Enhancement of A5/1 Stream Cipher Overcoming its Weakneses

Mahdi Madani, Salim Chitroub
Signal and Image Processing Laboratory
Electronics and Computer Science Faculty, USTHB
Algiers, Algeria
e-mail:{mmadani49@gmail.com, s_chitroub@hotmail.com}

***Abstract*—The Global system for Mobile (GSM) communication is still the most widely used cellular system in the world, with over the billion customers around the world; even the fourth and fifth generations are now operated in some countries. However, GSM bears numerous security vulnerabilities and for that reason it has seriously considered security threats. Although GSMs architecture is designed in such a way to provide various security features like authentication, data and signaling confidentiality, and the user secrecy, the GSM channel is yet susceptible to replaying, interleaving and man-in-the-middle attacks. The GSM voice calls are encrypting a family of algorithms collectively called A5. A5/1 is the stream cipher which encrypts the information transmitted from a mobile user. Initially, A5 algorithm was kept secret to ensure the security, but us algorithm was disclosed many cryptanalytic attacks that have made in evidence the weakness of A5 algorithm. In this paper, after the discussion on the A5/1 encryption algorithm, and the attacks that it has suffered, we will put in evidence the weakness of A5/1. Although attacks that can be used to break the A5/1 algorithm require incredible computing power so that not only certain people with certain computer could break the A5/1 algorithm, however, this does not preclude the necessity of improve safety A5/1. We propose here to enhance the A5/1 stream cipher by overcoming its weakness.**

Keywords-GSM Networks; Stream Cipher; Mobile Security; Encryption and Cryptography; Cryptanalysis; Non-Linear Boolean Functions; Linear Feedback Shift Gegister.

I. INTRODUCTION

Mobile communications have become now more popular and easier. Nowadays, people can communicate with each other on any place at any time. However, the openness of wireless communications poses serious security threats of communicating parties. How to provide secure communication channels is essential to the success of a mobile communication network. Encryption in wireless communication is essential to protect sensitive information, and to prevent fraud. Stream ciphers are symmetric-key ciphers that generate pseudo-random binary sequences, which are used to encrypt the message signals and data wireless communications. Stream ciphers can be designed to be exceptionally fast, much faster than any block cipher. While block ciphers operate on large blocks of data, stream ciphers typically operate on smaller units of plaintext, usually bits. A stream cipher generates a key-stream and encryption is provided by combing the key-stream with the plaintext, usually with the bitwise XOR operation.

GSM is a widely used mobile standard in the world that uses A5/1 stream cipher for protecting the privacy and secrecy of the subscriber's information over the air interface. However, recent research and studies show that it has some limitations owing to which it is cryptanalyzed by a number of cryptographic attacks. One of the weaknesses of A5/ is fixed feedback polynomial of Linear Feedback Shift Registers (LFSRs); and other is the weak clocking mechanism. A5/1 was initially cryptanalyzed by Golic [1], when only a rough outline of A5/1 was leaked. Then, A5/1 algorithm was cryptanalyzed by Biryukov, Shamir, and Wagner[2], Bihan and Dunkelman [3], Ekdahl and Johansson [4], Maximov et al. [5], and recently, by Barkan and Bihan [6]. Most of the attacks against A5/1 are known plaintext attacks and use security weakness in the clock-controlling unit [7].

The remainder of the paper is organized as follow. The following section summarizes the design of the A5/1 stream cipher. The third section describes how the algorithm works. The fourth section reviews some of the attacks that can be used against A5/1 algorithm. The fifth section shows the weakness of the A5/1. Our proposed method to enhance the security of the A5/1 stream cipher algorithm is exposed in the sixth section. The last section concludes this work.

