Generalized Correlation and Higher Order Nonlinearity for Probabilistic Algebraic Attacks Description

Sergiy Pometun

Institute of Physics and Technology,
National Technical University of Ukraine "KPI", Kiev
e-mail: pomu@mail.ru

Abstract. Algebraic attacks are relatively new and interesting subject in cryptanalysis. The algebraic attacks where introduced in [1], where several possible attack's scenarios where given. The big attention was paid to deterministic scenarios of those. In this paper, probabilistic scenarios are studied. Conception of conditional correlation and partial higher order nonlinearity of Boolean function where introduced (briefly definition of conditional correlation: $C(g, f \mid f = a) := \Pr(g = f \mid f = a) - \Pr(g \neq f \mid f = a)$) . It was shown, that the both types of scenarios can be seen as a one unified attack – higher order correlation attack, which uses conditional correlation. The clear criteria of vulnerability of Boolean function to both types of scenarios was given. Accordingly, the notion of the algebraic immunity was extended.

There are very vulnerable functions to probabilistic scenario. Calculations show that if a function with a very low partial higher order nonlinearity was used in the cipher like SFINKS [8], the simple attack would require only about 2^{42} operations and $32Kb$ of keystream. The question about relation between partial higher order nonlinearity and algebraic immunity remains open yet.

Keywords: cipher, algebraic attack, Boolean function, algebraic immunity, conditional correlation, partial higher order nonlinearity.

Introduction

Algebraic attacks are an attractive field for investigation. They were introduced in [1] where they were applied to the LFSRs based stream ciphers. During a few years, a number of papers were written which investigated, extended and improved algebraic attacks. These attacks were more or less efficiently applied to LFSRs based stream ciphers with memory, with several outputs, for E0 (Bluetooth) cipher, and for the block ciphers [1,2,9-16]. Paper [1] gives us in general, two types of attack's scenario: deterministic and probabilistic ones. But efforts [1,2,9-16], listed above, dealt with deterministic type only. And it really gives significant results.

On the other hand, there is another type of cryptanalytic attack – higher order correlation attack [2]. This attack rightly gives us equations, which are true only with some probability. Interesting, that paper [1], where algebraic attacks were introduced, is a logical extension of paper [2].

In this paper, the probabilistic scenarios of algebraic attacks are studied. We show that both (deterministic and probabilistic) scenarios can be shown as one unified attack – higher order correlation attack, which uses conditional correlation instead of the usual. And deterministic scenarios are the special case where approximation precision is equal to one. Certain steps in investigation of this attack (probabilistic algebraic attack) were made. We give clear criteria of vulnerability of Boolean function, show the existence of the very vulnerable functions, and give an example of such attack and estimation of its complexity.

The paper structure: The first chapter gives briefly the main idea of the article. The second chapter gives main results about probabilistic and deterministic scenarios unification. The third chapter gives proofs of correctness and benefits of conditional correlation. The fourth chapter introduces and does the same thing about higher order nonlinearity. The fifth chapter shows the existence of very vulnerable functions. The sixth chapter gives a simple probabilistic attack example and estimation of its complexity, if we use a very vulnerable function in a cipher like SFINKS. Finally, in the seventh chapter, concluding remarks are made and some important open questions are addressed.

1 The Main Idea, Briefly

The main idea of algebraic attack is to lower degree of the system of multivariate equations, which describes the functioning of the cipher and where unknowns are the bits of the key. Each equation is a

function from the certain linear combination of the key bits. The degree is lowered by multiplying each equation on the well chosen another function. For example, $f(x)=a \Rightarrow f(x)g(x)=ag(x)$. We propose to look at this idea in the following way: $f(x)=a \Rightarrow h(x)=a$. Where $h(x)$ has a low degree. Obviously, that $h(x)$ is the same as $f(x)$ at the subset of arguments, where $f(x)=a$. So conditional correlation between $h(x)$ and $f(x)$ is equal to $C(g,f \mid f=a)=1$. The functions $h(x)$, such that $C(h,f \mid f=a) \geq 1-\varepsilon$ - it are those functions, which obtained when probabilistic algebraic attack scenarios from [1] are used. Accordingly, the notion of partial r-th order nonlinearity is introduced, it is a (doubled) Hamming distance from $f(x)$ to subset of functions with degree no greater than r, but the distance is calculated only at the subset of arguments, where $f(x)=a$. For sure, it gives a much stronger version of the usual higher order correlation attack.

2 Conditional Correlation

In this section we introduce conditional correlation of Boolean functions and approximations in the terms of this correlation. We show that Boolean function annihilators are a particular case of such approximations. Also in terms of this correlation, may be for the first time, the probabilistic algebraic attack from [1] is explored.

Definition 1. *Let B_n be the set of all Boolean functions $f: GF(2)^n \to GF(2)$, $n \geq 0$. Let $1_f=\{x \in GF(2)^n \mid f(x)=1\}$ and $0_f=\{x \in GF(2)^n \mid f(x)=0\}$ are the subsets of arguments, where function $f(x)$ is equal to one or zero respectively.*

Definition 2. *Let us denote $\mid f \mid$ the number of arguments, where $f(x)$ is not equal to 0 (weight) of Boolean function $f \in B_n$. $\mid f \mid := \mid 1_f \mid$. Also let us use the notion of partial weight on the subset of arguments X, $\mid f \mid_X := \mid 1_f \cap X \mid, X \subset GF(2)^n$.*

Definition 3. *Let $f,g \in B_n$ be a Boolean functions. Correlation between f and g is a difference of probabilities of their equality and inequality* $C(f,g) := \Pr(f=g)-\Pr(f \neq g)=\frac{\mid f+g+1 \mid}{2^n}-\frac{\mid f+g \mid}{2^n}$

Definition 4. *Annihilator of Boolean function $f \in B_n$ is any function $h \in B_n$, which holds $f \cdot h \equiv 0$. Let $An(f) := \{h \in B_n \mid f \cdot h \equiv 0\}$ be the set of all annihilators of f.*

The function h is an annihilator of f then and only then, when $1_f \subset 0_h$ [16].

Definition 5. *Algebraic immunity $AI(f)$ of function $f \in B_n$ is defined as a lowest degree of all annihilators of f or $f+1$.* $AI(f) := \min\{\deg(h) \mid h \in An(f) \cup An(f+1)\}$

Where f is represented in algebraic normal form (ANF) - as a multivariate polynomial over $GF(2)$.

In general than to more low algebraic immunity of function, that it is more vulnerable to the algebraic attacks.

Let us introduce conditional correlation of Boolean functions:

Definition 6. *Let $f,g \in B_n$ be a Boolean functions, $a \in GF(2)$. We will define conditional correlation between f and g under condition $f=a$ as follows:*

$$C(g,f \mid f=a) := \Pr(g=f \mid f=a)-\Pr(g \neq f \mid f=a)=\Pr(g=a \mid f=a)-\Pr(g \neq a \mid f=a)= \\ =\frac{\mid a_g \cap a_f \mid}{\mid a_f \mid}-\frac{\mid \overline{a}_g \cap a_f \mid}{\mid a_f \mid}=\frac{\mid a_g \mid_{a_f}-\mid \overline{a}_g \mid_{a_f}}{\mid a_f \mid} \quad (1)$$

Where $\overline{a} = a + 1 \in GF(2)$, *and if* f *is balanced then* $|a_f| = 2^{n-1}$. *If* $a_f = \emptyset$, *then* $C(g, f \mid f = a) := 1, \forall g \in B_n$ *by definition.*

In fact, there are two correlations – one on the subset $0_f \subset GF(2)^n$ where f is equal to zero, and one on the subset $1_f \subset GF(2)^n$ where f is equal to one. Sometimes we will write them as $C_0(g, f)$ та $C_1(g, f)$. For example, an equality $C_0(g, f) = 1$ means that functions f and g are the same on the subset where f is zero (to be more precise both functions are zeros on this subset), or in other words an implication $f = 0 \Rightarrow g = 0$ takes place.

Lema 1. *Let* $f, h \in B_n$ *and* f *is balanced,* $a \in GF(2)$, $C_a(h, f) = 1 - \varepsilon$. *The number of mismatching points of* h *and* f *on the subset where* $f(x) = a$ *is equal to* $|h + f|_{a_f} = \varepsilon 2^{n-2}$.

Proof:

$$C_a(h, f) = \Pr(h = f \mid f = a) - \Pr(h \neq f \mid f = a) = 1 - 2\Pr(h \neq f \mid f = a) =$$
$$= 1 - 2\frac{|h + f|_{a_f}}{|a_f|} = 1 - 2\frac{|h + f|_{a_f}}{2^{n-1}} = 1 - 2^{-n+2}|h + f|_{a_f} = 1 - \varepsilon \Rightarrow$$
$$\Rightarrow |h + f|_{a_f} = \varepsilon 2^{n-2}$$

Let us have two Boolean functions $f, g \in B_n$, $a \in GF(2)$ and the correlation between them under condition $f = a$ is equal to $C_a(g, f) = 1 - \varepsilon$. We can say that the function g is an approximation of f in the terms of conditional correlation C_a with precision $1 - \varepsilon$. Approximation with precision 1 we will name precise.

Definition 7. *Let* $f \in B_n, a \in GF(2)$. *We denote by* $R_a(f, \varepsilon) := \{g \in B_n \mid C(g, f \mid f = a) \geq 1 - \varepsilon\}$ *the set of all approximations of function* f *in the terms of* C_a *with precision not less than* $1 - \varepsilon$. *Also we denote* $R(f, \varepsilon) := R_0(f, \varepsilon) \cup R_1(f, \varepsilon)$.

Let $G \subset B_n, h \in B_n$, we denote $G + g := \{g + h \mid g \in G\}$.

Proposition 1. *Let* $f \in B_n$. *There is a simple bijection between sets of annihilators* $An(f), An(f + 1) \subset B_n$ *and sets of precise approximations* $R_1(f, 0), R_0(f, 0) \subset B_n$ *of* f *in terms of conditional correlation. To be more precise* $An(f) = R_1(f, 0) + 1$, $An(f + 1) = R_0(f, 0)$.

Proof:

1. $An(f) = R_1(f, 0) + 1$

$$h \in An(f) \Leftrightarrow [f \cdot h \equiv 0] \Leftrightarrow [\forall x : f(x) = 1 \Rightarrow h(x) = 0] \Leftrightarrow$$
$$[\forall x : f(x) = 1 \Rightarrow h(x) + 1 = 1] \Leftrightarrow [\Pr(h + 1 = 1 \mid f = 1) = 1] \Leftrightarrow$$
$$[C(h + 1, f \mid f = 1) = 1] \Leftrightarrow [h + 1 \in R_1(f, 0)] \Leftrightarrow h \in R_1(f, 0) + 1$$

2. $An(f + 1) = R_0(f, 0)$ - in a similar way.

Corollary 1 $An(f + a + 1) = R_a(f, 0) + a$

Corollary 2 $\min\{\deg(h) \mid h \in An(f + a + 1)\} = \min\{\deg(h) \mid h \in R_a(f, 0)\}$

So the algebraic immunity can be represented by the precise approximations:

$$AI(f) = \min(\deg(h) \mid h \in An(f) \cup An(f + 1)\} =$$
$$= \min(\deg(h) \mid h \in R_1(f, 0) \cup R_0(f, 0)\} = \min(\deg(h) \mid h \in R(f, 0)\} \quad (2)$$

Now the main idea of lowering degree can be seen more clearly. Let h be a low degree function and $h \in R_a(f, 0)$ fore some $a \in GF(2)$. Then:

$$h \in R_a(f, 0) \Leftrightarrow [C_a(h, f) = 1] \Leftrightarrow [\Pr(h = a \mid f = a) = 1] \Leftrightarrow [f = a \Rightarrow h = a] \quad (3)$$

The equation $f(x) = a$ is substituted by $h(x) = a$, where $h \in R_a(f, 0)$.

The similar thing we can see with annihilators also:

$$f \cdot h \equiv 0 \Leftrightarrow [f = 1 \Rightarrow h = 0],\ (f+1)h \equiv 0 \Leftrightarrow [f = 0 \Rightarrow h = 0] \quad (4)$$

But our description allows compelling generalization for case of the not precise approximation – when $h \in R_a(f, \varepsilon)$, and $\varepsilon > 0$ - is a small but non-zero number. Then the following will be true:

$$h \in R_a(f,\varepsilon) \Leftrightarrow [C_a(h,f) \geq 1-\varepsilon] \Leftrightarrow [\Pr(h = a \mid f = a) - \Pr(h \neq a \mid f = a) \geq 1-\varepsilon] \Leftrightarrow$$
$$[2\Pr(h = a \mid f = a) - 1 \geq 1-\varepsilon] \Leftrightarrow [\Pr(h = a \mid f = a) \geq 1-\varepsilon/2] \Leftrightarrow \quad (5)$$
$$[f = a \Rightarrow h = a \quad , \Pr \geq 1-\varepsilon/2]$$

It turns out to be a description of probabilistic scenario(s) S4 of algebraic attack with theoretical possibility pointed out by Courtois in [1].

Let us remember this scenario generally. We need to lower the degree of equation $f(x) = a$. It is multiplied by function g. Then, we obtain equation $[f(x)g(x) = ag(x)] \Leftrightarrow [t(x) = 0]$ (which has already lower degree) is approximated with high precision by equation $h(x) = 0$ which has yet lower degree. The aim is to find $h(x)$. Although equations of type $h(x) = 0$ will only be true with some probability we have a chance for solution due to of their essentially low degree.

It is known [6] that the function t is an annihilator of $f + a + 1$. (really $[f(x)g(x) = ag(x)] \Leftrightarrow [(f+a)g = 0] \Leftrightarrow [t(x) = 0]$, but $(f+a+1)(f+a)g \equiv 0$, so $(f+a)g$ is an annihilator of $(f+a+1)$). By virtue of the fact that $An(f+a+1) + a = R_a(f, 0)$, $(f+a+1) + a \equiv t + a$ is a precise approximation of f in terms of C_a, or in other words $C(t+a, f \mid f = a) = 1$. For simplicity, we substitute functions $t + a$ and $h + a$ by t and h (in fact it means that $t(x) = a$ is approximated by $h(x) = a$). Correlation between t and h remains unchanged. Now we can formulate

Proposition 2. *Let*

1. $f, h, t \in B_n$ *and* f *is balanced*
2. $C(t, f \mid f = a) = 1$
3. $C(h,t) = 1 - \varepsilon$

Then: $h \in R_a(f, 2\varepsilon)$

To be more precise $1 - 2\varepsilon \leq C_a(h,f) \leq 1$ *and when* $\varepsilon \leq 1$ *the both equalities are can be reached.*

Proof:

From condition 3, it arises that $0 \leq |h+t|_{0_f} \leq \varepsilon 2^{n-1}$ and $0 \leq |h+t|_{1_f} \leq \varepsilon 2^{n-1}$.

(really $C(h,t) = 1 - 2P(h \neq t) = 1 - 2\frac{|h+t|}{2^n} = 1 - \varepsilon \Rightarrow |h+t| = \varepsilon 2^{n-1}$, but $|h+t| = |h+t|_{0_f} + |h+t|_{1_f}$). If f is balanced and $\varepsilon \leq 1$ then $|h+t|_{0_f}$ and $|h+t|_{1_f}$ separately can be in the range form 0 to $\varepsilon 2^{n-1}$ inclusively.

The conditional correlation between h and f can be expressed as follows:

$$C_a(h,f) = \Pr(h = f \mid f = a) - \Pr(h \neq f \mid f = a) = 1 - 2\Pr(h \neq f \mid f = a) =$$
$$= 1 - 2\frac{|h+f|_{a_f}}{|a_f|} = 1 - \frac{|h+f|_{a_f}}{2^{n-2}}$$

From condition 2: $C_a(h,f) = 1 - \frac{|h+f|_{a_f}}{2^{n-2}} = 1 - \frac{|h+t|_{a_f}}{2^{n-2}}$

Taking into account the inequality $0 \leq |h+t|_{a_f} \leq \varepsilon 2^{n-1}$ we have

$$1 - 2\varepsilon \leq 1 - \frac{|h+t|_{a_f}}{2^{n-2}} = C_a(h,f) \leq 1$$

So, the scenario S4 is reducing to finding low-degree approximations in terms of conditional correlation.

For describing the lowest degree among the functions $h \in R_a(f, 2\varepsilon)$, we can extend the notion of algebraic immunity:

$$AI(f,\varepsilon)=\min\{\deg(h)\mid h\in R_1(f,2\varepsilon)\cup R_0(f,2\varepsilon)\}=\min\{\deg(h)\mid h\in R(f,2\varepsilon)\} \quad (6)$$

This extension is correct:

$$AI(f,0)=\min\{\deg(h)\mid h\in R(f,0)\}=\min\{\deg(h)\mid h\in R_0(f,0)\cup R_1(f,0)\}=$$
$$=\min\{\deg(h)\mid h\in An(f+1)\cup An(f)\}=AI(f)$$

$AI(f,\varepsilon)$ is a minimal equation's degree, which can be obtained from $f(x)=0$ or $f(x)=1$ and will be true with probability not less than $1-\varepsilon$. When ε grows, $AI(f,\varepsilon)$ steps down

Corollary: we proved that the approximation in terms of conditional correlation describes both deterministic and probabilistic algebraic attack's scenarios simultaneously.

3 Approximations in Terms of Usual and Conditional Correlation

The important cryptographic property of Boolean functions is presence (or absence) of sufficiently close low degree approximations. Such approximations, in particular, use in the high order correlation attacks [2]. Associated with development of algebraic attacks some more works have evolved, where such approximations are investigated, for example [3], [4], [5].

It turns out that approximations in terms of conditional correlation in many cases can be used for cryptanalysis in the same way as usual approximations. They always are also not less precise (for the most part are more precise) than usual. So use of approximation in terms of conditional correlation instead of usual can improve many earlier results.

Really, let the Boolean function f is approximated by some function (of low degree or with any other property we need) g and $C(g,f)=1-\varepsilon$. From this, it follows that $\Pr(g\neq f)=\varepsilon/2$, it means that replacement of f by g in any expression will reduce to an expression that is wrong with probability no greater than $\varepsilon/2$. From these considerations are searched such approximations, that have the properties we need (for example a low degree) and as possible small value of ε. But often it makes replacement not in arbitrary expression, on frequent occasions it changes only $f(x)=a$ for $g(x)=a$ for some known $a\in GF(2)$. In this separate case we are interested in the probability $\Pr(g\neq f\mid f=a)$ instead of the general $\Pr(g\neq f)$. Here is a case when replacement of f by g at $\Pr(g\neq f)=\varepsilon/2$ can give expression which will be wrong with probability, that is slightly smaller than $\varepsilon/2$.

As $\Pr(g\neq f\mid f=a)=\dfrac{1-C_a(g,f)}{2}$ we come to approximation in terms of conditional correlation.

Remark

It is noteworthy, that

$C_a(g,f)=1-\varepsilon\Rightarrow[\Pr(g=a\mid f=a)=1-\varepsilon/2]\Rightarrow[f=a\Rightarrow g=a,\ \Pr=1-\varepsilon/2]$. The probability that equation $g(x)=a$ is wrong equals to $\varepsilon/2$, namely is the same as in general case by replacement of f by h in some expression, when $C(h,f)=1-\varepsilon$. It means that conditional correlation gives the same opportunities to writing equations as usual (except of specificity strictly speaking of replacement $f=a\Rightarrow g=a$). So we can see, that it is entered correctly, thus it makes sense to compare it with usual correlation.

Proposition 3. *Let $f,g\in B_n$, where f is balanced, then*

$$C(g,f)=\frac{C_0(g,f)+C_1(g,f)}{2}$$

Proof

$$\tfrac{1}{2}(C_0(g,f)+C_1(g,f))=\tfrac{1}{2}(2\Pr(g=f\mid f=0)-1+2\Pr(g=f\mid f=1)-1)=$$
$$=\Pr(g=f\mid f=0)+\Pr(g=f\mid f=1)-1=$$
$$=\Pr(g=f,f=0)/\Pr(f=0)+\Pr(g=f,f=1)/\Pr(f=1)-1=$$
$$=(\Pr(g=f,f=0)+\Pr(g=f,f=1))/(1/2)-1=2\Pr(g=f)-1=C(h,f)$$

Corollary $\min[C_0(g,f),C_1(g,f)]\le C(g,f)\le\max[C_0(g,f),C_1(g,f)]$

The simple example

Let $f(x) = f(x_1, x_2) = x_1 + x_2$, $g(x) = g(x_1, x_2) = x_1 x_2 + 1$, it is not difficult to see, that

$$C(g, f) = 1/2, \quad C_0(g, f) = 0, \quad C_1(g, f) = 1$$

The corollary shows that one of approximations in terms of C_0 or C_1 is always not less than it terms of usual correlation C. It gives reason for importance of functions properties research relative to these correlations. We attempt with this purpose in following section to generalize concept of higher order nonlinearity of the Boolean functions through using the entered correlations.

4 Partial Higher Order Nonlinearity

The higher order nonlinearity profile is important cryptographic property of Boolean functions [7]. Because of appearance of algebraic attacks there were works on the estimation of lower bound on the r-th order nonlinearity of function with given algebraic immunity [3], [4], [5]. In this section, the concept of the partial r-th order nonlinearity is introduced, thereby the r-th order nonlinearity is enhanced.

Let's review the exact definition of this characteristic:

Definition 8. The r-th order nonlinearity of a Boolean function $f \in B_n$ is the minimum Hamming distance from f to a class of functions with degree not more than r.

$$nl_r(f) := \min_{\deg(h) \le r} |f + h| \tag{7}$$

It is clear, that the more r, then less is nonlinearity, and, if $r = \deg(f)$ then $nl_r(f) = 0$. If function has small nonlinearity for some small value r, then it has close approximation (in terms of usual correlation) by some function of a degree not above r. This is a cryptographic weakness.

Let's try to enter concept of nonlinearity which uses conditional correlation advantages. Do this we shall prove before simple lemma.

Lemma 2. *Let $f, h \in B_n$ be Boolean functions, $C(h, f) = 1 - \varepsilon$.* Then Hamming distance between h and f is equal to $|h + f| = \varepsilon 2^{n-1}$

Proof:

$$C(h, f) = \Pr(h = f) - \Pr(h \ne f) = 1 - 2\Pr(h \ne f) =$$
$$= 1 - 2\frac{|h + f|}{2^n} = 1 - 2^{-n+1}|h + f| = 1 - \varepsilon \Rightarrow |h + f| = \varepsilon 2^{n-1}$$

Now the r-th order nonlinearity can be easily expressed in terms of correlation:

$$nl_r(f) = \min_{\deg(h) \le r} |h + f| = \min_{\deg(h) \le r} (1 - C(h, f))2^{n-1} =$$
$$= 2^{n-1} \min_{\deg(h) \le r} (1 - C(h, f)) = 2^{n-1}(1 - \max_{\deg(h) \le r} C(h, f)) \tag{8}$$

Analogously to (8) we will enter concept of partially r-order nonlinearity

Definition 9. *The partial r-th order nonlinearity of a Boolean function $f \in B_n$ we name the quantity*

$$nlp_{a,r}(f) := 2^{n-1}(1 - \max_{\deg(h) \le r} C_a(h, f)), \ a \in GF(2) \tag{9}$$

$$nlp_r(f) := \min[nlp_{0,r}(f), nlp_{1,r}(f)] \tag{10}$$

With regard to lemma 1 for every balanced f we have:

$$nlp_{a,r}(f) = 2^{n-1}(1 - \max_{\deg(h)\le r} C_a(h,f)) = 2^{n-1}(1 - \max_{\deg(h)\le r}(1 - \frac{|h+f|_{a_f}}{2^{n-2}})) =$$
$$= 2^{n-1}(\min_{\deg(h)\le r} \frac{|h+f|_{a_f}}{2^{n-2}}) = 2\min_{\deg(h)\le r} |h+f|_{a_f} \quad (11)$$

The formula (11) is similar to usual nonlinearity definition (7), but weight is considered not on all arguments just where f is equal only to zero (one). Also value of weight should be multiplied by 2 (it is clear since cardinal number of arguments has decreased twice, so mismatching on each argument becomes twice weighty).

Let's show, that partial nonlinearity gives the same opportunities for writing of the equations, as usual. That is, if $nl_r(f) = nlp_{a,r}(g)$ for the some $a \in GF(2)$, so it is possible to replace both f and g by the equations (best approximations) of degree, that is not above r and after this replacement the obtained equations will be true with the same probability. Actually it simply follows from (8) and (9). Really, definition of usual nonlinearity differs from partial nonlinearity only in using conditional correlation instead of usual in (9). Corresponding property of conditional correlation is already shown in section (3). In that section there were considered also replacement restrictions for conditional correlation.

It could be shown, that comparison of partial and usual nonlinearity is correct, in another way. Really, directly from definition of nonlinearity (7) follows, that probability of an error at the closest approximation (of a degree not above r) of function f is equal to $\frac{\min_{\deg(h)\le r} |f+h|}{2^n} = \frac{nl_r(f)}{2^n}$. The same probability at the closest approximation on subset of arguments a_f is equal to $\frac{\min_{\deg(h)\le r} |f+h|_{a_f}}{2^{n-1}} = \frac{2\min_{\deg(h)\le r} |f+h|_{a_f}}{2^n} = \frac{nlp_{a,r}(f)}{2^n}$. Now then it's reasonable to compare $nl_r(f)$ and $nlp_{a,f}$. This will be true only for balanced f.

Proposition 4. *Let $f \in B_n$ is balanced. Then for any $r \ge 0$*

$$nlp_r(f) \le nl_r(f)$$

Proof:

Follows from (8), (9) and (10) and from the fact, that $\forall g \in B_n \quad C(g,f) \le \max[C_0(g,f), C_1(g,f)]$.

Really,

$$nlp_r(f) := \min[nlp_{0,r}(f), nlp_{1,r}(f)] = \min[2^{n-1}(1 - \max_{\deg(h)\le r} C_0(h,f)), 2^{n-1}(1 - \max_{\deg(h)\le r} C_1(h,f))] =$$
$$= 2^{n-1}\min[(1 - \max_{\deg(h)\le r} C_0(h,f)), (1 - \max_{\deg(h)\le r} C_1(h,f))] = 2^{n-1}(1 - \max[\max_{\deg(h)\le r} C_0(h,f), \max_{\deg(h)\le r} C_1(h,f)] =$$
$$= 2^{n-1}(1 - \max_{\deg(h)\le r} \max[C_0(h,f), C_1(h,f)]) \le 2^{n-1}(1 - \max_{\deg(h)\le r} C(h,f)) = nl_r(f)$$

Let's gather some properties of partial nonlinearity in the table as a matter of convenience.

Table 1

Relation of Higher Order Nonlinearity, Partial Higher Order Nonlinearity, Algebraic Immunity and Degree for Balanced Boolean Function f of n arguments

r	0	1	2	...	$d = AI(f) \le \lceil \frac{n}{2} \rceil$	$d+1$	...	$k = \deg(f)$	$k+1$	...	n
$nl_r(f)$	$nl_0(f)$	$nl_1(f)$	$nl_2(f)$	...	$nl_d(f)$	$nl_{d+1}(f)$	...	$nl_k(f) = 0$	0	...	0
Comparison	II	II	VI	VI	VI	VI	VI	II	II	II	II
$nlp_r(f)$	$nlp_0(f)$	$nlp_1(f)$	$nlp_2(f)$	...	$nlp_d(f) = 0$	0	0	0	0	0	0

The notes to the Table 1:

1. Nonlinearity and partial nonlinearity decrease with increasing r (in the line of the right).

2. Definition of partial nonlinearity and formula (2) imply, that $d = AI(f)$ is the smallest number, where $nlp_d(f) = 0$
3. The fact that $nl_1(f) = nlp_1(f)$ for balanced f is not proved here (the statement 4 gives only that $nl_1(f) \geq nlp_1(f)$).

Thus partial higher order nonlinearity gives us the essential advantages over usual. The example of function with small partial and the sufficiently large usual second order nonlinearity will be given in the next section.

5 The Example of Vulnerable Function

Let us construct the function for illustration of concepts from previous chapters. We can imagine that this function is used as filtering function in some LFSR based stream cipher. Our design criteria are: not a very vulnerable against deterministic algebraic attack, not a very vulnerable against higher order correlation attack and a very vulnerable against probabilistic algebraic attack. The only additional criterion is balancing. (to go in advance $f(x)$ will have the six arguments, $AI(f) = 3$, $nl_2(f) = 12$, $nlp_2(f) = 2$).

Point out directly a subclass of functions with very low partial second order nonlinearity:

$$f(x) = x_1x_2 + x_1x_2...x_n + g(x)(x_1x_2 + 1) \qquad (12)$$

Proposition 5. *Let* $f \in B_n$ *have a form* $f(x) = x_1x_2 + x_1x_2...x_n + g(x)(x_1x_2 + 1)$, where $g \in B_n$ *- such function that makes* f *balanced. Then*

$$C_0(h, f) = 1 - \frac{1}{2^{n-2}}, \text{ where } h(x) = x_1x_2$$

Proof (uses Bayesian formula):

$$C_0(h, f) = 1 - 2\Pr(h \neq f \mid f = 0) =$$

$$1 - 2\frac{\Pr(h = 1, f = 0)}{\Pr(f = 0)} =$$

$$1 - 2\frac{\Pr(f = 0 \mid h = 1) \cdot \Pr(h = 1)}{\Pr(f = 0)} =$$

$$1 - 2\frac{\Pr(x_1x_2 + x_1x_2...x_n + g(x)(x_1x_2 + 1) = 0 \mid x_1x_2 = 1) \cdot \Pr(x_1x_2 = 1)}{1/2} =$$

$$1 - 4\Pr(x_1x_2 + x_1x_2...x_n + g(x)(x_1x_2 + 1) = 0 \mid x_1x_2 = 1) \cdot (1/4) =$$

$$1 - \Pr(1 + x_3x_4...x_n + g(x)(1 + 1) = 0 \mid x_1x_2 = 1) =$$

$$1 - \Pr(1 + x_3x_4...x_n = 0 \mid x_1x_2 = 1) =$$

$$1 - \Pr(x_3x_4...x_n = 1) = 1 - \frac{1}{2^{n-2}}$$

Corollary: $nlp_2(f) \leq 2$

(really $nlp_2(f) \leq nlp_{0,2}(f) = 2^{n-1}(1 - \max_{\deg(h) \leq 2} C_0(h, f)) \leq 2^{n-1} \cdot \frac{1}{2^{n-2}} = 2$)

Remark There is only the one point where $f(x) = 0$ and $h(x) = 1$, $x = (x_1, x_2, ..., x_n) = (1,1,...,1)$.

All that remains to be done – is to single out a function from this subclass with needed properties. We did not do this theoretically but by computer program for $n = 6$.

Table 2

Relation of Higher Order Nonlinearity, Partial Higher Order Nonlinearity, Algebraic Immunity and Degree for Some Balanced Boolean Function f of 6 Arguments

r	0	1	2	$3 = AI(f)$	4	$5 = \deg(f)$	6
$nl_r(f)$	32	18	12	6	2	0	0
$nlp_r(f)$	32	18	2	0	0	0	0

Where $$f(x) = f(x_1, x_2, x_3, x_4, x_5, x_6) = \begin{bmatrix} 0011001101111001 \\ 0101001110010001 \\ 0101100100110001 \\ 0001111100111010 \end{bmatrix}$$

In fact, there are many more functions with so low partial second order nonlinearity ($nlp_2(f) = 2$) than in subclass (12).

6 The Simple Probabilistic Attack Description and Calculation

Let us describe the simple probabilistic algebraic attack and make estimation of it complexity if the filtering function with partial nonlinearity $nlp_{0,2}(f) = 2$ is used. Let us have a LFSR of length n bits and the filtering function f uses k of them. Simplified functioning of our cipher can be written as follows:

$$\begin{cases} f(Px) = b_0 \\ f(PLx) = b_1 \\ f(PL^2x) = b_2 \\ \ldots\ldots\ldots\ldots \\ f(PL^Nx) = b_N \end{cases} \quad (13)$$

Where b_i - (know) keystream bits, $x = (x_1, x_2, ..., x_n)$ - unknown bits of the key, L - linear operator which describes LFSR functioning, P - projection operator which takes k arguments from n, f - our filtering function.

We will attack this system in such a manner. As $nlp_{0,2}(f) = 2$ then there is such function $h \in B_k$, $\deg(h) = 2$, that $C_0(h, f) = 1 - \frac{1}{2^{k-2}}$. Let us choose from system (13) t equations which have a right part $b_{i_r} = 0$. Do a replacement: $f(PL^{i_r}) = 0 \Rightarrow h(PL^{i_r}) = 0$. As a result we have a new system:

$$h(PL^{i_r}x) = b_{i_r}, \; r = 1, ..., t \quad (14)$$

System (14) – is a system of quadratic multivariate equations and each of them can be wrong with probability $\frac{1}{2^{k-1}}$. If there where no wrong equations it would enough about $\frac{(n+1)^2}{2}$ true equations for solution by simple linearization method. We can estimate the number of wrong equations among $\frac{(n+1)^2}{2}$ as $D = \frac{(n+1)^2}{2} \cdot \frac{1}{2^{k-1}}$.

It turns out this number can be quite small, even smaller than one.

Let us calculate D for the cipher SFINKS (if it had filtering function with so small partial second order nonlinearity). According to specification [8] SFINKS has LFSR with length $n = 256$ bits,

and its filtering function uses $k=17$ of them. So among $\frac{(n+1)^2}{2}$ equations $D=\frac{(n+1)^2}{2}\cdot\frac{1}{2^{k-1}}=\frac{257^2}{2}\cdot\frac{1}{2^{16}}\approx\frac{1}{2}$ will be wrong on average. The computational complexity of solution of such system will not be greater than complexity in the case when all equations are guaranteed true.

So the complexity of solution is roughly equal to $C=(\frac{(n+1)^2}{2})^{\log_2 7}\approx\frac{n^{5.6}}{7}$ operations. If $n=256$ it makes $C=2^{42}$ operations with only about $N=\frac{(n+1)^2}{2}\approx 32Kb$ of keystream. The best known attack on SFINKS (with its real filtering function) is algebraic attack and requires about 2^{71} operations and 2^{49} bits of keystream [9].

Hence, if the filtering function has very small partial higher (at least second) order nonlinearity, then efficient (in sense of computational complexity and amount of data) probabilistic algebraic attack is possible.

