Breaking NLM-MAC Generator

Mohammad Ali Orumiehchiha[1], Josef Pieprzyk[1], and Ron Steinfeld[2]

[1]Center for Advanced Computing – Algorithms and Cryptography
Department of Computing,
Faculty of Science, Macquarie University, Sydney, NSW 2109, Australia
(mohammad.orumiehchiha,josef.pieprzyk)@mq.edu.au
[2]Clayton School of Information Technology
Monash University, Clayton VIC 3800, Australia
ron.steinfeld@monash.edu

Abstract. NLM generator, designed by HoonJae Lee, SangMin Sung, HyeongRag Kim, is the strengthened version of the LM-type summation generator with two memory bits; which uses non-linear combination of linear feedback shift register and non-linear feedback shift register. Recently, the cipher along with a massage authenticate function have been proposed for a lightweight communication framework in wireless sensor networks. Also, the generator has been used in two different RFID mutual authentication protocols and a protocol to secure access in internet. This paper indicates some critical cryptographic weak points leading to the key recovery and forgery attack. We prove the internal state of NLM-n can be recovered with time complexity about $n^{log7\times 2}$ where the total length of internal state is $2 \cdot n + 2$ bits. The attack needs about n^2 key-stream bits. We also show attacker is able forge any MAC tag in real time by having only one pair (MAC tag, cipher-text). The proposed attacks are completely practical and break the scheme with negligible error probability.

Keywords: NLM Stream Cipher, MAC Function, Cryptanalysis, Key Recovery Attack, Forgery Attack.

1 Introduction

Stream ciphers are symmetric encryption algorithms, which play effective role to provide confidentiality. They accept two inputs: a secret key K and an initial value IV and use them to generate key stream bits. The key stream then can be applied for encryption and decryption. One of such stream ciphers is a summation generator designed by Rainer Rueppel in 1985. It produces key-stream bits by adding output bits of two linear feedback shift register (LFSR) and one bit carry bit of an adder. The cipher exhibits many desirable cryptographic propertied such as maximum period, near-maximum linear complexity and maximum order of

correlation immunity. But, the cipher is insecure against correlation and algebraic attacks.

In 2000, Hoon Jae Leea and Sang Jae Moon proposed an improved summation generator with 2-bit memory (LM-type Generator) [7]. The design was intended to enhance security properties by adding an extra bit of memory to the combining function. However, there were still some cryptographic weaknesses of the cipher. Due to a high correlation between the input variables and the output sequences of the combining function, the authors of [2, 11] showed that the cipher is vulnerable to correlation attacks. Also, an efficient attack recovering the internal state of the cipher in real time was published in [4].

The NLM stream cipher [6] is actually a modification of the LM-type generator proposed by Hoon Jae Lee, Sang Min Sung, and Hyeong Rag Kim in 2009. The main idea of NLM generator is to add a non-linear feedback shift register to the summation generator that strengthens the cipher. In addition, the authors of [8] have checked performance of NLM stream cipher for low power consumption applications and have confirmed that the cipher is suitable for implementations requiring a small number of gates.

1.1 Related Works

Message authentication codes (MACs) are cryptographic tools that provide integrity and authentication of messages. A typical MAC can be designed using a symmetric cipher. There are many constructions of stream ciphers with a built-in MAC functionality (see [1, 3, 12, 13, 5] as examples). Recently, Lee et al. in [5] have proposed a lightweight secure data communication framework based on the NLM stream cipher and a new MAC function combined with the cipher to enhance security in wireless sensor networks.

The NLM stream cipher has also been employed in two RFID mutual authentication protocols [10, 9] to encrypt sensitive data. In addition, Lee, Kim and Lee in [9] propose an internet protocol to establish secure access for mobile users based on functionality of the NLM generator. The cryptographic analysis presented in this paper shows weaknesses of the NLM generator and discusses their impact on the security of the protocols it supports.

The rest of the paper is structured as follows. Section 2 describes briefly the NLM stream cipher and the NLM-MAC function. Section 3 investigates the weaknesses of the cipher and proposes state recovery attacks on the NLM generator and a forgery attack on the MAC function.

Also, the section discusses the weaknesses of the whole scheme. Section 4 concludes the work.

2 Description of NLM-MAC Scheme

In this section, we first describe the NLM-128 generator. Then, we explain how the NLM-MAC algorithm works.

