ВЫЯСНЕНИЕ КРИПТОГРАФИЧЕСКИХ СВОЙСТВ БУЛЕВЫХ ФУНКЦИЙ ШИФРА «МАГМА»

С.Н. Шемякин, *доцент, Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, к.т.н., s4421764@yandex.ru;*
М.Э. Ахметшина, *Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, ahmet-mil@yandex.ru;*
А.И. Катасонов, *Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, ksasha716@yandex.ru;*
Р.М. Манкаев, *Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 24rasul1999@mail.ru.*

УДК 003.26

Аннотация. В данной статье описываются криптографические характеристики булевых функций, которые влияют на блок замен действующего алгоритма шифрования ГОСТ 28147-89 «Магма». Также разбирается таблица замен, определяются свойства используемых в ней булевых функций.

Ключевые слова: булевы функции; уравновешенность; алгебраическая степень; фиктивные переменные; корреляционная иммунность; *m*-оптимальность; нелинейность; лавинный критерий.

FINDING OUT THE CRYPTOGRAPHIC PROPERTIES OF BOOLEAN FUNCTIONS OF THE «MAGMA» CIPHER

Sergei Shemyakin, *Associate Professor, St. Petersburg state university of telecommunications n/a prof. M.A. Bonch-Bruevich;*
Milena Akhmetshina, *St. Petersburg state university of telecommunications n/a prof. M.A. Bonch-Bruevich;*
Alexander Katasonov, *St. Petersburg state university of telecommunications n/a prof. M.A. Bonch-Bruevich;*
Rasul Mankaev, *St. Petersburg state university of telecommunications n/a prof. M.A. Bonch-Bruevich.*

Annotation. This article describes the cryptographic characteristics of Boolean functions that affect the substitution block of the current encryption algorithm GOST 28147-89 «Magma». The substitution table is also analyzed, and the properties of Boolean functions used in it are determined.

Keywords: boolean functions; balance; algebraic degree; dummy variables; correlation immunity; *m*-optimality; non-linearity; avalanche criterion.

Введение

Булевы функции играют значительную роль при построении шифров, например, в качестве функций блоков замены. Следовательно, они должны удовлетворять ряду свойств, необходимых для построения надежной криптографической системы. Такими характеристиками функций выступают уравновешенность, алгебраическая степень, фиктивные переменные, корреляционная иммунность, *m*-оптимальность, нелинейность и строгий лавинный критерий.

Для начала введем основные определения и обозначения.

Булевы функции определяются на множестве из двух элементов, обозначающих ложность или истинность, которые обычно принимаются за значения 0 и 1. Определим булеву функцию, как функцию, аргументы которой и

сами ее значения принадлежат множеству $B \in \{0,1\}$. Булевым отображением от n переменных называется отображение $f: GF(2^n) \rightarrow GF(2^m)$ и обозначается через $f(x_1,x_2,\ldots,x_n)$. Другими словами, булева функция ставит в соответствие каждому упорядоченному набору из n числа элементов со значениями B.

Весом Хэмминга $wt(f)$ булевой функции f называется количество единиц в векторе значений этой функции. Булева функция f от n переменных называется уравновешенной, если ее вес составляет 2^{n-1}, то есть функция одинаково часто принимает значение 0 и 1 при случайном выборе аргумента x. Уравновешенные функции используются в нелинейных узлах усложнения шифров гаммирования и блоках замены блочных шифров. В первом случае обеспечиваются хорошие статистические свойства, а во втором обеспечивается взаимная однозначность дешифрования [1-3].

Полином Жегалкина – многочлен с коэффициентами из множества $B \in \{0,1\}$ над полем Галуа $GF(2)$. Для каждой функции он записывается через операции конъюнкции (&), сложения по модулю 2 ($\oplus$) и константы 1. Представление булевой функции в виде полинома Жегалкина единственно с точностью до перестановок слагаемых и порядка переменных в конъюнкциях.

