Uniform Binary Sequence Generated over Odd Characteristic Field

Yuta Kodera, Takeru Miyazaki, Takuya Kusaka, Ali Md. Arshad, Yasuyuki Nogami, and Satoshi Uehara

***Abstract*—This paper has focused on a pseudorandom binary sequence for the security applications. The secureness of random numbers, namely the randomness of the sequence, is evaluated by some specific properties such as period, its linear complexity and bits distribution. Especially from the viewpoint of the security, the latter two properties are widely used to evaluate the randomness of a sequence. The authors have proposed a pseudorandom sequence which achieves the maximum linear complexity in their previous works. However, the distribution of bits is not balanced in one period of an NTU sequence. Therefore, this paper aims to derive the uniformly distributed sequence without fewer effects on the other properties. The approach will help to achieve the uniform distribution with almost no calculation costs.**

***Index Terms*—ICT, secure database system, pseudorandom binary sequence, uniform bits distribution.**

I. Introduction

H The information and Communication Technology (ICT) has widely penetrated into our real life. For example, we use ICTs at home, on business, at hospitals and so on. Especially, hospitals keep many kinds of ordinary information such as a patient treatment information and surgery schedules as well as some private information regarding the patients. Therefore, the data must be access securely, and it should be kept secret.

However, evil users often try to obtain such valuable data. In fact, WannaCry [1], [2], one of ransomware, attacked 16 hospitals across the United Kingdom and encrypted all their data in computers, therefore the hospital authorities could not access their patients' valuable data. In this way, almost all the hospitals had lost their functionalities.

In this context, the authors concern about the importance of a secure database system which possesses the following three key points: 1) Confidentiality, 2) Integrity, and 3) Availability. Pseudorandom sequences are widely used in security protocols and cryptosystems; therefore, it is very crucial to ensure the randomness of a pseudorandom sequence.

The randomness of a pseudorandom sequence is often evaluated by the linear complexity and its distribution of bits, respectively. The authors' previous work on a pseudorandom sequence called NTU sequence [3], [4] has revealed the linear complexity of NTU sequence becomes maximum. In other words, NTU sequence shows the most difficult characteristics on predicting the next bit. Moreover, a theoretic proof for the period and the autocorrelation, and the cross-correlation have already given in [3], [4].

On the other hand, there still have problems that we cannot ignore. That is the efficiency of the calculation and the distribution of bits. The authors have proposed an efficient trace calculation [5] for the former one and it is sufficiently improved. The distribution of bits has been experimentally revealed [6], however, it is not uniformly distributed. Therefore, this paper proposes an approach for obtaining the balanced NTU sequence and observes its linear complexity and bits distribution.

Conventionally, NTU sequence has been generated by combining M-sequence [7] and Legendre sequence [8], [9] as follows: Let $f(x)$ and ω denote a primitive polynomial and a zero of. Then, ω becomes a generator of the extension field $\mathbb{F}_{p^m}$ and it can represent every non-zero element in $\mathbb{F}_{p^m}$ as ω^i, where $0 \leq i < p^m - 1$. For this polynomial sequence, we apply trace function defined as the sum of conjugates of ω^i. It outputs a prime field element which will be converted to $0,1$ or -1 by the Legendre symbol calculation. Finally, to binarize the sequence, $0,1$ is mapped to 0 and -1 is mapped to 1.

This mapping function cases the unbalanced sequence because of the following reasons. If a trace value is 0 or Quadratic Residue (QR) element in $\mathbb{F}_p$, then the sequence coefficient becomes 0. Otherwise, if it is Quadratic Non-Residue (QNR) element in $\mathbb{F}_p$, the coefficient becomes 1. The number of QR and QNR element are equal, therefore, 0 appears many times in one period of an NTU sequence.

