УДК 681.3.06:519.248.681

ПОСТРОЕНИЕ ПЕРЕОПРЕДЕЛЁННОЙ СИСТЕМЫ УРАВНЕНИЙ ДЛЯ ОПИСАНИЯ АЛГОРИТМА ШИФРОВАНИЯ «ЛАБИРИНТ»

Олейников Р.В, Казимиров А.В.

Изучаются показатели криптографической стойкости шифра «Лабиринт» к алгебраической атаке, строящейся на основе описания S-блоков алгоритма с помощью переопределенной системы уравнений. Показана возможность проведения атаки на модифицированный вариант шифра.

Cryptographic strength of symmetric block cipher “Labirinth” to algebraic attack based on cipher S-box description with overdefined system of equations is researched. It is shown the possibility of the attack to the modified version of the cipher.

Введение

В настоящее время в Украине проходит открытый конкурс, предполагающий выбор алгоритма-прототипа национального стандарта блочного симметричного шифра. Среди участников конкурса представлен алгоритм «Лабиринт», разработанный в ЗАО «Криптомаш» [1].

Анализ решений, использованных при построении этого шифра, позволяет заключить, что он во многом унаследовал принципиальные идеи победителя конкурса AES. Как и в Rijndael, основой линейного преобразования является матричное умножение в поле, построенное на использовании кодов МДР. Близкими по конструкции (и алгебраическими по структуре) оказываются и S-блоки шифра «Лабиринт». Как и в Rijndael, при их построении используется принцип, предложенный К. Ниберг, т.е. преобразования, использующего вычисление обратного элемента в поле $GF(2^8)$, с последующим аффинным преобразованием. Поэтому и для этого шифра открывается возможность его описания с помощью детерминированных [6] и вероятностных [8] систем алгебраических уравнений.

1. Алгебраическое описание раундовой функции

Шифр является итеративным, т.е. основу его процедуры шифрования составляет цикловое преобразование, которое повторяется заданное число раз. Для алгоритма «Лабиринт» каждый цикл состоит из двух идентичных итераций цепи Фейстеля, при этом для обновления одного блока требуется, как минимум, две итерации. Кроме повторяемого циклового преобразования процедура шифрования включает начальное (IT) и конечное (FT) преобразования.

В большинстве случаев, при анализе свойств S-блока рассматривается функциональная зависимость выходных значений битов от входных: $y_h = f_h(x_7, x_6, \ldots, x_0), h = 0,1,\ldots,7$ [1]. Для большинства современных шифров, включая «Лабиринт», степень такого уравнения равна семи. Тем не менее, используя общий вариант описания S-блока системой уравнений, где используются все входные и выходные переменные, можно получить более низкую степень термов. Более того, в некоторых случаях, такая система оказывается переопределённой [6].

Один из алгоритмов поиска такой системы предполагает построение матрицы, описывающей все возможные значения термов для всех вариантов входных переменных S-блока [2].

Для k битов на входе размерность такой матрицы составит:

$$|A| = \left(2^k\right) \times \left(C_{2k}^2 + 2k + 1\right),$$

где C_m^n – число сочетаний из m элементов по n. Здесь первый индекс элемента матрицы определяет вариант входа S-блока, второй – номер терма, включая все входные и выходные переменные, все комбинации второй степени и константу 1.

Пусть $X = \{x_7,x_6,x_5,x_4,x_3,x_2,x_1,x_0\}$ – байт, который подаётся на вход S-блока, а $Y = \{y_7,y_6,y_5,y_4,y_3,y_2,y_1,y_0\}$ - байт на выходе S-блока. Тогда строка матрицы A имеет вид { 1, x_7, ... , x_0, y_7, ... , y_0, x_7x_6, x_7x_5, ... , x_1x_0, x_7y_7, x_7y_6, ... , x_0y_0, y_7y_6, y_7y_5,…, y_1y_0 }.

Применение к этой матрице преобразования NullSpace [7] позволяет получить новую матрицу (если такая существует), которой будет описываться S-блок. Эта матрица используется для построения переопределённой системы уравнений.

Изложенная методика была применена к S-блоку шифра «Лабиринт». Размерность матрицы S-блока составила при этом $\left(2^8\right) \times \left(C_{16}^2 + 16 + 1\right) = 256 \times 137$, из которой была получена переопределённая система, состоящая из 39 уравнений.

