2. Ru-WireGuard: Использование российских криптографических алгоритмов в протоколе безопасности сетевого уровня WireGuard. https://bit.ly/3mChAdq.
3. Reference implementation of the Ru-WireGuard protocol in Go. https://github.com/bi-zone/ruwireguard-go.

УДК 519.7 DOI 10.17223/2226308X/14/19

АЛГЕБРАИЧЕСКИЙ КРИПТОАНАЛИЗ НИЗКОРЕСУРСНЫХ ШИФРОВ SIMON И SPECK[1]

А. В. Куценко, Н. Д. Атутова, Д. А. Зюбина, Е. А. Маро, С. Д. Филиппов

Представлены алгебраические атаки на шифры Simon и Speck — два семейства низкоресурсных блочных шифров, имеющих LRX- и ARX-структуры соответственно. Они были представлены Агентством национальной безопасности США в 2013 г., а затем стандартизированы ISO как часть стандарта радиоинтерфейса RFID. Шифры алгебраически кодируются и получаемые системы булевых уравнений решаются с помощью различных SAT-решателей, а также методов, основанных на линеаризации. Впервые к этим шифрам применяются подходы, использующие разреженность систем булевых уравнений. Оценены параметры линеаризации в системах уравнений для обоих шифров. Приведено сравнение эффективности используемых методов.

Ключевые слова: *алгебраический криптоанализ, блочный шифр, низкоресурсный шифр,* Simon, Speck.

Низкоресурсная криптография — направление исследований, представляющее интерес в настоящее время. Это связано с тем, что влияние и использование RFID-меток, ПЛИС, смарт-карт, мобильных телефонов, сенсорных сетей и других криптографических алгоритмов для устройств с ограничениями на используемые ресурсы постоянно растёт и приобретает всё большую важность. Низкоресурсные криптографические примитивы предназначены для обеспечения эффективности и безопасности при ограниченном объёме ресурсов. В этом случае возникает проблема поиска компромисса между безопасностью и эффективностью. В 2013 г. Агентство национальной безопасности США представило семейства Simon и Speck низкоресурсных блочных шифров. Шифр Simon был оптимизирован для производительности на аппаратных устройствах, а Speck — для производительности в программном обеспечении. Но было подчеркнуто, что оба семейства работают исключительно хорошо как в аппаратном, так и в программном обеспечении, обеспечивая гибкость платформы, требуемую будущими приложениями. По состоянию на октябрь 2018 г. шифры Simon и Speck были стандартизированы Международной организацией по стандартизации (ISO) в рамках стандарта радиоинтерфейса RFID (радиочастотной идентификации). Эти шифры являются представителями LRX- и ARX-структур блочных шифров, основой которых является явное использование нелинейных алгебраических операций вместо S-блоков. Это обуславливает интерес к алгебраическому анализу данных шифров. Алгебраический анализ Simon проведён в [1], комбинация алгебраического и усечённого дифференциального криптоанализа шифра Simon от малого числа раундов рассмотрена в [2]. Алгебраические атаки представлены SAT-решателем и алгоритмом ElimLin.

[1]Работа выполнена в рамках госзадания ИМ СО РАН (проект № 0314-2019-0017) при поддержке лаборатории криптографии JetBrains Research; работа первого автора выполнена при поддержке РФФИ (проект № 20-31-70043).

Основная идея алгебраического криптоанализа состоит в составлении сложной системы булевых уравнений, описывающих преобразование шифра. Система строится на основе полностью известного алгоритма шифрования. Зашифрование на неизвестном криптоаналитику ключе некоторого количества открытых текстов позволяет провести подстановку в уравнения системы значений векторов открытых текстов и шифртекстов. На следующем этапе осуществляется решение системы с помощью различных методов относительно битов ключа.

Для анализа шифров было автоматизировано построение системы уравнений, описывающей преобразования раундов шифров.