In this paper, the authors propose to replace the trace value with QR and QNR element evenly when trace value is equivalent to 0. It is found that the approach leads us to obtain an almost uniformly distributed sequence like M-sequence. Moreover, in the case of M-sequence, it's linear complexity is known to be minimum. On the other hands, the proposed NTU sequence seems to achieve high linear complexity. The period also looks like same as the conventional NTU sequence.

Practically, security applications require quite long periodic sequence such as 256-bit. NTU sequence can generate such a large periodic sequence by using small p and large m. For example, let $p = 17$ and $m = 50$, then the period of NTU sequence becomes almost 256-bit. In this sense, the proposed sequence will provide secure and long period sequence.

Manuscript received December 9, 2017; revised March 10, 2018.

Y. Kodera, T. Kusaka, A. M. Arshad, and Y. Nogami are with Graduate School of Natural Science and Technology, Okayama University, Japan (e-mail: yuta.kodera@s.okayama-u.ac.jp, kusaka-t@okayama-u.ac.jp, arshad@s.okayama-u.ac.jp, yasuyuki.nogami@okayama-u.ac.jp).

T. Miyazaki and S. Uehara are with Faculty of Environmental Engineering The University of Kitakyushu, Japan (e-mail: miyazaki@kitakyu-u.ac.jp, uehara@kitakyu-u.ac.jp).

doi: 10.18178/ijiee.2018.8.1.685

II. Preliminaries

This section firstly introduces some useful notations and reviews mathematical fundamentals of a pseudorandom binary sequence. Then, the properties of the pseudorandom binary sequence are described.

A. Notations

This paper uses S_λ and s_i to show a pseudorandom binary sequence of the period λ and its i-th coefficient, respectively.

For a sequence S_λ, we consider the distribution of every bit pattern. b^n is used to indicate an arbitrary n-bit pattern, for instance, when $n = 2$, $b^2 \in \{00,01,10,11\}$. Finally, for a bit pattern b^n, $Z(b^n)$ and $D_{S_\lambda}(b^n)$ show the number of zeros in b^n and the frequency of b^n in S_λ, respectively.

B. Mathematical Fundamentals

1) Primitive polynomial:

Let $f(x)$ be a polynomial of degree m over prime field $\mathbb{F}_p$. If $f(x)$ is not divisible by any smaller degree polynomials, it is called irreducible polynomial. Let $f(x)$ be an irreducible polynomial over $\mathbb{F}_p$ and $t \in \mathbb{Z}$ be the smallest positive integer such that $f(x) \mid (x^t - 1)$. If $t = p^m - 1$, then $f(x)$ is especially called a primitive polynomial. In what follows, $f(x)$ denotes a primitive polynomial of degree m over $\mathbb{F}_p$. Moreover, a basis in $\mathbb{F}_{p^m}$ is considered with a polynomial basis in this paper.

2) Trace function:

The trace function $Tr_{\mathbb{F}_{p^m}|\mathbb{F}_p}(x)$ is defined as the sum of conjugates of $x \in \mathbb{F}_{p^m}$ with respect to $\mathbb{F}_p$. It is calculated by

$$Tr_{\mathbb{F}_{p^m}|\mathbb{F}_p}(x) = \sum_{i=0}^{m-1} x^{p^i}. \tag{1}$$

As an example, consider $\mathbb{F}_{7^2}$ with an irreducible polynomial $g(x) = x^2 + 1$ over $\mathbb{F}_7$. Let ω be a zero of $g(x)$ and then $\{1, \omega\}$ forms a basis in $\mathbb{F}_{7^2}$. It is noted that ω satisfies $\omega^2 = -1$ because $g(\omega) = 0$. Let us consider an example with an element $2 + 4\omega \in \mathbb{F}_{7^2}$. Firstly, its p-th power is obtained as follows:

$$\begin{aligned}(2 + 4\omega)^7 &= 2 + 4\omega^7 = 2 + 4(\omega^2)^3\omega \\ &= 2 + 4(-1)^3\omega = 2 + 3\omega.\end{aligned}$$

