

УДК 003.26 519.7

**КВАДРАТИЧНЫЕ АППРОКСИМАЦИИ СПЕЦИАЛЬНОГО ВИДА
ДЛЯ ЧЕТЫРЕХРАЗЯДНЫХ ПОДСТАНОВОК В S-БЛОКАХ¹**

Н.Н. Токарева

*Институт математики им. С.Л. Соболева СО РАН,
Новосибирский государственный университет*

E-mail: tokareva@math.nsc.ru

Рассматриваются квадратичные аппроксимации (булевых функций) специального вида и возможность применения их в нелинейном криптоанализе блочных шифров. Для четырехразрядных подстановок, рекомендованных для использования в S-блоках алгоритмов ГОСТ 28147-89, DES, s³DES, показано, что почти во всех случаях существуют более вероятные (по сравнению с линейными) квадратичные соотношения специального вида на входные и выходные биты этих подстановок.

Ключевые слова: квадратичный криптоанализ, S-блок, k-бент-функция.

Метод линейного криптоанализа (для блочного шифра FEAL) был предложен М. Мацуи и А. Ямагиши в 1992 г., для блочного шифра DES – М. Мацуи [1] в 1993 г.; в настоящее время этот метод наряду с методом дифференциального криптоанализа [2] считается одним из наиболее эффективных. Большое число работ посвящено различным обобщениям и улучшениям метода линейного криптоанализа.

Общий подход к использованию в линейном криптоанализе нелинейных аппроксимаций предложили в 1996 г. Л. Кнудсен и М. Робшау [3], но он не получил пока должного развития в силу своей общности. В этом направлении можно отметить исследования Т. Шимоямы и Т. Канеко [4], связанные с поиском квадратичных соотношений для конкретных подстановок, использующихся в S-блоках DES; экспериментальные исследования Дж. Накахары и др. [5]; работу Ж. Тапиадора и др. [6] по применению эвристических алгоритмов для поиска хороших нелинейных аппроксимаций (с примерами для S-блоков шифра MARS). Вопросы нелинейных аппроксимаций булевых функций (с использованием их приведенного представления) рассматривались также А.В. Ивановым [7].

Мы рассматриваем [8] квадратичные аппроксимации (булевых функций) специального вида и возможность применения их в нелинейном криптоанализе блочных шифров. В работе [9] (см. также [10]) для каждого целого k , $1 \leq k \leq m/2$ (m четно), была определена бинарная операция $\langle \cdot, \cdot \rangle_k : Z_2^m \times Z_2^m \rightarrow Z_2$ на множестве векторов Z_2^m , которую, исходя из ее свойств, можно считать аналогом скалярного произведения векторов над Z_2 . Определение было дано в рамках теоретико-кодowego подхода; при этом по существу использовалась классификация Z_4 -линейных кодов типа Адамара, полученная Д.С. Кротовым [11, 12]. Аналитически операцию $\langle \cdot, \cdot \rangle_k$ можно задать следующим образом:

$$\langle u, v \rangle_k = \left(\bigoplus_{i=1}^k \bigoplus_{j=i}^k (u_{2i-1} \oplus u_{2i})(u_{2j-1} \oplus u_{2j})(v_{2i-1} \oplus v_{2i})(v_{2j-1} \oplus v_{2j}) \right) \oplus \langle u, v \rangle,$$

где $\langle u, v \rangle$ – обычное скалярное произведение векторов $u = (u_1, \dots, u_m), v = (v_1, \dots, v_m) \in Z_2^m$ над Z_2 . При фиксированном векторе $u \in Z_2^m$ функция $\langle u, v \rangle_k$ от переменных v_1, \dots, v_m является линейной или квадратичной.