7 Remarks and Open Questions

Remark 1

We have proved in proposition 2 from chapter two that scenario S4 of algebraic attack from [1] is reduced to approximation of filtering/combining function by low-degree approximations in terms of conditional correlation. We can see that such description in terms of approximation, in contrast to S4, is not redundant. Really, according to S4 equation $f(x)=a$ can be multiplied by a lot of different functions g, but then resulting equations of type $fg+ag=0$ may be approximated all by the same equation $h=0$. So there are different ways to obtain the same result. This fact can be considered as redundancy of S4. In our description we are interested directly by h.

Remark 2

Let f be a balanced function and h - such function that $C_a(h,f)>0$. We replace $f(x)=a$ by $h(x)=a$ and try to solve it by reason of $\Pr(h=a)>1/2$ due to positive conditional correlation between h and f. But it's not always possible in view that h can be a non-balanced function, and we can have even the situation when $\Pr(h=a)>1/2$. Fox example, let $h(x)=x_1x_2$ and $C_0(h,f)=1/2$. Then we have $\Pr(h=0\,|\,f=0)=3/4$. But equations of type $h(x)=0$ (which will be true with probability 0.75) give us no information by reason of $\Pr(h=0)=3/4$ just because $h(x)$ is not balanced. So the correlation $C_a(h,f)$ should be big enough to "outweigh" of non-equiprobability of h. The question about amount of information which can we get from equation $h(x)=a$ with non-balanced h needs further research. The vulnerable functions from chapter 5 are obviously not a subject of this problem so far as $P(h=0)=3/4$ and $C_0(h,f)\to 1$ quickly while number of arguments $n\to\infty$.

Remark 3

When we searched for vulnerable function, (for chapter 5) we observed an interesting fact: if partial second order nonlinearity of the function $f\in B_n$ is equal to $nlp_2(f)=2$ then its algebraic immunity $AI(f)\le 3$ (at least we couldn't find function with $AI(f)=4$ for $n=8$). Of course, it should exist the relation between algebraic immunity and the minimal possible partial higher order nonlinearity. At least similar relation for usual higher order nonlinearity was established (for example in [3], [4]). Maybe the high algebraic immunity will be a sufficient condition to prevent efficient probabilistic algebraic attack. This question is still open.

But for the functions with very low second order nonlinearity, probabilistic algebraic attack is more efficient than deterministic, in spite of low algebraic immunity. An example is an attack from section 6.

References

[1]. Nicolas Courtois and Willi Meier: *Algebraic Attacks on Stream Ciphers with Linear Feedback,* Eurocrypt 2003, Warsaw, Poland, LNCS 2656, pp. 345–359, Springer. An extended version is available at http://www.minrank.org/toyolili.pdf

[2]. Nicolas T. Courtois: *Higher Order Correlation Attacks, XL Algorithm and Cryptanalysis of Toyocrypt,* http://eprint.iacr.org/2002/087

[3]. Kolokotronis N., Limniotis K., Kalouptsidis N.: Best Quadratic Approximations of Cubic Boolean Functions, http://eprint.iacr.org/2007/037

[4]. Carlet C.: A lower bound on the higher order nonlinearity of algebraic immune functions, http://eprint.iacr.org/2005/469

[5]. Mesnager S.: Improving the lower bound on the higher order nonlinearity of boolean functions with prescribed algebraic immunity, http://eprint.iacr.org/2007/117

[6]. Willi Meier, Enes Pasalic, Claude Carlet: Algebraic attacks and decomposition of Boolean functions, Proceeding of Eurocrypt 2004, LNCS 3027, pp. 474-491, Springer, 2004.

[7]. 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.

[8]. An Braeken, Joseph Lano, Nele Mentens, Bart Preneel and Ingrid Verbauwhede, Sfinks specification and source code, April 2005, Available on ECRYPT Stream Cipher Project page, http://www.ecrypt.eu.org/stream/sfinks.html

[9]. Nicolas T. Courtois, Cryptanalysis of Sfinks, http://eprint.iacr.org/2005/243

[10]. Nicolas Courtois, Josef Pieprzyk: Cryptanalysis of block ciphers with over defined systems of equations, Proceedings of Asiacrypt 2002, LNCS 2501, pp. 267-287, Springer, 2002.

[11]. Frederik Armknecht, Matthias Krause: Algebraic attacks on Combiners with Memory, Proceedings of Crypto 2003, LNCS 2729, pp. 162-176, Springer, 2003.

[12]. Nicolas Courtois: Algebraic Attacks on Combiners with Memory and Several Outputs, Cryptology ePrint Archive, Report 2003/125, 2003. http://eprint.iacr.org/2003/125

[13]. Nicolas Courtois: Fast Algebraic Attacks on Stream Ciphers with Linear Feedback, Proceedings of Crypto 2003, LNCS 2729, pp. 177-194, Springer, 2003.

[14]. Frederik Armknecht: A Linearization Attack on the Bluetooth Key Stream Generator, Available on http://eprint.iacr.org/2002/191/ 13 December 2002

[15]. Nicolas Courtois: How Fast can be Algebraic Attacks on Block Ciphers? http://eprint.iacr.org/2006/168

[16]. Frederik Armknecht: On the Existence of low-degree Equations for Algebraic Attacks, http://eprint.iacr.org/2004/185/