Then, the trace of $2 + 4\omega$ is given as:

$$\begin{aligned}\mathrm{Tr}_{\mathbb{F}_{p^m}|\mathbb{F}_p}(2 + 4\omega) &= (2 + 4\omega) + (2 + 4\omega)^p \\ &= (2 + 4\omega) + (2 + 3\omega) = 4.\end{aligned}$$

Readers need to be conscious that a trace value always belongs to $\mathbb{F}_p$ and the trace function has linearity property over $\mathbb{F}_p$ as follows:

For arbitrary elements $a, b \in \mathbb{F}_p$ and $X, Y \in \mathbb{F}_{p^m}$, trace function $\mathrm{Tr}_{\mathbb{F}_{p^m}|\mathbb{F}_p}(aX + bY)$ satisfies

$$\mathrm{Tr}_{\mathbb{F}_{p^m}|\mathbb{F}_p}(aX + bY) = a\mathrm{Tr}_{\mathbb{F}_{p^m}|\mathbb{F}_p}(X) + b\mathrm{Tr}_{\mathbb{F}_{p^m}|\mathbb{F}_p}(Y)$$

3) Legendre symbol:

Let a be an arbitrary element in $\mathbb{F}_p$, the Legendre symbol defined as below is used for checking whether a has a square root in $\mathbb{F}_p$ or not. When a has a square root in $\mathbb{F}_p$, then a is called Quadratic Residue (QR) element. Otherwise, it is called Quadratic Non-Residue (QNR) element.

$$\begin{aligned}\left(a/p\right)_2 &= a^{\frac{p-1}{2}} \bmod p \\ &= \begin{cases} 0 & \text{if } a = 0, \\ 1 & \text{else if } a \text{ is QR in } \mathbb{F}_p^*, \\ -1 & \text{otherwise } a \text{ is a QNR in } \mathbb{F}_p^*. \end{cases}\end{aligned} \tag{2}$$

It is noted that $\mathbb{F}_p^*$ denotes a multiplicative group with elements $\{1,2,3,\ldots,p-1\}$.

4) Mapping function:

A pseudorandom binary sequence considered below uses a mapping function defined as follows to binarize the sequence generated by the Legendre symbol.

$$M_2(x) = \begin{cases} 0 & \text{if } x = 0 \text{ or } 1 \\ 1 & \text{otherwise } x = p - 1 \end{cases} \tag{3}$$

where *x* is an output of the Legendre symbol calculation.

C. NTU Sequence

This section introduces how to generate a pseudorandom binary sequence called NTU sequence [3], [4]. In addition, some kinds of properties are described here. In the following sections, $f(x)$ is a primitive polynomial of degree m over $\mathbb{F}_p$ and S_λ and s_i denote an NTU sequence of period λ and its i-th coefficient, respectively.

1) Generating procedure:

Let ω be a zero of $f(x)$, then there are $p^m - 1$ kinds of elements in $\mathbb{F}_{p^m}$ and they are represented by the powers of ω. The *i*-th coefficient of an NTU sequence is generated by the following calculation.

$$\begin{aligned}S = \{s_i\}, s_i &= M_2\left(\left(\mathrm{Tr}_{\mathbb{F}_{p^m}|\mathbb{F}_p}(x)/p\right)_2\right). \end{aligned} \tag{4}$$

where $i = 0,1,2,\ldots$. Therefore, the seed value of an NTU sequence is said to be an arbitrary element in.$\mathbb{F}_{p^m}$

2) Properties of NTU sequence

There are various kinds of viewpoints to understand the characteristics of a pseudorandom sequence such as period, autocorrelation, cross-correlation, linear complexity and distribution of bits. Almost all the above properties have been theoretically shown in the previous works [3], [4]. In detail, the period λ of NTU sequence is given by

$$\lambda = \frac{2(p^m - 1)}{p - 1} \tag{5}$$

Especially, the measurements of the randomness are inseparable aspects of security applications. The linear complexity and distribution of bits are two of them and the linear complexity of NTU sequence is known as maximum.

