УДК 681.03.245

Л.К. Бабенко, Е.А. Ищукова

ДИФФЕРЕНЦИАЛЬНЫЙ КРИПТОАНАЛИЗ АЛГОРИТМА ГОСТ 28147-89*

Исследована стойкость алгоритма ГОСТ к атаке на основе дифференциального криптоанализа. В результате исследования было определено, что существует ряд S-блоков, обладающих слабыми свойствами по отношению к дифференциальному криптоанализу. Использование таких блоков в алгоритме ГОСТ позволяет получать характеристики, обладающие довольно высокими вероятностями, которые можно использовать для проведения атаки. Так, при использовании одного и того же слабого блока замены, вероятность характеристики для 32 раундов ГОСТ может составлять 2^{-25}, что позволяет сравнительно легко получать правильные пары текстов для анализа.

ГОСТ; дифференциальный криптоанализ; входная разность; выходная разность характеристика; секретный ключ.

L.K. Babenko, E.A. Ischukova

DIFFERENTIAL CRYPTANALYSIS OF GOST ENCRYPTION ALGORITHM

In this article we explore the resistance of the GOST 28147-89 algorithm to the attack based on differential cryptanalysis. As the result of our research, we have found out that there is a number of S-boxes with weak properties with respect to differential cryptanalysis. The use of such elements in GOST allows obtaining features that have a fairly high probability that can be used to carry out attacks. So, if we use the same weak block replacement, the probability characteristics for the 32 rounds of GOST can reach 2^{-25}, which makes it relatively easy to get the right pair of texts for analysis.

GOST; differential cryptanalysis; input difference; output difference; characteristic; secret key.

Метод дифференциального криптоанализа, впервые предложенный Э. Бихамом (E. Biham) и А. Шамиром (A. Shamir) для анализа алгоритма DES [1, 2], базируется на прослеживании изменения разности двух сообщений при их прохождении через раунды шифрования. После появления работ [1, 2] большинство существовавших на тот момент алгоритмов шифрования были подвергнуты анализу с использованием данного метода. Исследования показали, что метод дифференциального криптоанализа является универсальным, т.е. может быть применен к анализу большинства известных симметричных криптосистем. Именно поэтому вновь создаваемые алгоритмы шифрования в первую очередь тестируются на устойчивость к данному виду анализа.

Данная работа посвящена изучению стойкости к методу дифференциального криптоанализа алгоритма ГОСТ, определенного в качестве государственного стандарта в Российской Федерации. До сих пор в открытой печати имеется сравнительно мало информации о возможных уязвимостях данного шифра. Одной из наиболее значимых работ является статья [3], в которой авторы предложили вариант анализа алгоритма ГОСТ с использованием дифференциального криптоанализа на связанных ключах (Related-Key Attack) при условии использования слабых блоков замены. В настоящей работе предлагается рассмотреть возможность осуществления атаки на алгоритм шифрования ГОСТ с помощью классического метода дифференциального криптоанализа и определить условия, при которых осуществление данной атаки возможно.

* Работа поддержана грантом РФФИ № 09-07-00245-а.

Отличительной чертой алгоритма ГОСТ является использование в его структуре нефиксированных блоков замены. Предполагается, что при любом заполнении S-блоков тридцати двух раундов шифрования будет достаточно для того, чтобы противостоять таким мощным методам анализа, как линейный и дифференциальный криптоанализ. В данной работе будет показано, что существуют слабые блоки замены, использование которых в алгоритме ГОСТ может привести к успешному осуществлению атаки на основе метода дифференциального криптоанализа. Долгое время считалось, что если оставлять S-блоки в секрете, то их можно рассматривать как дополнительный ключевой материал [6]. Однако в работе [5] был предложен метод, применение которого позволяет достаточно просто восстановить значения S-блоков, используемых для шифрования данных.

Наличие нефиксированных S-блоков в алгоритме ГОСТ приводит к тому, что каждый раз анализ необходимо начинать сначала. В отличие от алгоритмов с фиксированными блоками замены, где можно единожды определить хорошие характеристики и использовать их для анализа шифрованных данных. В данной статье будет показано, что при применении предложенных подходов к анализу, а также при использовании алгоритма поиска характеристик, предложенного ранее в работе [4], первый этап анализа, заключающийся в поиске хороших характеристик, не составляет проблемы. Хорошая характеристика может быть найдена за несколько секунд и даже меньше.

