On Algebraic Immunity and Annihilators

Xian-Mo Zhang[1], Josef Pieprzyk[1], and Yuliang Zheng[2]

[1] Centre for Advanced Computing - Algorithms and Cryptography
Department of Computing, Macquarie University
Sydney , NSW 2109, Australia
xianmo,josef@ics.mq.edu.au
[2] Department of Software & Information Systems
The University of North Carolina at Charlotte
9201 University City Blvd, Charlotte
NC 28223-0001, USA
yzheng@uncc.edu

Abstract. Algebraic immunity $AI(f)$ defined for a boolean function f measures the resistance of the function against algebraic attacks. Currently known algorithms for computing the optimal annihilator of f and $AI(f)$ are inefficient. This work consists of two parts. In the first part, we extend the concept of algebraic immunity. In particular, we argue that a function f may be replaced by another boolean function f^c called the algebraic complement of f. This motivates us to examine $AI(f^c)$. We define the extended algebraic immunity of f as $AI^*(f) = \min\{AI(f), AI(f^c)\}$. We prove that $0 \leq AI(f) - AI^*(f) \leq 1$. Since $AI(f) - AI^*(f) = 1$ holds for a large number of cases, the difference between $AI(f)$ and $AI^*(f)$ cannot be ignored in algebraic attacks. In the second part, we link boolean functions to hypergraphs so that we can apply known results in hypergraph theory to boolean functions. This not only allows us to find annihilators in a fast and simple way but also provides a good estimation of the upper bound on $AI^*(f)$.

Key Words: Algebraic Attacks, Algebraic Immunity, Hypergraph Theory, Greedy Algorithm.

1 Introduction to Algebraic Immunity

Recent algebraic attacks [4, 5, 3, 6, 14, 7, 8, 2, 1, 16, 9] have become a powerful tool that can be used for almost all types of cryptographic systems. Normally an algebraic attack is run in two stages. In the first stage, attackers build algebraic equations that reflect the relations between inputs, outputs and a secret key. In the second stage, attackers solve the algebraic equations in order to discover the secret key or restrict the secret key to a small domain (then exhaustively search the small domain). Algebraic attacks will be more efficient if algebraic equations have low degrees because the number of monomials (terms) of low degree is relatively small. Using annihilators is one of techniques to enable us to produce algebraic equations of low degree. Algebraic attacks have been used very successfully to analyse LFSR-based stream ciphers because all the algebraic equations preserve their algebraic degree. The concept of annihilators for

algebraic attacks was introduced by Courtois and Meier in [5]. For a boolean function f with n-bit inputs, $AN(f)$ is a set of boolean functions, defined as $AN(f) = \{g : (GF(2))^n \rightarrow GF(2) | f(x)g(x) = 0, \text{ for all } x \in (GF(2))^n\}$. Each function $g \in AN(f)$ is called an *annihilator* of f. Courtois and Meier [5] proposed three different scenarios to reduce the degree of algebraic equations. They discussed the relation among the three scenarios S3a, S3b and S3c in [6]. Later Meier at el. [14] showed that the scenario S3c can be replaced by the scenario S3a. Dalai et al [8] demonstrated that all the scenarios are equivalent to finding the union of two related annihilators, namely, $AN(f)$ and $AN(1 \oplus f)$ and then they defined the *algebraic immunity* $AI(f)$ as the minimum degree of nonzero boolean functions in $AN(f) \cup AN(1 \oplus f)$.

We now explain an application of annihilators to algebraic attacks. We may consider two types of algebraic equations, namely $f(x) = 0$ or $f(x) = 1$. For an algebraic equation $f(x) = 0$, multiplying the equation by g_1, such that $g_1 f = h$ is of a lower degree than the degree of f. Consequently, the attackers obtain a lower degree equation $h(x) = 0$. For the algebraic equation $f(x) = 1$, multiplying the equation by g_2 of a low degree such that $g_2 f$ is identical to the constant zero. Then the attackers obtain a lower degree equation $g_2(x) = 0$.

Courtois and Meier [5, 6] studied $AI(f)$ and proved that $AI(f) \leq \lceil n/2 \rceil$ where $\lceil c \rceil$ denotes the smallest integer that is equal to or bigger than c. The problem of finding function f, such that $AI(f) = \lceil n/2 \rceil$, was examined in [8, 2]. It is easy to observe that $AI(f)$ is never higher than its degree, i.e. $AI(f) \leq deg(f)$. This fact is true because $(1 \oplus f)f = 0$. In general, for any boolean function f of n variables, we have $AI(f) \leq \min\{deg(f), \lceil n/2 \rceil\}$. Very recently, Armknecht et al [1] presented a method by which the algebraic immunity of a random boolean function with n variables and degree d can be computed in $\mathcal{O}(D^2)$ steps where $D = \sum_{i=0}^{d} \binom{n}{i}$. This is an improvement on the previous best result $\mathcal{O}(D^3)$. This method is efficient for many classes of boolean functions including boolean functions of low degree. However D^2 will be as large as $\mathcal{O}(2^n)$ for random functions when d is larger than or close to $\frac{1}{2}n$.

2 Introduction to This Work

This work is composed of two parts. In the first part, we review the current definition of algebraic immunity and extend the concept. For a boolean function f, we create its algebraic complement f^c and define *extended algebraic immunity* of the function f as $AI^*(f) = \min\{AI(f), AI(f^c)\}$. We next prove that $0 \leq AI(f) - AI^*(f) \leq 1$. Since $AI(f) - AI^*(f) = 1$ holds for a large number of cases, the difference between $AI(f)$ and $AI^*(f)$ cannot be ignored in algebraic attacks. $AI^*(f)$ is applicable not only to LFSR-based stream ciphers but also to other ciphers whenever attackers can replace the original function f by f^c. In the second part, we apply the hypergraph theory to study annihilators. This new approach enables us to examine the relation among boolean functions f, $1 \oplus f$, f^c and $1 \oplus f^c$. The main tool we use here is the concept of transversals in the hypergraph theory. We can produce annihilators of a function f, and its related functions $1 \oplus f$, f^c and $1 \oplus f^c$ and obtain an upper bound on $AI^*(f)$ in

a fast and straightforward way. We also prove that the functions obtained in our approach must be annihilators, although they may not be optimal. Further we argue that the transversal number can be smaller than both $deg(f)$ and $\lceil n/2 \rceil$. This means the transversal number gives a new upper bound on $AI^*(f)$.

The rest of the paper is organised as follows. We review the definition of algebraic immunity $AI(f)$ and present the extended algebraic immunity $AI^*(f)$ in Section 3. We briefly introduce the hypergraph theory in Section 4. We describe the connection between boolean functions and hypergraphs in Section 5. In Sections 6 we show how to convert the problem of finding annihilators into the related problem of finding transversals in a hypergraph. Then in Section 7 we derive an upper-bound on $AI^*(f)$. In Section 8 we study boolean functions and their transversal numbers. In Section 9 we apply the well-known greedy algorithm in order to find annihilators for boolean functions in an efficient and straightforward way. Section 10 concludes the work. In the Appendix we elaborate how to use the greedy algorithm to obtain better annihilators.

