Designing resilient functions and bent function for stream ciphers

Belmeguenaï Aïssa[1], Doghman Nouredine[2]

[1]Department of electronics, Faculty of Science and Engineering, August 20 University- Skikda LP 26 El-hadeik Avenue
Email: belmeguenaiaissa@yahoo.fr
[2]Department of electronics, Faculty of Science and Engineering, Badji Mokhtar University - Annaba LP 12.
Email: ndoghmane1@yahoo.fr

Abstract:

In this paper, we study a certain class of resilient functions with highest possible algebraic immunity or with a reasonably high algebraic immunity which achieves nonlinearity better than that obtained by M. Lobanov, if the non linearity of initial functions achieves Sarkar et al's bound.

Keywords: *Stream cipher, Boolean Function, Algebraic Degree, Resiliency, Nonlinearity, Algebraic Immunity, Annihilator.*

I. Introduction

Boolean functions, when used in stream cipher, are required to have good cryptographic properties. Some of the important properties are balance, high algebraic degree, correlation immunity of reasonably high order and high non linearity; these properties ensure that the functions are resistant against correlation attacks [1] and linear cryptanalysis [2]. Many papers studied these properties. Unfortunately, they showed that certain properties all cannot be obtained simultaneously and the compromises must be found. The nonlinearity of a Boolean function does not exceed $2^{n-1} - 2^{\frac{n}{2}-1}$ [3]. A function achieving this nonlinearity can not be balanced. Siegenthaler showed in [4] that any *n*-variable t*-th* order correlation immune function ($0 \le t \le n$) has algebraic degree smaller than or equal to $n-t$, and that any *n*-variable t-resilient function ($0 \le t \le n$) has algebraic degree smaller than or equal to $n-t-1$ if $t \le n-2$ and equal to 1 if $t = n-1$.

Sarkar and Maitra demonstrated in [5] a divisibility bound on the Walsh transform values of an *n*-variable, t*-th* order correlation immune (resp. t-resilient) function, with $t \le n-2$: these values are divisible by 2^{t+1} (resp. by 2^{t+2}). This provided a nontrivial upper bound on the nonlinearity of resilient functions (and also of correlation immune functions, but non-balanced functions present less cryptographic interest), independently obtained by Tarannikov [6] and by Zheng and Zhang [7]: the nonlinearity of any *n*-variable, t-resilient function is upper bounded by $2^{n-1} - 2^{t+1}$. Tarannikov proved that resilient functions achieving this bound must have degree $n-t-1$ (that is, achieve Siegenthaler's bound); thus, they achieve best possible trade-offs between resiliency order, degree and nonlinearity.

Theses criteria were until recently the only requirements needed for the design of the Boolean functions used in a stream cipher systems as a combining functions or as a filtering one. The recent algebraic attacks [8], [9], [10], [11], [12], [13], [14] have complicated this situation by adding the new criterion of algebraic immunity, of considerable importance to this list. Today, it is known that to resist the attacks of the type algebraic attacks, and the selected Boolean functions must have a degree of algebraic immunity greater than seven. Unfortunately these criteria are generally incompatible what obliges the cryptograph to seek compromises. It is the reason for which research on the Boolean functions is very active and especially capital. The stake to seek compromises

between these criteria is essential to make it possible to better apprehend the safety of the cipher systems which use the Boolean functions as cryptographic primitives.