Simon — семейство низкоресурсных блочных шифров, разработанных для оптимальной производительности аппаратного обеспечения [3]. Имеет структуру классической схемы Фейстеля, на каждом раунде $2n$-битный вход раунда делится на две n-битные половины. К левой половине L применяется раундовая нелинейная небиективная функция $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$. К выводу функции применяется операция XOR с правой половиной R и ключом k, и две половины меняются местами (рис. 1).

Для шифра Simon, вводя новую переменную для каждого выхода побитовой операции $\odot$, для описания T раундов получаем $n(T-2)$ квадратичных уравнений с $n(T-2)+k$ неизвестными, где n — размер слова; T — количество раундов; k — длина ключа. При генерации ключа получается $n(T-m)$ уравнений. В результате для шифра Simon с T раундами генерируется $n(T-m)+n(T-2)$ уравнений. Количество ключей m зависит от размера входного блока $2n$ и количества раундов T.

Speck — семейство низкоресурсных блочных шифров, обеспечивающих отличную производительность как в аппаратном, так и в программном обеспечении, но оптимизированных для работы на микроконтроллерах [3]. В каждом раунде $2n$-битных входа делятся на две n-битные половины. Каждый раунд Speck применяет операции конъюнкции, циклического сдвига влево и вправо, а также сложения по модулю 2^n. Параметры имеют следующие значения: $\alpha = 7$ и $\beta = 2$, если $n = 16$ (размер блока равен 32) и $\alpha = 8$ и $\beta = 3$ в противном случае. На рис. 2 представлена схема шифрования данного шифра. Ключевое расписание шифра Speck использует раундовую функцию для генерации раундовых ключей.

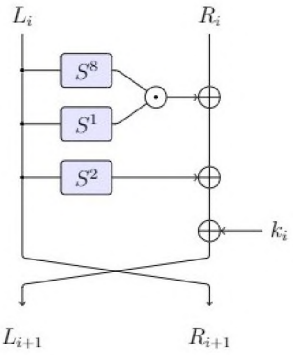

Рис. 1. Схема раундового преобразования шифра Simon [3]

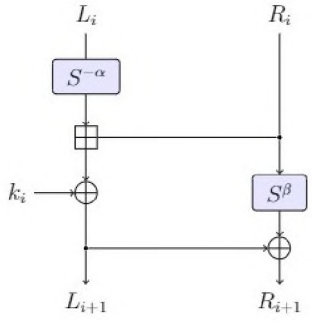

Рис. 2. Схема раундового преобразования шифра Speck [3]

В шифре Speck, вводя на каждом раунде новые переменные, получим следующие количества уравнений и неизвестных:

$$e = \begin{cases} (7n-3)(T-1) + (8n-3)(T-1) + 2n, & m = 1, \\ 2(8n-3)(T-1) + 2n, & m = 2,3,4, \end{cases}$$

$$u = \begin{cases} n(5T-4), & m = 1, \\ n(6T-5), & m = 2,3,4. \end{cases}$$

где e — число уравнений; u — число неизвестных.

1. Линеаризация

Проведение криптоанализа на небольшом количестве раундов (например, 3 и 4) с выбором стандартных характеристик нецелесообразно, так как ключи не строятся на основе исходных и между ними не будет никакой связи. Поэтому в данной работе рассматриваются шифры с $m = 1$ для $T \in \{3,4\}$.

Рассмотрены атаки, основанные на линеаризации. Идея простой линеаризации состоит в том, чтобы присвоить каждому одночлену исходной системы новую переменную. Система после переобозначения становится линейной и решается, например, методом Гаусса. Затем для решений линейной системы проверяется, являются ли они решениями исходной нелинейной системы уравнений.

Количество различных одночленов в исходной системе определяет количество переменных n' в системе линейных уравнений, эффективность линеаризации зависит от ранга r этой системы. Множество решений не пусто, его мощность равна $2^{n'-r} \geqslant 0$, поэтому для оценки производительности необходимо проанализировать границы значений n' и r.