На основе операции $\langle \cdot, \cdot \rangle_k$ в работе [9] определяются k -преобразование Уолша – Адамара

$W_f^{(k)}(u) = \sum_{v \in Z_2^m} (-1)^{\langle u, v \rangle_k \oplus f(v)}$ булевой функции f от m переменных и k -бент-функция как функция, для

которой $W_f^{(j)}(u) = \pm 2^{m/2}$ для каждого $j = 1, \dots, k$ и любого $u \in Z_2^m$. Заметим, что 1-бент-функции и бент-функции совпадают, а с ростом параметра k «нелинейные» свойства бент-функции усиливаются (см. подробнее [9] и продолжение исследований k -бент-функций в [13]).

¹ Исследование выполнено при финансовой поддержке интеграционного проекта СО РАН № 35 «Древовидный каталог математических Интернет-ресурсов www.mathtree.ru», Российского фонда фундаментальных исследований (проекты 07-01-00248, 08-01-00671) и Фонда содействия отечественной науке.

Через Δ_m обозначим класс всех функций от $v = (v_1, \dots, v_m)$ вида $\langle u, \pi(v) \rangle_k$, где π – любая перестановка на m элементах v_1, \dots, v_m , параметры u, k произвольны. В [8] предлагается аппроксимировать неизвестные булевы функции функциями из Δ_m . Класс Δ_m состоит из 2^m (т.е. всех) линейных функций и $\sum_{k=2}^{m/2} \binom{m}{2k} 2^{m-k} (2k-1)!!$ (что не превышает числа $2^{m(1+\log m)}$) квадратичных функций от m переменных.

Введем простые обозначения, относящиеся к блочным шифрам: m, m_{key} – четные числа; $P = (p_1, \dots, p_m) \in Z_2^m$ – открытый текст; $C = (c_1, \dots, c_m) \in Z_2^m$ – шифртекст; $K = (k_1, \dots, k_{m_{key}}) \in Z_2^{m_{key}}$ – ключ шифрования; $F(P, K)$ – функция шифрования.

Нами предложены [8] модификации алгоритмов 1 и 2 линейного криптоанализа Мацуи [1]. Основная идея модификаций заключается в использовании (линейных и квадратичных) соотношений вида

$$\langle a, \pi(P) \rangle_i \oplus \langle b, \sigma(C) \rangle_j = \langle d, \tau(K) \rangle_k, \quad (1)$$

выполняющихся с некоторой вероятностью $\text{Pr} = (1/2) + \varepsilon$, где параметр ε , называемый *преобладанием* соотношения, отличен от нуля. Здесь векторы $a, b \in Z_2^m$, $d \in Z_2^{m_{key}}$, перестановки π, σ, τ и целые числа i, j, k выбираются криптоаналитиком с целью максимизировать абсолютное значение преобладания ε (задача такого характера является общей для различных методов статистического криптоанализа). Далее при неизвестном K на основе набранной статистики – пар (P, C) , где $C = F(P, K)$, – с учетом параметра ε принимается решение о том, что соотношение $\langle d, \tau(K) \rangle_k = \delta$ (для некоторого δ из Z_2) является верным. С помощью полученного соотношения проводится дальнейший анализ шифра. В [8] приведены формулы для вычисления абсолютных значений преобладаний ε (связанные с вычислением k -коэффициентов Уолша – Адамара) и для расчета надежности алгоритмов. Показано, что использование k -бент-функций в качестве функций шифрования позволяет снижать максимальное абсолютное значение преобладания до его минимального значения, а следовательно, предельно повышать стойкость шифра к рассматриваемым квадратичным аппроксимациям. Приведены свойства аппроксимирующих функций, которые могут быть использованы в квадратичном криптоанализе при согласовании нелинейных раундовых аппроксимаций (см. подробнее [8]).

Известно, что стойкость блочного шифра напрямую зависит от стойкости используемых в нем S-блоков. Целью данной работы является рассмотрение примеров четырехразрядных подстановок для S-блоков алгоритмов ГОСТ 28147-89, DES, s^3 DES и доказательство того факта, что почти во всех случаях существуют более вероятные (по сравнению с линейными) квадратичные соотношения, аналогичные (1), на входные и выходные биты этих подстановок. Результаты получены с помощью компьютера.

