Deterministic Random Number Generator Algorithm for Cryptosystem Keys

Adi A. Maaita, Hamza A. A. Al_Sewadi

Abstract—One of the crucial parameters of digital cryptographic systems is the selection of the keys used and their distribution. The randomness of the keys has a strong impact on the system’s security strength being difficult to be predicted, guessed, reproduced, or discovered by a cryptanalyst. Therefore, adequate key randomness generation is still sought for the benefit of stronger cryptosystems. This paper suggests an algorithm designed to generate and test pseudo-random number sequences intended for cryptographic applications. This algorithm is based on mathematically manipulating a publically agreed upon information between sender and receiver over a public channel. This information is used as a seed for performing some mathematical functions in order to generate a sequence of pseudorandom numbers that will be used for encryption/decryption purposes. This manipulation involves permutations and substitutions that fulfill Shannon’s principle of “confusion and diffusion”. ASCII code characters were utilized in the generation process instead of using bit strings initially, which adds more flexibility in testing different seed values. Finally, the obtained results would indicate sound difficulty of guessing keys by attackers.

Keywords—Cryptosystems, Information Security agreement, Key distribution, Random numbers.

I. INTRODUCTION

INDEPENDENT, unpredictable and uniformly distributed numbers that cannot be reliably reproduced are referred to as random numbers [1]. They play a major part in practical implementation and strength of most cryptographic systems. They may be used as keys for symmetric crypto-systems, public key parameters, session keys, etc. [2]. Failure of obtaining strong keys definitely will end up with data security compromise. Therefore, strong random number generators that exhibit high statistical quality and can withstand cryptanalysis efforts are keenly sought. Such strong random number generators constitute an important building block in the design and testing of high quality-crypto-systems [3].

Generally random numbers can be truly random TRN, pseudo-random PRN or quasi-random QRN. Truly random numbers are unpredictable. Their generation stems from random physical or natural phenomena, such as radioactive decay, amplified noise generated by a resistor or a semiconductor diode, fed to a comparator or Schmitt trigger and then the output is sampled to get a series of bits which are statistically independent or random [4].

Pseudo-random numbers generators, also known as deterministic random bit generators, are computer programs that generate a sequence of numbers whose properties approximate the properties of sequences of random numbers [5], [6]. These sequences are not truly random as they are completely determined by a relatively small set of initial values called seeds; however, they are important in practice for their speed and reproducibility in number generation.

Quasi-random numbers are sequences in arbitrary dimensions which progressively cover a d-dimensional space with a set of points that are uniformly distributed. They are also known as low-discrepancy sequences [7]. The quasi-random sequence generators use an interface that is similar to the interface for random number generators, except that seeding is not required as each generator produces a single sequence.

Recently, a new type of generators which are called Lagged Fibonacci pseudo-random number generators [8], [9] have become increasingly popular generators for serial as well as scalable parallel machines. They are proved to be easy to implement, cheap to compute and they are performing reasonably well on standard statistical tests especially when the lag is sufficiently high.

After the brief definitions in Section I, related works will be summarized in Section II. Section III defines the important randomness tests that will be executed to examine the randomness of the generated keys. Section IV explains the proposed Pseudo-random generator scheme; Section V includes the implementation of the proposed scheme and the results of the randomness tests. Finally Section VI concludes the paper.

II. RELATED WORK

Random number generators may be classified into Integer generators, sequence generators, integer set generators, Gaussian generators, decimal fraction generators or row random byte generators depending if they generate integers, integer sequence, set of random integers, integers that fits normal distribution or numbers in the 0 and 1 range with configurable decimal places, respectively. Each of the mentioned types is useful for many cryptographic purposes [10]. Splitable pseudorandom number generators (PRNGs) were very useful for structuring purely functional programs that deal with randomness, because they allow different parts of the program to independently generate random values, thus avoiding random seed threading through the whole program [11].

The available pseudorandom number generators (PRNGs)
are either secure but slow, or fast but insecure [12]. Besides, they are either not efficient enough, have inherent flaws, or lack formal arguments for their randomness. Claessen and Palka [12] provided proofs in order to show strong guarantees of randomness under assumptions commonly made in cryptography.

Mixing secure and fast PRNGs in order to benefit from their respective qualities was sought recently, for example [14] proposed chaotic dynamical systems which appear to be good candidates to achieve this mixture for optimization by topological chaos and chaotic iterations for hash functions, [13]. PRNGs based on chaotic iterations were suggested also for watermarking application [14] and [15].