Рассматривая алгоритм шифрования, можем оценить количество различных мономов для каждого раунда шифра Simon. Для каждой операции вводятся новые переменные и проводится переобозначение при замене L_{i+1} и R_{i+1}; в результате получаем следующую оценку количества мономов M:

$$M \leqslant 6nT.$$

В шифре Speck основным методом сохранения степени является введение новых переменных для выходных битов нелинейных операций. В этом случае степень не будет превышать 2. На каждом раунде вводится $28n$ новых переменных. В системе уравнений, описывающей сложение по модулю 2^n, имеется всего $5(7n-8)$ мономов. На практике оказалось, что различных мономов в системе уравнений сложения по модулю 2^n не больше $25n - 18$. Таким, образом, количество различных мономов на каждом раунде шифра Speck не больше $28n - 18$. Итоговая оценка числа различных мономов, исключая такие, которые образуются при генерации ключей (все уравнения линейны), имеет вид

$$M \leqslant (28n-18)T.$$

XL-атака представлена в [4, 5]. На вход поступает система из m полиномиальных уравнений от n неизвестных степени d, выбирается степень $D > d$, все уравнения исходной системы умножаются на одночлены степени $D - d$ или меньше, система линеаризуется и на выходе получаем одно или несколько решений.

Для случая $d = 2$ и $D = d + 1$ анализ этой атаки [6] показывает, что единственное решение, вероятно, будет найдено, если $m \approx n^2/6$.

Алгоритм ElimLin описан в [7]. Его суть — поиск скрытых линейных уравнений, существующих в идеале, порождённом данной системой уравнений. Этот алгоритм состоит из двух последовательных шагов:

1) исключение Гаусса: в линейной оболочке исходной системы отыскиваются все линейные уравнения;
2) замена: переменные итеративно выражаются с помощью найденных линейных уравнений, получаемые выражения подставляются в исходную систему.

В табл. 1 и 2 приведены результаты для простой линеаризации, XL-метода и ElimLin. Полученные данные позволяют сравнить эффективность этих методов для Simon и Speck. Для XL-метода $D = 3$. Сложность полного перебора составляет 2^{16} (при $n = 16$, $m = 1$). Как видно из табл. 1, метод линеаризации начиная с 4–5 раундов даёт худшие результаты, чем атака полным перебором. Использование метода простой линеаризации для $T \geqslant 4$ и XL-метода для пяти раундов (шифра Simon) не улучшает поиск решения по сравнению с полным перебором.

Таблица 1

Результаты применения атак, основанных на линеаризации

Шифр, параметры	Метод	Кол-во уравнений	Кол-во переменных	Кол-во мономов	Кол-во решений
Simon, $T = 3$, $m = 1$	Линеаризация	48	32	48	4, только одно явл-ся ключом
	XL-метод	1584	32	992	1
Simon, $T = 4$, $m = 1$	Линеаризация	64	48	80	65536
	XL-метод	3136	48	2616	256, только одно явл-ся ключом
Simon, $T = 5$, $m = 1$	Линеаризация	80	64	112	2^{32}
	XL-метод	5200	64	5008	2^{336}
Speck, $T = 3$, $m = 1$	Линеаризация	500	176	1236	—
	XL-метод	88500	176	185216	—

Таблица 2

Результаты применения метода ElimLin

Шифр, параметры	(Кол-во уравнений, кол-во лин. уравнений)	(Кол-во уравнений, кол-во лин. уравнений) после ElimLin
Simon, $T = 3$, $m = 1$	(48, 32)	(48, 32)
Simon, $T = 5$, $m = 1$	(80, 32)	(80, 48)
Speck, $T = 3$, $m = 1$	(500, 132)	(307, 137)
Speck, $T = 5$, $m = 2$	(1032, 296)	(654, 297)

2. SAT-решатели

Задача булевой выполнимости (SAT) — это задача принятия решения, в которой для произвольной булевой формулы возникает вопрос, существует ли такое значение переменных, что формула имеет значение `true`. Эта задача является NP-трудной.