II. DATA ENCRYPTION IN GSM

The GSM standard has its specific algorithms for data encryption and authentication that are grouped in the A5/1 algorithms family. Only the A5/0, which is no an encryption algorithm, A5/1 and A5/2 are the two encryption algorithms stipulated by this standard, where the stream cipher A5/1 is used within Europe and most other countries. A5/2 is the internationally weaker version of A5/1 which has been developed (due to export restrictions) for deploying GSM outside of Europe. A5/3 is a new algorithm based on the Universal Mobile Telecommunications System (UMTS)/ Wide Code Division Multiple Access (WCDMA) and Kasumi algorithm [8]. All of these algorithms use 64-bits key [9]. Through terminal of both ciphers were kept secret, their designs were disclosed in 1999 by means of reverse engineering [12]. In this work, we focus on the A5/1, its performances and its weakness.

 ISBN: 978-1-61208-347-6

A5/1 is a synchronous stream cipher that uses a 64 bits session key and an initial vector of 2 bits derived from the 22 bits frame number, which is publicly known. It uses three LFSRs called R1, R2, and R3 of the length 19, 22, and 23 bits respectively, as it is shown in Figure 1. The combined length of the three LFSRs is then 64 bits. The rightmost bit in each register is labeled as bit zero. Three primitive feedback polynomials are used for the three LFSRs R1, R2 , and R3, which are: $x19 + x5 + x2 +x +1$, $x22 + x + 1$, and $x23 + x15 + x2 + x + 1$, respectively. Each register contain a set of bits called taps. The taps of the LFSRs correspond to primitive polynomials and, therefore the registers give sequences of maximum periods. The taps of R1 are at bit positions 13, 16, 17, 18; the taps of R2 are at bit positions 20, 21; and the taps of R3 are at bit positions 7, 20, 21, 22. R1, R2, and R3 are clocked irregularly based on the values of the clocking bits that are at positions 8, 10, and 10 of registers R1, R2, and R3, respectively.

III. HOW A5/1 IS WORKING?

A. Clocking

The stream cipher produces the key-stream by generating one bit at each clock cycle until reached the size of such key-stream. The registers are clocked in a stop/go clock control using a majority rule, which uses three clocking bits C1, C2 and C3 of registers R1, R2 and R3, respectively, and determines the value of the majority bit m using the formula: m = maj(C1, C2, C3) among the clocking bits, if two or more are 0 then the value of majority bit m is 0. Similarly, if two or more bits are 1, then majority bit m is 1. After that, if Ci=m then bi=1, otherwise bi=0. For example, let (C1, C2, C3) = (1, 0, 1) then, according to majority rule, m=1. Now, if Ci = m, then register Ri will be clocked (shifted left by 1 bit), where i=1, 2, 3. The probability of an individual LFSR being clocked is 3/4. At each clocking cycle, each LFSR generates one bit xi which are then combined by a linear combining function z(t), to produce one bit of the output key-stream, defined as follows:

$$Z(t) = x1 \oplus x2 \oplus x3. \quad (1)$$

All the possible cases of clocking process are summarized in Table 1.

B. Process

To generate the 228 bits of key-stream, initially, the 64 bits of the session key Kc and the 22 bits of frame number are used, then the process below is follow-up [11]:

- The first step consists to zeroed the three registers, then for 64 cycles (ignoring the stop/go clock control), the 64 bits of the secret key Kc (from LSB to MSB) are mixed using the XOR operator, in parallel with the least significant bit of each registers.
- Similarly, the second step consists to clock the registers for 22 additional cycles (ignoring the stop/go clock control). Like the first step, during this period the 22 bits of Frame number (Fn) (from LSB to MSB) are again mixed in parallel with the least significant bit of each register. The contents of the three registers at the end of this step are called the initial states of the frame.
- In the third step, the three register are clocked for 100 additional cycles with the stop/go clock control, but the output is discarded for these cycles. After this is completed, the cipher is now ready to produce two 114 bit sequences at output, the first 114 bits are for downlink, and the last 114 bits are for uplink.
- The last step consists to clock the three registers for 228 additional cycles with the stop/go clock control in order to produce the 228 output bits. At each clock cycle, one bit is produced as the result of the XOR operation of the most significant bit of the three registers.

