A Study of Pseudo Random Sequence Generated by Cascaded Trace and Legendre Symbol over Finite Field

September, 2019

Md. Arshad Ali

Graduate School of
Natural Science and Technology
(Doctor's Course)

Okayama University

Doctoral Thesis

A Study of Pseudo Random Sequence Generated by Cascaded Trace and Legendre Symbol over Finite Field

Author:
Md. Arshad Ali

Supervisor:
Yasuyuki Nogami

Co-supervisors:
Yoshitaka Toyota
Minoru Kuribayashi

A dissertation submitted to
Okayama University
in fulfillment of the requirements for the degree of
Doctor of Philosophy in Engineering
in the
Graduate School of Natural Science and Technology

August 20, 2019

To Whom It May Concern

We hereby certify that this is a typical copy of the original Doctoral dissertation of

Md. Arshad Ali

Thesis Title:

A Study of Pseudo Random Sequence Generated by Cascaded Trace and Legendre Symbol over Finite Field

---

Seal of Supervisor *Official Seal*

Professor Yasuyuki Nogami

Graduate School of
Natural Science and Technology

Declaration of Authorship

This dissertation and the work presented here for doctoral studies were conducted under the supervision of Professor Yasuyuki Nogami. I, Md. Arshad Ali, declare that this thesis titled, "A Study of Pseudo Random Sequence Generated by Cascaded Trace and Legendre Symbol over Finite Field" and the work presented in it are my own. I confirm that:

Signed: Md. Arshad Ali Student number: 51430703

Date: August 20, 2019

Abstract

Md. Arshad Ali

A Study of Pseudo Random Sequence Generated by Cascaded Trace and Legendre Symbol over Finite Field

In cryptography, randomness plays an important role in numerous applications due to the output of these applications must be seen by an adversary as a sequence of random values carrying secret information nevertheless revealing no clues regarding the precious information. It is required in fundamental tasks such as secret key generation, message encryption, masking the content of a certain protocol, internet gambling, initialization vector generation, computer simulation, radar ranging, and other areas. The security of these cryptographic algorithms and protocols is strongly related to the random number generators that exhibit high statistical quality. It is unrealistic to generate such random numbers in real life, instead of sequences of number or bits that appear to be random yet repeatable to employ them in real life applications. These sequences are so-called pseudo random sequences.

Since the importance of random numbers in cryptographic applications, extensive literature devoted to several pseudo random generators such as maximum length sequence (M-sequence), Legendre sequence, Sidelnikov sequence, gold sequence, Kasami sequence, blum blum shub (BBS), linear congruential generator (LCG), and mersenne twister (MT). The research trends on the generation of pseudo random sequence basically classified into two directions such as from the viewpoint of theoretical aspect (M-sequence and Legendre sequence) and their application point of view (MT and BBS). This dissertation concentrates on the theoretical aspect rather than the application of pseudo random sequences.

To determine the suitability of pseudo random sequences for security applications, the study of their properties is obligatory their period needs to be very long, low correlation, high linear complexity, and uniform distribution of bit patterns. An M-sequence is generated using linear feedback shift registers based on the Galois field theory. An M-sequence is well-known for its prominent maximum period, best periodic autocorrelation, and uniform distribution of bit patterns properties. On the other hand, the linear complexity of an M-sequence is minimum. Legendre sequence is generated by applying the Legendre symbol over the prime field and its linear complexity is known to be high, whereas its period is long.

The Legendre sequence and M-sequence are the base of the sequence research area considering their theoretical features. Their properties are already proven,

and many researchers attracted by those theoretic proof. The study of pseudo random sequence generation by combining the well-known M-sequence and Legendre sequence started with the NTU (Nogami-Tada-Uehara) sequence. Although the NTU sequence holds interesting features such as its long period and high linear complexity to be compatible with other pseudo random sequences, there exists some scope of its improvements to make it a perfect sequence. As for drawbacks of NTU sequence, its distribution of bit patterns does not become uniform and in terms of autocorrelation, there exists a small difference between the maximum peak and other peaks. The foremost contribution of this thesis is to overcome these drawbacks of the NTU sequence, as well as to theoretically prove these properties. Several innovative ideas adopted to improve the features of the NTU sequence, among them the use of the sub extension field during the sequence generation procedure, suited the most. It should be noted that to adopt the sub extension field, the conventional trace calculation modified, which is so-called the cascaded trace, as well as the Legendre symbol calculation also modified.

In this dissertation, the proposed pseudo random sequence is defined over the sub extension field, whereas all NTU_class previous sequences are defined over the prime field. Therefore, it is a new and innovative perception to consider the sub extension field during the sequence generation procedure. As an outcome, in case of correlation (both autocorrelation and cross-correlation) except the maximum peak, other peaks are suppressed to be very small, while the high peaks are appearing in the sequences generated over the prime field. Theoretic proof of the correlation explained in detail along with experimental and comparison results in this dissertation.

The distribution of bit patterns is an important measure to check the randomness of a sequence. In terms of this crucial property, binary sequence (defined over the sub extension field) holds much better (close to uniform) bit distribution compared to the sequence (defined over the prime field). In this dissertation, mathematical proof of this property included along with experimental and comparison results.

The linear complexity of a sequence is considered as the most important property because it is closely related to how difficult it is to guess the next bit by observing the previous bit pattern of a sequence. It is worth to know the linear complexity to judge the forward and backward unpredictability of a pseudo random sequence. In addition, to examine the randomness regarding a sequence, well-known statistical tests are available, among them NIST statistical test suite is prominent. This dissertation includes the experimental observation of the linear complexity and NIST statistical test.

There are several scopes to improve the proposed pseudo random sequence in this thesis. As future works, a uniformization algorithm can be used to make the distribution of bit patterns uniform without degrading the linear complexity. The author of this dissertation also would like to use proposed pseudo random sequence in some suitable real applications.

Acknowledgements

The last 4 years were one of the best times of my life that I would cherish forever. I am immensely blessed throughout this period for which I have many people to thank. I am grateful to many people who have directly and indirectly helped me finish this work.

First and foremost, I would like to express my deepest gratitude to my supervisor, Professor *Yasuyuki Nogami*, for his ceaseless supervision, innumerable counseling, and unrelenting conviction. I am indebted to *Nogami Sensei* for having me in his laboratory *(Information Security Laboratory)* as a master's student as well as a doctoral student and allowing to work in the field of my research interests. He not only teaches me cryptography and sequences but also reveals the beauty of knowledge. *Nogami Sensei's* invaluable encouragement and constructive criticism have been crucial in helping me to develop as a researcher. His counsel and expertise in the field resolved many difficulties that I encountered during my studies under his supervision. I thank *Nogami Sensei* for his advice and valued suggestions to all aspects of my research and beyond research. This dissertation would not have been possible without his inspiration and encouragement.

I am also most grateful to my doctoral course co-supervisors Professor *Yoshitaka Toyota* (Optical and Electromagnetic Waves Laboratory) and Associate Professor *Minoru Kuribayashi* (Distributed System Design Laboratory) for having their time to read my thesis draft. Their insightful comments and helpful advice helped to shape the thesis into this state.

I am obliged to Professor *Nobuo Funabiki* (Distributed System Design Laboratory) I must recall my experience of taking the course "System Security and Optimization". His strong passion for algorithmic problem solving during the lectures was not only inspiring but also contagious.

I would like to thank Professor *Satoshi Denno* during my days at *Secure Wireless System Laboratory* He provided me with the deep-seated idea of the research works as well as the Japan life. His questions, suggestions, and comments during the time of half-yearly progress meetings were very instinctive.

I want to express my gratitude to Senior Assistant Professor *Takuya Kusaka* (Information Security Laboratory) for the in-depth discussion of scientific topics. His strong work ethic and passion for research helped us to publish some of the remarkable collaborative works. He was always there to help me while any difficulty arose from attending a conference or to publish a paper.

I express my gratitude to Senior Assistant Professor *Hiroto Kagotani* of (Information System Design Laboratory) for his enlightening comments during the half-yearly progress report presentation.

I am also grateful to Assistant Professor *Kengo Iokibe* (Optical and Electromagnetic Waves Laboratory) for the collaborative work we had on masking on AES block cipher by using the pseudo random sequence.

I would like to extend my thanks to Professor *Robert H. Morelos-Zaragoza* of San Jose State University, USA for collaborating with my research. His valuable suggestions and comments during my visit to his laboratory inspired me a lot to do quality research work in the area of pseudo random sequence generation.

I am also thankful to Professor *Howon Kim* of Pusan National University, South Korea and his Ph.D. student Taehwan Park for great research collaboration on IoT security.

Thanks to Japanese Government *Monbukagakusho: MEXT* Scholarship, which fulfilled my dream to pursue my higher study (masters and doctoral) in Japan. I sincerely acknowledge all the funds that afforded me to join several international conferences (in Japan as well as abroad) and conduct research activities.

I am also grateful to all administrative officers of the Faculty of Engineering who directly or indirectly made an impact on my doctoral study. My special thanks to Ms. Yumiko Kurooka for her kind support in administrative works.

Special thanks also to my seniors, juniors, and friends in the laboratory for creating a great work atmosphere and their generous support. Thanks to sequence team members of my lab who are one of the brightest minds I've worked with. Thanks, all the laboratory members for the wonderful time we spent together.

I cannot thank enough to my wife and children for their endless love, support and sacrifices during my staying in Japan. I am indebted to my parents for their encouragements and sacrifices. They always supported me in my studies, my goings abroad, and any of my decisions. Without them, I would not be here.

At last, I would like to thank all the cryptographic researchers whose works keep inspiring students like me. I would like to extend my thanks to all my research collaborator, co-authors, and reviewers for making a memorable journey of my doctoral study.

Contents

List of Figures

List of Tables

List of Notations and Symbols

Notation	Description
p	$p > 3$ is an odd prime integer in this thesis.
m	Positive integer, mainly denotes the extension degree.
m'	One of the factors of m.
q	Power of odd prime $q = p^{m'}$.
k	Prime factor of $q-1$ such as $k \mid (q-1)$.
$x \bmod p$	Modulo operation. The least nonnegative residue of x modulo p.
$\mathbb{F}_p$	Prime field. The field of integers mod p.
$\mathbb{F}_{p^m}$	An extension field (base p and extension degree m).
$\mathbb{F}_q$	Sub extension field.
$\mathbb{F}_p^*$	The multiplicative group of the field $\mathbb{F}_p$.
$\mathbb{F}_q^*$	The multiplicative group of the sub extension field $\mathbb{F}_q$.
$\lfloor \cdot \rfloor$	The floor of $\cdot$ is the greatest integer less than or equal to $\cdot$. For example, $\lfloor 1 \rfloor = 1$ and $\lfloor 6.3 \rfloor = 6$.
$f(x)$	A primitive polynomial.
ω	Root of an irreducible polynomial.
g	Generator of a group.
ϵ_k	Primitive k-th root of unity.
$\mathcal{S}$	Proposed sequence in this paper.
n	Period of a sequence $\mathcal{S}$.
$\mathrm{R}_{\mathcal{S}}(\cdot)$	Autocorrelation of a sequence $\mathcal{S}$.
$\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(\cdot)$	Cross-correlation between the sequences $\hat{\mathcal{S}}$ and $\mathcal{S}$.

Dedicated to the people I owe the most. To my parents who brought me to this world, my wife and children who sacrificed the most during my Ph.D. journey.

Research Activities

Peer-Reviewed Journal Papers (First author)

1. **Md. Arshad Ali**, Yuta Kodera, Md. Fazle Rabbi, Takuya Kusaka, Yasuyuki Nogami, Satoshi Uehara, and Robert H. Morelos-Zaragoza. "Well Balanced Multi-value Sequence and its Properties Over Odd Characteristic Field". In: *Advances in Science, Technology and Engineering Systems Journal*, vol. 4, no. 4 (2019), pp. 188-196. DOI: 10.25046/aj040423.

2. **Md. Arshad Ali**, Yuta Kodera, Takuya Kusaka, Satoshi Uehara, Yasuyuki Nogami, and Robert H. Morelos-Zaragoza. "Multi-value Sequence Generated Over Sub Extension Field and Its Properties". In: *Journal of Information Security*, vol. 10, no. 3 (2019), pp. 130-154. DOI: 10.4236/jis.2019.103008.

3. **Md. Arshad Ali**, Yuta Kodera, Takuya Kusaka, Yasuyuki Nogami, Satoshi Uehara, and Robert H. Morelos-Zaragoza. "Distribution of Bit Patterns in Binary Sequence Generated Over Sub Extension Field". In: *Advances in Science, Technology and Engineering Systems Journal*, vol. 4, no. 2 (2019), pp. 370-379. DOI: 10.25046/aj040246.

4. **Md. Arshad Ali**, Yuta Kodera, Taehwan Park, Takuya Kusaka, Yasuyuki Nogami, and Howon Kim. "Relation Between the Irreducible Polynomials that Generates the Same Binary Sequence Over Odd Characteristic Field". In: *Journal of Information and Communication Convergence Engineering*, vol. 16, no. 3 (2018), pp. 166-172. DOI: 10.6109/jicce.2018.16.3.166.

5. **Md. Arshad Ali**, Emran Ali, Md. Ahsan Habib, Md. Nadim, Takuya Kusaka, and Yasuyuki Nogami. "Pseudo Random Ternary Sequence and Its Autocorrelation Property Over Finite Field". In: *International Journal of Computer Network and Information Security*, vol. 9, no. 9 (2017), pp. 54-63. DOI: 10.5815/ijcnis.2017.09.07.

Peer-Reviewed International Conference Papers (First author)

ACM Proceedings:

6. **Md. Arshad Ali**, Yuta Kodera, Shoji Heguri, Takuya Kusaka, Satoshi Uehara, and Robert H. Morelos-Zaragoza. "Bit Distribution of Binary

Sequence Generated by Trace Function and Legendre Symbol Over Sub Extension Field". In: *The 6th International Conference on Information Technology (ICIT 2018)*, Hong Kong, December 2018, pp. 92–96. DOI: 10.1145/3301551.3301562.

IEEE Xplore indexed:

7. **Md. Arshad Ali**, Takeru Miyazaki, Shoji Heguri, Yasuyuki Nogami, Satoshi Uehara, and Robert Morelos-Zaragoza. "Linear complexity of pseudo random binary sequence generated by trace function and Legendre symbol over proper sub extension field". In: *Eighth International Workshop on Signal Design and Its Applications in Communications (IWSDA) 2017*, Sapporo, Japan, September 2017, pp. 84–88. DOI: 10.1109/IWSDA.2017.8095741.

8. **Md. Arshad Ali**, Takeru Miyazaki, Yasuyuki Nogami, Satoshi Uehara, and Robert H. Morelos-Zaragoza. "Multi-value sequence generated by trace function and power residue symbol over proper sub extension field". In: *IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW) 2017*, Taipei, Taiwan, June 2017, pp. 251– 252. DOI: 10.1109/ICCE-China.2017.7991089.

9. **Md. Arshad Ali**, Hiroto Ino, Chiaki Ogawa, and Yasuyuki Nogami. "Linear complexity of signed binary sequence over odd characteristic field". In: *19th International Conference on Computer and Information Technology (ICCIT) 2016*, Dhaka, Bangladesh, December 2016, pp. 266-269. DOI: 10.1109/ICCITECHN.2016.7860207.

10. **Md. Arshad Ali**, Yasuyuki Nogami, Hiroto Ino, and Satoshi Uehara. "Auto and Cross Correlation of Well Balanced Sequence over Odd Characteristic Field". In: *Fourth International Symposium on Computing and Networking (CANDAR) 2016*, Hiroshima, Japan, November 2016, pp. 604-609. DOI: 10.1109/CANDAR.2016.0109.

11. **Md. Arshad Ali**, Yasuyuki Nogami, Chiaki Ogawa, Hiroto Ino, Satoshi Uehara, and Robert Morelos-Zaragoza. "A new approach for generating well balanced Pseudo-random signed binary sequence over odd characteristic field". In: *International Symposium on Information Theory and Its Applications (ISITA) 2016*, Monterey, CA, USA, November 2016, pp. 777-780. E-ISBN: 978-4-88552-309-0 .

12. **Md. Arshad Ali** and Yasuyuki Nogami. "A pseudo-random binary sequence generated by using primitive polynomial of degree 2 over odd characteristic field $\mathbb{F}_p$". In: *IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW) 2016*, Nantou, Taiwan, May 2016, pp. 1-2. DOI: 10.1109/ICCE-TW.2016.7520895.

IEICE/IEIE sponsored:

13. **Md. Arshad Ali**, Yuta Kodera, Takuya Kusaka, Takeru Miyazaki, Yasuyuki Nogami, and S. Uehara, and Robert H. Morelos-Zaragoza. "Linear Complexity of Pseudo Random Binary Sequence Generated Over Proper Sub Extension Field". In: *33rd International Technical Conference on Circuits / Systems, Computers and Communications, ITC-CSCC 2018*, Bangkok, Thailand, July 2018, pp. 448-451, IEIE, USB.

Peer-Reviewed Journal Papers (Co-author)

14. Yuta Kodera, **Md. Arshad Ali**, Takeru Miyazaki, Takuya Kusaka, Yasuyuki Nogami, Satoshi Uehara, and R. H. Morelos-Zaragoza. "Algebraic Group Structure of the Random Number Generator: Theoretical Analysis of NTU sequence(s)". In: *Special Issue on Information Theory and Its Applications, IEICE Transaction on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E102-A, no. 12 (2019).

15. Yuta Kodera, Takeru Miyazaki, Md. Al-Amin Khandaker, **Md. Arshad Ali**, Takuya Kusaka, Yasuyuki Nogami, and Satoshi Uehara. "Distribution of Digit Patterns in Multi-value Sequence over the Odd Characteristic Field". In: *Special Issue on Discrete Mathematics and Its Applications, IEICE Transaction on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E101-A, no. 9 (2018), pp. 1525–1536. DOI: 10.1587/transfun.E101.A.1525.

16. Shunsuke Ueda, Ken Ikuta, Takuya Kusaka, Md. Al-Amin Khandaker, **Md. Arshad Ali**, and Yasuyuki Nogami. "An Extended Generalized Minimum Distance Decoding for Binary Linear Codes on a 4-Level Quantization over an AWGN Channel". In: *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 101-A, no. 8 (2018), pp. 1235–1244. DOI: 10.1587/transfun.E101.A.1235.

17. Yuta Kodera, Takeru Miyazaki, Takuya Kusaka, **Md. Arshad Ali**, Yasuyuki Nogami, and Satoshi Uehara. "Uniform Binary Sequence Generated over Odd Characteristic Field". In: *Special Issue, International Journal of Information and Electronics Engineering*, vol. 8, no. 1 (2018), pp. 5–9. DOI: 10.18178/IJIEE.2018.8.1.685.

International Conference Papers (Co-author)

IEEE Xplore indexed:

18. Yuta Kodera, Takeru Miyazaki, Md. Al-Amin Khandaker, **Md. Arshad Ali**, Yasuyuki Nogami, and Satoshi Uehara. "Distribution of bit patterns on multi-value sequence over odd characteristics field". In: *IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW) 2017*, Taipei, Taiwan, June 2017, pp. 137–138. DOI: 10.1109/ICCE-China.2017.7991033.

IEICE sponsored:

19. Chiaki Ogawa, **Md. Arshad Ali**, Yasuyuki Nogami, Satoshi Uehara, Kazuyoshi Tsuchiya, and Robert Morelos-Zaragoza. "Pseudo Random Binary Sequence Generated by Trace and Legendre Symbol with Non-Primitive Element in $\mathbb{F}_{p^2}$". In: *International Symposium on Nonlinear Theory and Its Applications (NOLTA) 2016*, Yugawara, Japan, November 2016, pp. 242–244, IEICE, CD-ROM.

Chapter 1

Introduction

Everything we do to achieve privacy and security in the computer age depends on random numbers – Simon Cooper

Recently, information technologies and communications are an inseparable part of our daily life. To maintain social and business interactions, people are very much used to the internet, banking systems, mobile phones, and so on. This dependency brings the threat of vulnerabilities regarding their most sensitive information which can cause harm in a variety of ways such as unauthorized illegal actions. Consequently, the study of cryptology has been widespread due to it ensures secure transmission of data over the public channels [Sch15; Gu16]. Several types of cryptographic techniques are utilized to ensure authentication, privacy, and integrity during the use of the internet, financial, and social services. The idea of cryptography began thousands of years ago, however, it was basically used for the government and military organizations. Nowadays, cryptography becomes an integral part of our modern digitized society and random numbers are an important tool in cryptography. This dissertation concerns several properties (both theoretically and experimentally) of a pseudo-random sequence generator. This chapter introduces cryptography, random sequences and their importance in cryptography, motivation, contribution, and outline of this thesis.

1.1 Cryptography

Cryptography has a long and fascinating history. For over 4500 years, cryptography has existed as a means of secretly communicating information from one party to another. The *Force: Cracking the Data Encryption Standard* [Cur05], by Matt Curtin and *Codebreakers: The Story of Secret Writing* [Lar67], by David Khan are prominent books on the cryptographic history, provides interesting information on how cryptography has affected world events. The development of computers and communication systems in the 1960s brought a demand from the private sector to protect the information in digital form and to provide security services. In response, Feistel at IBM in the early 1970s and culminating in 1977 with the adoption as a U.S. Federal Information Processing Standard for encryption unclassified information, DES, the Data Encryption Standard,

is the most widely known cryptographic mechanism in history. Regarding the security concern, the DES remains the standard for many financial institutions around the world.

The contemporary form of cryptography was invented by Claude Shannon in 1949 [CE49]. In the area of electronic communications and digital computing, Shannon for his well-established fundamental theory for cryptography and it is counterpart, cryptanalysis. Shannon's proposal mainly relies on a unique shared secret, the *key*, that allowed two parties to communicate securely if this *key* was not revealed. This class of algorithms, known as a secret key, private key, or symmetric key, was the sole method of secure communication. The most conspicuous development in the history of cryptography came in 1976, when Whitefield Diffie and Martin Hellman proposed a revolutionary key distribution technique [DH76]. Diffie and Hellman's proposal led to the development of a new class of algorithms, the public key or asymmetric key, where a pair of mathematically related *keys* are utilized. Among these two keys, one is made public, which needs to share secretly between specific two parties. On the other hand, asymmetric cryptography also known as public key cryptography, uses public and private keys to encrypt and decrypt data. One key in the pair can be shared with everyone; it is called the public key. The other key in the pair is kept secret; it is called the private key. Either of the keys can be used to encrypt a message; the opposite key from the one used to encrypt the message is used for decryption. Nowadays, information systems use a hybrid approach, which combines both symmetric key and asymmetric key algorithms to ensure fast and secure communication.

Cryptography is the study of mathematical techniques related to aspects of information security such as confidentiality, data integrity, entity authentication, and data origin authentication. Modern cryptography concerns with the following basis:

- **Confidentiality:** Information cannot be observed by an unauthorized entity. Several approaches exist for providing confidentiality, ranging from physical protection to mathematical algorithms which make the information enigmatic for the unapproved entities.
- **Data integrity:** Transmitted or stored data cannot be altered due to error or by an unauthorized entity. To confirm data integrity, one must have the ability to identify data manipulation (such as insertion, deletion, substitution, and so on).
- **Authentication:** The term identification is synonymous with authentication. Authentication applies to both entities and transmitted information. Therefore, the entities should provide a certificate to identify the authenticity of the information. Such authentication can be done using a digital signature.
- **Non-repudiation:** A rule, which prevents an entity (both sender and receiver) from denying previous commitment or action. If such dispute

arises by the action of an entity, then resolve this situation is mandatory. This is accomplished via digital signature.

1.1.1 Importance of Randomness in Cryptography

Cryptography can provide a wide range of solutions for different security requirements. To do so, cryptography requires random sequences for many different purposes, therefore randomness takes a variety of roles as follows [KAI10; Kin+12].

One of the most important roles randomness plays in cryptography is represented by **cryptographic keys**, which responsible for the transformation of the plaintext into ciphertext and vice versa. Consider that both the encryption and the decryption algorithms are publicly known together with all the ciphertexts, the security of the whole cryptosystem is dependent on how the key information is managed (such as generated, applied, stored, and destroyed).

Both stream ciphers and block ciphers use **initialization vectors**. Although their implementation differs, the purpose of using the initialization vectors are the same. Initialization vectors ensure that the ciphers produce a unique output even under the same encryption key to avoid the laborious work of re-keying.

Random **nonce values** are generally used in entity authentication (challenge-handshake authentication protocols) and authentication key establishment protocols. Typically, a nonce is some values with time, in order to verify that specific values are not reused. A nonce can be a time stamp, a visit counter on a webpage or a special marker intended to limit or prevent the unauthorized reply or reproduction of a file.

Random nonce values are very important in **zero-knowledge proof** protocols. A zero-knowledge proof is a protocol that provides a scope for interactive proof with a prover and a verifier, where the prover convinces the verifier of a statement (with the high probability) without revealing any information about how to go about proving the statement.

The blind signature scheme is a special class of digital signature algorithm, which uses the **blinding values**. It provides additional features to authentication and non-repudiation, like the blindness the property which makes sure of signing a message without exposing the content to the signer. The blind signature scheme is of great importance in many security-sensitive applications (such as electronic voting, digital cash, and so on), where privacy is the major matter of concern.

The randomness plays many roles in cryptography and the importance of each role extensively relies on the strength of the randomness. This randomness can achieve by a good random number generator; therefore, this thesis concentrates on the generation of random sequence which holds desire randomness properties.

1.1.2 Random Number Generators

Many cryptographic applications require random numbers, for example, to generate secret keys, to encrypt messages or to mask the content of certain protocols by combining the content with a random sequence. Therefore, random number generators (RNGs) are part of many IT-security products. Inappropriate RNGs may totally weaken IT systems that are principally strong. Random number generators basically classified into two categories: true random number generators (TRNGs) and pseudo random number generators (PRNGs), a brief introduction regarding these classifications are introduced in the following section [KAI10; Kin+12; DiC12].

1.1.2.1 True Random Number Generators (TRNGs)

A true random number generator uses entropy[1] sources that already exist, no need to invent them. Some real-world events such as after flipping a coin, it is almost impossible to guess the result accurately. Thus, a coin flips have a high degree of entropy and the same thing happened for rolling a dice also. Flipping coins and rolling dice are two ways of entropy could be obtained for a generator, although the rate at which random numbers could be produced would be restricted. Such generators rely on some sort of hardware; thus, their unpredictability being guaranteed by the physical laws. TRNGs require some sort of hardware, which make the true random number generator more expensive to implement. In the case of TRNGs, without considering any kind of side channel attacks, physical devices are typically vulnerable to wear over time and errors in their construction that can naturally bias the generated sequences. To overcome this bias, most of the true random number generators adopt some post-processing (such as filtering, to reduce the possible correlation and making the output like a statistically perfect sequence) to compensate for it. Nonetheless, PRNGs need specialized hardware, the generators are expensive and some of them are slow and impractical.

FIGURE 1.1: Example of true random number generator.

[1] Entropy refers to the amount of uncertainty about an outcome.

1.1.2.2 Pseudo Random Number Generators (PRNGs)

Random number generator that does not rely on real-world phenomena to generate random stream is referred to as pseudo random number generators. PRNGs exact randomness comes from an initial value, called *seed*, which is expanded by using a recursive formula. Unlike the TRNGs, PRNGs generate random sequences using software methods only. As a result, the effective entropy depends on the seed value, therefore, to obtain the seed from a true random number sequence is often recommended.

The term forward security is used to describe a generator where knowing the internal state of a generator at a point in time will not help an attacker to know about the previous outputs [ZBT06]. This forward security becomes a vital issue when pseudo random number generators are being used for password creation. On the other hand, backward security denotes that an attacker who learns the state of the generator at a point in time will not be able to determine future numbers that will produce. Backward security can be achieved by introducing some level of entropy into the generator's equations. True random number generators have no deterministic components; thus, they always have forward and backward security. The practical features of PRNGs such as high generation speed, good statistical properties, and no need for additional hardware devices made these generators very attractive and most widely used RNGs. In cryptographic systems, they are called by the Cryptographically Secure Pseudo Random Number Generators (CSPRNGs) because they are based on cryptographic primitives or mathematical problems considered to be extremely difficult to solve. Most popular algorithms used in pseudo random number generator are linear feedback shift registers, linear congruential generator, and lagged Fibonacci generator.

