Correlations in RC6

Lars R. Knudsen
Department of Informatics, University of Bergen, N-5020 Bergen

Willi Meier
FH-Aargau, CH-5210 Windisch

July 29, 1999

Abstract

In this paper the block cipher RC6 is analysed. RC6 is submitted as a candidate for the Advanced Encryption Standard, it has 128-bit blocks and supports keys of 128, 192 and 256 bits, and is an iterated 20-round block cipher. Here it is shown that versions of RC6 with 128-bit blocks can be distinguished from a random permutation with up to 15 rounds; for some weak keys up to 17 rounds. Moreover, with an increased effort key-recovery attacks can be mounted on RC6 with up to 15 rounds faster than an exhaustive search for the key.

Keywords. Cryptanalysis. Block Cipher. Advanced Encryption Standard. RC6.

1 Introduction

RC6 is a candidate block cipher submitted to NIST for consideration as the Advanced Encryption Standard (AES). RC6 (see [10]) is an evolutionary development of RC5. Like RC5, RC6 makes essential use of data-dependent rotations. New features of RC6 include the use of four working registers instead of two, and the inclusion of integer multiplication as an additional primitive operation. RC6 is a parameterized family of encryption algorithms, where RC6-$w/r/b$ is the version with word size w in bits, with r rounds and with an encryption key of b bytes.

The AES submission is the version with $w = 32$, $r = 20$, and RC6 is a shorthand notation for this version, whereby the key length can be $b = 16, 24$, and 32 bytes, respectively. In [2, 3] the security of RC6 has been evaluated with respect to differential and linear cryptanalysis. It was concluded that RC6 is secure with respect to differential cryptanalysis for 12 or more rounds. For linear cryptanalysis, some variants are considered in [2]. It was found that a two-round iterative linear approximation leads to the most effective basic linear attack applicable up to 13 rounds. However, no specific method for key-recovery was given. Furthermore, in [2] some potential enhancements of linear attacks using multiple approximations and linear hulls are sketched, and it is estimated that 16 rounds of RC6 can be

attacked using about 2^{119} known plaintexts. These additional considerations on linear cryptanalysis were used to set a suitable number of rounds for RC6 to be $r = 20$.

In this paper we investigate two-round iterations which are quite different from those considered in [2]. Instead of tracing bitwise linear approximations, we consider input-output dependencies by fixing the least significant five bits in the first and third words of the input block. The correlations of the corresponding two 5-bit integer values at the output are caused by specific rotation amounts in the data dependent rotations and can be effectively measured by χ^2 tests. As confirmed by extensive experiments, this leads to an efficient statistical analysis which considerably improves over the basic linear attack. Estimates of the complexity of our analysis imply that reduced round versions of RC6 with up to 15 rounds are not random.

The linear attacks in [2] deal with correlations between input and output bits, but they do not involve key bits, whereas our statistical analysis can be used to develop a method to find all round subkeys.

This attack is faster than an exhautive key search for the 128-bit version of RC6 with up to 12 rounds, and for the 192-bit and 256-bit versions of RC6 with up to 14 and 15 rounds.

After completion of the first version of this report [7], our attention was drawn to an earlier result by Baudron et al in [1] where an attack similar to ours is outlined. (See also [4].) Their attack distinguishes RC6 up to 15 rounds from a random permutation with a complexity of 2^{125} for 15 rounds. Although based on the same idea, their attack is less efficient than the attacks in this paper and no key-recovery algorithm is reported.

In the following we briefly recall the description of RC6, see Figure 1. For a detailed description we refer to [10]. The user-key has length b bytes and the $4w$-bit plaintext block is loaded into words A, B, C, D. These four w-bit words also contain the ciphertext at the end. The key-schedule (see [10]) expands the user-key into subkeys $S[0], S[1], ..., S[2r+3]$. In our considerations we shall not make use of the detailed description of the key-schedule, but we assume the subkeys to be uniformly random. To describe the encryption algorithm the following notation is used: $(A, B, C, D) = (B, C, D, A)$ means the parallel assignment of values on the right to registers on the left. Moreover, $a \times b$ denotes integer multiplication modulo 2^w, $a <<< \lg w$ means fixed rotation of the w-bit word a by $\lg w$, the base-two logarithm of w, and $a <<< b$ denotes rotation of a to the left by the amount given by the least significant $\lg w$ bits of b.