Криптоанализ на основе SAT предполагает два этапа: на первом этапе обеспечивается кодирование SAT, например перевод данной системы из алгебраической нормальной формы (АНФ) в конъюнктивную нормальную форму. Мы используем конвертер anf2cnf [8] из библиотеки PolyBoRi, интегрированной в Sage. На втором этапе

полученный экземпляр SAT-задачи решается с помощью SAT-решателя. Для криптографических систем часто применяются SAT-решатели CryptoMiniSat [9] и Lingering (с его параллельными версиями Plingeling и Treengeling) [10]. Мы применяем SAT-решатели CryptoMiniSat (в Sage ver. 6.10) и Lingeling, Plingeling, Treengeling на ПК со следующими параметрами: Core i5-4690 CPU 3,5 ГГц (x4), 12 Гбайт оперативной памяти. Экспериментальные результаты для шифров Simon и Speck представлены в табл. 3 и 4. Рассмотрены два генератора систем уравнений в форме АНФ для шифра Simon: в одном все раундовые ключи являются независимыми переменными, в другом все они представлены алгоритмом ключевого расписания.

Таблица 3

Результаты SAT-решателя для шифра Simon

Параметры	Кол-во ур-ий	Кол-во неизв.	Параметры SAT	SAT	Время, с	RAM, Мбайт
$T = 8$, $m = 2$ (с раунд. ключом)	224	224	384 лит., 2528 клоз	CryptoMiniSat Plingeling Treengeling Lingeling	— 69811,9 4775,5 12702,81	— 120,5 260,3 182
$T = 8$, $m = 2$ (ключ. расписание)	128	128	368 лит., 4448 клоз	CryptoMiniSat Plingeling Treengeling Lingeling	— 845,4 1188,8 4426,12	— 26,6 169,2 95
$T = 9$, $m = 2$ (ключ. расписание)	144	144	480 лит., 6448 клоз	CryptoMiniSat Plingeling Treengeling Lingeling	— >260174,3 47799,2 24547,91	— >180,7 620,3 172
$T = 10$, $m = 2$ (ключ. расписание)	160	160	560 лит., 8096 клоз	CryptoMiniSat Plingeling Treengeling Lingeling	— — 17554,9 60776,91	— — 458,8 234

Таблица 4

Результаты SAT-решателя для шифра Speck

Параметры	Кол-во ур-ий	Кол-во неизв.	Параметры SAT	SAT	Время, с	RAM, Мбайт
$T = 3$, $m = 1$	500	176	1460 лит., 11020 клоз	CryptoMiniSat Plingeling Treengeling Lingeling	0,56 0,9 0,97 0,2	 9,6 4 1,9
$T = 4$, $m = 2$	782	320	2492 лит., 17380 клоз	CryptoMiniSat Plingeling Treengeling Lingeling	21,4 3,0 8,25 61,4	 17,3 15 14,8
$T = 5$, $m = 2$	1032	416	3312 лит., 23184 клоз	CryptoMiniSat Plingeling Treengeling Lingeling	— — 14448,17 —	— — 278 —
$T = 6$, $m = 2$	1282	512	4132 лит., 28988 клоз	CryptoMiniSat Plingeling Treengeling Lingeling	— — 123353,82 —	— — 546 —

Прочерки в таблицах означают, что SAT-решателю не удалось найти решение системы; для шифра SPECK CryptoMiniSat при $T = 3, 4$ не выдал размер файла.

3. Метод Раддума — Семаева

Данный подход к решению разреженных полиномиальных систем уравнений над полем $\mathbb{F}_2$ был представлен Г. Раддумом и И. Семаевым в работе [11]. Анализ и некоторые свойства можно найти в [12].

По исходной системе уравнений строится граф. Вершины соответствуют каждому уравнению (верхний набор вершин), также присутствуют вершины, образуемые пересечением наборов переменных соответствующих уравнений (нижний набор вершин). Каждой вершине приписан список возможных означиваний соответствующих переменных. Обработка и поиск решения осуществляется с помощью так называемой процедуры Agreeing-Gluing (согласования-склейки). Процедура согласования берет две соседние вершины и обновляет их списки, удаляя векторы, которые имеют разные подвекторы для общих переменных. Процедура склеивания заменяет две вершины новой вершиной с обновлённым списком означиваний.