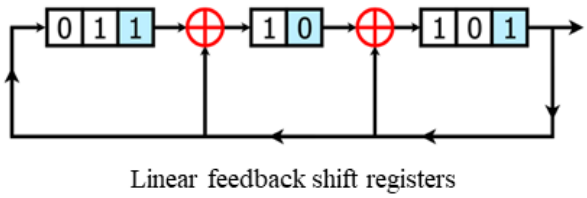

FIGURE 1.2: Example of pseudo random number generator.

1.2 Motivation

The research trend on pseudo random sequence basically classified into two directions: one is from the viewpoint of theoretical aspect (M-sequence [Ren04] and Legendre sequence [XG10; JS+96]), where mathematically explain the sequence features are major matter of concern and the other one is from their practical point of view (Mersenne Twister (MT) [MN98] and Blum-Blum-Shub

(BBS) [BBS86], where the throughput and ease of implementation are important; such as their usage as the key stream in stream ciphers, entity authentication mechanism, private keys for digital signature algorithms, quasi-Monte Carlo integration, and so on). In both cases, the randomness and unpredictability are evaluated by experimental tests. This dissertation concentrates on the theoretical aspect rather than the application point of view of pseudo random sequences. Few prominent pseudo random sequences (having a theoretical background) are briefly introduced here.

Maximum Length Sequence (M-sequence)

Maximum length sequences (also called M-sequence) are constructed based on Galois field theory. M-sequences are generated using linear feedback shift registers (LFSR) structures that implement linear recursion. Any LFSR can be represented as a polynomial of variable X, referred to as the generator polynomial of an M-sequence.

$$G(X) = g_m X^m + g_{m-1} X^{m-1} + g_{m-2} X^{m-2} + \ldots + g_2 X^2 + g_1 X + g_0.$$

The coefficients g_i represents the tap weights (for feedback connection depending on the value of 0 or 1), m be the number of LFSR stages, and the generator polynomial $G(X)$ be a primitive polynomial (it is a kind of polynomial that can generate a maximum length polynomial sequence). Let the generator polynomial $G(X)$ be $X^3 + X + 1$, then the design of LFSR is shown in **Figure** 1.3. Few

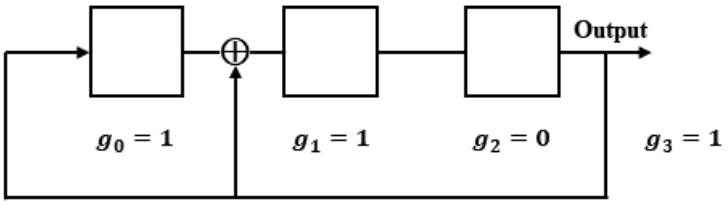

FIGURE 1.3: An M-sequence generated by 3 LFSRs.

notable properties regarding M-sequence are as follows:

- An m-bit register generates an M-sequence of period $2^m - 1$.
- An M-sequence contains exactly 2^{m-1} ones and $2^{m-1} - 1$ zeros.
- A cyclic shift (left-cyclic or right-cyclic) of an M-sequence is also an M-sequence.
- The sum of an M-sequence and a cyclic shift of itself (mod 2, term by term) is another M-sequence.
- The modulo-2 sum of an M-sequence and another phase of the same sequence yields yet a third phase of the sequence.

Furthermore, M-sequence has maximum period, the best periodic autocorrelation in terms of minimizing the maximum value of the out-of-phase autocorrelation, and uniform distribution of bit patterns. On the other hand, the linear complexity of M-sequence is minimum. All the properties regarding the M-sequence are already theoretically proven, thus many researchers attracted by those theoretic proofs [Nea59].

Legendre Sequence (L-sequence)

Legendre sequences (also called L-sequence) are generated by applying the Legendre symbol. These sequences are to be any prime length and are based on quadratic residues. The Legendre symbol is a multiplicative function. Let p be a prime, the Legendre sequence s^{∞} with respect to the prime p is defined by

$$s_i = \begin{cases} \frac{1+\left(i/p\right)}{2}, & \text{if } i \neq 0 \bmod p, \\ 0, & \text{otherwise,} \end{cases}$$

for each $i \geq 0$, where $\left(i/p\right)$ is the Legendre symbol. The Legendre sequence is a well-known pseudo random sequence with ideal periodic and aperiodic correlation functions and is also known for exhibiting large linear complexity. There is an extensive literature devoted to the Legendre sequence properties due to theoretic proof of its properties [HA06; Dam90; Din98; HK98; KS01; MS97].

The study of pseudo random sequence generation by combining the well-known M-sequence and Legendre sequence started with the NTU (Nogami-Tada-Uehara) sequence [NTU14]. Although NTU sequence holds interesting features (such as its long period and high linear complexity) to be compatible with other pseudo random sequences, there exists some scope of its improvements to make it a perfect sequence. As for drawbacks of NTU sequence, its distribution of bit patterns does not become uniform and in terms of autocorrelation, there exists a small difference between the maximum peak and other peaks. The key motivation of this work is to improve the drawbacks in the NTU sequence, as well as to explain these sequence features mathematically.

1.3 Contribution

As discussed above, there are several applications of pseudo random sequences in enormous applications in the area of information security, especially in cryptography. Such applications choose random sequences having unpredictability property and good statistical properties such as long period, low correlation, high linear complexity, and uniform distribution of bit patterns. Moreover, it also needs to be resistant against the cryptanalytic attacks. In addition, many cryptographic statistical tests have been proposed to measure the randomness in a pseudo random sequence, for example, the NIST statistical test

suite [Ruk+00]. A pseudo random sequence, that will pass all the statistical tests will be referred to as a cryptographically secure pseudo random sequence.

In the preceding section, it is already mentioned that the study on pseudo random sequences basically classified into the theoretical aspect and their applications in cryptography. This dissertation prefers to study pseudo random sequences from the theoretical aspect rather than their applications. The study of pseudo random sequence generation by combining the well-known M-sequence and Legendre sequence started with the NTU (Nogami-Tada-Uehara) sequence. To represent NTU sequence as a perfect pseudo random sequence, all its features should be up to the mark like other sequences. The foremost contribution of this thesis is to overcome these drawbacks such as its distribution of bit patterns does not become uniform and in terms of autocorrelation, there exists a small difference between the maximum peak and other peaks of the NTU sequence, as well as to mathematically prove these properties. Several innovative ideas adopted to improve the features of the NTU sequence, among them the use of the sub extension field during the sequence generation procedure, suited the most. It should be noted that to adopt the sub extension field, the conventional trace calculation modified, which is so-called the cascaded trace, as well as the Legendre symbol calculation also modified.

The proposed sequence in this thesis could not be generated without a few obligatory functions such as primitive polynomial, trace function, and Legendre function. At first, focus on the procedure for generating the proposed sequence very briefly. Initially, a primitive polynomial generates maximum length vector sequence, then the trace function maps an element of the extension field to an element of the sub extension field, and finally, the Legendre symbol binarizes the trace outputs to make a pseudo random binary sequence. It should be noted that the proposed sequences in this dissertation are defined over the sub extension field, whereas all the previous sequences proposed in [NTU14; Nog+16; Ali+16b; Ali+16a] are defined over the prime field. Therefore, it is a new and innovative perception to consider the sub extension field during the sequence generation procedure to overcome the drawbacks of the NTU sequence and the detail outcomes (in case of correlation, big differences between the peak values; higher linear complexity; close to uniform distribution of bit patterns) of this concept are introduced in later chapters in this thesis.

As previously stated, a primitive polynomial is required during the proposed pseudo random sequence generation procedure. It is well-known that a primitive polynomial is a special kind of polynomial in the extension field. Therefore, it requires an additional calculation to judge the primitiveness of a polynomial. This scenario becomes worse when the degree of a polynomial is large, thus finding a primitive polynomial surely a time-consuming calculation. It is another contribution of this dissertation to find a relationship between the polynomials and sequences. By using this relationship, same sequence (having the same property) by using an irreducible polynomial instead of a primitive polynomial. Thus, it contributes to increase the candidates for the selection of polynomials.

1.4 Outline of this Thesis

The remainder of this dissertation is structured as follows.

Chapter 2 describes and defines the mathematical concepts related to modular arithmetic, group, ring field, finite field, trace function, Legendre symbol, and cubic residue, which are fundamental mathematical concepts behind the generation of a pseudo random sequence. In addition, pseudo random sequence also introduces along with its properties such as period, autocorrelation, cross-correlation, linear complexity, distribution of bit patterns.

Chapter 3 introduces the proposed pseudo random sequence along with the NTU (Nogami-Tada-Uehara) sequence.

Chapter 4 contains a discussion of period, autocorrelation, and cross-correlation properties regarding a pseudo random sequence (including a binary and multi-value sequence) defined over the sub extension field. These properties are theoretically shown along with experimental results. Besides, one of the notable outcomes is proposed sequence defined over the sub extension field holds low correlation compared to the sequence defined over the prime field.

Chapter 5 introduces the distribution of bit patterns property in a binary sequence defined over the sub extension field. The distribution of bit patterns is an important measure to judge the randomness of a sequence. This chapter restricts the discussion on the distribution of bit patterns property. In terms of this crucial property, the binary sequence which defined over the sub extension field holds much better (close to uniform) bit distribution than the sequence defined over the prime field. This chapter also includes the theoretical proof of the distribution of bit patterns property.

Chapter 6 explains the linear complexity and NIST statistical test experimentally. It is worth to know the linear complexity and NIST test results to judge the unpredictability (both forward and backward) and randomness of a pseudo random sequence, respectively.

Chapter 7 presents how the same binary sequence can be generated by using an irreducible polynomial instead of a primitive polynomial. In an extension field, a primitive polynomial is a special kind of polynomial, thus it is a time-consuming calculation to find such polynomial. However, from the aspect of the practical use in cryptographic systems sequence needs to generate swiftly. The main contribution of this chapter is to find a relation between the generated sequence and irreducible polynomials. In addition, these relationships are explained both theoretically and experimentally.

Finally, **Chapter 8** concludes this dissertation along with an outline of the future works.

Chapter 2

Fundamental Mathematics

This chapter recalls the mathematical concepts related to modular arithmetic, group, ring field, finite field, trace function, Legendre symbol, and cubic residue, which are fundamental mathematical concepts behind generating a pseudo random sequence. In addition, pseudo random sequence also introduces along with its properties such as period, autocorrelation, cross-correlation, linear complexity, distribution of bit patterns. The mathematical fundamentals and pseudo random sequence (including its properties) related introduction presented in this chapter can be found in [Gol67; GG05; LN96].

2.1 Modular Arithmetic

Modular arithmetic is the fundamental tool for modern cryptography especially public key cryptosystems.

Definition 1 (Modular Arithmetic) *Let p be a positive integer named as the modulus and a and b are two arbitrary integers. If p divides $b - a$ then it can be written as,*

$$a \equiv b \pmod{p}$$

and express as a and b are congruent modulo p. ■

Example 2.1 *Let, $p = 7$, $a = 19$ and $b = 5$ then $19 \equiv 5 \pmod{7}$.*

Example 2.2 *Let, $p = 7$, $a = -17$ and $b = 11$. Then $-17 \pmod{7} = 4$ and $11 \pmod{7} = 4$. Therefore,*

$$-17 \equiv 11 \pmod{7}$$

and usually express -17 and 11 are congruent modulo 7.

2.2 Group, Ring, Field

2.2.1 Group

The concept of group is very fundamental to understanding cryptography. It is an algebraic system defined as follows.

Definition 2 (Group) *A group is a non-empty set $\mathbb{G}$ with a binary operation $\circ$ on its elements denoted as $\langle \mathbb{G}, \circ \rangle$, sometimes denoted by $\mathbb{G}$ only, which satisfies the following axioms.*

Closure *The group is closed under the operation $\circ$, i.e. $\forall a \in \mathbb{G}$, and $\forall b \in \mathbb{G}$ the result of $(a \circ b) = c \in \mathbb{G}$.* [1]

Identity element *There exist an **identity element** e also know as neutral element or unit element in $\mathbb{G}$ such that $\forall a \in \mathbb{G}$, $a \circ e = e \circ a = a$.*

Inverse element *For $\forall a \in \mathbb{G}$, there exists an element $b \in \mathbb{G}$ such that $a \circ b = e = b \circ a$, where b is called inverse element of a.*

Associativity *Elements in group $\mathbb{G}$ should follow associativity. i.e. $(a \circ b) \circ c = a \circ (b \circ c)$ for all $a, b, c \in \mathbb{G}$.*

Definition 3 (Commutative Group)

A group $\mathbb{G}$ will be commutative if $a \circ b = b \circ a$ for all $a, b \in \mathbb{G}$.

■

A commutative group is also called *abelian* group.

Example 2.3 *The set of integers $\mathbb{Z}$ forms a group under the group operation of addition $+$ denoted as $(\mathbb{Z}, +)$, where 0 is the identity element of the group.*

Example 2.4 *The set of positive integers $\mathbb{N}$ under addition does not form a group since elements have not inverse.*

Definition 4 (Order of a Group) *The order of a group $\mathbb{G}$ often denoted as $\#\mathbb{G}$ is the number of elements in the group $\mathbb{G}$.* ■

Remark 1 *Groups order can be finite and infinite. In example 2.3, $(\mathbb{Z}, +)$ has infinite order.*

Definition 5 (Order of group element) *For an element $a \in \mathbb{G}$, the smallest positive integer m such that $a^m = e$ is called the order of a, where e is the identity element in $\mathbb{G}$.* ■

Example 2.5 *Finite group: As shown in example 2.4, the set $\mathbb{N}$ under addition does not form a group since it does not satisfy the group axioms. Let us consider a set $\mathbb{N}_n$ under the operation* mod n *such that*

$$\mathbb{N}_n = \{0, 1, 2, 3, \ldots, n-1\},$$

[1] $\forall$ symbol bears this usual notation *"for all"*.

*where $n \in \mathbb{N}$. It means $\mathbb{N}_n$ is the set of remainders under "*mod n*". Recall the modular arithmetic that*

$$a + b \equiv c \mod n \qquad a, b \in \mathbb{N}_n,$$

means c is associated to a remainder on division by n, when $a + b = c \notin \mathbb{N}_n$. It makes c belongs to $\mathbb{N}_n$ making $(\mathbb{N}_n, +)$ forming a group. It also includes element 0 which acts as an identity element.

Definition 6 (Group generator) *For a given group $\mathbb{G}$ if there is an element $g \in \mathbb{G}$ such that for any $a \in \mathbb{G}$ there exist an unique integer i with $a = g^i$ then g will be called a generator of $\mathbb{G}$.* ■

Definition 7 (Cyclic Group) *A group $\mathbb{G}$ will be* cyclic *if there exist at least one generator $g \in \mathbb{G}$. Cyclic group usually expressed as $\mathbb{G} = \langle g \rangle$.* ■

Remark 2 *The number of generator in a group $\mathbb{G}$ of order n is defined by Euler's totient function $\phi(n)$[2]. If n is a prime p then the group $\mathbb{G}$ will be called prime order group and it will have $\phi(p) = p - 1$ generators.*

In this case, assume the notation $\langle \mathbb{G}, \circ \rangle$; there exists some ambiguity which operation we consider. Therefore, the following two types of group nations are prevalent in literature.

Definition 8 (Additive group) *A cyclic group is called additive if its group operation can be written in the same way of performing addition, that is*

$$f = g + x,$$

can also appear as $[x]g$ means applying $x - 1$ times addition operator $+$ on g. It is also common to write as $x \cdot g$. For example, 1 is one of generators in group $(\mathbb{Z}_5, +)$ under addition modular 5, then $1 \cdot 4$ can be written as,

$$4 = 1 + 1 + 1 + 1.$$

■

Definition 9 (Multiplicative group) *A cyclic group is called multiplicative if its group operation can be written in the same way of performing multiplication, that is*

$$f = g \cdot x \text{ or } f = g^x.$$

■

Remark 3 *In both notation the x is an integer called the discrete logarithm of h to the base g.*

Remark 4 *Unless otherwise stated, through out this thesis the notation xg used for ordinary addition e.g. $a + a = 2a$ and $a + a + a = 3a$ and for multiplicative notation, these will denoted by a^2, a^3.*

[2]When n is a positive integer, Euler's totient function $\phi(n)$ gives the number of positive integers less than or equal to n that are co-prime to n.

From the definition cyclic group, it can be see visualized that any elements in cyclic a group are generated with iterative operations of generator g. **Figure** 2.1 shows this schematically.

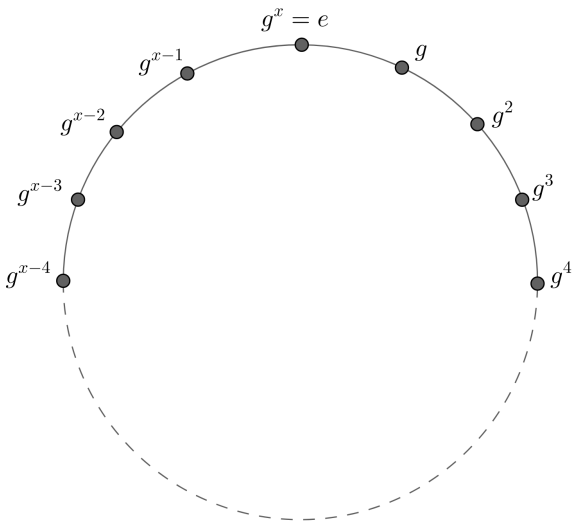

FIGURE 2.1: Cyclic group.

A well known practice of presenting a finite group's operation is *Cayley table* as shown in example 2.6. Cayley table shows all possible group operation that can be performed in a finite group.

Example 2.6 *The Cayley table for the group $\mathbb{Z}_4$ is:*

$\oplus_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

In the above example of group $(\mathbb{Z}_4, +)$, there are $\phi(4) = 2$ generators, 3 and 1.

Definition 10 (Subgroup) *Let $\mathbb{H}$ be a non-empty subset of group $\mathbb{G}$, $\mathbb{H}$ will be called subgroup of $\mathbb{G}$ if $\mathbb{H}$ itself follows group axioms and $\mathbb{H}$ has the same identity element of group $\mathbb{G}$.* ■

Theorem 1 (Lagrange's Theorem) *Let $\mathbb{G}$ be a finite abelian group and $\mathbb{H}$ is a subgroup of $\mathbb{G}$. The order of $\mathbb{G}$, $\#\mathbb{G}$ is divisible by the order of subgroup $\mathbb{H}$, $\#\mathbb{H}$ i.e. $\#\mathbb{H}|\#\mathbb{G}$.* ■

Theorem 2 (Fermat's Little Theorem) *Let p is a prime and $a \in \mathbb{Z}$, then*

$$a^p = a \pmod{p}.$$

■

Fermat's little theorem is a special case of Lagrange's theorem.

2.2.2 Ring

The concept of *ring* will not come as frequently as group and field in the subsequent chapters. However, it is relevant to define the ring to understand the related concept.

Definition 11 (Ring) *A* ***ring*** $\mathbb{R}$ *is an algebraic structure with two operations, i.e. addition* $+$ *and multiplication* $\cdot$ *usually denote as* $(\mathbb{R}, +, \cdot)$.

- $\mathbb{R}$ *is abelian group under addition operation.*
- *Under multiplication,* $\mathbb{R}$ *is closed and associative with identity element* 1.
- *Multiplication is distributive over addition:* $\forall a, b, c \in \mathbb{R} : a \cdot (b + c) = a \cdot b + a \cdot c$.

■

If multiplication operation is commutative, $\mathbb{R}$ forms a commutative ring.

Definition 12 (Multiplicative Inverse Modulo n) *Let* $\mathbb{Z}_n$ *be a set under modulo* n *and* $a \in \mathbb{Z}_n$. *The multiplicative inverse modulo* n *of* a *can be written as follows:*

$$a \cdot x \equiv 1 \bmod n.$$

The value x *is the multiplicative inverse modulo* n *of* a, *often written as* a^{-1}. ■

Such value of x only exists if $\gcd(x, n) = 1$. If $n = p$ is a prime, then every non-zero element in the set $\mathbb{Z}_p$ will have a multiplicative inverse. Such $(\mathbb{Z}_p, +, \cdot)$ will be a ring and having the above property it will form a field.

2.2.3 Field

Definition 13 (Field) *A field* $(\mathbb{F}, +, \cdot)$ *is a set that obeys two binary operations denoted by* $+$ *and* $\cdot$, *such that:*

- $\mathbb{F}$ *is a commutative group concerning* $+$ *having identity element* 0.
- *Let* $\mathbb{F}^*$*is a subset of* $\mathbb{F}$ *having only non-zero element of* $\mathbb{F}$ *i.e.* $\mathbb{F}^* = \mathbb{F} - \{0\}$. *Then* $\mathbb{F}^*$*will be called a commutative group respect to multiplication, where every element should have multiplicative inverse in* $\mathbb{F}^*$.
- *For all* $a, b, c \in \mathbb{F}$ *the distributive law will be followed, e.g.* $a \cdot (b + c) = a \cdot b + a \cdot c$ *and* $(b + c) \cdot a = b \cdot a + c \cdot a$.

■

Definition 14 (Subfield) *Let* $\mathbb{F}_1$ *is a subset of field* $\mathbb{F}$. $\mathbb{F}_1$ *will be called a subfield if* $\mathbb{F}_1$ *itself obeys the laws of field with respect to the field operation inherited from* $\mathbb{F}$. ■

Remark 5 *In Definition 14,* $\mathbb{F}$ *is called an* extension field *of* $\mathbb{F}_1$. *If* $\mathbb{F}_1 \neq \mathbb{F}$, *then* $\mathbb{F}_1$ *is a* proper subfield *of* $\mathbb{F}$.

Definition 15 (Order of Finite Field) *The order is the number of elements in $\mathbb{F}$. If the order of $\mathbb{F}$ is finite, $\mathbb{F}$ is called finite field.* ■

Definition 16 (Characteristic of Finite Field) *Let $\mathbb{F}$ be a field and smallest positive number n such that $n \cdot a = 0$ for every $a \in \mathbb{F}$. Such n is called characteristic. If there is no such n in $\mathbb{F}$ then $\mathbb{F}$ has characteristics 0.* ■

Most of the works presented in this dissertation deal with finite fields only. A common property of finite fields often used in cryptographic is following:

Theorem 3 *For every finite field $\mathbb{F}$, the multiplicative group $(\mathbb{F}^*, \cdot)$ is cyclic.* ■

Definition 17 (Prime Field) *Let p be a prime. The ring of integers modulo p is a finite field of characteristics p having field order p denoted as $\mathbb{F}_p$ is called a prime field.* ■

Remark 6 *A prime field contains no proper subfield.*

Theorem 4 *Every finite field has a prime field as a subfield.* ■

In this dissertation, finite fields classified into two types, i.e. prime field $\mathbb{F}_p$ and its extension field. **Section** 2.3 explains more of extension field. The prime field $\mathbb{F}_p$ has the order and characteristic as p. Using the modular arithmetic in the same way as Definition 2.3, fundamental operations of prime field can be defined as $\mathbb{F}_p = \{0, 1, 2 \ldots, p-1\}$.

Example 2.7 *The Cayley table for the two operations $+$ and $\cdot$ for elements in $\mathbb{F}_5$ are as follows:*

$\oplus_5$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\odot_5$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

It should be noted that the notations $\oplus_5$ and $\odot_5$ denote the addition mod 5 and multiplication mod 5, respectively. As described above, arithmetic operations in $\mathbb{F}_p$ define by modular operations (mod p) for integers. However, it does not work in an extension field $\mathbb{F}_{p^m}$. In the next section, arithmetic operations in extension field $\mathbb{F}_{p^m}$ is described in detail.

2.3 Extension Field

A subset $\mathbb{F}_0$ of a field $\mathbb{F}$ that is itself a field under the operations of $\mathbb{F}$ will be called a *subfield* of $\mathbb{F}$. In this case, $\mathbb{F}$ is called an *extension field* of $\mathbb{F}_0$. An extension field of a prime field $\mathbb{F}_p$ can be represented as m-dimensional vector space that has m elements in $\mathbb{F}_p$. Let the vector space be the m-th extension field ; it is denoted by $\mathbb{F}_{p^m}$. The order of extension fields $\mathbb{F}_{p^m}$ is given as p^m. In

what follows, let q be the power of p, the extension field of a prime field $\mathbb{F}_p$ is denoted by $\mathbb{F}_q$.

There are several methods to represent an element in extension fields, such as polynomial basis and normal basis. This thesis mostly used polynomial basis. Let ω be a root of m-th irreducible polynomial over $\mathbb{F}_q$, then the following m elements can be represented as follows.

$$\omega,\ \omega^q,\ \omega^{q^2},\ \ldots,\ \omega^{q^{m-1}}.$$

All elements in this set are conjugate to each other. When the set of the conjugates become linearly independent, this is called *normal basis*. Using normal basis, an element $\alpha \in \mathbb{F}_q$ is expressed as a polynomial by

$$\alpha = a_1\omega + a_2\omega^q + a_3\omega^{q^2} + \ldots + a_m\omega^{q^{m-1}}, \tag{2.1}$$

where $a_1,\ a_2,\ a_3, \cdots,\ a_m \in \mathbb{F}_q$.

Arithmetic operations in $\mathbb{F}_q$ are carried out with ordinary addition and multiplication for polynomial and modular reduction by an irreducible polynomial.

In very brief, it can be said that a prime field is a subset of sub extension field and sub extension field is also a subset of extension field, which schematically shown in **Figure** 2.2.

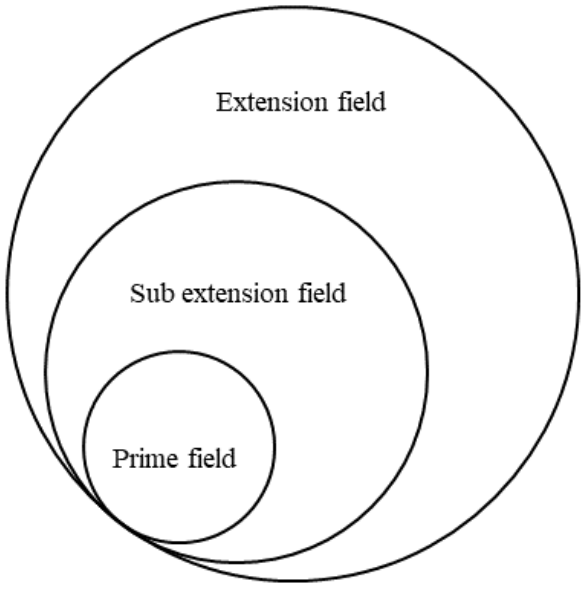

FIGURE 2.2: Prime field, sub extension field, and extension field.

2.4 Polynomial Arithmetic

A polynomial arithmetic is a branch of algebra dealing with some properties of polynomials which share strong analogies with properties of number theory relative to integers. It includes basic mathematical operations such as addition, subtraction, and multiplication, as well as more elaborate operations like euclidean division and properties related to roots of polynomial.

Let X be a variable over the field $\mathbb{F}$. Again let $\mathbb{F}[X]$ be a set of polynomials, which is represented as follows.

$$f(X) = a_0 + a_1X + \ldots + a_kX^k, \quad a_i \in \mathbb{F}_p, \quad i = 0, \ldots, k, \tag{2.2}$$

when $a_k \neq 0$, then the degree of the Eq.(2.2) becomes k. In addition, the degree of $f(X)$ be denoted as deg f.

Assume two polynomials as $f(X) = \sum f_iX^i$ and $g(X) = \sum g_iX^i$. Then the addition and subtraction operations are defined as follows.

$$f(X) + g(X) = \sum (f_i + g_i)X^i.$$
$$f(X) \cdot g(X) = \sum_i \left(\sum_{j=0}^{i} (f_i g_{i-j}) \right) X^i.$$

Particularly, scalar multiplication is defined as follows.

$$sf(X) = \sum (sf_i)X^i.$$

2.4.1 Irreducible and Primitive Polynomials

A polynomial is said to be irreducible , if it cannot be factorized into non-trivial polynomials over the same field. Let $f(X)$ be a polynomial over $\mathbb{F}[X]$ and the degree of $f(X)$ is more than 1. For such polynomial, if it cannot be factorized into smaller degree polynomials (including the scalar factor), it is said to be an irreducible polynomial.

For arbitrary polynomial $f(X), g(X) \in \mathbb{F}[X]$ and $g(X) \neq 0$ holds the following relation as,

$$f(X) = a(X)g(X) + r(X), \quad \deg g > \deg r, \tag{2.3}$$

where both of them are uniquely determined, such as $a(X), f(x) \in \mathbb{F}[X]$. In Eq.(2.3), when $r(X) = 0$, then $g(X)$ divides $f(X)$.

