МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ

DOI 10.17223/20710410/6/5

УДК 519.7

ДИФФЕРЕНЦИАЛЬНЫЙ КРИПТОАНАЛИЗ БЛОЧНОГО ШИФРА MARS

А. И. Пестунов

Институт вычислительных технологий СО РАН, г. Новосибирск, Россия

E-mail: pestunov@gmail.com

Предлагается дифференциальная атака на шифр MARS, который является финалистом конкурса AES. Эта атака является более эффективной, чем ранее известные, и позволяет провести криптоанализ урезанной версии шифра MARS, использующей 752 бита подключей, в то время как лучшая из ранее известных атак позволяет провести криптоанализ урезанной версии шифра MARS, использующей только 682 бита подключей. Вероятность успеха предлагаемой атаки составляет более 99 %, а ее сложность меньше сложности полного перебора ключей.

Ключевые слова: *блочный шифр, дифференциальный криптоанализ, Advanced Encryption Standard, MARS.*

Введение

Блочные шифры являются важным элементом в современных системах защиты информации, поэтому в последние годы имели место несколько проектов, направленных на их исследование [1 – 3]. Одним из наиболее влиятельных проектов является Advanced Encryption Standard (AES) [1]. Типичный блочный шифр преобразует открытый текст, представленный в виде последовательности битов (нулей и единиц), по блокам фиксированной длины s, чаще всего s равно 64 или 128 бит. Секретный ключ также представляет собой битовую последовательность (обычно длиной 128 или 256), из которой по определенному для каждого шифра механизму получается набор из R *подключей*. Процедура шифрования носит итеративный характер и заключается в R-кратном выполнении некоторого относительно «слабого» преобразования, которое зависит от своего подключа и называется *раундом* шифрования. Раунды могут быть как одинаковыми, так и различными. Чем больше их выполнено, тем надежней, но медленней становится шифр, поэтому его создатель определяет такое количество раундов, которое обеспечивает и безопасность, и быстродействие алгоритма.

Исследовать безопасность криптоалгоритмов призван раздел криптологии, называемый *криптоанализом*. Криптоанализ блочных шифров подобен соревнованию между криптоаналитиками, которые стараются найти неизвестный секретный ключ для упрощенных версий шифра, по возможности более близких к исходному варианту. Например, для шифра, состоящего из 20 раундов, криптоанализ 12 раундов считается продвижением в анализе шифра, если ранее удавалось найти ключ только для версии шифра, состоящей из 10 раундов. Заметим, что знание всех подключей блочного шифра эквивалентно знанию секретного ключа, поскольку в шифровании участвуют именно подключи, а секретный ключ используется только для их вычисления. Алгоритм нахождения ключа или подключей называется *атакой*. Даже если атака такова,

что она не осуществима на практике, но ее сложность меньше сложности полного перебора ключей, это является важным сертификационным недостатком шифра [4].

На сегодняшний день не существует цельной теории, в рамках которой можно было бы гарантировать безопасность того или иного блочного шифра, поэтому силу шифра можно оценить только на основе предложенных атак на него. Например, 20-раундовый шифр, имеющий атаку на 5 раундов, может быть признан безопасным, а имеющий атаку на 19 раундов вряд ли выдержит атаки криптоаналитиков в ближайшие годы.

Для финалиста конкурса AES шифра MARS [5] подсчитывать количество раундов не совсем корректно, поскольку половина раундов этого 32-раундового шифра — *раунды перемешивания* — не снабжаются подключами и, следовательно, не несут в себе такой криптографической силы, как раунды *ядра*, снабженные двумя подключами. Кроме того, в шифре присутствуют процедуры *пред-отбеливания* (pre-whitening) и *пост-отбеливания* (post-whitening), которые также используют подключи. По этим причинам при сравнении атаки, представленной в этой работе, с ранее известными результатами подсчитывается не количество раундов, а количество раскрытых битов подключей.

В таблице показано, что новая атака позволяет атаковать такой вариант шифра MARS, который использует 752 бита подключей, в то время как лучшая из ранее известных атак — только 682. Отметим, что шифр MARS имеет 256-битовый секретный ключ, и все атаки, приведенные в таблице, действуют быстрее, чем полный перебор ключей такой длины.