Четырехразрядные подстановки в S-блоках

Пример 1. В книге А.Г. Ростовцева и Е.Б. Маховенко [14] приведена серия экстремальных четырехразрядных подстановок S^1, \dots, S^{10} , рекомендованных для S-блоков стандарта ГОСТ 28147-89. Из каждой подстановки путем умножения ее на аффинные подстановки получается целый класс экстремальных подстановок. Все они были выбраны так, чтобы максимально повысить стойкость шифра к методам линейного и дифференциального криптоанализа.

Найдем наиболее вероятные квадратичные и линейные зависимости между входными и выходными битами произвольной такой подстановки, используя класс аппроксимирующих функций Δ_4 .

Число функций в классе Δ_4 равно 28. Из них 16 – линейные функции, 12 – квадратичные, которые можно перечислить следующим образом:

$$\begin{aligned} &\langle 0101, v_1 v_2 v_3 v_4 \rangle_2, \langle 0110, v_1 v_2 v_3 v_4 \rangle_2, \langle 1001, v_1 v_2 v_3 v_4 \rangle_2, \langle 1010, v_1 v_2 v_3 v_4 \rangle_2, \\ &\langle 0101, v_1 v_3 v_2 v_4 \rangle_2, \langle 0110, v_1 v_3 v_2 v_4 \rangle_2, \langle 1001, v_1 v_3 v_2 v_4 \rangle_2, \langle 1010, v_1 v_3 v_2 v_4 \rangle_2, \\ &\langle 0101, v_1 v_4 v_2 v_3 \rangle_2, \langle 0110, v_1 v_4 v_2 v_3 \rangle_2, \langle 1001, v_1 v_4 v_2 v_3 \rangle_2, \langle 1010, v_1 v_4 v_2 v_3 \rangle_2. \end{aligned}$$

Двоичному вектору $x = (x_1, x_2, x_3, x_4)$ сопоставим целое число $\tilde{x} = 8x_1 + 4x_2 + 2x_3 + x_4$ от 0 до 15. Рассмотрим соотношения $\langle a, \pi(P) \rangle_i \oplus \langle b, \sigma(C) \rangle_j = 0$, где при $i = 1$ вектору a соответствуют числа от 0 до 15 и тождественная перестановка π (что отвечает всем линейным комбинациям битов P); при $i = 2$ вектору a соответствуют числа 5, 6, 9, 10 и перестановки $\pi = \text{id}, (1324), (1342)$ (что отвечает квадратичным комбинациям битов P). Аналогично при $j = 1$ и 2 выбираем значения для b и σ . При данных условиях функции $\langle a, \pi(\cdot) \rangle_i, \langle b, \sigma(\cdot) \rangle_j$

$S^1 = (0, 13, 11, 8, 3, 6, 4, 1, 15, 2, 5, 14, 10, 12, 9, 7)$
$S^2 = (0, 1, 9, 14, 13, 11, 7, 6, 15, 2, 12, 5, 10, 4, 3, 8)$
$S^3 = (0, 1, 11, 13, 9, 14, 6, 7, 12, 5, 8, 3, 15, 2, 4, 10)$
$S^4 = (0, 1, 2, 4, 3, 5, 8, 10, 7, 9, 6, 13, 11, 14, 12, 15)$
$S^5 = (0, 1, 11, 2, 8, 6, 15, 3, 14, 10, 4, 9, 13, 5, 7, 12)$
$S^6 = (0, 1, 11, 2, 8, 3, 15, 6, 14, 10, 4, 9, 13, 5, 7, 12)$
$S^7 = (0, 4, 11, 2, 8, 6, 10, 1, 14, 15, 3, 9, 13, 5, 7, 12)$
$S^8 = (0, 4, 11, 2, 8, 3, 15, 1, 14, 10, 6, 9, 13, 5, 7, 12)$
$S^9 = (0, 11, 15, 9, 1, 5, 6, 8, 3, 10, 4, 12, 14, 13, 7, 2)$
$S^{10} = (0, 7, 10, 14, 9, 1, 13, 8, 12, 2, 11, 15, 3, 5, 4, 6)$