Let ω be an arbitrary element in $\mathbb{F}$. Now consider the relation between ω and $f(X)$. If $f(\omega) = 0$, then ω is said to be the root of $f(X)$. Alternatively, when $f(X) = 0$, then it is said to be that it has a root over $\mathbb{F}$. If an irreducible polynomial $f(X)$ has a root over $\mathbb{F}$ and this root is also a primitive element over $\mathbb{F}$, then $f(X)$ is said to be a *primitive polynomial* . For example, in case of

extension field $\mathbb{F}_{p^m}$, let the modular polynomial $f(X)$ be a primitive polynomial, then the root of $f(X)$ be ω, and it has a multiplicative order of p^m-1. Therefore, the primitive element ω (where $i = 0, 1, 2, \ldots, p^m-1$) holds the following relation.

$$\omega^i = 1 \text{ (only when } i = 0 \text{ or } i = p^m - 1 \text{ and if } i \neq j, \text{ then } \omega^i \neq \omega^j). \tag{2.4}$$

The above equation can be rewrite for $i = p^m - 1$ as follows.

$$\left(\omega^{\frac{p^m-1}{p-1}}\right)^{p-1} = 1. \tag{2.5}$$

furthermore, $\omega^{(p^m-1)/(p-1)}$ has a multiplicative order of $p-1$, which can be confirmed by the above equation. Therefore, $\omega^{(p^m-1)/(p-1)}$ is a primitive element of $\mathbb{F}_p$, that holds the following relation.

$$\omega^{\frac{p^m-1}{p-1}} = g, \tag{2.6}$$

where g is a primitive element of $\mathbb{F}_p$. According to Fermat's little theorem, the following property between $\mathbb{F}_{p^m}$ and its base field $\mathbb{F}_p$ holds [LN86].

Property 1 *Let g be a generator of $\mathbb{F}^*_{p^m}$, $g^{(q-1)/(p-1)}$ becomes a non-zero element in prime field $\mathbb{F}_p$ and is also a generator of $\mathbb{F}^*_p$.* ∎

(Proof) Since g is a generator of $\mathbb{F}^*_{p^m}$, its order is p^m-1. Let, i be a non-negative integer, the order of g^i is given by

$$\frac{p^m - 1}{\gcd(p^m - 1, i)}. \tag{2.7}$$

Therefore, the order of $g^{(p^m-1)/(p-1)}$ is $p-1$. It means that $g^{(p^m-1)/(p-1)}$ is a generator of $\mathbb{F}^*_p$. ∎

2.5 Trace Function

The trace function maps an element of extension field $X \in \mathbb{F}_{p^m}$ to an element of prime field $x \in \mathbb{F}_p$. In other words, the trace of X over $\mathbb{F}_p$ is the sum of the conjugates of X with respect to $\mathbb{F}_p$, where the conjugates can be given by X^{p^i} $(i = 0, 1, 2, \ldots, m-1)$ and the $\mathrm{Tr}(X)$ is given by the following equation.

$$x = \mathrm{Tr}_{p^m|p}(X) = \sum_{i=0}^{m-1} X^{p^i}. \tag{2.8}$$

An important point is that, the above trace becomes a scalar value and the trace function has a linearity property over the prime field $\mathbb{F}_p$ as follows.

$$\mathrm{Tr}(aX + bY) = a\mathrm{Tr}(X) + b\mathrm{Tr}(Y), \tag{2.9}$$

where $a, b \in \mathbb{F}_p$ and $X, Y \in \mathbb{F}_{p^m}$.

Property 2 *For each* $i = 1, 2, \ldots, p-1 \in \mathbb{F}_p$*, the number of elements in* $\mathbb{F}_q$ *whose trace with respect to* $\mathbb{F}_p$ *becomes* i *is given by* $q/p = p^{m-1}$ *and if* $i = 0$*, the number of elements is given by* $p^{m-1} - 1$*.* ∎

(Proof) Elements in $\mathbb{F}_q$ are the roots of $x^q - x$. It is factorized over $\mathbb{F}_p$ as follows:

$$\begin{aligned} x^q - x &= x^{p^m} - x \\ &= \prod_{i=0}^{p-1} (\mathrm{Tr}(x) - i)\,. \end{aligned} \tag{2.10}$$

Since the degree of $\mathrm{Tr}(x)$ is $p^m - 1$ and $\mathrm{Tr}(x)$ does not have any duplicate roots, this property is shown. ∎

Thus, the number of non-zero elements in $\mathbb{F}_q$ whose trace is zero given by $q/p - 1$.

Cascaded Trace Function

As mentioned earlier, to improve the drawbacks of the NTU sequence this thesis adopts the sub extension field during the sequence generation procedure. To do so, the conventional introduced in Eq.(2.8) (which maps the extension field $\mathbb{F}_{p^m}$ elements to the prime field $\mathbb{F}_p$ elements) modified as follows:

$$x = \mathrm{Tr}_{p^m|p^{m'}}(X) = \sum_{i=0}^{\frac{m}{m'}-1} X^{p^{im'}}, \tag{2.11}$$

where m' be one of the factors of the extension degree m. In this thesis, the modified trace calculation is so-called the cascaded trace , which maps the extension field $\mathbb{F}_{p^m}$ elements to the sub extension field $\mathbb{F}_{p^{m'}}$ elements (instead of the prime field $\mathbb{F}_p$ elements in the convention trace calculation) is utilized. In what follows, this cascaded trace will be simply called as the trace.

2.6 Quadratic Residue and Quadratic Non-residue

Assume that an integer a and a prime number p are relatively prime to each other. Here, a is said to be a quadratic residue (QR) and quadratic non-residue (QNR) (modulo p) if the following congruential expression has a solution or not, respectively.

$$x^2 \equiv a \pmod{p}.$$

Such property is so called a quadratic residue property. In this dissertation, the Legendre symbol for a element a over the prime field $\mathbb{F}_p$ return the values 1, −1, and 0 if a has a quadratic residue, quadratic non-residue, and equal to 0,

respectively. It is defined as,

$$\left(\frac{a}{p}\right) = a^{\frac{p-1}{2}} \bmod p = \begin{cases} 0, & \text{if } a = 0 \\ 1, & \text{else if } a \text{ is non-zero QR} \\ -1, & \text{otherwise } a \text{ is QNR} \end{cases} \tag{2.12}$$

Property 3 *Let a and b be non-zero elements in the prime field $\mathbb{F}_p$, then the power residue symbol holds the following relation as,*

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right). \tag{2.13}$$

■

2.6.1 Power Residue and Power Non-residue

Let us assume that $a\,(a \geq 1)$ be an element of the finite field $\mathbb{F}_p$. Then, a is said to be a k-th power residue and power non-residue when a has k-th root in $\mathbb{F}_p$ or not, respectively, where k is a prime factor of $p-1$ such that $k \mid (p-1)$. It return the values 1, ϵ_k^i, and 0 if a is a k-th power residue, power non-residue, equal to 0, respectively. The ϵ_k is a primitive k-th root of unity that exists in $\mathbb{F}_p$ and $0 \leq i < k$. In this dissertation, to generate a multi-value sequence (including binary sequence), the Legendre symbol is generalized as follows.

$$\left(\frac{a}{p}\right)_k = a^{\frac{p-1}{k}} \bmod p = \begin{cases} 0, & \text{if } a = 0 \\ 1, & \text{else if } a \text{ is non-zero } k\text{-th PR} \\ \epsilon_k^i, & \text{otherwise } a \text{ is } k\text{-th PNR} \end{cases} \tag{2.14}$$

where PR and PNR stand for Power Residue and Power Non-Residue respectively. It becomes a Legendre symbol when $k = 2$ [Ber15]. Note that, for a non-zero element a and a fixed ϵ_k, the exponent i in Eq.(2.14) is uniquely determined in the range of $0 \sim k-1$. Moreover, since $\epsilon_k^k = \epsilon_k^0 = 1$ and k is a prime number in this thesis, the exponents can be dealt with as elements in $\mathbb{F}_k$. This symbol is basically used for checking whether or not a is a k-th PR over $\mathbb{F}_q$ as shown above.

To represent the exponent i in Eq.(2.14), following notation is used.

$$i = \log_{\epsilon_k}\left(\left(\frac{a}{p}\right)_k\right) = \log_{\epsilon_k}\left(a^{(p-1)/k} \bmod p\right). \tag{2.15}$$

This power residue symbol maps an element in $\mathbb{F}_p$ to an element in $\mathbb{F}_k$. Regarding the power residue symbol $\left(\frac{a}{p}\right)_k$, the following property holds.

Property 4 *For each i from 0 to $k-1$, the number of non-zero elements in $\mathbb{F}_p$ such that*

$$\left(\frac{a}{p}\right)_k = \epsilon_k^i \tag{2.16}$$

is given by $(p-1)/k$. ■

(Proof) Non-zero elements in $\mathbb{F}_p$ are the roots of $x^{p-1} - 1$ over $\mathbb{F}_p$ without any duplicates. Since it is factorized as

$$x^{p-1} - 1 = \prod_{i=0}^{k-1} \left(x^{(p-1)/k} - \epsilon_k^i \right), \tag{2.17}$$

it is thus found that the number is given by $(p-1)/k$. ∎

2.7 Dual Bases

Dual basis that is used for some proofs shown in this thesis is defined as follows:

Definition 18 *Let $\mathcal{A} = \{\alpha_0, \alpha_1, \ldots, \alpha_{m-1}\}$ be a basis in $\mathbb{F}_{p^m}$, the basis $\mathcal{B} = \{\beta_0, \beta_1, \ldots, \beta_{m-1}\}$ such that*

$$\mathrm{Tr}\left(\alpha_i \beta_j\right) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise} \end{cases} \tag{2.18}$$

is called a dual basis of $\mathcal{A}$. ∎

The dual basis of an arbitrary basis is uniquely determined in [LN86]. In this thesis, the following property is important.

Property 5 *Let $\mathcal{A}$ and $\mathcal{B}$ be a basis and its dual basis of $\mathbb{F}_p$ over $\mathbb{F}_{p^m}$, respectively. Based on the definition of dual basis and the linearity of the trace function, if α_l be a basis of $\mathcal{A}$ in $\mathbb{F}_{p^m}$ is a non-zero prime field element, then,*

$$\mathrm{Tr}\left(\alpha_l \beta_j\right) = \alpha_l \mathrm{Tr}\left(\beta_j\right) = \begin{cases} 1, & \text{if } j = l, \\ 0, & \text{otherwise,} \end{cases} \tag{2.19}$$

where $0 \le l, j \le m-1$. Thus, when $\alpha_l = 1$, $\mathrm{Tr}_{p^m|q}(\beta_l) = 1$. ∎

(Proof) Based on the definition of the dual basis,

$$\mathrm{Tr}\left(\alpha_l \beta_j\right) = \begin{cases} 1, & \text{if } j = l, \\ 0, & \text{otherwise.} \end{cases} \tag{2.20}$$

Since the trace function has a linearity for a non-zero prime field element α_l such as,

$$\mathrm{Tr}\left(\alpha_l \beta_j\right) = \alpha_l \mathrm{Tr}\left(\beta_j\right). \tag{2.21}$$

Thus, we obtain Eq.(4.2). Especially for $j = l$,

$$\mathrm{Tr}\left(\alpha_l \beta_l\right) = \alpha_l \mathrm{Tr}\left(\beta_l\right) = 1. \tag{2.22}$$

Therefore, when $\alpha_l = 1$, $\mathrm{Tr}\left(\beta_l\right)$ is determined by 1. ∎

2.8 Pseudo Random Sequence and Its Properties

Sequences, which are generated by deterministic[3] algorithms, which takes seed as input and generate an output sequence of any desired length, is so called a pseudo random sequence. It will be indistinguishable from a true random sequence and should have some statistical properties such as long period, low correlation, high linear complexity, and uniform distribution of bit patterns [LG12; GG05; Rai86]. Let $\mathcal{S}$ be a pseudo random sequence, then $\mathcal{S}$ can be denoted as follows

$$\mathcal{S} = \{s_i\}, i = 0, 1, 2, \ldots, n-1, \ldots,$$

As pseudo random sequence, this dissertation considers both binary sequence and multi-value sequence and they are distinguished by the parameter k. More specifically, for binary case, $k = 2$ and $s_i \in \{0, 1\}$ and for multi-value case, $k > 2$ and $s_i \in \{0, 1, \ldots, k-1\}$. Above mentioned properties regarding a sequence are briefly introduced in the subsequent sections.

2.8.1 Period

The period of a pseudo random sequence is the shortest number of steps until the sequence begins to repeat. Repetitive use of the same period in same application (such as in stream cipher, where encryption is performed by the addition modulo two, thus all the strength of the encryption depends secrecy of the pseudo random sequence) brings security vulnerabilities. Considering the security threat, a basic requirement that a pseudo random sequence (which used for cryptographic applications) should have large period.

Definition 19 (Period) *A sequence $\mathcal{S} = (s_0, s_1, \ldots)$ is eventually periodic if there exists $n \in \mathbb{Z}, n > 0$, and $u \geq 0$ such that*

$$s_{i+n} = s_i \quad \text{for all} \quad i > u.$$

If $u = 0$, $\mathcal{S}$ is strictly periodic. The smallest n that satisfies the above equation is called the period of $\mathcal{S}$. ∎

2.8.2 Autocorrelation and Cross-correlation

The autocorrelation of a sequence shows how a certain sequence period correlates with the same sequence when it is shifted by x bits. The autocorrelation $\mathrm{R}_{\mathcal{S}}(x)$ of binary sequence $\mathcal{S} = \{s_i\}$ having a period of n shifted by x is generally defined as follows. Here, $s_i \in \{0, 1\}$.

$$\mathrm{R}_{\mathcal{S}}(x) = \sum_{i=0}^{n-1} (-1)^{s_{i+x} - s_i}. \tag{2.23}$$

[3] *Deterministic* here means that given the same initial seed, the generator will always produce the same output sequence.

In the case of multi-value sequence, the autocorrelation $\mathrm{R}_{\mathcal{S}}(x)$ evaluation equation becomes as follows.

$$\mathrm{R}_{\mathcal{S}}(x) = \sum_{i=0}^{n-1} \tilde{\epsilon}_k^{s_{i+x}-s_i}, \tag{2.24}$$

where $s_i \in \{0, 1, 2, \ldots, k-1\}$ and $\tilde{\epsilon}_k$ is a primitive k-th root of unity over the complex number $\mathbb{C}$ and it also follows that

$$\mathrm{R}_{\mathcal{S}}(0) = \sum_{i=0}^{n-1} \tilde{\epsilon}_k^{0} = n. \tag{2.25}$$

The autocorrelation evaluation equation for multi-value sequence is a generalized equation and it can be used for binary sequence also, because of, when $k = 2$ it corresponds to a binary sequence. After evaluating the autocorrelation by using the Eq.(2.24), it can be represented graphically. As an example, **Figure** 2.3 shows the autocorrelation of an M-sequence, whereas an M-sequence is well known for its minimum autocorrelation.

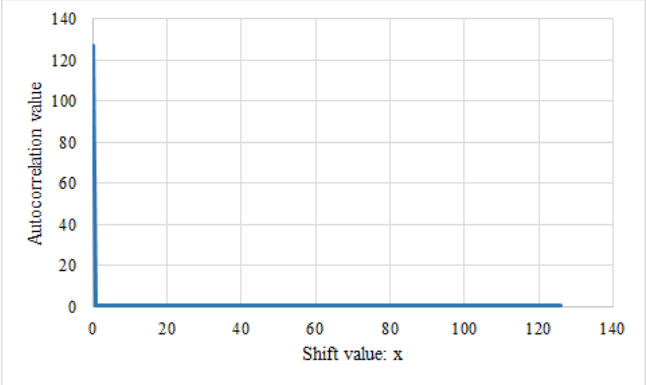

FIGURE 2.3: Autocorrelation of M-sequence.

The autocorrelation of a sequence explains the correlation for the same sequence when it is shifted by x bits. On the other hand, the **cross-correlation** explains how two different sequences of having a same period are correlates to each other when a sequence is shifted by x bits. Let $\mathcal{S} = \{s_i\}$ and $\mathcal{T} = \{t_i\}$ be different sequences of having the same period of n, then the cross-correlation between these two sequences shifted by x is generally defined as follows.

$$\mathrm{R}_{\mathcal{S},\mathcal{T}}(x) = \sum_{i=0}^{n-1} \tilde{\epsilon}_k^{s_{i+x}-t_i}, \tag{2.26}$$

where, $s_i, t_i \in \{0, 1, 2, \ldots, k-1\}$ and $\tilde{\epsilon}_k$ is a primitive k-th root of unity over the complex number $\mathbb{C}$. Alike the generalized autocorrelation evaluation equation introduced in Eq.(2.24), the above cross-correlation equation can be also used for binary case due to when $k = 2$, it corresponds to binary sequence.

2.8.3 Linear Complexity

The linear complexity (LC) of a sequence $\mathcal{S}$ is an important characteristic to judge its quality. Basically, LC is defined as the length of the shortest linear feedback shift register that can generate the sequence [LN96]. In other words, the linear complexity of a sequence is closely related to how difficult it is to guess the next bit after observing the bits of a sequence. Since this dissertation considers both binary and multi-value sequence, therefore generalized equation for evaluating the linear complexity of a sequence $\mathcal{S}$ having a period of n is defined as follows.

$$\mathrm{LC}(\mathcal{S}) = n - \deg\left(\gcd\left(x^n - 1, h_{\mathcal{S}}\left(x\right)\right)\right), \tag{2.27}$$

where $h_{\mathcal{S}}\left(x\right)$ of $\mathcal{S} = \{s_i\}$ is defined over $\mathbb{F}_k$ as,

$$h_{\mathcal{S}}(x) = \sum_{i=0}^{n-1} s_i x^i. \tag{2.28}$$

It should be noted that $\gcd(x^n - 1, h_{\mathcal{S}}(s))$ in Eq.(6.1) needs to be calculated over $\mathbb{F}_k$. Furthermore, in case of binary and multi-value sequence, the value of k becomes $\{0, 1\}$ and $\{0, 1, \ldots, k-1\}$, respectively. Pseudo random sequences with high linear complexity ($\mathrm{LC}(\mathcal{S}) > \mathrm{n}/2$, where n denotes the period of the sequence) are important for security applications. The linear complexity regarding a sequence can be represented graphically, as an example, the linear complexity of Legendre sequence is shown in **Figure** 2.4. It should be be noted that Legendre sequence is well known for its high linear complexity.

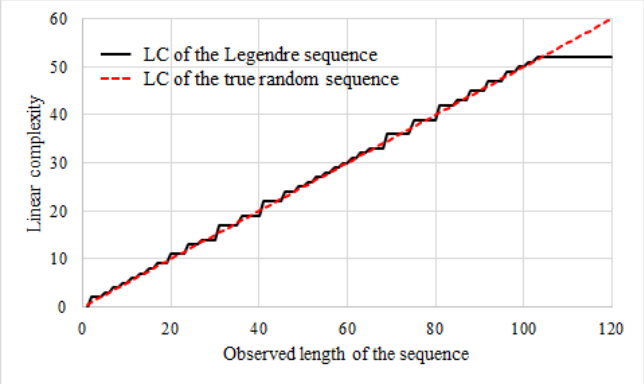

FIGURE 2.4: Linear complexity profile of Legendre sequence.

2.8.4 Distribution of Bit Patterns

From the viewpoint of security, the distribution of bit patterns is as important as the linear complexity. If a sequence holds uniform distribution of bit patterns, then it becomes difficult to guess the next bit after observing the previous bit patterns. For example, assume a binary sequence having a period of 12 as

$\mathcal{S}_{12} = \{1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0\}$. This sequence has uniform distribution of 1 and 0, considering the 1-bit pattern. In other words, 1 and 0 appears same in number. However, in case of 2-bit patterns on $\mathcal{S}_{12}$, it has two type of patterns (10 and 01) only. In this case, it is easy to predict the next bit patterns after observing the previous patterns. Therefore, it is also essential to evaluate the distribution of bit patterns of a sequence to confirm its randomness. An M-sequence is well known for its uniform distribution of bit patterns, therefore the bit distribution of M-sequence is shown in **Table** 2.1. In what follows, the notation $b^{(n)}$ in the remaining bit distribution tables is omitted.

TABLE 2.1: Bit distribution of M-sequence.[4]

n	$b^{(n)}$	$\mathrm{H_w}(b^{(n)})$	$Z(b^{(n)})$	$D_S(b^{(n)})$
1	0	0	1	7
	1	1	0	8
2	00	0	2	3
	01	1	1	4
	11	2	0	4
3	000	0	3	1
	001	1	2	2
	011	2	1	2
	111	3	0	2

2.9 Use Case of Pseudo Random Sequence

One of the most common applications of the pseudo random binary sequence is in a stream cipher. The basic structure of a stream cipher is shown in **Figure** 2.5. Basically, a stream cipher is divided into two classes: block cipher and stream cipher. Among these in case of a block cipher, the same key is used for both encryption and decryption of each block ($\geq$ 64 bits) of data. On the other hand, in case of a stream cipher, encryption and decryption are performed by the bitwise $\oplus$ (XOR) operation with a key stream. This section restricts the discussion of the proposed pseudo-random binary sequence in a stream cipher.

Few important considerations during the design of a stream cipher are the key (which used for both encryption and decryption) should have a large period, good randomness, and unpredictability properties due to the usage of the same key in both encryption and decryption. Here, the encryption is carried out by applying a bit-wise $\oplus$ (XOR) operation between the plain-text of byte stream

[4]In this table $b^{(n)}$, $\mathrm{H_w}(b^{(n)})$, $Z(b^{(n)})$ and $D_{\mathcal{S}}(b^{(n)})$ denotes a bit pattern of length n, the hamming weight of $b^{(n)}$, number of 0's in $b^{(n)}$, and number of appearance of $b^{(n)}$ in $\mathcal{S}$, respectively. For example, in a binary sequence of period 15, a 3-bit pattern $b = 101$ appears 4 times. Then, these notations become $b^{(3)} = 101$, $Z(b^{(3)}) = 1$, and $D_{\mathcal{S}}(b^{(3)}) = 4$.

M and encryption key K. Then, the cipher-text C transmitted through a network. On the other hand, during the decryption, after the bit-wise $\oplus$ operation between the cipher-text C and the same key K we will get the original plain-text M. In a stream cipher, a lot of sequences are assigned to several users, respectively. If these sequences have some correlation, then it will make some security vulnerabilities. Under this circumstance, it is important to observe the cross-correlation property between several sequences. Additionally, its linear complexity and distribution of bit patterns needs to be high and uniform, respectively to confirm its randomness. The proposed method in this dissertation can generate a long period pseudo random sequence with typical auto and cross-correlation, high linear complexity, and almost uniformly distributed bit patterns. After observing the experimental and comparison results from the preceding chapters, it can be concluded that the proposed sequence which defined over the sub extension field can be a prominent candidate for suitable applications (as like the other pseudo random sequence).

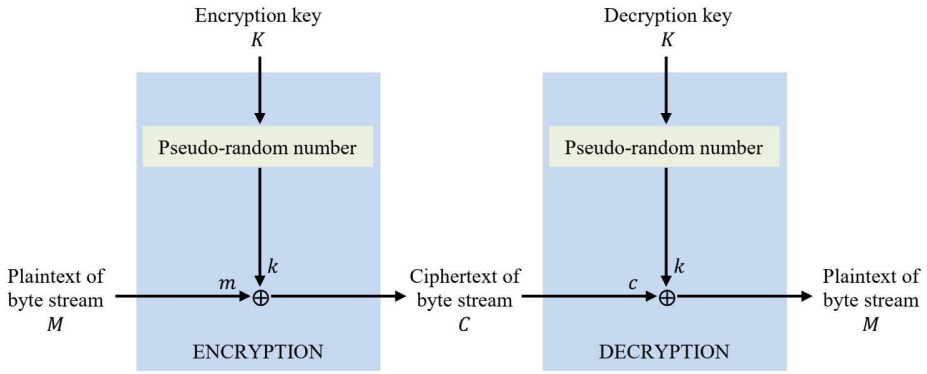

FIGURE 2.5: Use case of the pseudo random sequence in stream cipher.

2.10 Summary

This chapter defined the related mathematical fundamentals and introduced the notations for the subsequent chapters.

Chapter 3

Pseudo Random Sequence

Sequences of numbers generated by using an algorithm is referred to as a pseudo random sequence. This chapter proposes a pseudo random sequence along with NTU (Nogami-Tada-Uehara) sequence [NTU14].

3.1 Preparation

To adopt the sub extension field during the sequence generation procedure, the conventional trace calculation modified, which is so-called the cascaded trace. This section introduces the modified trace function (cascaded trace) and k-th power residue symbol for sub extension field.

3.1.1 Trace Function for Sub Field

This work utilizes this cascaded trace function to map an element of the extension field $X \in \mathbb{F}_{p^m}$ to an element of the sub extension field $x \in \mathbb{F}_q$ as,

$$x = \mathrm{Tr}_{p^m|q}(X) = \sum_{i=0}^{\frac{m}{m'}-1} X^{p^{im'}}. \tag{3.1}$$

A crucial point, the above cascaded trace becomes an arbitrary element in $\mathbb{F}_q$ and the trace function has a linearity property over the sub extension field $\mathbb{F}_q$ as follows,

$$\mathrm{Tr}_{p^m|q}(aX + bY) = a\mathrm{Tr}_{p^m|q}(X) + b\mathrm{Tr}_{p^m|q}(Y), \tag{3.2}$$

where $a, b \in \mathbb{F}_q$ and $X, Y \in \mathbb{F}_{p^m}$. In this chapter, the following property is important [LN86].

Property 6 For each arbitrary element $\alpha \in \mathbb{F}_q$, the number of elements in $\mathbb{F}_{p^m}$ whose trace with respect to $\mathbb{F}_q$ becomes α is given by $p^m/q = p^{m-m'}$ and the number of non-zero elements in $\mathbb{F}_{p^m}$ whose trace is zero is given by $(p^m/q) - 1 = p^{m-m'} - 1$. ∎

3.1.2 k-th Power Residue Symbol for Sub Field

As an extension of the Legendre symbol, this chapter considers the k-th power residue symbol $\left(a/q\right)_k$ for an arbitrary element a in $\mathbb{F}_q$ and a prime factor k of $q-1$ as follows:

$$\begin{aligned}\left(\frac{a}{q}\right)_k &= a^{(q-1)/k} \\ &= \begin{cases} 0, & \text{if } a = 0, \\ 1 = \epsilon_k^0, & \text{else if } a \text{ is a } k\text{-th PR in } \mathbb{F}_q^*, \\ \epsilon_k^i, & \text{otherwise } a \text{ is a } k\text{-th PNR in } \mathbb{F}_q^*, \end{cases}\end{aligned} \tag{3.3}$$

where PR and PNR stand for Power Residue and Power Non-Residue respectively. The ϵ_k is a primitive k-th root of unity that exists in $\mathbb{F}_q$ and $0 \le i < k$. It becomes a Legendre symbol when $k = 2$ [Ber15] and if $k|(p-1)$, it becomes the previous work [Ali+16b]. Note that, for a non-zero element a and a fixed ϵ_k, the exponent i in Eq.(3.3) is uniquely determined in the range of $0 \sim k-1$. Moreover, since $\epsilon_k^k = \epsilon_k^0 = 1$ and k is a prime number in this thesis, the exponents can be dealt with as elements in $\mathbb{F}_k$. This symbol is basically used for checking whether or not a is a k-th PR over $\mathbb{F}_q$ as shown above. The output of the k-th power residue symbol can be represented as an exponent of ϵ_k, where ϵ_k is a k-th primitive root. This thesis uses k-th power residue symbol to translate a trace sequence over $\mathbb{F}_q$ to a k values multi-value sequence such as $\{0, \epsilon_k^i\}$, where $i \in \{0, \ldots, k-1\}$.

To represent the exponent i in Eq.(3.3), this chapter uses the following notations and it should be noted that the following notation excludes the case of $a = 0$.

$$i = \log_{\epsilon_k}\left(\left(\frac{a}{q}\right)_k\right) = \log_{\epsilon_k}\left(a^{(q-1)/k}\right). \tag{3.4}$$

This chapter utilizes the power residue symbol to map an element in $\mathbb{F}_q$ to an element in $\mathbb{F}_k$. Regarding the power residue symbol $\left(a/q\right)_k$, the following property holds.