2.1 NLM-128 Stream Cipher

The NLM-128 stream cipher is based on summation generation which uses LFSR and NLFSR sequences and two memory bits; a carry bit (c_i), and a memory bit (d_i). Figure 1 depicts the cipher. The LFSR has primitive polynomial $P(x)$ as follows:

$P(x) = x_{127} \oplus x_{109} \oplus x_{91} \oplus x_{84} \oplus x_{73} \oplus x_{67} \oplus x_{66} \oplus x_{63} \oplus x_{56} \oplus x_{55} \oplus x_{48} \oplus x_{45} \oplus x_{42} \oplus x_{41} \oplus x_{37} \oplus x_{34} \oplus x_{30} \oplus x_{27} \oplus x_{23} \oplus x_{21} \oplus x_{20} \oplus x_{19} \oplus x_{16} \oplus x_{13} \oplus x_{12} \oplus x_7 \oplus x_6 \oplus x_2 \oplus 1$

NLFSR uses a non-linear feedback function $f(x)$ of degree 129 defined as

$$\begin{aligned} f(x) = {} & x_5 \oplus x_9 \oplus x_{13} \oplus x_{17} \oplus x_{21} \oplus \\ & x_{25} \oplus x_{29} \oplus x_{33} \oplus x_{37} \oplus x_{41} \oplus \\ & x_{45} \oplus x_{49} \oplus x_{53} \oplus x_{57} \oplus x_{61} \oplus \\ & x_{65} \oplus x_{69} \oplus x_{73} \oplus x_{77} \oplus x_{81} \oplus \\ & x_{85} \oplus x_{89} \oplus x_{93} \oplus x_{97} \oplus x_{101} \oplus \\ & x_{105} \oplus x_{109} \oplus x_{113} \oplus x_{117} \oplus x_{121} \\ & \oplus x_{125} \oplus x_{129} \oplus (x_1 \cdot x_2 \cdots x_{128} \cdot x_{129}). \end{aligned} \tag{1}$$

The carry bit c_j, and the additional memory bit d_j are updated according to the following relations:

$$c_j = a_j \cdot b_j \oplus (a_j \oplus b_j) \cdot c_{j-1} \tag{2}$$

$$d_j = b_j \oplus (a_j \oplus b_j) \cdot d_{j-1} \tag{3}$$

Finally, the key-stream bit z_j is generated as shown below

$$z_j = a_j \oplus b_j \oplus c_{j-1} \oplus d_{j-1} \tag{4}$$

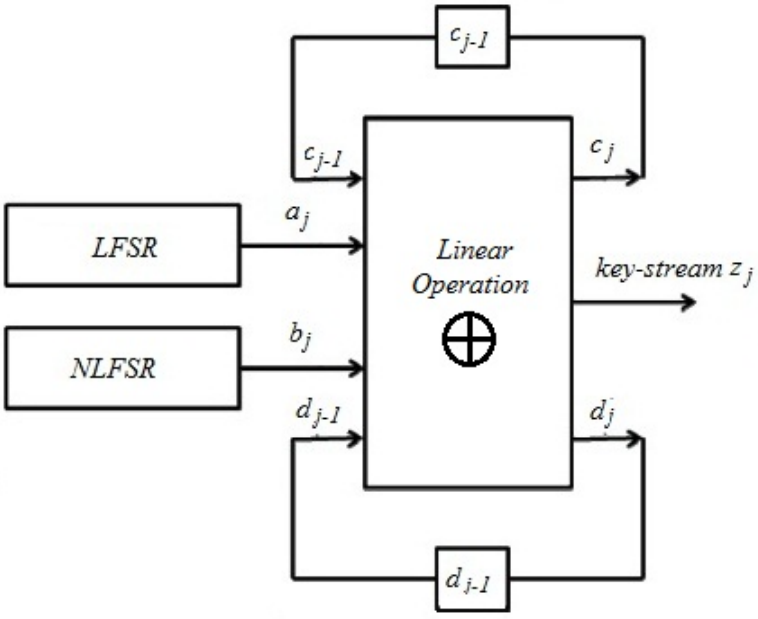

Fig. 1. NLM Family Stream Cipher

2.2 The NLM-MAC Function

The NLM message authentication code authenticates the two parties (a sender and a receiver) and verifies the integrity of transmitted messages as follows:

1. Sender encrypts a plain-text with an encryption key and an initialization vector and generates a corresponding ciphertext (CT) using the NLM-128 stream cipher.
2. Sender computes a MAC value for the ciphertext with MAC-Key (i.e., K_{mac}) according to the following steps:
   - 2.1 CT is split into 32-bit words and then the last word is padded with zeros if required.
   - 2.2 K_{mac} is fed through variables l, m, n, p and then K_{mac} is XORed with 32-bit CT words and with 32-bit of l.
   - 2.3 After Xoring all 32-bit CT words with l, the NLM-MAC will be generated as follows:

$$\text{NLM-MAC} = l \oplus m \oplus n \oplus p$$

3. The receiver checks the validity of the MAC tag and then decrypts authenticated ciphetext to obtain plaintext.

Note: The protocol uses a time stamp to check the freshness of the messages. The time stamp has no impact on our proposed attack.