Отметим, что для обратного S-блока также получается переопределённая система из 39 уравнений, что в совокупности даёт $39 \cdot 2 = 78$ уравнений.

Ещё 7 дополнительных уравнений позволяет получить алгебраическая структура подстановки, которая аффинно эквивалентна конструкции Ниберг-Динга и может быть записана как

$S(X) = M_Y \times \left\lfloor \left(M_X \times X \oplus V_X\right)^E \right\rfloor_B \oplus V_Y;$

$X, V_X, V_Y \in F_2^8; M_X, M_Y \in GL(8, F_2); E = 2^8 - 1 - 2^t$, где $0 \le t < 8$,

B – некоторый базис над $GF(2^8)$ [3,4], который определяется образующим (неприводимым) полиномом 8-й степени $f_s(x)$;

E – показатель степени;

M_X, M_Y – квадратные невырожденные матрицы размерностью 8x8.

Пусть $Z = M_Y^{-1} \times \left(S(X) \oplus Y_Y\right)$, $Z^{-1} = M_X \times X \oplus Y_X$. С учётом того, что для шифра «Лабиринт» E=254, получаем $Z * Z^{-1} = 1$. Последнее соотношение и даёт нам дополнительные 7 уравнений.

В результате S-блок можно описать переопределённой системой из $39 \cdot 2 + 7 = 85$ уравнений. Необходимо отметить, что уравнения справедливы для всех входных наборов, и при этом максимальная степень системы равна 2.

Для того чтобы описать всю F-функцию, необходимо представить биты на выходе S-блока через биты выхода F-функции. Пусть X – вектор-столбец длиной 8 байт, соответствующий слову-аргументу, а Y – вектор-строка, образованная в результате умножения вектора X на матрицу M, т.е. $Y = M \times X$ (рис. 1).

$$Y = M \times \begin{pmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{pmatrix} = \begin{pmatrix} 0C & 0F & 1F & 0A & 0F & 03 & 01 & 01 \\ 0F & 1F & 0A & 0F & 03 & 01 & 01 & 0C \\ 1F & 0A & 0F & 03 & 01 & 01 & 0C & 0F \\ 0A & 0F & 03 & 01 & 01 & 0C & 0F & 1F \\ 0F & 03 & 01 & 01 & 0C & 0F & 1F & 0A \\ 03 & 01 & 01 & 0C & 0F & 1F & 0A & 0F \\ 01 & 01 & 0C & 0F & 1F & 0A & 0F & 03 \\ 01 & 0C & 0F & 1F & 0A & 0F & 03 & 01 \end{pmatrix} \times \begin{pmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{pmatrix} = \begin{pmatrix} D_0 \\ D_1 \\ D_2 \\ D_3 \\ D_4 \\ D_5 \\ D_6 \\ D_7 \end{pmatrix}$$

Рисунок 1 – MBN-преобразование

Ввиду того, что MBN-преобразование является биективным, то для матрицы M (рис. 2) должна существовать обратная матрица M^{-1} (рис. 3).

$$M = \begin{pmatrix} 0C & 0F & 1F & 0A & 0F & 03 & 01 & 01 \\ 0F & 1F & 0A & 0F & 03 & 01 & 01 & 0C \\ 1F & 0A & 0F & 03 & 01 & 01 & 0C & 0F \\ 0A & 0F & 03 & 01 & 01 & 0C & 0F & 1F \\ 0F & 03 & 01 & 01 & 0C & 0F & 1F & 0A \\ 03 & 01 & 01 & 0C & 0F & 1F & 0A & 0F \\ 01 & 01 & 0C & 0F & 1F & 0A & 0F & 03 \\ 01 & 0C & 0F & 1F & 0A & 0F & 03 & 01 \end{pmatrix}$$

Рисунок 2 – матрица, которая используется при прямом MBN-преобразовании

$$M^{-1} = \begin{pmatrix} C0 & C7 & A9 & 9E & C6 & 2E & F3 & 35 \\ C7 & A9 & 9E & C6 & 2E & F3 & 35 & C0 \\ A9 & 9E & C6 & 2E & F3 & 35 & C0 & C7 \\ 9E & C6 & 2E & F3 & 35 & C0 & C7 & A9 \\ C6 & 2E & F3 & 35 & C0 & C7 & A9 & 9E \\ 2E & F3 & 35 & C0 & C7 & A9 & 9E & C6 \\ F3 & 35 & C0 & C7 & A9 & 9E & C6 & 2E \\ 35 & C0 & C7 & A9 & 9E & C6 & 2E & F3 \end{pmatrix}$$

