Вісник Харківського національного університету
Серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління»
№ , 2010, с.

УДК 621.3.06

Выбор *S*-блоков для симметричных криптографических алгоритмов на основе анализа алгебраических свойств

Р. В. Олейников, А. В. Казимиров
Харьковский национальный университет радиоэлектроники, Украина

An approach to protection of symmetric cryptographic primitives from algebraic cryptanalysis is considered in the article. A proof for the maximal obtainable power of overdefined system for arbitrary size *S*-boxes is given. There are proposed parameters for *S*-boxes comparison from algebraic cryptanalysis protection point of view. *S*-boxes of modern ciphers AES/Rijndael, Camellia, Kalyna and Labyrinth are compared according to proposed criterion.

1. Общая постановка задачи и её актуальность

Симметричные криптографические примитивы получили широкое распространение благодаря высокой производительности и низкой сложности реализации [1,2]. Кроме обеспечения конфиденциальности с помощью блочных и поточных шифров, симметричные примитивы используются для обеспечения целостности на основе кодов аутентификации сообщений и хэш-функций, как компоненты электронной цифровой подписи для защиты аутентичности, генерации псевдослучайных последовательностей, в составе протоколов подтверждения подлинности [3] и т.п.

В соответствии с известными принципами Шеннона [4], такие алгоритмы используют нелинейные операции для перемешивания и линейные преобразования для рассеивания. Последовательное многократное применение перемешивания и рассеивания позволяет добиться высокого уровня криптографической стойкости.

Узлы нелинейной замены для современных симметричных примитивов, как правило, реализуют в виде таблиц замены, или *S*-блоков. Учитывая, что большинство современных блочных алгоритмов (Rijndael [5], Camellia [6], DES [7] и др.) для введения раундовых ключей и комбинирования межраундовых значений используют единственную линейную операцию (сложение по модулю 2), *S*-блоки оказываются единственным элементом, определяющим нелинейность шифрующего преобразования и уровень его стойкости к криптоаналитическим атакам. Необходимое число раундов блочных шифров вычисляется на основе обеспечения стойкости к известным видам криптографического анализа при условии заданных свойств узлов нелинейной замены.

Многие поточные алгоритмы, криптографические хэш-функции, генераторы псевдослучайных последовательностей построены на базе блочных шифров или их конструктивных элементов. Таким образом, криптографическая стойкость большинства современных симметричных примитивов в значительной степени зависит от свойств выбранных *S*-блоков.

2. Истоки исследования авторов

Проблема построения или выбора *S*-блоков с заданными свойствами является достаточно известной [1,3]. Один из первых подходов к выбору узлов нелинейной замены предполагал оценку случайности выбранных перестановок [8,9]. В дальнейшем критерии выбора были расширены для анализа свойств, позволяющих защитить шифр от статистических атак [5], в том числе от дифференциального и линейного криптоанализа [10,11], а также их модификаций. В работах [12,13] были обоснованы предельные достижимые границы вероятностей преобразований разностей и линейных аппроксимаций для *S*-блоков. В частности, эти результаты были использованы при построении шифров Rijndael, Camellia, «Лабиринт», что обеспечило этим алгоритмам максимальный уровень запаса стойкости к традиционным статистическим видам криптоанализа.

3. Нерешенные проблемы и цели работы

Тем не менее, позднее в открытой литературе был представлен новый подход, позволяющий построить переопределенную систему уравнений над конечным полем для описания всего криптографического преобразования [14], названный алгебраическим криптоанализом. Оказалось, что применение *S*-блоков с предельно достижимыми показателями свойств для защиты от статистических видов криптоанализа приводит к низкой степени переопределенной системы, описывающей весь *S*-блок. А поскольку только эти элементы определяют нелинейность многих шифров, то ряд алгоритмов, таких как AES/Rijndael, Camellia, «Лабиринт», могут быть описаны разреженной переопределенной системой всего лишь второй степени [14,15]. Практическая стойкость криптографических систем, использующих эти шифры, обеспечивается только лишь отсутствием универсальных методов, позволяющих находить решения систем нелинейных уравнений над конечными полями.

В связи с этим актуальным становится вопрос построения или отбора *S*-блоков, обеспечивающих защиту симметричных криптографических примитивов от алгебраических атак.

4. Принцип построения алгебраической атаки

Основной идеей атаки является построение и решение системы уравнений низкой степени, описывающей всё шифрующее преобразование.

Алгебраическая атака состоит из двух основных этапов:

а) формирование максимального количества уравнений, описывающих преобразование с минимальной степенью термов;

б) решение полученной системы.