Property 7 *For each i from 0 to $k-1$, the number of non-zero elements in $\mathbb{F}_q$ such that*

$$\left(\frac{a}{q}\right)_k = \epsilon_k^i \tag{3.5}$$

is given by $(q-1)/k$. ∎

(Proof) Non-zero elements in $\mathbb{F}_q$ are the roots of $x^{q-1} - 1$ over $\mathbb{F}_q$ without any duplicates. Since it is factorized as

$$x^{q-1} - 1 = \prod_{i=0}^{k-1}\left(x^{(q-1)/k} - \epsilon_k^i\right), \tag{3.6}$$

it is thus found that the number is given by $(q-1)/k$. ∎

3.2 Proposed Pseudo Random Sequence

In this thesis, the proposed pseudo random sequence is generated by combining the features of an M-sequence and Legendre sequence. The purpose of combining these two well-known sequences is to adopt their good features (such as uniform distribution of bit patterns and maximum linear complexity from the M-sequence and Legendre sequence, respectively). Fortunately, the proposed sequence holds these good features.

The proposed pseudo random sequences (including both binary and multi-value sequence) are defined over the sub extension field. It is generated by applying a primitive polynomial, trace function, and Legendre symbol (or k-th power residue symbol, it is an extended version of the Legendre symbol) over the sub extension field $\mathbb{F}_q$. In details, the proposed sequence generation procedure is as follows: let p be an odd characteristic prime and m be the extension degree of a primitive polynomial $f(x)$ over the extension field $\mathbb{F}_{p^m}$. It is well known that using the primitive polynomial makes it possible to generate a maximum length vector sequence over $\mathbb{F}_{p^m}$. Let ω be a zero of the primitive polynomial $f(x)$ and it is a primitive element in $\mathbb{F}^*_{p^m}$. Then the sequence

$$\mathcal{T} = \{t_i \mid t_i = \mathrm{Tr}_{p^m|q}\left(\omega^i\right), i = 0, 1, 2, \ldots, p^m - 2\}$$

becomes a maximum length sequence of having a period of $p^m - 1$, where, p and q denote an odd prime number and its power $q = p^{m'}$, respectively, m' be one of the factors of the extension degree m, and $\mathrm{Tr}_{p^m|q}(\cdot)$ is the trace function over the sub extension field $\mathbb{F}_q$.

In the beginning, a primitive polynomial $f(x)$ generates a maximum length vector sequence over the extension field $\mathbb{F}_{p^m}$, then the trace function $\mathrm{Tr}_{p^m|q}(\cdot)$ maps an element of the extension field $\mathbb{F}_{p^m}$ to an element of the sub extension field $\mathbb{F}_q$. After the trace calculation, a non-zero constant element A is added to the trace values. This non-zero A can be any arbitrary element within the sub extension field $\mathbb{F}_q$ such as $A \in \{1, 2, \ldots, q-1\}$. Finally, the Legendre symbol (or k-th power residue symbol) maps the sub extension field elements generated by the trace function to a binary sequence (or multi-value sequence). Therefore, the generalized equation of the proposed pseudo random sequence in this thesis becomes as follows:

$$\mathcal{S} = \{s_i\}, s_i = f_k\left(\mathrm{Tr}_{p^m|q}\left(\omega^i\right) + A \Big/ q\right)_k. \tag{3.7}$$

Here k is a prime number as well as a factor of $q - 1$ such as $k \mid (q - 1)$. The sequence values s_i can be described by the exponent of ϵ_k such as ϵ_k^e, where ϵ_k is a primitive k-th root of unity that exists in $\mathbb{F}_q$. Finally, a mapping function $f_k(\cdot)$ is used to translate the vector sequence to a k-value sequence. This mapping

function $f_k(\cdot)$ is defined as follows:

$$f_k(x) = \begin{cases} 0, & \text{if } x = 0, \\ \log_{\epsilon_k}\left(\left(x/p\right)_k\right), & \text{otherwise.} \end{cases} \tag{3.8}$$

The period n of the proposed sequence is given by

$$n = p^m - 1. \tag{3.9}$$

In this thesis, both binary and multi-value sequences can be distinguished by the parameter k. For example, for binary case, $k = 2$ and sequence coefficients becomes {0 and 1}, on the other hand, for multi-value case, $k > 2$ and sequence coefficients becomes $\{0, 1, \ldots, k-1\}$.

The forthcoming chapters describe the proposed pseudo random sequence properties such as its period, autocorrelation, cross-correlation, distribution of bit patterns, and linear complexity.

3.3 NTU Sequence

The study of pseudo random sequence by combining the features of an M-sequence and Legendre sequence started from the NTU (Nogami-Tada-Uehara) sequence [NTU14]. It is generated by utilizing a primitive polynomial, trace function, and Legendre symbol. Moreover, the NTU sequence is defined over the prime field $\mathbb{F}_p$.

In brief, the NTU sequence is generated as follows: let p be an odd characteristic and m be the degree of primitive polynomial $f(x)$ over the prime field $\mathbb{F}_p$. A primitive polynomial generates a maximum length sequence of vectors, next the trace function maps these vectors as scalars (as elements of the prime field $\mathbb{F}_p$), the Legendre symbol binarize the trace outputs. Thus, the generalized equation of the NTU sequence is as follows:

$$\mathcal{S}_{\mathrm{NTU}} = \{s_i\}, s_i = M_2\left(\left(\mathrm{Tr}_{p^m|p}\left(\omega^i\right)/p\right)\right), \tag{3.10}$$

where ω is a primitive element in the extension field $\mathbb{F}_{p^m}$ and the mapping function $M_2(\cdot)$ is defined as follows.

$$M_2(x) = \begin{cases} 0, & \text{if } x = 0, 1 \bmod p, \\ 1, & \text{else if } x = -1 \bmod p. \end{cases} \tag{3.11}$$

The period n of the NTU sequence is given by

$$n = \frac{2\left(p^m - 1\right)}{p - 1}. \tag{3.12}$$

As a small example, when the parameters are set as $p = 7$, $m = 2$, and $f(x) = x^2 + 6x + 3$, then the generated sequence becomes as follows:

$$\mathcal{S} = \{0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 1\}. \tag{3.13}$$

Although NTU sequence holds interesting features (such as period, autocorrelation, distribution of bit patterns, and linear complexity) to be compatible with other pseudo random sequences, there exists some scope of its improvements to make it a perfect sequence. As for drawbacks of NTU sequence, its distribution of bit patterns does not become uniform (as shown in Table 3.1)and in terms of autocorrelation there exists a small difference between the maximum peak and other peaks (as shown in **Figure** 3.1). The foremost contribution of this thesis is to overcome these drawbacks of the NTU sequence, as well as to mathematically prove these properties.

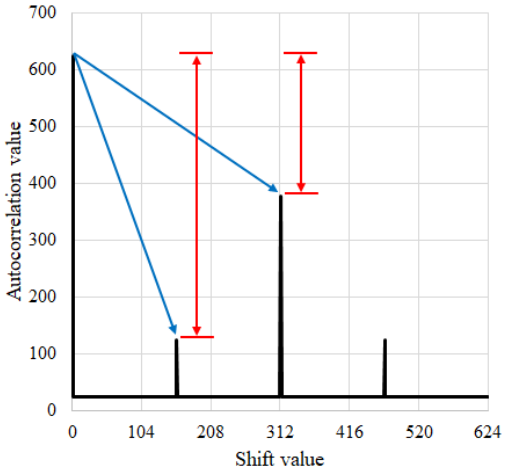

FIGURE 3.1: Autocorrelation of the NTU sequence.

The proposed sequence in this thesis combines the features of an M-sequence and Legendre sequence, as well as, it also extends the NTU sequence by considering the sub extension field during the sequence generation procedure. In addition, the proposed sequence also includes one more additional parameter A (non-zero arbitrary element in the sub extension field). The consideration of the sub extension field extends the mapping area just before the binarization, as shown in **Figure** 3.2. This extended mapping area contributes to the betterment of the proposed sequence properties (such as its correlation and distribution of bit patterns) compared to the NTU sequence. The details of these improvements are introduced in the forthcoming chapters.

TABLE 3.1: Bit distribution of the NTU sequence.

n	Bit patterns $b^{(n)}$	Appearance $D_{\mathcal{S}}(b^{(n)})$	%
1	0	156865	56.93
	1	117649	43.07
2	00	89637	32.53
	01	67228	24.40
	11	50421	18.30
3	000	51221	18.59
	001	38416	13.94
	011	28812	10.45
	111	21609	7.84
4	0000	29269	10.62
	0001	21952	7.96
	0011	16464	5.97
	0111	12348	4.48
	1111	9261	3.36

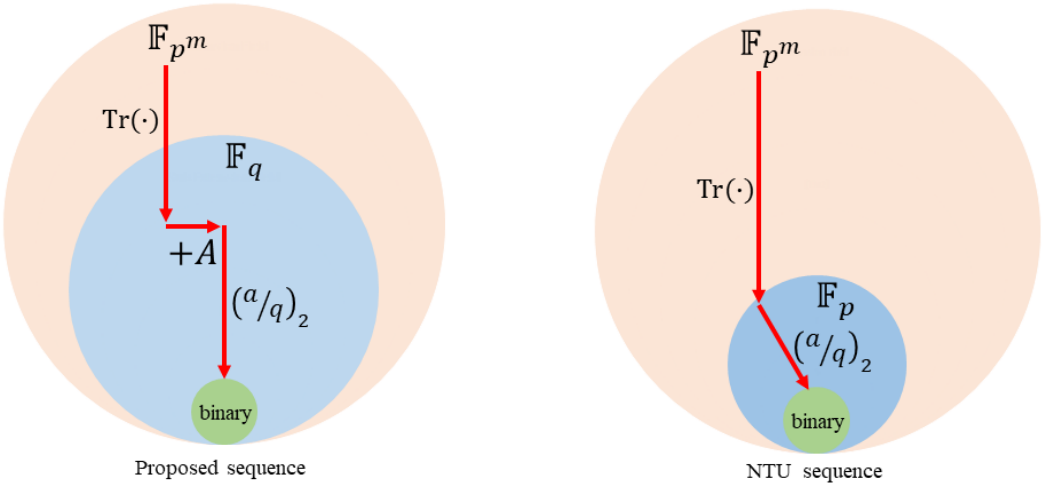

FIGURE 3.2: Difference in sequence generation procedure between the proposed sequence and NTU sequence.

Chapter 4

Period and Correlation of Pseudo Random Sequence

4.1 Introduction

Background and Motivation

Pseudo random sequences are inseparable parts in information technology as well as in the modern electronics. They are used in both communication (such as cellular telephones and GPS signals) and cryptographic applications (such as keystream for stream cipher, sampling data for simulations, timing measurements in radar systems, error correcting codes in satellite communications, and so on). In most cases it is important to have the reproduce ability of the pseudo random sequence [GK12]. As well as it should have many desirable characteristics such as a long period and low correlation [CD09], and statistical randomness [Raf17] to become a prominent candidate for information security and cryptographic related applications [Gol67; MVO96]. The randomness regarding a sequence considered as the key strength of the cryptographic systems [Oma+18]. The major substance for randomness is independency of values (or lack of correlation) [YG06; XQ06], unpredictability (or lack of predictability), and uniform distribution (or lack of bias) [Kin+12].

This chapter focuses on the period, autocorrelation, and cross-correlation properties regarding a pseudo random sequence, which includes both binary sequence and multi-value sequence. Both binary and multi-value sequences can be distinguished by the parameter k. More specifically, for binary case, $k = 2$ and sequence coefficients becomes {0 and 1}, on the other hand, for multi-value case, $k > 2$ and sequence coefficients becomes $\{0, 1, \ldots, k-1\}$.

Related Works

Other prominent geometric sequences having theoretical aspect are the Legendre sequence [Zie58], maximum length sequence (M-sequence) [Nea59], and Sidelnikov sequence [Sid71]. Generally, the typical features of a pseudo random sequence cannot be theoretically proven. However, if a sequence is defined over

the finite field, then those features are often proven. All these above-mentioned sequences generated based on some mathematics more specifically they are defined over the finite field. Therefore, most of their important properties are already theoretically proven.

Previous binary sequence [NTU14] generated by combining the features of M-sequence and L-sequence and its period, autocorrelation, and linear complexity properties have been theoretically proven. Previous works on multi-value sequence [Nog+16; Ali+16b] also defined over the prime field and its period, autocorrelation, and cross-correlation properties theoretically proven. Previous work on multi-value sequence [Ali+17b] considered on the sub extension field $\mathbb{F}_q$ characterized by four parameters however it has a shorter sequence period of

$$n = \frac{k(p^m - 1)}{q - 1}.$$

The period and autocorrelation properties of the proposed sequence explained based on some experimental results only.

Contribution

This chapter restricts the discussion on the period and correlation properties of a pseudo random sequence (including both binary sequence and multi-value sequence) which defined over the sub extension field $\mathbb{F}_q$. This chapter considers the sub extension field $\mathbb{F}_q$, whereas, previous works on NTU_class sequence [NTU14; Nog+16; Ali+16b] are considered in the prime field $\mathbb{F}_p$. The proposal of this chapter is an extension of the previous works [NTU14; Nog+16; Ali+16b] and if k satisfies the condition $k|(p-1)$, then it also includes the previous work [Ali+16b]. This proposal overcomes the shorter period shortcoming of the previous work [Ali+17b] by adding one more additional parameter A. In addition, the period, autocorrelation, and cross-correlation properties regarding the proposed sequence are explained both theoretically and experimentally. It also makes a comparison in terms of autocorrelation and it was found that the proposed sequence holds low correlation compared to the previous work [Nog+16].

4.2 Preparation

In order to give some theoretic proofs, the dual bases over sub extension field needs to be introduced.

4.2.1 Dual Bases for Sub Field

Dual basis that is used for some proofs shown in this chapter is defined as,

Definition 20 *$\mathbb{F}_{p^m}$ be a finite field and $\mathbb{F}_q$ be a finite extension of $\mathbb{F}_{p^m}$. Then the two bases $\mathcal{A} = \{\alpha_0, \alpha_1, \ldots, \alpha_{m-1}\}$ and $\mathcal{B} = \{\beta_0, \beta_1, \ldots, \beta_{m-1}\}$ of $\mathbb{F}_q$ over $\mathbb{F}_{p^m}$ are said to be the dual (or complementary) bases if*

$$\mathrm{Tr}_{p^m|q}\left(\alpha_i\beta_j\right) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise} \end{cases} \tag{4.1}$$

where $1 \le i, j \le m-1$. ■

The dual basis of an arbitrary basis is uniquely determined in [LN86]. In this chapter, the following property is important.

Property 8 *Let $\mathcal{A}$ and $\mathcal{B}$ be a basis and its dual basis of $\mathbb{F}_q$ over $\mathbb{F}_{p^m}$, respectively. Based on the definition of dual basis and the linearity of the trace function, if α_l be a basis of $\mathcal{A}$ in $\mathbb{F}_{p^m}$ is a non-zero sub extension field element, then,*

$$\mathrm{Tr}_{p^m|q}\left(\alpha_l\beta_j\right) = \alpha_l\mathrm{Tr}_{p^m|q}\left(\beta_j\right) = \begin{cases} 1, & \text{if } j = l, \\ 0, & \text{otherwise,} \end{cases} \tag{4.2}$$

where $0 \le l, j \le m-1$. Thus, when $\alpha_l = 1$, $\mathrm{Tr}_{p^m|q}(\beta_l) = 1$. ■

(Proof) Based on the definition of the dual basis,

$$\mathrm{Tr}_{p^m|q}\left(\alpha_l\beta_j\right) = \begin{cases} 1, & \text{if } j = l, \\ 0, & \text{otherwise.} \end{cases} \tag{4.3}$$

Since the trace function has a linearity for a non-zero sub extension field element α_l such as,

$$\mathrm{Tr}_{p^m|q}\left(\alpha_l\beta_j\right) = \alpha_l\mathrm{Tr}_{p^m|q}\left(\beta_j\right). \tag{4.4}$$

Thus, we obtain Eq.(4.2). Especially for $j = l$,

$$\mathrm{Tr}_{p^m|q}\left(\alpha_l\beta_l\right) = \alpha_l\mathrm{Tr}_{p^m|q}\left(\beta_l\right) = 1. \tag{4.5}$$

Therefore, when $\alpha_l = 1$, $\mathrm{Tr}_{p^m|q}(\beta_l)$ is determined by 1. ■

4.3 Cross-correlation

This section explains the cross-correlation property theoretically, before going to the details of it, this section begins with the idea of how the proposed sequence generated. Let ω be a primitive element in the extension field $\mathbb{F}_{p^m}$, n be the period of the proposed sequence, m be a composite number which denotes the extension degree of the primitive polynomial, and m' be one of the factors of m. This thesis proposes the following sequence $\mathcal{S}$ by utilizing the trace function and k-th power residue symbol as follows:

$$\mathcal{S} = \{s_i\}, s_i = f_k\left(\mathrm{Tr}_{p^m|q}\left(\omega^i\right) + A\middle/q\right)_k. \tag{4.6}$$

Here k is a prime number as well as a factor of $q-1$ such as $k \mid (q-1)$. To make the above equation more simpler, from here on $\mathrm{Tr}_{p^m|q}(\cdot)$ will be represented as $\mathrm{Tr}(\cdot)$. Therefore, the above equation becomes,

$$\mathcal{S} = \{s_i\}, s_i = f_k\left(\mathrm{Tr}\left(\omega^i\right) + A\Big/q\right)_k. \tag{4.7}$$

The sequence values s_i can be described by the exponent of ϵ_k such as ϵ_k^e, where ϵ_k is a primitive k-th root of unity that exists in $\mathbb{F}_q$. Considering the sub extension field $\mathbb{F}_q$, here ϵ_k is a vector rather than a scalar. For example, when $p = 5$, $k = 3$, and $f(x) = x^4 + 3x^3 + x + 2$ be a primitive polynomial over $\mathbb{F}_5$. Let ω be its zero, then the 3-rd primitive root vectors in $\mathbb{F}_5$ becomes $2\omega^2 + 3\omega + 1$ and $3\omega^2 + 2\omega + 3$. In this example, let us fix $2\omega^2 + 3\omega + 1$ as a primitive 3rd root. Then $2\omega^2 + 3\omega + 1$ and $3\omega^2 + 2\omega + 3$ vectors can be represented as the exponent of $2\omega^2 + 3\omega + 1$, this relation is developed as follows:

$$2\omega^2 + 3\omega + 1 = \epsilon_3^1. \tag{4.8a}$$

$$3\omega^2 + 2\omega + 3 = \epsilon_3^2. \tag{4.8b}$$

Finally, a mapping function $f_k(\cdot)$ is used to translate the vector sequence generated by the k-th power residue symbol to a k-value sequence. The mapping function $f_k(\cdot)$ is defined as follows:

$$f_k(x) = \begin{cases} 0, & \text{if } x = 0, \\ \log_{\epsilon_k}\left(\left(x/p\right)_k\right), & \text{otherwise.} \end{cases} \tag{4.9}$$

As mentioned in **Section** 3.1.2, $f_k(x)$ with a fixed ϵ_k maps an arbitrary element $x \in \mathbb{F}_q$ to an element in $\mathbb{F}_k$. For example, by utilizing the parameter $p = 5$ and $k = 3$, the sequence values will be in the range of $\{0, 1, 2\}$, all of these values are the elements of $\mathbb{F}_3$. In addition, let us fixed [1 4 3] be as a 3-rd primitive root of unity in $\mathbb{F}_q$. Then, all of the sequence values can be represented as a exponent of this primitive root ϵ_3. More details of this example is shown in Table 5.4. This mapping function $f_k(\cdot)$ holds the following property.

Property 9 Consider $x, y \in \mathbb{F}_q$. If $x \neq 0$ and $y \neq 0$,

$$f_k(x) \pm f_k(y) = f_k\left(xy^{\pm 1}\right). \tag{4.10}$$

■

Based on the **Section** 3.1.2 and **Property** 7, the mapping function also satisfies the following equation, here C is a non-zero element in $\mathbb{F}_q$.

$$\sum_{v=0}^{k-1} \tilde{\epsilon}_k^v = 0. \tag{4.11a}$$

TABLE 4.1: Mapping procedure of $f_k(\cdot)$ for 24 different trace $\mathrm{Tr}(\cdot)$ values.

	Output of $\mathrm{Tr}(\cdot)$	Output of $\left(a/p\right)_k$	ϵ_3 =[1 4 3][1]	Output of $f_k(\cdot)$
1	[0]	[3 1 2]	$[1\ 4\ 3]^2$	2
2	[1]	[1 4 3]	$[1\ 4\ 3]^1$	1
3	[2]	[3 1 2]	$[1\ 4\ 3]^2$	2
4	[4]	[1 4 3]	$[1\ 4\ 3]^1$	1
5	[0 1 2]	[1]	$[1\ 4\ 3]^0$	0
6	[0 2 4]	[1]	$[1\ 4\ 3]^0$	0
7	[0 3 1]	[1 4 3]	$[1\ 4\ 3]^1$	1
8	[0 4 3]	[3 1 2]	$[1\ 4\ 3]^2$	2
9	[1 1 2]	[1]	$[1\ 4\ 3]^0$	0
10	[1 2 4]	[3 1 2]	$[1\ 4\ 3]^2$	2
11	[1 3 1]	[1]	$[1\ 4\ 3]^0$	0
12	[1 4 3]	[3 1 2]	$[1\ 4\ 3]^2$	2
13	[2 1 2]	[1]	$[1\ 4\ 3]^0$	0
14	[2 2 4]	[1 4 3]	$[1\ 4\ 3]^1$	1
15	[2 3 1]	[3 1 2]	$[1\ 4\ 3]^2$	2
16	[2 4 3]	[1]	$[1\ 4\ 3]^0$	0
17	[3 1 2]	[1]	$[1\ 4\ 3]^0$	0
18	[3 2 4]	[3 1 2]	$[1\ 4\ 3]^2$	2
19	[3 3 1]	[3 1 2]	$[1\ 4\ 3]^2$	2
20	[3 4 3]	[1 4 3]	$[1\ 4\ 3]^1$	1
21	[4 1 2]	[0]	0	0
22	[4 2 4]	[1 4 3]	$[1\ 4\ 3]^1$	1
23	[4 3 1]	[1 4 3]	$[1\ 4\ 3]^1$	1
24	[4 4 3]	[1 4 3]	$[1\ 4\ 3]^1$	1

$$\sum_{u=1}^{p-1} \tilde{\epsilon}_k^{f_k(u)} = \sum_{u=1}^{p-1} \tilde{\epsilon}_k^{f_k(u^{-1})} = \left(\frac{p-1}{k}\right) \sum_{v=0}^{k-1} \tilde{\epsilon}_k^{v} = 0. \tag{4.11b}$$

$$\sum_{u=1}^{p-1} \tilde{\epsilon}_k^{f_k(Cu)} = \sum_{u=1}^{p-1} \tilde{\epsilon}_k^{f_k(Cu^{-1})} = 0. \tag{4.11c}$$

This section, firstly mathematically prove the cross-correlation property of the proposed k-value sequence, then it explains the autocorrelation property, and finally the period is introduced. Additionally, these properties are also observed based on some experimental results.

[1]This example fixed [1 4 3] as a primitive 3rd root of unity that exists in $\mathbb{F}_q$. Therefore, every element can be represented as a power of this 3rd primitive root ϵ_3.

The cross-correlation is calculated between two different sequences of having the same period. These two different sequences $\hat{\mathcal{S}}$ and $\mathcal{S}$ can be defined as,

$$\hat{\mathcal{S}} = \left\{ \hat{s}_i \mid \hat{s}_i = f_k\left(\mathrm{Tr}\left(\omega^i\right) + B \Big/ p \right)_k \right\}, \tag{4.12a}$$

$$\mathcal{S} = \left\{ s_i \mid s_i = f_k\left(\mathrm{Tr}\left(\omega^i\right) + A \Big/ p \right)_k \right\}. \tag{4.12b}$$

Here, A and B are non-zero elements in $\mathbb{F}_q$. They can be represented with a generator g that exists in the sub extension field $\mathbb{F}_q$ and they holds the following relation.

$$B = g^h A, \tag{4.13}$$

where the index term h satisfies $0 \leq h \leq q-2$ relation. In addition, here g needs to be given by $\omega^{(p^m-1)/(q-1)}$, which used in the following proofs[2]. The cross-correlation of these two sequences $\hat{\mathcal{S}}$ and $\mathcal{S}$ is calculated as,

$$\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(x) = \sum_{i=0}^{n-1} \tilde{\epsilon}_k^{f_k\left(\mathrm{Tr}\left(\omega^{i+x}\right)+g^h A\right) - f_k\left(\mathrm{Tr}\left(\omega^i\right)+A\right)}, \tag{4.14}$$

where n is the period of these two sequences and according to the following section, it is given by $p^m - 1$. Furthermore, when $h = 0$, then the value of A and B becomes exactly equal to each other, therefore, the cross-correlation becomes the autocorrelation of $\mathcal{S}$.

Theorem 5 *The cross-correlation between the sequence $\hat{\mathcal{S}}$ and $\mathcal{S}$ given by the Eq.(4.14) is as follows.*

$$\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(x) = \begin{cases} p^{m-m'} + \left(p^m - 1 - p^{m-m'}\right) \tilde{\epsilon}_k^{f_k(g^h)}, & \text{if } x = h\bar{n}, \\ p^{m-m'} \left(\tilde{\epsilon}_k^{f_k\left(A(g^h - g^j)\right)} + \tilde{\epsilon}_k^{-f_k\left(A(1-g^{h-j})\right)} - \tilde{\epsilon}_k^{f_k(g^j)} \right) - \tilde{\epsilon}_k^{f_k(g^h)}, & \text{else if } x = j\bar{n}, \\ p^{m-2m'} - \tilde{\epsilon}_k^{f_k(g^h)}, & \text{otherwise,} \end{cases} \tag{4.15}$$

where $\bar{n} = n/(q-1) = (p^m - 1)/(q-1)$ and h satisfies the relation in Eq.(4.13) as well as $0 \leq j \neq h \leq q-2$. ∎

The proof for each case of Eq.(4.15) is explained below. It should be noted that i holds the relation $0 \leq i < n = (p^m - 1)$ and it is mainly appeared at summations. Furthermore, in the following section m/m' is denoted as r.

4.3.1 The Case of $x = h\bar{n}$

In this case, the cross-correlation between the sequences $\hat{\mathcal{S}}$ and $\mathcal{S}$ becomes as follows:

$$\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(x) = \sum_{i=0}^{n-1} \tilde{\epsilon}_k^{f_k\left(\mathrm{Tr}\left(\omega^{i+h\bar{n}}\right)+g^h A\right) - f_k\left(\mathrm{Tr}\left(\omega^i\right)+A\right)} = \sum_{i=0}^{n-1} \tilde{\epsilon}_k^{f_k\left(g^h\left(\mathrm{Tr}\left(\omega^i\right)+A\right)\right) - f_k\left(\mathrm{Tr}\left(\omega^i\right)+A\right)}. \tag{4.16}$$

[2]Since ω is a generator of $\mathbb{F}_{p^m}^*$, therefore $g = \omega^{(p^m-1)/(q-1)}$ becomes a generator of $\mathbb{F}_q^*$.