Рисунок 3 – матрица, которая используется при обратном MBN-преобразовании

Поэтому Y и X связаны между собой соотношением $X = M^{-1} \times Y$ (рис. 4).

$$\begin{pmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{pmatrix} = M^{-1} \times \begin{pmatrix} D_0 \\ D_1 \\ D_2 \\ D_3 \\ D_4 \\ D_5 \\ D_6 \\ D_7 \end{pmatrix}$$

Рисунок 4 – обратное к MBN преобразование

2. Описание алгебраической атаки

Для реализации атаки была построена упрощённая схема основной функции (F-функции) алгоритма «Лабиринт» (рис. 5) с длиной блока 128 бит и длиной ключа 128 бит. От оригинала она отличается тем, что операция сложения по модулю 2^{64} была побитовым сложением (XOR). Компонента xc_t для данного случая равна 0, поэтому на рисунке 5 она не отображена.

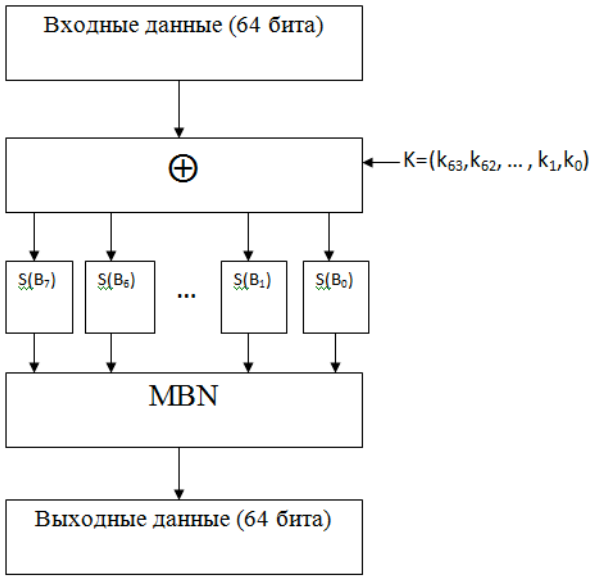

Рисунок 5 – Упрощённая F-функция

Для нахождения системы уравнений, описывающей F-функцию, рассмотрим по отдельности основные её компоненты: операцию XOR, S-блок и MBN преобразование.

Операция XOR представляет собой побитовую сумму по модулю 2. Пусть входное сообщение имеет вид A= (a_{63},a_{62}, … , a_1,a_0), а ключ - K= (k_{63},k_{62}, … , k_1, k_0), тогда сообщение после операции XOR примет вид B = ($a_{63} \oplus k_{63}$,$a_{62} \oplus k_{62}$, … , $a_1 \oplus k_1$,$a_0 \oplus k_0$).

На вход S-блока подаются 64 бита сообщения B, разбитые на 8 байт, т.е. X_7 = ($a_{63} \oplus k_{63}$, … , $a_{56} \oplus k_{56}$), X_6 = ($a_{55} \oplus k_{55}$, … , $a_{48} \oplus k_{48}$), … , X_0 = ($a_7 \oplus k_7$, … , $a_0 \oplus k_0$) (рис. 6).

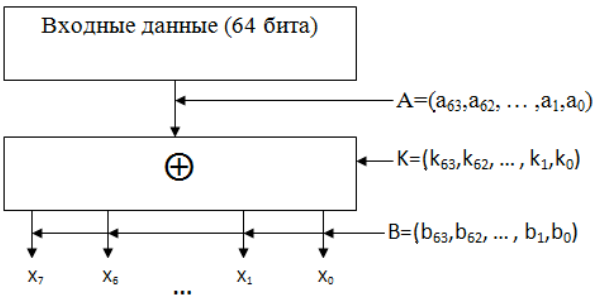

Рисунок 6 - Представление сообщения на входе S-блока

Пусть $D = (d_{63},d_{62}, \dots , d_1,d_0)$ сообщение на выходе F-функции. Для того чтобы получить сообщение на выходе S-блока, мы должны выполнить преобразование, изображённое на рисунке 4.