TABLE I. MAJORITY RULE RESULTS

b1	b2	b3	m	c1	c2	c3	Register(s) clocked
0	0	0	0	1	1	1	**R1, R2, R3**
0	0	1	0	1	1	0	**R1, R2**
0	1	0	0	1	0	1	**R1, R3**
0	1	1	1	0	1	1	**R2, R3**
1	0	0	0	0	1	1	**R2, R3**
1	0	1	1	1	0	1	**R1, R3**
1	1	0	1	1	1	0	**R1, R2**
1	1	1	1	1	1	1	**R1, R2, R3**

As we said above, the GSM conversations are send as sequence of the frames, and the objective is to encrypt each frame before its transmission. So, after generating this pseudo random of 228 bits, they will be mixed using the XOR operator with one frame of the plaintext and then sent the result (one frame of cipher-text) to the other part of the network. Then, the receiver can decrypt the messages using the same process.

IV. ATTACKS AGAINST A5/1

A certain number of serious weaknesses in its coding mechanism were identified. Thereafter, it was targeted by several attacks which we can classify in two categories. First, Brute force attacks are based on searching to decrypt the conversations by trying all the 2^{64} computations of Kc.

To decrypt one frame of 228 bits, the operation tacks about 3-6 years. Since the decryption is not in real-time, the conversation maybe meaningless after 3-6 years. Second, Look Up Table Attacks need in total of 2^{64} memories to store the look up table in total. That cost about 18446744 terabytes in total, and we also need a strong computing power to process such big data like that.

The security of the A5/1 encryption algorithm was analyzed in several papers [1][2][12][13]. The known attacks can be summarized as this:

- Briceno [12] found out that in all the deployed versions of the A5/1 algorithm, the ten least significant of the 64 bits of the keys were always set

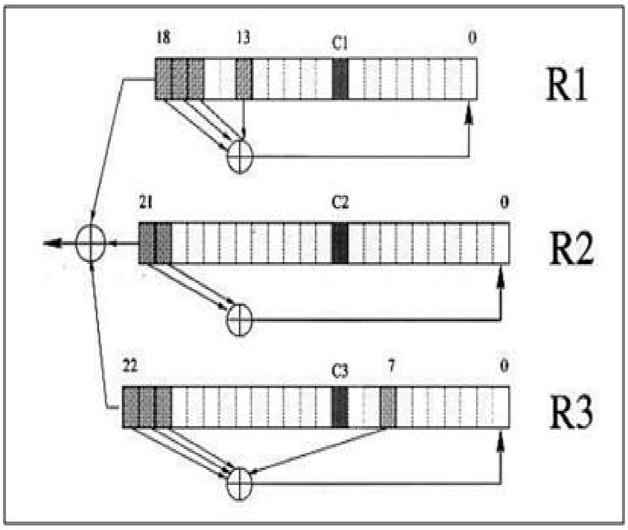

Figure 1. A5/1 stream cipher

to zero. The complexity of exhaustive search is thus reduced to $O(2^{54})$.