This paper is organized as follows: In section 2 we review χ^2 tests as a useful tool to detect nonuniformness in probability distributions. In Section 3 the relationship between small rotation amounts and correlation in RC6 is investigated and a class of weak keys is identified. In Section 4 distinguishing and key-recovery attacks are developed, and in Section 5 we draw some conclusions.

Input:	Plaintext stored in four w-bit registers A, B, C, D Number r of rounds w-bit round keys $S[0], ..., S[2r+3]$
Output:	Ciphertext stored in A, B, C, D

Procedure:	$B = B + S[0]$
	$D = D + S[1]$
	for $i = 1$ **to** r **do**
	{
	$t = (B \times (2B + 1)) <<< \lg w$
	$u = (D \times (2D + 1)) <<< \lg w$
	$A = ((A \oplus t) <<< u) + S[2i]$
	$C = ((C \oplus u) <<< t) + S[2i + 1]$
	$(A, B, C, D) = (B, C, D, A)$
	}
	$A = A + S[2r + 2]$
	$C = C + S[2r + 3]$

Figure 1: Encryption with RC6-$w/r/b$.

2 χ^2 tests

In this section we recall how to distinguish a random source with unknown probability distribution p_X from a random source with uniform distribution p_U. A common tool for this task is the χ^2 test, which is briefly recalled together with some useful facts (see e.g., [5], [6], [8], [11]). We shall later use χ^2 tests to detect correlation between specific input and output subblocks of r-round RC6.

Let $\mathbf{X} = X_0, X_1, ..., X_{n-1}$ be independent and identically distributed random variables taking values in the set $\{a_0, a_1, ..., a_{m-1}\}$ with unknown probability distribution. Then the χ^2 test is used to decide if an observation $X_0, X_1, ..., X_{n-1}$ is consistent with the hypothesis $Pr\{X = a_j\} = p(j)$ for $0 \leq j < m$, where $p_X = \{p(j)\}$ is a (discrete) probability distribution on a set of m elements. Let $N_{a_j}(\mathbf{X})$ denote the number of times the observation $\mathbf{X}$ takes on the value a_j. Then obviously $\sum_i N_{a_j}(\mathbf{X}) = n$. The χ^2 statistic is the random variable defined by

$$\chi^2 = \sum_{j=1}^{m} (N_{a_j}(\mathbf{X}) - np(j))^2 / np(j) \tag{1}$$

For the uniform distribution p_U, the χ^2 statistic is just $m/n \sum_i (N_{a_j}(\mathbf{X}) - n/m)^2$. In a χ^2 test, the observed χ^2 statistic is compared to $\chi^2_{a,m-1}$, the threshold for the χ^2 test with $m-1$ degrees of freedom and with significance level a. In our investigation of RC6, we shall specifically need the threshold values for 1023 degrees of freedom, as shown in Tables 1 and 2. For example,

the entry 1131 for 0.99 in Table 1 says that the expression $m/n \sum_i (N_{a_j}(\mathbf{X}) - n/m)^2$ for large n will exceed 1131 only in 1% of the time, provided the underlying distribution of the observation $\mathbf{X}$ is indeed uniform.

Level	0.50	0.60	0.70	0.80	0.90	0.95	0.99	0.999	0.9999
χ^2	1022	1033	1046	1060	1081	1098	1131	1168	1200

Table 1: Selected threshold values of the χ^2 distribution with 1023 degrees of freedom.

Level	$1-2^{-16}$	$1-2^{-24}$	$1-2^{-32}$	$1-2^{-48}$	$1-2^{-64}$
χ^2	1222	1280	1330	1414	1474

Table 2: Selected threshold values of the χ^2 distribution with 1023 degrees of freedom.