Nonlinearity is the most important criterion among those cryptographic criteria on Boolean functions used in cipher systems (stream ciphers, bloc ciphers). In [15], Lobanàov has improved a lower bound (between nonlinearity and algebraic immunity) obtained in [16] on the (first-order) nonlinearity of Boolean functions f a n-variables with given algebraic immunity [15], which was: $Nf \geq 2\sum_{i=0}^{AI(f)-2}\binom{n-1}{i}$. Moreover, by constructing a family of Boolean function achieving the equality $Nf = 2\sum_{i=0}^{AI(f)-2}\binom{n-1}{i}$, he proved that this lower bound cannot be improved further. Lobanov's result has been extended to the r th-order nonlinearity Nf_r of an n-variable Boolean function f in two different lower bounds [17], [18]. The result of [15] gives a new reason why one should not use functions f with low nonlinearity, since in that case $AI_n(f)$ would be low. However, they do not assure that if f has high algebraic immunity (for instance an optimum one $AI_n(f) = \left\lceil \frac{n}{2} \right\rceil$) then its nonlinearity will be high. Indeed, the result of [15] implies then that f has nonlinearity at least $2\sum_{i=0}^{\left\lceil \frac{n}{2} \right\rceil -2}\binom{n-1}{i}$, that is, $2^{n-1} - \binom{n-1}{\frac{n-1}{2}}$ if n is odd and $2^{n-1} - \binom{n-1}{\frac{n}{2}-1} - \binom{n-1}{\frac{n}{2}}$ if n is even.

In this paper, we give a construction method which can get a large class of Boolean functions with an important nonlinearity and algebraic immunity. It shows that we can construct reslient functions with a reasonably high algebraic immunity having non linearity better than that obtained by Lobanov in [15], when the non linearity of initial functions achieve Sarkar et al's bound. This construction does not increase the number of variables, contrary to the known general secondary constructions.

The paper is organized as follows. First, we will give some preliminaries of the paper. In section 3, we give a construction to get a large numerous Boolean functions with an important non linearity and algebraic immunity. In section 4 we study resilient functions with algebraic immunity $k \prec\prec \left\lceil \frac{n}{2} \right\rceil$, achieves nonlinearity better than that obtained by M. Lobanov.

II. Preliminaries

A Boolean function on n variables may be viewed as a mapping from F_2^n in to F_2. The set of all n-variable Boolean function is denoted by B_n. By $\oplus$ we denote sum modulo 2. The Hamming weight $wt(f)$ of a Boolean function f on F_2^n is the size of its *support* $\{x \in F_2^n ; f(x) = 1\}$. The Hamming distance $d(f,g)$ between two Boolean functions f and g is the Hamming weight of their difference $f \oplus g$, $d(f,g) = wt(f \oplus g)$.An n-variable Boolean function f has unique algebraic normal form (A.N.F): $f(x_1,...,x_n) = a_0 + \sum_{i=0}^{n} a_i x_i + \sum_{1 \leq i \prec j \leq n} a_{ij} x_i x_j + ... + \sum_{1 \leq i_1 \leq ... \leq i_{k-1} \leq n} a_{i_1 ... i_n} x_{i_1} ... x_{i_n}$.

The algebraic degree of Boolean function f, denoted by $d^\circ(f)$, is defined as the number of variables in the longest term of f. If algebraic degree of f is smaller than or equal to one then f is

called affine function. An affine function with a constant term equal to zero is called a linear function. Many properties of Boolean functions can be described by the Walsh-Hadamard transform. Let f be Boolean function on F_2^n. Then the Walsh-Hadamard transform of f is defined as:

$$\forall u \in F_2^n, Wf(u) = \sum_{x \in F_2^n} (-1)^{f(x)} (-1)^{u.x} . \quad (1)$$

Where $u.x$ denoted the usual scalar product of vectors u and x.
The nonlinearity Nf of an n variables function f is the minimum distance from the set of all n variables affine function, it equal to:

$$Nf = 2^{n-1} - \frac{1}{2} \max_{u \in F_2^n} |Wf(u)| . \quad (2)$$

Boolean functions used in cipher system must have high nonlinearity to prevent correlation and linear attacks [19], [20], [21], [1]. A Boolean function f on F_2^n is balanced if $wt(f) = wt(f \oplus 1)$. Otherwise, f is balanced if and only if $wt(f) = 2^{n-1}$. Correlation immune functions and resilient functions are two important classes of Boolean functions. Xiao and Massey [22] provided a spectral characterization of correlation immune and k-*th* resilient functions. A function f is k-*th* order correlation immune if and only if its Walsh transform f satisfies: $Wf(u) = 0$, for $1 \le wt(u) \le k$, where *wt(u)* denotes the Hamming weight of *u*, and f is k-*th* resilient if moreover $Wf(0) = 0$. $\forall u \in F_2^n$, $0 \le wt(u) \le k$.

