ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

Statistical Tests in Cryptography: an Analysis on the Statistical Aspects of Determining the Randomness of Generators

Jyotirmoy Das

Research Scholar, Department of Computer Science, Assam Down Town University, Assam, India

ABSTRACT: Statistical tests can help determine the randomness of pseudo-random generators which are very much important in the field of cryptography. Cryptography has evolved as one of the core techniques in providing Data Security in Computer Systems and various communication methods. Exciting developments in the field of Cryptography have been observed during the last few decades. Statistical tests have become an efficient tool for the cryptographers to gain at least some information about the generator. For any cipher to be secured, it must have the randomness property. Also, in cryptography, the pseudo-random generators possess a close relationship with the stream ciphers. The use of various statistical tests in the field of Cryptography has been playing crucial roles from time to time, but there are instances where they fail too. Statistical tests like the Diehard, Statistical Test Suite by National Institute of Standards and Technology (NIST) have been very popular to test the different ciphers. In this paper we will introduce and describe the use of statistical theory related to the randomness property in the field of cryptography.

KEYWORDS: Statistical Tests, Pseudo-random generators, Cryptography, Ciphers, Randomness

I. INTRODUCTION

This paper is prepared in order to examine the recent trends in Statistical tests that are almost inevitable in the field of Cryptography. Cryptography is the science and study of secret writing in numbers or codes. Cryptography seeks to study and create systems for ciphering and to verify and authenticate the integrity of data. The plaintext (or cleartext) is converted into what is known as a 'ciphertext' (sometimes called a cryptogram). The process of converting the plaintext into ciphertext is called encryption. The reverse process of transforming the ciphertext back to the plaintext is called encryption. Both the processes, encryption and decryption, are controlled by a cryptographic key or keys. There are different types of cryptographic techniques that are applicable for different kinds of applications. The ones that we are interested here is the stream ciphers. The security of a stream cipher completely depends on the key stream where randomness plays a major role.

The main objectives of this paper are as follows :

i. To understand the statistical theory related to randomness testing.
ii. To describe the statistical tests used for testing the randomness.
iii. To study and analyze the various tests that has been formulated.

II. FREQUENCY TEST

Frequency Test is considered as the most basic test for testing the randomness property. The Frequency Test compares the Hamming weight of the sequence with the expected weight (n/2) of random sequence. The test statistic is :

$$\sum_{i=1}^{n} X_i$$

X_i having Bernoulli distribution, and X having Binomial distribution, therefore

This test is the most basic of all other tests. While performing, if a sequence does not pass this test, it has an increasing chance of failing in other tests too.

III. FREQUENCY TEST WITHIN A BLOCK

Frequency Test within a Block divides n-bit sequence into m-bit blocks and determines whether the frequency of ones in an m-bit block is approximately m/2, as would be expected in a random sequence. Here, n = 512, we take m = 8 and determine probabilities of four cases for blocks of size 8,(Table 1). In test we will count number of blocks having weight 3, 4, 5 or different and finally x^2 test is applied. More precisely let n be the length of sequence and m be the length of block (sequence is divided into $k = \left\lfloor \frac{n}{m} \right\rfloor$ blocks and the remaining bits are discarded), π_1 = number of blocks with weight 0,1,2,6,7 or 8, π_2 = number of blocks with weight 3.

Hamming weight of 8-bit sequence	Probability
0, 1, 2, 6, 7, 8	p_1 = 0:2890625
3	p_2 = 0:21875
4	p_3 = 0:2734375
5	p_4 = 0:21875

Table 1: Probabilities of weights of 8-bit sequence.

π_3 = number of blocks with weight 4, π_4 = number of blocks with weight 5. The test statistic is

$$X^2 = \sum_{i=1}^{4} \frac{(\pi_i - kp_i)^2}{kp_i}$$

and has x^2 distribution with 3 degrees of freedom.