Chaotic systems have many advantages as unpredictability or disorder-like, which are required in building complex sequences, [16], [17]. This is why chaos has been applied to secure optical communications as suggested by [18]. However, chaotic systems of real-number or infinite bit representation realized in finite computing precision have the problems of non-ideal distribution and short cycle length. Hence infinite space of integers was considered lately leading to the proposition of using chaotic iterations (CIs) techniques. Such proposal has to a new family of statistically perfect and fast PRNGs, [19], where a new version of this family has been proposed. It uses decimation strategies that lead to improvements in both random number generation speed and statistical qualities. Other interesting PRNGs used a new Iteration Function System (IFS) that measures the sensitivity of the IFS to certain initial values in order to generate chaotic random numbers, [20].

PRNGs were also suggested based on iterative implementation of one-way functions utilizing a randomly selected start value with a key, [21]. Both the start value and the key for subsequent iterations were selected from the already generated random number in the previous iteration.

Another interesting PRNG method was suggested by [22]. It introduced a dynamic system to produce an interesting hierarchy of random numbers based on the review of random numbers characteristics and chaotic functions theory. The authors had carried out certain statistical tests on a series of numbers obtained from the introduced hierarchy.

Orue et al. [23] suggested cryptographically secure PRNGs that are based on the combination of the sequences generated by three coupled Lagged Fibonacci generators, mutually perturbed. The mutual perturbation method consists of the bitwise XOR cross-addition of the output of each generator with the right-shifted output of the nearby generator. The proposed generator has better entropy and much longer repetition period than the conventional Lagged Fibonacci Generator.

This paper proposes a pseudo-random number generation scheme that is based on Shannon’s concept of confusion and diffusion. The generated random numbers are to be used for cryptographic application. It suggests an in house scenario process that implements a one-time-pad key for secure communication. The scheme generates continuous strings of random bit sequences to be used as one time keys progressively for the subsequent messages.

### III. KEY RANDOMNESS TEST

The generated pseudorandom binary sequences can be tested for randomness by some of the statistical tests outlined by NIST [7]. These tests will focus on a variety of different types of non-randomness. The selected tests here include Frequency (Monobit) test, Frequency test within a Block, the Runs test, and the test for the Longest-Run-of-Ones in a Block and will be summarized below. For any of these tests the P-value is calculated and compared against the level of significance α (whose value is commonly set to about 0.01 for cryptographic applications). α is defined as the probability that the generated number is not random when it is really random and P-value is the probability that a perfect random number generator would have produced a sequence less random than the sequence that was tested, given the kind of non-randomness assessed by the test [7]. The criteria is if P-value=>α, the sequence appears to be random but if P-value<α, the sequence appears to be non-random.

#### A. Frequency (Monobit) Test:

It tests the proportion of zeroes and ones for the entire sequence. The purpose of this test is to determine whether the number of ones and zeros in a sequence are approximately the same as would be expected for a truly random sequence.

For the purpose of testing the randomness of a number with a bit string length of n bits, such that \( x = b_1, b_2, \ldots, b_n \), an observed value \( S_{\text{obs}} \) is used as a test statistic which is defined by (1).

\[
S_{\text{obs}} = \frac{S_n}{\sqrt{n}}
\]  
(1)

where \( S_n \) is the sum of all string bits after converting zeros and ones to -1 and +1, respectively.

The P-value for this test is calculated by (2).

\[
P\text{-value} = \text{erfc}\left(\frac{S_{\text{obs}}}{\sqrt{n}}\right)
\]  
(2)

where \( \text{erfc} \) is the complementary error function [7]

#### B. Frequency Test within a Block test:

It tests the proportion of 1’s within M-bit blocks. The purpose of this test is to determine whether the frequency of 1’s in an M-bit blocks is approximately M/2, as would be expected under an assumption of randomness. The P-value is calculated by (3).

\[
P\text{-value} = igame\left(\frac{N}{2}, \chi^2(\text{obs})/2\right)
\]  
(3)

where \( igame \) is the incomplete gamma function, \( N \) is the number of M-bit blocks to be tested, and \( \chi^2(\text{obs}) \) is the chi function of the observed proportion of 1’s within a given M-bit block given by (4).
\[ \chi^2(\text{obs}) = 4M \sum_{i=1}^{n} \left( \frac{e_i - M + 1}{M} \right) - \frac{1}{2} \]  

(4)

C. The Runs Test:

It tests the total number uninterrupted sequence of identical bits, i.e. whether the number of runs of 1’s and 0’s of various lengths is as expected for a random sequence. This test indicates the speed of 1’s and 0’s whether it is too fast or too slow. The \( P \)-value is calculated by (5):

\[ P - \text{value} = \text{erfc} \left( \frac{V_r(\text{obs}) - 2n\pi(1 - \pi)}{1\sqrt{2n\pi(1 - \pi)}} \right) \]