Лучшие результаты криптоанализа шифра MARS

Найденные биты подключей	Раунды			Отбеливания		Сложность			Источник
	Всего	Ядра	Перемешивающие	Пред-	Пост-	Блоки	Память, байт	Шифрования	
566	21	5	16	+	+	2^3	2^{236}	2^{232}	[6]
566	21	5	16	+	+	2^{50}	2^{197}	2^{247}	[6]
628	12	6	6	+	+	2^{69}	2^{73}	2^{197}	[6]
682	11	11	0	–	–	2^{65}	2^{69}	2^{229}	[7]
752	16	8	8	+	+	2^{105}	2^{109}	2^{231}	Данная работа

Используемые обозначения:

$A := B$	—	A присвоить B	δ_n	—	32-битовое слово, которое имеет 1 на n-й позиции и нули на остальных
$a \oplus b$	—	побитовый xor			
$a \boxplus b$	—	сложение по модулю 2^{32}	(a_0, a_1, a_2, a_3)	—	128-битовый блок, поделенный на четыре 32-битовых слова
$a \circledast b$	—	умножение по модулю 2^{32}			
$a \boxminus b$	—	вычитание по модулю 2^{32}	$(a_0, ?, ?, a_3)$	—	блок, где некоторые слова неизвестны
$a <<< n$	—	поворот a влево на n бит	31-й бит	—	старший бит в слове

1. Краткое описание дифференциального криптоанализа

Основной объект, который исследуется в дифференциальном криптоанализе, — это пары блоков текста A и B с определенной *разностью* $A \oplus B$ [8]. Если информация о том, как связаны входная разность (между блоками открытого текста) и выходная разность (между блоками шифртекста), отсутствует, то все выходные разности равновероятны. Однако если (как-то) удается установить, что некоторая входная разность Δ_{inp} приводит к некоторой выходной разности Δ_{out} с вероятностью p большей, чем остальные, то это может быть использовано для отыскания подключей шифра. Пара $(\Delta_{\text{inp}},\ \Delta_{\text{out}})$ называется *дифференциалом*, а совокупность дифференциалов на различных раундах

называется *характеристикой*. Если Δ_{out} содержит неизвестные биты, то дифференциал называется *усеченным* [9]. Далее дифференциал обозначается так:

$$\Delta_{\text{inp}} \underset{p}{\rightarrow} \Delta_{\text{out}}.$$

Например, для совершенного шифра со 128-битовым блоком при любой входной разности выходная разность примет некоторое фиксированное значение с вероятностью 2^{-128}. Таким образом, если в процессе анализа шифра обнаружится, что определенная входная разность приводит к определенной выходной разности с вероятностью больше чем 2^{-128} (например, 2^{-100}), то эта информация может быть использована с целью отыскания его подключей. Количество текстов, требуемых для реализации атаки, пропорционально $1/p$.

2. Описание шифра MARS

Опишем шифр MARS в объеме, достаточном для понимания статьи (см. полную спецификацию в [5]).

Общая структура шифра MARS. MARS оперирует со 128-битовыми блоками, секретный пользовательский ключ имеет длину до 256 бит. Перед шифрованием этот ключ преобразуется в массив из сорока 32-битовых подключей $K_0, ..., K_{39}$. Каждый блок открытого текста A представляется в виде четырех 32-битовых слов (a_0, a_1, a_2, a_3) и подвергается следующим преобразованиям:

— пред-отбеливание: $a_0 := a_0 \boxplus K_0;\ a_1 := a_1 \boxplus K_1;\ a_2 := a_2 \boxplus K_2;\ a_3 := a_3 \boxplus K_3$;
— 8 *прямых* раундов перемешивания;
— 8 прямых раундов ядра;
— 8 *обратных* раундов ядра;
— 8 обратных раундов перемешивания;
— пост-отбеливание: $a_0 := a_0 \boxminus K_{36};\ a_1 := a_1 \boxminus K_{37};\ a_2 := a_2 \boxminus K_{38};\ a_3 := a_3 \boxminus K_{39}$.

Описание функции E. Основную криптографическую силу шифру MARS придает функция E, которая в качестве аргументов принимает одно 32-битовое слово IN и два подключа: K^+ и $K^\times$. На выходе эта функция дает три 32-битовых слова: L, M и R (рис. 1). Важно заметить, что подключи $K^\times$ на всех раундах имеют только 30 неизвестных бит, поскольку согласно алгоритму их получения два младших бита этих подключей всегда равны 1. Слова L, M и R в течение выполнения функции E преобразуются следующим образом:

$R := ((IN <<< 13) \circledast K^\times) <<< 5$;
$M := IN \boxplus K^+;\ L := M$;
$M := M <<< (R \mod 32);\ M$ — выход;
$L := \mathrm{S}(L \mod 512) \oplus R$; (S — это S-бокс, отображающий 9 бит в 32 бита);
$R := R <<< 5;\ L := (L \oplus R) <<< (R \mod 32);\ L$ — выход; R — выход.

