УДК 681.3.06

Р.В. Сергиенко[1], И.В. Московченко[2]

[1] *Львовский институт Сухопутных войск НУ «Львовская политехника», Львов*
[2]*Харьковский институт танковых войск НТУ «ХПИ», Харьков*

ИССЛЕДОВАНИЕ КРИПТОГРАФИЧЕСКИХ СВОЙСТВ НЕЛИНЕЙНЫХ УЗЛОВ ЗАМЕН АЛГОРИТМА СИММЕТРИЧНОГО ШИФРОВАНИЯ ГОСТ 28147-89

Рассматривается алгоритм блочного симметричного шифрования ГОСТ 28147-89. С использованием математического аппарата булевой алгебры, методов корреляционного и спектрального анализа исследуются криптографические свойства нелинейных узлов замен алгоритма ГОСТ 28147-89.

алгоритм блочного симметричного шифрования

Введение

Постановка проблемы в общем виде и анализ литературы. В середине 1970-х годов стало ясно, что в криптографической защите информации нуждаются не только специальные структуры, но и различные промышленные, информационные и другие организации, а также частные лица.

По заказу Национального совета безопасности и Национального бюро стандартизации США корпорацией IBM был разработан и в 1977 году принят в качестве национального стандарт шифрования DES (Data Encryption Standard.) [1]. Еще при принятии DES в качестве национального стандарта шифрования возникли некоторые возражения – почти все они касались длины ключа, но некоторые исследователи также выражали озабоченность по поводу таблиц подстановок (нелинейных узлов замен, S-блоков) DES. Имеются ли в них «ловушки», секретная информация, которая позволила бы тому, кто ею владеет, легко осуществлять расшифрование, даже без знания ключа? Но немного информации об S-блоках появилось до того, как два израильских исследователя Эли Бихам и Ади Шамир, открыли дифференциальный криптанализ, – атаку, которая использует нелинейность S-блоков DES [2 – 4]. Используя атаку выбранного открытого текста, Бихам и Шамир получали информацию о битах ключа, использованных в некоторой сессии DES-зашифрования. Они доказали, что ключ может быть раскрыт атакой выбранного открытого текста 2^{47} зашифрованиями. В 1993 году Мицуру Мацуи открыл линейный криптоанализ, который использует линейные отношения между входными и выходными битами DES [5]. Он показал, что DES может быть взломан атакой открытого текста за 2^{43} зашифрований. Таким образом, актуальным является вопрос, насколько эффективен S-блок шифра, как хорошо он скрывает статистические закономерности и минимизирует корреляцию между входными и выходными векторами данных.

Разработка советского стандарта шифрования ГОСТ 28147-89 была своего рода ответом на принятие в качестве стандарта шифрования США алгоритма шифрования DES [6]. ГОСТ, также как и DES, использует итеративную схему Файстеля, но имеет бо́льшую длину ключа (256 бит) и бо́льшее количество раундов (32) [7 – 9]. Кроме того, ГОСТ 28147-89 использует секретные таблицы подстановок. Эти отличия, направленные на «экстенсивное» повышение надежности шифра, вероятно, и послужили причиной недостаточного внимания криптоаналитиков к анализу шифра ГОСТ 28147-89. **Целью данной статьи** является исследование криптографических свойств S-блоков шифра ГОСТ 28147-89, используемых в некоторых финансовых учреждениях Российской Федерации и которые недавно стали известны [9].

Методика исследований

Основными показателями эффективности нелинейных преобразований являются: сбалансированность выходной последовательности, нелинейность преобразования, корреляционный иммунитет, критерий распространения, алгебраическая степень функции. Для аналитического описания и исследования криптографических свойств S-блоков шифров в ряде работ предложен математический аппарат булевых функций [10 – 13]. Введем основные понятия и определения.

