Journal of Advanced Computer Applications
vol-01, issue-01, December, 2024

ISSN No.: XXXX-XXXX
Pages 32-39

JOURNAL OF ADVANCED COMPUTER APPLICATIONS

ISSN: XXXX-XXXX (Online)
Vol: 01 **Issue:** 01
Contents available at: https://www.swamivivekanandauniversity.ac.in/jaca/

A Study on Algebraic Attacks over Stream Cipher

Subrata Nandi

[1]*Department of Computer Science and Engineering, Swami Vivekananda University, Barrackpore-700121, WB, INDIA*

ABSTRACT

There are several known-plaintext attacks(KPA) on Stream cipher. Algebraic Attack is one kind of KPA. In this paper, we study the fundamental aspects of algebraic attack on Stream Cipher. We also study one of the important property related to Algebraic Attack, Algebraic Immunity.

Keyword: Algebraic attack, Stream Cipher, Algebraic Immunity.

I. INTRODUCTION

Stream ciphers play a crucial role in security in wireless communication. To produce the ciphertext bits, it does bitwise-XOR between plaintext bits and the pseudorandom bits(keystream). Stream cipher A5/1 was used in 2G mobile communication,and SNOW 3G, ZUC ciphers[1] are used for 4G and 5G mobile communication to restore confidentiality and integrity. The primary component of the Stream cipher is the keystream generator(KSG). Linear Feedback Shift Register(LFSR) is a very useful KSG. This is because of their low hardware cost, good statistical properties and good periods. Nonlinear Function isused with LFSR to resist the KSG from BMA attack.

Attacks against stream cipher are another threat. Algebraic attack is one of the attacks that Courtois and Meier[2] on EUROCRYPT 2003. There are two fundamental models of stream ciphers: combiner generator and filter generator, where a nonlinear boolean function f takes important roles to generate pseudorandom bits. It can be observed [3] that If a function f or its complement $1 + f$ has low degree annihilators, one can construct equations of degree equal to the degree of the annihilators. So, the designer should not use such boolean functions of having low-degree annihilators. To resist algebraic attack, algebraic immunity(AI) [4] takes a significant role. AI is nothing but the minimum degree annihilator between f or $1 + f$. Details ofthe study on algebraic immunity will be discussed in a later section.

* Authors for Correspondence

A) Literature Survey

Algebraic Attack was ideated by Courtois [2] in 2003. It found vulnerabilities in Toyocrypt, LILI-128 ciphers. This article[5] explains theoretical analysis regarding the algebraic immunity of nonlinear boolean functions. Later, Billet [6] explains the algebraic attacks on the cipher SNOW 2.0 in time complexity 2^{51}. In addition to that, [7] improved the attack with time complexity 2^{291}. Besides, [8] mentions the algebraic attack on the Welch-Gong family of stream ciphers. The article [9] attacks Bluetooth stream cipher E_0 with 2^{79} time complexity using SAT solver, Binary Decision Diagram and Grobner Basis.In this article, we explain the basics of Algebraic Attack.

II. PRE-REQUISITES

Here, we study some definitions and properties of Boolean functions. Weight $wt(x)$ of a vector x in F^n is the number of

one count in x. Let f be a n variable boolean function defined as follows

$f : V_n \rightarrow F_2$

where V_n is the domain of n dimensional vector space and F_2 is the binary field of 2 elements. The hamming distance between two boolean functions f and g of n variable is $\mathrm{wt}(f+g)$. The degree of a Boolean function is defined as the length of the longest monomial in its polynomial representation.

Algebraic normal form representation of a boolean function is defined as, if f is n variable Boolean functions in the polynomial form over the field F_2 with n many indeterminates $x_1, x_2, \dots, x_n, f$ can be represented in the ring $\mathbb{F}_2[x_1, x_2, \dots, s_n]/< x_1^2 - x_1, x_2^2 - x_2, \dots, x_n^2 - x_n >$ as follows:

$$f(x_1, x_2, \dots, x_n) = a_0 + \sum_{i=1}^{10} a_i x_i + \sum_{1 \le i < j \le n}^{10} a_i x_i + \cdots + a_{i_1,\dots,i_n} x_{i_1} \dots x_{i_{n-1}} + a_{1,\dots,n} x_1 \dots x_n$$

