УДК 621.391.7

Эффективная атака на блоковый шифр RC6

В. А. Монарев, А. М. Лубкин

Работа посвящена криптоанализу блокового шифра RC6. Предложенная атака основана на результатах Л. Кнудсена и В. Мейера и использует тест хи-квадрат. Схема атаки позволяет значительно уменьшить трудоёмкость нахождения секретного ключа. Ранее известные варианты атак, основанные на тесте хи-квадрат, имеют значительно бóльшую сложность. Например, для 5-раундового шифра RC6 вычислительная сложность атаки составляет 246, а для лучшего из ранее известных методов – 2101.

Ключевые слова: криптоанализ, блочный шифр, RC6.

1. Введение

Шифр RC6 – один из претендентов на звание AES, конкурса, проводимого в 2001 году. Существуют многочисленные работы, посвящённые анализу этого шифра. Большинство из них используют результаты атаки, предложенной в [1], и применимы к шифру RC6, но в несколько урезанном виде, то есть без операции так называемого «забеливания» («post–whitening» и «pre–whitening», см. [2], [3]). Эти атаки основаны на исследовании статистических свойств шифра RC6. Например, получен результат, согласно которому выходную последовательность шифра RC6 можно отличить от случайной последовательности при наличии подходящего количества выбранного шифротекста (для r-го раунда необходимо 2^{8r+10} текста). Последним и наилучшим результатом является работа [2], где описана атака на RC6, позволяющая найти секретный ключ для r–го (r = 2, 4, ...) раунда при использовании 2^{8r-1} блоков выбранного текста и перебора ключа, равному 2^{54}. Таким образом, уже для 4-го раунда временна́я сложность будет значительной – 2^{85}, что делает атаку практически нереализуемой.

В данной работе предлагается новая атака, которая позволяет найти секретный ключ для 5-го раунда, при этом временная сложность равна 2^{46}. Атака также основана на исследовании статистических свойств шифра RC6.

2. Применение теста хи-квадрат

В математической статистике широко известен тест хи-квадрат. Точная постановка задачи следующая. Некоторый источник порождает символы из алфавита $A = a_1, a_2, \ldots, a_s$. Требуется по выборке $x_1, x_2, \ldots, x_n$ проверить гипотезу H_0: $P(a_i) = P_i$. Для этого вычисляется статистика:

$$x^2 = \sum_{i=1}^{s} \frac{(\nu(a_i) - nP_i)^2}{nP_i}, \qquad (1)$$

где $\nu(a_i)$ – частота буквы a_i.

Известно, что распределение случайной величины x^2 асимптотически приближается к распределению χ^2 с $(s-1)$ степенью свободы при выполнении гипотезы H_0. В данной работе будет рассматриваться случай $A = 1,2,\ldots,1023$, $P_i = \frac{1}{1024}$, для всех $i < 1024$. В таблице 1 приведены квантили распределения χ^2 с 1023 степенями свободы.

Таблица 1. Квантили распределения χ^2

$Q_{0.95}$	1098.5208
$Q_{0.99}$	1131.1587
$Q_{0.999}$	1168.4972

Из таблицы, например, следует, что $P(x^2 > Q_{0.95}) = 0.05$.

3. Статистические свойства шифротекста RC6

Одно из обязательных условий, предъявляемых к современным шифрам, можно сформулировать следующим образом: любое зашифрованное сообщение должно быть «похоже» на реализацию бернуллиевского процесса, в которой вероятность порождения нуля и единицы равна 1/2 (дальше такие последовательности будем называть случайными). Этот раздел криптоанализа называется «distinguishing attack». Поэтому проверке статистических свойств шифров посвящено множество работ. Далее будет показано, как «distinguishing attack» может быть применена для поиска секретного ключа.

В данной работе рассматривается самый распространённый вариант шифра RC6, когда на вход подаются 4 блока по 32 бита. Перед описанием схемы шифрования введём необходимые обозначения.

$\mathrm{lsb}_n(X)$: младшие n бит в числе X;
$\oplus$: операция XOR
$a <<< b$: циклический сдвиг a на b бит влево;
$a >>> b$: циклический сдвиг a на b бит вправо;
S_i : i-ый подключ;
R : число раундов шифрования;
(A_i, B_i, C_i, D_i) : вход i-го раунда четыре 32-битных блока (128 бит);
$f(x)$: $x(2x+1)$;
$F(x)$: $f(x) <<< 5$;
$x \| y$: конкатенация x и y.