Прежде чем приступить к рассмотрению проблемы, введем несколько базовых определений, которые будут использованы при изложении материала. Под ***разностью*** будем понимать результат поразрядного сложения по модулю два (операция XOR) двух отдельно шифруемых на одном и том же секретном ключе текстов, находящихся в одной и той же позиции рассматриваемого алгоритма шифрования. Под ***входной разностью*** будем понимать разность, поступающую на вход рассматриваемого криптографического преобразования, а под ***выходной разностью*** – разность, образованную на выходе рассматриваемого преобразования. Под ***характеристикой*** будем понимать пару значений: входная разность – выходная разность для n раундов алгоритма шифрования. Под ***правильной парой текстов*** будем понимать две пары значений: открытое сообщение – шифрованное сообщение. Разность входных сообщений равна входной разности характеристики, а разность шифрованных сообщений равна выходной разности характеристики.

Описание алгоритма ГОСТ. Алгоритм шифрования ГОСТ 28147-89 является государственным стандартом Российской Федерации. Его использование обязательно для шифрования данных в государственных организациях РФ. Алгоритм ГОСТ является симметричным блочным шифром, построенным по схеме Фейстеля (Feistel). На вход алгоритма поступает 64-битовый блок данных, который под воздействием 256-битового ключа преобразуется в 64-битовый блок шифрованных данных. В каждом раунде правая часть шифруемого сообщения поступает на вход функции F, где преобразуется с использованием трех криптографических операций: сложения данных с раундовым подключом по модулю 2^{32}, замены данных с использованием S-блоков, циклического сдвига влево на 11 позиций. Выход функции F складывается по модулю 2 с левой частью шифруемого сообщения, после чего правая и левая части меняются местами. Алгоритм содержит 32 раунда, в последнем раунде шифрования правая и левая части местами не меняются.

В алгоритме шифрования ГОСТ используется 8 S-блоков, которые преобразуют 4 бита на входе в S-блок в 4 бита на выходе. В отличие от большинства алгоритмов шифрования ГОСТ не имеет фиксированных блоков замены и может использовать любые варианты блоков.

Секретный ключ шифрования содержит 256 битов и представляется в виде последовательности из восьми 32-битовых слов: К1, К2, К3, К4, К5, К6, К7, К8. В каждом раунде шифрования в качестве раундового подключа используется одно из этих 32-битовых слов. При определения раундового подключа руководствуются следующим принципом: с 1 по 24 раунды используются последовательно К1, К2, К3, К4, К5, К6, К7, К8, К1, К2 и т.д. С 25 по 32 раунды: К8, К7, К6, К5, К4, К3, К2, К1. Таким образом, получается, что в первом и последнем раундах используется один и тот же раундовый подключ К1.

Дифференциальные свойства криптографических преобразований алгоритма ГОСТ. ***Сложение по модулю 2^{32}.*** Метод дифференциального криптоанализа базируется на прослеживании изменения несхожести между двумя сообщениями. Для определения несхожести используется операция сложения по модулю два, которая в результате сложения дает ненулевые биты в тех позициях, в которых два исходных сообщения имели различные значения битов. Именно поэтому в алгоритме шифрования DES при рассмотрении прохождения разности через функцию раундового преобразования F значение ключа не влияет на изменение разности текстов, так как два одинаковых значения дадут 0 в результате сложения по модулю два. В алгоритме шифрования ГОСТ, в отличие от DES, сложение с ключом происходит по модулю 2^{32}. В результате исследования свойств данной операции методом индукции были выявлены правила определения вероятности того, что разность останется неизменной при прохождении через операцию целочисленного сложения по модулю 2^n:

1. Любое значение входной разности может отобразиться само в себя, т.е. остаться неизменным. Вероятность такого отображения определяется следующим образом:

$$p = \frac{1}{2^k}, \text{ если входная разность } \Delta_{вх} < 2^{n-1}; \quad (1)$$

$$p = \frac{1}{2^{k-1}}, \text{ если входная разность } \Delta_{вх} \geq 2^{n-1}, \quad (2)$$

где k – число ненулевых позиций входной разности.

2. Для входной разности равной Δвх=0 на выходе преобразования будет значение выходной разности Δвых = 0 с вероятностью p=1.

3. Для входной разности $\Delta вх = 2^{n-1}$ на выходе преобразования будет значение выходной разности $\Delta вых = 2^{n-1}$ с вероятностью p = 1.

Более детальное объяснение можно найти в работе [4].

Замена с помощью S-блоков. Несходство различных пар текстов при прохождении через криптографические операции приводит к несходству получаемых шифртекстов с определенной вероятностью. Для S-блоков эти вероятности можно определить, построив соответствующие таблицы анализа. Таблицы сроятся по следующему принципу: по вертикали располагаются все возможные комбинации входной разности ΔA для рассматриваемого S-блока, по горизонтали – все возможные комбинации выходной разности ΔC для этого же блока, а на пересечении – число соответствий данного значения ΔC данному значению ΔA (либо отношение полученного числа соответствий к возможному числу соответствий; возможное число соответствий равно 2^n, где n – число битов, поступающих на вход S-блока). Пары разностей ΔA и ΔC, имеющие наибольшую вероятность (или вероятность, близкую к максимальной), могут быть использованы для анализа с целью нахождения секретного ключа шифрования.