However, the distribution of bits is ununiform [6] due to the mapping function Eq. (3). For example, when we focus on 3-bit patterns in the following NTU sequence S_{114} with $p = 7$ and $m = 3$, the distribution $D_{S_{114}}(b^3)$ is given as shown in Table I.

1,0,1,0,0,1,0,0,1,0,1,0,1,0,1,1,1,1,1,0,0,0,0,
1,0,1,0,0,0,0,0,0,1,0,1,1,0,0,0,1,0,0,0,0,0,1,
1,1,0,0,1,1,0,0,0,0,1,0,1,0,1,1,0,0,1,0,0,0,0,
0,1,0,0,0,0,0,0,0,0,1,1,0,1,0,0,1,1,1,1,1,1,0,1,0,
0,1,1,0,0,1,1,0,1,1,0,0,0,1,1,0,0,1,1,1,1,1,0.

TABLE I: THE APPEARANCE NUMBER OF N-BIT PATTERN WHEN $P = 7, M = 3$ AND $n = 1,2,3$

	$b^{(n)}$	$Z(b^{(n)})$	$D_{S_{57}}(b^{(n)})$
$n = 1$	0	1	65
	1	0	49
$n = 2$	00	2	37
	01	1	28
	10	1	28
	11	0	21
$n = 3$	000	3	21
	001	2	16
	010	2	16
	011	1	12
	100	2	16
	101	1	12
	110	1	12
	111	0	9

Table I shows that the number of b^n depends on the number of zeros in the bits. If the distribution of NTU sequence is improved, NTU sequence will be more suitable pseudorandom sequence for security applications. Therefore, this paper aims to improve the distribution of NTU sequence without losing the superior properties of NTU sequence as mentioned above.

III. PROPOSED METHOD

This section introduces a technique for obtaining the uniform distribution in the generating procedure of NTU sequence.

A. A Technique for the Uniform Distribution

As shown in Section II-C2, the number of 0s is larger than that of 1s in a period of NTU sequenceS_λ. This is happening, when $\mathrm{Tr}_{\mathbb{F}_{p^m}|\mathbb{F}_p}(\omega^i)$ is equivalent to 0 or QR element in $\mathbb{F}_p$, s_i becomes 0. Otherwise, $s_i = 1$. In detail, 0 and 1 appear $\left(\frac{p+1}{2} \cdot p^{m-1} - 1\right)$ times and $\left(\frac{p-1}{2} \cdot p^{m-1}\right)$ times inS_λ, respectively. In addition, the number of QR and QNR element in $\mathbb{F}_p$ are equal. Therefore, the main idea of proposition is just replacing a trace value by QR element and QNR element evenly when $\mathrm{Tr}_{\mathbb{F}_{p^m}|\mathbb{F}_p}(\omega^i) = 0$.

It is found that the above idea has been theoretically guaranteed to have the same number of 0 and 1 in one period of an NTU sequence. However, the replacement must be done randomly and efficiently. This paper proposes to use a coefficient of ω^i for the replacement because there are 2 advantages. Firstly, ω^i is already calculated, thus the replacement can be done easily. Secondly, coefficients of ω^i changes by the choice of a primitive polynomial.

In detail, let $\omega^i = \sum_{j=0}^{m-1} c_j \omega^j$. When $\mathrm{Tr}_{\mathbb{F}_{p^m}|\mathbb{F}_p}(\omega^i) = 0$, we refer the smallest degree non-zero coefficient in ω^i. Then, the coefficient is used as an input for the Legendre symbol calculation instead of $\mathrm{Tr}_{\mathbb{F}_{p^m}|\mathbb{F}_p}(\omega^i) = 0$. The flowchart of this procedure is drawn as Fig. 1, where c_j denotes j-th coefficient of ω^i

.