For practical experiments the question arises how large the size n of the observation should be in order to detect that a distribution p_X is nonuniform. In order to estimate n, consider the bias of a probability distribution p_X defined by the distance measure

$$||p_X - p_U|| = \sum_j (p_X(j) - p_U(j))^2 \tag{2}$$

From [5] we quote the expected value of the χ^2 statistic (1) of a distribution p_X, as well as some useful conclusions:

$$\mathbf{E}_X \chi^2 = nm||p_X - p_U|| + m - m||p_X|| \tag{3}$$

For the case of the uniform distribution this implies $\mathbf{E}_U \chi^2 = m-1$. Moreover it follows that for $n = c/||p_X - p_U||$ the expected value is $\mathbf{E}_X \chi^2 = cm + m - m||p_X||$. Since in practical cases often $||p_X|| \approx ||p_U||$, this simplifies to $\mathbf{E}_X \chi^2 \approx (c+1)m - 1$. Thus $\mathbf{E}_X \chi^2$ differs from $\mathbf{E}_U \chi^2$ significantly, if $c = \Omega(1)$. As a conclusion, the size $n = c/||p_X - p_U||$ of the observation suffices to distinguish a source with distribution p_X from a source with uniform distribution. Clearly, the constant c needs to be larger for higher significance level a.

3 Correlations in RC6

In [2], under the title of Type I Approximations, a two-round linear approximation has been studied which is based on small rotation amounts in the data dependent rotations. This linear approximation is described by $(A \cdot e_t) \oplus (C \cdot e_s) = (A'' \cdot e_u) \oplus (C'' \cdot e_v)$. Here A and C are the first and third words of some intermediate data, A'' and C'' are the first and third words of the intermediate data after a further two rounds of encryption in RC6, and e_t denotes the 32-bit word with a single one in the t^{th} least significant bit position. It has been noticed that for $t = s = u = v = 0$ the case where

both rotation amounts are zero in the first of the two rounds leads to a bias of 2^{-11}. This is derived by using the piling-up lemma and the fact that the second and fourth words remain unchanged in the second round. If t, s, u, v are nonzero but less than 5, there is a smaller bias, which depends on the values of t, s, u, v. Note that no key bits are involved in the approximation.

In our approach we do not consider the XOR of single bits in the first and third words. Instead we fix each of the least significant five bits in words A and C of the input and investigate the statistics of the 10-bit integer obtained by concatenating each of the least significant five bits in words A'' and C'' every two rounds later. This is motivated by the fact that the least significant five bits in A and C altogether are not changed by the xor and data dependent rotation if both rotation amounts are zero. More generally, we can expect a bias for amounts smaller than five. As we shall demonstrate, this leads to much stronger biases which can be iterated over many rounds, just as linear approximations. In this way we can consider small rotation amounts as a single event, in which amounts near zero from the negative, like 30 or 31, prove to be useful as well.

3.1 Small Rotation Amounts

To see the effect of small rotation amounts on the values of the least significant five bits in the first and third words in RC6, we implemented the following tests with 2 rounds:

Let us denote by (a, b) the two amounts in the data dependent rotations in the first round. To measure the effect on the distribution of the target bits, we forced the values of a and b by taking appropriate plaintexts and we computed the χ^2-value of the 10-bit integers after 4 rounds. For each experiment we took 2^{18} texts, although about 2^{13} texts would do (see Section 3.2), but we deliberately chose a big χ^2-value to clearly measure the effect.

a, b	χ^2
0,0	2775
0,31	2107
0,1	1998
31,31	1715
1,1	1643
0,30	1633
0,2	1572
30,31	1388
1,2	1326
0,3	1306
0,4	1145
0,5	1053

Table 3: Statistical effect of small rotation amounts

By the symmetry in the design of RC6, it can be expected that (a, b) gives

the same χ^2-value as (b, a).

We observe that the χ^2-values for all pairs (a, b) with $|a| < 5$ and $|b| < 5$ are significantly higher than the expected value 1023 for uniform 10-bit integers. Note that these tests suggest that we get similar χ^2-values for constant values of the "distance" $|a - 32| + |b - 32|$ (mod32): The pairs (0,31) and (0,1) both have distance 1 and similar χ^2-values, and the pairs (1,1),(0,2), (31,31), and (0,30) all have distance 2 and have similar χ^2-values.