Циклический сдвиг. Операция сдвига в алгоритме шифрования ГОСТ 28147-89 является одной из трех основных операций, составляющих раундовую функцию пре-

образования. При рассмотрении дифференциального криптоанализа речь обычно идет о разностях двух текстов. Пусть у нас имеется два текста: A и B, тогда их разность равна (A ⊕ B). Если сдвинуть каждое из значений A и B циклически влево на s разрядов, то получим разность (A << s) ⊕ (B << s), которая обладает следующим свойством:

$$(A << s) \oplus (B << s) = (A \oplus B) << s. \quad (3)$$

Если рассмотреть эту операцию применительно к алгоритму ГОСТ 28147-89, то получается:

$$(A << 11) \oplus (B << 11) = (A \oplus B) << 11,$$

т.е. для получения правильной разности на выходе операции циклического сдвига, необходимо входную разность циклически сдвинуть влево на 11 позиций.

Слабые блоки для алгоритма ГОСТ: поиск и анализ. Алгоритм ГОСТ содержит 8 S-блоков. При этом сами блоки не являются фиксированными. То есть теоретически считается, что могут быть использованы блоки замены, сформированные случайным образом. Криптографическую стойкость алгоритму должно обеспечить достаточно большое число раундов шифрования. Долгое время считалось, что если держать S-блоки в секрете, то можно рассматривать их как дополнительный ключевой материал. Однако в работе [5] было показано, что S-блоки, используемые в алгоритме шифрования, можно достаточно просто восстановить.

В связи с этим разумно рассмотреть слабые S-блоки для алгоритма ГОСТ и оценить степень сложности атаки на основе дифференциального криптоанализа при их использовании.

Было определено, что в качестве слабых выступают те блоки замены, для которых входная разность ΔA, содержащая одну единицу (т.е. возможно всего 4 варианта такой разности ΔA = 1, ΔA = 2, ΔA = 4, ΔA = 8) заменяется на выходную разность ΔC, также содержащую одну единицу (ΔC = 1, ΔC = 2, ΔC = 4, ΔC = 8).

Если считать, что разность ΔA получается путем сложения двух сообщений X и X1, поступающих на вход S-блока, то можно сопоставить вместе все варианты для получения разностей ΔA так, как показано в табл. 1. В табл. 1 в скобках указаны значения входов X и X1 в десятичном виде.

Таблица 1

Варианты входных значений для получения разностей ΔA

Вход X	Вход X1 ΔA=0001 (1)	Вход X1 ΔA=0010 (2)	Вход X1 ΔA=0100 (4)	Вход X1 ΔA=1000 (8)
0000 (0)	0001 (1)	0010 (2)	0100 (4)	1000 (8)
0001 (1)	0000 (0)	0011 (3)	0101 (5)	1001 (9)
0010 (2)	0011 (3)	0000 (0)	0110 (6)	1010 (10)
0011 (3)	0010 (2)	0001 (1)	0111 (7)	1011 (11)
0100 (4)	0101 (5)	0110 (6)	0000 (0)	1100 (12)
0101 (5)	0100 (4)	0111 (7)	0001 (1)	1101 (13)
0110 (6)	0111 (7)	0100 (4)	0010 (2)	1110 (14)
0111 (7)	0110 (6)	0101 (5)	0011 (3)	1111 (15)
1000 (8)	1001 (9)	1010 (10)	1100 (12)	0000 (0)
1001 (9)	1000 (8)	1011 (11)	1101 (13)	0001 (1)
1010 (10)	1011 (11)	1000 (8)	1110 (14)	0010 (2)
1011 (11)	1010 (10)	1001 (9)	1111 (15)	0011 (3)
1100 (12)	1101 (13)	1110 (14)	1000 (8)	0100 (4)
1101 (13)	1100 (12)	1111 (15)	1001 (9)	0101 (5)
1110 (14)	1111 (15)	1100 (12)	1010 (10)	0110 (6)
1111 (15)	1110 (14)	1101 (13)	1011 (11)	0111 (7)

Сопоставив значение каждого входа i с соответствующим ему выходом a_i, представим результат в виде графа связей так, как показано на рис. 1. На рис. 1 стрелками обозначены связи для входов i, которые используются для получения значений ΔA = 1, ΔA = 2, ΔA = 4, ΔA = 8. Римские цифры над горизонтальными стрелками и слева от вертикальных стрелок указывают номер связи (связь I – для ΔA = 1, связь II – для ΔA = 2, связь III – для ΔA = 4 и связь IV – для ΔA = 8). Так, например, значение ΔA, равное 2 (связь II), может быть получено, если X = 0, а X1 = 2 или если X = 1, а X1 = 3 и т.д.