3 Extended Algebraic Immunity

In this section we present the concept of extended algebraic immunity. Throughout the paper we are going to use the following notations. The vector space of n-tuples of elements from $GF(2)$ is denoted by $(GF(2))^n$. We write all vectors in $(GF(2))^n$ as $(0,\ldots,0,0) = \alpha_0$, $(0,\ldots,0,1) = \alpha_1$, ..., $(1,\ldots,1,1) = \alpha_{2^n-1}$, and call α_i the *binary representation* of integer i, $i = 0, 1, \ldots, 2^n - 1$. A boolean function f is a mapping from $(GF(2))^n$ to $GF(2)$ or simply, a function f on $(GF(2))^n$. The *Hamming weight* of f, denoted by $HW(f)$, is defined as $HW(f) = \#\{\alpha \in (GF(2))^n,\ f(\alpha) = 1\}$, where $\#$ denotes the cardinality of a set. We express f as $f(x) = f(x_1, \ldots, x_n)$ where $x = (x_1, \ldots, x_n) \in (GF(2))^n$. The function f can be uniquely represented by a polynomial $f(x_1, \ldots, x_n) = \bigoplus_{\alpha \in (GF(2))^n} g(a_1, \ldots, a_n) x_1^{a_1} \cdots x_n^{a_n}$ where $\alpha = (a_1, \ldots, a_n)$, and g is also a function on $(GF(2))^n$. The polynomial representation of f is called the *algebraic normal form* (ANF) of the function and each $x_1^{a_1} \cdots x_n^{a_n}$ is called a *monomial (term)* in ANF of f. The *algebraic degree*, or simply *degree*, of f, denoted by $deg(f)$, is defined as the number of variables in the longest monomial of f, i.e., $deg(f) = \max\{HW(\alpha) \mid g(\alpha) = 1,\ \alpha \in (GF(2))^n\}$.

As an example, we consider stream ciphers based on LFSRs (Linear Feedback Shift Registers [10]). A such stream cipher is composed of two parts: a single LFSR defined by a *connection function* L and a *nonlinear filter* (boolean function) f on $(GF(2))^n$, where both L and f are known. A secret vector state K is also called the *initial state*. The stream cipher generates a sequence of keystream bits b_i as follows:

$$b_i = f(L^i(K)), \quad i = 0, 1, \ldots \tag{1}$$

In a typical attack, adversaries wish to find the initial state K knowing the structure of the cipher (i.e. functions L and f) and a sequence of keystream bits b_i for some (not necessarily consecutive) clocks i. Since L is a linear transformation, $L^i(0) = 0$, $i = 0, 1, \ldots$. Therefore K must NOT be the all-zero state.

Notation 1 *Set* $\Delta(x) = (1 \oplus x_1) \cdots (1 \oplus x_n)$ *where* $x = (x_1, \ldots, x_n) \in (GF(2))^n$.

It is easy to prove the following lemma.

Lemma 1. *The function* $\Delta(x)$ *has the following properties. (i)* $\Delta(\alpha) \neq 0$ *if and only if* $\alpha = 0$*, (ii)* $h(x)\Delta(x)$ *is identical with the constant zero for any boolean function* $g \in (GF(2))^n$ *with* $h(0) = 0$*, (iii)* $h(x)\Delta(x)$ *is identical with* $\Delta(x)$ *for any boolean function* $g \in (GF(2))^n$ *with* $h(0) = 1$.

Notation 2 *Given a function* f *on* $(GF(2))^n$*. We define an algebraic complement of* f*, denoted by* f^c*, as the function that contains all monomials* $x_1^{a_1} \cdots x_n^{a_n}$*, where each* $a_j \in \{0, 1\}$*, that are not in ANF of the function* f.

The following properties of the algebraic complement are obvious: (1) $(f^c)^c = f$ for any function f; (2) any pair of functions (f, f^c) does not have any monomials in common.

Lemma 2. *Let* f *be a function on* $(GF(2))^n$*. Then (i)* $f^c(x) = \Delta(x) \oplus f(x)$ *for all* $x \in (GF(2))^n$*, (ii)* $f^c(x) = f(x)$ *for all nonzero* $x \in (GF(2))^n$.

Proof. It is easy to verify that ANF of $\Delta(x)$ contains all $2^n - 1$ possible monomials $x_1^{a_1} \cdots x_n^{a_n}$. Thus the statement (i) is true. Using Lemma 1, we can say that the statement (ii) holds.

Due to (ii) of Lemma 2, f can be replaced by f^c. This leads us to the following theorem.

Theorem 1. *Let the connection function* L *be nonsingular (i.e.* $L(\alpha) \neq L(\alpha')$ *if* $\alpha \neq \alpha'$*). Then Equation (1) is true if and only if*

$$b_i = f^c(L^i(K)), \quad i = 0, 1, \ldots \tag{2}$$

holds.

Proof. It is noted that the secret K must be nonzero. Since L is linear and nonsingular, $L^i(K) \neq 0$, $i = 0, 1, \ldots$. According to Lemma 2, we have proved the theorem. □

Note that there exists no guarantee that $AI(f)$ and $AI(f^c)$ are equal. This can be seen from a large number of evidences, for instance

Example 1. Let $f(x_1, x_2, x_3) = x_2x_3 \oplus x_2 \oplus x_3 \oplus x_1 \oplus 1$. Then its algebraic complement is $f^c(x_1, x_2, x_3) = x_1x_2x_3 \oplus x_1x_2 \oplus x_1x_3$. It is easy to check that $AI(f) = 2$ but $AI(f^c) = 1$. □

Clearly, in an algebraic attack, adversaries are going to compute both $AI(f)$ and $AI(f^c)$ and find annihilators for both functions f and f^c. Obviously, they can apply the annihilator whose degree is lowest. This is the reason why we need to revise the concept of algebraic immunity.

Definition 1. *Given a function f on $(GF(2))^n$. The* extended algebraic immunity *of f, denoted by $AI^*(f)$, is the minimum degree of nonzero boolean functions in $AN(f) \cup AN(1 \oplus f) \cup AN(f^c) \cup AN(1 \oplus f^c)$, or in other words, $AI^*(f) = \min\{AI(f), AI(f^c)\}$.*

Example 2. In Example 1, the extended algebraic immunity $AI^*(f) = 1$ but the algebraic immunity $AI(f) = 2$. □

Theorem 2. *Let f a function on $(GF(2))^n$. Then*

(i) $|AI(f) - AI(f^c)| \leq 1$,
(ii) $0 \leq AI(f) - AI^*(f) \leq 1$ *and* $0 \leq AI(f^c) - AI^*(f) \leq 1$.

Proof. Let $g \in AN(f) \cup AN(1 \oplus f)$ such that $deg(g) = AI(f)$. It is easy to see that there exists some i_0 with $1 \leq i_0 \leq n$ such that $1 \oplus x_{i_0}$ is not a factor of g, or in other words, g cannot be expressed as $g(x) = (1 \oplus x_{i_0})g'(y)$ where g' is a boolean function on $(GF(2))^{n-1}$. Hence $x_{i_0}g(x)$ is a nonzero function. Due to Lemma 1, $x_{i_0}\Delta(x)$ is identical with the constant zero. There exist two cases to be considered: $g \in AN(f)$ and $g \in AN(1 \oplus f)$. Consider the first case: $g \in AN(f)$. The function gf is identical with the constant zero. Therefore $x_{i_0}g(x)f^c(x)$ or $x_{i_0}g(x)(f(x) \oplus \Delta(x))$ is identical with the constant zero. This implies that $x_{i_0}g(x) \in AN(f^c)$ and thus $AI(f^c) \leq 1 + AI(f)$. We next consider the second case: $g \in AN(1 \oplus f)$. The function $g(1 \oplus f)$ is identical with the constant zero. Therefore $x_{i_0}g(x)(1 \oplus f^c(x))$ or $x_{i_0}g(x)(1 \oplus f(x) \oplus \Delta(x))$ is identical with the constant zero. This implies that $x_{i_0}g(x) \in AN(1 \oplus f^c)$ and thus $AI(f^c) \leq 1 + AI(f)$. We then have proved that $AI(f^c) \leq 1 + AI(f)$ in both cases. Since $(f^c)^c = f$, we know that $AI(f) \leq 1 + AI(f^c)$. $AI(f^c) \leq 1 + AI(f)$ and $AI(f) \leq 1 + AI(f^c)$ together imply that $-1 + AI(f) \leq AI(f^c) \leq 1 + AI(f)$, i.e., $|AI(f) - AI(f^c)| \leq 1$. Thus we have proved the relation (i) of the theorem. The relation (ii) is true due to (i) and the definition of $AI^*(f)$. □