Let us take a closer look at how the above observations lead to a nonuniform distribution. Assume that the least significant five bits of plaintext words A and C are fixed, e.g., to zero bits. Let us denote by X the concatenation of the least significant five bits of the ciphertext words A and C after two rounds of encryption. In this example, for illustration, we will ignore the addition of the subkeys in the output transformation, and also we will assume that the least significant five bits of both round keys $S[2]$ and $S[3]$ are zero. Denote by t_5 and u_5 the least significant five bits of t and u, see Figure 1. Then in the first round, if $t_5 = u_5 = 0$, then X will be zero. Since the function f is a permutation, t_5 and u_5 will be zero with a probability of 2^{-5} each. If we assume that for $t_5 \geq 5$ and $u_5 \geq 5$, the values of X will be distributed uniformly at random, the probability that X is zero is at least $2^{-10} + (27/32 \cdot 1/32)^2 \simeq 2^{-10} + 2^{-10.5}$. With rotations $t_5 = 1, u_5 = 0$, X will take the possible values (in bits) $0000b00001$, where 'b' is a random bit. With rotations $t_5 = 0, u_5 = 1$, X will take the possible values (in bits) $000010000b$. Thus, $X = 0000100001$ with probabilty at least $2 \cdot 2^{-11} + (27/32 \cdot 1/32)^2$. Note that both these estimates are lower bounds. E.g., in the case where $t_5 = u_5 = 4$, X will take the possible values (in bits) $0b_1b_2b_3b_40b_5b_6b_7b_8$, and in the case where $t_5 = 1, u_5 = 16$, X will take the possible values (in bits) $b_1b_2b_3b_4b_50000b_6$, where the b_is are random bits. Thus, X can take both the values 0000000000 and 0000100001 also in these cases.

It has been clearly demonstrated that the distribution of X is nonuniform. Note that although it was assumed that the involved subkey bits were zero, it follows easily that the nonuniformity remains when these key bits are randomly chosen.

3.2 χ^2 statistic of RC6

Here, we investigate the nonrandomness of r-round versions of RC6. This analysis is based on systematic experiments on increasing numbers of rounds of RC6 with varying word length w. Our method is used to demonstrate that detecting and quantifying nonrandomness is experimentally feasible up to 6 rounds of RC6.

For this purpose, the least significant lg w bits in words A and C of the input are fixed to zero. Depending on the experiment and the number of rounds, the remaining input bits are either chosen randomly, or more of the remaining input bits are suitably fixed so that one (or both) of the data dependent rotations are zero. In our tests, we persue the χ^2 statistic of the integer of size twice lg w bits as obtained by concatenating the least

significant lg w bits in words A'' and C'' every two rounds later.

In the experiments, we consider versions of RC6 with word size $w = 8, 16$ and 32 bits, respectively ($w = 32$ corresponding to the AES candidate RC6). It is instructive to see that the general behaviour of the χ^2 test for increasing numbers of rounds in all three cases is very similar. To judge the outcome of these χ^2 tests note that for the word sizes w as considered, 6-bit, 8-bit and 10-bit integers are tested at the output. Hence the numbers of freedom are 63, 255 and 1023 respectively, and these numbers coincide with the expected value of the χ^2 statistic, provided the distribution to be tested is uniform.

Subsequently we discuss the results of implemented tests in more detail, where the keys are chosen at random.