3 Cryptanalysis of NLM-MAC Scheme

In this section, we reveal weak points of the NLM algorithm and demonstrate the attack's details. Also, we prove that an attacker not only can break the NLM stream cipher but he is also able to generate valid MAC tags for fake messages in real time.

3.1 Cryptanalysis of NLM generator

The NLM generator is strengthen version of LM generator family. The generator aims to prevent the attacks published in [4, 2, 11] by using NLFSR to make the design resistant against correlation and algebraic attacks. First, we identify weaknesses of the cipher.

1. The algebraic degree of key-stream output bits when the cipher uses two LFSRs is 2. Han and Lee show in [4] that the algebraic degree of LM-type generators can be kept constant (with degree 2). To show this we take Equations (2) and (3) and get

$$c_j \oplus d_j = a_j b_j \oplus b_j \oplus (a_j \oplus b_j)(c_{j-1} \oplus d_{j-1}).$$

   If we put $c_{j-1} \oplus d_{j-1} = z_j \oplus (a_j \oplus b_j)$ to Equation (4), we obtain the following equation

$$c_j \oplus d_j = a_j b_j \oplus b_j \oplus (a_j \oplus b_j)(z_j \oplus (a_j \oplus b_j)) \tag{5}$$

   Substituting $(j + 1)$ for j in Equation (4) and using Equation (5), we finally have

$$z_{j+1} = a_j + 1 \oplus b_j + 1 \oplus a_j \oplus a_j b_j \oplus (a_j \oplus b_j) z_j. \tag{6}$$

   Equation (6) creates equations of degree 2 connecting 2 output bits and the register outputs.
2. The NLM designers have believed that replacing LFSR by a NLFSR strengthens the design and makes it resistant against algebraic analysis. To keep the desirable properties of LM cipher, they have used a NLFSR that has the full period with feedback function (relation 1). Although the algebraic degree of feedback function (1) is high and equal to 129, the non-linearity is surprisingly low. The attacker can

approximate the non-linear feedback function with a linear function with the following probability:

$$Pr(f(x) = L(x)) = 1 - 2^{-129}$$

where $L(x) = (x_5 \oplus x_9 \oplus x_{13} \oplus x_{17} \oplus x_{21} \oplus x_{25} \oplus x_{29} \oplus x_{33} \oplus x_{37} \oplus x_{41} \oplus x_{45} \oplus x_{49} \oplus x_{53} \oplus x_{57} \oplus x_{61} \oplus x_{65} \oplus x_{69} \oplus x_{73} \oplus x_{77} \oplus x_{81} \oplus x_{85} \oplus x_{89} \oplus x_{93} \oplus x_{97} \oplus x_{101} \oplus x_{105} \oplus x_{109} \oplus x_{113} \oplus x_{117} \oplus x_{121} \oplus x_{125} \oplus x_{129}$.

The second weak point lets attacker replace the NLFSR with the LFSR defined by the feedback function $L(x)$. The cipher can be broken in the two following steps.

1. The attacker constructs the non-linear algebraic system based on equations of the form (6). The number of variables equals total length of the shift registers and two memory bits (*e.g.* $n = 258$)). Han and Lee [4] prove that the time complexity of solving the system is $O(n^{5.6})$ and the attacks needs about n^2 bits.
2. Next the attacker checks the validity of the recovered internal state. To this end, attacker needs to generate additional output bits by using the recovered internal state. The probability of recovering incorrect internal state equals to $2^{-129} \times n^2 = 2^{-129} \times (258)^2 = 2^{-111}$, which is still a negligible probability. In addition, one can repeat the attack on the next n^2 bits of key-stream and find the internal state to verify the previous result.

3.2 Analysis on NLM-MAC Function

The most critical point is that the NLM-MAC function is totally linear. It means all relations between the MAC secret key K_{mac} and cipher-text are constructed linearly. So, one can compute the linear relation of K_{mac} words by having only one MAC tag and its corresponding cipher-text. This leakage reveals the linear relation of l, m, n, p which are enough to compute valid MAC value for every arbitrary cipher-text.