Раунды ядра шифра MARS. Каждый из 16 раундов ядра (рис. 1) осуществляет следующее преобразование блока: слово a_0 вводится в функцию E (как IN) и затем поворачивается на 13 бит влево. Три выходных слова этой функции (R, M и L) изменяют оставшиеся три слова блока:

$a_1 := a_1 \boxplus L;\ a_2 := a_2 \boxplus M;\ a_3 := a_3 \oplus R$; (прямой раунд ядра);
$a_1 := a_1 \oplus R;\ a_2 := a_2 \boxplus M;\ a_3 := a_3 \boxplus L$; (обратный раунд ядра);

завершается раунд перестановкой слов: $(a_0, a_1, a_2, a_3) := (a_1, a_2, a_3, a_0)$.

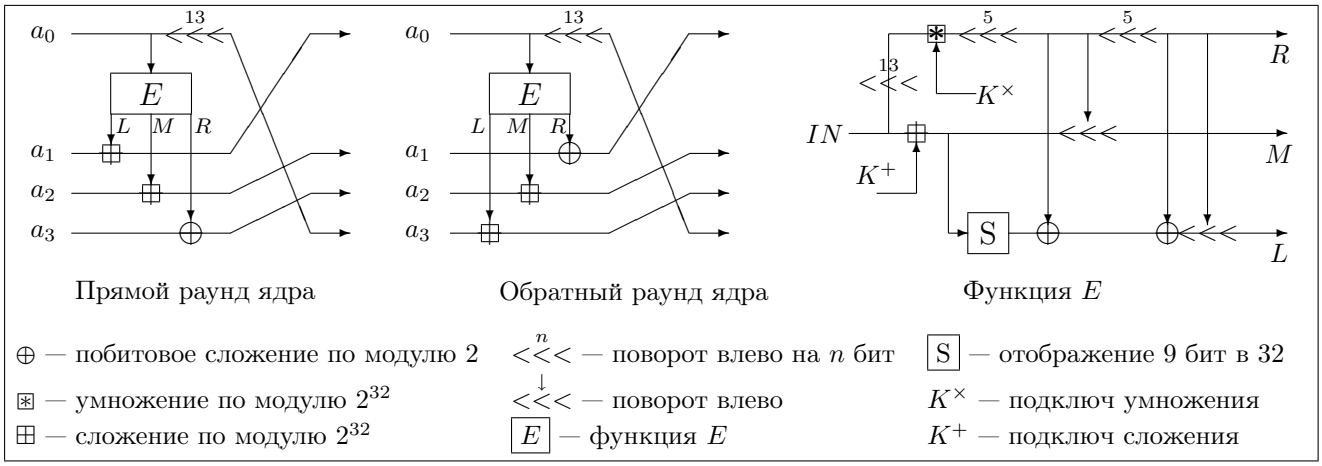

Рис. 1. Раунды ядра и функция E шифра MARS

3. Дифференциальная характеристика для восьми раундов

Опишем построение дифференциальной характеристики

$$(0,0,0,2^{18}) \xrightarrow[p=1/2]{pre-whitening\ +\ 3\ rounds} (2^{18},0,0,0) \xrightarrow[p=1]{1\ round} (2^9,?,?,2^{31}) \xrightarrow[p=1]{1\ round} (?,?,?,2^{22}) \underset{p=2^{-96}}{=} (0,0,0,2^{22}) \xrightarrow[p=1/2]{3\ rounds} (2^{22},0,0,0), \tag{1}$$

которая покрывает 8 раундов ядра шифра MARS с вероятностью 2^{-98}. Первые три раунда могут быть как прямыми, так и обратными, последние пять — только обратными. Эта характеристика получается путем объединения нескольких дифференциалов.