- Briceno [12] found out that in all the deployed versions of the A5/1 algorithm, the ten least significant of the 64 bits of the keys were always set to zero. The complexity of exhaustive search is thus reduced to $O(2^{54})$.
- Anderson and Roe [13] proposed an attack based on guessing the 41 bits in the shorter registers, R1 and R2, and deriving the 23 bits from the longer register, R3, of the output. However, they occasionally have to guess additional bits to determine the majority-based clocking sequence, and thus the total complexity of the attack is about $O(2^{45})$. This attack needs more than one month to find one key.
- Golic [1] described an improved attack which requires $O(2^{40})$ steps. This attack is based on the solution of a system of linear equations. It takes, however, more time than the previous algorithm.
- Golic [1] describes a general time-memory tradeoff attack on stream ciphers. It was independently discovered by Babbage [1] two years earlier. Golic [1] concludes that it is possible to find the key, of the A5/1 algorithm, in $O(2^{22})$ probes into random locations of a computed table with 242128 bit entries. Since such a table requires a hard disk of 64 terabytes, the space requirement is unrealistic. Alternatively, it is possible to reduce the space requirement to 862 gigabytes, but then, the number of probes increases to $O(2^{28})$. Since random access to the fastest commercially available PC disks requires about 6 milliseconds, the total probing time is almost three weeks. And it can only be used to attack GSM phone conversations which last more than 3 hours, and so it is unrealistic.
- Biryukov et al. [2] proposed two new attacks cryptanalytic on A5/1, in which a simple PC can extract the key from conversation in real time starting from some output generated. The first attack (called the biased birthday attack) requires two minutes of data and one second of processing time, whereas the second attack (called the random subgraph attack) requires two seconds of data and several minutes of processing time.

V. WEAKNESS OF A5/1

The major problem of stream ciphers, as the algorithm A5/1, is the difficulty of generating a long unpredictable bit pattern (key-stream). According to the results reached by the operations and the cryptanalysis made by the researchers, in addition to the various attacks listed in preceding section, the algorithm A5/1 suffers from several weaknesses. For this reason, the A5/1 stream cipher was classified in the column of low level of security algorithms. The set of huge weaknesses is the following.

- The first weakness is in the generation of the output sequence which is ensured by a linear function (simple XOR), which is considered weak to crack by a linear cryptanalysis.
- The second one is about the short period problem: without stop/go operation, the period clocking of sum of the three LFSRs is given by:

$$(2^{19}\text{-}1)(2^{22}\text{-}1)(2^{23}\text{-}1). \qquad (2)$$

However, the experiments show that the period of A5/1 is around $(4/3)\,(2^{23} - 1)$.

- The third weakness is caused by the collision problem: the different seeds (i.e., different initial states of three LFSRs) may result in the same key-stream and so, only about 70% of seeds that really generate different key-streams.
- The forth weakness is the majority function that is the worst function in terms of correlation with all affine functions.

VI. IMPROVED MODEL FOR A5/1

Before describing in detail our proposed improved model for A5/1, we recall that the basic model suffers from fore large weaknesses. In this paper, we propose two solutions. The first one is for improving the clocking mechanism, which is considered as one of the major weaknesses of A5/1. The second solution is concerning the problem of the linear output function that is reversible.

Therefore, the proposed model aims at improve those two weaknesses. The clocking mechanism is improved by a new function and the linear combining function of A5/1 is replaced by a more cryptographically better nonlinear function to strengthen the cipher. The architecture of the proposed scheme is shown in the Figure 2, and the model is detailed as follows:

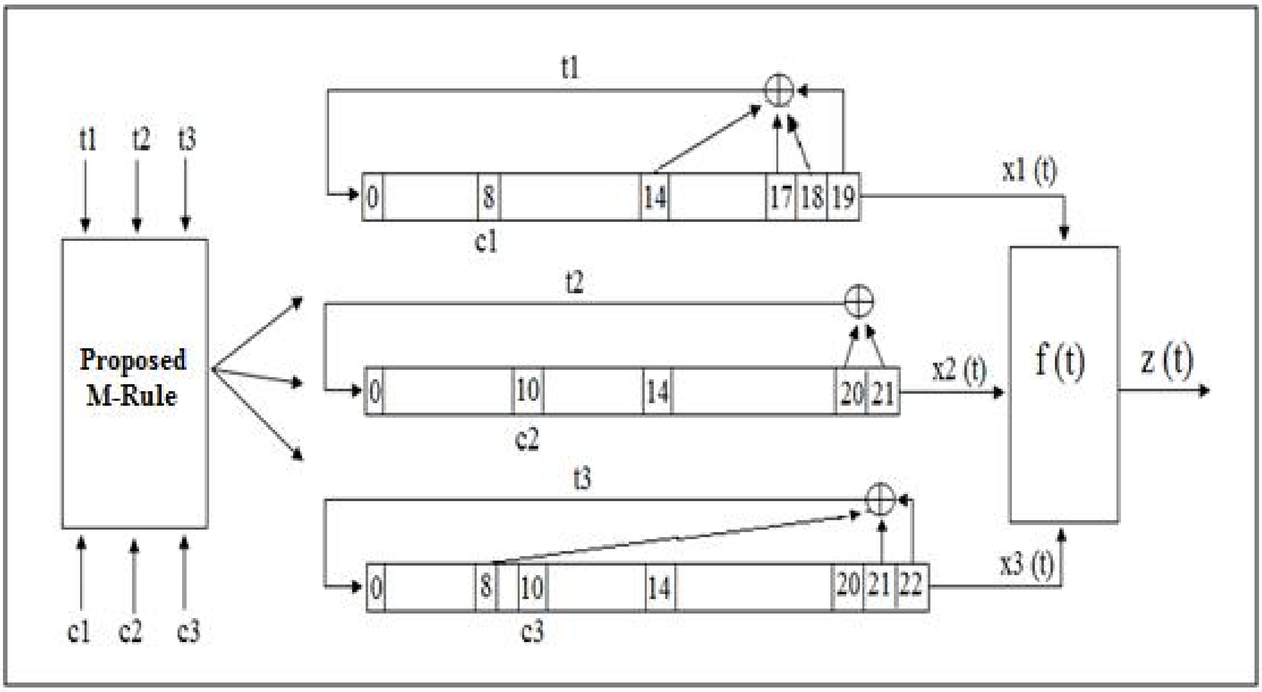

Figure 2. Proposed modified A5/1

A. Improved Unit of Clock-Controlling

As we mentioned in the previous section, we propose a new model to ameliorate the flaw of the clocking mechanism. The clock-controlling unit of the proposed scheme uses a new different rule to clock the registers. Instead of the existing M-rule, which is based on the one clocking taps bit (only one bit by register), the proposed rule consists of clocking each register by the output of its taps of the bits in combination with its clocking taps as well.

This mechanism can be governed by our proposed M-rule, which is defined by the following formula:

$$m = C1 . t1 \oplus C2 . t2 \oplus C3 . t3. \quad (3)$$

where "." represents the AND operation, and $\oplus$ represents the XOR operation. t1, t2, t3 are the outputs of the taps of the bits of the three registers (see Figure 2). They are defined by the three following formulas:

$$t1 = R1(13) \oplus R1(16) \oplus R1(17) \oplus R1(18).$$

$$t2 = R2(19) \oplus R2(20).$$

$$t3 = R3(7) \oplus R3(19) \oplus R3(20) \oplus R3(21).$$

where C 1, C 2, and C 3 are the clocking taps of the registers R1, R2 and R3, respectively. Once the majority bit is defined, we compare this bit with the t bit of each register, and then the decision will be taken. The register will be clocked only if the two bits are much. In other words, if m = t(i), for i = 1, 2 or 3, R(i) will be clocked. The proposed M-rule is illustrated in the Figure 3.

B. New Non-linear Combination Function

A5/1 stream cipher generates one bit at each clocking cycle. As we mentioned in the previous sections, the output of A5/1 is ensured by combining each the output bits of the three LFSRs using the XOR operation (see Figure 4). But the attackers prove that this function is weak and it is easy to recover the inputs data. In fact, this function is weak like all the linear functions because any linear function is reversible. This problem is treated by the searchers in the literature.
Sarkar and Maitra [15] say that the linear combination functions are cryptographically weak functions.