(5)

Where \( V_r(\text{obs}) \) is the total number of run across \( n \) and \( \pi \) is the pre-test proportion in the input sequence given by (6):

\[ \pi = \frac{\sum e_j}{n} \]

(6)

D. Longest-Run-of-Ones in a Block Test

It tests the longest run of 1’s within M-bit blocks, and show whether it is consist with the length of the longest run of 1’s that would be expected in a random sequence. The \( P \)-value is calculated by (7) and \( \chi^2(\text{obs}) \) is a measure of matching between observed longest run length within M-bit blocks with the expected longest length within M-bit blocks, given by (8):

\[ P\text{-value} = \frac{K}{2} \frac{\chi^2(\text{obs})}{2} \]

(7)

\[ \chi^2(\text{obs}) = \sum_{i=1}^{v} \left( \frac{v_i - N\pi_i^2}{N\pi_i} \right) \]

(8)

Where \( v_i \) is the frequencies of the longest runs of 1’s in each block categorized for \( i=0 \) to \( K \) while the values of \( K \) and \( N \) are determined by the value of \( M \) in accordance with the pre-set Table I.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>PRE-SET VALUES FOR M, K AND N WITHIN THE NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum key length</td>
<td>M</td>
</tr>
<tr>
<td>128</td>
<td>8</td>
</tr>
<tr>
<td>6272</td>
<td>128</td>
</tr>
<tr>
<td>750 000</td>
<td>10000</td>
</tr>
</tbody>
</table>

IV. PSEUDORANDOM NUMBER GENERATION METHODOLOGY

This work proposes a scheme for pseudorandom number generator (PRNG) that combines bitwise logical operation and bits manipulation in order to fulfill the confusion and diffusion principle. It starts with a randomly selected input key (seed) that consists of any combination of letters (lower or upper case) and numbers (i.e. a b c ... z, A B C ... Z, _ , 0 1 2 ... 9). This seed can be exchanged between sender and receiver publicly but would even better if it is exchanged secretly. They are replaced by their ASCII code binary representation as they enter to the PRNG. The length of this key is decided by the cryptographic system that is going to be implemented, for example, the 64 bits key for DES needs 8 characters and the 128 bits key for AES needs 16 characters, and so on. The work flow diagram of suggested PRNG scheme is illustrated in Fig. 1.

Prior to the operation of the proposed PRNG, the characters of the supplied key are converted to ASCII codes representation, and then the following steps are performed according to the work flow of Fig. 1.

Step 1. The selected key characters are first replaced by their binary representations. Let the entered key consists of \( n \)-characters, then the length of this key will be \( 8n \) (i.e. \( n \)-bytes).

Step 2. Abitwise XOR operations are performed on the bit blocks of each two successive bytes, i.e. (1st XOR 2nd) replaces the 1st byte, (2nd XOR 3rd) replaces 2nd character, etc., until the last byte where it is XOR’ed with the first one, or (\( n^\text{th} \)XOR1st) replaces the \( n^\text{th} \)-byte, as given by (9):

\[ i^\text{th} \text{ XOR } (i+1)^\text{th} \mod n \text{ , for } i=1 \text{ to } n \]

(9)

Step 3. Successive bytes are exchanged with each other in pairs. However, if \( n \) is odd number, then the last byte is left unaltered.

Step 4. The resulting bit string of the previous step is divided into half, left and right each of \( 4n \) bits length. The resulting bit sequence of \( 8n \) can be taken as the first pseudorandom random key \( K_1 \).

Step 5. The generated key in step 4 can be fed back as an input to step 2 in order to generate next random key.

Step 6. In order to generate more keys, steps 2-5 can be repeated as many as required.

A computer program is written to perform these steps written in C# language and tabulate the results in excel sheet together with their randomness tests in order to be ready for use in any cryptographic system.

Security of the key agreement:

Mixing of bitwise Boolean operations (XOR), bitwise and
operations oriented (manipulation) serves to avoid purely algebraic attacks and the purely bit oriented attacks and prevents the mathematical behavior of the scheme from being shaped easily. They both contribute to a great mathematical complexity together with high computational efficiency. Moreover, only very efficient operations are used, bitwise Boolean operations, and bits displacements, and they are both of easy implementation either by hardware or by software.

