Practical Attack on NLM-MAC Scheme

Mohammad Ali Orumiehchiha[a], Josef Pieprzyk[b], Ron Steinfeld[c]

[a] *Center for Advanced Computing – Algorithms and Cryptography, Department of Computing, Macquarie University, Sydney, NSW 2109, Australia*
[b] *Center for Information Security, School of Electrical Engineering and Computer Science, Queensland University of Technology, Brisbane, QLD 4000, Australia*
[c] *Clayton School of Information Technology, Monash University, Clayton VIC 3800, Australia*

Abstract

The NLM stream cipher designed by HoonJae Lee, SangMin Sung, HyeongRag Kim is a strengthened version of the LM summation generator that combines linear and non-linear feedback shift registers. In recent works, the NLM cipher has been used for message authentication in lightweight communication over wireless sensor networks and for RFID authentication protocols.

The work analyses the security of the NLM stream cipher and the NLM-MAC scheme that is built on the top of the NLM cipher. We first show that the NLM cipher suffers from two major weaknesses that lead to key recovery and forgery attacks. We prove the internal state of the NLM cipher can be recovered with time complexity about $n^{log7\times2}$, 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 adversary is able forge any MAC tag very efficiently by having only one pair (MAC tag, cipher-text). The proposed attacks are practical and break the scheme with a negligible error probability.

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

1. Introduction

The summation generator (10) designed by Rainer Rueppel produces keystream bits by adding output bits of two linear feedback shift registers (LFSRs) and the carry bit of an adder. The cipher exhibits many desirable cryptographic propertied. It produces a key stream of maximum period, achieves near maximum linear complexity and maximum order of correlation immunity. Unfortunately, it is insecure against the correlation and algebraic attacks. In 2000, Hoon Jae Leea and Sang Jae Moon proposed an improved summation generator with 2-bit memory (5). We call it the LM generator. The analysis published in (1; 9) shows that the cipher is vulnerable to correlation attacks. This weakness is the result of a high correlation between the input variables and the output sequences of the combining function. Also, an efficient attack that recovers the internal state of the cipher in real time has been published in (2).

In 2009 Hoon Jae Lee, Sang Min Sung, and Hyeong Rag Kim published another version of the LM generator (see (4)). They called it the NLM stream cipher. The main feature of the cipher that was supposed to strengthen the cipher is the addition of a non-linear feedback shift register. Using the NLM stream cipher, Lee et al. have recently proposed a lightweight secure data communication framework - see (3) for details. A part of the framework is a new MAC function that is intended to enhance security of wireless sensor networks. The NLM stream cipher is suitable for implementation requiring a small number of gates as confirmed by the work of Lee and Lee (6). Because of the attractive features, the NLM stream cipher

Email addresses: mohammad.orumiehchiha@mq.edu.au (Mohammad Ali Orumiehchiha), josef.pieprzyk@qut.edu.au (Josef Pieprzyk), ron.steinfeld@monash.edu (Ron Steinfeld)

has been deployed in two RFID authentication protocols (8; 7). Also Lee, Kim and Lee in (7) use the NLM cipher in their Internet protocol that establishes a secure access for mobile users.

The analysis presented in this paper shows weaknesses of the NLM cipher and discusses their impact on the security of the protocols it supports. We conclude that the NLM-MAC scheme is completely insecure and the NLM cipher is not recommended.

2. Description of NLM-MAC Scheme

In this section, we first describe the NLM-128 cipher. Then, we explain how the NLM-MAC function works. For details, the reader is referred to the original paper (4).

2.1. NLM-128 Stream Cipher

The cipher is based on the summation generation, which uses the LFSR and NLFSR sequences and the two bits, namely a carry bit (c_i) and a memory bit (d_i). Figure 1 shows the overall structure of the cipher. The LFSR is defined by its primitive polynomial $P(x)$ of the following form:

$$\begin{aligned} P(x) &= x^{127} + x^{109} + x^{91} + x^{84} + x^{73} + x^{67} + x^{66} + x^{63} \\ &+ x^{56} + x^{55} + x^{48} + x^{45} + x^{42} + x^{41} + x^{37} + x^{34} \\ &+ x^{30} + x^{27} + x^{23} + x^{21} + x^{20} + x^{19} + x^{16} + x^{13} \\ &+ x^{12} + x^{7} + x^{6} + x^{2} + 1 \end{aligned}$$

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

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

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

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

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

Finally, the keystream bit z_j is generated as shown below:

$$z_j = a_j \oplus b_j \oplus c_{j-1} \oplus d_{j-1} \quad (4)$$

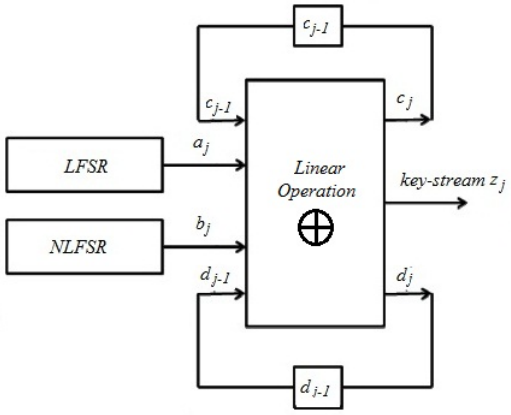

Figure 1: The NLM Cipher

2.2. NLM-MAC Function

Given a ciphertext (CT) and a MAC key (denoted by K_{mac}), the MAC value is computed according to the following steps:

1. CT is split into 32-bit words and then the last word is padded with zeros if required.

2. K_{mac} is fed through four 32-bit variables l, m, n, p, where they are 32-bit words initialised by the MAC key, where the sequence $(\{l, m, n, p\} = K_{mac}(128 - bit))$. Then K_{mac} is xor-ed with 32-bit CT words and with 32-bit of l.

3. After xor-ing all 32-bit CT words with l, the NLM-MAC value will be generated as follows:

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

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

3. Analysis of NLM-MAC Scheme

In this section, we reveal weak points of the NLM algorithm and describe details of our attack. We prove not only that it is possible to break the NLM cipher but also an adversary is able to create valid MAC tags for fake messages very efficiently.

3.1. Analysis of NLM Cipher

First, we identify weaknesses of the cipher.

1. In their work (2), Han and Lee show that the algebraic degree of LM generators can be kept constant and equal to 2. This fact can be established as follows. We start from two Equations (2) and (3) and add them. This gives us the following relation

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

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

$$c_j \oplus d_j = a_j b_j \oplus b_j \oplus (a_j \oplus b_j)(z_j \oplus (a_j \oplus b_j)) \quad (5)$$

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

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

Equation (6) creates equations of degree 2 connecting 2 output bits (that can be observed by the adversary) and the register outputs.

2. The NLM designers believe that replacing LFSR by NLFSR strengthens the design and makes it resistant against algebraic analysis. To keep the desirable properties of LM cipher, they have used NLFSR that has the maximum period (NLFSR is characterised by its feedback function given by Equation 1). Although the algebraic degree of feedback function (1) is high and equal to 129, the non-linearity is surprisingly low. The adversary can approximate the non-linear feedback function with a linear function with the following probability:

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

where

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

The second weak point lets the adversary replace NLFSR with LFSR defined by the feedback function $L(x)$.

3.2. Attack stages

The cipher can be broken in the two following steps.

1. The adversary constructs the non-linear algebraic system using the observations of the output bits and creating system of relations derived from Equation (6). The number of variables equals to the total length of the shift registers and two memory bits (*e.g.* $n = 258$). In the work (2), it is shown that the time complexity of solving the system is $O(n^{5.6})$ and the attacks needs about n^2 bits.
2. Next the adversary checks the validity of the recovered internal state. To this end, adversary needs to generate additional output bits by using the recovered internal state. The probability of recovering incorrect internal state equals to $2^{-129} \times n^2 = 2^{-129} \times (258)^2 \approx 2^{-111}$, 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.

3.3. Analysis on NLM-MAC Function

The most critical point in the NLM-MAC function is that the function is totally linear. It means that all relations between the MAC secret key K_{mac} and the ciphertext 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 ciphertext. Suppose the adversary possesses a valid MAC tag and its corresponding ciphertext (CT). Then he splits the ciphertext into 32-bit words and makes one 32-bit by xoring all the ciphertext words with the Mac tag. The new computed words are

actually the initial value for $l \oplus m \oplus n \oplus p$. After this stage, the adversary is able to compute another valid MAC value for every ciphertext.

3.4. Attack on NLM Scheme

Now, we show how we can launch a key recovery attacks on the NLM cipher and forge MAC values. What the adversary needs is about 2^{16} bits of keystream and a MAC tag and its corresponding ciphertext. The attack proceeds as follows.

1. For a ciphertext of length n^2 bits, where n is the number of internal bits, the adversary finds the internal state of the cipher with a negligible error probability.
2. For the pair (ciphertext, MAC tag), the adversary applies the attack from Section 3.3.
3. The adversary can send an arbitrary ciphertext along with a valid MAC tag or by adding new plaintext bits following the original plaintext, he can compute ciphertext and update the new MAC value. Another approach is to replace the original plaintext with an arbitrary text and compute corresponding ciphertext 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 the 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 adversary generate arbitrary ciphertexts along with a valid MAC tag. In conclusion, we can say that the proposed scheme is totally insecure and it is not recommended to be used.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] H. J. Lee and S.-J. Moon, *On an improved summation generator with 2-bit memory*, Signal Processing, 80 (2000), pp. 211–217.

[6] 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.

[7] 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.

[8] 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.

[9] J. C. Mex-Perera and S. J. Shepherd, *Cryptanalysis of a summation generator with 2-bit memory*, Signal Process., 82 (2002), pp. 2025–2028.

[10] R. A. Rueppel, *Correlation immunity and the summation generator*, in Advances in Cryptology,

CRYPTO '85, London, UK, 1986, Springer-Verlag,
pp. 260–272.