На начальном этапе секретный 128- (192, 256)-битный ключ преобразуется в массив размера $(2r+4)$ 32-битных подключей $S_0, S_1, ..., S_{r+3}$. С помощью этого массива происходит шифрование данных по следующему алгоритму.

Алгоритм 1 (шифрования)

1. $A_1 = A_0$; $B_1 = B_0 + S_0$; $C_1 = C_0$; $D_1 = D_0 + S_1$;
2. for $i = 1$ to r do:
   $t = F(B_i)$; $u = F(D_i)$; $A_{i+1} = B_i$;

$B_{i+1} = ((C_i \oplus u) <<< (t \bmod 32)) + S_{2i+1}$;

$C_{i+1} = D_i$; $D_{i+1} = ((A_i \oplus t) <<< (u \bmod 32)) + S_{2i}$;

3. $A_{r+2} = A_{r+1} + S_{2r+2}$; $B_{r+2} = B_{r+1}$; $C_{r+2} = C_{r+1} + S_{2r+3}$; $D_{r+2} = D_{r+1}$.

Пункты алгоритма 1, 3 называются операциями «забеливания». У шифра RC6 существуют различные режимы шифрования. Принято обозначение RC6-*w*/*r*/*b* – шифр, на вход которого подается четыре *w*-битных слова, *r* – число раундов шифрования и *b* – длина ключа в байтах. В работе исследовался RC6-32/*r*/16, для различного числа раундов. Перейдем к описанию предлагаемого метода статистического анализа последовательностей, получаемых на выходе шифра.

Алгоритм 2 (тестирование)

1. На вход шифра подаётся последовательность 128-битных блоков данных

$$(0,0,0,0),(1,0,0,0),\ldots,(2^{32}-1,0,0,0),(0,0,1,0),(1,0,1,0),(2,0,1,0),\ldots \qquad (2)$$

2. Каждый блок из (2) шифруется по алгоритму 1 на *r* раундов. Из каждого 128-битного блока шифротекста $(A_{r+2}, B_{r+2}, C_{r+2}, D_{r+2})$ извлекаются 10 бит – $\mathrm{lsb}_5(A_{r+2}) \| \mathrm{lsb}_5(C_{r+2})$ и к полученной выборке 10-битных чисел применяется тест хи-квадрат. То есть вычисляем статистику x^2 по формуле (1). Эксперимент проводился на 1000 случайных ключах, для 4-го и 5-го раундов шифрования. Из результатов, приведённых в таблице 1, видно, что статистика x^2 на 4-ом раунде при длине выборки 2^{12} превышает квантиль $Q_{0.99}$ в 963 случаях из 1000 (под размером выборки подразумевается количество 128-битных блоков), хотя среднее число превышений квантили распределения должно было быть равно 10.

Таблица 2. Статистические свойства шифра RC6

r	Размер выборки	Среднее значение x^2	$x^2 > Q_{0.95}$	$x^2 > Q_{0.99}$	$x^2 > Q_{0.999}$
4	2^{10}	1342.8	730	608	488
4	2^{11}	1604.5	908	831	713
4	2^{12}	2068.8	986	963	930
5	2^{24}	3716.3	692	624	559
5	2^{25}	5744.3	801	711	635
5	2^{26}	9364.0	921	862	789

Это означает, что на 4-ом раунде выходная последовательность шифра неравномерная, что может быть обнаружено на выборке 2^{12} для более чем 98 % ключей. Если значение статистики x^2 для одной последовательности больше, чем для другой, то будем говорить, что эта последовательность «менее случайна». Из результатов можно сделать вывод, что случайность последовательности (2) после её шифрования тем больше, чем больше число раундов у шифра.

4. Описание атаки на RC6 и результаты

Отметим, что данная атака является атакой третьего типа или атакой по «выбранному шифротексту». Основная идея атаки близка к атакам, предложенным в [1]–[4]. В этих атаках

используется различие между последовательностями, зашифрованными на r и на $(r+1)$ раунда, которое было отмечено в предыдущем разделе.