32-bit RC6. First consider a version of RC6 with block length of 32 bit. This corresponds to the case $w = 8$, which is shown in Table 4. For $r = 2$ and $r = 4$ rounds more than one entry is given. The first entry shows a number of texts, measured in powers of two, which is necessary to detect that the mean of the χ^2 values over 20 tests is higher than the expected value 63 if the distribution would be random. The other entries show a significant increase of this mean if the number of plaintexts is doubled, thus a strong deviation from the uniform distribution. For 2^8, 2^{17} and 2^{26} texts and correspondingly for 2, 4 and 6 rounds, the χ^2 values are approximately the same. Thus we have to increase the number of plaintexts by the same factor 2^9 for every two more rounds to get a comparable statistical deviation as measured by the χ^2 test. For this small version of RC6 we cannot go beyond 6 rounds, as we have to fix 6 input bits, and for 6 rounds we already need 2^{26} random texts.

r	#Texts	χ^2	#Tests
2	2^8	77	20
2	2^9	107	20
4	2^{16}	68	20
4	2^{17}	73	20
4	2^{18}	83	20
6	2^{26}	78	20

Table 4: RC6 with 32-bit blocks and r rounds. Expected χ^2 for a random function is 63.

64-bit RC6. Next consider the version of RC6 with word size $w = 16$, i.e. RC6 with 64-bit blocks. The results are shown in Table 5. Here the expected value of the χ^2 statistic is 255. Again a substantial increase is observed in the mean for χ^2-values if the number of texts is doubled. We notice that passing from 2 to 4 to 6 rounds, the averaged χ^2-values increase slightly if the corresponding number of plaintexts is increased by a constant factor of 2^{13}.

128-bit RC6. Consider now r-round versions of RC6 with word size 32 bits, i.e. with round function as in the AES proposal. Table 6 shows the

r	#Texts	χ^2	#Tests
2	2^{10}	283	100
2	2^{11}	308	100
2	2^{12}	364	100
4	2^{23}	286	100
4	2^{24}	318	100
6	2^{36}	298	10

Table 5: RC6 with 64-bit blocks and r rounds. Expected χ^2 for a random function is 255.

results of implemented tests for $r = 2$ and $r = 4$ rounds. Recall that for 10-bit integers the expected value of the χ^2 statistic is 1023, and according to Table 1 the 95% significance level is 1098 and the 99% significance level is 1131. Thus all tests as reported in Table 6 are very unlikely to be produced by uniformly distributed 10-bit integers. In fact for 4 rounds and 2^{33} texts almost twice the expected value for a uniform distribution is achieved.

r	#Texts	χ^2	#Tests
2	2^{13}	1096	20
2	2^{14}	1196	20
2	2^{15}	1332	20
2	2^{16}	1649	20
2	2^{17}	2208	20
4	2^{29}	1096	20
4	2^{30}	1163	20
4	2^{31}	1314	20
4	2^{32}	1527	20
4	2^{33}	2054	20

Table 6: RC6 with 128-bit blocks and r rounds. Expected χ^2 for a random function is 1023.

Table 7 shows the results of tests with up to 6 rounds but with one or both data dependent rotations in the first round to be fixed to zero. The last entry is the result of a test run on eight processors of a Cray Origin 2000 computer. Both, the experiments in Table 6 and in Table 7 demonstrate that for up to 6 rounds each additional two rounds require roughly 2^{16} times as many texts to get about the same χ^2-value on average.

r	#Texts	χ^2	#Tests	Comments
4	2^{22}	1124	20	zero rotation in 1. round at word D
4	2^{23}	1228	20	zero rotation in 1. round at word D
6	2^{38}	1106	1	zero rotation in 1. round at word D

Table 7: RC6 with 128-bit blocks and r rounds. Expected χ^2 for a random function is 1023.

3.3 A possible analytical explanation

In this subsection we make an attempt to analytically predict the complexities of the χ^2 tests on RC6.

In the following, let X be the random variable representing the 10 bits as considered in the ciphertexts after 2 rounds of encryption with RC6 in the tests from the preceding section. Also, let Y and Z be the random variables representing these 10 bits in the ciphertexts after 4 respectively 6 rounds of encryption. It follows from the description of RC6, that the 10 bits in the ciphertexts after six rounds are not the exclusive-or of 10 biased bits from the first two rounds and 10 biased bits from the next two rounds. This is due to the fact that the data-dependent rotations in RC6 are performed after the exclusive-or with the data from the previous rounds. Thus, a parallel to the Piling-Up Lemma used by Matsui [9] does not seem to be applicable.