The algebraic immunity of a Boolean function f is the smaller degree of non null function g such that $f * g = 0$ or $(1+f) * g = 0$. Otherwise, the minimum value of d such that f or $f+1$ admits an annihilator of degree d. We denoted by $AI(f)$ the algebraic immunity of a Boolean function f. It is shown in [13] and [23] that algebraic immunity of a Boolean function f is at most $\lceil \frac{n}{2} \rceil$. By $AN(f)$ we mean the set of annihilators of f.

III. Construction of Booleans functions

The idea of our construction comes from the following.

Construction 1: *Let k, n be any two positive integers such that $k \prec\prec \left\lfloor \frac{n}{2} \right\rfloor$. Let g and f be two Booleans functions of B_n with the following conditions.*

1. *g equal zero for $wt(x) \prec k$ and $wt(x) \succ n-k$,*
2. $$f(x) = \begin{cases} 0 & \text{if } wt(x) \prec k \\ g(x) & \text{if } k \le wt(x) \le n-k \\ 1 & \text{if } wt(x) \succ k-n \end{cases}$$

Then we have the following important result.

Lemma 1 *Let $f \in B_n$ be a function as described in Construction 1. Then $IA_n(f) \ge k$.*

Proof: We will show that the functions f and $1\oplus f$ have not a nonzero annihilator of degree less than k.
Writ the possible annihilator h of the functions f and $1\oplus f$ of degree at most $k-1$ by means of indeterminate coefficients:

$$h = a_0 + \sum_{i=0}^{n} a_i x_i + \sum_{1\le i\prec j\le n} a_{ij} x_i x_j + ... + \sum_{1\le i_1\le ...\le i_{k-1}\le n} a_{i_1...i_{k-1}} x_{i_1}...x_{i_{k-1}}.$$

1. The function h is the annihilator of f if only if $f(x)=1$ follows $h(x)=0$. we obtain the system of homogeneous linear equations on the coefficients of the function h: $h(x)=0$.
For all vectors x of Hamming weight greater than $n-k$.
Since $h(x)=0$, we obtain that all coefficients of h are zeros, hence, $h\equiv 0$.

2. The function h is the annihilator of $1\oplus f$ if only if $1\oplus f(x)=1$ follows $h(x)=0$. We obtain the system of homogeneous linear equations on the coefficients of the function h: $h(x)=0$.
For all vectors x of Hamming weight less than or equals $k-1$.
Since $h(0,...,0)=0$, we have $a_0=0$. Since $h(x)=0$ if $wt(x)=1$, we have $a_i=a_0=0$. Applying the induction on the weight of vectors we obtain that all coefficients of h are zeros, hence, $h\equiv 0$.
Hence, according to item 1 and item 2, we have $IA_n(f)\ge k$.

Lemma 2 *Let $f\in B_n$ be a function as described in Construction 1. Then*

$$\sum_{i=0}^{k-1}\binom{n}{i} \le wt(f) \le \sum_{i=0}^{n-k}\binom{n}{i}. \qquad (3)$$