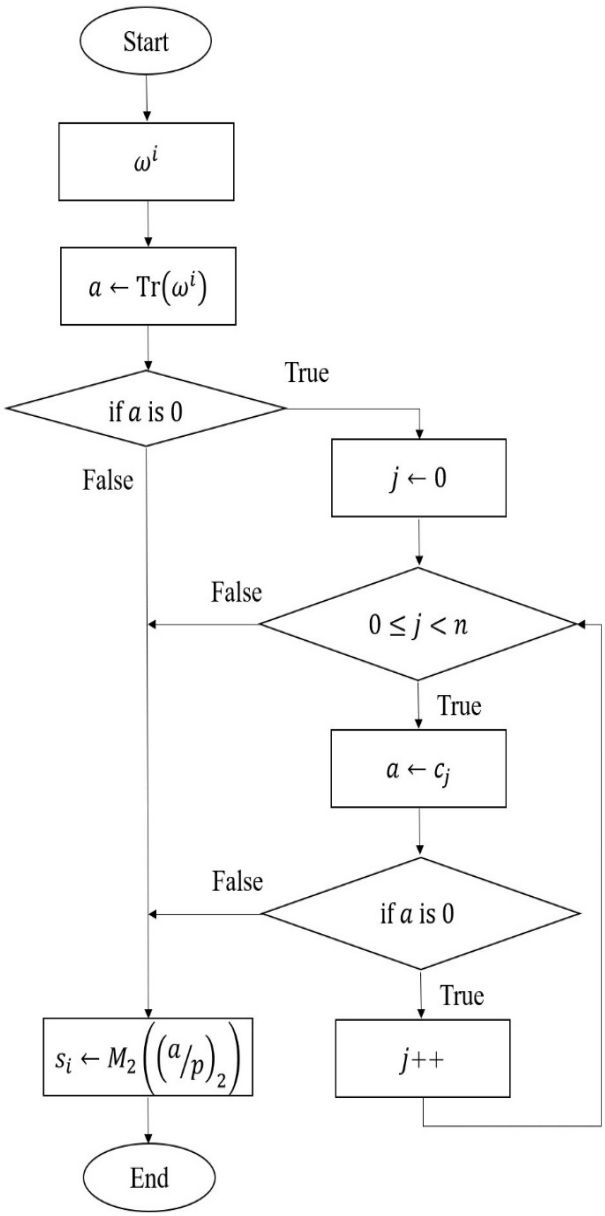


Fig. 1. The generating procedure of the proposed NTU sequence.

IV. EXPERIMENT AND CONSIDERATION

A small example of the proposed NTU sequence and its distribution are shown in this section. In addition, the properties of the proposed sequence are given here.

A. Experimental Result

Here, we show the bits distribution of the proposed NTU sequence and observe experimental results for the linear complexity profile. The result of the bits distribution with $p = 13, m = 7$ is shown in Table II and that of the linear complexity with $p = 7, m = 5$ is shown in Fig. 2.

TABLE II: THE NUMBER OF N-BIT PATTERNS IN ONE PERIOD OF THE PROPOSED NTU SEQUENCE WHEN $p = 13, m = 7$ AND $n = 1,2,3,4$

	$b^{(n)}$	$D_{S_\lambda}(b^{(n)})$
$n = 1$	0	5229043
	1	5229043
$n = 2$	00	2615710
	01	2613333
	10	2613333
	11	2615710
$n = 3$	000	1308495
	001	1307215
	010	1306118
	011	1307215
	100	1307215
	101	1306118
	110	1307215
	111	1308495
$n = 4$	0000	654557
	0001	653938
	0010	653347
	0011	653868
	0100	653347
	0101	652771
	0110	653277
	0111	653938
	1000	653938
	1001	653277
	1010	652771
	1011	653347
	1100	653868
	1101	653347
	1110	653938
	1111	654557
$n = 5$	00000	327403
	00001	327154
	00010	326858
	00011	327080
	00100	326829
	00101	326518
	00110	326769
	00111	327099
	01000	326839
	01001	326508
	01010	326253
	01011	326518
	01100	326788
	01101	326489
	01110	326784
	01111	327154
	10000	327154
	10001	327154
	10010	326784
	10011	326489
	10100	326788
	10101	326518
	10110	326508
	10111	326839
	11000	327099
	11001	326769
	11010	326518
	11011	326829
	11100	327080
	11101	326858
	11110	327154
	11111	327403