This method generates pseudo random numbers by selecting a certain piece of secure information. For the users involved in this system to generate these random numbers, they must agree upon this information in advance to enable them to generate their own random numbers which are then used in cryptographic systems.

The users also agree upon the predefined structure of this secret information or the key. The structure in this method consists of a certain number of digits constituting certain digital data that can either be represented as characters or as a bit sequence. The agreement upon the size (n) of this key, which depends practically on the cryptographic system that is intended to be used with, represents one of the constraints of this structure. And another constraint is the method used for splitting this piece of information for the various operations. The secure information is dealt with at byte level segments. Obviously each byte contains different digital contents.

The algorithm generates the first pseudorandom key at the end of the first run, however to generate more random numbers, further runs can be done. Each of the successive runs takes the previously generated random key as the input. Therefore, the difficulty and complexity of the generated keys will increase. The acceptance strength of the generated random numbers will be decided by the tests that will be performed on the results in the following section.

V. IMPLEMENTATION AND RESULTS

A Computer program is written in C# language for the proposed PRNG algorithm illustrated in Fig. 1. It is designed to accept a seed of any number of characters and generate as many random numbers as practically required with any length of bit sequence. It has been experimented for the generation and testing of random keys of 64, 128, and 512 bits lengths. An example demonstrating the implementation steps for the algorithm execution and elementary randomness tests are included in the next subsection, then four elaborate tests were carried out and their results were listed in next subsection.

A. PRNG Implementation

The following example can illustrate the implementation of the scheme for a 64 bits key sequence, i.e. the seed consists of 8 characters. Let us assume that step 1 produce the binary representation of the seed characters to be:

```
"10011001 00100110 11010011 10100010 00110111 01100001 00101001 10110001 10011001 00100110 11010011 10100010 00110111 01100001 00101001 10110001 10011001 00100110 11010011 10100010 00110111 01100001 00101001 10110001 10011001 00100110 11010011 10100010 00110111 01100001 00101001 10110001 10011001 00100110 11010011 10100010 00110111 01100001 00101001 10110001 10011001 00100110 11010011 10100010"
```

And finally step 4 will produce:

```
"00101111 01100001 00101001 10110001 10100010 00100110 11010011 10100010"
```

This is the generated pseudorandom number of the first round. Then more random numbers can be generated by repeating the PRNG algorithm using the generated number of a round as input to the algorithm for the next round.

To test the randomness of the generated random number, the following simple frequency test can be conducted. This test uses the chi-function formula given in (10):

\[
\chi^2 = \frac{(n_0 - n_1)^2}{n}
\]

where \(n_0\) and \(n_1\) are the numbers of 0’s and 1’s in the generated key sequence, respectively. Good sequence in the generated random number should have \(\chi^2\) values in the range \(0 < \chi^2 < 3.84\) [24].

Since \(\chi^2 = (29-35)/64 = 0.5625\) which is \(< 3.84\) for the above example, therefore it is considered as acceptable random number.

Counting the 0’s and 1’s in the generated key in the above example resulted into \(n_0 = 34\) and \(n_1 = 30\), therefore \(\chi^2 = (34-30)^2/64 = 0.25\), and since this value is \(< 3.84\), hence it is considered as acceptable random number.

Other test like frequency (monobit) can also be conducted, for the above example above.

The string length \(n = 64\) and the sum of all number string bits sequence after converting 0’s and 1’s to -1 and +1, respectively, \(S_n\) is determined as:

\[
S_n = \sum_{i=0}^{n-1} \text{sign}(x_i) = -1+1+1-1+1+1+1-1-1+1-1-1+1-1 = -4.
\]

Then from (1), the test statistic observed value

\[
S_{\text{obs}} = -6\sqrt{64}/\sqrt{64} = 0.5.
\]

Now applying (2), the \(P\)-value is calculated and it is here

\[
P\text{-value} = \text{erfc}(0.5/8) = 0.9296
\]

Then since 0.9296is > 0.01, the number is random otherwise it is not.

Another series test for the frequency of occurrence of sequences of two bits, i.e. 00, 01, 10, and 11 can be conducted. For the above example are \(n_{00} = 16\), \(n_{01} = 16\), \(n_{10} = 17\) and \(n_{11} = 10\) and the number of 1’s and 0’s in the key sequence \(n_1 = 30\) and \(n_0 = 34\).

Applying the chi-function for this test which is given by (11) [24]:

\[
\chi^2 = \frac{4}{n-1} \left( n_{00}^2 + n_{01}^2 + n_{10}^2 + n_{11}^2 - \frac{n_1^2 + n_0^2}{2} \right) + 1
\]

The criterion for good randomness sequence is \(\chi^2 \leq 5.99\).