Дифференциал 1. Построим дифференциал

$$(0,0,0,2^{18}) \xrightarrow[p=1/2]{pre-whitening\ +\ 3\ rounds} (2^{18},0,0,0), \tag{2}$$

который покрывает пред-отбеливание и три раунда ядра (прямые или обратные — значения не имеет). Для этого рассмотрим пару блоков A и B с разностью $(0,0,0,2^{18})$ на входе шифра перед пред-отбеливанием, т. е. $a_3 \oplus b_3 = 2^{18}$. В этом случае выходная разность после пред-отбеливания и трех раундов ядра будет $((a_3 \boxplus z) \oplus (b_3 \boxplus z), 0, 0, 0)$, где z — это сумма первого подключа и выходов функции E на трех расматриваемых раундах. Если $a_3 \oplus b_3 = 2^{18}$, то вероятность того, что $(a_3 \boxplus z) \oplus (b_3 \boxplus z) = 2^{18}$, равна $1/2$ (данный вопрос рассматривался, например, в [10]). Этот дифференциал является начальным в (1).

Дифференциал 2. Аналогичные рассуждения позволяют построить следующий дифференциал:

$$(0,0,0,2^{22}) \xrightarrow[p=1/2]{3\ rounds} (2^{22},0,0,0). \tag{3}$$

Этот дифференциал является заключительным в (1).

Дифференциал 3. Дифференциал

$$(2^{18},0,0,0) \xrightarrow[p=1]{1\ round} (2^9,?,?,2^{31}) \tag{4}$$

начинается с выходной разности $(2^{18}, 0, 0, 0)$ первого дифференциала и покрывает один обратный раунд ядра.

Для его построения заметим, что разность 2^{18} находится в первом слове, которое попадает в функцию E. Эта разность поворачивается на 13 бит влево и становится равной 2^{31}. Далее она проходит через умножение на подключ с вероятностью 1, т. е. после умножения на подключ двух слов с разностью 2^{31} они будут иметь такую же разность с вероятностью 1. За умножением следуют два поворота на 5 бит, т. е. суммарный поворот осуществляется на 10 бит, что приводит к разности $(?, ?, 2^9)$ на выходе функции E.

Выходные разности для двух оставшихся слов M и L могут быть частично установлены, но после следующего раунда уже невозможно будет сказать ничего определенного, их точные значения в дальнейшем не требуются, поэтому считаем их неизвестными. В итоге после рассмотренного одного обратного раунда ядра получим выходную разность $(2^9, ?, ?, 2^{31})$.

Дифференциал 4. Дифференциал

$$(2^9, ?, ?, 2^{31}) \xrightarrow[p=1]{1\ round} (?, ?, ?, 2^{22}) \tag{5}$$

начинается с выходной разности $(2^9, ?, ?, 2^{31})$ третьего дифференциала и покрывает один обратный раунд ядра. Этот дифференциал получается, если проследить за перестановкой слов в обратном раунде ядра.

Объединение дифференциалов. Чтобы получить характеристику (1), необходимо объединить построенные дифференциалы. Перед этим заметим, что вероятность того, что $(?, ?, ?, 2^{22}) = (0, 0, 0, 2^{22})$, равна 2^{-96}. Данный факт следует из того, что информации о неизвестной 96-битовой разности нет, поэтому можно лишь подразумевать, что она принимает любое значение с вероятностью 2^{-96} (подобные идеи использовались в [11] при построении атаки на AES).

Таким образом, вероятность характеристики (1) равна произведению вероятностей входящих в нее дифференциалов и 2^{-96}, что составляет 2^{-98}.

4. Атака на MARS

Опишем атаку на шифр MARS, состоящий из пред- и пост-отбеливаний, 8 обратных раундов ядра и 8 обратных раундов перемешивания. Таким образом, атакуется версия шифра, состоящая из отбеливаний и 16 раундов (с 17-го по 32-й). Данная версия шифра MARS использует 752 бита подключей. Атака, направленная на их отыскание, основана на дифференциальной характеристике (1).

Атака выполняется в два этапа: сначала — с целью определения подключей пост-отбеливания, затем — с целью определения подключей раундов ядра.

4.1. Атака на подключи пост-отбеливания

Для удобства рассмотрим четыре 32-битовых подключа для пост-отбеливания как один 128-битовый подключ. Первым и самым трудоемким этапом атаки является определение этого подключа. Положим также, что $X = \mathrm{MARS}(A)$ — это шифрование блока A в блок X с помощью шифра MARS, состоящего из пред-отбеливания, восьми раундов ядра, которые покрываются характеристикой, восьми обратных раундов перемешивания и пост-отбеливания.