Thus, it is necessary to replace these functions with another one more secure and more complex. Thus, to increase the linear complexity of the A5/1 stream cipher and to overcome the weaknesses, due to the use of the linear function for combining, a new non linear combination function, which is more cryptographically better, is proposed here. The scheme of the proposed function is shown in Figure 5. Such proposed function is defined by the following formula:

$$Z(t) = f(x1, x2, x3) = (x1 . x2) \oplus (x1 \oplus x3) . (x2 . x3). \quad (4)$$

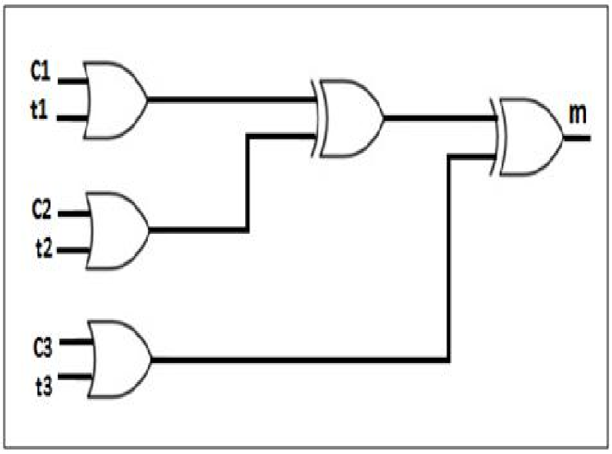

Figure 3. Proposed M-rule

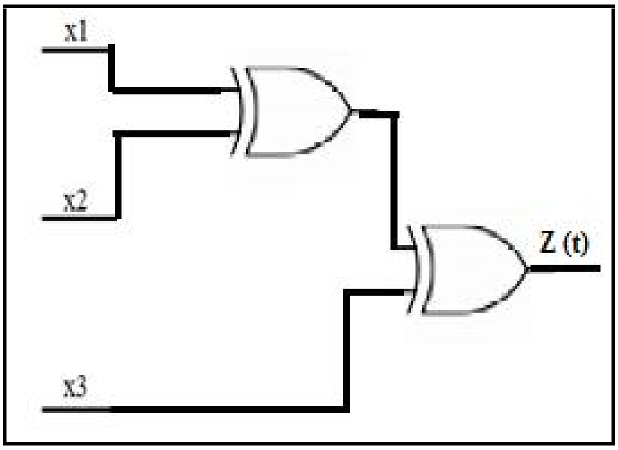

Figure 4. Original linear Function

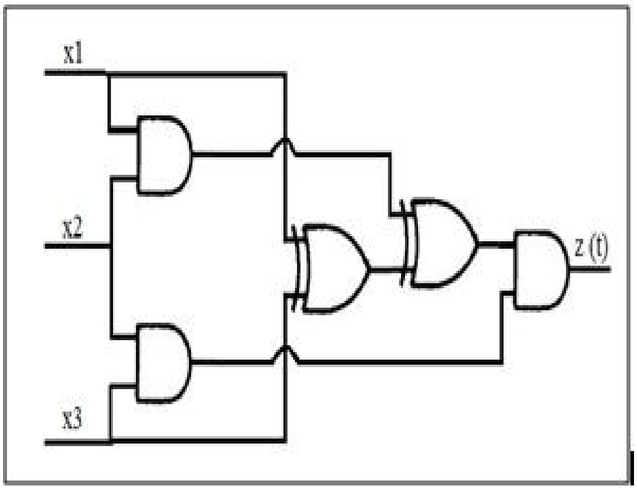

Figure 5. Proposed Non-linear Combination Function

where "." represents the AND operation, and ⊕ represents the XOR operation. x1, x2, x3 are the output bits of the three registers after clocking cycle, and z(t) is the output of the proposed scheme.

Like in the original algorithm, one bit is generated in each clocking cycle. Our proposed scheme is also based on the output of the three LFSRs, but the difference appears in the combining function, when the linear XOR operation is replaced by a new non-linear function f(x).

Now, we will show that the proposed scheme is more secure than the existing one because of its complexity and its non linearity. In fact, because of the clocking mechanism that becomes more irregular, the new nonlinear combining function is more secure than the existing one. However, because of the restriction of the number of the pages for this paper, we cannot detail this point here, but it will be included in the future works.