Ниже в формуле 1 представлен общий вид разложения по всем переменным. Далее $\boldsymbol{x} = x_1, x_2, \ldots, x_n$ обозначает наборы, которые принимаются на вход функции, x_i - компоненты из Z_2.

$$f(x_1, x_2, \ldots, x_n) = \sum_{a_1 \ldots a_n \in Z_2^n} \oplus (x_1 \oplus a_1 \oplus 1) \cdot \ldots \cdot (x_n \oplus a_n \oplus 1) f(a_1, \ldots, a_n) \quad (1)$$

Алгебраической степенью булевой функции f называется максимальное количество переменных в самом длинном слагаемом полинома Жегалкина, имеющем ненулевой коэффициент, и обозначается $deg(f)$. Функция, степень которой не превышает 1, называется аффинной. Если в полиноме Жегалкина свободный член равен 0, то функция называется линейной. Понятно, что подобные функции нельзя использовать в нелинейных преобразованиях шифров, таких как S-блоки.

Переменная x_i в функции $f(x_1,\ldots,x_i,\ldots,x_n)$ является фиктивной или несущественной переменной, если имеются такие значения $a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_n$ переменных $x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n$, что выполняется соотношение 2.

$$f(a_1, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_n) = f(a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_n) \quad (2)$$

То есть значения функций совпадают. В противном случае x_i называется существенной переменной функции f. Следовательно, если булева функция g зависит от аргумента x_k фиктивно, то $g(x_1,\ldots,x_n,x_k) = f(x_1,\ldots,x_n)$, тогда и вес функций соотносятся как $wt(g)=2wt(f)$. При вычислениях в криптосистемах используются функции, которые существенно зависят от переменных, то есть те, в которых присутствуют только существенные переменные. В ином случае уменьшается вычислительная стойкость шифра.

Булевы функции f_1,f_2 называются статистически независимыми, если для любых $a,b \in B$ события $f_1(\boldsymbol{x})=a$ и $f_2(\boldsymbol{x})=b$ являются независимыми, то есть при случайном равновероятном выборе аргумента $\boldsymbol{x}$ из множества Z_2^n выполняется равенство 3.

$$\Pr[f_1(\boldsymbol{x}) = a, f_2(\boldsymbol{x}) = b] = \Pr[f_1(\boldsymbol{x}) = a] \cdot \Pr[f_2(\boldsymbol{x}) = b], \quad (3)$$

где: *Pr* – вероятность. Под функцией *f'* булевой функции *f* от переменных $x_1,\dots,x_n$ называется булева функция, получаемая из *f* подстановкой вместо некоторых переменных $x_{i1},\dots x_{im}$, где $0<m\leq n$ конкретных констант $a_1,\dots,a_m \in B$.

Функция *f* от *n* переменных статистически не зависит от подмножества своих переменных $x_{i1},\dots,x_{im}$, при условии $0<m\leq n$, если для любой ее подфункции *f'* имеет место равенство $Pr[f'=1]=Pr[f=1]$ и означает, что знание m входных бит не дает информации о значении функции.

Булева функция называется корреляционно-иммунной порядка m и $0<m\leq n$, если она статистически не зависит от любого *m*-элементного подмножества своих аргументов, то есть вес подфункции *f'* удовлетворяет соотношению $wt(f')=wt(f)/2^m$. Обозначается *cor(f)*. Данная характеристика показывает, что выход булевой функции, не зависит от комбинации любых *m* входов.

Согласно неравенству Зигенталера:

- если $cor(f)=m$, то выполняется $deg(f)+m \leq n$;
- если *f* – уравновешенная функция, $cor(f)=m$ и $m\leq n-2$, то выполняется $deg(f)+m \leq n-1$.

Из теоремы становится понятно, что чем выше алгебраическая степень булевой функции, тем меньше ее корреляционная иммунность и наоборот. При этом желательно, чтобы оба данных параметра были достаточно высокими, чтобы противостоять атакам на шифр.