where $a_0, a_1, \dots, a_{1,\dots,n} \in \mathbb{F}_2$ are called the coefficient of the respective monomials. Boolean Function f_1 and f_2 can be defined as $d(f_1, f_2) = |\{x \in \mathbb{F}_2^n | f_1(x) \neq f_2(x)|$. The Walsh coefficient of a vector plays a crucial role in the cryptographic boolean function. It can be defined as for any vector $u \in \mathbb{F}_2^n$ the value:

$$W_f(u) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x) \oplus <u,x>}$$

The nonlinearity of a Boolean function $nl(f)$ is defined as

$$nl(f) = \{\min d(f, l) \mid \deg(l) \le 1\}$$

It can also be defined to walsh coefficient like the following:

$$nl(f) = 2^{n-1} - \frac{1}{2}\max_{u\in\mathbb{F}_2^n}|W_f(u)|$$

A) Algebraic Attack on NLFSR

This section will describe the existing literature on generating the algebraic equation of low degree. With two basic models of LFSR-based stream ciphers like a nonlinear combiner and a nonlinear filter generator, this algebraic attack can be possible. Let the model use k bit LFSRs, and each time, it is updated by a linear update function denoted by

$$L: \mathbb{F}_2^n \to \mathbb{F}_2^n$$

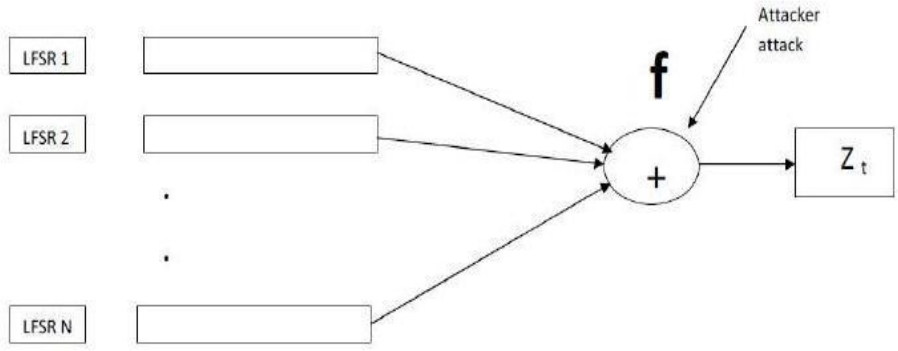

Fig. 1. Nonlinear Combiner Generator

Let the initial state be $\boldsymbol{S^0} = \{s_0, s_1, \dots, s_{k-1}\}$. At the t - th clock, the keystream output will be $z_t = f(S^t), t \geq 0$, where f is the nonlinear function. $S^t = L^t(S^0)$ denotes the state when the linear function L will be operated t times on the state S^0. The problem is to recover the initial state $\boldsymbol{S^0} = \{s_0, s_1, \dots, s_{k-1}\}$. If an adversary exploit the known plaintext attack, some l many keystream bits(say, $z_{k_1}, z_{k_2}, \dots, z_{k_l}$) are known. So it is easier to generate a system of equations of degree equal to $\deg(f)$ as follows:

$$f\left(L^{k_1}(S^0)\right) = z_{k_1}$$
$$f\left(L^{k_2}(S^0)\right) = z_{k_2}$$
$$\vdots$$
$$f\left(L^{k_l}(S^0)\right) = z_{k_l}$$

The time complexity of solving the system of equations increases if the degree of the nonlinear functions f is high. One may like to generate low-degree equations using some weakness in the internal structure of nonlinear functions. We know that $f\left(L^t(S^0)\right) = f(S^t) = z_t$. The main idea[10] is to use low-degree multiples and annihilators of the nonlinear function f to generate a low-degree equation. So multiplying $f(S^t)$ (usually high degree) with a well-chosen function $g(S^t)$ such that the degree of fg is reduced.

1. if $z_t = 1$, any function g in $AN(f)$ leads to $g\left(L^t(S^0)\right) = 0$.
2. if $z_t = 0$, any function h in $AN(1+f)$ leads to $h\left(L^t(S^0)\right) = 0$.

So if we can collect the relations to all functions of degree at most d (obviously $< \deg(f)$) in $AN(f) + AN(1+f)$ for known L keystream bits, we obtain a smaller degree equation on n variables $x_1, x_2, \ldots, x_n$. So, we can recover the bits of the initial state by solving the multivariate polynomial system.