В предлагаемом методе (по аналогии с методом, описанным в [4]) на вход шифра подаётся последовательность, далёкая от случайной, что позволяет находить отклонения от случайности на существенно меньших выборках, чем в методах, описанных в [1]–[3]. Также изменён алгоритм перебора ключей. Это позволяет найти секретный ключ со сложностью $T\times 2^{32}$ (T – необходимое количество выбранного шифротекста или количество 128-битных блоков данных) вместо $T\times 2^{54}$, что существенно уменьшает количество операций. Для начала опишем алгоритм, который позволит понять основную идею.

Утверждение 1. *Если в алгоритме шифрования RC6 известны S_0 и S_1, то можно подобрать текст таким образом, что шифротекст (выходная последовательность) будет в той же степени случаен, что и шифротекст при $(r-1)$–раундовом шифровании последовательности (2).*

Доказательство. Обозначим через (a_n, b_n, c_n, d_n) последовательность вида (2). Покажем, что если известны S_0 и S_1, то можно выбрать текст таким образом, что $A_2 = a_n$, $B_2 = b_n + S_3$, $C_2 = c_n$, $D_2 = d_n + S_2$. Действительно, полагая $D_0 = c_n - S_1$, $B_0 = a_n - S_0$,

$A_0 = (d_n >>> (u \bmod 32)) \oplus t$, $C_0 = (b_n >>> (t \bmod 32)) \oplus u$, где $u = F(c_n)$ и $t = F(a_n)$, получим требуемое. Поскольку на вход второго раунда шифра поступает последовательность $(a_n, b_n + S_3, c_n, d_n + S_2)$, то фактически мы получим на выходе последовательность (a_n, b_n, c_n, d_n), зашифрованную на $(r-1)$ раунд.

Перейдём к описанию полной версии алгоритма атаки.

Алгоритм 3 (атака со сложностью перебора 2^{64})

1. Для всех возможных пар S_0' и S_1' вычисляем последовательность, описанную в утверждении 1, которую подаём на вход шифра.
2. Для каждой из зашифрованных последовательностей вычисляем статистику x^2, используя алгоритм 2.
3. Находим пару S_0' и S_1', для которых значение x^2 максимально. Полагаем их равными искомым S_0 и S_1.

Поясним основную идею атаки. На вход подаётся последовательность, которая зависит от значений S_0 и S_1. Если значения подключей найдены правильно, то во второй раунд шифра будет поступать последовательность (2), следовательно, отклонение от случайности может быть зафиксировано в соответствии с таблицей 2. Если же значения будут неправильно подобраны, то случайность выходной последовательности будет выше. Действительно, в таблице 3 приведены результаты тестирования.

Таблица 3. Оценка эффективности подбора пары подключей

Номер ключа	r	Размер выборки	x^2 для прав. ключа (S_0|S_1)	Макс. x^2 из 1000 непр. ключей	x^2 для 1000 непр. ключей
1	5	2^{12}	1344.5	1162.5	1023.5
2	5	2^{12}	1869.0	1179.5	1021.5
3	5	2^{12}	2943.0	1164.0	1023.3
4	5	2^{12}	2457.5	1211.5	1021.8
5	5	2^{12}	1196.0	1182.5	1024.7
6	6	2^{26}	17054.2	1187.7	1024.1
7	6	2^{26}	2415.8	1183.6	1024.2
8	6	2^{26}	2670.8	1163.4	1020.6
9	6	2^{26}	17510.4	1166.5	1021.9
10	6	2^{26}	2873.0	1170.6	1023.6

Результаты получены для 5 случайных ключей и 5 и 6-раундового RC6. Для каждого правильного ключа было проверено 1000 неправильных случайных ключей, подсчитана статистика x^2 для правильного ключа и x^2 для неправильного, также подсчитано среднее значения x^2 для неправильных ключей. Из таблицы 3 видим, что величина x^2 для правильного ключа больше, чем максимальное значение x^2 среди 1000 случайно выбранных неправильных ключей. Как показали эксперименты, при относительно небольшом увеличении размера выборки эта разница становится ещё более существенной. Заметим, что подбор подключей S_0 и S_1 можно производить по отдельности. В этом случае сложность перебора составит 2^{32}.

Алгоритм 4 (атака со сложностью перебора 2^{32})

1. Для каждого $S_0^{'}$ вычисляем последовательность $D_0 = c_n$, $B_0 = a_n - S_0^{'}$, $A_0 = b_n \oplus t$, $C_0 = d_n$, где $t = F(a_n)$, которую подаём на вход шифру.
2. Для каждой последовательности вычисляем статистику x^2, используя алгоритм 2.
3. Находим значение $S_0^{'}$, для которого значение x^2 максимально. Полагаем его равным искомому S_0.