Булева функция от *n* переменных называется *m*-оптимальной, если она является уравновешенной, степень устойчивости, которой удовлетворяет условию $m\leq n-2$, а степень ее полинома Жегалкина равна $n-m-1$. Корреляционная иммунность и связанное с ней понятие оптимальности являются криптографически значимыми свойствами булевых функций, так как позволяют обеспечивать стойкость к некоторым видам корреляционных атак.

Нелинейностью булевой функции *f* от *n* переменных называется величина N_f, которая равна расстоянию Хэмминга от функции *f* до класса аффинных функций $\mathcal{A}_n$. Множество аффинных функций обычно состоит из аффинных и линейных функций. Аффинная добавка к функции не изменит величину ее нелинейности. Расстояние между *f* и линейной функцией выражается соотношением 4.

$$N_f = 2^{n-1} - W_f(a)/2, \tag{4}$$

где: $W_f(a)$ – преобразование Уолша-Адамара – некоторая целочисленная функция, которая сопоставляется каждой булевой функции и более наглядно характеризует ее свойства, определяется уравнением 5.

$$W_f(a) = \sum_{x\in Z_2^n}(-1)^{f(x)\oplus(a,x)}, \tag{5}$$

где: *(a,x)* – скалярное произведение двух векторов длины n, принимает значение $a_1x_1 \oplus \dots \oplus a_nx_n$.

Чем больше нелинейность булевой функции, тем больше расстояние до класса линейных функций, тем соответственно больше будет устойчивость к линейному криптоанализу.

Каноническим базисом векторного пространства Z_2^n будем называть множество: векторов $e_j = (0,\dots,1,\dots,0)$, где *1* является *j*-ой координатой, $j=1,\dots,n$. Для булевой функции выполняется строгий лавинный критерий (СЛК), если для любого вектора канонического базиса e_j из Z_2^n выполняется равенство (6).

$$\left\| f(x) + f(x + e_j) \right\| = 2^{n-1} \quad (6)$$

Левая часть этого равенства показывает количество совпадений и несовпадений значений функций *f(x)* и ее значения при изменении бита *j*-ой позиции на противоположный $f(x+e_j)$. Если выполняется условие лавинности, то изменение одного бита входной последовательности функции f приведет к тому, что значение функции изменится с вероятностью *1/2* (примерно половина бит изменится, что наблюдается из самого определения), а именно реализуется один из принципов Шеннона – рассеивание. Единичная ошибка в аргументе функции тоже изменит ее значения с той же вероятностью. Для нахождения булевых функций, отвечающих строгому лавинному критерию, удобно пользоваться функцией взаимной корреляции (7).

$$\Delta_{f,g}(v) = \sum_{x \in Z_2^n} (-1)^{f(x) \oplus g(x \oplus v)}, \quad (7)$$

где: $f,g \in P_2(n)$. В нашем случае формула является автокорреляционной функцией $\Delta_{f,f}$ (АКФ), будем обозначать через Δ_f, u является каноническим вектором e_j. Функция *f* будет удовлетворять строгому лавинному критерию, если $\Delta_f(v) = 0$, при условии, что вес вектора u принимает значение *1*. Данный вывод делается из определения СЛК и АКФ: так как на выходе должна измениться половина значений функции, то количество изменившихся бит равно количеству нетронутых на выходе бит.

Функция f будет удовлетворять строгому лавинному критерию порядка *m*, если любая подфункция *f'* от *n-m* переменных удовлетворяет СЛК. Функции порядка меньше чем *m* будут также удовлетворять условию СЛК. Криптографический алгоритм, обладающий достаточным лавинным эффектом, будет более устойчив к предположениям о входной информации.

Алгоритм шифрования, описанный в ГОСТ 28147-89, до сих пор является действующим, и носит название «Магма». Схема данного алгоритма представлена на рис. 1. Он предназначен для аппаратной и программной реализации, удовлетворяет необходимым криптографическим требованиям, следовательно, его анализ является актуальной задачей [4-6].

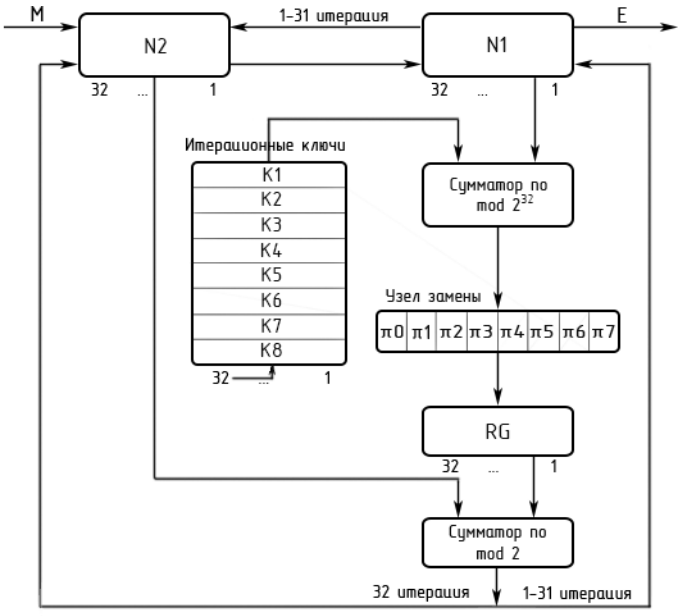

Рисунок 1

Нас будут интересовать узел замены, поэтому более подробно опишем нелинейные преобразования с использованием *S*-блоков. Другое название данной части алгоритма – табличные замены. 32-битный аргумент, после попадания в *S*-блок, разбивается на 8 последовательных векторов в каждом из которых 4 бита.

Каждый 4-битный входной вектор преобразуется с помощью узла замен в выходной вектор, также состоящий, из 4 бит. Каждый узел замены можно представить в виде таблицы-перестановки шестнадцати чисел от 0 до 15, представленных в виде двоичных векторов длины 4. Посредством данной таблицы происходят нелинейные преобразования. В конце восемь преобразованных *S*-блоками векторов последовательно соединяются в 32-битовый вектор. *S*-блок является единственным нелинейным элементом в шифре. От него будет зависеть насколько шифр сможет противостоять линейному и дифференциальному криптоанализу.

В Таблица 1 представлена сравнительная характеристика функций *S*-блока «Магма». Первые два столбца показывают распределение функций по *S*-блокам и их порядок следования в них. Далее в столбцах перечисляются свойства: уравновешенность – *bal*, алгебраическая степень – *deg*, присутствие фиктивных переменных – фикт., корреляционная иммунность – *cor*, оптимальность – *opti*, нелинейность – N_f, лавинность – СЛК. «+» или «-» обозначают наличие или присутствие названного свойства у функции, цифра показывает степень соответствующей характеристики функции. Например, функция «8*D*72», находящаяся в первом блоке, является уравновешенной, ее алгебраическая степень равна 3, фиктивные переменные отсутствуют, присутствует корреляционная иммунность 1 порядка, имеет нелинейность 4, строгий лавинный критерий отсутствует.

Таблица 1.

№*S*-	*f*	*bal*	*deg*	фикт.	*cor*	*opti*	N_f	СЛК
1	4*B*65	+	3	-	0	0	4	-
	*D*839	+	3	-	0	0	4	-
	8*D*72	+	2	-	1	1	4	-
	96*E*1	+	3	-	0	0	2	-
2	26*DA*	+	3	-	0	0	4	-
	5*A*99	+	2	-	1	1	4	-
	691*E*	+	3	-	0	0	2	-
	9*D*92	+	3	-	0	0	4	-
3	929*D*	+	3	-	0	0	4	-
	7*A*49	+	3	-	0	0	4	-
	*B*46*C*	+	3	-	0	0	4	-
	*E*61*A*	+	3	-	0	0	4	-
4	*D*962	+	3	-	0	0	2	-
	5*A*96	+	2	-	1	1	4	-
	4*CF*1	+	3	-	0	0	4	-
	9*E*58	+	3	-	0	0	4	-
5	*B*4*C*6	+	3	-	0	0	4	-
	*E*827	+	3	-	0	0	4	-
	92A*D*	+	3	-	0	0	4	-
	994*B*	+	3	-	0	0	4	-
6	964*D*	+	3	-	0	0	2	-
	2*ED*8	+	3	-	0	0	4	-
	5879	+	3	-	0	0	4	-