В результате мы получим сообщение $C = (c_{63}, c_{62}, \dots , c_1, c_0)$, которое разбивается на блоки по восемь байт – $Y_7 = (y_{63}, y_{62}, \dots ,y_{56})$; $Y_6 = (y_{55}, y_{54}, \dots ,y_{48})$; … ; $Y_0 = (y_7, y_6, \dots ,y_0)$ (рис. 7).

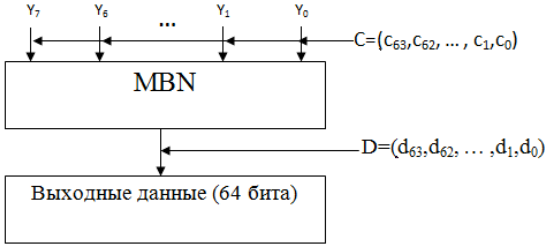

Рисунок 7– Представление сообщения на выходе S-блока

Следовательно, всю F-функцию (рис. 8) можно представить в виде системы уравнений.

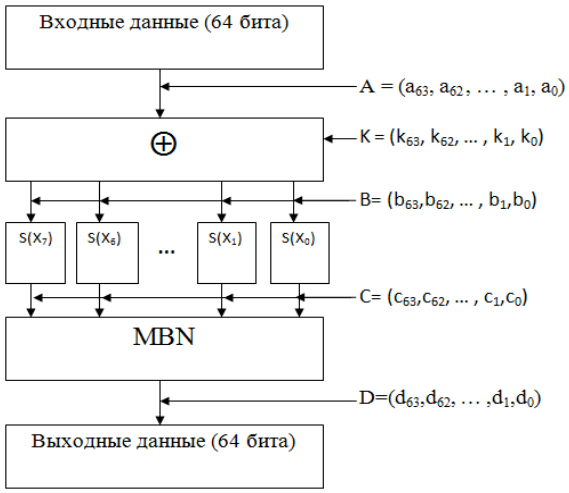

Рисунок 8 – Представление F-функции

Таким образом, F-функцию удается описать с помощью $8 \cdot 85 = 680$ уравнений второй степени с 64 неизвестными (с учётом того, что значения на входе и выходе известны).

На рисунке 9 представлено два раунда цепи Фейстеля, где $L^{(0)}_{<64>}, R^{(0)}_{<64>}$ – открытое сообщение, $L^{(2)}_{<64>}, R^{(2)}_{<64>}$ – зашифрованное сообщение, $L^{(1)}_{<64>}, R^{(1)}_{<64>}$ – промежуточное значение сообщения, $K^{(0)}, K^{(1)}, K^{(2)}, K^{(3)}$ – подключи шифрования.

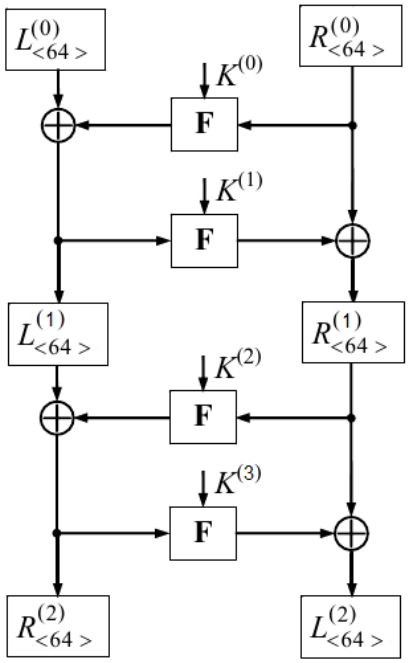

Рисунок 9 – два раунда цепи Фейстеля

Для нахождения ключей $K^{(0)}, K^{(1)}, K^{(2)}, K^{(3)}$ необходимо записать систему уравнений, которая бы описывала бы два раунда цепи Фейстеля. Как указано выше, F-функцию можно описать системой с 680 уравнениями, а значит, для описания 2 раундов цепи Фейстеля будем иметь систему с $680 \cdot 4 = 2720$ уравнениями и $4 \cdot 64 + 2 \cdot 64 = 384$ переменными. Переменными в данной системе являются $K^{(0)}, K^{(1)}, K^{(2)}, K^{(3)}$, $R^{(1)}_{<64>}, L^{(1)}_{<64>}$. Следует также отметить, что при разворачивании ключа в шифре «Лабиринт» используется F-функция, что позволяет получить дополнительные $680 \cdot 2 = 1360$ уравнения. Конечная система уравнений состоит из 4080 уравнений и 384 переменнных.