В качестве результатов использования этого алгоритма для атаки на SIMON и SPECK мы приводим только максимальное количество раундов, для которых алгоритм завершился за допустимое время. Стоит отметить, что временная сложность сильно зависит от эвристики, используемой для запуска процесса согласования, будь то (частичное) разделение или склейка.

Для шифра SIMON максимальное число переменных в уравнении зависит от количества раундов и ключей. Для шести переменных количество уравнений равно $n(T-2)+n(T-m)$.

Благодаря введению новых переменных в каждый раунд шифра SPECK количество переменных на каждом раунде не зависит от количества раундов T и ключей m. Максимальное количество переменных в одном уравнении равно 6; количество уравнений и переменных представлено в табл. 5 для $m = 1$ и табл. 6 для $m = 2, 3, 4$.

Таблица 5

Количество переменных каждого уравнения шифра SPECK, $m = 1$

Кол-во переменных	Кол-во уравнений
6	$2(T-1)(2n-4)$
5	$2(T-1)n$
4	$6(T-1)n+(T-2)n$
3	$2(n+1)(T-1)$
2	$3n$

Таблица 6

Количество переменных каждого уравнения шифра SPECK, $m = 2, 3, 4$

Кол-во переменных	Кол-во уравнений
6	$2(T-1)(2n-4)$
5	$2(T-1)n$
4	$6(T-1)n+(T-2)n$
3	$2(n+1)(T-1)$
2	$(T-1)n+3n$

Алгоритм Agreeing-Gluing был запущен для Simon до 9 раундов, для Speck — до 6 (табл. 7).

Таблица 7

Параметры алгоритма Раддума — Семаева для шифров Simon и Speck

Параметры	Количество уравнений	Количество неизвестных	Верхний набор	Нижний набор
Шифр Simon				
$T = 7$, $m = 2$	112	112	112	800
$T = 8$, $m = 2$	128	128	128	1072
$T = 9$, $m = 2$	144	144	144	1600
Шифр Speck				
$T = 3$, $m = 1$	500	176	500	558
$T = 4$, $m = 2$	782	320	782	749
$T = 5$, $m = 2$	1032	416	1032	1005
$T = 6$, $m = 2$	1282	512	1282	1229

4. Анализ полученных результатов и заключение

В работе предпринята попытка оценить устойчивость шифра Speck к алгебраическому криптоанализу с помощью различных методов. Экспериментальные результаты показывают, что методы алгебраического анализа являются перспективным способом анализа надёжности современных шифров (в частности, низкоресурсных). Применительно к шифрам Simon и Speck показано, что методы, основанные на линеаризации, неэффективны уже при малом количестве раундов. С использованием SAT-решателя для шифра Simon решение найдено до 10 раундов включительно, для шифра Speck — до 6 раундов. Применение алгоритма Раддума — Семаева даёт результат для шифра Simon до 9 раундов, Speck — до 6. Результаты алгебраического анализа показывают, что включение дополнительных нелинейных операций (например, операции сложения по модулю 2^n) значительно увеличивает время атаки и объём используемой памяти. Поэтому рассмотренные методы более эффективны для криптоанализа шифра Simon, чем для Speck. В то же время разреженность систем уравнений, описывающих шифры Simon и Speck, достаточно высока, что приводит к мысли о целесообразности использования метода Раддума — Семаева, разработанного специально для таких систем.

В дальнейшем планируется провести теоретическую оценку сложности алгебраического анализа для полнораундовых шифров Simon и Speck, а также оценить эффективность использования алгоритма Бухбергера.