№*S*-	*f*	*bal*	*deg*	фикт.	*cor*	*opti*	N_f	СЛК
	1*BE*4	+	2	-	1	1	4	-
7	992*D*	+	3	-	0	0	4	-
	5*A*99	+	2	-	1	1	4	-
	8679	+	3	-	0	0	2	-
	4*B*63	+	3	-	0	0	4	-
8	2*AE*5	+	3	-	0	0	4	-
	9*C*27	+	3	-	0	0	4	-
	8771	+	3	-	0	0	4	-
	*B*16*C*	+	3	-	0	0	4	-

По результатам исследования все представленные функции для «Магма» удовлетворяют условию уравновешенности, в них отсутствуют фиктивные переменные, однако 5 функций имеют небольшую нелинейность, как видно, в таблице есть функции с другими аналогичными свойствами, которые имеют нелинейность 4. Следовательно, можно найти похожие функции с более хорошими характеристиками. Из всех 32 функций только у 5-ти алгебраическая степень равна 2, зато у этих функций присутствует оптимальность и корреляционная иммунность порядка 1, что очевидно является плюсом. Ни одна функция не удовлетворяет строгому лавинному критерию, что ослабляет шифр.

Одна функция «5*A*99» дважды используется, во 2 и в 7 *S*-блоках, имеет преимущества в виде корреляционной иммунности и оптимальности, нелинейность 4 порядка.

В случае для таблицы выше предположительно сначала строились блоки с лучшими характеристиками, так как в каждом *S*-блоке присутствует либо функция, обладающая корреляционной иммунностью и оптимальностью 1 порядка, либо все функции блока имеют относительно большую нелинейность. Для построения более надежной системы следует использовать в составе *S*-блока функции, обладающие несколькими положительными свойствами, рекомендуется использование лавинных функций.

Заключение

Были определены характеристики булевых функций, которые входят в состав нелинейного элемента действующего шифра «Магма», влияющего на стойкость шифра. Перечислены характеристики функций действующей таблицы замен, на ее примере описаны возможные недостатки, в том числе приведены рекомендации.

Литература

1. Коржик В.И., Яковлев В.А. Основы криптографии: учебное пособие – СПб., ИЦ Интермедия, 2016. – 296 с.
2. Панкратова И.А. Булевы функции в криптографии: учебное пособие – Томск: Издательский Дом Томского государственного университета, 2014. – 88 с.
3. Коржик В.И., Просихин В.П. Основы криптографии: учебное пособие по специальности 210403 "Защищенные телекоммуникационные системы связи". – Санкт-Петербург, 2008.
4. Тихонов С.В., Коржик В.И. Метод защиты аппаратной реализации шифра гост от атаки измерения потребляемой мощности в цепи питания // Проблемы информационной безопасности. Компьютерные системы, 2013. – № 3. – С. 62-72.
5. Лось А.Б., Нестеренко А. Ю., Рожков М. И. Криптографические методы защиты информации для изучающих компьютерную безопасность: учебник для

академического бакалавриата // 2-е изд., испр. – М.: Издательство Юрайт, 2019. – 473 с.
6. Коржик В.И., Тихонов С.В. О возможности взлома аппаратной реализации шифра гост // Проблемы информационной безопасности. Компьютерные системы, 2012. – № 3. – С. 53-62.