Булевой функцией f от n переменных является функция, осуществляющая отображение из поля $GF(2^n)$ всех двоичных векторов $x = (x_1, \dots, x_n)$ длины n в поле GF(2). Обычно булевы функции представляются в алгебраической нормальной форме и рассматриваются как сумма произведений составляющих координат. Поле $GF(2^n)$ состоит из 2^n векторов α_i: $\alpha_0 = (0,\dots,0,0)$, $\alpha_1 = (0,...,0,1),...,$ $\alpha_{2^n-1} = (1,...,1,1)$, $\alpha_i \in V_n$, где V_n – векторное пространство в $GF(2^n)$.

Последовательность функции f является сбалансированной, если ее (0,1)-последовательность ((1,–1)-последовательность) содержит одинаковое количество нулей и единиц (единиц и минус единиц). Функция f является сбалансированной, если сбалансирована ее последовательность. Сбалансированные функции наиболее стойки к прямым статистическим атакам. Эквивалентное определение сбалансированности [10 – 13]: функция f над $GF(2^n)$ является сбалансированной, если ее выходные значения являются равновероятными:

$$| \{x \mid f(x) = 0 \} | = | \{x \mid f(x) = 1 \}| = 2^{n-1}.$$

Аффинной функцией f называется функция вида $f = a_1x_1 \oplus \dots \oplus a_nx_n \oplus c$, где a_j, $c \in GF(2)$, $j = 1, 2,..., n$. Функция f называется линейной, если $c = 0$.

Весом Хэмминга вектора α ((0,1)-последовательности α), обозначаемым как $W(\alpha)$, является количество единиц в векторе (последовательности).

Расстоянием Хэмминга d(f,g) между последовательностями двух функций f и g является количество позиций, в которых различны последовательности этих функций.

Нелинейность функции N_f – минимальное расстояние Хэмминга N_f между функцией f и всеми аффинными функциями над $GF(2^n)$:

$$N_f = \min \{d(f,\varphi)\},$$

где φ – множество аффинных функций.

Для произвольной функции f нелинейность N_f над $GF(2^n)$ может достигать:

$$N_f \le 2^{n-1} - 2^{n/2-1}.$$

Функция f обладает *корреляционным иммунитетом* порядка k, если выходная последовательность функции $y \in Y$ статистически не зависит от любого подмножества из k входных координат:

$$\forall\{x_1, \dots, x_k\} \qquad P(y \in Y / \{x_1, \dots, x_k\} \in X) = P(y \in Y).$$

Эквивалентное определение корреляционного иммунитета в терминах преобразования Уолша [10 – 13]: функция f над полем $GF(2^n)$ имеет корреляционный иммунитет порядка k, КИ(k), если ее преобразование Уолша удовлетворяет равенству $F(\omega) = 0$ для всех $\omega \in V_n$ таких, что $1 \le W(\omega) \le k$:

$$\forall\omega \in V_n \qquad F(\omega) = 0 \qquad КИ(f) = k.$$

Корреляционно-эффективной функцией является функция, для которой не меньше чем на поло-

вине векторов веса $1 \le w \le q$ значения компонентов спектра ПУ равны 0.

Преобразование Уолша $F(\omega)$ функции f над полем $GF(2^n)$ определяется как принимающая действительные значения функция

$$F(\omega) = 2^{-n} \sum_x (-1)^{f(x) \oplus \langle \omega, x \rangle},$$

где $\omega \in V_n$, $f(x)$, $\langle \omega, x \rangle \in N$ ($\langle \omega, x \rangle$ – скалярное произведение $w_1x_1 \oplus \ldots \oplus w_nx_n$).