Theorem 3. *Let f be a function on $(GF(2))^n$. Then $AI^*(f) = AI(f)$ if there exists some h in $AN(f) \cup AN(1 \oplus f)$ with $deg(h) = AI(f)$ and $h(0) = 0$, and, there exists some g in $AN(f^c) \cup AN(1 \oplus f^c)$ with $deg(g) = AI(f^c)$ and $g(0) = 0$.*

Proof. Let $h \in AN(f) \cup AN(1 \oplus f)$ with $deg(h) = AI(f)$ and $h(0) = 0$. Due to Lemma 1, the function $h(x)\Delta(x)$ is identical with the constant zero. Thus $h(x)f^c(x) = h(x)(f(x) \oplus \Delta(x)) = h(x)f(x)$. Similarly, $h(x)(1 \oplus f^c(x)) = h(x)(1 \oplus f(x) \oplus \Delta(x)) = h(x)(1 \oplus f(x))$. Consequently, h is either an annihilator of f^c or an annihilator of $1 \oplus f^c$ and then $AI(f^c) \leq AI(f)$. Symmetrically, $AI(f) \leq AI(f^c)$. Thus $AI(f^c) = AI(f)$ and thus $AI(f^*) = AI(f)$. □

Due to Theorem 3, $AI(f) - AI^*(f) = 0$ may hold sometimes. However the next example indicates that $AI(f) - AI^*(f) = 1$ can also hold.

Example 3. Let $\Delta(y)$ be the function on $(GF(2))^p$ defined in Notation 1 and $\beta_j \in (GF(2))^p$ be the binary representation of positive integer j, $j = 1, \ldots, 2^p - 1$. Let $q \geq 2^p - 1$ be another integer. Thus there exist $2^p - 1$ linearly independent

linear functions $\psi_1, \ldots, \psi_{2^p-1}$ on $(GF(2))^q$. Define a function on $(GF(2))^{p+q}$: $f(x) = \bigoplus_{j=1}^{2^p-1} \Delta(y \oplus \beta_j)\psi_j(z) \oplus \Pi_{i=1}^{q+p}(1 \oplus x_i)$ where $y = (x_1, \ldots, x_p)$, $z = (x_{p+1}, \ldots, x_{p+q})$ and $x = (y, z)$. Then $f^c(x) = \bigoplus_{j=1}^{2^p-1} \Delta(y \oplus \beta_j)\psi_j(z)$. It is not hard to verify that $AI(f) \geq p+1$ and $AI(f^c) \geq p$. Since $\Delta(y)\Delta(y \oplus \beta)$ is identical with the constant zero for any nonzero $\beta \in (GF(2))^p$, $\Delta(y)$ is an annihilator of f^c. Thus $AI(f^c) \leq p$. $AI(f^c) \geq p$ and $AI(f^c) \leq p$ together imply that $AI(f^c) = p$. Due to $AI(f^c) = p$, $AI(f) \geq p+1$ and Theorem 2, we have $AI(f) = p+1$. Hence we have proved that $AI(f) = p+1$ but $AI^*(f) = p$. □

Due to Example 3, $AI(f) - AI^*(f) = 1$ holds for a large number of boolean functions. Therefore the difference between $AI(f)$ and $AI^*(f)$ cannot be ignored in algebraic attacks. Observe that the extended algebraic immunity $AI^*(f)$ is not only relevant to LFSR-based stream ciphers but in general, to any ciphers whose initial states do not contain the zero vector.

4 Brief Introduction to Hypergraph

Hypergraph theory is a part of combinatorics. The word "hypergraph" was introduced in 1966. Let $X = \{x_1, \ldots, x_n\}$ be a finite set. Set $E = \{e_1, \ldots, e_m\}$, where each e_j is a subset of X. A *hypergraph*, denoted by $\aleph$, is the pair $\aleph = (X, E)$. Each x_j is called a *vertex*, $j = 1, \ldots, n$ and each e_j is called an *edge*; $j = 1, \ldots, m$. It should be noted that repeated edges are permitted. An edge $e \in E$ is called a *loop* if $\#e = 1$. The *rank* of $\aleph$ is defined as $\max\{\#e | e \in E\}$. In particular, the hypergraph $\aleph$ is called a *graph* if the rank of $\aleph$ is less or equal to 2. Graph theory was formed much earlier than hypergraph theory. Let X' be a subset of X and E' be a subset of E. If there exists some $e_j \in E'$ such that $X' \cap e_j \neq \emptyset$, where $\emptyset$ denotes the empty set, we simply say that X' and E' are *associated*. A star centered at a vertex x_j is a family of edges of $\aleph$ associated with x_j. The *degree* of vertex x_j, denoted by $\Delta_\aleph(x_j)$, is the size of the star centered at x_j. The maximum value of degrees of vertices is denoted by $\Delta(\aleph)$. Let $X' \subseteq X$, define $\aleph - X'$ as a hypergraph whose vertex set is $X - X'$ and whose edge set consists of all edges in E with all vertices in $X - X'$. A sequence $x_1 e_1 x_2 e_2 \cdots x_p e_p x_{p+1}$ is called a *path* of length p joining x_1 to x_{p+1}, where $p > 1$, all the e_j are distinct, x_j with $1 \leq j \leq p$ are distinct, and $x_j, x_{j+1} \in e_j$, $j = 1, \ldots, p$. In particular, if $x_1 = x_p$ then the path is called a *cycle* of length p. A subset of X, say S, is a *stable set* of $\aleph$, if $e_j \not\subseteq S$ for each $j = 1, \ldots, m$. The maximum cardinality of a stable set is called the *stability number* of $\aleph$ and it is denoted by $\varsigma(\aleph)$. A subset of X, say T, is a *transversal* of $\aleph$, if $T \cap e_j \neq \emptyset$ for each $j = 1, \ldots, m$. The minimum cardinality of a transversal is called the *transversal number* of $\aleph$ and it is denoted by $\tau(\aleph)$. A subset of E, say $M = \{e_{j_1}, \ldots, e_{j_q}\}$, is a *matching* of $\aleph$, if $e_{j_u} \cap e_{j_v} = \emptyset$, for $u \neq v$. The maximum number of edges in a matching is called the *matching number* of $\aleph$, denoted by $\nu(\aleph)$.