Как следует из табл. 1, любые линейные соотношения на входные и выходные биты S^9 выполняются с вероятностью не большей $3/4$, тогда как существуют 6 квадратичных соотношений, вероятность которых составляет $7/8$. Среди них, например, такие соотношения

$$\langle 1100, c_1 c_2 c_3 c_4 \rangle_1 \oplus \langle 0110, p_1 p_3 p_2 p_4 \rangle_2 = c_1 \oplus c_2 \oplus p_1 p_2 \oplus p_1 p_4 \oplus p_2 p_3 \oplus p_3 p_4 \oplus p_1 \oplus p_4 = 0,$$

$$\langle 1110, c_1 c_2 c_3 c_4 \rangle_1 \oplus \langle 0101, p_1 p_2 p_3 p_4 \rangle_2 = c_1 \oplus c_2 \oplus c_3 \oplus p_1 p_3 \oplus p_1 p_4 \oplus p_2 p_3 \oplus p_2 p_4 \oplus p_1 \oplus p_3 = 0,$$

которые, что интересно, линейны относительно битов шифртекста. Аналогично для S^9 можно выбрать соотношения, линейные относительно битов открытого текста. Этот пример показывает, что использование соотношений вида $\langle a, \pi(P) \rangle_i \oplus \langle b, \sigma(C) \rangle_j = 0$ в составе систем уравнений с неизвестными битами (входными или выходными) может приводить к более вероятным аппроксимациям неизвестных битов, не усложняя при этом решение системы (система по-прежнему может оставаться линейной относительно неизвестных битов).

Все подстановки S^1, \dots, S^{10} имеют параметр нелинейности NL , равный 4, тогда как параметр неквадратичности NQ для каждой из них равен 2. В табл. 2 приведены количества наиболее вероятных (квадратичных) соотношений на P и C для подстановок S^1, \dots, S^{10} и S^{11}, \dots, S^{18} (для них также $NL = 4$, см. далее). Кроме того, в табл. 2 указано, сколько среди них соотношений, линейных по битам шифртекста и по битам открытого текста (очевидно, одновременной линейности быть не может, так как все соотношения квадратичные).

Таблица 2

Число квадратичных соотношений на входные и выходные биты подстановок S^1, \dots, S^{18} , выполняющихся с вероятностью $7/8$

S : $NL(S) = 4$	Число квадр. соотн. $P_T = 7/8$	Линейных по битам C	Линейных по битам P
S^1	4	2	2
S^2	7	2	2
S^3	4	2	2
S^4	5	0	2
S^5	2	1	1
S^6	1	0	1
S^7	4	1	1
S^8	4	1	1
S^9	6	3	1
S^{10}	6	2	2
S^{11}	5	2	1
S^{12}	7	1	4
S^{13}	2	0	0
S^{14}	6	4	0
S^{15}	7	1	4
S^{16}	8	0	0
S^{17}	7	2	4
S^{18}	0	0	0

Пример 2. В книге Б. Шнайера [16] приведены восемь четырехразрядных подстановок, использовавшихся при шифровании методом ГОСТ в приложении для ЦБ РФ, а также в однонаправленной хэш-функции ГОСТ. Все они имеют параметр NL , равный 2 (кроме подстановки S^{11} , для которой $NL=4$). Для каждой подстановки имеем $NQ = 2$, и в среднем добавляется 5 – 6 новых наиболее вероятных квадратичных соотношений специального вида на входные и выходные биты каждой подстановки. Данные о подстановке S^{11} , вызывающей особый интерес, приведены в табл. 2.