Proof: Let h be a Boolean function of degree $k-1$. Let the ANF of h equal $h = a_0 + \sum_{i=0}^{n} a_i x_i + \sum_{1\le i\prec j\le n} a_{ij} x_i x_j + ... + \sum_{1\le i_1\le ...\le i_{k-1}\le n} a_{i_1...i_{k-1}} x_{i_1}...x_{i_{k-1}}$. The function h is an annihilator of f if only if $f(x)=1$ follows $h(x)=0$. Then in order to provide $IA_n(f)\succ k$, it is necessary that obtained homogeneous system of linear equations on coefficients $a_0, a_1, ...$ has the only zero solution. For this it is necessary that the number of unknowns $\sum_{i=0}^{k-1}\binom{n}{i}$ does not exceed the number equations $Wt(f)$. Hence, the left inequality is proved. Applying the same reasoning to $1\oplus f$ we obtain $wt(1\oplus f)\ge\sum_{i=0}^{k-1}\binom{n}{i}$. This gives, $wt(f)\le\sum_{i=0}^{n}\binom{n}{i}-\sum_{i=0}^{k-1}\binom{n}{i}$, i.e $wt(f)\le\sum_{i=0}^{n-k}\binom{n}{i}$.

Lemma 3 *Let $f\in B_n$ be a function as described in Construction 1. Then the value of the Walsh transform of f at every $u\in F_2^n$ equals:*

$$Wf(u) = Wg(u) \qquad (4)$$

Proof: for every $u\in F_2^n$, we have

$$Wf(u) = \sum_{x\in F_2^n}(-1)^{f(x)\oplus u\cdot x} = \sum_{x\in F_2^n / wt(x)\prec k}(-1)^{0\oplus u\cdot x} + \sum_{x\in F_2^n / k\le wt(x)\le n-k}(-1)^{g(x)\oplus u\cdot x} + \sum_{x\in F_2^n / wt(x)\succ n-k}(-1)^{1\oplus u\cdot x}$$

$$= \sum_{i=0}^{k-1}\binom{n}{i}\delta_0(u) + Wg(u) - \sum_{i=n-k+1}^{n}\binom{n}{i}\delta_0(u) = \sum_{i=0}^{k-1}\binom{n}{i}\delta_0(u) + Wg(u) - \sum_{i=0}^{k-1}\binom{n}{i}\delta_0(u) = Wg(u).$$

Where $\delta_0(u)$ is Dirac function at zero.

IV. Construction of resilient functions

We use now the result of Lemma 3 to build resilient functions with reasonably high algebraic immunity and high nonlinearities.

Theorem 1: *Let $f \in B_n$ be a function as explained in Construction 1. Then, if g is t-th order correlation immune (res. t-resilient). Then, the function f is t-th order correlation immune (res. t-resilient). Moreover:*

$$Nf = 2^{n-1} - \frac{1}{2}\max_{u\in F_2^n}|Wg(u)| \tag{5}$$

Proof: Relation (4) and the fact that for every non zero vector $u \in F_2^n$ of Hamming weight at most t, we have the value $|Wg(u)|$ equal to zero. This implies that $Wf(u) = 0$. Thus, f *is* t-th order correlation immune. Same property occurs for $u = 0$ in the case g is t-resilient.

The relation (4) implies $\max_{u\in F_2^n}|Wf(u)| = \max_{u\in F_2^n}|Wg(u)|$, that is, using relation (2), we have $Nf = 2^{n-1} - \frac{1}{2}\max_{u\in F_2^n}|Wg(u)|$, that is relation (5).

Remark 1: The family of functions described by construction 1 is very general. It is easy to see that construction 1 makes possible to define more numerous Boolean functions and resilient functions with an important non linearity and algebraic immunity. Thus, this construction permits to design a large class of resilient functions with algebraic immunity $k \prec\prec \left\lfloor \frac{n}{2} \right\rfloor$ or with a reasonably high algebraic immunity having non linearity better than that obtained by Lobanov [15], when the non linearity of initial functions achieve Sarkar et al's bound.
In the following corollary 1, we will show that the nonlinearity of function $f \in B_n$ as described in construction 1 can achieve the best possible nonlinearity (Sarkar et al's bound), better than obtained by Lobanov[15].