Чтобы решить систему, можно воспользоваться линеаризацией. Для ее осуществления необходимо, чтобы выполнялось условие:

$$m \le \frac{n^2}{2},$$

где m – количество уравнений;

n – количество переменных.

Для нашего случая $4080 \ge \frac{384^2}{2}$, что не удовлетворяет отмеченному условию. Однако здесь для решения нашей системы можно воспользоваться XL методом [5]. Этот метод предполагает определение параметра D, который вычисляется по формуле:

$$D \ge \frac{n}{\sqrt{m}}$$

и последующее умножение каждого из начальных уравнений на все комбинации переменных вплоть до

термов степени $D-2$. Для нашего случая $D=7$, следовательно, $D-2=5$.

В результате может быть получено необходимое количество уравнений, теоретически позволяющих выполнить линеаризацию.

Нужно также отметить, что применение в шифре цепи Фейстеля позволяет сократить количество уравнений и переменных. Для этого необходимо подать на вход шифра несколько значений открытых текстов, таких, что $R^0_{<64>}=\text{const}$, а $L^0_{<64>}$ выбирать произвольным образом, при этом получим $L^{(1)new}_{<64>}=L^{(0)new}_{<64>}\oplus L^{(0)old}_{<64>}\oplus L^{(1)old}_{<64>}$.

Это позволит увеличить количество исходных уравнений на 4080 при каждом новом входе, в то время как число переменных будет увеличиваться всего лишь на 64.

Так, если на вход модифицированного шифра подать пять значений открытых текстов, то будем иметь $n=384+64\cdot 4=640$, $m=4080\cdot 5=20400$ и $D=5$. Тогда применение XL метода позволит сократить размер системы примерно в 200 раз.

Для представления всего алгоритма системой уравнений необходимо описать начальное и конечное преобразования.

В функции IT и FT преобразований (рис. 10, рис. 11) сумма по модулю 2^{128} также заменена суммой по модулю 2 (XOR).

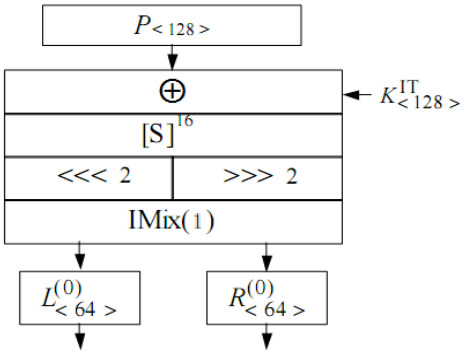

Рисунок 10 – Модифицированное начальное преобразование.

При размере блока в 128 бит преобразование IMix имеет вид:

$$L'_0=L_0\oplus\sum_0,\quad R'_0=R_0\oplus\sum_0;$$
$$\sum_0=(L_0\oplus R_0)<<<28.$$

Ввиду того, что операция циклического сдвига является линейной (переименование переменных), то преобразование IMix является линейным относительно операции XOR.

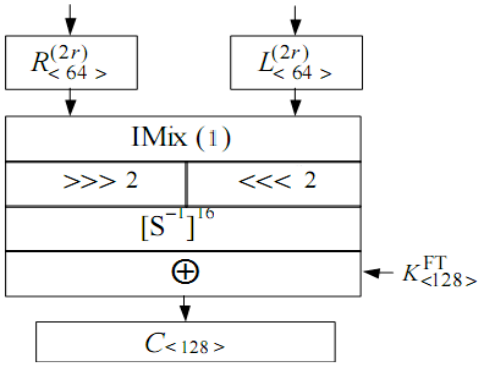

Рисунок 11 – Модифицированное конечное преобразование

Чтобы представить IT-преобразование в виде системы уравнений, необходимо связать биты открытого текста с битами выхода преобразования IMix. Пусть $P=\{p_{127}, p_{126}, \dots, p_0\}$ – открытое сообщение, $K=\{k_{127}, k_{126}, \dots, k_0\}$ – ключ K^{IT}, тогда $U=\{p_{127}\oplus k_{127}, p_{126}\oplus k_{126}, \dots, p_0\oplus k_0\}$ – сообщение на входе S-блоков.