5 Relating Hypergraphs to Boolean Functions

Definition 2. *Let f be a function on $(GF(2))^n$. If the constant monomial in the ANF of f is zero (one) we say f to be* 0-CM *(*1-CM*).*

Definition 3. *Let $f(x)$ or $f(x_1, \ldots, x_n)$ be a 0-CM boolean function on $(GF(2))^n$, where $x = (x_1, \ldots, x_n)$. We now define a hypergraph $\aleph(f)$ associated with the function f as follows. The vertex set $X(f)$ of $\aleph(f)$ consists of all variables of the function f, i.e. $X(f) = \{x_1, \ldots, x_n\}$. A subset $e = \{x_{j_1}, \ldots, x_{j_s}\}$ over $X(f)$ is an edge of $\aleph(f)$ if and only if $x_{j_1} \cdots x_{j_s}$ is a monomial in ANF of f. Denote the collection of edges of $\aleph(f)$ by $E(f)$. The hypergraph $\aleph(f) = (X(f), E(f))$ is called the* hypergraph of the 0-CM boolean function f. *We define the* hypergraph of the 1-CM boolean function f *as the hypergraph of $1 \oplus f$ and use the same notation $\aleph(f) = (X(f), E(f))$.*

According to Definition 3, for any boolean function f, there uniquely exists a hypergraph $\aleph$ such that $\aleph = \aleph(f)$, but, for any hypergraph $\aleph$, there are precisely two boolean functions f and $1 \oplus f$ whose hypergraphs are identical, i.e. $\aleph = \aleph(f) = \aleph(1 \oplus f)$. Denote the stability number, the transversal number and the matching number of $\aleph(f)$ simply by $\varsigma(f)$, $\tau(f)$ and $\nu(f)$ respectively. In this way we can apply the known results in the hypergraph theory in our study of annihilators. The relation between boolean functions and hypergraphs was first introduced by Zheng et al in [20]. Note, however, that the authors of [20] used hypergraphs to examine the nonlinearity of boolean functions while in this work we use hypergraphs to study annihilators and extended algebraic immunity. It should also be noted that the relation between boolean functions and hypergraphs established in [20] contains a minor inaccuracy because 1-CM boolean functions do not correspond to any hypergraph. Note also that 0-CM and 1-CM can be united by the definition of algebraic immunity based on the scenarios S3a and S3b: let h have a lower degree than f, then h is an annihilator of f if and only if $h(1 \oplus f) = h$, while, h is an annihilator of $1 \oplus f$ if and only if $hf = h$.

6 Annihilators versus Transversals

In this section we relate transversals to annihilators.

Lemma 3. *For a given 0-CM function f on $(GF(2))^n$, let $T = \{x_{j_1}, \ldots, x_{j_t}\}$ be a subset of $X(f)$. Then the following equation holds*

$$(1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t}) \cdot f = (1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t}) \cdot f|_{x_{j_1}=0,\ldots,x_{j_t}=0}.$$

Proof. Note that $a(1 \oplus a) = 0$ holds for any $a \in GF(2)$. Let $x_{i_1} \cdots x_{i_v}$ be a monomial in ANF of f. If $\{x_{i_1}, \ldots, x_{i_v}\} \cap \{x_{j_1}, \ldots, x_{j_t}\} \neq \emptyset$ then $x_{i_1} \cdots x_{i_v} \cdot (1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t})$ turns out to be the zero boolean function. If $\{x_{i_1}, \ldots, x_{i_v}\} \cap \{x_{j_1}, \ldots, x_{j_t}\} = \emptyset$ then $x_{i_1} \cdots x_{i_v} \cdot (1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t})$ will be different from zero. Note that the monomials of f that have empty intersection with T are uniquely identified by $f|_{x_{j_1}=0,\ldots,x_{j_t}=0}$. So the result follows. □

Note that Lemma 3 is relevant to the proof of Proposition 1 of [8] or the proof of Proposition 2 of [2]. The authors of [2, 8] indicated that the algebraic immunity of a boolean function will be low if it has a sub-function of low degree. Since the authors of [2, 8] did not determine how many variables or which variables are involved in such a sub-function, their claims need more investigation.

Lemma 4. *Let f be a 0-CM function on $(GF(2))^n$. Let $T = \{x_{j_1}, \ldots, x_{j_t}\}$ be a subset of $X(f)$. Then the following statements are equivalent: (i) T is a transversal of $\aleph(f)$, (ii) $(1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t})$ is an annihilator of the function f, (iii) $f|_{x_{j_1}=0,\ldots,x_{j_t}=0}$ vanishes or is identical with the constant zero.*

This lemma establishes a relation between annihilators of f and transversals of $\aleph(f)$. Due to Lemma 4, we can introduce the following equivalence.

Definition 4. *If $T = \{x_{j_1}, \ldots, x_{j_t}\}$ is a transversal of a 0-CM boolean function f then $(1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t})$ is called the annihilator of the function f,* corresponding to *the transversal T.*

7 Upper-bound on Extended Algebraic Immunity

Theorem 4. *For any boolean function f on $(GF(2))^n$, the extended algebraic immunity of f is upper-bounded by its transversal number, i.e.,*

$$AI^*(f) \leq \min\{\tau(f), \tau(f^c)\}.$$

Proof. According to Lemma 4, $AI(f) \leq \tau(f)$ and $AI(f^c) \leq \tau(f^c)$. Then $AI^*(f) \leq \min\{\tau(f), \tau(f^c)\}$. □

In the hypergraph theory (see Section 3 of [11]), $\tau(\aleph) + \varsigma(\aleph) = n$, where $\varsigma(\aleph)$ is the stability number of $\aleph$. This equality and Theorem 4 imply that the following statement is true.

Corollary 1. *For any boolean function f on $(GF(2))^n$, the following upper bound on extended algebraic immunity holds:*

$$AI^*(f) \leq \min\{\lceil n/2 \rceil, deg(f), deg(f^c), \tau(f) = n - \varsigma(f), \tau(f^c) = n - \varsigma(f^c)\}.$$

According to Corollary 1, a large transversal number $\min\{\tau(f), \tau(f^c)\}$ is necessary for resistance against algebraic attacks. In the next section, we show a large number of boolean functions whose transversal numbers are less than both $deg(f)$ and $\lceil n/2 \rceil$. Therefore the new bound in Theorem 4 or Corollary 1 is non-trivial.

8 Boolean Functions with Low Transversal Number

Throughout this section, we discuss only f however we can do the same for f^c and then study $AI^*(f)$. We indicate that there exist a large number of boolean functions with small transversal number. It is known that the inequality $\nu(\aleph) \leq \tau(\aleph)$ holds for every hypergraph [11] where $\nu(\aleph)$ is the matching number of $\aleph$. The hypergraph $\aleph$ is said to satisfy the *König property* if $\nu(\aleph) = \tau(\aleph)$. We say that a boolean function f satisfies the König property if its hypergraph does.

Theorem 5. *Let f be a 0-CM boolean function on $(GF(2))^n$ satisfying the König property. Let M be a matching of $\aleph(f)$ such that $\#M = \nu(f)$. Let us denote $\lambda_M = \frac{1}{\nu(f)} \sum_{e \in M} \#e$. Then $AI(f) \leq \lfloor n/\lambda_M \rfloor$, where $\lfloor c \rfloor$ denotes the maximum integer less than or equal to c.*

Proof. It is noted that any two distinct $e \in M$ and $e' \in M$ are disjoint because M is a matching of $\aleph(f)$. Thus $\lambda_M \nu(\aleph) = \sum_{e \in M} \#e \leq n$. It follows that $\nu(\aleph) \leq n/\lambda_M$. Due to the König Property $\tau(f) = \mu(f)$, we know that $\tau(f) \leq n/\lambda_M$. Since $\tau(f)$ is an integer, $\tau(f) \leq \lfloor n/\lambda_M \rfloor$. We have proved the theorem. □

The following is a consequence of Theorem 5.

Corollary 2. *Let f be a 0-CM boolean function on $(GF(2))^n$ satisfying the König property. Let M be a matching of $\aleph(f)$ such that $\#M = \nu(f)$. Let $m_0 = \min\{\#e \mid e \in M\}$. Then $AI(f) \leq \lfloor n/m_0 \rfloor$.*

In Corollary 2, if $m_0 > \min\{2, n/deg(f)\}$ then $AI(f) < \min\{n/2, deg(f)\}$. Therefore the König property of a function may result in a lower algebraic immunity.

