# Randomness of Binary Sequences with Long Period by Combining m-Sequence and Knuth's Quadratic Congruential Sequence

Kohei KAWASE[†], Takeru MIYAZAKI[†], Shunsuke ARAKI[‡],
Satoshi UEHARA[†] and Yasuyuki NOGAMI[††]
[†]The University of Kitakyushu, Fukuoka, Japan
[‡]Kyushu Institute of Technology, Japan
[††]Okayama University, Japan

***Abstract*—Cryptography is one of the most important elements on the information security. In this paper, we propose a construction method of a long-period binary sequence with good randomness which plays an important role in cryptography. We generate the sequence by combining two types of binary sequences; one is the m-sequence and the other is a quadratic congruential sequence proposed by Knuth. It is well known that the m-sequence has good statistical properties, and Knuth's method can generate a long period sequence. However, both of these sequences cannot pass the almost NIST statistical tests by stand-alone. We derive a binary sequence combining these two sequences. In addition; we also consider the reasons why the proposed sequence has good randomness properties.**

## I. Introduction

In recent years, with the progress of network technology, threats to information assets such as information leakage and unauthorized access are increasing. Information security is necessary for them. Pseudorandom numbers becomes one of the import factors for an information security technology. Generally, a long-period and unpredictable sequence is one of the most important elements to construct a pseudorandom number generator. Hence, we focus on the quadratic congruential (QC) sequence with long period proposed by Knuth [1]. This sequence has good properties on "Linear complexity test" and "Rank test", but the other tests of it are not good enough. In this paper, we propose a new pseudorandom binary sequence by using a QC sequence and an m-sequence. The m-sequence has good properties of randomness, but the linear complexity is too smaller than the length of its period. Then, we improve their randomness properties by combining these two sequences and evaluate them by NIST statistical random test. We also consider the reason to improve the properties of our method.

## II. Preparations

In this section, we introduce the two sequences, the QC sequence and the m-sequence.

### A. The Quadratic Congruential Sequence

The QC sequence is a pseudorandom number sequence obtained by the method proposed by Knuth [1]. We generate a QC sequence with the following recurrence formula [1],

$$X_{n+1} = (dX_n^2 + aX_n + c) \pmod{N_q} \qquad (n \geq 0) \tag{1}$$

for $0 \leq d, a, c, X_0 < N_q$, where this sequence has the maximum period $N_q$ in case that Eq. (1) satisfies the following conditions.

i) $\gcd(N_q, c) = 1$,

ii) $d$ and $a-1$ are both multiples of $p$, for all odd primes $p$ dividing $N_q$,

iii) $d$ is even, and $d \equiv a-1 \pmod 4$, if $N_q$ is a multiple of 4,

iv) $d$ is even, and $d \equiv a-1 \pmod 2$, if $N_q$ is a multiple of 2, and

v) $d \not\equiv 3c \pmod 9$, if $N_q$ is a multiple of 9.

Depending on the value of the parameter constituting the recurrence formula, the QC sequence may have perfect regularity in the vicinity of the least significant bit. However, they have no regularity depending on the selection of parameters, sometimes the randomness is not impaired. Furthermore, parameters satisfying the condition of the maximum period depends on the value of $N_q$, and the selection of other parameters is restricted accordingly. Since the QC sequence has many choices of parameters, we can have many sequences with the fixed period $N_q$. In this paper, we generate a binary sequence. Let $l_0$ and $l$ be the bit length of $N_q$ and the arbitrary extracted bit length, respectively. The following inequality holds

$$l \leq l_0 = \lceil \log_2 N_q \rceil. \tag{2}$$

### B. The m-Sequence

The m-sequence is the one with the longest period among the sequences generated by the linear recurrence formula over the Galois field. If the degree of a primitive polynomial over $\mathrm{GF}(p)$ is equal to $n$, it generate an m-sequence with period $N_m = p^n - 1$.

## III. Combination of Binary Quadratic Congruential Sequence and Binary m-Sequence

In this section, we describe a binary sequence generated by combining a QC sequence and an m-sequence. Let $Q$ be a binary sequence by concatenating the lower $l$ bits of each element in a QC sequence. Then the length $l_q$ of $Q$ is equal to $l \times N_q$. Let $M$ be a binary m-sequence with the length $l_m = N_m$. We take a binary sequence $Q_M$ by bitwise exclusive OR of each element of these two sequence.

## IV. Evaluation of QC sequence

In this section, we try to evaluate the randomness properties of our proposal sequence $Q_M$ compared with the sequence $Q$. First, we evaluate statistical randomness of the both sequences $Q$ and $Q_M$ with the NIST SP800-22 test [2].

Table I shows the the numbers of failure tests in the NIST tests for these sequences, where $N_q = 31^5$ and $N_m = 2^7 - 1$.

