4. Test of hypothesis

Evaluation of the generators is performed as follows. For the generator Maple 13, which has generated a sample Ω , value $\chi^2(\text{Freq}\{\Omega_r\})$ falls into the next half-interval, the boundaries of which are determined from the table of values of the random variable χ^2, Figure 5: $1.649 < \chi^2(\text{Freq}\{\Omega_r\}) = 1,8597.. < 2.195, r = 6$, Figure 4. According to the table in [3] for the distribution function of a random variable χ^2 with $r - 2 = 4$ degrees of freedom, accidental hit of values χ^2 into the interval $[1.859.., \infty)$ is probably more than 80 % and less than in 90 % cases, Figure 4. Therefore, the hypothesis $\mathbb{H}$ is not rejected.

This study is supported by grants No. AP05130928, AP05132262.

References

1 Колесова Н.А. Оценка качества генераторов последовательностей случайных чисел / Н.А. Колесова // Вестник АГТУ. Сер. Управление, вычислительная техника и информатика. — 2011. — № 1. — С. 119–123.

2 Ажмухамедов И.М. Методика оценки качества последовательности случайных чисел / И.М. Ажмухамедов, Н.А. Колесова // Вестник АГТУ. Сер. Управление, вычислительная техника и информатика. — 2010. — № 2. — С. 141–163.

3 Крамер Г. Методы математической статистики / Г. Крамер. — М.: Мир, 1975. — 625 с.

4 Бухштаб А.А. Основы теории чисел / А.А. Бухштаб. — М.: Просвещение, 1966. — 385 с.

М.Т. Жиеналиев, А.Т. Нұртазин, З.Г. Хисамиев

Криптожүйе кілттері генераторларының кездейсок қасиетін бағалау

Осы зерттеудің мақсаты RSA криптожүйе кілттері құрылатын кездейсоқ элементтерін білдіретін бүтін мәнді кездейсоқ шаманы жасау және оның ықтималдықтарын үлестіру кестесін анықтау үшін алгоритмді құрастыру болып табылады. Құрастырылған кездейсоқ шамасының мәндерінің ықтималдықтары берілген генераторды өңдейтін, сол мәндерін ықтималдықтарымен бірдей болатынын болжамдайды. Авторлар алдын ала белгілген жартылай аралығындағы жай сандарды қабылдайтын, біркелкі үлестірілген кездейсок шама болған жағдайда, осы болжамды тексеруге арналған есептеу кестесін келтіреді.

Кілт сөздер: RSA-шифр, кілттер, кездейсоқ шама, жай сан, ықтималдық, ықтималдықтарды үлестіру.

М.Т. Дженалиев, А.Т. Нуртазин, З.Г. Хисамиев

Оценка свойства случайности генераторов ключей криптосистемы

Целью исследования является разработка алгоритма построения целочисленной случайной величины, представляющей случайные составляющие элементы ключей криптосистемы RSA, и определение таблицы распределения её вероятностей. Построенная случайная величина используется для оценки гипотезы о том, что данный генератор вырабатывает свои значения с такой же вероятностью, с какой случайная величина принимает эти значения. Приведен алгоритм проверки гипотезы, который сопровождается таблицей расчетов для равномерно распределенной случайной величины, принимающей простые числа из заданного полуинтервала.

Ключевые слова: RSA-шифр, ключи, случайная величина, простое число, вероятность, распределение вероятностей.

References

1 Kolesova, N.A. (2011). Otsenka kachestva heneratorov posledovatelnostei sluchainykh chisel [Evaluation of the quality of random number sequence generators]. *Vestnik AHTU. Seriia Upravlenie, vychislitelnaia tekhnika i informatika — Bulletin of Astrakhan state technical university. Series Management, Computer Science and Informatics, 1,* 119–123 [in Russian].

2 Azhmukhamedov, I.M. & Kolesova, N.A. (2010). Metodika otsenki kachestva posledovatelnosti sluchainykh chisel [Method of estimating the quality of a sequence of random numbers]. *Vestnik AHTU. Seriia Upravlenie, vychislitelnaia tekhnika i informatika – Herald of Astrakhan state technical university. Series Management, Computer Science and Informatics, 2,* 141–163 [in Russian].

3 Kramer, G. (1975). *Metody matematicheskoi statistiki [Mathematical methods of statistics].* Moscow: Mir [in Russian].

4 Bukhshtab, A.A. (1966). *Osnovy teorii chisel [Fundamentals of number theory].* Moscow: Prosveshchenie [in Russian].