Каждому входному значению ΔA может быть противопоставлено одно из четырех выходных значений ΔC (ΔC = 1, ΔC = 2, ΔC = 4, ΔC = 8). Таким образом, для четверки входных разностей ΔA = 1, ΔA = 2, ΔA = 4, ΔA = 8 можно противопоставить 24 комбинации из четверок выходных разностей (табл. 2).

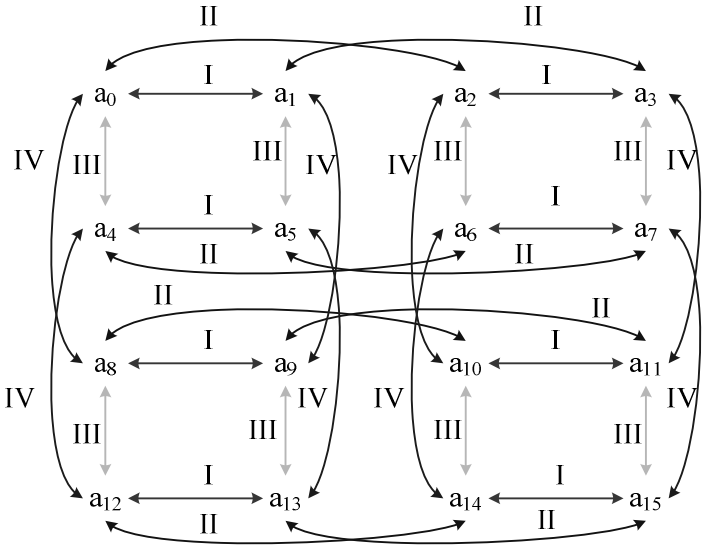

Рис. 1. Граф связей для возможных выходов S-блока

Таблица 2

Варианты соответствия выходных разностей ΔC

	№1 ΔC	№2 ΔC	№3 ΔC	№4 ΔC	№5 ΔC	№6 ΔC	№7 ΔC	№8 ΔC	№9 ΔC	№10 ΔC	№11 ΔC	№12 ΔC
ΔA=1	1	1	1	1	1	1	2	2	2	2	2	2
ΔA=2	2	2	4	4	8	8	1	1	4	4	8	8
ΔA=4	4	8	2	8	2	4	4	8	1	8	1	4
ΔA=8	8	4	8	2	4	2	8	4	8	1	4	1
	№13 ΔC	№14 ΔC	№15 ΔC	№16 ΔC	№17 ΔC	№18 ΔC	№19 ΔC	№20 ΔC	№21 ΔC	№22 ΔC	№23 ΔC	№24 ΔC
ΔA=1	4	4	4	4	4	4	8	8	8	8	8	8
ΔA=2	1	1	2	2	8	8	1	1	2	2	4	4
ΔA=4	2	8	8	1	1	2	2	4	1	4	1	2
ΔA=8	8	2	1	8	2	1	4	2	4	1	2	1

Обозначим ΔC_j значение выходной разности для ΔA = j. Тогда в соответствии с рис. 1 можно определить 15 формул для нахождения заполнения искомой таблицы:

$a_1 = a_0 \oplus \Delta C_1$;
$a_2 = a_0 \oplus \Delta C_2$;
$a_4 = a_0 \oplus \Delta C_3$;
$a_8 = a_0 \oplus \Delta C_4$;
$a_3 = a_1 \oplus \Delta C_2$;
$a_5 = a_1 \oplus \Delta C_3$;
$a_9 = a_1 \oplus \Delta C_4$;
$a_6 = a_2 \oplus \Delta C_3$;
$a_{10} = a_2 \oplus \Delta C_4$;
$a_7 = a_3 \oplus \Delta C_3$;
$a_{11} = a_3 \oplus \Delta C_4$;
$a_{12} = a_4 \oplus \Delta C_4$;
$a_{14} = a_6 \oplus \Delta C_4$;
$a_{15} = a_7 \oplus \Delta C_4$;
$a_{13} = a_{15} \oplus \Delta C_2$;

Вообще таким образом можно построить 28 формул, однако пятнадцати, представленных выше, достаточно для нахождения S-блока, остальные формулы будут дублировать друг друга.

Меняя значение a_0 от 0 до 15, можно получить 16 разных S-блоков для каждой из возможных комбинаций ΔC. Таким образом, можно получить 16*24 = 384 различных S-блоков, обладающих слабыми свойствами.

В качестве примера была использована комбинация №11 из табл. 2 и значение $a_0 = 1$. Применив формулы, получим блок замены, представленный в табл. 3.