Функция f над полем $GF(2^n)$ удовлетворяет:

– *критерию распространения относительно вектора* α, $KP(\alpha)$, если функция $f(x) \oplus f(x \oplus \alpha)$ является сбалансированной, $x \in V_n$, где $x = (x_1, x_2, \ldots, x_n)$:

$$P(f(x) = f(x \oplus \alpha)) = 0{,}5;$$

– *критерию распространения степени* k, KP(k), если удовлетворяется критерий распространения относительно всех векторов $\alpha \in V_n$ при $1 \le W(\alpha) \le k$:

$$P(f(x) = f(x \oplus \alpha)) = 0{,}5 \quad \forall \alpha : 1 \le W(\alpha) \le k.$$

Алгебраическая степень deg(f) является степенью самого длинного слагаемого функции, представленной в алгебраической нормальной форме.

Коэффициент корреляции функции со множеством всех аффинных функций определяется как

$$c_i(f, L_w) = 2^{-n} \sum_x (-1)^{f(x)} (-1)^{wx} = 2^{-n} \hat{F}(w).$$

Функция f над $GF(2^n)$ называется бент-функцией, если $2^{-n/2} \sum_{x \in V_n} (-1)^{f(x) \oplus \langle \beta, x \rangle} = \pm 1$ для всех $\beta \in V_n$.

Результаты исследований

Каждый из 8 S-блоков ГОСТ 28147-89 отображает входной 4-битный вектор в выходной вектор такой же размерности. Другими словами, этот узел подстановки представляет собой таблицу $2^4 = 16$ строк, которые включают по 4 бита заполнения. Входной вектор представляет адрес строки в таблице замен, а заполнение этой строки – выходной 4-битный вектор. Таким образом, S-блок можно представить объединением четырех компонентных булевых функций от четырех переменных $f_i(x_1, x_2, x_3, x_4)$, затем исследовать криптографические и статистические свойства этих функций. Проиллюстрируем первый S-блок шифра ГОСТ 28147-89 (табл. 1). Проанализируем первую булеву функцию F_1 первого S-блока. Исследуемая функция имеет равное количество нулей и единиц: по восемь, поэтому функция сбалансирована.

Для определения других необходимых параметров функции (количества термов функции, количества термов, содержащих каждую переменную, алгебраической степени функции, алгебраической степени каждой переменной и др.) необходимо восстановить алгебраическую нормальную форму функции (АНФ). Для этого составим таблицы истинности каждого терма, и затем для нахождения коэффициентов a_i при каждом терме решим систему уравнений вида:

Таблица 1

Таблица истинности компонентных булевых функций первого S-блока алгоритма ГОСТ28147-89

Sb1	4	10	9	2	13	8	0	14	6	11	1	12	7	15	5	3
F_1	0	0	1	0	1	0	0	0	0	1	1	0	1	1	1	1
F_2	0	1	0	1	0	0	0	1	1	1	0	0	1	1	0	1
F_3	1	0	0	0	1	0	0	1	1	0	0	1	1	1	1	0
F_4	0	1	1	0	1	1	0	1	0	1	0	1	0	1	0	0

$$\sum_{i=0}^{15} a_i T_i = F_i(X),$$

где T_i – i-й одночлен АНФ функции.

В результате получим:

$$F_1(X) = x_2 + x_3 + x_1x_2 + x_1x_3 + x_1x_4 + x_1x_2x_4 + x_2x_3x_4.$$

Из полученной АНФ функции простым подсчетом мы определяем следующие криптографически важные свойства функции:

– количество термов в функции: $term(f) = 7$;

– количество термов в функции, содержащих определенную переменную:

$$term_{x_1}(f) = 4,\ term_{x_2}(f) = 4,$$

$$term_{x_3}(f) = 3,\ term_{x_4}(f) = 3;$$

– алгебраическая степень функции (нелинейный порядок функции), $deg(f) = 3$;

– алгебраическая степень каждой переменной: $deg(f, x_i) = 3$ для всех i.

Перечисленные выше показатели характеризуют стойкость S-блока и шифра в целом к алгебраическим атакам, которые используют недостаточную сложность математического описания шифра.