Описание атаки. Атака работает наподобие шестислойного фильтра. Перебираются все возможные 2^{128} подключей пост-отбеливания и пропускаются через этот

фильтр. Решающее свойство фильтра заключается в том, что правильный подключ успешно проходит все шесть слоев, а ни один из неправильных подключей-кандидатов этого сделать не может. Атака реализуется в несколько шагов:

1. Сформировать 6 групп различных пар блоков $A_i^b,\ B_i^b$ $(b = 1, ..., 6;\ i = 1, ..., 2^{101})$ с фиксированной разностью $A_i^b \oplus B_i^b = (0, 0, 0, 2^{18})$. Эта разность является входной разностью характеристики (1).
2. Для каждой из пар $A_i^b,\ B_i^b$ запросить пару шифртекстов $X_i^b = \mathrm{MARS}(A_i^b)$ и $Y_i^b = \mathrm{MARS}(B_i^b)$, полученные $6 \cdot 2^{101}$ пар шифртекстов сохранить в памяти.
3. Перебрать все возможные значения искомого подключа $\texttt{key} \in \{0, ..., 2^{128} - 1\}$ и для каждого из них выполнить следующие действия:
   а) $b := 1$;
   б) с помощью пост-отбеливания подключом `key` и 8 обратных раундов перемешивания расшифровать сохраненные в памяти пары шифртекста из группы b и получить пары P_i^b и Q_i^b;
   в) если $P_i^b \oplus Q_i^b \neq (0, 0, 0, 2^{22})$ для всех $i = 1, ..., 2^{101}$, то `key` — это неправильный подключ-кандидат. Отбросить его и перейти к п. 3, где выбрать следующий подключ-кандидат;
   если хотя бы одна из пар обеспечивает условие $P_i^b \oplus Q_i^b = (0, 0, 0, 2^{22})$, то перейти к п. *г*;
   г) если $b < 6$, то $b := b + 1$ и перейти к п. *б*; иначе — к п. *д*;
   д) $b = 6$, и подключ-кандидат прошел все шесть слоев фильтра. Это правильный подключ.

Вероятность успеха атаки. Вычислим вероятность того, что все неправильные подключи-кандидаты отброшены, а правильный 128-битовый подключ остался.

Поскольку пара $X_i^b,\ Y_i^b$, частично расшифрованная с неправильным подключом, не подчиняется дифференциалу

$$(0, 0, 0, 2^{18}) \xrightarrow[p=2^{-98}]{pre-whitening\ +\ 8\ rounds} (2^{22}, 0, 0, 0),$$

то разность $P_i^b \oplus Q_i^b$ принимает любое значение (включая и $(2^{22}, 0, 0, 0)$) равновероятно, т. е. с вероятностью 2^{-128}. Вероятность получить такую разность среди 2^{101} пар приблизительно равна 2^{-27}, а вероятность получить такую разность сразу в шести группах — 2^{-162}. Значит, вероятность того, что хотя бы один неправильный подключ-кандидат из $2^{128} - 1$ пройдет шесть слоев фильтра, составляет примерно 2^{-34}.

Пара $X_i^b,\ Y_i^b$, частично расшифрованная с правильным подключом, подчиняется указанному дифференциалу, следовательно, $P_i^b \oplus Q_i^b = (0, 0, 0, 2^{22})$ с вероятностью 2^{-98}, и вероятность получить такую разность хотя бы раз среди 2^{101} пар равна примерно 0,999665 (см. распределение Пуассона [12]). Значит, вероятность получить такую разность во всех шести группах сразу приблизительно равна 0,99799.

Сложность атаки на подключи пост-отбеливания. Для реализации описанной части атаки необходимо примерно 2^{233} частичных 8-раундовых операций расшифрования, что эквивалентно 2^{231} операциям шифрования полным шифром MARS. Также требуются 2^{105} выбранных открытых текстов и 2^{109} байт памяти, чтобы их хранить.