$$S^{11} = (4, 10, 9, 2, 13, 8, 0, 14, 6, 11, 1, 12, 7, 15, 5, 3)$$

Пример 3. Для всех 32 подстановок на 16 элементах, используемых в S-блоках алгоритма DES (см., например, [16]), параметры NL и NQ совпадают и равны 2. Отметим, что для каждой подстановки добавляется от 0 до 11 (в среднем 4 – 5) новых, наиболее вероятных квадратичных соотношений специального вида на входные и выходные биты.

$$S^{12} = (8, 2, 11, 13, 4, 1, 14, 7, 5, 15, 0, 3, 10, 6, 9, 12)$$

$$S^{13} = (10, 5, 3, 15, 12, 9, 0, 6, 1, 2, 8, 4, 11, 14, 7, 13)$$

$$S^{14} = (5, 10, 12, 6, 0, 15, 3, 9, 8, 13, 11, 1, 7, 2, 14, 4)$$

$$S^{15} = (3, 9, 15, 0, 6, 10, 5, 12, 14, 2, 1, 7, 13, 4, 8, 11)$$

$$S^{16} = (15, 0, 10, 9, 3, 5, 4, 14, 8, 11, 1, 7, 6, 12, 13, 2)$$

$$S^{17} = (12, 6, 3, 9, 0, 5, 10, 15, 2, 13, 4, 14, 7, 11, 1, 8)$$

$$S^{18} = (13, 10, 0, 7, 3, 9, 14, 4, 2, 15, 12, 1, 5, 6, 11, 8)$$

Пример 4. Рассмотрим 32 подстановки (см., например, [16]) в S-блоках модифицированного алгоритма s^3DES [17], [18], которые считаются устойчивыми к методам дифференциального и линейного криптоанализа. Среди них только 7 подстановок (это подстановки S^{12}, \dots, S^{18}) обладают нелинейностью $NL = 4$, для 25 остальных параметр

NL равен 2. Для шести из семи подстановок с нелинейностью $NL = 4$ выполняется $NQ = 2$, и в среднем для каждой такой подстановки имеется около 6 квадратичных соотношений с вероятностью $7/8$. И лишь для одной подстановки S^{18} квадратичный криптоанализ не дает новой информации: имеем $NL = NQ = 4$.

Заключение

Мы рассмотрели несколько примеров четырехразрядных подстановок с предельно высокой нелинейностью $NL = 4$. Среди 18-ти таких подстановок лишь для одной нельзя построить более вероятные (по сравнению с линейными) квадратичные соотношения специального вида на входные и выходные биты. Для 13-ти из 17-ти остальных подстановок такие соотношения не только существуют, но среди них есть соотношения линейные относительно битов шифртекста, что по существу может использоваться при решении систем линейных уравнений, возникающих при анализе шифров.

Интересным направлением для дальнейшего исследования (и приложения в криптоанализе) является развитие подхода по использованию аналогов линейных свойств у нелинейных операций $\langle \cdot, \cdot \rangle_k$. Поясним сказанное. Если при фиксированном k всем функциям $\langle u, v \rangle_k \oplus \delta$, где $u \in Z_2^m$, $\delta \in Z_2$, сопоставить их векторы значений длины 2^m , то получится двоичный код A_m^k с параметрами кода Адамара. На этом коде достаточно просто можно определить операцию «сложения» • векторов (отличную от \oplus и зависящую от k), относительно которой код образует абелеву группу, причем эта операция согласована с метрикой Хэмминга $d(\cdot, \cdot)$: для любых $x, y \in A_m^k$ выполняется $d(x, y) = d(x \cdot y^{-1}, 0)$, где $y \cdot y^{-1} = 0$. Это обстоятельство и проявляет «линейные свойства» нелинейного скалярного произведения $\langle \cdot, \cdot \rangle_k$.