Notation 3 *Let f be a 0-CM function on $(GF(2))^n$ Let $f^{[i]}$, where $i = 1, \ldots, n$, denote the 0-CM function composed of all terms of f with degree at least i and $f_{[i]}$ denote the 0-CM function on $(GF(2))^n$ composed of all terms of f with degree at most $i-1$. Clearly $f = f^{[i]} \oplus f_{[i]}$.*

Lemma 5. *Let f be a 0-CM boolean function on $(GF(2))^n$ and $\aleph(f^{[i_0]})$ satisfy the König property for an integer i_0 with $2 \leq i_0 \leq n-1$. Then there exists a transversal $T = \{x_{j_1}, \ldots, x_{j_t}\}$ of $\aleph(f^{[i_0]})$ such that $t \leq \lfloor n/i_0 \rfloor$ and*

$$(1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t}) \cdot f = (1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t}) \cdot f_{[i_0]}|_{x_{j_1}=0,\ldots,x_{j_t}=0} \qquad (3)$$

where the degree of $(1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t}) \cdot f_{[i_0]}|_{x_{j_1}=0,\ldots,x_{j_t}=0}$ is at most $\lfloor n/i_0 \rfloor + i_0 - 1$, or, $(1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t}) \cdot f_{[i_0]}|_{x_{j_1}=0,\ldots,x_{j_t}=0}$ is identical with the constant zero.

Proof. Applying the proof of Theorem 5 to $\aleph(f^{[i_0]})$, we know that $\tau(f^{[i_0]}) \leq \lfloor n/\lambda_M \rfloor$, where λ_M is defined for $f^{[i_0]}$. Since $i_0 \leq \lambda_M$, we know that $\tau(f^{[i_0]}) \leq \lfloor n/i_0 \rfloor$. Thus there exists a transversal $T = \{x_{j_1}, \ldots, x_{j_t}\}$ of $f^{[i_0]}$ such that $\#T = t = \tau(f^{[i_0]}) \leq \lfloor n/i_0 \rfloor$. Therefore, from $f = f^{[i_0]} \oplus f_{[i_0]}$, we know that the equality (3) holds. If T is also a transversal of $f_{[i_0]}|_{x_{j_1}=0,\ldots,x_{j_t}=0}$ then $(1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t}) \cdot f_{[i_0]}|_{x_{j_1}=0,\ldots,x_{j_t}=0}$ will be identical with the constant zero. If T is not a transversal of $f_{[i_0]}|_{x_{j_1}=0,\ldots,x_{j_t}=0}$ then the degree of $(1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t}) \cdot f_{[i_0]}|_{x_{j_1}=0,\ldots,x_{j_t}=0}$ is at most $t + i_0 - 1 = \tau(f^{[i_0]}) + i_0 - 1 \leq \lfloor n/i_0 \rfloor + i_0 - 1$. We have proved the lemma. □

Corollary 3. *Let f be a 0-CM boolean function on $(GF(2))^n$. Let there be a subset X' of $X = \{x_1, \ldots, x_n\}$ such that $\aleph(f) - X'$ satisfies the König property. Let M be a matching of $\aleph(f) - X'$ such that $\#M = \nu(\aleph(f) - X')$. Denote $\lambda_M = \frac{1}{\nu(f)} \sum_{e \in M} \#e$, then $AI(f) \leq \#X' + \lfloor (n - \#X')/\lambda_M \rfloor$.*

Proof. Applying Theorem 5 to the hypergraph $\aleph(f) - X'$, we know that there exists a transversal $T' = \{x_{j_1}, \ldots, x_{j_t}\}$ of $\aleph(f) - X'$ such that $\#T' \leq \lfloor (n - \#X')/\lambda_M \rfloor$. Denote $T = X' \cup T'$. Clearly T is a transversal $\aleph(f)$ and $\#T = \#X' + \lfloor (n - \#X')/\lambda_M \rfloor$. Then the corollary holds. □

Corollary 3 shows that if a hypergraph does not satisfy the König property but its a sub-hypergraph obtained by removing some vertices does, then the transversal number τ can be small.

Example 4. In Lemma 5, if $\#i_0 \approx \sqrt{n}$ then the degree of $(1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t}) \cdot f_{[i_0]}|_{x_{j_1}=0,\ldots,x_{j_t}=0}$ is approximately $2\sqrt{n} - 1$. This shows a possible lower algebraic immunity. In fact, it is easy to verify that $AI(f) < \min\{n/2, deg(f)\}$ when $n \geq 12$ and $2\sqrt{n} - 1 < deg(f)$. It is noted that the real-valued function $\varphi(t) = n/t + t - 1$ reaches its minimum value $\varphi(\sqrt{n}) = 2\sqrt{n} - 1$. □

There exist many sufficient conditions for the König property. For example, a hypergraph $\aleph$ will satisfy the König property if it does not have a cycle of odd length [11].

9 Annihilators by Greedy Algorithm

Throughout this section, we discuss only the original function f and symmetrically we can do the same for the algebraic complement f^c. The *greedy algorithm* [11] is widely used in combinatorial optimisation. It is based on the natural principle of building up a solution from best choices that are made locally.

9.1 Annihilators of 0-CM Boolean Functions by Greedy Algorithm

Let f be a 0-CM boolean function over the set $X = \{x_1, \ldots, x_n\}$ of variables and its hypergraph $\aleph(f)$. We would like to find the transversal T of $\aleph(f)$. Below we give the description of such algorithm.

greedy algorithm (finds a transversal T of $\aleph(f)$)

1. Set $T_0 = \emptyset$.
2. For $k = 1, 2, \cdots$ do {
   - choose a vertex $x_{j_k} \in \aleph(f) - T_{k-1}$ where

$$\Delta_{\aleph(f)-T_{k-1}}(x_{j_k}) = \Delta(\aleph(f) - T_{k-1}),$$

   - set $T_k = T_{k-1} \cup \{x_{j_k}\}$,
   - if $\aleph(f) - T_k$ is empty return the transversal T_k and exit. }

Let $T = \{x_{j_1}, \ldots, x_{j_t}\}$ be a transversal obtained from the greedy algorithm. According to Lemma 4, we know that $(1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t})$ is an annihilator of f, i.e., $f \cdot (1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t})$ is identical with the constant zero.

Note that the greedy algorithm does not guarantee that the resulting transversal T is optimal. An optimal transversal should satisfy $\#T = \tau(f)$. However, we still use the greedy algorithm to obtain a "reasonable" solution, and the greedy algorithm is often used in practice. We also note that there may exist two or more resulting transversals of $\aleph(f)$ by the greedy algorithm because, for example, there may exist two or more monomials whose degrees are equal to the $deg(f)$. Using the results from [12, 19, 13], the following statement can be proved.

Theorem 6. *Let f be a 0-CM boolean function with n variables. Then for any transversal T of $\aleph(f)$ obtained by the greedy algorithm, there is an upper bound on the cardinality of T and*

$$\#T \leq \tau(f)(1 + 1/2 + \cdots + 1/deg(f)).$$

Our considerations are illustrated on a boolean function f that was used as the filter function in the LILI-128 stream cipher [18]. Although the function f in the next example was studied in [6], in this work we use it to illustrate the greedy algorithm.