В таблице 4 приведены результаты проверки эффективности алгоритма 4. Для каждого из 10 ключей было проверено 220 случайных неправильных ключей. Например, для 2-го ключа значение статистики x^2 равно 1996.9 для правильного ключа, в тоже время 1283.1 – это максимальное значение x^2 из 220 неправильных ключей (для «самого подозрительного» ключа). Таким образом, в данном случае ключ S_0 будет найден правильно.

Таблица 4. Оценка эффективности подбора одного подключа

Номер ключа	r	Размер выборки	x^2 для правильного ключа (S_0|S_1)	Макс. x^2 для 2^{20} неправильных ключей	Среднее x^2 для 2^{20} неправильных ключей
1	5	2^{14}	2360.4	1257.9	1022.95
2	5	2^{14}	1996.9	1283.1	1022.96
3	5	2^{14}	1774.0	1264.5	1023.01
4	5	2^{14}	1961.4	1245.8	1023.01
5	5	2^{14}	1728.3	1252.6	1023.00
6	5	2^{14}	2994.9	1250.9	1023.02
7	5	2^{14}	1381.1	1259.6	1022.97
8	5	2^{14}	1193.4	1262.9	1022.94
9	5	2^{14}	1544.5	1269.9	1023.10
10	5	2^{14}	1475.6	1289.8	1023.00

Аналогично можно найти подключ S_1. Для того чтобы найти правильно S_0 и S_1, мы рекомендуем из 2^{32} подключей выбрать 2^{15} «самых подозрительных» (для которых статистика x^2 больше, чем для других подключей) и проверить пары «самых подозрительных» по той же схеме, что и в алгоритме 3. Ясно, что общее число операций для такого перебора составит $T \times 2^{30}$, что несущественно увеличит общую сложность, но повысит эффективность атаки. Из таблицы 4 видно, что правильный S_0 будет (с большой вероятностью) содержаться среди «самых подозрительных» ключей из 2^{32} возможных ключей (исключение – 8-й эксперимент). Отметим, что сложность атаки для 5-го раунда составит 2^{46}. Для этого необходима выборка 2^{14} для каждого из 2^{32} ключей.

Литература

1. Knudsen L., Meier W. Correlations in RC6 with a reduced number of rounds // FSE 2000. LNCS 1978(2000). Springer–Verlag. P. 94–108.
2. Miyaji A., Nonaka M. Evaluation of the security of RC6 against the χ^2-attack // IEICE Trans. Fundamentals. Vol.E88–A, No.1. 2005.
3. Isogai N., Matsunaka T., Miyaji A. Optimized χ^2-attack against RC6 // ACNS 2003. LNCS 2846(2003). Springer–Verlag. P. 199–211.
4. Рябко Б.Я., Монарев В.А., Шокин Ю.И. Новый тип атак на блоковые шифры // Проблемы передачи информации. т. 41, н.4. 2005. с.181– 128.
5. Miyaji A., Nonaka M. Cryptanalysis of the reduced–round RC6 // Proc. ICICS 2002. LNCS 2513. Springer. P. 480–494.

Статья поступила в редакцию 20.11.2010.

Монарев Виктор Александрович
к.ф.-м.н., научный сотрудник ИВТ СО РАН, (630090, Новосибирск, Академика Лаврентьева пр., 6), e-mail: viktor.monarev@gmail.com.

Лубкин Алексей Михайлович
ассистент СибГУТИ, (630102, Новосибирск, ул. Кирова, 86) тел. (383) 2-698-272, e-mail: realsoft@mail.ru.

Effective attack on block cipher RC6

V. Monarev, A.Lubkin

The paper is devoted to new cryptanalysis of block cipher RC6. The offered attack is based on results of L. Knudsen and W. Meier and uses the chi-square test. The attack scheme allows to reduce the complexity of finding a secret key. Earlier known variants of attacks based on the chi-square test, have considerably greater complexity. For example, for the 5-round cipher RC6, the complexity of the suggested attack is 2^{46} operations, while for the best of earlier known methods it is 2^{101} operations.

Keywords: cryptanalysis, block cipher, RC6.