VII. SIMULATION RESULTS

This section presents the simulation results of the system that we have implemented using the proposed new non-linear combination function that we have detailed in Section VI. Note that for implementing our proposed system, we must go through several steps that we have described in details in the preceding sections. We have the preliminary results are shown in Figure 6. The difference between the linear output signal (above) and non-linear output signal (below), in terms of irregularities and complexities, is remarkable. It is due to the random aspect of the new non-linear function. In the future works, we intend to make the test of the effectiveness of our proposed system by securing data in GSM networks.

VIII. CONCLUSION AND DISCUSION

The aim behind the proposed enhancements of the A5/1 encryption algorithm, used in GSM standard, is to increase the linear complexity of the generated output sequence, and to accelerate the clocking mechanism. So, we have resolved two from the four problems listed in this paper that from which the A5/1 stream cipher suffers.

The proposed scheme can also retain the speed and it can be used for transmitting data in real time. Even though the algorithm became complex, it is easy in its implementation. We note that there is no explanation regarding the reason for selecting bits in the LFSR registers as tap bits because of the original architecture of A5/1, which is until this time unknown. There is no official description of A5/1 in the literature. The details of the algorithm have been published anonymously on a mailing list on the Internet, presumably due to a reverse engineering. So, we can conclude that the proposed scheme is better against to the cryptographic attacks compared to the conventional A5/1 stream cipher, since it generates cryptographically better binary sequence than the conventional one with slight increase in the hardware implementation.