Example 5. Let f be the out filter function of LILI-128 (called f_d in [18]) that is a balanced, highly nonlinear and 3rd correlation immune boolean function of degree 6 on $(GF(2))^{10}$ constructed using design criteria given in [17]. ANF of f is taken from [6]. We next list all the monomials of f. They are

$x_2, x_3, x_4, x_5, x_6x_7, x_1x_8, x_2x_8, x_1x_9, x_3x_9, x_4x_{10}, x_6x_{10}, x_3x_7x_9, x_4x_7x_9, x_6x_7x_9,$
$x_3x_8x_9, x_6x_8x_9, x_4x_7x_{10}, x_5x_7x_{10}, x_6x_7x_{10}, x_3x_8x_{10}, x_4x_8x_{10}, x_2x_9x_{10}, x_3x_9x_{10},$
$x_4x_9x_{10}, x_5x_9x_{10}, x_3x_7x_8x_{10}, x_5x_7x_8x_{10}, x_2x_7x_9x_{10}, x_4x_7x_9x_{10}, x_6x_7x_9x_{10},$
$x_1x_8x_9x_{10}, x_3x_8x_9x_{10}, x_4x_8x_9x_{10}, x_6x_8x_9x_{10}, x_4x_6x_7x_9, x_5x_6x_7x_9, x_2x_7x_8x_9,$
$x_4x_7x_8x_9, x_4x_6x_7x_9x_{10}, x_5x_6x_7x_9x_{10}, x_3x_7x_8x_9x_{10}, x_4x_7x_8x_9x_{10}, x_4x_6x_7x_8x_9,$
$x_5x_6x_7x_8x_9, x_4x_6x_7x_8x_9x_{10}, x_5x_6x_7x_8x_9x_{10}.$

We apply the greedy algorithm to $\aleph(f)$. Since $\Delta_{\aleph(f)}(x_9) = \Delta(\aleph(f)) = 30$, we set $T_1 = \{x_9\}$. We then have $\aleph(f) - T_1 = \{x_2,\ x_3,\ x_4,\ x_5,\ x_6x_7,\ x_1x_8,\ x_2x_8,\ x_4x_{10},\ x_6x_{10},\ x_4x_7x_{10},\ x_5x_7x_{10},\ x_6x_7x_{10},\ x_3x_8x_{10},\ x_4x_8x_{10},\ x_3x_7x_8x_{10},\ x_5x_7x_8x_{10}\}$. As $\Delta_{\aleph(f)-T_1}(x_{10}) = \Delta_{\aleph(f)-T_1} = 9$, we set $T_2 = T_1 \cup \{x_{10}\}$. Observe that $\aleph(f) - T_2 = \{x_2,\ x_3,\ x_4,\ x_5,\ x_6x_7,\ x_1x_8,\ x_2x_8\}$. Although we can continue the greedy algorithm until we find a transversal of f, we now stop the algorithm and multiple f by $(1 \oplus x_9)(1 \oplus x_{10})$. According to Lemma 3, we get

$$\begin{gathered}(1 \oplus x_9)(1 \oplus x_{10}) \cdot f \\ = (1 \oplus x_9)(1 \oplus x_{10}) \cdot (x_2 \oplus x_3 \oplus x_4 \oplus x_5 \oplus x_6x_7 \oplus x_1x_8 \oplus x_2x_8)\end{gathered}$$

Thus multiplying the equation $f(x_1, \ldots, x_{10}) = 1$ by $(1 \oplus x_9)(1 \oplus x_{10})$, we have

$$(1 \oplus x_9)(1 \oplus x_{10}) \cdot (1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus x_5 \oplus x_6x_7 \oplus x_1x_8 \oplus x_2x_8) = 0$$

Similarly, multiplying the equation $f(x_1, \ldots, x_{10}) = 0$ by $(1 \oplus x_9)(1 \oplus x_{10})$, we receive

$$(1 \oplus x_9)(1 \oplus x_{10}) \cdot (x_2 \oplus x_3 \oplus x_4 \oplus x_5 \oplus x_6x_7 \oplus x_1x_8 \oplus x_2x_8) = 0$$

Therefore we have reduced degree of the equations from 6 to 4. □

9.2 Greedy Algorithm on 1-CM Boolean Functions

Let f be a 1-CM function on $(GF(2))^n$. Then $1 \oplus f$ is 0-CM. We apply the greedy algorithm to $\aleph(1 \oplus f)$ and obtain a transversal $T = \{x_{j_1}, \ldots, x_{j_t}\}$. According to Lemma 4, we know that $(1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t})$ is an annihilator of $1 \oplus f$, i.e., $(1 \oplus f) \cdot (1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t})$ is identical with the constant zero, or in other words, $f \cdot (1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t}) = (1 \oplus x_{j_1}) \cdots (1 \oplus x_{j_t})$.

9.3 Complexity of the Greedy Algorithm for Annihilators

We next investigate the complexity of the greedy algorithm for an annihilator.

Theorem 7. *For any function f on $(GF(2))^n$, by using the greedy algorithm, we can obtain an annihilator of f or $1 \oplus f$ in $n(n+1)$ steps.*

Proof. For the case that f is 0-CM, we first compute $\Delta_{\aleph(f)}(x_j) = d_j$, $j = 1, \ldots, n$. Thus it takes n steps to obtain $d_1, \ldots, d_n$. Set $p_1 = d_1$. Assume we have had p_k. Set p_{k+1}=max$\{p_k, d_{k+1}\}$. We then get p_n. Clearly $p_n = \max\{d_1, \ldots, d_n\}$. Thus we only need n steps to find p_n or x_{j_0} such that $\Delta_{\aleph(f)}(x_{j_0}) = \Delta_{\aleph(f)}$. Concluding, the computation takes at most $2n$ steps on $\aleph(f)$ to find $d_1, \ldots, d_n$, and p_n. Similarly, we compute the degree of each vertex of $\Delta_{\aleph(f)-\{x_{j_0}\}}$, and then find x_{j_1} such that $\Delta_{\aleph(f)-\{x_{j_0}\}}(x_{j_1}) = \Delta_{\aleph(f)-\{x_{j_0}\}}$. The computation takes at most $2(n-1)$ steps on $\aleph(f) - \{x_{j_0}\}$. By using the greedy algorithm, we can find an annihilator of a 0-CM function f with n variables in at most $2n+2(n-1)+\cdots + \leq n(n+1)$ steps. Since we can apply the greedy algorithm to $1 \oplus f$ when f is 1-CM, we then have proved the theorem. □

According to Theorem 7, the greedy algorithm is always fast. The algorithm guarantees the resulting function must be an annihilator although it may not be best (with minimum degree). The greedy algorithm will be refined in the Appendix.

10 Conclusions

We have argued that in algebraic attacks, boolean functions f may be replaced by their algebraic complements f^c. We then have introduced the extended algebraic immunity $AI^*(f) = \min\{AI(f), AI(f^c)\}$. We prove that $0 \leq AI(f) - AI^*(f) \leq 1$. We have also indicated that $AI(f) - AI^*(f) = 1$ holds for a large number of boolean functions. Therefore the difference between $AI(f)$ and $AI^*(f)$ cannot be ignored in algebraic attacks. We have established a relation between annihilators of boolean functions and traversals of hypergraphs. The relation allows us to find annihilators in a fast and effective way provided ANF of the function is known. In addition, we establish a new upper-bound on $AI^*(f)$. The new upper-bound and the algorithms together show that the new approach is helpful in analysis of the extended algebraic immunity $AI^*(f)$ and in finding annihilators.

Acknowledgment

The first two authors were supported by Australian Research Council grants DP0345366, DP0451484 and DP0663452. We would like to thank the referees for helpful suggestions.

References

1. F. Armknecht, C. Carlet, P. Gaborit, S. Künzli, W. Meier, and O. Ruatta. Efficient computation of algebraic immunity for algebraic and fast algebraic attacks. In *Advances in Cryptology - Eurocrypt'06*, volume 4004 of *Lecture Notes in Computer Science*, pages 147–164. Springer-Verlag, Berlin, Heidelberg, New York, 2006.
2. C. Carlet, D. Dalai, K. Gupta, and S. Maitra. Algebraic immunity for cryptographically significant boolean functions: Analysis and construction. *IEEE Transactions on Information Theory*, IT-xx No. x:xxx–xxx, 2006.
3. N. Courtois. Fast algebraic attacks on stream ciphers with linear feedback. In *Advances in Cryptology - Crypto'03*, volume 2729 of *Lecture Notes in Computer Science*, pages 176–194. Springer-Verlag, Berlin, Heidelberg, New York, 2003.
4. N. Courtois. Higher order correlation attacks, XL algorithm and cryptanalysis of Toyocrypt. In *The 5th International Conference on Information Security and Cryptology (ICISC'02), Seoul, Korea*, volume 2587 of *Lecture Notes in Computer Science*, pages 182–199. Springer-Verlag, Berlin, Heidelberg, New York, 2003.
5. N. Courtois and W. Meier. Algebraic attacks on stream ciphers with linear feedback. In *Advances in Cryptology - Eurocrypt'03*, volume 2656 of *Lecture Notes in Computer Science*, pages 345–359. Springer-Verlag, Berlin, Heidelberg, New York, 2003.
6. N. Courtois and W. Meier. Algebraic attacks on stream ciphers with linear feedback. (`http://www.nicolascourtois.net/toyolili.pdf`), 2003.
7. D. Dalai, K. Gupta, and S. Maitra. Results on algebraic immunity for cryptographically significant boolean functions. In *Proceedings of Indocrypt 2004*, volume 3348 of *Lecture Notes in Computer Science*, pages 92–106. Springer-Verlag, Berlin, Heidelberg, New York, 2004.
8. D. Dalai, K. Gupta, and S. Maitra. Cryptographically significant boolean functions: Construction and analysis in term of algebraic immunity. In *Proceedings of Fast Encryption 2005*, volume 3557 of *Lecture Notes in Computer Science*, pages 98–111. Springer-Verlag, Berlin, Heidelberg, New York, 2005.
9. F. Didier and J. Tillich. Computing the algebraic immunity efficiently. In *Proceedings of Fast Encryption 2006*, volume xxxx of *Lecture Notes in Computer Science*, pages xxx–xxx. Springer-Verlag, Berlin, Heidelberg, New York, 2006.
10. S. W. Golomb. *Shift Register Sequences.* Laguna Hills, CA: Aegean Park, 1982.
11. R. L. Graham, M. Grötschel, and L. Lovász. *Handbook of Combinatorics*, volume I. Elsevier Science B. V., 1995.
12. D. S. Johnson. Approximation algorithms for combinatorial problems. *J. Comput. System. Sci.*, 9:256–298, 1974.
13. L. Lovász. On the ratio of optimal fractional and integral covers. *Discrete Math.*, 13:383–390, 1975.
14. W. Meier, E. Pasalic, and C. Carlet. Algebraic attacks and decomposition of boolean functions. In *Advances in Cryptology - Eurocrypt'04*, volume 3027 of *Lecture Notes in Computer Science*, pages 474–491. Springer-Verlag, Berlin, Heidelberg, New York, 2004.
15. A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. *Handbook of Applied Cryptography.* CRC Press, Boca Raton, London, New York, Washington, D.C., 1997.
16. Y. Nawaz, G. Gong, and K. Gupta. Upper bounds on algebraic immunity of power functions. In *Proceedings of Fast Encryption 2006*, volume xxxx of *Lecture Notes in Computer Science*, pages xxx–xxx. Springer-Verlag, Berlin, Heidelberg, New York, 2006.

17. P. Sarkar and S. Maitra. Nonlinearity bounds and constructions of resilient boolean functions. In *Advances in Cryptology - CRYPTO2000*, volume 1880 of *Lecture Notes in Computer Science*, pages 515–532. Springer-Verlag, Berlin, Heidelberg, New York, 2000.
18. L. Simpson, E. Dawson, J. Golic, and W. Millan. LILI keystream generator. In *Selected Areas in Cryptography, 7th Annual International Workshop, SAC2000*, volume 2012 of *Lecture Notes in Computer Science*, pages 248–261. Springer-Verlag, Berlin, Heidelberg, New York, 2001.
19. S. K. Stein. Two combinatorial covering theorems. *Journal of Combinatorial Theory A*, 16:391–397, 1974.
20. Y. Zheng, X. M. Zhang, and Hideki Imai. Restrictions, terms and nonlinearity of boolean functions. *Theoretical Computer Science*, 226:207–223, 1999.

Appendix: Multiple Greedy Algorithms for Annihilators

Throughout this section, we discuss only f and symmetrically we can do the same for f^c. It is noted that $1 \oplus f$ is an annihilator of f where f is any boolean function. Thus, if $\tau(f) > deg(f)$ then the greedy algorithm will fail. For this reason, in this section we strengthen the greedy algorithm in order to obtain better annihilators. By the improved algorithm, we may obtain a better annihilator of f even $\tau(f) > deg(f), \lceil n/2 \rceil$.

Let f boolean function on $(GF(2))^n$ (0-CM or 1-CM), $X = \{x_1, \ldots, x_n\}$ be the set of variables. Applying greedy algorithm, described in Section 9 to f or $1 \oplus f$, according to f is 0-CM or 1-CM, we obtain a transversal $T = \{x_{j_1}, \ldots, x_{j_t}\}$ of $\aleph(f)$ or $\aleph(1 \oplus f)$, where x_{j_1} is produced earliest in the algorithm, x_{j_2} is produced second earliest, ..., x_{j_t} is produced last. Based on the transversal T, we next present the *Multiple greedy algorithm* in a series of notations.

Notation 4 *We define a function D_β on $(GF(2))^k$, where $1 \leq k \leq r = \min\{\frac{1}{4}n, t-2\}$, as follows: $D_\beta(y) = (1 \oplus b_1 \oplus x_{j_1}) \cdots (1 \oplus b_k \oplus x_{j_k})$ where $y = (x_{j_1}, \ldots, x_{j_k})$, $\beta = (b_1, \ldots, b_k)$, $\{x_{j_1}, \ldots, x_{j_k}\} \subseteq T = \{x_{j_1}, \ldots, x_{j_t}\}$. We define $f_\beta(z) = f(x)|_{x_{j_1}=b_1,\ldots,x_{j_k}=b_k}$ where $z = (x_{i_1}, \ldots, x_{i_{n-k}})$ satisfying $\{x_{j_1}, \ldots, x_{j_k}\} \cup \{x_{i_1}, \ldots, x_{i_{n-k}}\} = \{x_1, \ldots, x_n\}$ with $i_1 < \cdots < i_{n-k}$.*

It is easy to see that

$$f(x) = \bigoplus_{\beta \in (GF(2))^k} D_\beta(y) f_\beta(z) \tag{4}$$

Due to the greedy algorithm, it should be noted that $y = (x_{j_1}, \ldots, x_{j_k})$ does not necessarily imply $j_1 < \cdots < j_k$.

Definition 5. *(4) is called* the kth greedy decomposition *of f with respect the transversal $T = \{x_{j_1}, \ldots, x_{j_t}\}$ of $\aleph(f)$. Each $f_\beta(z)$ is called a* subfunction *of f in the greedy decomposition (4).*

Notation 5 *Let k be a fixed integer with $1 \leq k \leq r$, where $r = \min\{\frac{1}{4}n, t-2\}$ and $t = \#T$. We write the kth greedy decomposition of f in the form (4). If $f_\beta(z)$ is a non-constant function, we apply the greedy algorithm to $f_\beta(z)$ (when $f_\beta(z)$ is 0-CM) or $1 \oplus f_\beta(z)$ (when $f_\beta(z)$ is 1-CM), and then we obtain the transversal $T_{k,\beta}$ of $f_\beta(z)$ or $1 \oplus f_\beta(z)$. Clearly $T_{k,\beta}$ is a subset of $\{x_{i_1}, \ldots, x_{i_{n-k}}\}$. We define an integer $\rho_{k,\beta} = \begin{cases} 0 & \text{if } f_\beta(z) \text{ is the constant one or zero} \\ \min\{deg(f_\beta(z)), \#T_{k,\beta}\} & \text{otherwise} \end{cases}$.*
We also define an integer $\rho_k = \min\{\rho_{k,\beta} | \beta \in (GF(2))^k\}$.

Notation 6 *If exists some k such that $\rho_k = 0$, we define $k^* = \min\{k | \rho_k = 0\}$. In this case, by definition, there exists some $\beta_{j^*} \in (GF(2))^{k^*}$ such that $f_{\beta_{j^*}}(z)$ is the constant zero or one. Otherwise, $\rho_k > 0$, $k = 1, \ldots, r$, we define $k^{**} + \rho_{k^{**}} = \min\{k + \rho_k | 1 \leq k \leq r\}$, where $r = \min\{\frac{1}{4}n, t-2\}$ and $t = \#T$. In this case, by definition, there exists some $\beta_{j^{**}} \in (GF(2))^{k^{**}}$ such that $\rho_{k^{**},\beta_{j^{**}}} = \rho_{k^{**}}$.*

Theorem 8. *Let f be a function on $(GF(2))^n$ (0-CM or 1-CM).*

(i) if the first case (in Notation 6) occurs then $D_{\beta_{j^*}}(y)$ is an annihilator of f or $1 \oplus f$, where $deg(D_{\beta_{j^*}}) = k^*$,
(ii) if the second case (in Notation 6) occurs then there exists a function g on $(GF(2))^{n-k^{**}}$ such that $D_{\beta_{j^{**}}}(y)g(z)$ is an annihilator of f or $1 \oplus f$, where $deg(g) = \min\{deg(f_{\beta_{j^{**}}}(z)), \#T_{k,\beta_{j^{**}}}\}$,
(iii) both annihilators in (i) and (ii) have a degree is less than or equal to $t = \#T$.

Proof. We first prove (i) of the theorem. It is noted that $D_{\beta'}(y) \cdot D_{\beta''}(y)$ is identical with the constant zero when $\beta' \neq \beta''$. Therefore, according to (4), we know that $D_{\beta_{j^*}}(y)f(x) = D_{\beta_{j^*}}(y)f_{\beta_{j^*}}(z)$. Thus, if $f_{\beta_{j^*}}(z)$ is the constant zero then $D_{\beta_{j^*}}(y)$ is an annihilator of f, and, if $f_{\beta_{j^*}}(z)$ is the constant one then $D_{\beta_{j^*}}(y)$ is an annihilator of $1 \oplus f$. We next prove (ii). Similarly to the proof of (i), we have $D_{\beta_{j^{**}}}(y)f(x) = D_{\beta_{j^{**}}}(y)f_{\beta_{j^{**}}}(z)$. When $\#T_{k^{**},\beta_{j^{**}}} < deg(f_{\beta_{j^{**}}}(z))$, there exists an annihilator g of $f_{\beta_{j^{**}}}(z)$ or $1 \oplus f_{\beta_{j^{**}}}(z)$, where the annihilator g is corresponding to (see Definition 4) the transversal $T_{k^{**},\beta_{j^{**}}}$. Therefore $D_{\beta_{j^{**}}}(y)g(z)f(x) = D_{\beta_{j^{**}}}(y)g(z)f_{\beta j^{**}}(z) = 0$ if $f_{\beta_{j^{**}}}(z)$ is 0-CM, or, $D_{\beta_{j^{**}}}(y)g(z)f(x) = D_{\beta_{j^{**}}}(y)g(z)f_{\beta j^{**}}(z) = D_{\beta_{j^{**}}}(y)g(z)$, i.e., $D_{\beta_{j^{**}}}(y)g(z)(1 \oplus f(x)) = 0$ if $f_{\beta_{j^{**}}}(z)$ is 1-CM. This proves that $D_{\beta_{j^{**}}}(y)g(z)$ is an annihilator of f or $1 \oplus f$. When $\#T_{k^{**},\beta_{j^{**}}} \geq deg(f_{\beta_{j^{**}}}(z))$, we set $g = 1 \oplus f_{\beta_{j^{**}}}(z)$. Therefore $D_{\beta_{j^{**}}}(y)g(z)f(x) = D_{\beta_{j^{**}}}(y)g(z)f_{\beta j^{**}}(z)$ that is identical with the constant zero. This proves that $D_{\beta_{j^{**}}}(y)g(z)$ is an annihilator of f. We have completed the proof of (ii). We finally prove (iii). The degree of annihilator in (i) is equal to k^*. According the the multiple greedy algorithm, $k \leq r$ where $t = \#T$. The degree of annihilator in (ii) is equal to $k^{**} + \rho_{k^{**},\beta_{j^{**}}} \leq k^{**} + \#T_{k^{**},\beta_{j^{**}}} \leq k^{**} + \#T_{k^{**},0}$. Recall that T is the transversal of $\aleph(f)$. Therefore $k^{**} + \#T_{k^{**},0} = \#T$. We have proved (iii). □

Definition 6. *We call the algorithm in this section the* multiple greedy algorithm. *To avoid confusion, we call the algorithm in Section 9 the* single greedy algorithm.

Theorem 9. *Let f be a function on $(GF(2))^n$ (0-CM or 1-CM). Let T with $\#T = t$ be a transversal of f by the Greedy Algorithm. Then an annihilator of f or $1 \oplus f$ can be computed by using the multiple greedy algorithm in $2^{r+1} \cdot n^2$ computing operations, where $r = \min\{\frac{1}{4}n, t-2\}$ and $t = \#T$.*

Proof. Due to the multiple greedy algorithm, for each k with $1 \leq k \leq r$, we do single greedy algorithm for at most 2^k functions on $(GF(2))^{n-k}$. According to Theorem 7, the computing operations is at most $\sum_{k=1}^{r} 2^k \cdot (n-k)(n-k+1) \leq n^2 \sum_{k=1}^{r} 2^k \leq n^2 \cdot 2^{r+1}$. □

The following statement is obvious.

Corollary 4. *In the multiple greedy algorithm, for any k with $1 \leq k \leq r$, where $r = \min\{\frac{1}{4}n, t-2\}$ and $t = \#T$, and any $\beta \in (GF(2))^k$, we have $AI(f) \leq k + AI(f_\beta) \leq k + \tau(f_\beta)$.*

According to Corollary 4, any degenerate subfunction is not desirable.

The main difference between the multiple and single greedy algorithms is that the multiple greedy algorithm contains many single greedy algorithms. It is noted that many subfunctions f_β are involved in the algorithm. This is helpful for the algebraic attacks because the subfunctions have less variables than original function f and some subfunctions may have a low degree or a small transversal number or satisfy the König property. Of course, The multiple greedy algorithm needs more computing times than the single greedy algorithm, but it results in better annihilators.

Note that Proposition 1 of [8] or Proposition 2 of [2] previously indicated that the algebraic immunity of a boolean function will be low if it has a subfunction of low degree. The main difference between the multiple greedy algorithm and the previous result is that the formula (4) is based on a transversal T of f, produced by the single greedy algorithm. Also the single Greedy Algorithm is further applied to each subfunction f_β in (4).

By the same reasoning, we can apply the multiple Greedy Algorithm to f^c and obtain an annihilator of f^c. Comparing the degree of the annihilator of f^c by the multiple greedy algorithm, to the degree of the annihilator of f by the same algorithm, we choose one with smaller degree between the two annihilators as the final result.