Для обеспечения криптографической стойкости шифра блоки замен должны быть высоконелинейными. Построим таблицу истинности всех возможных аффинных функций, и определим расстояние Хемминга до каждой из этих функций. Минимальное расстояние равно 4, следовательно, значение нелинейности функции равно 4.

Существенным усилением свойства сбалансированности БФ является требование сбалансированности всех частных функций, полученных из исходной функции фиксированием любых ее *k* или менее переменных. Указанное требование позволяет обеспечить стойкость криптографических преобразований к статистическим атакам при фиксированных значениях битов на входе преобразования. Данное свойство связано с показателем корреляционной имунности (КИ).

Корреляционно-эффективная функция – функция, для которой не меньше чем на половине векторов веса $1 \le w \le q$ значения компонентов спектра ПУ равны 0.

Определим значения $F(\omega)$ для всех возможных значений ω. Результаты сведем в табл. 2.

Таблица 2

Значения F(ω) для всех возможных значений ω

ω	<x,ω>	F(ω)	ω	<x,ω>	F(ω)	ω	<x,ω>	F(ω)
(1,0,0,0)	x_1	–4	(1,1,0,0)	$x_1 + x_2$	4	(1,1,1,0)	$x_1+x_2+x_3$	8
(0,1,0,0)	x_2	0	(1,0,1,0)	$x_1 + x_3$	0	(1,1,0,1)	$x_1+x_2+x_4$	–4
(0,0,1,0)	x_3	4	(1,0,0,1)	$x_1 + x_4$	-4	(1,0,1,1)	$x_1+x_3+x_4$	0
(0,0,0,1)	x_4	8	(0,1,1,0)	$x_2 + x_3$	4	(0,1,1,1)	$x_2+x_3+x_4$	4
			(0,1,0,1)	$x_2 + x_4$	0	(1,1,1,1)	$x_1+x_2+x_3+x_4$	0
			(0,0,1,1)	$x_3 + x_4$	-4			

Из анализа таблицы видно, что функция не является корреляционно-имунной, и даже не является корреляционно-эффективной, так как она имеет всего лишь 5 нулевых значений (≈31%).

При синтезе БФ, обеспечивающих высокую стойкость к дифференциальному, линейному и корреляционному криптоанализу, большое значение имеет автокорреляционная функция (АКФ) БФ. Определим значения АКФ для всех возможных значений s. Результаты сведем в табл. 3.

Очевидно, автокорреляция функции равна 8.

По рассмотренной выше методике проведены исследования криптографических свойств компонентных булевых функций, составляющих блоки замен шифра ГОСТ 28147-89. Все компонентные функции блоков замен являются сбалансированными,

Таблица 3

Значения АКФ

s	АКФ	s	АКФ	s	АКФ
(1,0,0,0)	0	(1,1,0,0)	8	(1,1,1,0)	0
(0,1,0,0)	0	(1,0,1,0)	0	(1,1,0,1)	0
(0,0,1,0)	0	(1,0,0,1)	8	(1,0,1,1)	0
(0,0,0,1)	0	(0,1,1,0)	–8	(0,1,1,1)	0
		(0,1,0,1)	–8	(1,1,1,1)	–8
		(0,0,1,1)	–8		

имеют высокую алгебраическую степень 3, подавляющее большинство функций (94%) имеют нелинейность 4 и более. Тем не менее, практически все функции не имеют корреляционной имунности и не удовлетворяют строгому лавинному критерию. Основные результаты исследования сведены в табл. 4.