According to **Property** 9 and depending on the condition of whether or not $\mathrm{Tr}\left(\omega^i\right) + A = 0$, the above equation can be rewritten as follows:

$$\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(x) = \sum_{\mathrm{Tr}\left(\omega^i\right)+A=0} \tilde{\epsilon}_k^{f_k(g^h\cdot 0)-f_k(0)} + \sum_{\mathrm{Tr}\left(\omega^i\right)+A\neq 0} \tilde{\epsilon}_k^{\left(g^h\left(\mathrm{Tr}\left(\omega^i\right)+A\right)\right)-f_k\left(\mathrm{Tr}\left(\omega^i\right)+A\right)}. \tag{4.17}$$

Thus, the above equation becomes as,

$$\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(x) = \sum_{\mathrm{Tr}\left(\omega^i\right)+A=0} \tilde{\epsilon}_k^{0} + \sum_{\mathrm{Tr}\left(\omega^i\right)+A\neq 0} \tilde{\epsilon}_k^{f_k(g^h)}. \tag{4.18}$$

It should be noted that $g^h \neq 0$. Therefore, according to **Property** 6, the cross-correlation between the sequence $\hat{\mathcal{S}}$ and $\mathcal{S}$ for the case of $x = h\bar{n}$ holds the following relation.

$$\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(x) = p^{m-m'} + \left(p^m - 1 - p^{m-m'}\right)\tilde{\epsilon}_k^{f_k(g^h)}. \tag{4.19}$$

4.3.2 The Case of $x = j\bar{n}$, $j \neq h$

In this case, the cross-correlation between the sequences $\hat{\mathcal{S}}$ and $\mathcal{S}$ becomes as follows:

$$\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(x) = \sum_{i=0}^{n-1} \tilde{\epsilon}_k^{f_k\left(\mathrm{Tr}\left(\omega^{i+j\bar{n}}\right)+g^hA\right)-f_k\left(\mathrm{Tr}\left(\omega^i\right)+A\right)} = \sum_{i=0}^{n-1} \tilde{\epsilon}_k^{f_k\left(g^j\mathrm{Tr}\left(\omega^i\right)+g^hA\right)-f_k\left(\mathrm{Tr}\left(\omega^i\right)+A\right)}. \tag{4.20}$$

According to **Property** 9, depending on the condition whether or not $\mathrm{Tr}\left(\omega^i\right) + A = 0$ and $g^j\mathrm{Tr}\left(\omega^i\right) + g^hA = 0$ following relation is obtained.

$$\begin{aligned}\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(x) = &\sum_{\substack{g^j\mathrm{Tr}\left(\omega^i\right)+g^hA\neq 0\\ \mathrm{Tr}\left(\omega^i\right)+A=0}} \tilde{\epsilon}_k^{f_k\left(A(g^h-g^j)\right)} + \sum_{\substack{g^j\mathrm{Tr}\left(\omega^i\right)+g^hA=0\\ \mathrm{Tr}\left(\omega^i\right)+A\neq 0}} \tilde{\epsilon}_k^{-f_k\left(A(1-g^{h-j})\right)}\\ &+ \sum_{\substack{g^j\mathrm{Tr}\left(\omega^i\right)+g^hA\neq 0\\ \mathrm{Tr}\left(\omega^i\right)+A\neq 0}} \tilde{\epsilon}_k^{f_k\left(\left(g^j\mathrm{Tr}\left(\omega^i\right)+g^hA\right)\left(\mathrm{Tr}\left(\omega^i\right)+A\right)^{-1}\right)}.\end{aligned} \tag{4.21}$$

For example, if $\mathrm{Tr}\left(\omega^i\right) + A = 0$ and $j \neq h$, then,

$$g^j\mathrm{Tr}\left(\omega^i\right) + g^hA = A(g^h - g^j) \neq 0. \tag{4.22}$$

Depending on **Property** 6, first and second summations in Eq.(4.21) respectively becomes as follows:

$$\sum_{\substack{g^j\mathrm{Tr}\left(\omega^i\right)+g^hA\neq 0\\ \mathrm{Tr}\left(\omega^i\right)+A=0}} \tilde{\epsilon}_k^{f_k\left(A(g^h-g^j)\right)} = p^{m-m'}\tilde{\epsilon}_k^{f_k\left(A(g^h-g^j)\right)}, \tag{4.23}$$

$$\sum_{\substack{g^j\mathrm{Tr}\left(\omega^i\right)+g^hA=0\\ \mathrm{Tr}\left(\omega^i\right)+A\neq 0}} \tilde{\epsilon}_k^{-f_k\left(A(1-g^{h-j})\right)} = p^{m-m'}\tilde{\epsilon}_k^{-f_k\left(A(1-g^{h-j})\right)}, \tag{4.24}$$

where the following facts and conditions should be noted for the above two summations:

- In this thesis, the parameter A is not 0 and $A \in \mathbb{F}_q$.
- The case of $\mathrm{Tr}\left(\omega^i\right)+A=0$, $g^j\mathrm{Tr}\left(\omega^i\right)+g^hA\neq 0$.
- While $g^j\mathrm{Tr}\left(\omega\right)^i+g^hA=0$, $\mathrm{Tr}\left(\omega^i\right)+A\neq 0$

Assume, $X_i=\mathrm{Tr}\left(\omega^i\right)+A\neq 0$. Then the third summation in Eq.(4.21) becomes as follows:

$$\sum_{\substack{g^j\mathrm{Tr}\left(\omega^i\right)+g^hA\neq 0\\ \mathrm{Tr}\left(\omega^i\right)+A\neq 0}} \tilde{\epsilon}_k^{f_k\left(\left(g^j\mathrm{Tr}\left(\omega^i\right)+g^hA\right)\left(\mathrm{Tr}\left(\omega^i\right)+A\right)^{-1}\right)} = \sum_{\substack{g^j\mathrm{Tr}\left(\omega^i\right)+g^hA\neq 0\\ \mathrm{Tr}\left(\omega^i\right)+A\neq 0}} \tilde{\epsilon}_k^{f_k\left(g^j+A(g^h-g^j)X_i^{-1}\right)}. \tag{4.25}$$

Now all of the possible values of $X_i \in \mathbb{F}_q$ needs to be consider to resolve the Eq.(4.25). According to **Property** 6 and considering the exceptions for the first and second summations in Eq.(4.21), following relations are obtained,

$$\#\{i|X_i=0\} = p^{m-m'}, \tag{4.26a}$$

$$\#\{i|X_i=A(1-g^{-j})\} = p^{m-m'}, \tag{4.26b}$$

$$\#\{i|X_i=A\} = p^{m-m'}-1, \tag{4.26c}$$

$$\#\{i|X_i=u\} = p^{m-m'}, \tag{4.26d}$$

here $0\le i<n$ and for each $u\in\mathbb{F}_q-\{0,A,A(1-g^{h-j})\}$. The cases of Eq.(4.26a) and Eq.(5.13) respectively comply the first and second summations in Eq.(4.21).

Furthermore, assume $Y_i=g^j+A(g^h-g^j)X_i^{-1}$ this is the input of mapping function $f_k(\cdot)$ as defined in Eq.(4.25). Hence, considering the cases of $X_i=A(1-g^{h-j})$ and $X_i=0$, the value of Y_i in Eq.(4.25) can not be 0 and g^j, respectively. These two cases already separated in Eq.(4.21) as the first and second summations. As a consequence, Eq.(4.25) can be rewritten as in Eq.(4.27). In order to conform, the case of $Y_i=g^j$ *part* (B) is added in Eq.(4.27). Furthermore, *part* (C) in Eq.(4.27) is for adjusting the number of cases of $X_i=A$, which mentioned in Eq.(4.26c). Therefore, Eq.(4.11b) holds at *part* A in Eq.(4.27).

$$\begin{aligned}
\sum_{\substack{g^j\mathrm{Tr}(\omega^i)+g^hA\neq 0\\ \mathrm{Tr}(\omega^i)+A\neq 0}} \tilde{\epsilon}_k^{f_k\left(g^j+A(g^h-g^j)X_i^{-1}\right)} &= \left(\tilde{\epsilon}_k^{f_k(g^h)}-\tilde{\epsilon}_k^{f_k(g^h)}\right)+\left(p^{m-m'}\tilde{\epsilon}_k^{f_k(g^j)}-p^{m-m'}\tilde{\epsilon}_k^{f_k(g^j)}\right)\\
&\quad+\sum_{\substack{g^j\mathrm{Tr}(\omega^i)+g^hA\neq 0\\ \mathrm{Tr}(\omega^i)+A\neq 0}} \tilde{\epsilon}_k^{f_k\left(g^j+A(g^h-g^j)X_i^{-1}\right)}\\
&= \underbrace{\left[\sum_{\substack{g^j\mathrm{Tr}(\omega^i)+g^hA\neq 0\\ \mathrm{Tr}(\omega^i)+A\neq 0}} \tilde{\epsilon}_k^{f_k\left(g^j+A(g^h-g^j)X_i^{-1}\right)}+\underbrace{p^{m-m'}\tilde{\epsilon}_k^{f_k(g^j)}}_{(\mathrm{B})}+\underbrace{\tilde{\epsilon}_k^{f_k(g^h)}}_{(\mathrm{C})}\right]}_{(\mathrm{A})}-p^{m-m'}\tilde{\epsilon}_k^{f_k(g^j)}-\tilde{\epsilon}_k^{f_k(g^h)}\\
&= \underline{0}_{(\mathrm{A})}-p^{m-m'}\left(\tilde{\epsilon}_k^{f_k(g^j)}\right)-\tilde{\epsilon}_k^{f_k(g^h)}=-p^{m-m'}\left(\tilde{\epsilon}_k^{f_k(g^j)}\right)-\tilde{\epsilon}_k^{f_k(g^h)}. \qquad (4.27)
\end{aligned}$$

Hence, the cross-correlation of the sequence $\hat{\mathcal{S}}$ and $\mathcal{S}$ becomes as follows for the case of $x = j\bar{n}, j \neq h$,

$$\begin{aligned}
\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(x) &= p^{m-m'}\tilde{\epsilon}_k^{f_k\left(A(g^h-g^j)\right)}+p^{m-m'}\tilde{\epsilon}_k^{-f_k\left(A(1-g^{h-j})\right)}-p^{m-m'}\tilde{\epsilon}_k^{f_k(g^j)}-\tilde{\epsilon}_k^{f_k(g^h)}\\
&= p^{m-m'}\left(\tilde{\epsilon}_k^{f_k\left(A(g^h-g^j)\right)}+\tilde{\epsilon}_k^{-f_k\left(A(1-g^{h-j})\right)}-\tilde{\epsilon}_k^{f_k(g^j)}\right)-\tilde{\epsilon}_k^{f_k(g^h)}. \qquad (4.28)
\end{aligned}$$

4.3.3 Otherwise

In this case, the cross-correlation between the sequences $\hat{\mathcal{S}}$ and $\mathcal{S}$ becomes as,

$$\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(x)=\sum_{i=0}^{n-1}\tilde{\epsilon}_k^{f_k\left(\mathrm{Tr}\left(\omega^{i+x}\right)+g^hA\right)-f_k\left(\mathrm{Tr}\left(\omega^i\right)+A\right)}. \qquad (4.29)$$

Here, x is not divisible by $\bar{n}$ and ω^x does not belongs to $\mathbb{F}_q$. We assume the following basis $\mathcal{G}$ in $\mathbb{F}_{p^m}$, by using this ω^x as,

$$\mathcal{G}=\{\omega^x,1,\gamma_2,\gamma_3,\ldots,\gamma_{r-1}\}. \qquad (4.30)$$

Again let $\mathcal{T}$ be the dual basis of $\mathcal{G}$.

$$\mathcal{T}=\{\theta_0,\theta_1,\theta_2,\theta_3,\ldots,\theta_{r-1}\}. \qquad (4.31)$$

Assume that ω^i can be represented with θ as follows:

$$\omega^i=\sum_{l=0}^{r-1}v_{i,l}\theta_l. \qquad (4.32)$$

Then, ω^{i+x} is given by

$$\omega^{i+x}=\sum_{l=0}^{r-1}v_{i,l}\theta_l\omega^x. \qquad (4.33)$$

Based on **Property** 8, initial value of $\mathrm{Tr}\left(\omega^i\right)$ is as,

$$\mathrm{Tr}\left(\omega^i\right) = v_{i,1}. \tag{4.34}$$

As previously mentioned that, $\mathcal{W}$ and $\mathcal{B}$ are the dual bases to each other, therefore $\mathrm{Tr}\left(\omega^{i+x}\right)$ can be expressed as follows:

$$\mathrm{Tr}\left(\omega^{i+x}\right) = v_{i,0}. \tag{4.35}$$

After substituting these trace values, Eq.(4.29) becomes as follows.

$$\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(x) = \sum_{i=0}^{n-1} \tilde{\epsilon}_k^{f_k(v_{i,0}+g^hA)-f_k(v_{i,1}+A)}. \tag{4.36}$$

Based on the Eq.(4.11), the above equation is rewritten as,

$$\begin{aligned}\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(x) &= \sum_{\substack{v_{i,0}+g^hA=0\\ v_{i,1}+A=0}} \tilde{\epsilon}_k^{0} + \sum_{\substack{v_{i,0}+g^hA=0\\ v_{i,1}+A\neq 0}} \tilde{\epsilon}_k^{-f_k(v_{i,1}+A)} + \sum_{\substack{v_{i,0}+g^hA\neq 0\\ v_{i,1}+A=0}} \tilde{\epsilon}_k^{f_k(v_{i,0}+g^hA)} \\ &+ \sum_{\substack{v_{i,0}+g^hA\neq 0\\ v_{i,1}+A\neq 0}} \tilde{\epsilon}_k^{f_k\left((v_{i,0}+g^hA)(v_{i,1}+A)^{-1}\right)}.\end{aligned} \tag{4.37}$$

According to Eq.(4.11b) and ω^i holds the relation $0 \leq i < n$, which actually represents every non-zero element in $\mathbb{F}_{p^m}$, therefore, the second and third summations holds the following relations.

$$\sum_{\substack{v_{i,0}+g^hA=0\\ v_{i,1}+A\neq 0}} \tilde{\epsilon}_k^{-f_k(v_{i,1}+A)} = 0. \tag{4.38a}$$

$$\sum_{\substack{v_{i,0}+g^hA\neq 0\\ v_{i,1}+A=0}} \tilde{\epsilon}_k^{f_k(v_{i,0}+g^hA)} = 0. \tag{4.38b}$$

In addition, by considering the sub extension field $\mathbb{F}_q$ and fixing the values of $v_{i,0}$ and $v_{i,1}$ the first summation holds the following relation as,

$$\sum_{\substack{v_{i,0}+g^hA=0\\ v_{i,1}+A=0}} \tilde{\epsilon}_k^{0} = p^{m-2m'}. \tag{4.38c}$$

Considering the same calculation procedure of Eq.(4.27), the fourth summation in Eq.(4.37) becomes as follows:

$$\sum_{\substack{v_{i,0}+g^hA\neq 0\\ v_{i,1}+A\neq 0}} \tilde{\epsilon}_k^{f_k\left((v_{i,0}+g^hA)(v_{i,1}+A)^{-1}\right)} = p^{m-2m'} \sum_{a=1}^{p-1}\sum_{b=1}^{p-1} \tilde{\epsilon}_k^{f_k(ab^{-1})} - \tilde{\epsilon}_k^{f_k(0+g^hA)-f_k(0+A)}. \tag{4.39}$$

Since ω^i can not represent the zero vector, the number of vectors such that $v_{i,0} = 0$ and $v_{i,1} = 0$ is one less than that of the other combinations like $v_{i,0} = 0$ and $v_{i,1} = 1$. That is why, the last subtraction $\tilde{\epsilon}_k^{f_k(0+g^hA)-f_k(0+A)}$ is required in Eq.(4.39). According to the condition from the Eq.(4.11b), the first summation in Eq.(4.39) becomes 0. Therefore, the following relation is obtained,

$$\sum_{\substack{v_{i,0}+g^hA\neq 0\\ v_{i,1}+A\neq 0}} \tilde{\epsilon}_k^{f_k\left((v_{i,0}+g^hA)(v_{i,1}+A)^{-1}\right)} = -\tilde{\epsilon}_k^{f_k(g^h)}. \tag{4.40}$$

Therefore, the cross-correlation of the sequences $\hat{\mathcal{S}}$ and $\mathcal{S}$ becomes as follows for this case,

$$\mathrm{R}_{\hat{\mathcal{S}},\mathcal{S}}(x) = p^{m-2m'} - \tilde{\epsilon}_k^{f_k(g^h)}. \tag{4.41}$$

Finally, the cross-correlation of the sequences $\hat{\mathcal{S}}$ and $\mathcal{S}$, that is in Eq.(4.15), is proven.

4.3.4 Autocorrelation and Period

If the value of $h = 0$, then $\hat{\mathcal{S}}$ and $\mathcal{S}$ becomes the same sequence. In this case, the cross-correlation in Eq.(4.15) becomes the autocorrelation after replacing the value $h = 0$.

$$\mathrm{R}_{\mathcal{S}}(x) = \begin{cases} p^m - 1, & \text{if } x = h\bar{n}, \\ p^{m-m'}\left(\tilde{\epsilon}_k^{f_k\left(A(1-g^j)\right)} + \tilde{\epsilon}_k^{-f_k\left(A(1-g^{-j})\right)} - \tilde{\epsilon}_k^{f_k(g^j)}\right) - 1, & \text{else if } x = j\bar{n}, \\ p^{m-2m'} - 1, & \text{otherwise.} \end{cases} \tag{4.42}$$

Corresponding to the above autocorrelation equation, the period of the proposed k-value sequence explicitly given by $p^m - 1$.

4.4 Examples and Discussions

This section experimentally observes the properties of the proposed sequence such as period, autocorrelation, and cross-correlation along with some examples. Throughout this section, $|x|$ provides the absolute value of a complex number x. In addition, the notation $\mathcal{S}_3$ denotes the proposed sequence with the parameter $A = 3$.

4.4.1 $p = 5, m = 4, m' = 2, k = 2,$ and $A = 3, 4$

Let $f(x) = x^4 + 2x^3 + 2x^2 + 2x + 2$ be a primitive polynomial over $\mathbb{F}_5$. In this case, the period of this sequence becomes $p^m - 1 = 624$. Then the sequence $\mathcal{S}_3$

is shown in Eq.(4.43) and its autocorrelation becomes as follows and **Figure** 4.1 shows its autocorrelation graph.

$$\begin{aligned}\mathcal{S}_3 = \{&0010001110100000000100000000110101001100011001001101110101011010100001001001010000\\&1010010011010001100001000011101011010100100001000001101110111101001111011000000010\\&1100010000011110011001001000011000010010111101110111100011101110010100011101011111\\&1001001110001100110000000011110000111110110000111110110110000000110110001110100001 1\\&1001010101001111100010001000011100110111100101101110101010001011100100000111001001\\&1010101000111110101000010111010110101101111011010100011100000100101111010001001 01\\&1011011110101000000100001011001001001100001110000110000111111110101010101100110110\\&110111110100011010011111101100110001001101100101001100110\}.\end{aligned} \tag{4.43}$$

$$|\mathrm{R}_{\mathcal{S}_3}(x)| = \begin{cases} 624, & \text{if } x = 0 \\ 24, & \text{else if } x = 26, 78, 104, 130, 156, 182, 234, 286, 312, 338, 390, 442, 468, 494, 520, 546, 598, \\ 76, & \text{else if } x = 52, 208, 260, 364, 416, 572, \\ 0, & \text{otherwise} \end{cases} \tag{4.44}$$

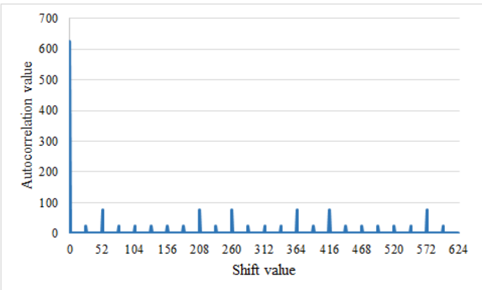

FIGURE 4.1: $|\mathrm{R}_{\mathcal{S}_3}(x)|$ with $p = 5, m = 4, m' = 2, k = 2$, and $A = 3$.

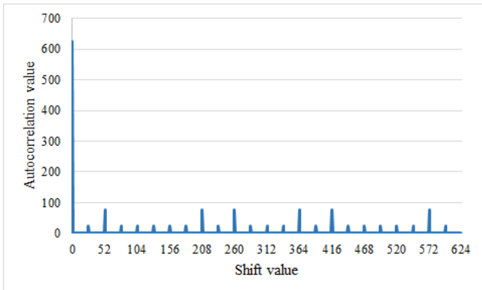

FIGURE 4.2: $|\mathrm{R}_{\mathcal{S}_4}(x)|$ with $p = 5, m = 4, m' = 2, k = 2$, and $A = 4$.

On the other hand, it should be noted that $\mathcal{S}_4$ is different from $\mathcal{S}_3$ and its autocorrelation is given as follows and **Figure** 4.2 shows its autocorrelation graph.

$$|\mathrm{R}_{\mathcal{S}_4}(x)| = \begin{cases} 624, & \text{if } x = 0 \\ 24, & \text{else if } x = 26, 78, 104, 130, 56, 182, 234, 286, 312, 338, 390, 442, 468, 494, 520, 546, 598, \\ 76, & \text{else if } x = 52, 208, 260, 364, 416, 572, \\ 0, & \text{otherwise} \end{cases} \tag{4.45}$$

The cross-correlation of $\mathcal{S}_3$ and $\mathcal{S}_4$ becomes as follows and **Figure** 4.3 shows its cross-correlation graph.

$$|\mathrm{R}_{\mathcal{S}_3,\mathcal{S}_4}(x)| = \begin{cases} 24, & \text{if } x = 0, 26, 52, 78, 130, 182, 234, 260, 286, 312, 338, 390, 442, 468, 494, 546, 598, \\ 76, & \text{else if } x = 104, 208, 364, 416, 520, 572, \\ 624, & \text{else if } x = 156, \\ 0, & \text{otherwise} \end{cases} \tag{4.46}$$

4.4.2 $p = 7, m = 4, m' = 2, k = 3,$ and $A = 3, 4$

Let $f(x) = x^4 + 4x^3 + 3x^2 + 5x + 3$ be a primitive polynomial over $\mathbb{F}_7$. In this case, the period of the sequence becomes $p^m - 1 = 2400$. **Figure** 4.4, **Figure** 4.5, and **Figure** 4.6 show the autocorrelation graphs of $\mathcal{S}_3$, $\mathcal{S}_4$, and the cross-correlation between the $\mathcal{S}_3$ and $\mathcal{S}_4$, respectively.

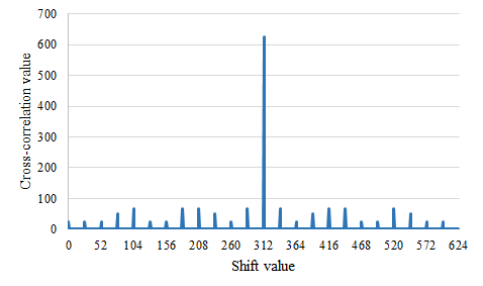

FIGURE 4.3: $|\mathrm{R}_{\mathcal{S}_3,\mathcal{S}_4}(x)|$ with $p = 5, m = 4, m' = 2$, and $A = 3, 4$.

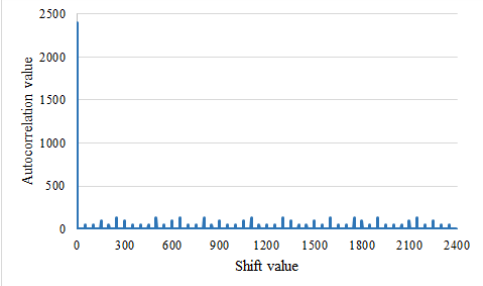

FIGURE 4.4: $|\mathrm{R}_{\mathcal{S}_3}(x)|$ with $p = 7, m = 4, m' = 2, k = 3$, and $A = 3$.

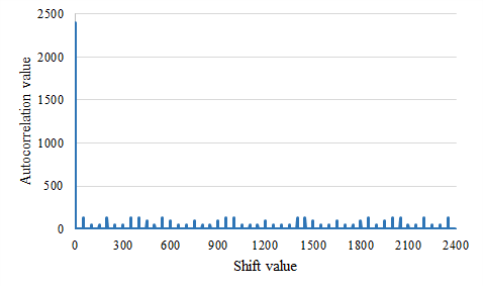

FIGURE 4.5: $|\mathrm{R}_{\mathcal{S}_4}(x)|$ with $p = 7, m = 4, m' = 2, k = 3$, and $A = 4$.

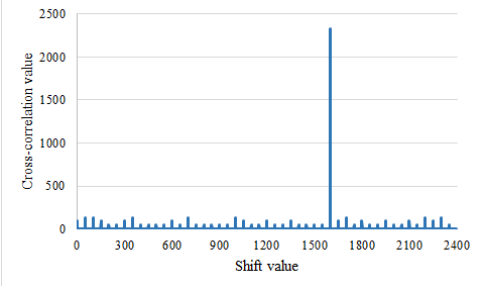

FIGURE 4.6: $|\mathrm{R}_{\mathcal{S}_3,\mathcal{S}_4}(x)|$ with $p = 7, m = 4, m' = 2$, and $A = 3, 4$.

By observing the experimental results, it is found that in every case, the cross-correlation graph has exactly $q-1$ number of peaks. Among those, only one has a maximum value. For example, in **Figure** 4.3, the maximum cross-correlation value is 624, which corresponds to the first case of $x = h\bar{n}$, the remaining $q - 2$ smaller peaks conform the second case of $x = j\hat{n}$, and except these $q - 1$ peaks the remaining part in the graph always holds a constant value of 0, which corresponds the case third case in Eq.(4.15). It means that all this cross-correlation graph can be explained by the Eq.(4.15). It is also observed that by changing all the parameter values does not have any impact in the cross-correlation evaluation. On the other hand, as like the cross-correlation, the autocorrelation graph also has $q - 1$ number of peaks. Among them, only one holds the maximum value, the others have small values, the remaining part always holds a constant value of 1, and all these autocorrelation graphs can be explained by the Eq.(4.42).

4.5 Comparison with Previous Work

Although nowadays multi-value sequence does not have enough application except the binary sequence (specially in security applications). Therefore, this section will emphasis on the binary case of the proposed sequence. Even though the proposed sequence is a multi-value sequence, it can be easily mapped into binary sequence by setting the parameter value $k = 2$. This section will introduce a comparison of the proposed sequence (binary case) with the previous work [Nog+16] in terms of autocorrelation property. In this section, the previous sequence proposed in [Nog+16] will be called as NTU (Nogami-Tada-Uehara) sequence.

Autocorrelation

The autocorrelation of a sequence is a measure for how much the sequence differs from its each shift value. In addition, by evaluating this property some special characteristics about the sequence such as its period, some pattern of the sequence, and so on can be also found and the value of the autocorrelation always preferred to be as low as possible [PVO91]. The autocorrelation of the proposed sequence (defined over sub extension field) and the previous sequence (NTU) (defined over prime field) is shown in **Figure** 4.1 and **Figure** 4.7, respectively. By observing the autocorrelation graph, it was found that on one hand, the number of peak values is increases in the sub field sequence, on the other hand, the difference between the maximum peak value with the smaller peak values are much smaller in the proposed sequence compared to the previous sequence. Moreover, in the proposed sequence except the peaks remaining autocorrelation value always remains at 0. It should be noted that in case of correlation evaluation, the less difference between the peak values are more crucial rather than the number of peaks [PVO91].

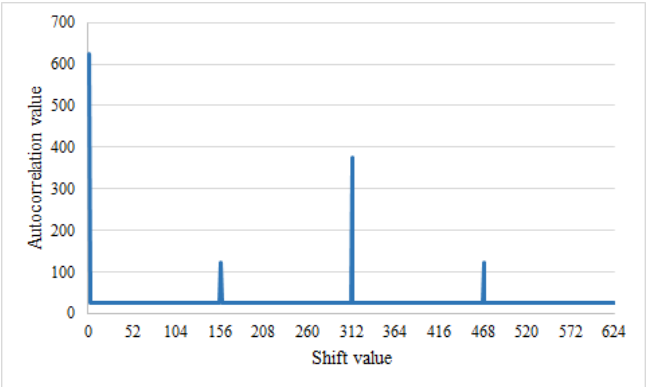

FIGURE 4.7: Autocorrelation of NTU sequence.

As mentioned previously, NTU sequence proposed in [Nog+16] is defined over the prime field and proposed sequence in this thesis is defined over the sub extension field. After the comparison results it is concluded that in terms

of correlation the proposed sequence holds low correlation compared to NTU sequence.

4.6 Summary

This chapter theoretically explained the period and correlation properties regarding a pseudo random sequence (which defined over the sub extension field).

- This is an extension of the previous works [NTU14; Nog+16; Ali+16b].
- This work overcomes the shorter period shortcoming of the previous work [Ali+17b].
- The period, autocorrelation, and cross-correlation properties regarding the proposed sequence are theoretically explained.
- This chapter makes a comparison in terms of correlation and according to the comparison result, the proposed sequence holds low correlation compared to the previous work [Nog+16].

Chapter 5

Distribution of Bit Patterns of Pseudo Random Sequence

5.1 Introduction

Background and Motivation

In this IoT era, we communicate with each other through the internet. Therefore, secure communication is the major matter of concern. We use symmetric cryptosystems (Advanced Encryption Standard (AES) [JV02]) and asymmetric cryptosystems (Rivest Shamir Adleman (RSA) [Mol02], and Elliptic Curve Cryptography (ECC) [Hen+05]) to establish a secure communication. More specifically, in case of cryptography, to generate the keys (public key, private, session key, and so on) a pseudo random number generator is used. Thus, it is mandatory to evaluate the randomness of a sequence before utilized them in any cryptosystems. Basically, two crucial properties namely the linear complexity [Ede14] and the distribution of bit patterns regarding a sequence are nowadays well-known to check the randomness of a pseudo random sequence. This chapter restricts the discussion on the distribution of bit patterns property to evaluate the randomness of a sequence which generated over the sub extension field.