ЛИТЕРАТУРА

1. Matsui M. Linear cryptanalysis method for DES cipher // Advances in Cryptology – EUROCRYPT'93. Workshop on the theory and application of cryptographic techniques (Lofthus, Norway. May 23–27, 1993). Proc. Berlin: Springer, 1994. P. 386 – 397 (LNCS V. 765).
2. Biham E., Shamir A. Differential cryptanalysis of DES-like cryptosystems // J. Cryptology. 1991. V. 4. No. 1. P. 3 – 72.
3. Knudsen L.R., Robshaw M.J.B. Non-linear approximation in linear cryptanalysis // Advances in Cryptology – EUROCRYPT'96. Workshop on the theory and application of cryptographic techniques (Saragossa, Spain. May 12–16, 1996). Proc. Springer Verlag, 1996. P. 224 – 236 (LNCS V. 1070).
4. Shimoyama T., Kaneko T. Quadratic relation of S-box and its application to the linear attack of full round DES // Advances in Cryptology – CRYPTO'98, 18th Annual International Cryptology Conference. (Santa Barbara, California. USA. August 23 – 27, 1998). Proc. Springer, 1998. P. 200 – 211 (LNCS V. 1462).
5. Nakahara J., Preneel B., Vandewalle J. Experimental Non-Linear Cryptanalysis // COSIC Internal Report. Katholieke Universiteit Leuven. 2003. 17 p.
6. Tapiador J. M. E., Clark J. A., Hernandez-Castro J. C. Non-linear Cryptanalysis Revisited: Heuristic Search for Approximations to S-Boxes // 11th IMA International Conference (Cirencester, UK. December 18 – 20, 2007). Springer, 2007. P. 99 – 117 (LNCS V. 4887).
7. Иванов А.В. Использование приведенного представления булевых функций при построении их нелинейных аппроксимаций // Вестник ТГУ. Приложение. 2007. № 23. С. 31 – 35.
8. Токарева Н.Н. О квадратичных аппроксимациях в блочных шифрах // Пробл. передачи информ. 2008. Т. 44. № 3. С. 105 – 127.
9. Токарева Н.Н. Бент-функции с более сильными свойствами нелинейности: k -бент-функции // Дискр. анализ и исслед. операций. 2007. Сер. 1. Т. 14. № 4. С. 76 – 102.
10. Tokareva N.N. On k -bent functions // Вестник ТГУ. Приложение. 2007. № 23. С. 74 – 76.
11. Кротов Д.С. Z_4 -линейные совершенные коды // Дискрет. анализ и исслед. операций. Сер. 1. 2000. Т. 7. № 4. С. 78–90.
12. Krotov D.S. Z_4 -linear Hadamard and extended perfect codes // Proc. of the Int. Workshop on Coding and Cryptography WCC 2001, Jan. 8 – 12, 2001. Paris, France, 2001. P. 329 – 334.
13. Токарева Н.Н. Описание k -бент-функций от четырех переменных // Дискр. анализ и исслед. операций. 2008. В печати.
14. Ростовцев А.Г., Маховенко Е.Б. Введение в теорию итерированных шифров. СПб.: НПО «Мир и Семья», 2003.
15. Heys H.M., Tavares S.E. Substitution-permutation networks resistant to differential and linear cryptanalysis // J. Cryptology. 1996. V. 9. No. 1. P. 1 – 19.
16. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си. М.: Триумф, 2002.
17. Kim K., Park S., Lee S. Reconstruction of s2DES S-Boxes and their Immunity to Differential Cryptanalysis // Korea – Japan Workshop on Information Security and Cryptography. (Seoul, Korea. October 24 – 26, 1993): Proc. P. 282 – 291.
18. Biham E., Biryukov A. How to strengthen DES using existing hardware // Advances in Cryptology – ASIACRYPT '94, 4th International Conference on the Theory and Applications of Cryptology. (Wollongong, Australia. November 28 – December 1, 1994): Proc. Springer, 1995. P. 398 – 412 (LNCS V. 917).