IV. RUNS TEST

A run is an uninterrupted maximal sequence of identical bits. In the Run Test the number of runs in the sequence is compared with the expected number of runs in a random sequence. For example a sequence 0010111001 has 6 runs: 00, 1, 0, 111, 00, 1. The purpose of the Runs test is to observe whether ones and zeros are not changing too fast (like sequence 0101010101) or too slow (like 0000011111). Here, we present a slightly different modification of what is usually considered as the Runs test. The usual Runs test computes the distribution of runs for random sequence with weight w. Here we derive the probability distribution for a sequence which has the length n and k runs, with no requirement for the weight of the sequence. Denote a and b as different bits and the test statistic (number of runs) as X. For k = 1 we have two sequences: aa : : : a and bb : : : b, hence

$$\Pr[X = 1] = \frac{2}{2^n}$$

For k = 2 the first and last bit is determined: acc : : : cb where c it either a or b, and hence we will have n-1 possibilities, where we switch from a to b, and so

$$\Pr[X = 2] = 2\frac{n-1}{2^n}$$

Now for an arbitrary k we can imagine our sequence with partition after every bit: a|c|c| : : : |b. Again the first and the last bit are determined and there are $\binom{n-1}{k-1}$ possibilities how to choose the partition where we switch from ones to zeros or from zeros to ones, hence

$$\Pr[X = k] = 2\frac{\binom{n-1}{k-1}}{2^n}, for\ k = 1,2,\dots,n$$

V. PRIME NUMBER TEST

The Prime Number Test is a proposal of a new test, which is related to integers. It means that our bit sequence is transformed into integers a these integers are then tested. This test measures if a sequence produces prime numbers with frequency that is expected from a random sequence.

Let n be the length of sequence and m is the length of block (sequence is divided into $k = \left\lfloor \frac{n}{m} \right\rfloor$ blocks and the remaining bits are discarded). Each block is converted to number between 0 and 2^m - 1 and is determined whether is prime or not. We will create a new sequence from old one as follows: if the number is prime we will append 1 and 0 otherwise. Example 2.3. Let M = 0111011001001 be the tested sequence, hence n = 13. We choose m = 3 and convert M: 0112 = 3, 1012 = 5, 1002 = 4, 1002 = 4 and the last 1 is discarded. Then we convert 3,5,4,4 into 1100.

Let $= \frac{number\ of\ primes\ between\ 0\ and\ 2^m-1}{2^m}$, then we have just created a sequence with the Bernoulli distribution with parameter p, hence sum of these bits has the Binomial distribution B_i(p,k). For large values of m one can use the Prime Number Theorem which states that the number of primes less than or equal to 2^m - 1 can be approximated by $\frac{2^m-1}{log(2^m-1)} \approx \frac{2^m-1}{m}$. But there are two practical issues: 1) We have to estimate the error in the Prime Number Theorem and 2) determining primality of large numbers can be done quickly only with probabilistic algorithms. With that in mind, we recommend to use this test only with a small values of m, where the primality of numbers less than 2^m can be done by the trivial division or with a table of prime numbers less than 2^m.

VI. CONCLUSION

In this paper, we have been dealing with statistical testing of cryptographic primitives which is an important part of cryptanalysis. Statistical tests play a significant role in evaluation of ciphers and hash functions security although they do not consider the inner structure of primitives. Failing in the statistical tests could be a sign of a poor inner structure or of dependencies inside the cipher, which should be a motivation for harder cryptanalysis and modification of the cipher.

ACKNOWLEDGEMENT

I would like to acknowledge the Department of Computer Science, Assam down town University for encouraging me in preparing this research paper.

REFERENCES

1. Petr Niznansky, ‘Probability testing and application of statistical tests in cryptography’
2. Juan Soto,’Statistical Testing of Random Number Generators’, National Institute of Standards & Technology
3. Song-Ju Kim, Ken Umeno, and Akio Hasegawa. Corrections of thenist statistical test suite for randomness. Cryptology ePrint Archive, Report 2004/018, 2004.
4. Alfredo Rizzi, ‘Statistical Methods for Cryptography’
5. Jonathan Katz and Yehuda Lindell,’ Introduction to Modern Cryptography’
6. Christof Paar · Jan Pelzl, ‘Understanding Cryptography’
7. Lawrence E. Bassham, III, Andrew L. Rukhin, Juan Soto, James R. Nechvatal, Miles E. Smid, Elaine B. Barker, Stefan D. Leigh, Mark
8. Levenson, Mark Vangel, David L. Banks, Nathanael Alan Heckert,James F. Dray, and San Vo. Sp 800-22 rev. 1a. a statistical test suite for random and pseudorandom number generators for cryptographic applications. Technical report, Gaithersburg, MD, United States, 2010.

BIOGRAPHY

Jyotirmoy Das is a Research Scholar in the Department of Computer Science, Assam down town University, India. He received his Master of Computer Application (MCA) degree in 2012 from NERIM under Dibrugarh University, India.