Таблица 3

Сформированный блок замены

Вход	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Выход	1	3	9	11	0	2	8	10	5	7	13	15	4	6	12	14

На первый взгляд может показаться, что полученный S-блок не может быть использован для осуществления атаки, так как в таблице анализа, построенной для него, все вероятности равны либо 0, либо 1. Однако это не так. Дело в том, что для алгоритма ГОСТ S-блоки не единственное преобразование, оказывающее влияние на изменение вероятности характеристики. Целочисленное сложение данных по модулю 2^{32} также оказывает влияние на преобразование разности. При этом вероятность того, что разность останется неизменной, напрямую зависит от количества ненулевых позиций в разности, поступающей на вход преобразования. Используя блоки замены, подобные ему, которые представлены в табл. 3, можно легко находить характеристики для анализа алгоритма. При этом вероятность прохождения разности через блоки замены будет равна единице, в то время как общая вероятность характеристики будет зависеть от операции сложения по модулю 2^{32}. Далее будет показано, как S-блок из табл. 3 может быть использован для анализа.

Нахождение характеристик для алгоритма ГОСТ. Для упрощения объяснения в рассматриваемом алгоритме ГОСТ будет использован один и тот же S-блок, представленный в табл. 3. Это не значит, что для успешного анализа алгоритм ГОСТ должен использовать один и тот же S-блок. Возможно достаточно большое число комбинаций из 384 слабых блоков, которые могут привести к подобному результату.

Прежде чем приступить к нахождению хороших характеристик, необходимо определить, какие характеристики не могут быть использованы для анализа. В соответствии с правилами преобразования разности при прохождении ее через операцию целочисленного сложения по модулю 2^n, если входная разность $\Delta вх = 2^{n-1}$, то на выходе преобразования будет значение выходной разности $\Delta вых = 2^{n-1}$ с вероятностью p = 1. Таким образом, если правая часть выходной разности характеристики будет равна 80000000_x, то такую характеристику нельзя будет использовать для поиска секретного ключа, так как вероятность прохождения характери-

стики через последний раунд шифрования будет равна p = 1. Более того, в соответствии с техникой поиска раундовых подключей, которая будет рассмотрена ниже, предположения о значении секретного ключа делаются по тетрадам, соответствующим данным при прохождении через S-блоки. Поэтому тетрады правой части выходной разности характеристики, содержащие значения 0_x или 8_x, также не могут быть использованы для извлечения информации о секретном ключе.

Получается, что необходимо найти такую характеристику, которая в правой части выходной разности в каждой тетраде будет содержать значения, отличные от 0_x или 8_x. В силу свойства операции целочисленного сложения по модулю 2^{32} нахождение единственной характеристики, имеющей ненулевые значения в каждой тетраде, влечет за собой довольно маленькое значение вероятности, так как в каждом раунде через операцию целочисленного сложения будет проходить разность, содержащая минимум 8 единиц. Это не оправдано. Гораздо проще получить несколько характеристик с довольно большими вероятностями, в каждой из которых будет задействованы одна или несколько тетрад.

В работе [4] была представлена разработка универсального алгоритма, позволяющего находить хорошие характеристики для алгоритма ГОСТ. На вход данного алгоритма поступает значение входной разности. После этого рассматриваются все варианты преобразования разности (с учетом свойств таблиц анализа для S-блоков) и отбираются те выходные разности, для которых вероятность будет либо максимальна, либо не ниже заданной (в соответствии с начальными установками алгоритма). В рассматриваемом варианте алгоритма ГОСТ для S-блока замены, определенного в табл. 3, таблица анализа содержит только максимальные вероятности (p=1). Поэтому для входной разности характеристики возможность раундового преобразования определена единственным образом. В связи с этим поиск выходной разности при заданной входной разности занимает время меньше секунды. Данный алгоритм был использован для опробования всех возможных вариантов входных разностей, в тетрадах которых может находиться одно из пяти значений: 0_x, 1_x, 2_x, 4_x и 8_x. Так как входной блок в алгоритм ГОСТ содержит 64 бита, т.е. 16 тетрад, то всего было опробовано $5^{16} \approx 2^{37,15}$. Анализ 32 раундов алгоритма ГОСТ показал, что существует достаточно большой объем пар входной – выходной разности, для которых характеристика обладает достаточно большой вероятностью (в пределах от 2^{-25} до 2^{-33}) и может быть использована для анализа. Как и ожидалось, входные разности, для которых несколько тетрад (более трех) содержали ненулевые тетрады не дали хороших вероятностей. Таким образом, можно сделать вывод, что нет необходимости делать полный перебор для входных разностей. Достаточно рассмотреть несколько вариантов, где в разных тетрадах будут содержаться разные значения, и использовать их в дальнейшем для анализа. Подобный отбор с использованием алгоритма, приведенного в работе [4], может занять от нескольких минут до получаса.