TABLE I
The Numbers of Failures in NIST Statistical Tests of $Q$ and $Q_M$.

| Failed number of sequence | $l = 25$ | $l = 24$ | $l = 23$ |
|---|---|---|---|
| Sequence $Q$ | 155 | 143 | 96 |
| Sequence $Q_M$ | 0 | 0 | 0 |

In the case of the periodic sequence $Q$, the items "Rank test" and "Linear complexity test" are only passed in the NIST statistical tests, and the other items are failed. We know that the randomness of $Q$ is not good from these results. Since $N_q$ is denoted as 25 bits, there exists some biases in the upper bits. As $l$ shrinks from 25 to 23, these biases are reduced and the balance of the runs are good, but they are not good enough.

On the contrary, in case of the sequence $Q_M$, the all of the NIST tests are passed. This means that the randomness of the sequence $Q_M$ is pretty good. Also, when $N_m$ is a prime number, improvement of randomness is good. Moreover, we evaluate some parameters with $N_q$ as a power of an odd prime. Even if $N_m$ is a different value, it shows results similar to the results of this experiment.

Next, we investigate in the change of statistical properties due to exclusive OR operation into the QC sequence and m-sequence. Concretely, we evaluate a frequency distribution for all 7-tuple bit patterns in these sequences. Since the 7-tuple bit patterns are constructed from '0000000' to '1111111', the number of the patterns becomes $2^7 = 128$. When we consider that a sequence is true random, the distribution becomes a binomial distribution. Then if the length of the sequence is enough, this distribution is close to a normal distribution. Therefore, if we assume that the distribution of each bit pattern obtained from the appearance frequency of the pseudorandom number bit string is close to the normal distribution, this pseudorandom number bit string shows the same properties as the true random number. In this example, the number of each bit pattern existing between the width of 0.1 times the standard deviation when it is regarded as a true random number and the normal distribution when a true random number is input is compared in a graph. Then, it is judged whether the distribution is close or not. Figure 1 shows its distribution.

The distribution of sequence $Q$ is far from the theoretical value. The result shows that sequence $Q$ has a bias of occurrence on the 7-tuple bit patterns.

On the other hand the distribution of sequence $Q_M$ is roughly close to the normal distribution. The distribution of sequence $Q_M$ is close to the theoretical value and $Q_M$ has a property likes to a true random number. Even if $N_m$ and $N_q$ is a different value, it shows results similar to the results of this experiment.

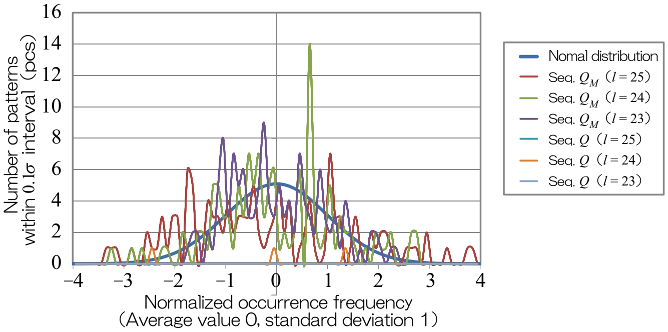

Fig. 1. Distribution of sequence bit patterns.

## V. Considerations

The sequence $Q_M$ is a binary sequence with high randomness, and we can easily generate it. We also confirm that the proposed sequence becomes long period, when $N_q$ is a power of an odd prime number and $N_m$ is a prime number. The sequence $Q_M$ is generated by an extremely simple generation method of the bitwise exclusive OR of the sequence $Q$ and the sequence $M$, and it is possible to generate at high speed. Then compare the speed with the binary sequence of the same period obtained in another way. Only when $N_m = 2^7 - 1$, the sequence $Q_M$ shows a significant improvement. Also, the period $N_q$ of the sequence $Q_M$ is expressed as follows.

$$N_{qm} = l_{qm} = \mathrm{lcm}(l_q, l_m) \tag{3}$$

When the value of $N_q$ is larger, the number of parameter candidates also increases. Therefore, we can expect many sequence numbers. If we want to make a bigger improvement, we guess that we need to meet some conditions. We state the condition as below.

- $N_q$ has few prime elements,
- $N_m$ is a prime.

## VI. Conclusion

In this paper, we proposed a new binary pseudorandom sequence combined a quadratic congruential sequence and an m-sequence by exclusive OR operation. Then we also evaluated its randomness. In the future, we will analyze the conditions theoretically and other conditions for improving randomness.

We will explain other conditions for improving randomness in oral presentation.

*Acknowledgement*

This work was partly supported by JSPS KAKENHI Grant-in-Aid for Scientific Research (A) 16H01723.

## References

[1] Donald. E. Knuth, "The Art of Computer Programming Volume 2 Seminumerical Algorithms Third Edition," Soc3.2.2, pp. 26-34, Apr. 1997.

[2] NIST, "A statistical test suite for random and pseudorandom number generators for cryptographic aplications," NIST Special Publication 800-22, 2001.