The trace calculation is an important step during the proposed sequence generation procedure. At first, focus on the important aspect regarding this calculation. In case of prime field $\mathbb{F}_p$, the trace function maps an element of the extension field $\mathbb{F}_{q^M}$ to an element of the prime field $\mathbb{F}_p$. Therefore, the number of possible trace outputs will be in the range of $\{0 \sim p-1\}$. In other words, if we calculate the trace over the prime field, then it will output p kinds of values. On the other hand, in case of the sub extension field $\mathbb{F}_q$, the cascaded trace function maps an element of the extension field $\mathbb{F}_{q^M}$ to an element of the sub extension field $\mathbb{F}_q$ and the number of possible trace outputs will be in the range of $\{0 \sim q-1\}$ which means the trace outputs q kinds of values. It should be noted that here $M = m/m'$, $q = p^{m'}$, and m' be one of the factors of m. From the theoretical perspective, more variation in the trace values contribute to the better appearance of bits (0 and 1) in a sequence. This is one of the important aspects to consider the sub extension during the sequence generation procedure.

Related Works

The Legendre sequence [XG10; JS+96] has a long period, high linear complexity, and the distribution of bit pattern is known to be close to uniform [Din98; Mor61]. On the other hand, M-sequence has a maximum period but minimum linear complexity. In addition, M-sequence [Ren04] is well-known for its uniform distribution of bit patterns [Hel11]. Previous work on geometric sequence [NTU14] combines the features of the Legendre sequence and M-sequence. The distribution of bit patterns in [NTU14] doesn't reaches up to the mark alike the Legendre sequence and M-sequence, whereas this property is an important measure to evaluate the randomness of a sequence. Hence, it is a scope to improve the distribution of bit patterns in previous sequence.

Contribution

This chapter observed the distribution of bit patterns in a binary sequence which generated by a primitive polynomial, trace function, and Legendre symbol over the sub extension field. After observing many experimental results, this chapter concluded that the number of appearances of each bit pattern is related to the number of zeros contained in each bit pattern. Furthermore, this chapter includes the theoretical prove of the distribution of bit patterns equation and a comparison with previous work [NTU14]. According to the comparison result, binary sequence (defined over sub extension field) holds much better (close to uniform) distribution of bit patterns than the previous binary sequence [NTU14]. Finding this improvement by considering the sub extension field is the major contribution of this chapter.

5.2 Preparation

This section briefly explains Legendre symbol for sub field. Then, binary sequence is introduced along with its period and distribution of bit patterns properties. In addition, the fundamental concepts are already introduced in **Chapter** 2.

5.2.1 Legendre Symbol for Sub Field

The Legendre symbol $\left(a/q\right)_2$ is used to check the quadratic residue for any arbitrary element a in $\mathbb{F}_q$. It is defined as,

$$\begin{aligned}\left(a/q\right)_2 &= a^{(q-1)/2} \\ &= \begin{cases} 0, & \text{if } a = 0, \\ 1, & \text{else if } a \text{ is a QR in } \mathbb{F}_q^*, \\ p-1, & \text{otherwise } a \text{ is a QNR in } \mathbb{F}_q^*. \end{cases}\end{aligned} \tag{5.1}$$

Here, QR and QNR stand for Quadratic Residue (QR) and Quadratic Non-Residue (QNR), respectively. These QR and QNR in $\mathbb{F}_q$ holds the following important property.

Non-zero elements are the roots of $x^{q-1}-1$ in $\mathbb{F}_q^*$ over $\mathbb{F}_q$ without any duplicates. Since it is factorized as follows:

$$x^{q-1}-1=\left(x^{(q-1)/2}-1\right)-\left(x^{(q-1)/2}-1\right). \tag{5.2}$$

It is thus found that the number of QR's and QNR's in $\mathbb{F}_q^*$ are the same and it is given by $(q-1)/2$. In addition, these numbers are important part in proving the theorem in the later section of this chapter.

The pseudo random sequence proposes in this chapter is a binary sequence and it is generated without the parameter A (non-zero arbitrary element in the sub extension field $\mathbb{F}_q$). Therefore, the modified equation for generating the proposed sequence generation is introduced in Eq.(5.3). In this case, the period of the sequence becomes shorter and it is denoted as λ throughout this chapter.

Binary Sequence Generation Procedure

Let ω be a primitive element in the extension field $\mathbb{F}_{q^M}$, where $M=m/m'$, m be a composite number which denotes the extension degree of the primitive polynomial, and m' be one of the factors of m. Then, by utilizing the trace function and Legendre symbol a binary sequence $\mathcal{S}$ is generated as follows:

$$\mathcal{S}=\{s_i\}, s_i=f_2\left(\left(\mathrm{Tr}_{q^M|q}\left(\omega^i\right)\Big/q\right)\right), \tag{5.3}$$

where $i=(0,1,2,\ldots,\lambda-1,\ldots)$, $s_i\in 0,1$ and $f_2(\cdot)$ be a mapping function, which translates the 0, 1, and $p-1$ values sequence generated by the Legendre symbol to a pseudo random binary sequence. This mapping function is defined as follows:

$$f_2(s)=\begin{cases}0, & \text{if } x=0,1 \bmod q,\\ 1, & \text{otherwise.}\end{cases} \tag{5.4}$$

After observing many experimental results, this chapter derive the equation for the period λ of the binary sequence as,

$$\lambda=\frac{2(q^M-1)}{q-1}. \tag{5.5}$$

Distribution of Bit Patterns

From the viewpoint of security, the distribution of bit patterns is as important as the linear complexity. This chapter counts the number of appearances for each n-bit patterns (where $1\le n\le(m/m')$). After observing many experimental results, this chapter concluded that the number of appearances of each bit

pattern is related to the number of zeros contained in each bit pattern. As a reference, an M-sequence is well-known for its uniform distribution of bit patterns as shown in **Table** 2.1. A uniform distribution of bit patterns means all the bit patterns (1-bit pattern, 2-bit patterns, 3-bit patterns and so on) should appear the same in number. If every bit pattern appears same in number in a sequence, then it will be difficult to guess the next bit after observing the previous bit patterns. Therefore, to judge the randomness of a sequence, the study of its bit distribution is essential.

5.3 Distribution of Bit Patterns in Binary Sequence

This section introduces the bit distribution of binary sequence which generated over the sub extension field. In addition, bit distribution of M-sequence and Legendre sequence is also introduced here. Throughout this section $b^{(n)}$, $Z(b^{(n)})$ and $D_{\mathcal{S}_\lambda}(b^{(n)})$ denotes a bit pattern of length n, number of 0's in $b^{(n)}$, and number of appearances of $b^{(n)}$ in $\mathcal{S}_\lambda$, respectively. For example, in a binary sequence of period 15, a 3-bit pattern $b = 101$ appears 4 times. Then, these notations become $b^{(3)} = 101$, $Z(b^{(3)}) = 1$, and $D_{\mathcal{S}_{15}}(b^{(3)}) = 4$. Furthermore, $\mathrm{H_w}$ denotes the hamming weight.

5.3.1 Bit Distribution of M-sequence

The M-sequence [Ren04] is generated by a linear recurrence relation over the finite field. M-sequence has a maximum period and uniform distribution of bit pattern except for the case of $Z(b^{(n)}) = n$ but it has minimum linear complexity. Let, $f(x) = x^4 + x + 1$ be a primitive polynomial over $\mathbb{F}_2$, then using the linear recurrence relation a M-sequence of period 15 becomes as follows.

$$\mathcal{S}_{15} = \{1, 0, 0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 0, 0\}. \tag{5.6}$$

The distribution of n-bit pattern in (5.6) is shown in Table 5.1, here $1 \leq n \leq m$. In the case of M-sequence, except the all-zero pattern, every pattern appears same in number. For example, when $n = 3$ all patterns appear 2 times (except 000 pattern). In other words, they are uniformly distributed. Every M-sequence has such good distribution of bit pattern feature.

5.3.2 Bit Distribution of Legendre Sequence

Legendre sequence [XG10; JS+96] is generated by applying the Legendre symbol over the odd characteristic field. Legendre sequence has a long period, high linear complexity, and the distribution of bit pattern is close to uniform. Let,

TABLE 5.1: Bit distribution of the M-sequence $\mathcal{S}_{15}$.

n	$\mathrm{H_w}(b^{(n)})$	$Z(b^{(n)})$	$D_{S_{15}}(b^{(n)})$
1	0	1	7
	1	0	8
2	0	2	3
	1	1	4
	2	0	4
3	0	3	1
	1	2	2
	2	1	2
	3	0	2

$p = 23$, then the Legendre sequence of period 23 becomes as follows.

$$\mathcal{S}_{23} = \{0,0,0,0,0,1,0,1,0,0,1,1,0,0,1,1,0,1,0,1,1,1,1\}. \tag{5.7}$$

The distribution of n-bit pattern in (5.7) is shown in Table 5.2. In case of Legendre symbol, bit patterns appearance is close to uniform.

TABLE 5.2: Bit distribution of the Legendre sequence $\mathcal{S}_{23}$.

n	$\mathrm{H_w}(b^{(n)})$	$Z(b^{(n)})$	$D_{S_{23}}(b^{(n)})$
1	0	1	12
	1	0	11
2	0	2	6
	1	1	6
	2	0	5
3	0	3	3
	1	2	3
	2	1	2
	3	0	2

5.3.3 Bit Distribution of the Proposed Binary Sequence

Let $\mathcal{S}_\lambda$ be a binary sequence of having a period of λ. Again, let $b^{(n)}$, $Z(b^{(n)})$, and $D_{\mathcal{S}_\lambda}(b^{(n)})$ denotes a bit pattern of length n, number of 0's in $b^{(n)}$, and number of appearance of $b^{(n)}$ in $\mathcal{S}_\lambda$, respectively. Then, the distribution of bit patterns in the binary sequence which defined over the sub extension field can be given by the following theorem.

Theorem 6

$$D_{S_\lambda}\left(b^{(n)}\right) = \begin{cases} q^{M-(n\cdot m')} \cdot \left(\frac{q-1}{2}\right)^{n-Z\left(b^{(n)}\right)-1} \cdot \left(\frac{q+1}{2}\right)^{Z\left(b^{(n)}\right)} & (5.8a) \\ \qquad\qquad \text{when } 0 \le Z\left(b^{(n)}\right) < n, & \\ \lambda - \sum_{u=0}^{Z\left(b^{(n)}\right)} {}_nC_u \cdot D_{S_\lambda}\left(b^{(n)}\right) & (5.8b) \\ \qquad\qquad \text{when } Z\left(b^{(n)}\right) = n. & \end{cases}$$

Let ω be a primitive element in the extension field $\mathbb{F}_{q^M}$, where $M = m/m'$, m be a composite number which denotes the extension degree of the primitive polynomial, and m' be one of the factors of m. Then, utilizing the trace function and Legendre symbol one period of a binary sequence is generated as follows.

$$\mathcal{S}_\lambda = \{s_i\}, s_i = f_2\left(\left(\mathrm{Tr}_{q^M|q}\left(\omega^i\right)\Big/p\right)\right), i = 0, 1, 2, \ldots, \lambda - 1, \ldots, \tag{5.9}$$

Here λ be the period of the sequence and it is given by the following equation,

$$\lambda = \frac{2(q^M - 1)}{q - 1}. \tag{5.10}$$

At first, a primitive polynomial is used, then the trace value is calculated, then the Legendre symbol outputs zero, QR or QNR in $\mathbb{F}_q$, and finally the sequence coefficients s_i is given by the mapping function $f_2(\cdot)$.

This chapter observes the distribution of n-bit patterns in a binary sequence. It should be noted that here n satisfies $1 \le n \le (m/m')$ relation. The distribution of n-bit patterns evaluated by observing the consecutive sequence coefficients $(s_i, s_{i+1}, \ldots, s_{i+(n-1)})$. Particularly,

$$\begin{aligned} s_{i+0} &= f_2\left(\left(\mathrm{Tr}\left(\omega^i \cdot \omega^0\right)\Big/p\right)\right), \\ s_{i+1} &= f_2\left(\left(\mathrm{Tr}\left(\omega^i \cdot \omega^1\right)\Big/p\right)\right), \\ &\vdots \\ s_{i+(n-1)} &= f_2\left(\left(\mathrm{Tr}\left(\omega^i \cdot \omega^{n-1}\right)\Big/p\right)\right), \end{aligned}$$

where $0 \le i \le (q^M - 2)$. By observing the above sequence coefficients, the distribution of bit patterns $D_{\mathcal{S}_\lambda}$ is determined by the following trace values.

$$\mathrm{Tr}\left(\omega^i \cdot \omega^0\right), \mathrm{Tr}\left(\omega^i \cdot \omega^1\right), \ldots, \mathrm{Tr}\left(\omega^i \cdot \omega^{n-1}\right). \tag{5.11}$$

Let $\mathcal{A} = \{\alpha_0, \alpha_1, \ldots, \alpha_{m-1}\}$ be a basis, ω be a primitive element and with this basis ω^i is represented as,

$$\omega^i = \sum_{j=0}^{m-1} a_{i,j}\alpha_j, \text{where } a_{i,j} \in \mathbb{F}_q \text{ and } 0 \leq i \leq q^M - 2. \tag{5.12}$$

Again let $\mathcal{B} = \{\omega^0, \omega^1, \ldots, \omega^{n-1}, \beta_n, \ldots, \beta_{m-1}\}$ be a dual basis of $\mathcal{A}$ in $\mathbb{F}_q$ over $\mathbb{F}_{q^M}$. Then we also have

$$\omega^t = \omega^t + \sum_{j=0}^{\frac{m}{m'}-1} 0 \cdot \beta_j, \text{where } 0 \leq t < n. \tag{5.13}$$

Since $\mathcal{A}$ and $\mathcal{B}$ are dual bases to each other, then $\mathrm{Tr}\left(\omega^i \cdot \omega^t\right)$ be calculated as,

$$\begin{aligned} \mathrm{Tr}\left(\omega^i \cdot \omega^t\right) &= \mathrm{Tr}\left(\sum_{j=0}^{m-1} a_{i,j}\alpha_j \cdot \left(\omega^t + \sum_{j=0}^{\frac{m}{m'}-1} 0 \cdot \beta_j\right)\right) \\ &= a_{i,t}. \end{aligned}$$

Therefore, by using the dual basis, the distribution of bit patterns $D_{\mathcal{S}_\lambda}\left(b^{(n)}\right)$ determined by the trace values becomes as follows.

$$\begin{aligned} &\mathrm{Tr}\left(\omega^i \cdot \omega^0\right), \mathrm{Tr}\left(\omega^i \cdot \omega^1\right), \ldots, \mathrm{Tr}\left(\omega^i \cdot \omega^{n-1}\right) \\ &= \left(a_{i,0}, a_{i,1}, \ldots, a_{i,n-1}\right). \end{aligned}$$

Thus, instead of using sequence coefficients $(s_i, s_{i+1}, \ldots, s_{i+(n-1)})$, we can consider the dual basis representation of these coefficients as $\left(a_{i,0}, a_{i,1}, \ldots, a_{i,(n-1)}\right)$. Additionally, all the above trace values belong to the sub extension field $\mathbb{F}_q$.

Furthermore, $\omega^i\left(0 \leq i \leq q^M - 2\right)$ in (Eq.(5.12)) represents every non-zero vectors in the extension field $\mathbb{F}_{q^M}$ as,

$$\left\{\mathrm{Tr}\left(\omega^0\right), \mathrm{Tr}\left(\omega^1\right), \mathrm{Tr}\left(\omega^2\right), \mathrm{Tr}\left(\omega^3\right), \ldots, \mathrm{Tr}\left(\omega^{q^M-2}\right)\right\}. \tag{5.14}$$

According to the trace property, non-zero $\mathbb{F}_q$ elements appear $q^{M-m'}$ times and zero appears one less than the other elements, in other words $q^{M-m'} - 1$ times in the above equation.

5.3.3.1 Dependency of the Sequence Coefficients

Depending on the three different types of trace values (0, QR, and QNR), the Legendre symbol outputs three different values (0, 1, and $p-1$), and finally the mapping function outputs 0 and 1 as sequence coefficients s_i. This dependency between the trace and Legendre symbol is explained as follows.

According to the above table, the sequence coefficient 0 comes from the two cases: one is for the $\mathrm{Tr}(0)$ case and another one is for the QR in $\mathbb{F}_q^*$ case. To

TABLE 5.3: Relation between the sequence coefficients with trace and Legendre symbol calculation -I.

s_i	$\mathrm{Tr}\left(\omega^i\right)$
0	0 or QR in $\mathbb{F}_q^*$
1	QNR in $\mathbb{F}_q^*$

deal with this two cases uniquely, denote **0** and 0 for the first and second cases, respectively. In addition, 1 comes for QNR in $\mathbb{F}_q^*$ case. Thus the above table can be further modified as follows. To distinguish the appearance of 0, this chapter

TABLE 5.4: Relation between the sequence coefficients with trace and Legendre symbol calculation -II.

s_i	$\mathrm{Tr}\left(\omega^i\right)$
0	0
0	QR in $\mathbb{F}_q^*$
1	QNR in $\mathbb{F}_q^*$

uses the notation **0**, when zero comes from $\mathrm{Tr}(0)$ and 0 when zero comes from QR. Let the number of **0** be denoted by u and $T_{u,n}$ denotes the number of bit patterns including u times **0** and $Z(b^{(n)})-u$ times 0. Thus, T_n can be considered as follows:

$$T_n = \sum_{u=0}^{Z\left(b^{(n)}\right)} T_{u,n}. \tag{5.15}$$

In the following section, the distribution of bit patterns in the binary sequence defined over the sub extension field theoretically proven.

5.3.3.2 Proof of (5.8a)

The period of the binary sequence is given by the following equation as,

$$\lambda = \frac{2(q^M-1)}{q-1}. \tag{5.16}$$

After rewriting the above equation we obtain,

$$q^M - 1 = \lambda \cdot \left(\frac{q-1}{2}\right). \tag{5.17}$$

To observe the distribution of bit patterns, the above relation becomes as,

$$D_{\mathcal{S}_{q^M-1}}\left(b^{(n)}\right) = D_{\mathcal{S}_\lambda}\left(b^{(n)}\right) \cdot \left(\frac{q-1}{2}\right). \tag{5.18}$$

Thus, we must consider two cases of the sequence length such as $\mathcal{S}_{q^M-1}$ and $\mathcal{S}_\lambda$. Hence, we will observe the distribution of bit patterns in $\mathcal{S}_{q^M-1}$ as $D_{\mathcal{S}_{q^M-1}}\left(b^{(n)}\right)$ and $\mathcal{S}_\lambda$ as $D_{\mathcal{S}_\lambda}\left(b^{(n)}\right)$.

In the previous section, we explained that n-bit patterns can be considered as $b^{(n)} = (a_{i,0}, a_{i,1}, \ldots, a_{i,n-1})$. On the other hand, the remaining $(m-(n\cdot m'))$-bit patterns are composed of $(a_{i,nm'}, a_{i,nm'+1}, \ldots, a_{i,m-1})$ coefficients of ω^i, which is given by the (5.13). In addition, the number of combinations of $(a_{i,nm'}, a_{i,nm'+1}, \ldots, a_{i,m-1})$ becomes $q^{M-nm'}$. It should be noted that here ω^i represents all of the non-zero coefficients in the extension field $\mathbb{F}_{q^M}$.

As mentioned previously, when the trace value is equal to 0 or QR, then the sequence coefficients becomes $\mathbf{0}$ and 0, respectively. In addition, if the trace value is equal to QNR, then the sequence coefficients becomes 1. Additionally, u denotes the number of $\mathbf{0}$ in $b^{(n)}$ $\left(\text{where } 0 \le u \le Z\left(b^{(n)}\right)\right)$ from $\mathrm{Tr}(0)$, then the other 0's comes from $Z\left(b^{(n)}\right) - u$ QR's, and finally 1's comes from $n - Z\left(b^{(n)}\right)$ QNR's. Therefore, by separating 0, $T_{u,n}$, and T_n the combination of n-bit patterns can be given as follows.

$$T_{u,n} = {}_nC_u \cdot {}_{n-u}C_{Z\left(b^{(n)}\right)-u} \cdot \left(\frac{q-1}{2}\right)^{Z\left(b^{(n)}\right)} \times {}_{n-Z\left(b^{(n)}\right)}C_{n-Z\left(b^{(n)}\right)} \cdot \left(\frac{q-1}{2}\right)^{n-Z\left(b^{(n)}\right)} \tag{5.19}$$

Furthermore, T_n can be derived as,

$$\begin{aligned} T_n &= \sum_{u=0}^{Z\left(b^{(n)}\right)} T_{u,n} \\ &= \sum_{u=0}^{Z\left(b^{(n)}\right)} {}_nC_u \cdot {}_{n-u}C_{Z\left(b^{(n)}\right)-u} \cdot \left(\frac{q-1}{2}\right)^{Z\left(b^{(n)}\right)} \times \left(\frac{q-1}{2}\right)^{n-Z\left(b^{(n)}\right)} \end{aligned} \tag{5.20}$$

According to the above equation, T_n can be calculated by $Z\left(b^{(n)}\right)$. In addition, there are ${}_nC_{Z\left(b^{(n)}\right)}$ possible bit patterns that have the same $Z\left(b^{(n)}\right)$. To calculate the $D_{\mathcal{S}_{q^M-1}}\left(b^{(n)}\right)$ for each $b^{(n)}$, T_n needs to be divided by ${}_nC_{Z\left(b^{(n)}\right)}$.

$$\begin{aligned} D_{\mathcal{S}_{q^M-1}}\left(b^{(n)}\right) &= q^{M-(n\cdot m')} \cdot \frac{T_n}{{}_nC_{Z\left(b^{(n)}\right)}} \\ &= q^{M-(n\cdot m')} \sum_{u=0}^{Z\left(b^{(n)}\right)} \frac{{}_nC_u \cdot {}_{n-u}C_{Z\left(b^{(n)}\right)-u}}{{}_nC_{Z\left(b^{(n)}\right)}} \\ &\times \left(\frac{q-1}{2}\right)^{Z\left(b^{(n)}\right)-u} \cdot \left(\frac{q-1}{2}\right)^{n-Z\left(b^{(n)}\right)} \end{aligned} \tag{5.21}$$

The above equation can be further modified as follows.

$$\begin{aligned}
&\sum_{u=0}^{Z\left(b^{(n)}\right)} {}_nC_u \cdot {}_{n-u}C_{Z\left(b^{(n)}\right)-u} = \frac{n!}{\left(n-Z\left(b^{(n)}\right)\right)!} \cdot \sum_{u=0}^{Z\left(b^{(n)}\right)} \cdot \frac{1}{u!\left(Z\left(b^{(n)}\right)-u\right)!} \\
&= \frac{n!}{\left(n-Z\left(b^{(n)}\right)\right)!} \cdot \sum_{u=0}^{Z\left(b^{(n)}\right)} \cdot \frac{1}{u!\left(Z\left(b^{(n)}\right)-u\right)!} \times \frac{\left(Z\left(b^{(n)}\right)\right)!\left(n-Z\left(b^{(n)}\right)\right)!}{n!} \\
&= \sum_{u=0}^{Z\left(b^{(n)}\right)} \frac{(Z(b^{(n)}))!}{u!(Z(b^{(n)}))!} = {}_{Z\left(b^{(n)}\right)}C_{Z\left(b^{(n)}\right)-u}. \qquad (5.22)
\end{aligned}$$

Thus, (5.21) becomes as follows:

$$\begin{aligned}
D_{\mathcal{S}_{q^M-1}}\left(b^{(n)}\right) &= q^{M-(n\cdot m')} \sum_{u=0}^{Z\left(b^{(n)}\right)} {}_{Z\left(b^{(n)}\right)}C_{Z\left(b^{(n)}\right)-u} \\
&\quad \times \left(\frac{q-1}{2}\right)^{Z\left(b^{(n)}\right)-u} \cdot \left(\frac{q-1}{2}\right)^{n-Z\left(b^{(n)}\right)} \qquad (5.23)
\end{aligned}$$

By using the bilinear theorem [AST95], the above equation can be rewritten as,

$$\begin{aligned}
D_{\mathcal{S}_{q^M-1}}\left(b^{(n)}\right) &= q^{M-(n\cdot m')} \cdot \left(\frac{q-1}{2}\right)^{n-Z\left(b^{(n)}\right)} \times \left(\frac{q-1}{2}+1\right)^{Z\left(b^{(n)}\right)} \\
&= q^{M-(n\cdot m')} \cdot \left(\frac{q-1}{2}\right)^{n-Z\left(b^{(n)}\right)} \cdot \left(\frac{q+1}{2}\right)^{Z\left(b^{(n)}\right)}. \qquad (5.24)
\end{aligned}$$

From the (5.18), $D_{\mathcal{S}_\lambda}\left(b^{(n)}\right)$ holds the following relation as follows,

$$D_{\mathcal{S}_\lambda}\left(b^{(n)}\right) = D_{\mathcal{S}_{q^M-1}}\left(b^{(n)}\right) \cdot \left(\frac{q-1}{2}\right)^{-1}. \qquad (5.25)$$

Therefore, using the (5.24), $D_{\mathcal{S}_\lambda}\left(b^{(n)}\right)$ can be given by the following relation as,

$$\begin{aligned}
D_{\mathcal{S}_\lambda}\left(b^{(n)}\right) &= q^{M-(n\cdot m')} \cdot \left(\frac{q-1}{2}\right)^{n-Z\left(b^{(n)}\right)} \times \left(\frac{q+1}{2}\right)^{Z\left(b^{(n)}\right)} \cdot \left(\frac{q-1}{2}\right)^{-1} \\
&= q^{M-(n\cdot m')} \cdot \left(\frac{q-1}{2}\right)^{n-Z\left(b^{(n)}\right)-1} \cdot \left(\frac{q+1}{2}\right)^{Z\left(b^{(n)}\right)}. \qquad (5.26)
\end{aligned}$$

Thus, the first part of the (5.8a) is proven.

5.3.3.3 Proof of (5.8b)

Consider the case that $Z\left(b^{(n)}\right) = n$. Therefore, the combination of n-bit patterns except the all-zero patterns is given as follows:

$$ {}_nC_{Z\left(b^{(n)}\right)}. \tag{5.27} $$

Thus, the distribution of all-zero patterns becomes

$$ D_{\mathcal{S}_\lambda}\left(b^{(n)}\right) = \lambda - \sum_{u=0}^{n-1} {}_nC_u \cdot D_{\mathcal{S}_\lambda}\left(b^{(n)}\right). \tag{5.28} $$

Thus, the second part of the (5.8b) is proven. In addition, the theorem in (5.8) is also proven.

5.4 Result and Discussion

This section explains the distribution of bit patterns in the binary sequence which generated over the sub extension field based on some experimental results. Then, a comparison between the binary sequence defined over the sub extension field and the previous geometric sequence [NTU14] also introduces in terms of the distribution of bit patterns property. Throughout this section, the symbol $\mathrm{H_w}$ denotes the hamming weight.

5.4.1 Experimental Results

Consider the distribution of bit patterns in the binary sequence which generated over the sub extension field in the following examples.