Definition II.1. A Boolean function g over $\mathbb{F}_2^n$ is an annihilator $AN(f)$ of a Boolean function f over $\mathbb{F}_2^n$ if

$$fg = 0$$

Definition II.2. The algebraic immunity $AI(f)$ of a Boolean function f over $\mathbb{F}_2^n$ is the degree of the Boolean function g over $\mathbb{F}_2^n$ where g is a nonzero function of minimum degree such that $fg = 0$ or $(1+f)g = 0$.

It is known[7] that for any function f over $\mathbb{F}_2^n AI(f) \leq \left\lceil \frac{n}{2} \right\rceil$.

Solving the system of multivariate algebraic equations is an important area in computational algebraic geometry and commutative algebra. The problem is NP-complete even if all the equations are quadratic and the base field is $\mathbb{F}_2$. Some existing techniques to solve those multivariate equations are XL, XSL, and Grobner basis algorithms (F_4, F_5).

B) Theoretical Results on Algebraic Immunity

Theorem II.1. [4] Let $f \in B_n$ (set of n variable Boolean functions) and $AI_n(f) > d$. Then

$$\sum_{i=0}^{d} \binom{n}{i} \leq wt(f) \leq \sum_{i=0}^{n-(d+1)}$$

Proof. Let f has an annihilator g of degree d. Let the ANF of g is

$$a_0 + \sum_{i=1}^{10} a_i x_i + \sum_{1 \leq i < j \leq n}^{10} a_i x_i + \cdots + a_{i_1,\ldots,i_d} x_{i_1} \ldots x_{i_d}$$

where a's are from $\mathbb{F}_2$. We know that $f(x) = 1$ implies $g(x) = 0$, since $g \in AN(f)$. We will get $wt(f)$ many homogeneous equations on the a's.

Solving the system of homogeneous linear equations, we can find annihilators g of degree $\leq d$ on nontrivial solutions. In trivial case, we will get all a's are equal to zero in which we are not interested as we are interested in nonzero g.

Here, we have $\sum_{i=0}^{d} \binom{n}{i}$ number of variables and $wt(f)$ many equations. If the number of variables exceeds the number of equations, we will get nontrivial solutions. Thus, f has no annihilator g of degree d, implying the number of equations is greater than or equal to the number of variables. so there must be at least $\sum_{i=0}^{d} \binom{n}{i}$ number of equations, i.e., $wt(f) \geq \sum_{i=0}^{d} \binom{n}{i}$. Similarly, when considering $1 + f$, we get $wt(1 + f) \geq \sum_{i=0}^{d} \binom{n}{i}$. From this we can say, $wt(1 + f) \leq 2^n - \sum_{i=0}^{d} \binom{n}{i}$, i.e., $wt(f) \leq \sum_{i=0}^{n-(d+1)} \binom{n}{i}$.

It also gives alternative proof [4]$AI(f) \leq \left\lceil \frac{n}{2} \right\rceil$. The inequality in the above theorem will not be satisfied if $d > n - (d + 1) \Rightarrow d > \frac{n-1}{2} \Rightarrow d \geq \left\lceil \frac{n}{2} \right\rceil$. It is observed that for any f the inequality in the above theorem will not be satisfied if $AI_n(f) > d \geq \left\lceil \frac{n}{2} \right\rceil$.

The reverse of the theorem is not always true. for example, the affine functions are balanced, i.e., their weight is $2^{(}n - 1)$, but they have linear annihilators.

Based on the above theorem, the following results give bound on $wt(f)$, where f of $1 + f$ do not have annihilators of degree less than $\left\lceil \frac{n}{2} \right\rceil$.

Corollary II.1.1. $AI_n(f) = \left\lceil \frac{n}{2} \right\rceil$ implies

1. f is balanced when n is odd
2. $\sum_{i=0}^{\frac{n}{2}-1} \binom{n}{i} \leq wt(f) \leq \sum_{i=0}^{\frac{n}{2}}$ when n is even.

Theorem II.2. If $nl(f) < \sum_{i=0}^{d} \binom{n}{i}$, then $AI_n(f) \leq d + 1$[4].

Theorem II.3. [11]Let $f \in B_n$ and $AI_n(f) = k$. Then $nl(f) \geq 2^(n-1) - \sum_{i=k-1}^{n-k} \binom{n-1}{i} = 2\sum_{i=0}^{n-k} \binom{n-1}{i}$.