With the test results of the preceding section and the estimate from Sec. 2, that with $n = c/||p_X - p_U||$ texts one can expect a χ^2-value of $(c+1)m$, it is possible to compute estimates of $||p_X - p_U||$, $||p_Y - p_U||$, and $||p_Z - p_U||$.

64-bit RC6. The results of the tests in Table 5 yield the following estimates for the distances:
$||p_X - p_U|| = 2^{-13.25}$, $||p_Y - p_U|| = 2^{-26.03}$, $||p_Z - p_U|| = 2^{-38.57}$. Thus, this is a clear indication that $||p_Y - p_U|| > ||p_X - p_U||^2$, and that $||p_Z - p_U|| > ||p_X - p_U|| \cdot ||p_Y - p_U||$.

This gives perhaps more convincing evidence, that passing from s to $s+2$ rounds in the tests of the preceding section, requires an increase in the texts needed of a factor of a little less than 2^{13}.

128-bit RC6. The results of the tests in Table 6 with a χ^2-value greater than 1300 yield the following estimates for the distances:
$2^{-16.79} \leq ||p_X - p_U|| \leq 2^{-16.71}$, $2^{-33.02} \leq ||p_Y - p_U|| \leq 2^{-32.81}$. Again with a clear indication that $||p_Y - p_U|| > ||p_X - p_U||^2$.

This confirms the estimate from the preceding section that passing from s to $s+2$ rounds in the χ^2-tests, requires an increase in the texts needed of a factor of a little more than 2^{16}. Later, we will use the factor $2^{16.2}$.

3.4 Weak keys

The test results from the previous sections were given as an average over tests using randomly chosen keys. There was some deviation of the single results, e.g., the χ^2-values of the tests for RC6 with 128-bit blocks and 4 rounds using 2^{33} texts varied from 1731 to 2595 with an average of 2044. Thus, for some keys the deviation is bigger than expected, for other keys it is lower than expected. In this section we report on some weak keys, which perform better than the average key. In Sec. 3.1 it was explained why there is a nonuniform distribution of the 10 target bits, and why for two rounds the involved key bits have no influence on the nonuniformity of the target bits in the χ^2-tests. However, when iterating the tests to several rounds, the modular additions of round-key bits introduce carry bits which affect the

nonuniformity. For 4 rounds, the key bits that may affect the nonuniformity are the five least significant bits of the round keys $S[2]$ and $S[3]$. When these bits are set to zeros, the χ^2-value increases. Similarly, for 6 rounds the least significant five bits of the subkeys 2,3,6, and 7 may influence the nonuniformity.

This is illustrated by a series of tests, the results of which are shown in Table 8. For 4 rounds the "distance" to a uniform distribution is about

r	#Texts	χ^2	#Tests	Comments
4	2^{30}	1398	20	1 in 2^{10} keys
6	2^{30}	1093	10	zero rotation in 1.round at B and D
6	2^{30}	1368	10	same, for 1 in 2^{20} keys

Table 8: RC6 with 128-bit blocks and r rounds for weak keys.

$2^{-31.5}$ which is more than a factor of two higher than for the results averaged over all keys. For 6 rounds the distance to the uniform distribution is about $2^{-33.87}$ for the second test of Table 8, and about $2^{-31.57}$ for the third test using weak keys. Thus, a factor of more than 4.

4 Attacks on RC6

4.1 Distinguishing attacks

It is possible to exploit the findings in the previous sections to distinguish RC6 with a certain number of rounds from a permutation randomly chosen from the set of all permutations. In the previous sections we fixed bits in the first and third plaintext words. As we shall see in the next section this makes good sense when implementing key-recovery attacks. In a distinguishing attack it is advantageous to fix the least significant five bits in the second and fourth words instead. It follows that after one round of encryption the least significant five bits in the first and third words of the ciphertext are constant. Table 9 lists the result of tests implemented for RC6 with 128-bit blocks with 3 and 5 rounds. It follows that $2^{13.8}$ texts are sufficient to distinguish the 3-round encryption permutation from a randomly chosen permutation in 90% of the cases. We estimate that for RC6 with $3 + 2r$ rounds similar results will hold using $2^{13.8+r\times 16.2}$ texts, which is confirmed by tests implemented on RC6 with 5 rounds.