Corollary 1 *Let $f \in B_n$ be t-resilient function as described in construction 1. If g is t-resilient achieve nonlinearity $2^{n-1} - 2^{t+1}$. Then f is t-resilient function achieve nonlinearity $2^{n-1} - 2^{t+1}$.*

Note that in [15] it was constructed the balanced function f *of* n-variables with the maximum possible algebraic immunity $k = \left\lceil \frac{n}{2} \right\rceil$ and nonlinearity $Nf = 2^{n-1} - \sum_{i=k-1}^{n-k} \binom{n-1}{i}$, that is $2^{n-1} - \binom{n-1}{\frac{n-1}{2}}$ if n is odd and $2^{n-1} - \binom{n-1}{\frac{n}{2}-1} - \binom{n-1}{\frac{n}{2}}$ if n is even. Our corollary 1 proved that it is possible to design balanced function with algebraic immunity $k \prec\prec \left\lceil \frac{n}{2} \right\rceil$ achieve nonlinearity $2^{n-1} - 2$, better than that obtained by Lobanov [15].

In the following theorem, we will use the observation of lemma 3 for construction bent functions. Recall that any n-variables bent function f (n even), admit a dual $\tilde{f}$ defined as: for every vector $u \in F_2^n$, we have $Wf(u) = 2^{\frac{n}{2}}(-1)^{\tilde{f}(u)}$.

Theorem 2: *Let k,n be any two positive integers such that $k \prec\prec \left[\frac{n}{2}\right]$. Let*

$$f(x) = \begin{cases} 0 & if \quad wt(x) \prec k \\ g(x) & if \quad k \le wt(x) \le n-k \\ 1 & if \quad wt(x) \succ k-n \end{cases}$$

be Boolean function of B_n. Than, if g is bent, than $f \in B_n$ is bent and dual of f is equal dual of g. Moreover

$$Nf = 2^{n-1} - 2^{\frac{n}{2}-1} \tag{6}$$

Proof: By hypothesis for every vector $u \in F_2^n$, we have $Wg(u) = 2^{\frac{n}{2}}(-1)^{\tilde{g}(u)}$. Relation (4) implies the relation $Wf(u) = Wg(u) = 2^{\frac{n}{2}}(-1)^{\tilde{g}(u)}$, what concludes the demonstration.

The relation (4) implies $\max_{u \in F_2^n} |Wf(u)| = \max_{u \in F_2^n} |Wg(u)| = 2^{\frac{n}{2}}$, that is, using relation (2), we have $Nf = 2^{n-1} - 2^{\frac{n}{2}-1}$, that is relation (6).

References

[1] T. Siegenthaler. Decrypting a class of stream ciphers using cipher text only. *IEEE Transactions on Computers*, C-34, N°1:81–85, January 1985.

[2] C. Ding, G. Xiao, and W. Shan. The stability theory of stream ciphers, *Number 561*, Lecture Notes in Computer Science, Springer Verlag, August 1991.

[3] O. S. Rothaus. On bent functions, *Journal of Combinatorial Theory*, Series A20, pp. 300-305.

[4] T. Siegenthaler. Correlation-immunity of nonlinear combining functions for cryptographic applications. *IEEE Transactions on Information Theory*, IT-30, N°5:776–780, September 1984.

[5] P. Sarkar and S. Maitra. Nonlinearity bounds and construction of resilient Boolean functions. In: *Advances in Cryptology - EUROCRYPT 2000*, vol. 1880 in Lecture Notes in Computer Science, pages 515–532. Springer Verlag, 2000.

[6] Y. V. Tarannikov. On resilient Boolean functions with maxim um possible nonlinearity. *Proceedings of INDOCRYPT 2000, lecture Notes in Computer Science 1977, pp19-30, 2000.*

[7] Y. Zheng et X. M. Zhang. Improving upper bound on the non linearity of high order correlation immune functions. *Proceedings of Selected Areas in Cryptography 2000*, Lecture Notes in computer Science 2012, pp262-274, 2001.