Путь $E=\{e_{127}, e_{126}, \dots, e_0\}$ – сообщение на выходе IMix преобразования, $Q=\{q_{127}, q_{126}, \dots, q_0\}$ – сообщение на выходе подстановок S-блоков, тогда сообщение Q может быть записано как:

$$L'_0=\{e_{127},e_{126},\dots,e_{64}\};$$
$$R'_0=\{e_{63},e_{62},\dots,e_0\};$$
$$\sum_0=\left(L'_0\oplus R'_0\right)<<<28=\{e_{35}\oplus e_{99},e_{34}\oplus e_{98},\dots,e_{36}\oplus e_{100}\};$$
$$L_0=\left(L'_0\oplus\sum_0\right)>>2=\{e_{37}\oplus e_{65}\oplus e_{101},e_{36}\oplus e_{64}\oplus e_{100},\dots,$$
$$e_{38}\oplus e_{66}\oplus e_{102}\};$$
$$R_0=\left(R'_0\oplus\sum_0\right)<<2=\{e_{33}\oplus e_{61}\oplus e_{97},e_{32}\oplus e_{60}\oplus e_{96},\dots,$$
$$e_{34}\oplus e_{62}\oplus e_{98}\};$$
$$Q=L_0\,\|\,R_0.$$

Подставляя с систему уравнений, описывающую S-блок, значения после сложения с ключом и значения бит Q получаем систему, которая описывает IT-преобразование.

Для описания FT-преобразования системой уравнений применяется аналогичная процедура, как и для IT-преобразования.

В начальном и конечном преобразованиях применяется 16 S-блоков, что позволяет описать эти преобразования с помощью $2\cdot 16\cdot 85=2720$ уравнений второй степени с $2\cdot 128=256$ неизвестными. Стоит отметить, что ключ $K^{IT}=k_7\,\|\,k_4$ и $K^{FT}=k_6\,\|\,k_3$ и в дальнейшем они не будут учитываться как дополнительные неизвестные.

Выводы

Таким образом, описание всего шифра «Лабиринт» (с заменой модульного сложения операцией XOR) возможно при помощи системы второй степе-

ни, состояыщей из 16800 уравнений с 1664 неизвестными.

Как и в случае с алгебраическим анализом шифра AES/Rijndael, на текущий момент неизвестен эффективный способ решения таких систем над конечным полем. Тем не менее, поскольку из всех видов криптоанализа для AES алгебраический является одним из наиболее перспективных, в мире продолжаются исследования в данной области.

Отметим, что описанная атака в общем виде не может быть применена к исходной версии алгоритма „Лабиринт” из-за применения операции модульного сложения. Однако для достаточно большого множества аргументов операции сложения по модулю и XOR дают одинаковый результат. Использование этого свойства позволяет в перспективе применить вероятностный алгебраический анализ к шифру „Лабиринт”.

Литература: 1. Головашич С.А. Алгоритм блочного симметричного шифрования «Лабиринт». Харьков, 2007. 2. Elizabeth Kleima. The XL and XSL attacks on Baby Rijndael. Ames, Iowa 2005. 3. Лидл Р., Нидеррайтер Г. Конечные поля: В 2-ч т. Т. 1. Пер. с англ. - М.: Мир, 1988. - 430 с. 4. Белоусов А.И., Ткачёв С.Б. Дискретная математика: Учеб. для вузов / Под ред. В.С. Зарубина, А.П. Крищенко. - 3-е изд., стереотип. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2004. - 744 с. 5. N. Courtis, A. Klimov, J. Patarin, A. Shamir. Efficient Algorithms for solving Overdefined System of Multivariate Polynomial Equations. Eurocrypt'2000, Springer, LNCS 1807. 6. Nicolas T. Courtois, Josef Pieprzyk. Cryptoanalysis of Block Cipher with Overdefined System of Equations. Asiacrypt 2002, Springer, LNCS 2501. 7. http://en.wikipedia.org/wiki/Nullspace 8. Nicolas T. Courtois. Higher Order Correlation Attacks, XL Algorithm and Cryptanalysis of Toyocrypt, http://eprint.iacr.org/2002/087