From the above discussion, we get $wt(f)$ many homogeneous linear equations using the a's. Let us denote the coefficient matrix of this system of equations by M. then M has $wt(f)$ many rows and $\sum_{i=0}^{d} \binom{n}{i}$. The rank(say, r) of the matrix M, $r \leq \min\left\{wt(f), \sum_{i=0}^{d} \binom{n}{i}\right\}$

1. If $r = \sum_{i=0}^{d} \binom{n}{i}\Big\}$, then there is no annihilator of degree $\leq d$.
2. If $r < \sum_{i=0}^{d} \binom{n}{i}\Big\}$, then there are annihilators of degree $\leq d$. There will be $\sum_{i=0}^{d} \binom{n}{i} - r$ many linearly independent annihilators having degree $\leq d$.

For any Boolean function f, the number of annihilators and linearly independent annihilators are $2^{wt(1+f)} - 1$ and $wt(1 + f)$. Suppose $M_{n,d}(f)$ is the matrix representation of boolean function f of n variables and algebraic degree d. Number of rows and columns of the matrix are $wt(f)$ and $\sum_{i=0}^{d} \binom{n}{i}$ respectively. An algorithm for Algebraic Immunity (AI) of Boolean function f is given below.

Algorithm 1 Find Algebraic Immunity of a Boolean Function

```
for i = 1 → ⌈n/2⌉ do
  Find the rank R_1 of the matrix M_{n,i}(f).
  Find the rank R_2 of the matrix M_{n,i}(1 + f).
  if min{R_1, R_2} ≤ Σ_{j=0}^{i} (n choose j) then
    Output i
  end if
end for
```

If f is a balanced boolean function, the time complexity of the above algorithm is approximately $(2^{n-2})^3$.

III. CONCLUSION

In this article, we study Algebraic attacks on Stream Cipher and the Algebraic Immunity property. We understand that the Boolean function with a low algebraic degree is prone to cryptanalysis. As the algorithm mentioned above, to find algebraic immunity of the Boolean function, is exponential, it could be a better problem to propose an efficient algorithm than the existing one.

REFERENCES

[1] A. Mufeed, "A different algebraic analysis of the zuc stream cipher," *Proceedings of the 4th international conference on Security of information and networks, Brisbane, Australia*, pp. 191–198, 2011.

[2] T. Nicolas, Courtois, and M. Willi, "Algebraic attacks on stream ciphers with linear feedback," *In Advances in Cryptology*

- *Eurocrypt 2003, number 2656 in Lecture Notes in Computer Science*, pp. 345–359, 2003.

[3] W. Meier and O. Staffelbach, "Nonlinearity criteria for cryptographic functions," in *Proceedings of the Workshop on the Theory and Application of Cryptographic Techniques on Advances in Cryptology*, ser. EUROCRYPT'89. New York, NY, USA: Springer-Verlag New York, Inc., 1990, pp. 549–562. [Online]. Available: http://dl.acm.org/citation.cfm?id=111563.111614

[4] D. K. Dalai, "On boolean functions to resist algebraic attacks: Some necessary conditions," Vdm Verlag Dr. Miller, 2010.

[5] D. K. Dalai, K. C. Gupta, and S. Maitra, "Cryptographically significant boolean functions: Construction and analysis in terms of algebraic immunity," in *International Workshop on Fast Software Encryption*. Springer, 2005, pp. 98–111.

[6] O. Billet and H. Gilbert, "Resistance of snow 2.0 against algebraic attacks," in *Topics in Cryptology–CT-RSA 2005: The Cryptographers' Track at the RSA Conference 2005, San Francisco, CA, USA, February 14-18, 2005. Proceedings*. Springer, 2005, pp. 19–28.

[7] N. T. Courtois and B. Debraize, "Algebraic description and simultaneous linear approximations of addition in snow 2.0." in *Information and Communications Security: 10th International Conference, ICICS 2008 Birmingham, UK, October 20-22, 2008 Proceedings 10*. Springer, 2008, pp. 328–344.

[8] S. Rønjom, "Improving algebraic attacks on stream ciphers based on linear feedback shift register over f 2ˆ k f 2 k,"

[9] *Designs, Codes and Cryptography*, vol. 82, pp. 27–41, 2017.

[10] R. La Scala, S. Polese, S. K. Tiwari, and A. Visconti, "An algebraic attack to the bluetooth stream cipher e0," *FiniteFields and Their Applications*, vol. 84, p. 102102, 2022.

[11] G. M., R. and J. D., *Computers and Intractability*. W H Freeman publisher, 1999.

[12] L. Mikhail, "Tight bounds between algebraic immunity and nonlinearities of high orders," Cryptology ePrint Archive, Report 2007/444, 2007, http://eprint.iacr.org/.