4.2. Атака на подключи раундов ядра

После того как 128-битовый подключ пост-отбеливания установлен, все хранящиеся в памяти шифртексты необходимо расшифровать на пост-отбеливании и на 8 не

снабженных подключами обратных раундах перемешивания. После этого шифртексты оказываются зашифрованными с помощью шифра MARS, состоящего из пред-отбеливания и восьми раундов ядра. Применяя дифференциалы, образующие характеристику (1), нужно реализовать аналогичную атаку, перебирая подключи раундов ядра один за другим. Для определения подключа последнего раунда необходимо использовать дифференциал

$$(0,0,0,2^{18}) \xrightarrow[p=2^{-98}]{pre-whitening + 7\ rounds} (0,2^{22},0,0),$$

для определения подключа предпоследнего раунда необходимо использовать дифференциал

$$(0,0,0,2^{18}) \xrightarrow[p=2^{-98}]{pre-whitening + 6\ rounds} (0,0,2^{22},0),$$

затем —

$$(0,0,0,2^{18}) \xrightarrow[p=1/2]{pre-whitening + 5\ rounds} (?,?,?,2^{22}),$$

$$(0,0,0,2^{18}) \xrightarrow[p=1/2]{pre-whitening + 4\ rounds} (2^{9},?,?,2^{31}),$$

$$(0,0,0,2^{18}) \xrightarrow[p=1/2]{pre-whitening + 3\ rounds} (2^{18},0,0,0),$$

$$(0,0,0,2^{18}) \xrightarrow[p=1/2]{pre-whitening + 2\ rounds} (0,2^{18},0,0),$$

$$(0,0,0,2^{18}) \xrightarrow[p=1/2]{pre-whitening + 1\ round} (0,0,2^{18},0),$$

$$(0,0,0,2^{18}) \xrightarrow[p=1/2]{pre-whitening} (0,0,0,2^{18}).$$

После определения всех раундовых подключей подключи для пред-отбеливания вычисляются путем решения системы линейных уравнений.

Сложность атаки на подключи раундов ядра и сложность всей атаки. В каждый раунд ядра вводятся два 32-битовых подключа, в одном из которых два младших бита всегда равны 1, следовательно, перебирать нужно 62-битовые подключи. Это существенно меньше, чем перебор 128-битовых подключей пост-отбеливания, поэтому сложность поиска раундовых подключей ничтожно мала по сравнению с поиском подключа пост-отбеливания, и суммарная сложность атаки определяется сложностью отыскания подключей пост-отбеливания.

Количество пар текстов, необходимых для отыскания подключей раундов ядра, также не превзойдет того количества, которое необходимо для атаки на ключ пост-отбеливания. Данный факт следует из того, что вероятности дифференциалов, используемых в атаке на подключи раундов ядра, меньше либо равны вероятности дифференциала первой части атаки.

Заключение

В данной работе описана атака на шифр MARS, которая является более эффективной, чем ранее известные. Эта атака основана на 8-раундовой дифференциальной характеристике.

ЛИТЕРАТУРА

1. `http://csrc.nist.gov/encryption/aes` — Advanced Encryption Standard (AES) project 1997–2000.
2. `https://www.cosic.esat.kuleuven.be/nessie` — New European Schemes for Signatures, Integrity, and Encryption Deliverables of the NESSIE project 2003.
3. `http://www.cryptrec.go.jp/english/` — CRYPTREC project 2000–2002.
4. *Schneier B.* A self-study course in block-cipher cryptanalysis // Cryptologia. 2000. V. 24. No. 1. P. 18–34.
5. *Burwick C. et al.* MARS — a candidate cipher for AES // AES submission. 1999. `http://www.research.ibm.com/security/mars.pdf`.
6. *Kelsey J., Schneier B.* MARS attacks! Preliminary cryptanalysis of reduced-round MARS variants // Proc. of the Third AES Candidate Conf. 2000. `http://www.schneier.com/paper-mars-attacks.pdf`.
7. *Kelsey J., Kohno T., Schneier B.* Amplified boomerang attacks againts reduced-round MARS and Serpent // LNCS. 2001. V. 1978. P. 75–93.
8. *Biham E., Shamir A.* Differential cryptanalysis of DES-like cryptosystems // J. Cryptol. 1991. V. 4. P. 3–72.
9. *Knudsen L.* Truncated and higher order differentials // LNCS. 1995. V. 1008. P. 196–211.
10. *Biryukov A., Kushilevitz E.* Improved cryptanalysis of RC5 // LNCS. 1998. V. 1403. P. 85–99.
11. *Biryukov A.* The boomerang attack on 5 and 6-Round reduced AES // LNCS. 2005. V. 3373. P. 11–15.
12. *Боровков А. А.* Теория вероятностей. М.: Наука, 1976. 352 с.