ЛИТЕРАТУРА

1. *Raddum H.* Algebraic analysis of the Simon block cipher family // LNCS. 2015. V. 9230. P. 157–169.
2. *Courtois N., Mourouzis T., Song G., et al.* Combined algebraic and truncated differential cryptanalysis on reduced-round Simon // 11th Intern. Conf. Security Cryptogr. 2014. P. 399–404
3. *Beaulieu R., Shors D., Smith J., et al.* The Simon and Speck Families of Lightweight Block Ciphers. Cryptology ePrint Archive, Report 2013/404, 2013.
4. *Courtois N., Shamir A., Patarin J., and Klimov A.* Efficient algorithms for solving overdefined systems of multivariate polynomial equations // LNCS. 2000. V. 1807. P. 293–407.

5. *Courtois N.* The Security of Cryptographic Primitives based on Multivariate Algebraic Problems. Ph.D. Thesis, Paris, 2001.
6. *Bard G.* Algebraic Cryptanalysis. Springer, 2009. 356 p.
7. *Courtois N. and Bard G. V.* Algebraic cryptanalysis of the data encryption standard // LNCS. 2007. V. 4887. P. 152–169.
8. *Albrecht M., Brickenstein M., and Soos M.* An ANF to CNF Converter using a Dense/Sparse Strategy. `https://doc.sagemath.org/html/en/reference/sat/sage/sat/converters/polybori.html`.
9. *Soos M.* The CryptoMiniSat 5 set of solvers at SAT competition 2016 // Proc. SAT Competition. Helsinki, 2016. P. 28.
10. *Biere A.* CaDiCaL, Lingeling, Plingeling, Treengeling, YalSAT entering the SAT Competition 2017 // Proc. SAT Competition. Helsinki, 2017. P. 14–15.
11. *Raddum H. and Semaev I.* New Technique for Solving Sparse Equation Systems. IACR Cryptology ePrint Archive, 2006/475, 2006.
12. *Biere A.* New technique for solving sparse equation systems // Des. Codes Cryptogr. 2008. V. 49. No. 1–3. P. 47–60.

УДК 512.64, 519.21, 519.72 DOI 10.17223/2226308X/14/20

К ЗАДАЧЕ ОПИСАНИЯ МИНИМАЛЬНЫХ ПО ВКЛЮЧЕНИЮ СОВЕРШЕННЫХ ШИФРОВ

Н. В. Медведева, С. С. Титов

Исследуются совершенные по Шеннону (абсолютно стойкие к атаке по шифртексту) шифры. На множестве ключей шифра определён граф эквивалентности ключей. Для шифра доказано достаточное условие его минимальности по включению. Построены примеры.

Ключевые слова: *совершенные шифры, эндоморфные шифры, неэндоморфные шифры.*

Рассматривается вероятностная модель Σ_B шифра [1]. Пусть X, Y — конечные множества соответственно шифрвеличин и шифробозначений, с которыми оперирует некоторый шифр замены, K — множество ключей, причём $|X| = \lambda$, $|Y| = \mu$, $|K| = \pi$, где $\lambda > 1$, $\mu \geqslant \lambda$. Открытые и шифрованные тексты представляются словами (ℓ-граммами, $\ell \geqslant 1$) в алфавитах X и Y соответственно. Согласно [2, 3], под *шифром* Σ_B будем понимать совокупность множеств правил зашифрования и правил расшифрования с заданными распределениями вероятностей на множествах открытых текстов и ключей. Шифры, для которых апостериорные вероятности открытых текстов совпадают с их априорными вероятностями, называются *совершенными*.

Получение строгих доказуемых оценок стойкости для каждого конкретного шифра — это очень сложная, актуальная и до конца не решённая проблема криптоанализа. Различают теоретическую и практическую стойкость шифров (оцениваемую через ресурсы, требуемые на взлом), которую естественно описывать как вероятность успеха в противодействии атакам различного вида. Здесь впечатляющим результатом представляется теорема Шеннона [1] о совершенных (абсолютно стойких к атакам по шифртексту) шифрах, которую можно (нестрого) сформулировать так: атака по шифртексту на совершенный шифр бессмысленна, так как пассивный злоумышленник, перехватив зашифрованный текст, не получает никакой информации (кроме длины сообщения) об исходном открытом тексте. Но совершенный шифр не стоек к атакам