Решения системы уравнений представляют собой биты ключей развёрнутого мастер-ключа (раундовые ключи) и, если вводились новые переменные, промежуточные состояния алгоритма шифрования.

Кратко опишем методику, которая позволяет представить алгоритм шифрования в виде системы уравнений:

а) декомпозиция алгоритма шифрования на составляющие операции;

б) алгебраическое описание каждого из элементов;

в) связывание входных и выходных значений каждого из элементов с другими элементами и битами ключа, открытого и шифрованного текста.

Под элементом понимаются линейные и нелинейные операции, применяемые в современных шифрах [3,1]. Декомпозиция представляет собой разбиение алгоритма шифрования на более мелкие, в большинстве случаев отдельные, части. Отдельно группируются линейные и нелинейные операции.

Под алгебраическим описанием понимается представление преобразования в виде системы, которая связывает входные и выходные значения преобразования. После представления отдельных преобразований необходимо записать общую систему, описывающую полный алгоритм шифрования и его схему разворачивания ключей.

На сегодняшний день существует достаточное большое количество методов решения уравнений над полем $GF(2)$, например, метод Гаусса, Коновальцева [16], XL, T', ElimLim, F4, F5, преобразование в SAT [17]. Однако сложность решения этих методов зависит от разреженности системы. Это позволяет сделать вывод о том, что разреженность уравнений, описывающих S-блок симметричного примитива, напрямую влияет на сложность решения конечной системы уравнений.

5. Описание метода построения системы уравнений

S-блок может быть представлен как набор булевых функций вида: $y_h = f_h\left(x_n, x_{n-1}, \ldots, x_0\right), h = 0,1,\ldots,k$. Однако существует и альтернативное описание – системой уравнений, где используются все возможные произведения комбинаций входных и выходных переменных над полем $GF(2)$. При таком представлении можно получить более низкую степень термов, чем при функциональной зависимости выхода от входа. В работе [14] показаны принципы построения такой системы для случая, когда степень каждого из термов не превышает вторую. В общем случае, для описания S-блока можно использовать произвольную степень.

Для поиска системы можно воспользоваться алгоритмом, основанным на построении матрицы, описывающей все возможные значения термов для всех вариантов комбинаций входных переменных S-блока.

Пусть X – значение, которое подаётся на вход S-блока, а Y – значение на выходе S-блока. Тогда строка матрицы A, необходимой для формирования системы уравнений, будет включать в себя все возможные сочетания битов X и Y до максимальной степени термов системы и константу 1. Так, для нахождения матрицы, описывающей S-блок, с термами не выше третей степени, строка будет содержать все возможные произведения комбинаций входных и выходных переменных третей степени, второй степени, биты входа, выхода S-блока и константу 1.

Размерность матрицы описывающей S-блок с n битами на входе и m битами на выходе равна:

$$\left|A\right| = \left(2^n\right) \times \left(N_C\right); \tag{1}$$

$$N_C = \sum_{i=0}^{d} C_{n+m}^{i}, \tag{2}$$

где C_g^h – число сочетаний из g элементов по h;

d – максимальная степень искомой системы.

Применение к этой матрице преобразования NullSpace позволяет получить новую матрицу – систему уравнений (если такая существует), описывающую S-блок.

В работе [14] показано, что необходимое количество уравнений r, при помощи которых можно описать S-блок $n \times m$, при $N_C > 2^n$ можно найти из формулы:

$$r \geq N_C - 2^n. \tag{3}$$

Однако формула (3) описывает лишь частный случай. Из теории матриц известно, что:

$$m = NullSpace(A) + Rank(A), \tag{4}$$

где m – количество столбцов матрицы A;

$Rank(A)$ – ранг матрицы A.

Таким образом, в общем виде r находится по следующей формуле:

$$r = N_C - Rank(A), \tag{5}$$

Следствие. Для того, чтобы найти минимальную степень матрицы (системы) d_{min}, при которой описываются все S-блоки $GF(2^n) \mapsto GF(2^m)$, необходимо чтобы выполнялось условие:

$$N_C > Rank(A). \tag{6}$$

Так как максимальный ранг матрицы $A = 2^n$, то d_{min} находится из соотношения:

$$\left(\sum_{i=0}^{d_{\min}} C_{2k}^i \right) > 2^k. \tag{7}$$

Приведём пример на основе подстановки «байт в байт». Для выбранного размера S-блока n = m = 8. Если d_{min} = 2, то N_c = 137, что меньше чем 256, следовательно, найдутся подстановки, которые невозможно будет описать системой уравнений второй степени. Если же взять d_{min} = 3, то N_c = 697, что больше 256 на 441. Это означает, что любой S-блок «байт в байт» может быть описан, по меньшей мере, 441 уравнениями.