Example 5.1 *Let* $p = 5, m = 4$, *and* $m' = 2$, *then the sequence having a period of* 52 *becomes as follows its distribution of* n*-bit patterns is shown in Table 5.5.*

$$ \mathcal{S}_{52} = \{0100110100000010101111001110110010111110010100001100\}. \tag{5.29} $$

Example 5.2 *Let* $p = 3, m = 6$, *and* $m' = 2$, *then the sequence having a period of* 182 *becomes as follows its distribution of* n*-bit patterns is shown in Table 5.6.*

Example 5.3 *Let* $p = 7, m = 9$, *and* $m' = 3$, *then the sequence having a period of* 235986 *becomes as follows its distribution of* n*-bit patterns is shown in Table 5.7.*

TABLE 5.5: Bit distribution of the binary sequence $\mathcal{S}_{52}$ with $p = 5, m = 4$, and $m' = 2$.

n	$\mathrm{H_w}(b^{(n)})$	$Z(b^{(n)})$	$D_{S_{52}}(b^{(n)})$
1	0	1	27
	1	0	25
2	0	2	14
	1	1	13
	2	0	12

TABLE 5.6: Bit distribution of the binary sequence $\mathcal{S}_{182}$ with $p = 3, m = 6$, and $m' = 2$.

n	$\mathrm{H_w}(b^{(n)})$	$Z(b^{(n)})$	$D_{S_{182}}(b^{(n)})$
1	0	1	101
	1	0	81
2	0	2	56
	1	1	45
	2	0	36
3	0	3	31
	1	2	25
	2	1	20
	3	0	16

Observation

It was found that the experimental results explicitly support (5.8). In addition, the number of appearance of each bit pattern is related to the number of zeros contained in each bit pattern. Moreover, $D_{\mathcal{S}_\lambda}(b^{(n)})$ increases in proportion to

TABLE 5.7: Bit distribution of the binary sequence $\mathcal{S}_{235986}$ with $p = 7, m = 9$, and $m' = 3$.

n	$\mathrm{H_w}(b^{(n)})$	$Z(b^{(n)})$	$D_{S_{235986}}(b^{(n)})$
1	0	1	118337
	1	0	117649
2	0	2	59341
	1	1	58996
	2	0	58653
3	0	3	29757
	1	2	29584
	2	1	29412
	3	0	29241

$Z(b^{(n)})$. To confirm this, check the **Example** 5.2 with $n = 3$.

$$Z(b^{(3)}) = 0 :$$
$$D_{S_{182}}(b^{(3)} = 111) = 3^{6-(3\times2)} \cdot 4^{3-0-1} \cdot 5^0 = 16.$$

$$Z(b^{(3)}) = 1 :$$
$$D_{S_{182}}(b^{(3)} = 011) = 3^{6-(3\times2)} \cdot 4^{3-1-1} \cdot 5^1 = 20,$$
$$D_{S_{182}}(b^{(3)} = 101) = 3^{6-(3\times2)} \cdot 4^{3-1-1} \cdot 5^1 = 20,$$
$$D_{S_{182}}(b^{(3)} = 111) = 3^{6-(3\times2)} \cdot 4^{3-1-1} \cdot 5^1 = 20.$$

$$Z(b^{(3)}) = 2 :$$
$$D_{S_{182}}(b^{(3)} = 001) = 3^{6-(3\times2)} \cdot 4^{3-2-1} \cdot 5^2 = 25,$$
$$D_{S_{182}}(b^{(3)} = 010) = 3^{6-(3\times2)} \cdot 4^{3-2-1} \cdot 5^2 = 25,$$
$$D_{S_{182}}(b^{(3)} = 100) = 3^{6-(3\times2)} \cdot 4^{3-2-1} \cdot 5^2 = 25.$$

$$Z(b^{(3)}) = 3 :$$
$$D_{S_{182}}(b^{(3)} = 000) = 182 - (1 \times 16 + 3 \times 20 + 3 \times 25) = 31.$$

5.4.2 Comparison With Previous Work

Previous work [NTU14] proposed a geometric sequence, namely NTU (Nogami-Tada-Uehara) sequence. It always holds long period, low correlation, high linear complexity properties which are the important considerations to use any sequence in cryptographic applications. Another crucial consideration before utilizing them in any secure applications, is to judge the randomness of a sequence. To do so, we need to evaluate the distribution of bit patterns property in a sequence. After the experimental observation, it was found that in terms of distribution of bit patterns NTU sequence is not uniformly distributed. In

other words, in case of binary NTU sequence, there is much difference in appearance between 0 and 1. To improve this drawback, instead of prime field (which used in the NTU sequence generation procedure), we focused on the sub extension field during the sequence generation procedure in this research work. As a result, after utilizing the sub extension field, the distribution of bit patterns becomes close to uniform. This comparison is shown in the following tables (Table 5.8 and Table 5.9).

It should be noted that the NTU sequence is controlled by 2 parameters (p and m), on the other hand the sequence over the sub extension field is controlled by 3 parameters (p, m, and m'). Therefore, it is not possible to make the comparison between these two sequences in terms of the same length (in other words, the same period λ). Although this chapter kept this difference as minimum as possible.

One of the most notable outcomes of this comparison result is the NTU sequence holds higher difference in terms of the appearance between the 'all zero' and 'all one' patterns. In other words, it also confirms the ununiform distribution of bit patterns. On the other hand, sequence defined over the sub extension field, that is the proposed sequence, minimizes this difference to make it close to uniform. This comparison is graphically shown in **Figure** 5.1.

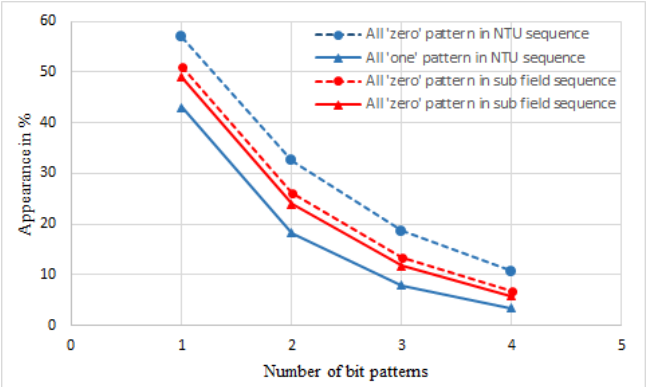

FIGURE 5.1: Appearance of 'all zero' and 'all one' bit patterns in the NTU and sub field sequence (proposed sequence).

Recently, there are lots of considerations to use a long period pseudo random sequence in cryptographic applications. The use of binary sequence in a stream cipher is one of the most common application. Before applying a sequence in such applications, the linear complexity and distribution of bit patterns are considered as the most important properties regarding a sequence to check its randomness. Among these two, the linear complexity property already observed in [Ali+17a] and it always holds a maximum value of the linear complexity. As a continuation, this chapter focuses on the distribution of bit patterns property. According to the comparison results, the binary sequence generated over the sub extension field holds much better (close to uniform) compared to previous binary sequence in terms of distribution of bit patterns. Therefore, the binary sequence defined over the sub extension field can be a suitable candidate for some cryptographic applications.

TABLE 5.8: Comparison in bit distribution between the sub field binary sequence and NTU sequence -I.

n	$\mathrm{H_w}(b^{(n)})$	$D_{S_{182}}(b^{(n)})$	%	$D_{NTU_{242}}(b^{(n)})$	%
1	0	101	55.49	161	66.52
	1	81	44.51	81	33.48
2	0	56	30.76	107	44.21
	1	45	24.72	54	22.31
	2	36	19.78	27	11.15
3	0	31	17.03	71	29.33
	1	25	13.73	36	14.87
	2	20	10.98	18	7.43
	3	16	8.79	9	3.71

5.5 Summary

This chapter observed the distribution of bit patterns in a binary sequence which defined over the sub extension field. The number of appearances is related to the number of zeros contained in each bit pattern. Furthermore, this chapter theoretically prove the distribution of bit patterns property. In addition, this chapter also made a comparison between the binary sequence defined over the sub extension field and prime field based on distribution of bit patterns property. According to the comparison results, the binary sequence generated over the sub extension field (proposed sequence) holds much better (close to uniform) compared to the binary sequence defined over the prime field.

TABLE 5.9: Comparison in bit distribution between the sub field binary sequence and NTU sequence -II.

n	$\mathrm{H_w}(b^{(n)})$	$D_{S_{240200}}(b^{(n)})$	%	$D_{NTU_{275514}}(b^{(n)})$	%
1	0	122551	51.02	156865	56.93
	1	117649	48.98	117649	43.07
2	0	62526	26.03	89637	32.53
	1	60025	24.98	67228	24.40
	2	57624	23.99	50421	18.30
3	0	31901	13.28	51221	18.59
	1	30625	12.74	38416	13.94
	2	29400	12.23	28812	10.45
	3	28224	11.75	21609	7.84
4	0	16276	6.77	29269	10.62
	1	15625	6.50	21952	7.96
	2	15000	6.24	16464	5.97
	3	14400	5.99	12348	4.48
	4	13824	5.75	9261	3.36

Chapter 6

Linear Complexity and NIST Test of Pseudo Random Sequence

6.1 Introduction

This chapter introduces experimental observation regarding the linear complexity and NIST statistical test suite. It is worth to know the linear complexity to judge the forward and backward unpredictability of a pseudo random sequence [LN96]. On the other hand, the NIST statistical test suite is a prominent statistical test to evaluate the randomness of a sequence [Ruk+00].

6.2 Linear Complexity

The unpredictability of a sequence can be measured by the length of the shortest Linear Feedback Shift Register (LFSR) which can generate the given sequence. This approach is particularly appealing since there exists an efficient procedure (it is so called the Berlekamp-Massy algorithm [AS07]) for finding the shortest LFSR. This length is referred to as the linear complexity associated with the sequence. The linear complexity property regarding a sequence is an important parameter which tells how difficult it is to predict the next bit pattern by observing the previous bit pattern of a sequence.

The linear complexity (LC) of a sequence is closely related to how difficult it is to guess the next bit after observing the previous bits of a sequence. The linear complexity of a binary sequence $\mathcal{S}$ having a period of n is defined as follows.

$$\mathrm{LC}(\mathcal{S}) = n - \deg\left(\gcd\left(x^n - 1, h_{\mathcal{S}}\left(x\right)\right)\right), \tag{6.1}$$

where $h_{\mathcal{S}}\left(x\right)$ of $\mathcal{S} = \{s_i\}$ is defined over $\mathbb{F}_2$ as,

$$h_{\mathcal{S}}(x) = \sum_{i=0}^{n-1} s_i x^i. \tag{6.2}$$

It should be noted that $\gcd(x^n - 1, h_S(s))$ in Eq.(6.1) needs to be calculated over $\mathbb{F}_2$. It is said that the linear complexity of the pseudo random sequence for security applications is preferred to be high.

The Berlekamp-Massy algorithm [AS07] is an efficient algorithm for determining the linear complexity of a finite binary sequence $\mathcal{S}_t$ of length t. This algorithm takes t iterations, with the T-th iteration computing the linear complexity of the sub-sequence $\mathcal{S}_T$ consisting of the first T terms of the sequence $\mathcal{S}_t$. The returned value of the linear complexity is the length of the shortest linear feedback shift register sequence that generates all bits in the sequence $\mathcal{S}_t$. According to the Berlekamp-Massy algorithm, the maximum linear complexity of a sequence is given by the following equation as follow.

$$\text{Linear Complexity (LC)} = \frac{l}{2},$$

where l denotes the length of the observed sequence. The value of the linear complexity $l/2$ is half of the observed sequence length l and it is maximum due to this algorithm observes the sequence $\mathcal{S}$ up to 2 times of its period. Although it well-known that an M-sequence is prominent for its uniform distribution of bit patterns, its linear complexity is very small (minimum). For example, the linear complexity of an M-sequence of becomes 6, whereas its period is 63 is shown in **Figure** 6.1. On the other hand, the Legendre sequence is renowned for its high (maximum) linear complexity. The **Figure** 6.2 shows the maximum linear complexity of a Legendre sequence (both of its period and linear complexity are 61).

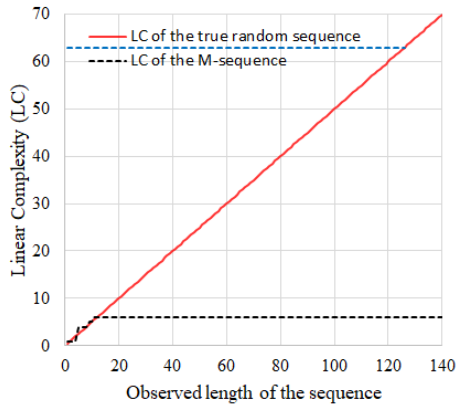

FIGURE 6.1: Linear complexity of an M-sequence.

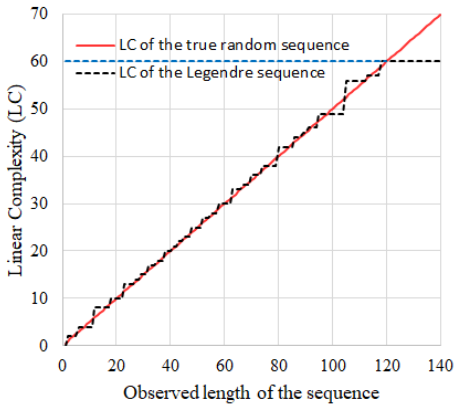

FIGURE 6.2: Linear complexity of the Legendre sequence.

It should be noted that the proposed sequence also holds the high linear complexity. As an example, the linear complexity of a sequence having a period of 14640 is shown in **Figure** 6.3.

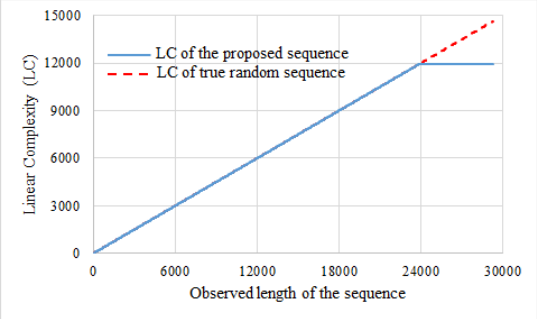

FIGURE 6.3: Linear complexity of the proposed sequence.

6.2.1 Comparison with Previous Work

The linear complexity of the proposed sequence (defined over sub extension field) and previous sequence (NTU) (defined over prime field) is shown in **Figure** 6.4 and **Figure** 6.5, respectively. By observing their linear complexity graph, it was found that the proposed sequence (which defined over the sub extension field) always hold high linear complexity compared to the NTU sequence. In other words, in terms of linear complexity the sequence defined over the sub extension field hold higher linear complexity than the sequence defined over the prime field.

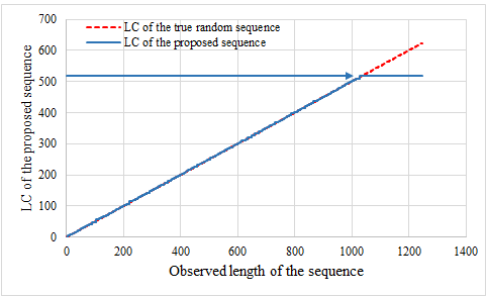

FIGURE 6.4: Linear complexity of the proposed sequence.

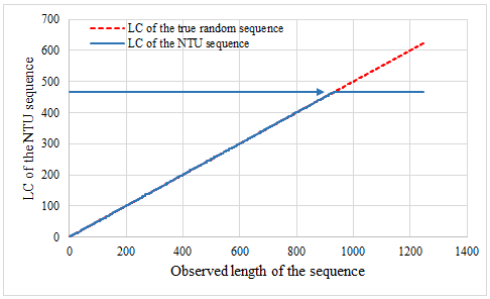

FIGURE 6.5: Linear complexity of the NTU sequence.

6.3 NIST Statistical Test

The need for pseudo random sequences (or numbers) arises in many cryptographic applications such as the generation of key materials. Sequences suitable for use in such applications may need to meet stronger requirements than for other applications. The randomness regarding sequences is one of the requirements, which ensures the unpredictability of their outputs in the absence of knowledge of the inputs. Various statistical tests can be applied to a sequence to attempt to compare and evaluate the sequence to a truly random sequence. Basically, randomness is a probabilistic property, which means the properties of a sequence can be characterized and described in terms of probability. This

section discusses the randomness testing of pseudo random sequences, which is so-called the National Institute of Standards and Technology (NIST) statistical test suite [Ruk+00].

The NIST test suite is a statistical package consisting of 16 tests that were developed to test the randomness of a binary sequence. These tests focus on a variety of different types of non-randomness that could exist in a sequence. These tests are as follows:

- The frequency (monobit) test,
- Frequency test within a block,
- The runs test,
- Test for the longest-run-of-ones in a block,
- The binary matrix rank test,
- The discrete Fourier transform test,
- The non-overlapping template matching test,
- The overlapping template matching test,
- Maurer's "universal statistical" test,
- The Lempel-Ziv compression test,
- The linear complexity test,
- The serial test,
- The approximate entropy test,
- The cumulative sums test,
- The random excursions test, and
- The random excursions variant test.

In the subsequent section, provides a brief introduction about the testing strategy and interpretation of these test results. Before going to the details of these tests, few relevant definitions are as follows.

Null Hypothesis: A statistical test is formulated to test a specific null hypothesis. For the purpose of the following section, null hypothesis under test is that the sequence being tested is random.

P-value: The test statistic is used to calculate a P-value that summarizes the strength of the evidence against the null hypothesis. In addition, each P-value is the probability that a perfect random number generator would have produced a sequence, less random than the sequence that was tested. If a P-value for a test is determined to be equal to 1, then the sequence appears

to have perfect randomness. A P-value of *zero* indicates that the sequences appears to be completely non-random.

Level of Significance: A significance level (α) can be chosen for the tests. If P-value $\geq \alpha$, then the null hypothesis is accepted; i.e. the sequence appears to be random. If P-value $< \alpha$, then the null hypothesis is rejected; i.e. the sequence appears to be non-random.

Frequency Test within a Block

The purpose of this test is to determine whether the frequency of ones in an M-bit block is approximately $M/2$, (where M is the number of bits in a substring (block) being tested), as would be expected under an assumption of randomness. If the computed P-value is < 0.01, then it can be concluded that the sequence is non-random. Otherwise, conclude that the sequence is random. Since the P-value obtained in sub field NTU sequence is ≥ 0.01 (i.e., P-value = 0.090936), the conclusion is that the sequence is random.

Runs Test

The focus of this test is the total number of runs in the sequence, where a run is an uninterrupted sequence of identical bits. A run of length r consists of exactly r identical bits. The purpose of the runs test is to determine whether the number of runs of ones and zeros of various lengths is as expected for a random sequence. If the computed P-value is < 0.01, then it can be concluded that the sequence is non-random. Otherwise, conclude that the sequence is random. Since the P-value obtained in sub field NTU sequence is ≥ 0.01 (i.e., P-value = 0.834308), the conclusion is that the sequence is random.

Test for the Longest Run of Ones in a Block

The focus of this test is the longest run of ones within M-bit blocks. Basically, the purpose of this test is to determine whether the length of the longest run of ones within the tested sequence is consistent with the length of the longest run of ones that would be expected in a random sequence. If the computed P-value is < 0.01, then it can be concluded that the sequence is non-random. Otherwise, conclude that the sequence is random. Since the P-value obtained in sub field NTU sequence is ≥ 0.01 (i.e., P-value = 0.699313), the conclusion is that the sequence is random.

Binary Matrix Rank Test

The focus of this test is on the rank of disjoint sub-matrices of the entire sequence. This test checks for linear dependence among fixed length sub-strings of the original sequence. If the computed P-value is < 0.01, then it can be concluded that the sequence is non-random. Otherwise, conclude that the sequence is random. Since the P-value obtained in sub field NTU sequence is ≥ 0.01 (i.e., P-value = 0.224821), the conclusion is that the sequence is random.

Discrete Fourier Transform Test

The purpose of this test is to find the peak height in the discrete Fourier transform of the sequence. It can detect periodic features (such as repetitive patterns that are near each other) in the tested sequence that would indicate a deviation from the assumption of randomness. The intention is to detect whether the number of peaks exceeding the 95% threshold is significantly different than 5%. If the computed P-value is < 0.01, then it can be concluded that the sequence is non-random. Otherwise, conclude that the sequence is random. Since the P-value obtained in sub field NTU sequence is ≥ 0.01 (i.e., P-value = 0.851383), the conclusion is that the sequence is random.

Non-overlapping Template Matching Test

The focus of this test is the number of occurrences of pre-specified target strings. This test can detect generators that produce too many occurrences of a given non-periodic (aperiodic) pattern. For this test and over lapping template matching test, an m-bit window is used to search for a specific m-bit pattern. If the pattern is not found, the window slides by one bit position, on the other hand, if the pattern is found, the window is reset to the bit after the found pattern, and the search resumes. If the computed P-value is < 0.01, then it can be concluded that the sequence is non-random. Otherwise, conclude that the sequence is random. Since the P-value (for all the non-overlapping template matching) obtained in sub field NTU sequence is ≥ 0.01, the conclusion is that the sequence is random.

Overlapping Template Matching Test

The purpose of the overlapping template matching test is the number of occurrences of pre-specified target strings. In case of non-overlapping template matching test, if the pattern is not found, the window slides one bit position, whereas, overlapping template matching test slides the window one bit position, when it finds the pattern before resuming the search. If the computed P-value is < 0.01, then it can be concluded that the sequence is non-random. Otherwise, conclude that the sequence is random. Since the P-value obtained in sub field

NTU sequence is ≥ 0.01 (i.e., P-value = 0.035174), the conclusion is that the sequence is random.

Linear Complexity Test

The focus of this test in on the length of a linear feedback shift register (LFSR). This test can determine whether the sequence is complex enough to be considered random. Basically, random sequences can be characterized by longer LFSRs. If the computed P-value is < 0.01, then it can be concluded that the sequence is non-random. Otherwise, conclude that the sequence is random. Since the P-value obtained in sub field NTU sequence is ≥ 0.01 (i.e., P-value = 0.021999), the conclusion is that the sequence is random.

Serial Test

The focus of this test is the frequency of all possible overlapping m-bit patterns across the entire sequence. This test can determine whether the number of occurrences of the 2^m m-bit overlapping patterns is approximately the same as would be expected for a random sequence. Random sequences have uniformity, which means every m-bit pattern has the same chance of appearing as every other m-bit pattern. If the computed P-value is < 0.01, then it can be concluded that the sequence is non-random. Otherwise, conclude that the sequence is random. Since the P-value obtained in sub field NTU sequence is ≥ 0.01 (i.e., P-value = 0.798139), the conclusion is that the sequence is random.

Random Excursions Test

The focus of this test is the number of cycles having exactly K visits in a cumulative sum random walk. This random walk is derived from partial sums after the binary sequence (having the values $0, 1$) is transferred to the appropriate $(-1, +1)$ sequence. A cycle of random walk consists of a sequence of steps of unit length taken at random that begin at the return to the origin. This test is a series of eight tests (one test and conclusion for each of the four states). According to the random excursions test result, sub field NTU sequence is random.

Random Excursions Variant Test

The focus of this test is on the total number of times that a state is occurs in a cumulative sum random walk. This test can detect deviations from the expected number of visits to various states in the random walk. Random excursions variant test is a series of eighteen tests (one test and conclusion for each of the nine states). According to the random excursions variant test result, sub field NTU sequence is random.

A summary of the NIST statistical test suite is shown in the following table (**Table** 6.1).

TABLE 6.1: A summary of the NIST statistical test results of the proposed sequence and NTU sequence.

Name of the test	(proposed sequence) pass (◯) / fail (☒)	(NTU sequence) pass (◯) / fail (☒)
Block frequency test	◯	◯
Runs test	◯	◯
Longest run test	◯	◯
Rank test	◯	◯
DFT test	◯	◯
Non-overlapping template test	◯	◯
Overlapping template test	◯	◯
Linear complexity test	◯	◯
Serial test	◯	◯
Random excursion test	◯	◯
Random excursion variant	◯	◯

6.4 Summary

This chapter experimentally observes the linear complexity and NIST statistical test results. In terms of linear complexity, the proposed sequence holds high linear complexity. From the comparison viewpoint, the proposed sequence which defined over the sub extension field holds much higher linear complexity compared to the sequence defined over the prime field. To judge the randomness of a sequence, the NIST statistical test suite is well-known. The proposed sequence in this dissertation mostly passed all the steps in the NIST test.

Chapter 7

Relationship Between Sequences and Polynomials

7.1 Introduction

Background and Motivation

Pseudo-random sequences have been widely employed in the field of information security and cryptography as a key stream of one-time pads, secret keys of symmetric cipher system, public key parameters, and so on [TCA98; Gol67; GG05; MVO96]. To ensure the security of these cryptosystems, the pseudo-random sequence should have unpredictable random quantities, as well as sequence needs to generate very rapidly. The unpredictability of a sequence can be achieved by using some nonlinear mathematical calculation during the sequence generation procedure. On the other hand, to generate a sequence very swiftly, the expensive calculation needs to be avoided.

Related Works

Other pseudo-random sequences, such as M-sequence [Nea59; CTW98] and Legendre sequence [Zie58; JS+96] have been researched well due to most of their important properties already theoretically proven. Previous work [NTU14] proposed a new binary sequence by combining the features of an M-sequence and Legendre sequence. The generation procedure of the previous sequence includes three steps. Firstly, it utilized a primitive polynomial primitive polynomial to generate maximum length vector sequence. This idea comes from M-sequence. Then, the trace function is applied to transform all the vectors to scalars. Finally, Legendre symbol applied to the scalars to generate the binary sequence. It should be noted that Legendre symbol is a non-linear function, that's why previous sequence holds unpredictability quality. This sequence has some prominent features also such as period, autocorrelation, and linear complexity, whereas, all these properties have been theoretically proven in [NTU14].

Contribution

As previously mentioned, during the sequence generation procedure, in the beginning, a primitive polynomial used. However, finding a primitive polynomial takes much calculation cost when the degree or the characteristic is large. In other words, it takes much longer time to find a higher degree primitive polynomial. To overcome this complication (finding a primitive polynomial during the sequence generation procedure), this chapter focused on the properties of the irreducible polynomials (which generates the proposed sequence) to find some prominent features, by which we can find an efficient way of generating the sequence. This dissertation accomplishes some experiment on generating a sequence by using the previous method. After careful observation, we found a relation between the irreducible polynomial and a sequence. The purpose of finding this relation is to generate the same sequence without the primitive polynomial. After the experimental observation, this chapter found that there are $(p-1)/2$ kinds of polynomial exist, which generates the same sequence and some of these irreducible polynomials are non-primitive polynomials. In other words, according to the observation of this chapter, the same binary sequence can be generated without using a primitive polynomial. This idea can overcome the drawback (finding primitive polynomial, which is an expensive calculation) of the previous work [NTU14]. The relation between the polynomial and sequence are mentioned as theorems in this chapter. Furthermore, these theorems are theoretically proven as well justified with the aid of some example.

7.2 Consideration of the Polynomials

This chapter claims that there are $(p-1)/2$ kinds of different irreducible polynomials that can generates the same sequence (which generated by the previous method with a certain primitive polynomial [NTU14]). This section firstly introduces the relation between the polynomials with the aid of some examples. In addition, these relations also explained mathematically as theorems in the later part of this section.

7.2.1 Examples

This chapter considers the relation between the irreducible polynomials over the odd characteristic field. Therefore, to make the better understanding (the concept of this chapter) for the readers, this chapter utilizes a small prime number 7 as p and its extension degree 2 as m to construct an odd characteristic field. Let $p = 7$, $m = 2$, and $f(x) = x^2 + 6x + 3$, then the generated sequence is shown in Eq.(3.13). If we utilize other irreducible polynomials such as x^2+5x+5 or $x^2 + 3x + 6$, instead of $x^2 + 6x + 3$, then the generated sequence also becomes the same. In other words, the following polynomials generate the same binary

sequence.

$$\begin{aligned} f_0(x) &= x^2 + 6x + 3, \\ f_1(x) &= x^2 + 5x + 5, \\ f_2(x) &= x^2 + 3x + 6. \end{aligned} \tag{7.1}$$

Again, let $p = 11$, $m = 2$, and $f(x) = x^2 + 5x + 2$, then the generated sequence becomes as follows.

$$\mathcal{T} = \{1, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 0\}. \tag{7.2}$$

Alike the former case, there are more polynomials which can generate the same sequence which is shown in Eq.(7.2). These polynomials are as follows.

$$\begin{aligned} f_0(x) &= x^2 + 5x + 2, \\ f_1(x) &= x^2 + 9x + 10, \\ f_2(x) &= x^2 + 3x + 6, \\ f_3(x) &= x^2 + x + 8, \\ f_4(x) &= x^2 + 4x + 7. \end{aligned} \tag{7.3}$$

From these examples, it should be noted that there are $(p-1)/2$ kinds of irreducible polynomials for the case of $p = 7$ and 11 which generates the same sequence. Moreover, it is also found that there exists a non-primitive polynomial for each case. Actually, $x^2 + 3x + 6$ in Eq.(7.1) and $x^2 + 9x + 10$ in Eq.(7.3) are non-primitive polynomials.