B. Consideration

It was found from Table II that the distribution of NTU sequence has been surely improved by the proposed method.

And if two bits are complemented on two, then $D_{S_\lambda}(b^n)$ seems to become the same in number. The linear complexity of the proposed sequence becomes the half of the conventional NTU sequence. This implies that the mapping function used in the conventional NTU sequence had enhanced the linear complexity. Moreover, according to [10], the linear complexity profile of true random numbers follows $n/2$, where n denotes the observation length. The linear complexity of the proposed sequence seems to be likely to follow a half of the period.

From the adequate observation, the period of the proposed NTU sequence seems to be same as the conventional NTU sequence. Namely, the period λ will be given as follows:

$$\lambda = \frac{2(p^m - 1)}{p - 1}$$

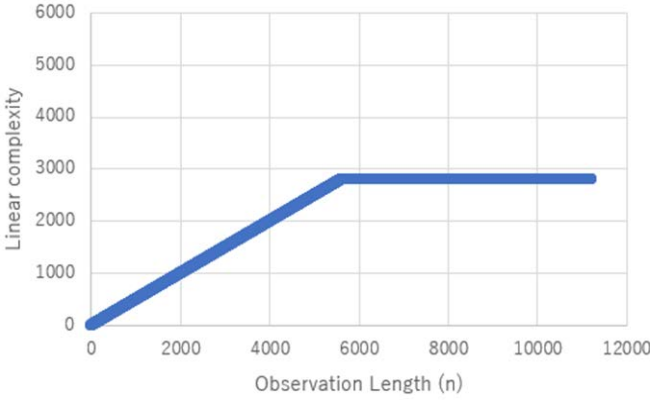


Fig. 2. The linear complexity profile of the proposed sequence.

V. CONCLUSIONS AND FUTURE WORK

The proposed approach overcomes the drawback of the conventional NTU sequence by introducing a special mapping when $\mathrm{Tr}_{\mathbb{F}_{p^m}|\mathbb{F}_p}(\omega^i) = 0$. This approach makes the uniform distribution of bits. Moreover, it seems to take over the superior properties of the conventional NTU sequence. In this sense, this proposed sequence is a prime candidate for the security applications.

As a future work, we would like to show the theoretic proof for each property such as the period, linear complexity and the distribution of bit patterns so that we can strongly recommend using it for the security applications.

REFERENCES

[1] S. K. Sahi, "A study of wannacry ransomware attack," *IJERCSE*, vol. 4, issue 9, 2017.
[2] S. Mohurle and M. Patil, "A brief study of Wannacry Threat: Ransomware Attack 2017," *International Journal of Advanced Research in Computer Science*, vol. 8, no. 5, 2017.
[3] Y. Nogami, K. Tada, and S. Uehara, "A geometric sequence binarized with legendre symbol over odd characteristic field and its properties," *IEICE Trans.*, vol. E97-A, no. 1, pp. 2336-2342, 2014.
[4] Y. Nogami, S. Uehara, K. Tsuchiya, N. Begum, H. Ino, and R. H. Morelos-Zaragoza, "A multi value sequence generated by power residue symbol and trace function over odd characteristic field," *IEICE Trans.*, vol. E99-A, no. 12, pp. 2226-2237, 2016.
[5] Y. Kodera, T. Kusaka, T. Miyazaki, Md. A. Khandaker, Y. Nogami, and S. Uehara, "An efficient implementation of trace calculation over finite field for a pseudorandom sequence," *The Fifth International Symposium on Computing and Networking*, 2017.
[6] Y. Kodera, T. Miyazaki, Md. A. Khandaker, Md. A. Ali, Y. Nogami, and S. Uehara, "Distribution of bit patterns on multi-value sequence over odd characteristics field," *IEEE 2017 ICCE-TW*, 2017.
[7] S. W. Golomb, "Shift register sequences," *Holden-Day*, San Francisco, 1967.
[8] N. Zierler, "Legendre sequence," *M.I.T. Lincoln Publications*, 1958.