Таблица 4

Криптографические свойства компонентных булевых функций блоков замен шфира ГОСТ

	S-блок 1				S-блок 2				S-блок 3				S-блок 4			
	f_1	f_2	f_3	f_4	f_1	f_2	f_3	f_4	f_1	f_2	f_3	f_4	f_1	f_2	f_3	f_4
Сбалансирован.	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Нелинейность N_f	4	4	4	4	4	6	6	6	4	4	6	4	6	6	6	2
Порядок КИ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Корреляц. эф-сть	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+
КЭ 1-го поряд.	–	–	–	–	–	+	–	–	–	–	–	+	–	+	+	+
Порядок КР	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0
AC(f)	8	8	8	8	8	8	8	8	8	8	8	8	8	16	8	16
Кол-во термов	7	7	8	8	5	6	8	9	9	9	10	7	9	7	6	4
Кол-во термов с x_1	4	5	5	4	4	2	5	3	4	3	4	4	4	3	1	2
Кол-во термов с x_2	4	2	5	3	4	3	3	3	3	5	4	3	5	2	4	2
Кол-во термов с x_3	3	3	3	5	2	4	2	5	4	6	5	3	3	5	3	1
Кол-во термов с x_4	3	4	3	3	1	1	4	4	5	6	4	3	4	2	1	1
Алг. степень deg(f)	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
$deg(f_{x1})$	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
$deg(f_{x2})$	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
$deg(f_{x3})$	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
$deg(f_{x4})$	3	3	3	3	3	2	3	3	3	3	3	3	3	3	2	2
Сбалансирован.	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Нелинейность N_f	4	6	4	4	4	2	6	4	6	4	4	4	6	4	4	4
Порядок КИ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Корреляц. эф-сть	–	–	–	–	–	+	–	–	+	–	–	–	–	–	–	–
КЭ 1-го поряд.	–	+	–	+	+	+	–	–	+	+	+	+	–	–	–	+
Порядок КР	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0
AC(f)	8	8	8	8	8	16	8	8	16	8	8	8	8	8	8	8
Кол-во термов	8	7	8	6	8	7	9	9	8	8	9	9	8	8	9	7
Кол-во термов с x_1	3	2	4	3	3	4	4	6	2	5	3	3	4	4	5	4
Кол-во термов с x_2	5	2	4	4	4	1	6	4	6	4	3	2	5	5	5	2
Кол-во термов с x_3	6	4	3	3	3	3	2	5	3	5	4	5	4	5	4	4
Кол-во термов с x_4	4	3	6	3	5	3	4	4	3	2	5	4	4	4	4	3
Алг. степень deg(f)	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3

Окончание табл. 4

	S-блок 1				S-блок 2				S-блок 3				S-блок 4			
	f_1	f_2	f_3	f_4	f_1	f_2	f_3	f_4	f_1	f_2	f_3	f_4	f_1	f_2	f_3	f_4
$\deg(f_{x1})$	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
$\deg(f_{x2})$	3	3	3	3	3	1	3	3	3	3	3	2	3	3	3	3
$\deg(f_{x3})$	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
$\deg(f_{x4})$	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3

Выводы

На основании полученных результатов можно сделать следующие выводы:

1. Отдельные блоки замен шифра ГОСТ 28147-89 имеют недостаточно высокие показатели, в частности в четвертом и шестом блоках по одной компонентной БФ имеют показатель нелинейности 2, что снижает также показатель нелинейности до 2 блока замен в целом [14]. Кроме того, все БФ не имеют корреляционной иммунности, и только три компонентные БФ корреляционно эффективны. Это обосновывается тем, что существует взаимосвязь между алгебраической степенью функции и порядком k ее корреляционного иммунитета [10 – 13]: для сбалансированных функций: $k + \deg(f) \le m - 1$. В данном случае $\deg(f) = 3$, $m = 4$, следовательно $k = 0$.

2. Для более «чувствительного» подхода к оценке корреляционного иммунитета БФ малого количества переменных (до 4) целесообразно ввести показатель *корреляционной эффективности 1-го порядка*. Из табл. 4 видно, что 14 (44%) компонентных БФ удовлетворяют этому показателю. Кроме того, они рассредоточены по блокам замен неравномерно, что снижает статистические свойства шифра. Следовательно, целесообразным было бы заменить функции, имеющие низкую нелинейность (2) высоко нелинейными функциями, удовлетворяющими показателю корреляционной эффективности 1-го порядка и расположить их так, чтобы в каждом блоке замен присутствовало по 2 функции с корреляционной эффективностью 1-го порядка в различных позициях (1…4). Во второй функции 6 блока переменная x_2 имеет низкую алгебраическую степень 1, что может создать предпосылки для алгебраической атаки шифра, которая может быть спровоцирована довольно простой процедурой расширения ключа.