7.2.2 Relation Between the Polynomials

As an example, the polynomials in Eq.(7.3) are considered in this section. Let ω_0, ω_1, ω_2, ω_3, ω_4 be the zeros of $f_0(x)$, $f_1(x)$, $f_2(x)$, $f_3(x)$, $f_4(x)$, respectively. Then, they hold the following relations.

$$\begin{aligned} \omega_0 &= \omega_0^{0\times 24+1} = 4^0\omega_0 \bmod f_0(\omega_0), \\ \omega_1 &= \omega_0^{1\times 24+1} = 4^1\omega_0 \bmod f_0(\omega_0), \\ \omega_2 &= \omega_0^{2\times 24+1} = 4^2\omega_0 \bmod f_0(\omega_0), \\ \omega_3 &= \omega_0^{3\times 24+1} = 4^3\omega_0 \bmod f_0(\omega_0), \\ \omega_4 &= \omega_0^{4\times 24+1} = 4^4\omega_0 \bmod f_0(\omega_0). \end{aligned} \tag{7.4}$$

In this case, we could find that $f_2(x)$ becomes the minimum polynomial of $\omega_0^{2\times 24+1} = \omega_0^{49}$ as,

$$\begin{aligned} f_2\left(\omega_0^{2\times 24+1}\right) &= \left(4^2\omega_0\right)^2 + 3\left(4^2\omega_0\right) + 6 \\ &= 3\omega_0^2 + 4\omega_0 + 6 \\ &= 3\left(\omega_0^2 + 5\omega_0 + 2\right) \\ &\equiv 0 \bmod f_0(\omega_0). \end{aligned} \tag{7.5}$$

Furthermore, it is also shown that the other polynomials become the minimal polynomial of given ω_0 as,

$$\begin{aligned} f_1\left(4^1\omega_0\right) &= 5\omega_0^2 + 3\omega_0 + 10 \\ &= 5\left(\omega_0^2 + 5\omega_0 + 2\right) \\ f_3\left(4^3\omega_0\right) &= 4\omega_0^2 + 9\omega_0 + 8 \\ &= 4\left(\omega_0^2 + 5\omega_0 + 2\right) \\ f_4\left(4^4\omega_0\right) &= 9\omega_0^2 + \omega_0 + 7 \\ &= 9\left(\omega_0^2 + 5\omega_0 + 2\right). \end{aligned} \tag{7.6}$$

Thus, $\omega_0^{5\times 24+1}$ becomes ω_0^1 because of the period of the primitive element in this case is 120, which means there are no more polynomials that satisfies these relations. Therefore, it can be found that there are $5\,(= (11-1)/2)$ kinds of polynomials that generate the same binary sequence.

Additionally, it is found that the non-primitive polynomial $f_1(x)$ holds the following relation with its zero $\omega_1 = \omega_0^{1\times 24+1}$ as,

$$\gcd(1 \cdot 24 + 1, 120) \neq 1. \tag{7.7}$$

In the next section, such relations are considered from the theoretical aspect.

7.2.3 Theorems and Its Proofs

After a lot of experimental observation, following theorems are obtained. This section mathematically prove these theorems.

Theorem 7 *Let $f_0(x)$ and ω_0 be a primitive polynomial and its zero, respectively. The polynomial $f_k(x)$, (where $k = 0, 1, \ldots, ((p-1)/2) - 1$) generates the same sequence, if it's zero ω_k satisfies the following condition.*

$$\omega_k \equiv \omega_0^{2k\left(\frac{q-1}{p-1}\right)+1} \bmod f_0(\omega_0). \tag{7.8}$$

■

(Proof) Let $\mathcal{T}_0 = \{t_{0,i}\}$ be the sequence which is generated by the polynomial $f_0(x)$. Again let $\mathcal{T}_k = \{t_{k,i}\}$ be the sequence generated by the polynomial $f_k(x)$.

Then, these sequences can be described as follows.

$$\begin{aligned} t_{0,i} &= M_2\left(\left(\mathrm{Tr}\left(\omega_0^i\right)\Big/p\right)\right), \\ t_{k,i} &= M_2\left(\left(\mathrm{Tr}\left(\omega_k^i\right)\Big/p\right)\right). \end{aligned} \tag{7.9}$$

According to the condition of the theorem and the property of a primitive element (**Property** 1), $t_{k,i}$ can be further modified as follows.

$$\begin{aligned} t_{k,i} &= M_2\left(\left(\mathrm{Tr}\left(\omega_0^{\left(2k\left(\frac{q-1}{p-1}\right)+1\right)i}\right)\Big/p\right)\right) \\ &= M_2\left(\left(\mathrm{Tr}\left(\omega_0^{\left(\frac{q-1}{p-1}\right)2ki}\cdot\omega_0^i\right)\Big/p\right)\right) \\ &= M_2\left(\left(\mathrm{Tr}\left(g^{2ki}\cdot\omega_0^i\right)\Big/p\right)\right). \end{aligned} \tag{7.10}$$

where g is a generator in $\mathbb{F}_p^*$. Then, from the linearity of the trace function (**Property** 2) and the calculation of the Legendre symbol (**Property** 3), the above formula can be described by the multiplicative property of the Legendre symbol as,

$$\begin{aligned} t_{c,i} &= M_2\left(\left(\left(g^{ki}\right)^2\mathrm{Tr}\left(\omega_0^i\right)\Big/p\right)\right) \\ &= M_2\left(\underline{\left(\left(g^{ki}\right)^2\Big/p\right)}\left(\mathrm{Tr}\left(\omega_0^i\right)\Big/p\right)\right). \end{aligned} \tag{7.11}$$

Because of $\left(g^{ki}\right)^2$ is a QR, thus the value of the underline part in the above equation becomes 1. Then, finally the sequence $t_{k,i}$ can be obtained as follows.

$$t_{k,i} = M_2\left(\left(\mathrm{Tr}\left(\omega_0^i\right)\Big/p\right)\right), \tag{7.12}$$

which means that the formula is same as $t_{1,i}$ for arbitrary i, therefore, the theorem has been proven. ∎

Let ω_0 be a primitive element of $g_0(x)$ (where $g_0(x)$ be a primitive polynomial) which belongs to the extension field $\mathbb{F}_{p^m}$ and it can generate a maximum of p^m-1 elements as like the red circle as shown in **Figure** 7.1. Therefore, ω_0 holds the following relation as,

$$\omega_0^{q-1} = \omega_0^{\frac{2(q-1)}{p-1}\times\frac{p-1}{2}} = 1.$$

Let the underlined part in the above equation be denoted as k. Then the above equation becomes as follows.

$$\omega_0^{\frac{2(q-1)}{p-1}\cdot k} = 1, (\text{where } k = 0, 1, 2, \ldots, \frac{p-1}{2} - 1).$$

It should be noted that each value of k will generate different elements $(\omega_1, \omega_2, \ldots)$ in $\mathbb{F}_{p^m}$, which can represent as a power of ω and these will become the zeros of different polynomials $(g_1(x), g_2(x), \ldots)$ as shown in **Figure** 7.1. Furthermore, all these polynomials are irreducible polynomials. It should be noted that a primitive polynomial is a special kind of irreducible polynomial.

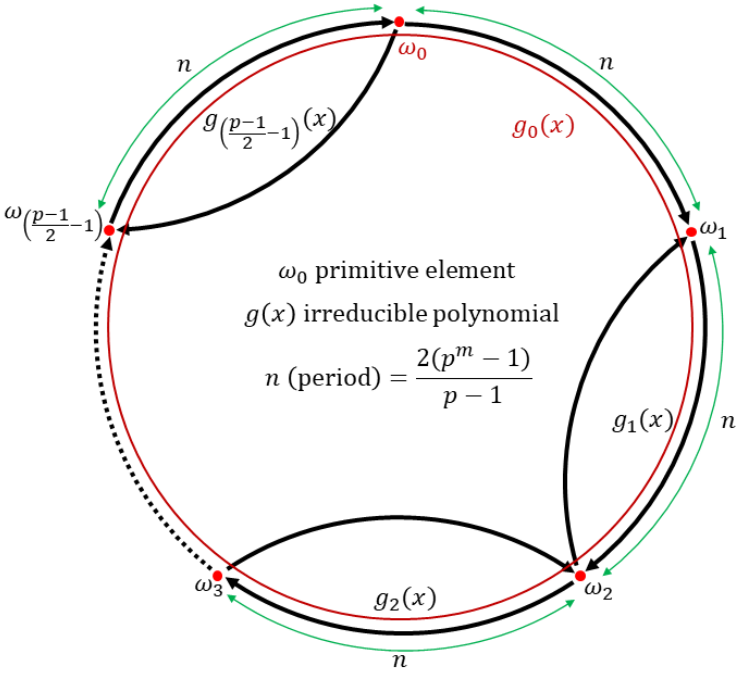

FIGURE 7.1: Relation between the irreducible polynomials.

Lemma 1 *In the* ***Theorem*** *7,* $f_k(x)$ *for every* k *becomes an irreducible polynomial of degree* m. ■

(Proof) According to the condition of **Theorem** 7, the following relation should be satisfied.

$$f_k(\omega_k) = f_k\left(g^{2k}\omega_0\right) \equiv 0 \bmod f_0(\omega_0). \tag{7.13}$$

If $f_0(\omega_0)$ be an irreducible polynomial of degree m, then $f_k\left(g^{2k}\omega_0\right)$ should be also degree m irreducible polynomial to satisfy the above equation. When $f_k(x)$ is given as,

$$f_k(x) = x^m + c_{m-1}x^{m-1} + \ldots + c_1x + c_0, \tag{7.14}$$

where c_i are the coefficients of $f_k(x)$. Then, $f_k\left(g^{2k}\omega_0\right)$ also becomes an irreducible polynomial of degree m as,

$$f_k\left(g^{2k}\omega_0\right) = \left(g^{2k}\right)^m \omega_0^m + c_{m-1}\left(g^{2k}\right)^{m-1}\omega_0^{m-1} + \ldots + c_1 g^{2k}\omega_0 + c_0. \tag{7.15}$$

■

Lemma 2 *In the* ***Theorem*** *7, $f_k(x)$ becomes a non-primitive polynomial if its zero $\omega_0^{2k((q-1)/(p-1))+1}$ holds the following relation as,*

$$\gcd\left(2k\left(\frac{q-1}{p-1}\right)+1, q-1\right) \neq 1. \tag{7.16}$$

■

(Proof) This lemma is for the case that ω_k is not a primitive element of $f_k(x)$ and it can be proven by the following contradiction.

Let ω_k be an element in the proper subfield $\mathbb{F}_{q'}$ (where $\mathbb{F}_{q'} \subset \mathbb{F}_q$). According to the **Property** 1, ω_k needs to be represented by the generator g and zero of a polynomial ω_0 as,

$$\omega_k = g^{2k}\omega_0. \tag{7.17}$$

The above equation can be rewritten as follows.

$$\omega_0 = g^{-2k}\omega_k. \tag{7.18}$$

The RHS of the above equation be an element of the subfield $\mathbb{F}_q$ and ω_0 also belongs to $\mathbb{F}_q$ ($\omega_0 \notin \mathbb{F}_{q'}$).

This is a contradiction. Thus, the above lemma is proven. ■

7.3 Summary

Previous work uses a primitive polynomial during the sequence generation procedure. Although the calculation cost for finding a primitive polynomial is high. However, from the viewpoint of security, if we use a sequence in some application, then the sequence needs to be generated very rapidly. This chapter concludes that the same sequence (which generated by the previous work [NTU14]) can be generated by some non-primitive polynomial. In addition, the following facts have been considered by the experimental observations and these are mathematically proven also.

- There are $(p-1)/2$ irreducible polynomials which generates the same sequence (here, this number $(p-1)/2$ related to the number of QR in $\mathbb{F}_p$).
- Some of the polynomials can be a non-primitive polynomial.

As mentioned previously, if the degree or the characteristic is large, then in that case finding a primitive polynomial takes much calculation cost. A primitive polynomial is one kind of special polynomial which can generate maximum

length vector sequence (for example, if $p = 11$ and $m = 2$, then a primitive polynomial can generate $p^m - 1 = 120$ distinct vectors without any duplication). There are only a few primitive polynomials are existing in the extension field. Therefore, when the characteristic field or the degree becomes large, then, to find a primitive polynomial becomes a time-consuming operation. Thus, the calculation cost will be reduced in previous work, if the sequence can generate without using a primitive polynomial.

Chapter 8

Conclusion and Future Works

This dissertation proposed a pseudo random sequence generated by cascaded trace function and Legendre symbol over the sub extension field. To confirm the randomness and unpredictability, sequence properties such as period, correlation, linear complexity, and distribution of bit patterns should be well studied. This thesis concentrated on the theoretical aspect rather than the application of pseudo random sequences. Consequently, properties regarding a sequence are theoretically proven.

Chapter 1 introduces the background, motivation, and contributions of the study in this thesis. **Chapter** 2 describes and defines the fundamental mathematical concepts. **Chapter** 3 introduces the proposed pseudo random sequence along with NTU sequence. In **Chapter** 4, the period, autocorrelation, and cross-correlation properties of the proposed sequence are theoretically proven along with experimental and comparison results. Considering the correlation, except the maximum peak, other peaks are suppressed to be very small in the proposed sequence, while the high peaks are appearing in the sequences generated over the prime field. **Chapter** 5 contains the theoretical proof of the distribution of bit patterns property. Regarding the distribution of bit patterns, sequence (defined over the sub extension field) holds much better (close to uniform) bit distribution compared to the sequence (defined over the prime field). **Chapter** 6 explains the linear complexity and NIST statistical test experimentally to judge the randomness of the proposed sequence. **Chapter** 7 finds a relationship between the generated sequences and polynomials to increase the candidates for the selection of polynomials. This relationship is justified both theoretically and experimentally.

As future works, a uniformization algorithm can be utilized to make the distribution of bit patterns uniform without degrading the linear complexity, which will improve the proposed pseudo random sequence in this thesis. Apart from this, mathematically prove the linear complexity property and to introduce more efficient calculation instead of the power residue symbol like expensive calculation will be another important future works. The ultimate target is to use proposed pseudo random sequence in some suitable real applications.

Bibliography

[Ali+16a] Md. Arshad Ali, Yasuyuki Nogami, Chiaki Ogawa, Hiroto Ino, Satoshi Uehara, Robert H. Morelos-Zaragoza, and Kazuyoshi Tsuchiya. "A new approach for generating well balanced Pseudo-random signed binary sequence over odd characteristic field". In: *2016 International Symposium on Information Theory and Its Applications (ISITA)*. 2016, pp. 777–780.

[Ali+16b] Md. Arshad Ali, Yasuyuki Nogami, Hiroto Ino, and Satoshi Uehara. "Auto and Cross Correlation of Well Balanced Sequence over Odd Characteristic Field". In: *2016 Fourth International Symposium on Computing and Networking (CANDAR)*. 2016, pp. 604–609. DOI: 10.1109/CANDAR.2016.0109.

[Ali+17a] Md. Arshad Ali, Takeru Miyazaki, Shoji Heguri, Yasuyuki Nogami, Satoshi Uehara, and Robert H. Morelos-Zaragoza. "Linear complexity of pseudo random binary sequence generated by trace function and Legendre symbol over proper sub extension field". In: *2017 Eighth International Workshop on Signal Design and Its Applications in Communications (IWSDA)*. 2017, pp. 84–88. DOI: 10.1109/IWSDA.2017.8095741.

[Ali+17b] Md. Arshad Ali, Takeru Miyazaki, Yasuyuki Nogami, Satoshi Uehara, and Robert H. Morelos-Zaragoza. "Multi-value sequence generated by trace function and power residue symbol over proper sub extension field". In: *2017 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW)*. 2017, pp. 249–250. DOI: 10.1109/ICCE-China.2017.7991089.

[AS07] Alexandra Alecu and Ana Sălăgean. "Modified Berlekamp-Massey Algorithm for Approximating the k-Error Linear Complexity of Binary Sequences". In: *Cryptography and Coding*. Ed. by Steven D. Galbraith. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 220–232. ISBN: 9783540772729. DOI: 10.1007/978-3-540-77272-9_14.

[AST95] M. Mathai Arak, B. Provost Serge, and Hayakawa Takesi. *Bilinear Forms and Zonal Polynomials*. Springer-Verlag. Springer, New York, 1995. ISBN: 9780387945224. DOI: 10.1007/978-1-4612-4242-0.

[BBS86] Lenore Blum, Manuel Blum, and Michael Shub. "A Simple Unpredictable Pseudo-Random Number Generator". In: *SIAM Journal on Computing* 15.2 (1986), pp. 364–383. DOI: 10.1137/0215025.

[Ber15] Elwyn R Berlekamp. *Algebraic Coding Theory*. World Scientific, 2015. ISBN: 9789814635899. DOI: 10.1142/9407.

[CD09] Ying Cai and Cunsheng Ding. "Binary Sequences with Optimal Autocorrelation". In: *Theoretical Computer Science* 410.24-25 (2009), pp. 2316–2322. ISSN: 0304-3975. DOI: 10.1016/j.tcs.2009.02.021.

[CE49] Shannon Claude Elwood. "Communication theory of secrecy systems". In: *The Bell System Technical Journal* 28.4 (1949), pp. 656–715. ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1949.tb00928.x.

[CTW98] Ding Cunsheng, Hesseseth Tor, and Shan Weijuan. "On the linear complexity of Legendre sequences". In: *IEEE Transactions on Information Theory* 44.3 (1998), pp. 1276–1278. ISSN: 0018-9448. DOI: 10.1109/18.669398.

[Cur05] Matt Curtin. *Brute Force: Cracking the Data Encryption Standard*. 1st. Copernicus Books, Springer-Verlag, New York, 2005. ISBN: 9780387201092. DOI: 10.1007/b138699.

[Dam90] Ivan Bjerre Damgård. "On The Randomness of Legendre and Jacobi Sequences". In: *Advances in Cryptology — CRYPTO' 88*. Ed. by Shafi Goldwasser. New York, NY: Springer New York, 1990, pp. 163–172. ISBN: 9780387347998. DOI: 10.1007/0-387-34799-2_13.

[DH76] Whitfield Diffie and Martin E. Hellman. "New directions in cryptography". In: *IEEE Transactions on Information Theory* 22.6 (1976), pp. 644–654. ISSN: 0018-9448. DOI: 10.1109/TIT.1976.1055638.

[DiC12] David DiCarlo. "Random Number Generation: Types and Techniques". PhD thesis. Liberty University, 2012.

[Din98] Cunsheng Ding. "Pattern distributions of Legendre sequences". In: *IEEE Transactions on Information Theory* 44.4 (1998), pp. 1693–1698. ISSN: 0018-9448. DOI: 10.1109/18.681353.

[Ede14] Vladimir Edemskiy. "On the linear complexity of interleaved binary sequences of period $4p$ obtained from Hall sequences or Legendre and Hall sequences". In: *IET Electronics Letters* 50.8 (2014), pp. 604–604. DOI: 10.1049/el.2014.0568.

[GG05] Solomon W. Golomb and Guang Gong. *Signal Design for Good Correlation: For Wireless Communication, Cryptography, and Radar*.

Cambridge University Press, 2005. ISBN: 9780511546907. DOI: 10.1017/CBO9780511546907.

[GK12] Mark Goresky and Andrew Klapper. *Algebraic feedback shift registers*. Cambridge University Press, 2012. DOI: 10.1017/CBO9781139057448.

[Gol67] Solomon W. Golomb. *Shift Register Sequences*. Holden-Day series in information systems. San Francisco : Holden-Day, 1967.

[Gu16] Ting Gu. "Statiscal Properties of Pseudorandom Sequences". PhD thesis. University of Kentucky, 2016. DOI: 10.13023/ETD.2016.159.

[HA06] Aly Hassan and Winterhof Arne. "On the k-error linear complexity over $\mathbb{F}_p$ of Legendre and Sidelnikov sequences". In: *Designs, Codes and Cryptography* 40.3 (2006), pp. 369–374. DOI: 10.1007/s10623-006-0023-5.

[Hel11] Tor Helleseth. *Maximal-Length Sequences*. Ed. by Henk C. A. Van Tilborg and Sushil Jajodia. Boston, MA: Springer US, 2011, pp. 763–766. ISBN: 9781441959065. DOI: 10.1007/978-1-4419-5906-5_359.

[Hen+05] Cohen Henri, Frey Gerhard, Avanzi Roberto, Doche Christophe, Lange Tanja, Nguyen Kim, and Vercauteren Frederik. *Handbook of Elliptic and Hyperelliptic Curve Cryptography*. Taylor & Francis Group. Chapman and Hall CRC, 2005. ISBN: 9780429139659. DOI: 10.1201/9781420034981.

[Her05] Doreen Hertel. "Cross-Correlation Properties of Perfect Binary Sequences". In: *Sequences and Their Applications - SETA 2004*. Ed. by Tor Helleseth, Dilip Sarwate, Hong-Yeop Song, and Kyeongcheol Yang. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 208–219. ISBN: 9783540320487. DOI: 10.1007/11423461_14.

[HK98] Tor Helleseth and P. Vijay Kumar. *Sequences with low correlation," in Handbook in Coding Theory*. Vol. II. Elsevier Science B.V, 1998, pp. 1765–1853.

[JS+96] No Jong-Seon, Lee Hwan-Keun, Chung Habong, Song Hong-Yeop, and Yang Kyeongcheol. "Trace representation of Legendre sequences of Mersenne prime period". In: *IEEE Transactions on Information Theory* 42.6 (1996), pp. 2254–2255. ISSN: 0018-9448. DOI: 10.1109/18.556617.

[JV02] Daemen Joan and Rijmen Vincent. *The Design of Rijndael AES-The Advanced Encryption Standard*. Springer-Verlag. Springer, Berlin, Heidelberg, 2002. ISBN: 9783662047224. DOI: 10.1007/978-3-662-04722-4.

[KAI10] Marton Kinga, Suciu Alin, and Iosif Ignat. "Randomness in Digital Cryptography: A Survey". In: *Romanian Journal Of Information Science and Technology* 13.3 (2010), pp. 219–240. URL: https://www.romjist.ro/content/pdf/kmarton.pdf.

[Kin+12] Marton Kinga, Suciu Alin, Sacarea Christian, and Cret Octavian. "Generation and Testing of Random Numbers for Cryptographic Applications". In: *Proceedings of Romanian Academy, Series A* 13.4 (2012), pp. 368–377.

[KS01] Jeong-Heon Kim and Hong-Yeop Song. "Trace Representation of Legendre Sequences". In: *Designs, Codes and Cryptography* 24.3 (2001), pp. 343–348. ISSN: 1573-7586. DOI: 10.1023/A:1011287607979.

[Lar67] Karl Larew. *The Codebreakers: The Story of Secret Writing*. 1st. Macmillan Publishing Co., Inc., New York, 1967. ISBN: 9780684831305.

[LG12] Chen Lidong and Gong Guang. *Communication System Security*. 1st. Chapman and Hall/CRC, New York, 2012. ISBN: 9780429094293. DOI: 10.1201/b12078.

[LN86] Rudolf Lidl and Harald Niederreiter. *Introduction to Finite Fields and Their Applications*. New York, NY, USA: Cambridge University Press, 1986. ISBN: 0521307066.

[LN96] Rudolf Lidl and Harald Niederreiter. *Finite Fields*. 2nd ed. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1996. DOI: 10.1017/CBO9780511525926.

[MN98] Makoto Matsumoto and Takuji Nishimura. "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator". In: *ACM Transactions on Modeling and Computer Simulation* 8.1 (1998), pp. 3–30. ISSN: 1049-3301. DOI: 10.1145/272991.272995.

[Mol02] Richard A. Mollin. *RSA and Public-Key Cryptography*. Taylor & Francis Group. Chapman and Hall CRC, 2002. ISBN: 9780429139659. DOI: 10.1201/9781420035247.

[Mor61] Boris Zelikovich Moroz. "The distribution of power residues and nonresidues (In Russian with English summary)". In: *Vestnik Leningrad Univ. Math* 16 (1961), pp. 164–169.

[MS97] Christian Mauduit and András Sárközy. "On finite pseudorandom binary sequences I: Measure of pseudorandomness, the Legendre symbol". In: *Acta Arithmetica* 82.4 (1997), pp. 365–377. DOI: 10.4064/aa-82-4-365-377.

[MVO96] Alfred J. Menezes, Scott A. Vanstone, and Paul C. Van Oorschot. *Handbook of Applied Cryptography*. 1st. Boca Raton, FL, USA:

CRC Press, Inc., 1996. ISBN: 0849385237. DOI: 10.1201/9780429/466335.

[Nea59] Zierler Neal. "Linear Recurring Sequences". In: *Journal of the Society for Industrial and Applied Mathematics* 7.1 (1959), pp. 31–48. DOI: 10.1137/0107003.

[Nog+16] Yasuyuki Nogami, Satoshi Uehara, Kazuyoshi Tsuchiya, Nasima Begum, Hiroto Ino, and Robert H. Morelos-Zaragoza. "A Multi-Value Sequence Generated by Power Residue Symbol and Trace Function over Odd Characteristic Field". In: *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* E99.A.12 (2016), pp. 2226–2237. DOI: 10.1587/transfun.E99.A.2226.

[NTU14] Yasuyuki Nogami, Kazuki Tada, and Satoshi Uehara. "A Geometric Sequence Binarized with Legendre Symbol over Odd Characteristic Field and Its Properties". In: *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* E97.A.12 (2014), pp. 2336–2342. DOI: 10.1587/transfun.E97.A.2336.

[Oma+18] Salhab Omar, Jweihan Nour, Jodeh Mohammed Abu, Taha Mohammed Abu, and Farajallah Mousa. "Survey paper: Pseudo Random Number Generators and Security Tests". In: *Journal of Theoretical and Applied Information Technology* 96.7 (2018), pp. 1951–1970. ISSN: 1992-8645.

[PVO91] Kumar P. Vijay and Moreno Oscar. "Prime-phase sequences with periodic correlation properties better than binary sequences". In: *IEEE Transactions on Information Theory* 37.3 (1991), pp. 603–616. ISSN: 0018-9448. DOI: 10.1109/18.79916.

[Raf17] Hamza Rafik. "A novel pseudo random sequence generator for image-cryptographic applications". In: *Journal of Information Security and Applications* 35 (2017), pp. 119–127. ISSN: 2214-2126. DOI: 10.1016/j.jisa.2017.06.005.

[Rai86] A. Rueppel Rainer. *Analysis and Design of Stream Ciphers*. Springer, Berlin, Heidelberg, 1986. ISBN: 9783642828652. DOI: 10.1007/978-3-642-82865-2.

[Ren04] Jian Ren. "Design of Long Period Pseudo-Random Sequence from the Addition of m-sequences over $\mathbb{F}_p$". In: *EURASIP Journal on Wireless Communication and Networking* 2004.1 (2004), pp. 802–851. DOI: 10.1155/s1687147204405052.

[Ruk+00] Andrew Rukhin, Juan Sota, James Nechvatal, Miles Smid, Elaine Barker, Stefan Leigh, Mark Levenson, Mark Vangel, David Banks, Alan Heckert, James Dray, and San Vo. *A statistical test suite for random and pseudorandom number generators for cryptographic*

applications. Tech. rep. 2000. DOI: 10.6028/nist.sp.800-22. URL: https://doi.org/10.6028%2Fnist.sp.800-22.

[Sch15] Bruce Schneier. *Applied Cryptography: Protocols, Algorithms, and Source Code in C*. New York, NY, USA: Wiley; 20th Anniversary edition, 2015. ISBN: 1119096723.

[Sid71] V. M. Sidelnikov. "The mutual correlation of sequences". In: *Dokl. Akad. Nauk SSSR* 196 (1971), pp. 531–534. ISSN: 0002-3264.

[TCA98] Cusick Thomas, Ding Cunsheng, and Renvall Ari. *Stream Ciphers and Number Theory*. North-Holland Mathematical Library. Elsevier Science, 1998. ISBN: 9780080541846. DOI: 10.1016/S0924-6509(98)80054-8.

[XG10] Tang Xiaohu and Gong Guang. "New Constructions of Binary Sequences With Optimal Autocorrelation Value/Magnitude". In: *IEEE Transactions on Information Theory* 56.3 (2010), pp. 1278–1286. ISSN: 0018-9448. DOI: 10.1109/TIT.2009.2039159.

[XQ06] Hong Xu and Wen-Feng Qi. "Autocorrelations of Maximum Period FCSR Sequences". In: *SIAM J. Discret. Math.* 20.3 (2006), pp. 568–577. ISSN: 0895-4801. DOI: 10.1137/050633974.

[YG06] Nam Yul Yu and Guang Gong. "Crosscorrelation Properties of Binary Sequences with Ideal Two-Level Autocorrelation". In: *Sequences and Their Applications – SETA 2006*. Ed. by Guang Gong, Tor Helleseth, Hong-Yeop Song, and Kyeongcheol Yang. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 104–118. ISBN: 9783540445241. DOI: 10.1007/11863854_9.

[ZBT06] Gutterman Zvi, Pinkas Benny, and Reinman Tzachy. "Analysis of the Linux random number generator". In: *2006 IEEE Symposium on Security and Privacy (S P'06)*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, 15 pp.–385. DOI: 10.1109/SP.2006.5.

[Zie58] Neal Zierler. *Legendre Sequences*. Holden-Day series in information systems. MIT Lincoln Laboratory Publications Office, 1958. URL: https://apps.dtic.mil/dtic/tr/fulltext/u2/268167.pdf.

Index