Note that the χ^2 numbers of Table 9 for 3 rounds are slightly lower than the numbers of Table 6 for 2 rounds. This stems from the fact that in the latter tests, the least significant five bits of the first and third words of the plaintexts were fixed to zeros. In a distinguising attack, one gets the first round "for free", by fixing totally 10 bits of the second and fourth words. However, as these words are added modular 2^{32} to subkeys in the input transformation, the least significant five bits of the first and third words in the inputs to the second round are nonzero, but constant, and there is

an effect of carry bits by the addition of subkeys after the second-round approximation.

We estimate that for keys where the least significant five bits of each of the two subkeys in every second round are zeros, the attack improves with more than a factor of two for each 2 rounds. This leads to the estimate that for one in 2^{80} keys, 17 rounds of RC6 with 128-bit blocks can be distinguished from a randomly chosen permutation.

r	#Texts	χ^2	Comments
3	2^{13}	1079	Implemented, average 20 tests
3	$2^{13.8}$	1100	Implemented, average 20 tests
3	2^{14}	1141	Implemented, average 20 tests
5	2^{29}	1054	Implemented, average 20 tests
5	2^{30}	1099	Implemented, average 20 tests
7	$2^{46.2}$		Estimated.
9	$2^{62.4}$		Estimated.
11	$2^{78.6}$		Estimated.
13	$2^{94.8}$		Estimated.
15	$2^{111.0}$		Estimated.
17	$\leq 2^{118}$		Estimated. For 1 in every 2^{80} keys.

Table 9: Complexities for distinguishing RC6 with 128-bit blocks and r rounds from a random function.

4.2 Key-recovery

As confirmed by several experiments, the χ^2-value is significantly higher if inputs are suitably fixed so that one (or both) of the data dependent rotations in the first round of RC6 are zero. Clearly, the choice of the right input depends on knowledge of the subkey $S[0]$ (or $S[1]$, respecively). We now describe how the considerations and experimental results of previous sections can be exploited for key recovery. Thereby we restrict to 128-bit RC6 with word size 32 bits.

In the following we will assume that to get similar values in a χ^2-test on $s+2$ rounds compared to s rounds requires a factor of $2^{16.2}$ additional plaintexts. Recall that we always fix the least significant five bits in words A and C. In addition suppose we fix inputs so that the data dependent rotation is zero in the first round at word D. Then with a factor of about $2^{8.1}$ less plaintexts we achieve a similar χ^2-value as for random inputs at word D (e.g. compare row 7 in Table 6 with the first row in Table 7). For symmetry reasons, the same holds if inputs at word B are fixed.

With regard to inputs at word D (or B), some comments related to the multiplication in RC6 are in order. The data dependent rotation amounts are determined by the five leading bits of the output of the permutation as given by the multiplication $D \times (2D+1)$ (see Figure 1). The permutation function restricted to these five output bits is therefore balanced. Rather

than fixing inputs we can restrict to inputs leading to these five bits being zero, resulting in more freedom for choosing plaintexts. For efficiency we can prepare a table T of the 2^{27} inputs to the permutation giving zero rotation. Thus for the correct key $S[1]$ we can choose 2^{27} different inputs at word D, all leading to zero rotation in the right half of the first round. (Alternatively, we can enlarge the table, and also accept inputs giving rotation amount 1 or -1, which still lead to increased χ^2-values.) To test a fixed trial key $S[1]$ we thus can roughly choose amongst 2^{113} plaintexts at random.

The attack goes as follows, choose plaintexts such that the least significant five bits of the first and third words are zeros. Prepare an array with 2^{10} entries for each value of the subkey $S[1]$. For each plaintext use the table T as prepared, to determine the values of $S[1]$ which lead to a zero rotation at word D. For each such value, update each array by incrementing the entry corresponding to the value obtained from the 10 target bits of the ciphertext. