On Boolean Ideals and Varieties with Application to Algebraic Attacks

Alexander Rostovtsev[1] Alexey Mizyukin[2]
Department of Information security of computer systems,
St. Petersburg State Polytechnic University, Russia
[1]alexander.rostovtsev@ibks.ftk.spbstu.ru [2]mizyukin@gmail.com

March 22, 2012

Abstract

Finding the key of symmetric cipher takes computing common zero of polynomials, which define ideal and corresponding variety, usually considered over algebraically closed field. The solution is the point of the variety over prime field; it is defined by a sum of the polynomial ideal and the field ideal that defines prime field. Some authors use partitioning of this sum and reducing syzygies of polynomial ideal modulo field ideal. We generalize this method and consider polynomial ideal as a sum of two ideals, one of them is given by short polynomials, and add this ideal to the field ideal. Syzygies are reduced modulo this sum of ideals. Accuracy of definition of the substitution ideal by short polynomials can be increased using affine equivalence of ideals. This method decreases degree and length of syzygies and reduces complexity of Groebner basis computation.

1 Introduction

Symmetric cipher, one-way hash function is described by a set of Boolean functions or polynomials in normal algebraic form (NAF). Any set of polynomials forms an ideal. Zeroes of an ideal (usually considered over algebraically closed field) form a variety.

Problem of computing the key of a cipher for some known plaintext/ciphertext (and problem of hash-function inverting) takes solving polynomial equations. If the cipher has some rounds, then variables are formed from key bits and intermediate text bits. Those polynomial equations form the ideal $\mathfrak{A}$, and solving system of equations means finding a point of variety of that ideal. Usually variety defined by the polynomial ideal has very large number of points over algebraically closed field. Hence we need to find a point of the variety over prime field. For this purpose we join field polynomials, that have zero for all points over prime field, to the ideal $\mathfrak{A}$.

There are some methods for solving systems of polynomial equations: Courtois, Klimov, Patarin and Shamir [5], Courtois and Pieprzik [6], Faugere methods [9], method based on resultants [17], Wu's characteristic set method [15], Raddum and Semaev agreeing-gluing method [13].

These methods are called algebraic attacks and are similar. Their common property is that initial data and final data are simple, but intermediate data is complex and takes exponential memory any hence exponential time.

Since NAF polynomial has degree at most 1 for each variable, it can be written as $f = f_0 + f_1 x$, where f_0, f_1 do not depend on given variable x. If $g = g_0 + g_1 x$, then variable x can be eliminated if the first equation is multiplied by g_1 and the second one is multiplied by $f_1 : fg_1 + gf_1 = f_0 g_1 + f_1 g_0$. Hence if $f = 0, g = 0$, then the determinant is zero: $\begin{vmatrix} f_0 & f_1 \\ g_0 & g_1 \end{vmatrix} = 0$. Any NAF polynomial is a zero divisor, so the multiplication of polynomial by zero divisor can give false solutions. Notice that if two polynomials correspond to unique x, then elimination of x is correct. The multiplication of polynomials takes the multiplication of monomials, and hence that elimination method is reduced to Groebner basis and XL algorithms, which are equivalent [16]. We consider the process of computing common zeroes of polynomials as Groebner basis computing, but proposed approach can also be used in the characteristic set method.

Modern ciphers are constructed by applying the XOR operation to text and key, substitutions, which are defined by non-linear polynomials, and linear diffusion maps (XSL-ciphers). This follows that it is hard to define a metric that shows how tested key is close to the required one. In statistical attacks this metric is defined at average, for large number of plaintexts and corresponding ciphertexts [1, 3, 11].

The complexity of solving Boolean equation system mainly depends on non-linear equations. Defining the ideal of substitution by polynomials of small degree increases complexity of solving [5, 6, 9]. Besides of that [6] proposed to separate the ideal in two parts. First ideal consists of NAF polynomials that define the cipher. Second ideal consists of field polynomials. Syzygies of polynomials of the first ideal are reduced modulo the second ideal.

We propose two methods for solving a system of Boolean equations. The first method generalizes known approach and adds short polynomials (monomials, binomials and trinomials, etc.) from the first ideal to the second ideal. Since reduction modulo short polynomials is easy (if monomial of the second ideal divides monomial of the syzygy, it can be deleted), it can be fast executed by a specialized logical chip.

The second method defines the substitution ideal by short polynomials (exactly or approximately) and Groebner basis or other known method is applied to the ideal, defined by these polynomials. Since syzygy of polynomial and binomial has the same length as the polynomial, the length of syzygy of polynomial and trinomial increases at most 1, complexity of the algorithm can be reduced comparatively to the original methods.

Accuracy of defining the ideal of substitution by short polynomials can be increased using the affine equivalence of ideals.

2 Polynomial ideals and varieties

In this technical report we denote: $+$ as addition of ring elements, $\oplus$ as addition of ideals, $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{P}, \mathfrak{I}$ as ideals[1], $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{x}, \mathbf{y}$ as vectors, $\mathbb{F}_2$ as field of two elements.

Let K be a field, $K[x_1, \ldots, x_n]$ — ring of polynomials, $\mathbb{A}^n(K)$ — affine space. Any ideal $\mathfrak{A} \subseteq K[x_1, \ldots, x_n]$ defines algebraic set $V(\mathfrak{A})$ — set of points $P \in \mathbb{A}^n(K)$, in which $\mathfrak{A} = 0$. Since $\mathfrak{A}, \mathfrak{A}^2, \mathfrak{A}^3, \ldots$ define the same variety but $\mathfrak{A} \supset \mathfrak{A}^2 \supset \mathfrak{A}^3 \supset \ldots$, there exists the largest ideal (radical) for given variety. Farther we consider radical ideals. The variety of the maximal ideal consists of one point. The intersection (product) of ideals corresponds to the union of corresponding varieties, the sum of ideals corresponds to the intersection of corresponding varieties. Any ideal of $K[x_1, \ldots, x_n]$ is finitely generated.

For the radical ideal $\mathfrak{A} \subseteq K[x_1, \ldots, x_n]$ with corresponding variety $V(\mathfrak{A})$ there exists affine coordinate ring $K[x_1, \ldots, x_n]/\mathfrak{A}$, its elements are polynomials (regular functions) on $V(\mathfrak{A})$. If the ring $K[x_1, \ldots, x_n]/\mathfrak{P}$ has no zero divisors, the ideal $\mathfrak{P}$ is prime. Maximal ideals are prime, but inverse is not true. Any ideal of $K[x_1, \ldots, x_n]$ is uniquely presented by the intersection of degrees of prime ideals [2].

A prime ideal can contain other prime ideal. For example in the ring $K[x, y, z]$ we have $(x) \subset (x, y)$, where (x) and (x, y) are both prime ideals. The maximal length of ascending chain of prime ideals of the ring is its Krull dimension.

The ring $K[x_1, \ldots, x_n]$ is Noetherian because ascending chains of its ideals are finite. If all descending chains of ideals in the ring are finite, this ring is Artinian (and Noetherian), its Krull dimension is zero [2].

The set of ideals is closed under binary operation: intersection $\mathfrak{A} \cap \mathfrak{B}$ (it is the intersection of sets of ideal polynomials), product $\mathfrak{A}\mathfrak{B}$ (it is generated by polynomials $fg, f \in \mathfrak{A}, g \in \mathfrak{B}$), and sum $(\mathfrak{A}, \mathfrak{B}) = \mathfrak{A} \oplus \mathfrak{B}$ (it is the smallest ideal that contains $\mathfrak{A}$ and $\mathfrak{B}$). If $\mathfrak{A} \supseteq \mathfrak{B}$ (or equivalently $V(\mathfrak{B}) \supseteq V(\mathfrak{A})$), then $\mathfrak{A}$ divides $\mathfrak{B}$ and the congruence $\mathfrak{B} \equiv 0 \pmod{\mathfrak{A}}$ holds. So $\mathfrak{A} \equiv 0 \pmod{\mathfrak{A} \oplus \mathfrak{B}}$. If the ideal $\mathfrak{A}$ is irreducible, the set $V(\mathfrak{A})$ is called a variety.

In the ring $K[x_1, \ldots, x_n]$ the division of polynomial by ideal is defined non-uniquely, its result depends on a sequence of division the polynomial by elements of the ideal basis. The division is defined correctly if the ideal is given by Groebner basis.

Unique division is obvious for monomials, it is defined by an ordering of monomials by their multidegree: $x_1^{c_1} x_2^{c_2} \ldots x_n^{c_n}$ is divisible by $x_1^{d_1} x_2^{d_2} \ldots x_n^{d_n}$ if $c_1 \geq d_1,\ c_2 \geq d_2, \ldots,\ c_n \geq d_n$. The polynomial f is divisible by the monomial m, if any monomial of f is divisible by m.

Groebner basis computation depends on ordering of variables [7]. Let $\mathrm{LT}(f)$, $\mathrm{LT}(g)$ be the leading terms of polynomials f, g respectively, and $\mathrm{LCM}(f, g)$ — least common multiple of the leading monomials. For computing Groebner basis Buchberger's algorithm computes syzygy

$$S(f, g) = \frac{\mathrm{LCM}(f, g)}{\mathrm{LT}(f)} f - \frac{\mathrm{LCM}(f, g)}{\mathrm{LT}(g)} g$$

[1] Euclid Fractur font.

for all pairs f, g of the ideal basis and joins $S(f,g)$ to the basis. The leading terms of f, g become zero in the syzygy. Since number of monomials in the ideal basis is finite, this process will terminate after finite number of steps. The redundant polynomials of the basis can be deleted. If syzygy of any two polynomials of the basis is divided by some polynomial of the basis, it is Groebner basis. The different orderings of variables lead to the different residues of polynomial modulo ideal, but if the division is exact, the zero residue is obtained for any ordering.

Groebner basis is used for solving systems of polynomial equations in the ring $K[x_1, \ldots, x_n]$, because any set of polynomials defines some ideal $\mathfrak{A}$ and corresponding variety $V(\mathfrak{A})$. Hence solving the system of polynomial equations is equivalent to finding $V(\mathfrak{A})$ (or some its point). If the field K is algebraically closed, $V(\mathfrak{A})$ usually has infinite (or finite but very large) number of points. If wanted solution is in the finite subfield $K_1 \subset K$, the ideal $\mathfrak{A}$ is to be changed by $\mathfrak{A} \oplus \mathfrak{F}$, where the ideal $\mathfrak{F}$ gives field K_1. If polynomials of the ideal $\mathfrak{A} \oplus \mathfrak{F}$ have unique zero $(e_1, \ldots, e_n), e_i \in K_1$, then Groebner basis of the ideal $\mathfrak{A} \oplus \mathfrak{F}$ is $(x_1 - e_1, \ldots, x_n - e_n)$.

Define the length of a polynomial as number of its terms. Some ideals of $K[x_1, \ldots, x_n]$ have a basis (not necessary Groebner) of short polynomials (monomials, binomials, etc). The monomial ideal is generated by monomials, the binomial ideal — by binomials, the trinomial ideal — by trinomials. Notice that the problem of recognition whether the ideal is binomial yet is not solved [8]. The problem of recognizing the trinomial ideal seems to be hard too.

3 Boolean rings, their ideals and varieties

Boolean ring consists of idempotent elements, which satisfy the equality $a^2 = a$. Boolean ring has characteristic 2 due to the equalities $a + a = (a + a)^2 = a^2 + 2a + a^2 = a + 2a + a$, hence $2a = 0$. This ring is commutative due to the equalities $(a + b) = (a + b)^2 = a^2 + ab + ba + b^2 = a + b + ab + ba$, so $ab = ba$.

The elements of finite Boolean ring are Boolean functions or their quotients[2]. Any Boolean function f of n variables $\mathbf{x} = (x_1, \ldots, x_n)$ is defined by its 2^n-dimension vector $\mathbf{f}$ of values for n-tuples $((0, \ldots, 0), (0, \ldots, 0, 1), (0, \ldots, 0, 1, 0), \ldots, (1, \ldots, 1))$. Besides of that Boolean function can be given by the polynomial in normal algebraic form (NAF):

$$f = a_0 + a_n x_n + a_{n-1} x_{n-1} + a_{n,n-1} x_n x_{n-1} + a_{n-2} x_{n-2} + \ldots + a_{n,n-1,\ldots,1} x_n x_{n-1} \ldots x_1,$$

i.e. by the vector of its coefficients $\mathbf{a} = (a_0, a_n, a_{n-1}, a_{n,n-1}, a_{n-2}, \ldots, a_{n,n-1,\ldots,1})$ of length 2^n.

NAF polynomials form the ring $\mathfrak{G}_n[\mathbf{x}], \mathbf{x} = (x_1, \ldots, x_n)$ of 2^{2^n} elements, it is defined as the quotient ring $\mathfrak{G}_n[\mathbf{x}] = \mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 + x_1, \ldots, x_n^2 + x_n)$, where $\mathbb{F}_2 = \{0, 1\}$. Since

[2]For example, the ring of subsets of 3-element set with symmetric difference and intersection operations has 8 elements, it is isomorphic to the quotient ring of NAF polynomials with variables x, y modulo ideal (xy).

$f^2 + f = f(f+1) = 0$, any non-constant element of $\mathfrak{G}_n[\mathbf{x}]$ divides zero, hence the ideal (0) is not prime. The ring $\mathfrak{G}_n[\mathbf{x}]$ has unique invertible element — the constant 1.

Since the ring $\mathfrak{G}_n[\mathbf{x}]$ is finite, it is Artinian and has next properties [2]:

1. Product of ideals coincides with their intersection.
2. $\mathfrak{G}_n[\mathbf{x}]$ has dimension 0.
3. Prime ideal is maximal.
4. Any ideal is radical.
5. $\mathfrak{G}_n[\mathbf{x}]$ possesses unique factorization: the ideal $\mathfrak{A}$ is a product of different prime ideals. Each of them corresponds to a point of $V(\mathfrak{A})$ and hence is binomial.
6. Product of all prime ideals is (0).

We will call the set of zeroes of the ideal of the ring $\mathfrak{G}_n[\mathbf{x}]$ as a variety, because, as it is shown below, we can define the division in different ways.

The prime ideal that corresponds to the point $(e_1, \ldots, e_n)$, can be given by the polynomial $1 + \prod_{i=1}^{n}(x_i + e_i + 1)$. Hence any ideal of the ring $\mathfrak{G}_n[\mathbf{x}]$ can be given by one polynomial and back, any polynomial defines the principal ideal. It is a bijective correspondence. Indeed, if we assume that two different polynomials define the same ideal, then non-zero sum of polynomials defines zero ideal, it is impossible.

Unique factorization allows defining exact division of ideals and polynomials. The ideal $\mathfrak{B}$ is divided by the ideal $\mathfrak{A}$, if $V(\mathfrak{B}) \supseteq V(\mathfrak{A})$. Similarly we can define the division of polynomials, that defines principal ideals. Hence it is not necessary to use Groebner basis for dividing polynomials. If $\mathfrak{A} = \prod_{i\in I}\mathfrak{P}_i$, $\mathfrak{B} = \prod_{i\in J}\mathfrak{P}_i$ — prime factorization, then $\mathrm{GCD}(\mathfrak{A}, \mathfrak{B}) = \prod_{i\in I\cap J}\mathfrak{P}_i$, $\mathrm{LCM}(\mathfrak{A}, \mathfrak{B}) = \prod_{i\in I\cup J}\mathfrak{P}_i$.

The sum of ideals $(\mathfrak{A}, \mathfrak{B}) = (\mathfrak{A} \oplus \mathfrak{B})$ — ideal whose variety satisfies the equation $V(\mathfrak{A} \oplus \mathfrak{B}) = V(\mathfrak{A}) \cap V(\mathfrak{B})$. For principal ideals we obtain $(f) \oplus (g) = (f + g + fg)$.

Almost any ideal can be represented as a sum of different ideals similarly to its product. Additively irreducible ideals cannot be represented as a sum of two different ideals. They are analogs of prime ideals for multiplication. If $\mathfrak{P}$ is a prime ideal, then its complement $1 + \mathfrak{P}$ is additively irreducible ideal.

The ring $\mathfrak{G}_n[\mathbf{x}]$ is isomorphic to the ring $\mathfrak{V}(2^n)$ of 2^n-dimension binary vectors with operations coordinate-wise addition and multiplication. Zero and unit elements are vectors that consist of zeroes and units correspondingly. From interpolating Lagrange formula we can find the vector of coefficients $\mathbf{a}$ of NAF polynomial for known vector $\mathbf{f} \in \mathfrak{V}(2^n)$ of values of Boolean function in points $((0,\ldots,0),(0,\ldots,0,1),(0,\ldots,0,1,1),\ldots,(1,\ldots,1))$: $\mathbf{a} = L_n\mathbf{f}$, where $L_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $L_{i+1} = \begin{pmatrix} L_i & 0 \\ L_i & L_i \end{pmatrix}$. Since $L_n = L_n^{-1}, \mathbf{f} = L_n\mathbf{a}$. Since both $\mathbf{f}$ and $\mathbf{a}$ are binary vectors of the same size, the ring $\mathfrak{V}(2^n)$ (and hence $\mathfrak{G}_n[\mathbf{x}]$) admits two different multiplications: bit-wise multiplication and convolution, which corresponds to

the multiplication of polynomials. Rings defined by these two multiplications are obviously isomorphic[3].

The ring $\mathfrak{G}_n[\mathbf{x}]$ has two exact division operations (without remainder) for multiplication of polynomials. The first one is usual division of polynomials in infinite integral ring $\mathbb{F}_2[x_1, \ldots, x_n]$: if $f = gh$, then $\deg(f) = \deg(g) + \deg(h)$, and the quotient $h = \frac{f}{g}$ is uniquely defined. The second one (division in sense of algebraic geometry) is defined by varieties of corresponding polynomials: $f = gh$, if $V(f) = V(g) \cup V(h)$, and h for given f, g is not uniquely determined. The second division generalizes the first one. Correspondingly we can define two divisions for ideals, because any ideal can be given by one polynomial.

Computing of Groebner basis uses first division and needs ordering of variables. The second division does not use the ordering of variables (the order of variables can be changed during execution of algorithm F4, F5).

These two different divisions use the field ideal

$$\mathfrak{F} = (x_1^2 + x_1, \ldots, x_n^2 + x_n) = (x_1^2 + x_1) \oplus \ldots \oplus (x_n^2 + x_n),$$

that gives the field $\mathbb{F}_2$, in different manner. For the first division the ideal $\mathfrak{F}$ is external with respect to the ring, for the second division the ideal $\mathfrak{F}$ is internal.

Each of these two divisions defines the remainder of polynomial modulo ideal and hence the remainder of ideal modulo ideal. For polynomial division the remainder is well-defined if the ideal is given by Groebner basis, but even if it is so, the remainder depends on ordering of variables. Notice that the ideal can be given in different ways (by different number of polynomials of basis and by different polynomials if its number ≥ 2 is the same). This follows that the remainder of polynomial (or ideal) modulo ideal is not uniquely defined.

For the second division we also can define the remainder of polynomial (or ideal) modulo polynomial (or ideal). It is sufficient to consider the remainder of polynomial modulo polynomial. Let $f = gh + f_1$, then $f \equiv f_1 (\mathrm{mod}(g))$. But since for polynomials f, g the quotient polynomial h is not unique, the remainder f_1 is not unique too. There is a bijection between ideals and polynomials, so the remainder for polynomial division is defined in a similar way. If the ideal if given as the sum of ideals, it is sufficient to consider the ideal as a principal one.

For some reasons, explained later, we are interested in computing the remainder $\mathfrak{A}$ (mod $\mathfrak{I}$) for $\mathfrak{I} \equiv 0 \pmod{\mathfrak{A}}$. In this case we can write $\mathfrak{A} = \mathfrak{I} \oplus \mathfrak{B}$ and $\mathfrak{A} \equiv \mathfrak{B} \pmod{\mathfrak{I}}$. For principal ideals $(f), (g), (h)$ and corresponding binary vectors of values we obtain $(f) = (g) \oplus (h)$ and $\mathbf{f} = \mathbf{g} \vee \mathbf{h}$, where $\vee$ denotes a logical OR. Last equality shows that there are many remainders of f modulo g. Hence one can obtain suitable remainder modulo ideal for both divisions.

Let the cipher be given by the set of polynomials. These polynomials form ideal $\mathfrak{A}$ and variety $V(\mathfrak{A})$, which in algebraic geometry is usually considered over algebraic closure of $\mathbb{F}_2$. Even if the key is uniquely defined by plaintexts and corresponding ciphertexts,

[3]Two multiplications define duality of the ring and its ideals. Among the set of ideals there exist self-dual ideals, which do not change under changing the multiplication.

$\#V(\mathfrak{A})$ is extremely large. In order to obtain a point of $V(\mathfrak{A})$ over $\mathbb{F}_2$ we must consider the ideal $\mathfrak{A} \oplus \mathfrak{F}$ instead of $\mathfrak{A}$.

For simplification it was suggested to partite the ideal $\mathfrak{A} \oplus \mathfrak{F}$ in two parts: find solutions of the ideal $\mathfrak{A}$ and reduce intermediate polynomials modulo the ideal $\mathfrak{F}$. Reduction modulo $\mathfrak{F}$ is very easy [6, 9, 5].

We generalize this method and represent the ideal $\mathfrak{A}$ as a sum $\mathfrak{A} = \mathfrak{B} \oplus \mathfrak{I}$, where the ideal $\mathfrak{I}$ contains short polynomials of $\mathfrak{A}$. Here initial polynomials of the basis of the ideal $\mathfrak{A}$ and syzygies can be changed by their images in the quotient ring $\frac{\mathfrak{A}}{\mathfrak{I} \oplus \mathfrak{F}}$ (or by polynomials over the variety $V(\mathfrak{I} \oplus \mathfrak{F})$).

4 Ideal of substitution and its representation as a sum of ideals

In this section we consider ideals of finite ring of NAF polynomials.

Usually the cipher is a composition of non-linear substitutions that act on short (3 – 8 bit) words and linear diffusion maps. Complexity of solving of such system of equations is defined mainly by non-linear polynomials of substitution. In [6, 9, 5] authors wrote the substitution as a set of implicit functions of degree 2, the number of linearly independent polynomials exceeds the number of variables.

Let $\mathbf{y} = S(\mathbf{x})$ be a substitution, that acts on n-bit words $\mathbf{x} = (x_1, \ldots, x_n), \mathbf{y} = (y_1, \ldots, y_n)$. Ideal of the substitution $\mathfrak{A}_S \subset \mathfrak{G}_{2n}[\mathbf{x}, \mathbf{y}]$ is the set of polynomials that take zero if equality $\mathbf{y} = S(\mathbf{x})$ holds. Hence $\mathrm{V}(\mathfrak{A}_S)$ has 2^n points.

It is common that for increasing strength with respect to linear [10] and differential [3] cryptanalysis, substitutions are to be chosen in a special way. Let $\mathbf{x}, \mathbf{x}'$ be a pair of inputs and $\mathbf{y}, \mathbf{y}'$ — corresponding outputs of the substitution. The differential of substitution is a pair $(\Delta\mathbf{x}, \Delta\mathbf{y})$, where $\Delta\mathbf{x} = \mathbf{x} + \mathbf{x}', \Delta\mathbf{y} = \mathbf{y} + \mathbf{y}'$, is characterized by its probability. The most likely differential must have minimal probability. Also probabilities of equalities $\mathbf{ax} + \mathbf{by} = 0$ must be close to 0.5. In [14] it is shown that special substitutions apparently have no advantage with respect to random ones.

The system of polynomial equations that describes the cipher is hard to solve for the next reason. The initial basis of ideal is given by short polynomials. The final basis of ideal is simple too. But intermediate polynomials are very complex (have large length and large degree) and take exponential memory. The number of syzygies also increases. For example if two polynomials have degree 2 and lengths l_1, l_2 respectively, their syzygy usually has degree 4 and its length is $l_1 + l_2 - 2$.

But if one of those polynomials is binomial, the length of syzygy does not increase. If one of polynomials is trinomial, the length of syzygy increases at most by 1. This shows that it is useful to define the ideal of substitution by short polynomials (monomials, binomials, trinomials, quadrinomials) even if its number will be large.

The ideal of substitution as Boolean function can be given by one, two, or more polynomials. Typically the ideal of substitution is given by polynomials $y_i + S_i(x)$, where S_i

is Boolean function that defines i-th output bit of substitution. Polynomials for inverse substitution can be used similarly. But such polynomials usually are not the shortest ones.

We are interested in defining the ideal of substitution by short polynomials, precisely or approximately. It is obvious that the ideal of substitution is not monomial. There are substitutions with binomial ideals, for example, the ideal of linear substitution defined by bit permutation is generated by binomials $x_i + y_j$.

Consider representation of the ideal as a sum of principal ideals. Let $\mathfrak{A} = (f) = (f_1, \ldots, f_k) = (f_1) \oplus \ldots \oplus (f_k)$ be such representation.

Theorem 1. *Any ideal can be uniquely represented as a sum of additively irreducible ideals.*

Proof. If the ideal (g) is prime, the ideal $(1 + g)$ is additively irreducible. From equality $\mathbf{g}_1 \vee \ldots \vee \mathbf{g}_k = 1 + (1 + g_1) \ldots (1 + g_k)$ and unique factorization in the ring $\mathfrak{G}_{2n}[\mathbf{x}, \mathbf{y}]$ we obtain the conclusion of theorem. □

Ideals of the ring $\mathfrak{G}_{2n}[\mathbf{x}, \mathbf{y}]$ form a monoid under addition, this monoid is isomorphic to the monoid $\mathfrak{V}(2^n)$ of binary vectors under the binary operation OR.

Define XOR-length (Xl) of the ideal $\mathfrak{A}$. If $\mathfrak{A} = (f_1) \oplus \ldots \oplus (f_k)$, there exists polynomial f_i of maximum length. $\mathrm{Xl}(\mathfrak{A})$ is a minimum of maximal lengths of f_i for all representations of $\mathfrak{A}$ as a sum of principal ideals. XOR-length of the ideal of substitution does not exceed length of polynomials, which define Boolean functions of substitution. But sometimes XOR-length of the ideal of substitution can be increased. Computing XOR-length of the ideal seems to be a hard problem.

It is clear that the monomial ideal as Boolean function can be represented as disjunctive normal form (DNF) in such a way that all conjunctions have no inversions. Boolean functions as DNF form a commutative semiring. Consider the relation between binomial ideals and Boolean functions in DNF.

Theorem 2. *The ideal is binomial iff it can be given by DNF where each conjunction has at most one inversion.*

Proof. At first we prove that if DNF consists of conjunctions which have at most one inversion, it represents the binomial ideal. Let m be a monomial. Since $m\overline{x} = m + mx$ is a binomial, such DNF gives binomial ideal. Back, let $(f) = (f_1, \ldots, f_k)$ and all f_i be binomials. Show that any binomial is DNF where all conjunctions have at most one inversion. Let $m_1 + m_2 = m_1\overline{m}_2 \vee \overline{m}_1 m_2$, where m_1, m_2 — monomials. Since $\overline{x_1 \ldots x_l} = \overline{x}_1 \vee \ldots \vee \overline{x}_l$, we obtained the required conclusion. □

Boolean function is symmetric if it does not change under arbitrary permutation of variables. Let $S_i(x_1, \ldots, x_n) \in \mathfrak{G}_n[\mathbf{x}]$ be i-th elementary symmetric polynomial. Define elementary symmetric DNF: $T_i(x_1, \ldots, x_n), 0 \le i \le n$, — disjunction of all conjunction that contain exactly i inversions (all conjunctions contain the same variables, inverse or not). For example, $T_1(x_1, x_2, x_3) = \overline{x}_1 x_2 x_3 \vee x_1 \overline{x}_2 x_3 \vee x_1 x_2 \overline{x}_3$. Here $T_i T_j = 0$ if $i \neq j$. For a semiring of DNF analog of symmetric polynomials the theorem is true.

Theorem 3. *Symmetric Boolean function can be represented as a disjunction of elementary symmetric DNF.*

Proof. It is sufficient to consider non-constant functions. Use the lexicographic ordering for those arguments which give unity values of function f so that the leading sets contain many inversions. Let m be the number of inversions in leading set. Then $f = 1$ for all sets of variables with m inversions, i.e. $f = T_m \vee \ldots$. After that find the next set of variables with m_1 inversions, etc. So $f = T_m \vee T_{m_1} \vee \ldots \vee T_{m_k}$. □

Consider trinomial ideals.

Theorem 4. *DNF is trinomial in the ring $\mathfrak{G}_n[\mathbf{x}]$ iff this DNF can be represented as $f = T_0(C_1, C_2, C_3) \vee T_2(C_1, C_2, C_3)$, where C_1, C_2, C_3 are conjunctions without inversions.*

Proof. Directly check that $f = 1$ iff $C_1 + C_2 + C_3 = 1$ and hence $f = 0$ iff $C_1 + C_2 + C_3 = 0$. This method gives arbitrary trinomial ($C_i = 1$ is possible). □

Theorem 5. *The ideal is trinomial iff it is given by next DNF:*

$$f = \vee_i (T_0(C_{1i}, C_{2i}, C_{3i}) \vee T_2(C_{1i}, C_{2i}, C_{3i})),$$

where C_{ji} are conjunction without inversions.

Proof. The proof follows immediately from the theorem 4. □

Consider quadrinomials. Let C_1, C_2, C_3, C_4 be monomials (conjunctions without inversions). The sum $C_1 + C_2 + C_3 + C_4$ takes 1 iff one or three terms take 1. So $C_1 + C_2 + C_3 + C_4 = T_1(C_1, C_2, C_3, C_4) + T_3(C_1, C_2, C_3, C_4)$.

Defining the ideal as a sum of principal ideals of the ring $\mathfrak{G}_n[\mathbf{x}]$ is not unique and is described in terms of Boolean function with binary operations AND, OR, XOR and the constant 1. Monomials are obtained by AND-multiplication of variables. XOR sums of monomials are polynomials that define principal ideals. OR operation (a disjunction of polynomials) defines a sum of principal ideals. Such Boolean function admits minimization. We are interested in minimization largest number of XOR operations in polynomials. Consider some algebraic properties of Boolean functions with these operations.

Theorem 6. *Let S_i be i-th elementary symmetric NAF polynomial. Next equalities hold.*

1. $a \vee (a + b) = a \vee b$.
2. $a_1 \vee \ldots \vee a_n = S_1(a_1, \ldots, a_n) + \ldots + S_n(a_1, \ldots, a_n)$.
3. $(a_1 + b) \vee \ldots \vee (a_n + b) = S_1(a_1, \ldots, a_n) + \ldots + S_n(a_1, \ldots, a_n) + b(1 + S_1(a_1, \ldots, a_n) + \ldots + S_{n-1}(a_1, \ldots, a_n))$.

Proof. The first equality and the second, third equalities for $n = 2$ can be proved directly. The case $n > 2$ can be proved by induction. □

This material allows the constructing of the algorithm for computing short polynomials that define basis of ideal of substitution (exactly or approximately). Notice that the direct test of short polynomials is hard. For example, for $n = 4$ number of trinomials is $2.7 \cdot 10^6$, number of quadrinomials is $1.75 \cdot 10^8$, for $n = 8$ number of binomials, trinomials, quadrinomials is $2.1 \cdot 10^9, 4.7 \cdot 10^{13}, 7.7 \cdot 10^{17}$ correspondingly. For speeding-up the algorithm number of possible terms in polynomials of basis of ideal is to be limited. This is achieved by one of the next algorithms.

Algorithm 1. *Short polynomials that define ideal of substitution S.*

1. *Find the variety $V(S)$ and its complement $\overline{V(S)}$, compute the polynomial $f(S)$ that gives the principal ideal of the substitution.*

2. *Find the list of monomials that take 0 in $V(S)$. Delete monomials that are divided by some other monomials of the list (first division, for polynomials). Obtain the monomial ideal $\mathfrak{A}_1$. Find the variety $V_1 = \overline{V(S)} \cap V(\mathfrak{A}_1)$.*

3. *Compute $f_2 = f(S) \pmod{\mathfrak{A}_1}$, find the list T_2 of monomials of f_2 and monomial divisors of that monomials.*

4. *Find binomials as sums of two monomials of the list T_2, that take 0 in all points of $V(S)$, and take 1 in some points of V_1. Join binomials to the list of monomials and find the ideal $\mathfrak{A}_2 = \mathfrak{A}_1 \oplus \{$obtained binomials$\}$ and its variety $V(\mathfrak{A}_2)$. Delete redundant binomials that does not change $V(\mathfrak{A}_2)$. Find the variety $V_2 = \overline{V(S)} \cap V(\mathfrak{A}_2)$.*

5. *Compute $f_3 = f_2 \pmod{\mathfrak{A}_2}$, find the list T_3 of monomials of f_3 and monomial divisors of that monomials.*

6. *Find trinomials of the list T_3 that take 0 in all points $V(S)$ and take 1 in some points of V_2. Join trinomials to the basis of ideal and find $\mathfrak{A}_3 = \mathfrak{A}_2 \oplus \{$obtained trinomials$\}$. Delete redundant trinomials that do not change $V(\mathfrak{A}_3)$. Find the variety $V_3 = \overline{V(S)} \cap V(\mathfrak{A}_3)$.*

7. *Repeat two last steps for computing quadrinomials, etc. The algorithm is stopped if the set V_k becomes empty or its cardinality becomes small with respect to $V(S)$. In the second case the ideal of substitution is approximated by computed short polynomials.*

Algorithm 2. *Short polynomials that define ideal of substitution S. Let M be a list of all monomials, the size of M is 2^{2n} for n-bit substitution.*

1. *Find the variety $V(S)$ and its complement $\overline{V(S)}$, compute the polynomial $f(S)$ that gives the principal ideal of the substitution.*

2. *Find the list of monomials that take 0 in $V(S)$. Delete monomials that are divided by some other monomials of the list (first division, for polynomials). Obtain the monomial ideal $\mathfrak{A}_1$. Find the variety $V_1 = \overline{V(S)} \cap V(\mathfrak{A}_1)$ and the list $M_1 = M \backslash \mathfrak{A}_1$.*

3. *Compute* $f_2 = f(S) \pmod{\mathfrak{A}_1}$.

4. *Find binomials as sums of two monomials of the list* M_1*, that take* 0 *in all points of* $V(S)$*, and take* 1 *in some points of* V_1*. Join binomials to the list of monomials and find the ideal* $\mathfrak{A}_2 = \mathfrak{A}_1 \oplus$ *{obtained binomials} and its variety* $V(\mathfrak{A}_2)$*. Let* M_2' *be a list of monomials used in leading terms of the ideal* $\mathfrak{A}_2$ *and find the list* $M_2 = M_1 \backslash M_2'$*. Delete redundant binomials that does not change* $V(\mathfrak{A}_2)$*. Find the variety* $V_2 = \overline{V(S)} \cap V(\mathfrak{A}_2)$.

5. *Compute* $f_3 = f_2 \pmod{\mathfrak{A}_2}$.

6. *Find trinomials of the list* M_2 *that take* 0 *in all points* $V(S)$ *and take* 1 *in some points of* V_2*. Join trinomials to the basis of ideal and find* $\mathfrak{A}_3 = \mathfrak{A}_2 \oplus$ *{obtained trinomials}. Let* M_3' *be a list of monomials used in leading terms of the ideal* $\mathfrak{A}_3$ *and find the list* $M_3 = M_2 \backslash M_3'$*. Delete redundant trinomials that do not change* $V(\mathfrak{A}_3)$*. Find the variety* $V_3 = \overline{V(S)} \cap V(\mathfrak{A}_3)$.

7. *Repeat two last steps for computing quadrinomials, etc. The algorithm is stopped if the set* V_k *becomes empty or its cardinality becomes small with respect to* $V(S)$*. In the second case the ideal of substitution is approximated by computed short polynomials.*

Both algorithms give an approximation of substitution. The algorithm 1 is fast but the algorith 2 gives better approximation. Experiments show that monomials of ideal of n-bit substitution have degree close to n. The transition from monomials to binomials, trinomials, etc, cause incrementing of degree of corresponding polynomials with respect to more short ones (see section 7).

5 Affine equivalence of ideals

The set of n-bit substitutions can be partitioned by affine equivalence: $S_1 \sim S_2$, if $S_1 = AS_2B$, where A, B are affine substitutions, $A(\mathbf{x}) = L\mathbf{x} + \mathbf{c}, L$ is invertible matrix. Affine equivalence is a tool of cryptanalysis [4]. Affine equivalence does not change probabilities of most likely differentials and absolute biases of most likely linear sums.

The complexity of computing Groebner basis does not depend on which variables correspond to the input and to the output of substitution. Hence we can arbitrary define which bits are input and output of substitution. This reason shows that we can mix input/output bits and then partite it in two parts by some way. So the ideal of substitution can correspond to at least 2 substitutions (initial and inverse).

The ideal $\mathfrak{A}_S \subset \mathfrak{G}_{2n}[\mathbf{x}, \mathbf{y}]$, which has 2^n zeroes, is an ideal of n-bit binary map with input $\mathbf{u}$, if we can choose n variables $\{u_1, \ldots, u_n\} \subset \{x_1, \ldots, x_n, y_1, \ldots, y_n\}$ in such way,that all zeroes of the ideal correspond to different sets of input variables.

Some substitutions possess the property that their ideals correspond to many substitution or binary maps. For example the identity substitution corresponds to 16 substitutions.

The ideal of affine substitution $\mathbf{y} = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ corresponds to 44 four-bit binary maps, 24 of them are substitutions. We tested ideals of 10 substitutions from DES S-box. All of them correspond to $4-9$ binary maps. The ideal of AES substitution corresponds only to the initial substitution and its inverse.

Define affine equivalence of ideals of the ring $\mathfrak{G}_{2n}[\mathbf{x}, \mathbf{y}]] : \mathfrak{A} \sim \mathfrak{B}$, if $\mathfrak{A}(\mathbf{x}, \mathbf{y}) = \mathfrak{B}(L \cdot (\mathbf{x}, \mathbf{y}) + \mathbf{c})$, where invertible $2n \times 2n$ over $\mathbb{F}_2$ matrix L is multiplied by the column $(\mathbf{x}, \mathbf{y})$. Affine equivalence of ideals generalizes affine equivalence of substitution, where the matrix L is block-diagonal. The ideal that is affine equivalent to the ideal of substitution can be not an ideal of binary map. Affine equivalent ideals have varieties of the same cardinality, and polynomials that define principal ideals have the same degree.

Considering ideals instead of substitutions allows to generalize the differential and the non-linearity of substitution. The probability of the differential $(\Delta\mathbf{x}, \Delta\mathbf{y})$ of the ideal $\mathfrak{A} \subset \mathfrak{G}_{2n}[\mathbf{x}, \mathbf{y}]$ is defined by averaging on variety $V(\mathfrak{A})$. The non-linearity of the ideal is the smallest non-linearity of Boolean function contained in the ideal. Affine equivalence preserves probabilities of most likely differentials and non-linearity of ideals.

Affine equivalence of ideals is a useful tool for solving systems of Boolean equations. For example, choosing appropriate affine equivalence, we can obtain more suitable polynomials.

Experiments show that short polynomials give a more accurate approximation to the ideal if the length of the polynomial that defines the principal ideal is minimized. Hence we can change the ideal of substitution by affine equivalent ideal defined by short polynomial. For that it is sufficient to find the shift vector $\mathbf{c}$ (it shows which variables x_i, y_j need to be replaced by $1 + x_i, 1 + y_j$). After that we find rows of the matrix L by the steepest descent. Since the group of matrices is generated by transvections [10], search can be based on transvections.

Algorithm 3. *Affine equivalence that minimizes the length of the principal ideal of substitution.*

1. *Compute the polynomial $f(S)$ that defines the principal ideal of substitution.*
2. *Compute the vector $\mathbf{c}$ that minimizes the length of $f(S)$. Change corresponding variables x_i, y_j by $1 + x_i, 1 + y_j$. Compute new $f(S)$.*
3. *Find the linear change of variables $x_1 \leftarrow x_1 + d_2x_2 + \ldots + d_{2n}y_n$ that minimizes the length of $f(S)$. Compute new $f(S)$.*
4. *In turn for $x_2, \ldots, y_n$ do step 3.*
5. *If the minimum is not obtained, repeat steps 3, 4. The result affine equivalence is defined as the vector $\mathbf{c}$ and the product of matrices of linear change of variables, $f(S)$ gives affine equivalent principal ideal.*

The algorithm 3 computes the affine change of variables that minimizes the length of the principal ideal. This affine equivalent ideal is quasi-optimal for minimization XOR-length of ideal.

6 Preparing a system of Boolean equations and its solution

Symmetric cipher is defined by a system of NAF polynomials. Non-linear polynomials describe the substitution. Usually polynomials that define Boolean functions for each bit of the substitution are used. But such system of polynomial is not suitable for solving.

Solving can be speeded-up if the system is over defined. If number of linearly independent polynomials that define the ideal of the substitution is large, complexity of computing Groebner basis decreases. Some authors [6, 9, 5] suggested using polynomials of small degree. For example, AES 8-bit substitution is described by 24 linearly independent quadratic polynomials.

Let $\mathfrak{A}$ be the ideal that defines encryption process and $\mathfrak{F}$ is a field ideal. In [6, 9, 5] authors suggest to decompose the initial ideal $\mathfrak{A} \oplus \mathfrak{F}$ into two ideals $\mathfrak{A}$ and $\mathfrak{F}$. Intermediate polynomials obtained from the ideal $\mathfrak{A}$ are reduced modulo $\mathfrak{F}$.

We propose to use other decomposition. We define the approximate ideal $\mathfrak{I}$ (for the ideal of substitution or its affine equivalent) that is generated by short polynomials. Instead of the field ideal $\mathfrak{F}$ we use the ideal $\mathfrak{F} \oplus \mathfrak{I}$. Notice that the basis of $\mathfrak{I}$ is not necessary Groebner. Initial polynomials of the ideal $\mathfrak{A}$ and intermediate polynomials are reduced modulo $\mathfrak{F} \oplus \mathfrak{I}$ (here we essentially use polynomials on the variety $V(\mathfrak{F} \oplus \mathfrak{I})$ instead of usual NAF polynomials).