3. Применение блоков замен, которые удовлетворяют основным криптографически важным показателям, позволило бы отказаться от их засекречивания во многих случаях, так как данный шифр (ГОСТ 28147-89) имеет достаточную длину ключа (256 бит). С другой стороны, наличие «несовершенных» блоков замен можно объяснить тем, что эти блоки могут быть секретными, то есть изменяться между сеансами шифрования. Стремление к использованию блока замен с очень хорошими показателями (которых не так уж много) может значительно сузить варианты перебора S-блоков.

Список литературы

1. National Institute of Standards and Technology, "FIPS-46-3: Data Encryption Standard." Oct. 1999.

2. Biham E. and Shamir A. Differential cryptanalysis of DES-like cryptosystems // Journal of Cryptology. – 1991. – Vol. 4, No. 1. – P. 3-72.

3. Biham E. New types of cryptanalytic attacks using related keys // Advances in Cryptology, Proceedings Eurocrypt'93, LNCS 765, T. Helleseth, Ed., Springer-Verlag. – 1993. – P. 398-409.

4. Biham E., Dunkelman O., Keller N. Enhancing differential-linear cryptanalysis. – NESSIE, 2002. NES/DOC/TEC/WP5/017.

5. Matsui M. Linear cryptanalysis method for DES cipher // Advances in Cryptology, Proc. Eurocrypt'93, LNCS 765, T. Helleseth, Ed., Springer-Verlag. – 1994. – P. 386-397.

6. ГОСТ 28147-89. Системы обработки информации. Защита криптографическая. Алгоритм криптографического преобразования. – М.: Изд-во стандартов, 1989. – 20 с.

7. Винокуров А., Применко Э. Сравнение российского стандарта шифрования, алгоритма ГОСТ 28147-89, и алгоритма Rijndael, выбранного в качестве нового стандарта шифрования США // Системы безопасности. – М.: Гротэк. – 2001. – №№ 1, 2.

8. Винокуров А.. ГОСТ не прост, а очень прост. – М.: Монитор. – 1995. – № 1. – С. 60-73.

9. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке СИ. – М.: Триумф, 2003. – 816 с.

10. Maier W., Staffelbach O. Nonlinearity criteria for cryptographic functions // In Advances in Cryptology – EUROCRYPT'89, vol.434, Lecture Notes in Computer Science, Springer-Verlag. – 1990. – P. 549-562.

11. .Maitra S, Pasalic E. Further constructions of resilient Boolean functions with very high nonlinearity // Accepted in SETA, May, 2001, Norway.

12. Pasalic E., Johansson T., Maitra S., Sarkar P. New constructions of resilient and correlation immune Boolean functions achieving upper bounds of nonlinearity // In Workshop of Coding and Cryptography, Electronic Notes in Discrete Mathematics. Elsevier, January 2001.

13. Millan W., Clark A., Dawson E. Heuristic Design of Cryptographically Strong Balanced Boolean Functions // In Advances in Cryptology EUROCRYPT'98, Springer Verlag LNCS 1403. – 1998. – P. 489-499.

14. Pieprzyk J. On Bent Permutations, Technical report SC91/11, Department of computer science, University of New South Wales; presented on International Conference, Las Vegas 1991.

Поступила в редколлегию 3.08.2007

Рецензент: д-р техн. наук, проф. Ю.В. Стасев, Харьковский университет Воздушных Сил им. И. Кожедуба, Харьков.