3.3 Attack on NLM Scheme

Now, we show how we can launch a key recovery attacks on the NLM cipher and forge MAC value. What the attacker needs is about 2^{16} bits of key-stream and a MAC tag and its corresponding cipher-text. The attack can work as follows:

1. For a ciphertext of length n^2 bits, where n is the number of internal bits, the attacker finds internal states of the cipher with negligible error probability.
2. For the pair (ciphertext, MAC tag), the attacker applies the explained attack in Section 3.2.
3. The attacker can send an arbitrary ciphertext along with a valid MAC tag, or by adding new plaintext bits following the original plaintext, he can compute cuphertext and update new MAC value. Another approach is to replace the original plaintext with an arbitrary text and compute corresponding cipher-text and MAC tag.

4 Conclusions

In this paper, we analysed the NLM-MAC scheme proposed for lightweight applications such as wireless sensor networks. We discovered some weaknesses leading to two successful cryptographic attacks. The first attack allows to recover an internal state with time complexity about $2^{44.86}$ and the required output bits about 2^{16}. The second attack permits to forge a MAC tag for every ciphertext in real time. Finally, we proposed an attack on the protocol, which lets attacker generate arbitrary ciphertexts along with a valid MAC tag. As a conclusion, we can say that the proposed scheme is totally insecure and it is not recommended to be used.

References

1. A. Braeken, J. Lano, N. Mentens, B. Preneel, and I. Verbauwhede, *Sfinks: A synchronous stream cipher for restricted hardware environments*, in SKEW - Symmetric Key Encryption Workshop, 2005.
2. C.-K. Chan and L. M. Cheng, *Correlation properties of an improved summation generator with 2-bit memory*, Signal Process., 82 (2002), pp. 907–909.
3. N. Ferguson, D. Whiting, B. Schneier, J. Kelsey, S. Lucks, and T. Kohno, *Helix - fast encryption and authentication in a single cryptographic primitive*, in Proc. Fast Software Encryption 2003, volume 2887 of LNCS, Springer-Verlag, 2003, pp. 330–346.
4. D. Han and M. Lee, *An algebraic attack on the improved summation generator with 2-bit memory*, Inf. Process. Lett., 93 (2005), pp. 43–46.
5. P. Kumar and H.-J. Lee, *Nlm-mac: Lightweight secure data communication framework using authenticated encryption in wireless sensor networks, applied cryptography and network security*, applied cryptography and network security, (2012), pp. 153–168.
6. H. Lee, S. Sung, and H. Kim, *Nlm-128, an improved lm-type summation generator with 2-bit memories*, in Proceedings of the 2009 Fourth International Conference on Computer Sciences and Convergence Information Technology, ICCIT '09, Washington, DC, USA, 2009, IEEE Computer Society, pp. 577–582.

7. H. J. Lee and S.-J. Moon, *On an improved summation generator with 2-bit memory*, Signal Processing, 80 (2000), pp. 211–217.
8. S. Y. Lee and H. Lee, *Hardware implementation and performance analysis of nlm-128 stream cipher*, in 6th International Conference Convergence and Hybrid Information Technology, ICHIT (2), vol. 206 of Communications in Computer and Information Science, Springer, 2011, pp. 446–453.
9. Y. S. Lee, T. Y. Kim, and H.-J. Lee, *Mutual authentication protocol for enhanced rfid security and anti-counterfeiting*, in Advanced Information Networking and Applications Workshops (WAINA), 2012 26th International Conference on, 2012, pp. 558–563.
10. Y. S. Lee, Y. Park, S. Lee, T. Kim, and H.-J. Lee, *Rfid mutual authentication protocol with unclonable rfid-tags*, in Mobile IT Convergence (ICMIC), 2011 International Conference on, 2011, pp. 74–77.
11. J. C. Mex-Perera and S. J. Shepherd, *Cryptanalysis of a summation generator with 2-bit memory*, Signal Process., 82 (2002), pp. 2025–2028.
12. D. Whiting, B. Schneier, S. Lucks, and F. Muller, *Phelix: Fast encryption and authentication in a single cryptographic primitive*, in eSTREAM, ECRYPT Stream Cipher Project Report 2005/027, 2005.
13. B. Zoltak, *Vmpc one-way function and stream cipher*, in FSE, B. K. Roy and W. Meier, eds., vol. 3017 of Lecture Notes in Computer Science, Springer, 2004, pp. 210–225.