[8] J. Y. Cho and J. Pieprzyk. Algebraic Attacks on SOBER-t32 and SOBER-128. In *FSE 2004*, number 3017 in Lecture Notes in Computer Science, pages 49–64. Springer Verlag, 2004.

[9] N. Courtois and J. Pieprzyk. Cryptanalysis of block ciphers with overde-fined systems of equations. In *Advances in Cryptology – ASIACRYPT 2002*, number 2501 in Lecture Notes in Computer Science, pages 267–287. Springer Verlag, 2002.

[10] N. Courtois and W. Meier. Algebraic Attacks on Stream Ciphers with Linear Feedback. *Advances in cryptology– EUROCRYPT 2003*, Lecture Notes in Computer Science 2656, pp. 345-359, Springer, 2003.

[11] N. Courtois. Fast Algebraic Attacks on Stream Ciphers with Linear Feedback. *advances in cryptology–CRYPTO 2003*, Lecture Notes in Computer Science 2729, pp. 177-194, Springer, 2003.

[12] D. H. Lee, J. Kim, J. Hong, J. W. Han and D. Moon. Algebraic Attacks on Summation Generators. In *FSE 2004*, number 3017 in Lecture Notes in Computer Science, pages 34–48. Springer Verlag, 2004.

[13] W. Meier, E. Pasalic and C. Carlet. Algebraic attacks and decomposition of Boolean functions. In *Advances in Cryptology - EUROCRYPT 2004*, number 3027 in Lecture Notes in Computer Science, pages 474–491. Springer Verlag, 2004.

[14] F. armknecht. Improving Fast algebraic Attacks. *In FSE 2004*, number 3017 in lecture Notes in computer Science, pages 65-82. Springer Verlag, 2004.

[15] M.Lobanov. Tight bound between nonlinearity and algebraic immunity. Paper 2005/441 in http://eprint.iacr.org/.

[16] D. K. Dalai, K. C. Gupta and S. Maitra. Results on Algebraic Immunity for Cryptographically Significant Boolean Functions. *Indocrypt 2004, Chennai*, India, December 20–22, pages 92–106, number 3348 in Lecture Notes in Computer Science, Springer Verlag, 2004

[17]. C. Carlet. On the higher order nonlinearities of algebraic immune Boolean functions, *CRYPTO 2006*, Lecture notes in Computer Science, vol. 4117, 2006, pp. 584–601.

[18]. C. Carlet, D. K. Dalai, K. C. Gupta, and S. Maitra. Algebraic immunity for cryptographically significant Boolean functions: Analysis and construction, *IEEE Transactions on Information Theory 52 (2006)*, no. 7, 3105–3121.

[19] A. Canteaut and M. Trabbia. Improved fast correlation attacks using parity-check equations of weight 4 and 5, *Advanced in Cryptology-EUROCRYPT 2000*. Lecture notes in computer science 1807 (2000), pp. 573-588.

[20] Johansson, T. and J¨onsson, F. Improved fast correlation attack on stream ciphers via convolutional codes. *Advances in Cryptology - EUROCRYPT'99*, number 1592 in Lecture Notes in Computer Science (1999), pp. 347–362.

[21] T. Johansson and F. Jonsson. Fast correlation attacks based on turbo code techniques. *Advances in Cryptology - CRYPTO'99*, number 1666 in Lecture Notes in Computer Science (1999), pp. 181–197.

[22] G.-Z. Xiao and J. L. Massey. A spectral characterization of correlation-immune combining functions. *IEEE Trans. Inf. Theory*, Vol IT 34, n° 3, pp. 569-571, 1988.

[23] N. Courtois and W. Meier. Algebraic attacks on stream ciphers with linear feedback. *In Advances in Cryptology - EUROCRYPT 2003*, number 2656 in Lecture Notes in Computer Science, pages 345–359. Springer Verlag, 2003.

Article received: 2007-10-01