Applying (11) results into

\[
\chi^2 = \frac{4}{64-1} (15^2 + 18^2 + 17^2 + 11^2) - \frac{2}{64} (30^2 + (-34)^2) + 1 = -10.96429 \leq 5.99
\]

Therefore, this key passes the randomness test.
B. PRNG Testing

The potential problems with deterministic generators are their failure in statistical pattern-detection tests. These problems include lack of distribution uniformity, correlation of successive values; output sequence has poor dimensional distribution, and shorter seed state. The random sequences generated by the proposed PRNG are tested for randomness by the four tests processes outlined in Section III. Although, the tests were done for 64, 128, 1nd 512 bits key sequences, however, it can be practically done for any key length.

Table II lists the calculated percentages of successfully generated random key sequences when the used seed for generating random numbers is only numeric’s, i.e. digits 0, 1, 2, … , 9. The tests included are explained four tests in section 3 namely; frequency (monobit) test, Frequency within a Block test, Runs test, and Longest-Run-of-Ones in a Block test. This table lists the results for three commonly required random key sequences; 64, 128, and 512 bits. However, Table III lists the percentages of successfully random keys generated using alphanumeric seed (i.e. both upper and lower case letters in addition to the numbers). It is also conducted for the same lengths and tests selected.

<table>
<thead>
<tr>
<th>Test</th>
<th>PERCENTAGE OF SUCCESSFUL RANDOMNESS (NUMERIC SEED ONLY)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=64</td>
<td>n=128</td>
</tr>
<tr>
<td>Frequency (monobit)</td>
<td>73%</td>
</tr>
<tr>
<td>Frequency within a block</td>
<td>84%</td>
</tr>
<tr>
<td>The Runs test</td>
<td>91%</td>
</tr>
<tr>
<td>Longest run of 1’s in a block</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test</th>
<th>PERCENTAGE OF SUCCESSFUL RANDOMNESS (ALPHANUMERIC SEEDS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=64</td>
<td>n=128</td>
</tr>
<tr>
<td>Frequency (monobit)</td>
<td>96%</td>
</tr>
<tr>
<td>Frequency within a block</td>
<td>98%</td>
</tr>
<tr>
<td>The Runs test</td>
<td>100%</td>
</tr>
<tr>
<td>Longest run of 1’s in a block</td>
<td>N/A</td>
</tr>
</tbody>
</table>

The results listed in this section have shown that the proposed PRNG algorithm effectively generates key sequences with a considerably acceptable randomness. The randomness tests showed that increasing the number of rounds to generate more keys enhances the efficiency of random key generation for cryptographic system applications. The scheme is characterized by its simple yet practical design as it does not have complicated and lengthy exponentiation processes. This leads to more efficient software and hardware implementations.

VI. CONCLUSION

An algorithm for computationally fast, cryptographically secure pseudorandom key generator has been proposed and described in this paper. It is based on mixing bitwise Boolean operations with bits manipulations and displacements for secret splitting. The implemented randomness tests have shown that the generated sequences were unpredictable and passed successfully stringent test suites. It obviously relies on the sender/receiver agreement protocol regarding the cryptosystem they are using, the key length, number of keys and way of their generation. Therefore the required random keys will be securely generated accordingly.

The algorithm was programmed in C# language on a 64 bits word length computer using only bitwise XOR and blocks exchanges. Hence excellent performance was achieved. Besides, the program produces as many number of random sequence keys as required, in which each random sequence is generated from a random one.

REFERENCES

[10] Random.org (Randomness and Integrity Service LTD),


Hamza A. A. Al_Sewadi is currently a professor at the Faculty of Information Technology, Isra University (Jordan). He got his B.Sc. degree in 1968 from Basrah University, Iraq, then M.Sc. and Ph.D. degrees in 1973 and 1977 respectively, from University of London(UK). He worked as associate professor at various universities such as Basrah University (Iraq), Zarqa University and Isra University (Jordan), visiting professor at University of Aizu (Japan). His research interests include Cryptography, Steganography, Information and Computer Network Security, Artificial Intelligence and Neural Networks.

Adi A. Maaita is currently an assistant professor at the Faculty of Information Technology, Isra University. He received his B.Sc. degree in 2002 from the University of Jordan (Jordan), his M.Sc in 2003 from the New York Institute of Technology (Jordan). Then received his Ph.D. from the University of Leicester (UK) in 2008. His research interests include Cryptography, Steganography, Information and Computer Network Security, Genetic Algorithms, Neural Networks, and Software Modeling.