Рассмотрим более подробно одну из найденных характеристик. На вход алгоритма шифрования подается входная разность $\Delta X1 = 80400000\ 00000000_x$. В соответствии со значением $\Delta X1$, на вход функции F первого раунда шифрования поступает значение разности 00000000_x, которое остается неизменным с вероятностью p1 = 1. После этого выход функции F складывается по модулю два с левой частью разности $\Delta X1$, т.е. со значением 80400000_x, и части меняются местами. В результате на вход второго раунда шифрования поступает значение разности $\Delta X2 = 00000000\ 80400000_x$. Здесь уже на вход функции F поступает ненулевое значение разности 80400000_x. Первым преобразованием функции F является операция целочисленного сложения по модулю 2^{32}. Так как значение $80400000_x > 2^{31}$, то, воспользовавшись формулой (2), определяем вероятность, с которой значение

разности 80400000_x останется неизменным p2 = $\frac{1}{2}$. Общая вероятность для двухраундовой характеристики будет равна p = p1*p2 = $\frac{1}{2}$. Следующим шагом является преобразование с помощью S-блоков. Известно, что значение 8_x меняется единственно возможным способом на значение 4_x, а значение 4_x – на значение 1_x. Таким образом, после замены с помощью S-блоков разность преобразуется к виду 40100000_x. Последним преобразованием является циклический сдвиг влево на 11 позиций, в результате которого получаем: (40100000_x <<< 11) = 80000200_x. Таким образом, на выходе F функции второго раунда появится значение разности 80000200_x с вероятностью $\frac{1}{2}$. После сложения с левой частью и обмена местами получим, что на вход третьего раунда шифрования поступит значение ΔX3 = 80400000 80000200x. Все дальнейшие преобразования выполняются аналогичным образом. Примечательно, что после пяти раундов преобразование разностей начинает циклически повторяться. Кроме того, в каждом раунде преобразование разности при прохождении через операцию целочисленного сложения по модулю 2^{32} выполняется с вероятностью не меньше, чем $\frac{1}{2}$. А в 7 раундах из 32 (раунды номер 1, 6, 11, 16, 21, 26, 31) на вход функции F поступает значение разности, равное 00000000_x, которое остается неизменным с вероятностью p = 1. Таким образом, итоговая вероятность для полученной характеристики равна $\frac{1}{2^{25}}$.

Приемы, используемые для поиска секретного ключа. Ранее уже упоминалось, что довольно легко можно найти несколько раундовых характеристик (в идеале характеристик должно быть столько, чтобы в результате не осталось ни одной тетрады в правой части выходной разности, равной 0_x или 8_x), для которых вероятность будет лежать в диапазоне от 2^{-25} до 2^{-33}. Такие значения вероятностей позволяют надеяться на сравнительно легкое нахождение правильных пар текстов, пригодных для анализа. Согласно Парадоксу Дней Рождений, для нахождения правильной пары текстов, соответствующей характеристике, которая выполняется с вероятностью $\frac{1}{2^{25}}$, необходимо в среднем проанализировать $2^{33,5}$ пар текстов.

При поиске секретного ключа для каждой правильной пары текстов необходимо рассматривать первый и последний раунды шифрования. В соответствии с классической структурой алгоритма ГОСТ в первом и последнем раунде используется один и тот же раундовый подключ K1. Таким образом, анализ первого и последнего раундов позволит сделать предположение о значении первого раундового подключа.

То, что найдена правильная пара текстов, соответствующая заданной характеристике, позволяет предполагать, что разность при прохождении через раунды шифрования преобразовывалась именно так, как было определено при построении характеристики. Таким образом, если найдена правильная пара текстов, то известны правые части исходных сообщений XR и XR1, которые поступили на вход функции F первого раунда шифрования. Также известно, что после прохождения через операцию целочисленного сложения по модулю 2^{32} разность этих текстов останется неизменной. Для первого раундового подключа K1 необходимо рассматривать восемь тетрад k1, k2, k3, k4, k5, k6, k7, k8 в соответствии с количеством используемых блоков замены. Для каждой тетрады надо производить опробование 16 вариантов возможных значений фрагмента ключа (от 0000 до 1111). Те варианты, которые будут сохранять значение разности в тетрадах XR и XR1

после целочисленного сложения, считаются возможными вариантами значений для фрагмента раундового подключа. Кроме того, необходимо учитыватьь тот факт, что при сложении тетрады сообщения (XR или XR1) с тетрадой раундового подключа может возникнуть перенос из младших разрядов. Поэтому для всех тетрад, кроме самой младшей, необходимо рассмотреть вариант, когда к результату сложения тетрады сообщения и тетрады ключа будет добавляться еще 1 в младший разряд тетрады. При анализе правильных пар текстов некоторые значения для каждой тетрады раундового подключа будут встречаться чаще других, что позволит сделать предположения о возможных значениях раундового подключа.