В таблице 1 представлены практические расчёты для S-блоков шифров «Калина» [18], AES/Rijndael, «Лабиринт» и для двух случайно сгенерированных перестановок. Все S-блоки представляют собой подстановки 8 в 8 бит.

Как видно из таблицы, S-блоки шифров AES/Rijndael и «Лабиринт» можно представить системой уравнений второй степени, а S-блоки остальных шифров – только третьей.

Таблица 1 – Количество уравнений, описывающих S-блок

S-блок	d	$\sum_{i=0}^{d} C_8^i$	Rank(A)	r
AES/Rijndael	2	137	98	39
«Калина»/S1	3	697	256	441
«Лабиринт»	2	137	98	39
Случайный №1	3	697	256	441
Случайный №2	3	697	256	441

6. Показатели выбора S-блоков на основе анализа алгебраических свойств

Исходя из увеличения сложности выполнения алгебраического анализа, предлагается критерий, который позволяет производить выбор оптимального S-блока из некоторого набора. Применение выбранной подстановки позволит обеспечить максимальный уровень стойкости симметричного преобразования к алгебраической атаке по сравнению с остальными S-блокам.

Лучшую защиту от алгебраического анализа обеспечивает подстановка, которая:

а) имеет более высокую степень ($d_{\min}$) при описании переопределенной системой;

б) является менее разреженной (количество нулевых термов в системе меньше).

Эти два критерия не исключают полностью алгебраическую атаку, однако позволяют значительно усложнить её.

Увеличение параметра d на единицу влечёт за собой и увеличение системы на C_{n+m}^{d}, а следовательно большей производительности и требуемого объёма памяти для обработки системы уравнений. Сложность решения конечной системы прямо зависит от её разреженности, соответственно, снижение разреженности ведет к значительному усложнению криптоанализа.

7. Сравнительная характеристика *S*-блоков по алгебраическим показателям

Рассмотрим *S*-блоки распространённых симметричных блочных алгоритмов, в том числе представленных на национальный открытый конкурс [19]:

- «Калина» (*S*-блоки S1-S8, обозначенные в таблице К1-К8);
- «Лабиринт» (Л);
- Camellia (*S*-блоки S1-S4, обозначенные С1-С4);
- AES/Rijndael.

Для каждого *S*-блока были построены переопределенные системы уравнений 2-й и 3-й степени. Результаты расчёта показателей алгебраических свойств приведены в таблице 2.

Для *S*-блоков, применяемых в шифре «Калина», минимальная степень системы, которая описывает подстановку, равна 3 (теоретический максимум). Для алгоритмов «Лабиринт», Camellia и AES/Rijndael минимальная степень системы равна 2, но также возможно и построение системы 3-й степени.

Соответственно, для *S*-блоков этих шифров в разных строках таблицы 2 приведены результаты по системам 2-й и 3-й степени.

Таблица 2 – Результаты расчёта показателей алгебраических свойств S-блоков

S-блок	Всего возможно термов	Число ненул. термов в системе	Нулевые термы, %	Кол-во свобод. членов	Число термов 1-й степени	Число термов 2-й степени	Число термов 3-й степени
К1	307377	56939	0.815	219	3552	26537	26631
К2	307377	57066	0.814	214	3529	26681	26642
К3	307377	56657	0.816	230	3491	26284	26652
К4	307377	56978	0.815	230	3508	26432	26808
К5	307377	56992	0.815	217	3531	26402	26842
К6	307377	56996	0.815	227	3551	26444	26774
К7	307377	56736	0.815	221	3544	26438	26533
К8	307377	56868	0.815	225	3536	26451	26656
Л	5343	1625	0.696	16	316	1293	-
	328287	42262	0.871	202	3244	16363	22453
С1	5343	1615	0.698	18	317	1280	-
	328287	43495	0.868	239	3548	17152	22556
С2	5343	1647	0.692	16	333	1298	-
	328287	43837	0.866	242	3622	17388	22585
С3	5343	1656	0.690	22	313	1321	-
	328287	43400	0.868	203	3535	17023	22639
С4	5343	1676	0.686	15	326	1335	-
	328287	44325	0.865	220	3684	17642	22779
AES	5343	1661	0.689	16	303	1342	-
	328287	42219	0.871	194	3181	16342	22502

Как следует из таблицы, наибольший уровень защищённости симметричного криптографического алгоритма обеспечивается *S*-блоками алгоритма «Калина» (по показателями степени системы и наименьшей части нулевых термов в системе). Среди *S*-блоков шифра «Калина» лучшим с точки зрения алгебраического критерия является S2 (с минимальным отрывом от других подстановок этого алгоритма).