 ISBN: 978-1-61208-347-6

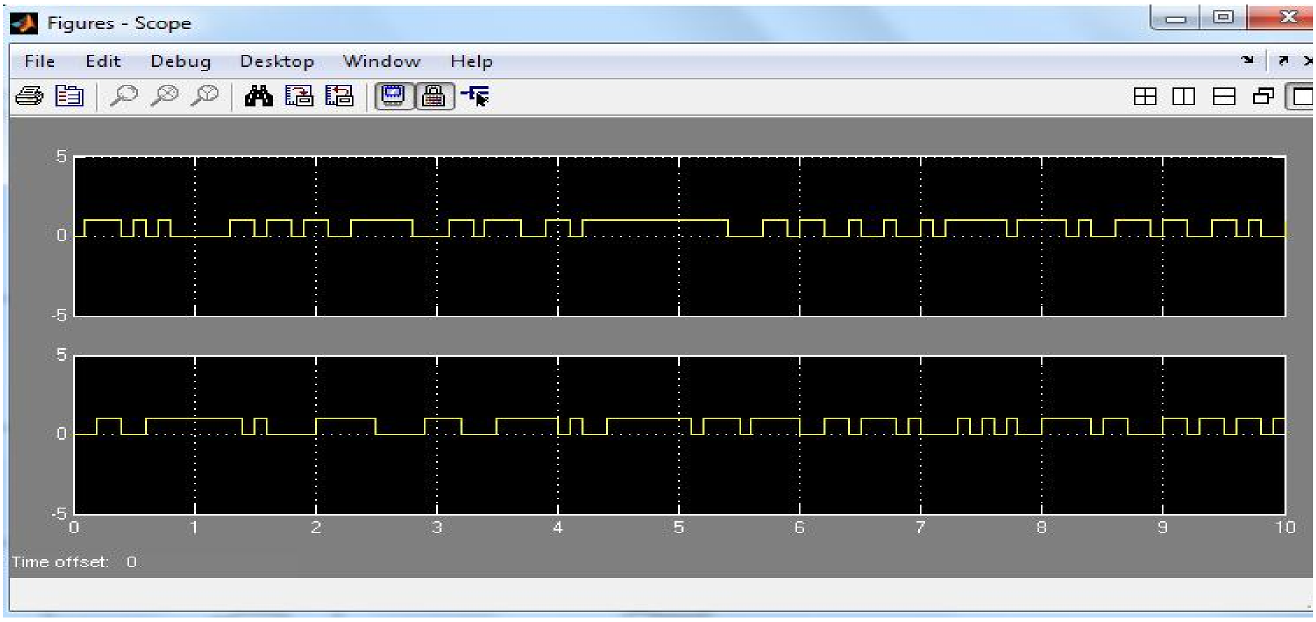

Figure 6. Comparaison between linear and non linear output signal

Regarding the cryptanalysis of the methods used in A5/1 algorithm versus proposed methods, the linear Boolean function-based attacks are not possible because our proposed majority function is a non-linear function. While for the non-linear function-based attacks, it will be very difficult since our majority function is based on new rule that is unknown.

We note that there are still other possibilities to improve the A5/1 algorithm to render it more secure without decreasing its speed. These possibilities will be the subject of our future works.

REFERENCES

[1] J. Golic, "Cryptanalysis of alleged A5 stream cipher," Advances in Cryptology. EUROCRYPT97. LNCS, vol. 1233, 1997, pp. 239-255, Springer-Verlag.

[2] A. Biryukov, A. Shamir, and D. Wagner, " Real time cryptanalysis of A5/1 on a PC," Advances in Cryptology. Fast Software Encryption. LNCS, vol.1978, 2001, pp. 1-18, Springer-Verlag.

[3] E. Biham and O. Dunkelman, "Cryptanalysis of the A5/1 GSM stream cipher," Progress in Cryptology, proceedings of INDOCRYPT00. LNCS, 2000, pp. 43-51, Springer-Verlag.

[4] P. Ekdahl and T. Johansson, "Another attack on A5/1," IEEE Transactions on Information Theory, vol. 49, Jan. 2003, pp. 284-289.

[5] A. Maximov, T. Johansson, and S. Babbage, "An improved correlation attack on A5/1," SAC 2004. LNCS, vol. 3357, Aug, 2004, pp. 1-18.

[6] E. Barkan and E. Biham, "Conditional estimators: an effective attack on A5/1," SAC 2005. LNCS, vol. 3897, 2006, pp. 1-19, Springer-Verlag.

[7] S. E. AlAschkar and M. T. El-Hadidi, "Known attacks for the A5/1 algorithm," International Conference on Information and Communications Technology (ICICT03), 2003, pp. 229-251.

[8] R. Alpesh Sankaliya, V. Mishra, and A. Mandloi. " Implementation of Cryptographic Algorithm for GSM and UMTS Systems," International Journal of Network Security & Its Applications (IJNSA), Vol.3, No.6, November 2011, pp. 81-88.

[9] G. Rose, "A précis of new attacks on GSM encryption," Qualcomm, Australia, Sept. 2003.

[10] A. Canteaut. Available from: http://www.springerreference.com/ [retrieved: December, 2013]

[11] J. Fernando, "Attacks on A5/1 Cryptography Algorithm," Makalah IF3058 Kriptografi Sem. II Tahun, [retrieved: December, 2013].

[12] M. Briceno, I. Goldberg, and D. Wagner, "A pedagogical implementation of A5/1". [Online]. Available from: http://cryptome.org/jya/a51-pi.htm,/ [retrieved: December, 2013].

[13] R. Anderson and M. Roe, A5. Available from http://cryptome.org/jya/crack-a5.htm,/[retrieved: April, 2014].

[14] S. Babbage, "A Space/Time Tradeoff in Exhaustive Search Attacks on Stream Ciphers," European Convention on Security and Detection. IEEE Conference publication, No. 408, May 1995, pp. 216-224.

[15] P. Sarkar and S. Maitra, " Construction of nonlinear Boolean functions with important Cryptographic properties," Advances in Cryptology EU- ROCRYPT00. LNCS, vol. 1807, 2000, pp. 485-506, Springer-Verlag.

 ISBN: 978-1-61208-347-6