Повысить результаты анализа позволяет прием, который использовали Э. Бихам (E. Biham) и А. Шамир (A. Shamir) при анализе алгоритма DES [1]. При поиске правильных пар текстов для полного 16-раундового DES они рассматривали лишь правую часть выходной характеристики. Левая часть выходной характеристики могла быть равна любому значению. В этом случае вероятность прохождения разности через последний раунд шифрования не учитывалась при определении общей вероятности характеристики. Тот факт, что правая часть правильных пар текстов совпадает с правой частью рассматриваемой характеристики, позволяет предположить, что разность текстов при прохождении через раунды шифрования преобразовывалась именно так, как это было определено при построении характеристики. И лишь в последнем раунде разность преобразовывалась одним из возможных способов.

Таким образом, если при поиске правильных пар текстов считать, что критерием для их отбора является только правая часть выходной разности, то это позволит проводить анализ последнего раунда шифрования более развернуто. В этом случае необходимо будет получить значение разности на выходе операции целочисленного сложения по модулю 2^{32} функции F последнего раунда шифрования. Для этого необходимо разность, образованную левыми частями шифрованных сообщений правильных пар текстов, сложить по модулю два с левой частью предполагаемого значения разности на входе последнего раунда шифрования. Таким образом, будет получено значение разности, которое появилось на выходе функции F последнего раунда шифрования. После этого полученное значение разности необходимо сдвинуть вправо на 11 позиций и в соответствии с таблицей анализа для блоков замены определить значение разности, которое поступало на вход S-блоков. Это и будет ожидаемое значение разности на выходе операции целочисленного сложения по модулю 2^{32}. Дальнейшее опробование ключей необходимо выполнять так, как это было описано для первого раунда шифрования.

В результате такого анализа правильных пар текстов будет сформировано некоторое количество возможных значений (возможно, их будет несколько тысяч) раундового подключа K1. После этого необходимо провести анализ тех же правильных пар текстов, но уже с использованием не фрагментов подключа, а с выделенными возможными значениями полного 32-битового раундового подключа. В результате такого опробования останется всего несколько ключей (меньше десяти), один из которых и будет являться истинным раундовым подключом.

Результаты. Для того, чтобы убедиться в действенности метода, была сымитирована атака на алгоритм ГОСТ, использующий один слабый блок замены, а также сокращенный до 12 раундов. При этом были использованы 8 раундовых подключей так, чтобы в первом и последнем раунде использовался один и тот же раундовый подключ. В качестве блоков замены был использован слабый S-блок, определенный в табл. 3. Атака на полный ГОСТ будет выглядеть аналогичным образом, за тем исключением, что потребуется опробовать большее число пар текстов, прежде чем правильная пара текстов будет найдена.

Итак, первоначально было найдено 10 характеристик для 12 раундов алгоритма ГОСТ, которые приведены в табл. 4. Из таблицы видно, что правые части выходных разностей ΔY имеют от одной до трех ненулевых тетрад. При этом каждая тетрада хотя бы один раз имеет значение 1_x, 2_x или 4_x.

Таблица 4

Характеристики для 12 раундов алгоритма ГОСТ

№	ΔX	ΔY	ΔYL^{-1}	p
1	00001000 00000001	00010000 10000000	00010100	$\frac{1}{2^{18}}$
2	00000800 00000004	00000004 00200804	00200800	$\frac{1}{2^{18}}$
3	00020000 00000008	00000008 04002008	04002000	$\frac{1}{2^{18}}$
4	00010000 00000010	00100000 00000001	00101000	$\frac{1}{2^{18}}$
5	00040000 00000020	00000020 08040020	08040000	$\frac{1}{2^{18}}$
6	00008000 00000040	00000040 02008040	02008000	$\frac{1}{2^{18}}$
7	00020000 00000080	00000080 40020080	40020000	$\frac{1}{2^{18}}$
8	00100000 00000100	01000000 00000010	01010000	$\frac{1}{2^{18}}$
9	00400000 00000200	00000200 80400200	80400000	$\frac{1}{2^{13}}$
10	80000000 00400000	00400000 80400200	80000200	$\frac{1}{2^{11}}$

Для каждой характеристики из табл. 4 было опробовано 100 000 пар текстов с целью поиска правильных пар текстов. Полученные правильные пары были проанализированы с использованием техники, описанной выше. Так, в результате первичного отбора тетрад секретного ключа было выделено 16 384 возможных ключей из общего числа значений 2^{32}. Последующее опробование полного раундового подключа оставляло в среднем 5 возможных значений (в зависимости от числа найденных правильных пар текстов это значение колебалось от 2 до 10). Полный анализ 12 раундов алгоритма ГОСТ с использованием 10-раундовых характеристик, приведенных в табл. 4, занимал в среднем от 1 до 2 минут (исследования проводились на процессоре Intel Celeron M CPU 530 1.73 GHz, RAM 1007Mb). Было проведено около 1 000 экспериментов с использованием различных значений раундовых подключей, и результат всегда был положителен.