7. Выводы по результатам

В статье рассмотрен подход, позволяющий обеспечить защиту перспективных криптографических симметричных примитивов от алгебраического анализа. Приведено обоснование наибольшей достижимой степени переопределённой системы, описывающей табличное преобразование «*n* битов в *m* битов». В частности, доказано, что для наиболее распространённого размера *S*-блоков современных шифров («8-в-8») максимальная степень переопределённой системы равна 3. Предложены показатели, по которым возможно сравнивать различные *S*-блоки для выбора лучшего с точки зрения защиты криптографического преобразования от

алгебраической атаки. Проведенное сравнение показало преимущества *S*-блоков шифра «Калина», представленного на национальный конкурс криптографических алгоритмов, для защиты от перспективных атак на симметричные шифры.

ЛИТЕРАТУРА

1. B. Schneier, Applied Cryptography. New York: Wiley, 1996.
2. R. A. Rueppel, “Stream ciphers,” in Contemporary Cryptology: The Science of Information Integrity, G. Simmons, Ed. New York: IEEE Press, 1991.
3. A. Menezes, P. van Oorschot, and S. Vanstone, Handbook of Applied Cryptography. Boca Raton, FL: CRC Press, 1997.
4. 1. C.E. Shannon. “Communication Theory of Secrecy Systems”, Bell Syst. Tech. Journal, Vol.28, 1949.
5. J. Daemen, V. Rijmen. The design of Rijndael. AES –The Advanced Encryption Standard. Springer-Verlag, Berlin, 2002.
6. K. Aoki, T. Ichikawa, M. Kanda et al. Camellia: A 128-Bit Block Cipher Suitable for Multiple Platforms. Nessie. September 26, 2000. http://www.cryptonessie.org.
7. FIPS 46-3. Data Encryption Standard (DES). National Bureau of Standards, USA, 1993.
8. Лисицкая И.В. К вопросу построения долговременных ключей для алгоритма ГОСТ-28147-89. Информационно-управляющие системы на железно-дорожном транспорте. - 1997. - №3. С.54-57.
9. Горбенко И.Д., Лисицкая И.В. Критерии отбора случайных таблиц подстановок для алгоритма шифрования по ГОСТ 2847-89. Радиотехника. Всеукраинский Межвед. Науч. техн. сб. 1997. - Вып.103.
10. E. Biham and A. Shamir. Differential Cryptanalysis of the Data Encryption Standard. Springer-Verlag, New York, 1993. – 312p.
11. M. Matsui. Linear Cryptanalysis Method for the DES Cipher. Lecture Notes in Computer Science, Advances in Cryptology, proceedings of Eurocrypt '93, Springer-Verlag, Berlin, 1993.
12. K. Nyberg. Differentially uniform mappings for cryptography. Advances in cryptology – Eurocrypt'93, Springer-Verlag, Berlin, 1993.
13. C. Carlet, C. Ding. Highly nonlinear mappings. Journal of Complexity. Volume 20, Issues 2, Springer-Verlag, Berlin, 1998.
14. N.T.Courtois, J. Pieprzyk. Cryptanalysis of block ciphers with overdefined systems of equations. Proceedings of Asiacrypt’02, LNCS. Springer-Verlag, Berlin, 2002.
15. Олейников Р.В, Казимиров А.В. построение переопределённой системы уравнений для описания алгоритма шифрования «Лабиринт». // Прикладная радиоэлектроника. – Харьков: ХНУРЭ. – 2009. Том. 8, № 3.
16. А.В. Бабаш, Г.П. Шанкин. Криптография – М.: СОЛОН-Р, 2002. – 512 с.
17. N. Courtis, A. Klimov, J. Patarin, A. Shamir. Efficient Algorithms for solving Overdefined System of Multivariate Polynomial Equations, – Eurocrypt 2002, Springer-Verlag, Berlin, 2002.
18. І.Д. Горбенко, В.І. Долгов, Р.В. Олійников, В. І. Руженцев. Перспективний блоковий симетричний шифр “Калина” – основні положення та

специфікація. Прикладна радіоелектроніка. Тематичний випуск, присвячений проблемам забезпечення безпеки інформації. – 2007. Том 6, №2. – Харків, ХНУРЕ, 2007. С. 158-173.

19. Положення про проведення відкритого конкурсу криптографічних алгоритмів [Електронний ресурс] / ДССЗІУ. - Режим доступу : www/ URL: http://www.dstszi.gov.ua/dstszi/control/uk/publish/article?art_id=48383&cat_id=38710 - 05.06.2009 г.