Such reduction can be easily performed by specialized logical chips. If the monomial C of syzygy is divided by monomial of $\mathfrak{I}$, C can be neglected. The same is true for binomials and trinomials. Besides of deleting binomials the other technique can be used. Let C be a monomial of polynomial to be reduced, $A + B$ is a binomial of $\mathfrak{I}$ and $C = AD$. Then $A \equiv B \pmod{\mathfrak{I}}$ and we can change $C \leftarrow BD$. Notice that the ordering of monomials and variables is not necessary; we can delete leading monomials in arbitrary order and change this order during the computation. If polynomial has the trinomial $A + B + C$ and $A + B \equiv 0 \pmod{\mathfrak{I}}, A + C \equiv 0 \pmod{\mathfrak{I}}$, then $B + C \equiv 0 \pmod{\mathfrak{I}}$, and $A + B + C \equiv A \equiv B \equiv C \pmod{\mathfrak{I}}$, we can change the trinomial by any of these monomials.

For example, the ideal of AES 8-bit substitution has 256 zeroes among 65536 points of affine space. The length of the polynomial that defines the principal ideal is 25465. The ideal of substitution contains 432 monomials such that any of them does not divide any other (in the sense of polynomial division): 36 monomial of degree $6, 324$ monomials of degree $7, 71$ monomial of degree $8, 1$ monomial of degree 9. This monomial ideal has 58988 zeroes, so it is bad approximation. But reduction of the principal ideal of substitution modulo monomial ideal and field ideal has length 16564. The problem of computing approximating ideal, generated by short polynomials, and affine minimization of the ideal

of AES substitution may be hard for a personal computer. But 8-bit fixed substitution can be changed by a network of at most 244-bit substitutions and linear maps. Indeed, AES substitution is a composition of powering to degree 254 in $\mathbb{F}_{256}$ and affine map. Since $254 = 2 + 4 + 8 + 16 + 32 + 64 + 128$ and the powers of 2 are defined by linear maps over $\mathbb{F}_2$, 6 multiplications in $\mathbb{F}_{256}$ are required. Any such multiplication is a non-linear map and can be represented by at most 4 multiplications of 4-bit words (or 4-bit substitutions).

The use of affine equivalent ideals is based on unproved hypothesis: joining affine polynomials to a set of non-linear polynomials does not significantly increase the complexity of computing Groebner basis.

Notice that the computing of Groebner basis can be stopped if the intermediate basis is given by linear polynomials. Farther linear algebra methods can be applied. In a such strategy significant part of linear polynomials that define affine equivalence are to be used in the last turn of the computation process.

Solving a system of polynomial equations takes two stages. In the first stage we prepare the ideals of substitution using algorithms $1, 2$ in such a way that the ideal is defined by short polynomial with best accuracy and find monomials, binomials, trinomials, quadrinomials that give the basis of the ideal $\mathfrak{I}$, approximate to the initial ideal.

The second stage can be performed in two different ways. Consider the first way.

In the second stage we reduce the initial basis of the ideal (defined by known methods) and obtain syzygies modulo $\mathfrak{F} \oplus \mathfrak{I}$. It is suitable to use a special logical apparatus for these reductions. Reduced syzygies have less degree and are shorter then syzygies in the original algorithm. This allows decreasing complexity of solving a system of Boolean equations. If the solution is not obtained (for example, the solving process gives the trivial equality $0 = 0$, i.e. the solution is in $V(\mathfrak{F} \oplus \mathfrak{I})$, syzygies of the ideal $\mathfrak{I}$ (mod $\mathfrak{F}$) are needed. But Groebner basis of the ideal $\mathfrak{I}$ (mod $\mathfrak{F}$) can be computed relatively easy.

The second way can be applied if $V(\mathfrak{A})$ and $V(\mathfrak{I})$ differ in a small number of points. In this case the ideal $\mathfrak{I}$ obtained in the first stage may not satisfy the congruence $\mathfrak{I} \equiv 0$ (mod $\mathfrak{A}$). We use the ideal $\mathfrak{I}$ instead of the initial ideal. The computing of Groebner basis of the ideal $\mathfrak{I}$ seems to be easier then the such computing for the original ideal $\mathfrak{A}$. In order to relax the influence of probability of wrong solutions, the solving process must be repeated for sufficiently large number of plaintexts/ciphertexts, as in statistical attacks.

Increasing a degree of monomial of syzygy increases the probability that it will be deleted during the reduction modulo $\mathfrak{F} \oplus \mathfrak{I}$. Hence the initial "'natural"' distribution of monomial probabilities changes: low degree monomials become more likely. Polynomials that have only low degree terms can be considered as "'smooth"', similarly to sieve methods[4]. So the answer to the question "'Is the complexity of computing a key of XSL cipher exponential or is it subexponential?"' is not obvious in spite of some difficulties concerned with the transpose from previous encryption round to next one.

[4]Remember that sieve methods have subexponential complexity.

7 Examples of preparing an ideal of substitution

1. The substitution $S = \{00, 1d, 2b, 38, 43, 56, 64, 71, 8f, 92, a5, be, ca, dc, e9, f7\}$. Here the first tetrad corresponds to the input of substitution, the second tetrad corresponds to the output of substitution. The variety of ideal of substitution coincides with S.

   This substitution has the best possible cryptographic properties: the most likely differentials have probability 0.25 and the maximal absolute bias of linear sum is 0.25.

   The ideal of this substitution is approximated by trinomials and quadrinomials with probabilities 0.46 and 0.59 correspondingly (approximate ideal is divided by the ideal of substitution, i.e. $V(S)$ is a subset of the variety of approximate ideal). The approximating monomial ideal contains 13 monomials.

   The principal ideal of substitution is defined by the polynomial of length 161. Minimization of length of the principal ideal by affine equivalence gives the variety $\{6a, 6c, 6d, 7b, 9e, b7, bd, be, d6, d7, df, e3, e6, ec, f7, fa\}$ (the first tetrad corresponds to the variables $\mathbf{x}$, the second tetrad corresponds to the variables $\mathbf{y}$), the length of that affine equivalent ideal $\mathfrak{A}$ is 17:

$$f(\mathfrak{A}) = 1 + x_2x_3y_1y_2 + x_2x_3x_4y_1y_2 + x_2x_3y_1y_3 + x_1x_2x_3y_1y_3 + x_2x_3x_4y_1y_3 + \\ + x_1x_2x_3y_2y_3 + x_1x_2x_4y_2y_3 + x_1x_4y_1y_2y_3 + x_1x_2x_3y_1y_2y_4 + x_1x_3x_4y_1y_2y_4 + \\ + x_1x_2x_3y_3y_4 + x_1x_2x_3x_4y_3y_4 + x_2x_3y_1y_3y_4 + x_1x_3x_4y_2y_3y_4 + x_2x_3y_1y_2y_3y_4 + \\ + x_1x_4y_1y_2y_3y_4.$$

   This ideal is defined by trinomials and quadrinomials with probabilities 0.40 and 0.84 correspondingly. Let $\mathfrak{I}$ be the approximating ideal for the ideal $\mathfrak{A}$. The reduced basis of the ideal $\mathfrak{I}$ (it is not Groebner) contains 1 monomial of degree 5, 16 binomials of average degree 3.7, 16 trinomials of average degree 2.7 and 6 quadrinomials of average degree 1 (all quadrinomials have maximal degree 2 and contain only one quadratic term).

   The reduction of the principal ideal $\mathfrak{A}$ modulo $\mathfrak{I} \oplus \mathfrak{F}$ decreases average degree of monomials of $f(\mathfrak{A})$ from 5 to 2.5. Notice that the reduction admits a large number of results. Our result was performed using MATHEMATICA package, and possibly it is not optimal.

2. The substitution $S = \{09, 14, 2a, 3b, 4d, 51, 68, 71, 86, 92, a0, b3, cc, de, ef, f7\}$ of SAES (short AES, simplified model of AES [12]).

   This substitution has also the best possible cryptographic properties: the most likely differentials have probability 0.25 and the maximal absolute bias of linear sum is 0.25.

   The ideal of this substitution is approximated by trinomials and quadrinomials with probabilities 0.52 and 0.53 correspondingly (approximate ideal is divided by the ideal of substitution).

The principal ideal of substitution is defined by the polynomial of length 107. Minimization of length of the principal ideal by affine equivalence gives the variety $\{1f, 5f, 6d, 6f, 9f, af, b9, ba, be, d3, d6, db, ed, f2, f3, fd\}$ (the first tetrad corresponds to the variables $\mathbf{x}$, the second tetrad corresponds to the variables $\mathbf{y}$). The length of the polynomial that defines the principal ideal is 15:

$$f(\mathfrak{A}) = 1 + x_1x_2x_3x_4y_3 + x_1x_3x_4y_1y_3 + x_1x_2x_4y_2y_3 + x_1x_2x_4y_1y_2y_3 + x_1x_3x_4y_1y_4 + \\ + x_1x_2x_3x_4y_1y_4 + x_2x_3y_1y_2y_4 + x_1x_3x_4y_1y_2y_4 + x_2x_3x_4y_1y_2y_4 + x_1x_2x_4y_3y_4 + \\ + x_1x_2x_3x_4y_3y_4 + x_1x_3y_1y_2y_3y_4 + x_4y_1y_2y_3y_4 + x_3x_4y_1y_2y_3y_4.$$

This ideal is approximated by trinomials and quadrinomials with probability 0.76 and 0.89 correspondingly — significantly better comparatively to the initial ideal of substitution. Let $\mathfrak{I}$ be the approximating ideal for the ideal $\mathfrak{A}$. Initially basis of $\mathfrak{I}$ contains 8 monomials of degree 5, but reducing of the basis shows that all they are redundant. The reduced basis of the ideal $\mathfrak{I}$ (it is not the Groebner) contains 14 binomials of average degree $3.6, 17$ trinomials of average degree $2.6, 3$ quadrinomials of average degree 1.5.

The reduction of the principal ideal $\mathfrak{A}$ modulo $\mathfrak{I} \oplus \mathfrak{F}$ decreases the average degree of monomials of $f(\mathfrak{A})$ from 5.2 to 2.

References

[1] Martin Albrecht and Carlos Cid. Algebraic techniques in differential cryptanalysis. Cryptology ePrint Archive, Report 2008/177, 2008. `http://eprint.iacr.org/2008/177`.

[2] M. Atiyah and I. Macdonald. *Introduction to commutative algebra.* Addison-Wesley series in mathematics. Westview Press, 1969.

[3] Eli Biham and Adi Shamir. Differential cryptanalysis of des-like cryptosystems. In Alfred Menezes and Scott A. Vanstone, editors, *Advances in Cryptology - CRYPTO 90, 10th Annual International Cryptology Conference, Santa Barbara, California, USA, August 11-15, 1990, Proceedings*, volume 537 of *Lecture Notes in Computer Science*, pages 2–21. Springer, 1990.

[4] Alex Biryukov, Christophe De Canniere, An Braeken, and Bart Preneel. A toolbox for cryptanalysis: Linear and affine equivalence algorithms. In Eli Biham, editor, *Advances in Cryptology - EUROCRYPT 2003, International Conference on the Theory and Applications of Cryptographic Techniques, Warsaw, Poland, May 4-8, 2003, Proceedings*, volume 2656 of *Lecture Notes in Computer Science*, pages 33–50. Springer, 2003.

[5] Nicolas Courtois, Alexander Klimov, Jacques Patarin, and Adi Shamir. Efficient algorithms for solving overdefined systems of multivariate polynomial equations. In Bart

Preneel, editor, *Advances in Cryptology - EUROCRYPT 2000, International Conference on the Theory and Application of Cryptographic Techniques, Bruges, Belgium, May 14-18, 2000, Proceeding*, volume 1807 of *Lecture Notes in Computer Science*, pages 392–407. Springer, 2000.

[6] Nicolas Courtois and Josef Pieprzyk. Cryptanalysis of block ciphers with overdefined systems of equations. Cryptology e-print archive, report 2002/044, 2002. `http://e-print.iacr.org/2002/044`.

[7] David A. Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms - an introduction to computational algebraic geometry and commutative algebra (2. ed.)*. Undergraduate texts in mathematics. Springer, 1997.

[8] B. Sturmfels D. Eisenbud. Binomial ideals. *Duke mathematical journal*, 84, 1996.

[9] Jean-Charles Faugere. A new efficient algorithm for computing grobner bases without reduction to zero (f4). *Journal of Pure and Applied Algebra*, 139:61–88.

[10] M. I. Kargopolov and Y. I. Merzliakov. *Basics of group theory.* Nauka, M., 1982.

[11] M. Matsui. *Linear cryptanalysis method for DES cipher*, volume 765. EUROCRYPT '93, LNCS, 1994.

[12] M. Musa, E. F. Schaefer, and S. Wedig. A simplied aes algorithm and its linear and differential cryptanalysis. *Cryptologia*, 27(2):148–177, 2003.

[13] Havard Raddum and Igor Semaev. New technique for solving sparse equation systems. Cryptology ePrint Archive, Report 2006/475, 2006. `http://eprint.iacr.org/2006/475`.

[14] Alexander Rostovtsev. Changing probabilities of differentials and linear sums via isomorphisms of ciphers. Cryptology ePrint Archive, Report 2009/117, 2009. `http://eprint.iacr.org/2009/177`.

[15] Xiao shan Gao and Zhenyu Huang. Efficient characteristic set algorithms for equation solving in finite fields and application in analysis of stream ciphers. Cryptology ePrint Archive, Report 2009/637, 2009. `http://eprint.iacr.org/2009/637`.

[16] M. Sugita, M. Kawazoe, and H. Imai. Relation between xl algorithm and groebner bases algorithms. Cryptology ePrint Archive, Report 2004/112, 2004. `http://eprint.iacr.org/2004/112`.

[17] Xijin Tang and Yong Feng. A new efficient algorithm for solving systems of multivariate polynomial equations. Cryptology ePrint Archive, Report 2005/312, 2005. `http://eprint.iacr.org/2005/312`.