В данной статье рассмотрена возможность проведения атаки на алгоритм ГОСТ с использованием метода дифференциального криптоанализа при условии, что алгоритм использует слабые S-блоки. Было показано, какие блоки могут считаться слабыми, а также был предложен способ быстрого поиска таких блоков. В результате использования слабых S-блоков для полного 32-раундового алгоритма ГОСТ становится возможным быстрое нахождение характеристик с вероятностями, лежащими в диапазоне от 2^{-25} до 2^{-33}, что позволяет надеяться на успешное проведение атаки. Также предложен способ проведения анализа правильных пар текстов с целью определения секретного ключа шифрования.

Для того чтобы проиллюстрировать возможность анализа, была проведена имитация атаки на 12-раундовый алгоритм ГОСТ, использующий слабые блоки. Высокая скорость анализа позволяет надеяться на успех анализа полного 32-раундового алгоритма.

В работе рассмотрено построение характеристик и проведение анализа для алгоритма ГОСТ, использующего один и тот же блок замены. Продолжением данной работы будет исследование влияния на вероятность характеристики сочетания различных слабых блоков замены и выявление тех из них, которые могут привести к вскрытию шифра.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Biham E., Shamir A.* Differential Cryptanalysis of the Full 16-round DES, Crypto'92, Springer-Velgar, 1998. – P. 487.
2. *Biham E., Shamir A.* Differential Cryptanalysis of DES-like Cryptosystems, Extended Abstract, Crypto'90, Springer-Velgar, 1998. – P. 2
3. *Kelsey J., Schnier B., Wagner D.*, Key-Schedule Cryptanalysis of IDEA, G-DES, GOST, SARER, and Triple-DES // http://www.schnier.com – 1996.
4. *Бабенко Л.К., Ищукова Е.А.* Применение рекурсивного алгоритма поиска в Б-деревьях для дифференциального криптоанализа алгоритма шифрования ГОСТ 28147-89 // Материалы IX Международной научно-практической конференции «Информационная безопасность». Ч. 2. – Таганрог: Изд-во: ТТИ ЮФУ, 2007 – С. 92-97.
5. *Saarien M.-J.* A Chosen Key Attack Against the Secret S-boxes of GOST // http://www.m.-js.com – Helsinki University of Technology, Finland.
6. *Панасенко С.П.* Алгоритмы шифрования. Специальный справочник. – СПб.: БХВ-Петербург, 2009. – 576 с.

Статью рекомендовал к опубликованию д.т.н., профессор Я.Е. Ромм.

Бабенко Людмила Климентьевна
Технологический институт федерального государственного автономного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге.
E-mail: blk@fib.tsure.ru.
347928, г. Таганрог, ул. Чехова, 2.
Тел.: 88634312018.
Кафедра безопасности информационных технологий; профессор.

Ищукова Евгения Александровна
E-mail: jekky82@mail.ru.
Тел.: 88634371905.
Кафедра безопасности информационных технологий; доцент.

Babenko Lyudmila Klimentevna
Taganrog Institute of Technology – Federal State-Owned Autonomy Educational Establishment of Higher Vocational Education “Southern Federal University”.
E-mail: blk@fib.tsure.ru.
2, Chekhov Street, Taganrog, 347928, Russia.
Phone: +78634312018.
The Department of Security of Information Technologies; Professor.

Ischukova Evgeniya Aleksandrovna
E-mail: jekky82@mail.ru.
Phone: +78634371905.
The Department of Security of Information Technologies; Associate Professor.

УДК 519.7

С.А. Диченко, А.К. Вишневский, О.А. Финько

РЕАЛИЗАЦИЯ ДВОИЧНЫХ ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ЛИНЕЙНЫМИ ЧИСЛОВЫМИ ПОЛИНОМАМИ

Разработан алгоритм распараллеливания генерации псевдослучайных двоичных последовательностей (ПСП) на основе представления систем порождающих рекуррентных логических формул посредством линейных числовых полиномов. Линейные числовые полиномы, в отличие от общей (нелинейной) формы, обеспечивают высокую скорость вычислений. «Арифметизация» генераторов ПСП позволяет применить известные арифметиче-