[9] C. Ding, T. Helleseth, and W. Shan, "On the linear complexity of legendre sequences," *IEEE Trans. on Inform. Theory*, vol. 44, pp. 1276-1278, 1998.
[10] A. H. Chan and R. A. Games, "On the linear span of binary sequences obtained from *q*-ary m-sequences, *q* odd," *IEEE Transactions on Information Theory*, vol. 36, no. 3, pp. 548-552, 1990.

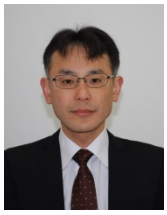
Yuta Kodera graduated from Okayama University in 2017. He is now a master's course student under the supervision of Professor Dr. Yasuyuki NOGAMI in the Graduate School of Natural Science and Technology, Okayama University, Japan.

His research interests are finite field theory and its applications such as Pseudo Random Number Generator and recent Public Key cryptographies. Currently, he is studying about pseudo random number generator and lattice-based cryptography.

Takeru Miyazaki received the B.E. and the M.E. degrees in information engineering from the Kyushu Institute of Technology in 1997 and 1999, and the Dr.Eng. degree of Engineering from the University of Kitakyushu in 2012, respectively.

From 2001, he was an engineering advisor at the University of Kitakyushu.

His research interests are pseudorandom number generator and cryptography.



Takuya Kusaka was born in 1970. He received the B.E. degree in electric engineering from Kobe University in 1994, and he received M.E. and Ph.D. degrees in information science from the Graduate School of Information Science, Nara Institute of Science and Technology in 1996 and 1999, respectively. In 2004, he joined Okayama University.

His current research interests include coding theory and information security.

Yasuyuki Nogami graduated from Shinshu University in 1994 and received the PhD degree in 1999 from Shinshu University. He is now a professor at Okayama University.

His main fields of research are finite field theory and its applications such as recent public key cryptographies. He is now studying about elliptic curve cryptography, pairing-based cryptography, Lattice-based cryptography, pseudo-random number generator, Advanced Encryption Standard, and homomorphic encryptions.

Recently, he is a member of security research group at Okayama university and particularly focusing on IoT security from the viewpoints of software and hardware implementations. He is a member of IEICE and IEEE.

Satoshi Uehara received the B.E. degree from Saga University and the M.E. degree from Kyushu University, and the Dr. Eng. degree in computer science and system engineering from Kyushu Institute of Technology, in 1987, 1989 and 1998, respectively. From 1989 to 2000, he was a research associate at Kyushu Institute of Technology. From 2000, he was with the Department of Information and Media Engineering, The University of Kitakyushu as an associate professor, and became a professor in 2009.

He is engaged in research on sequence design for cryptography and spread spectrum applications.

Ali Md. ARSHAD was born in 1986. He received the B.S. degree in computer science and engineering from Hajee Mohammad Danesh Science and Technology University (HSTU), Bangladesh in 2007. In 2009, he joined HSTU as a part-time teacher, and he worked as a lecturer from 2010 to 2013, he became an assistant professor in 2013. Currently, he is a master's course student at Okayama University, Japan.

His research interest includes information security, advanced encryption standard, pseudo-random binary sequence, elliptic curve cryptography, and homomorphic encryption.

He is a member of IEEE.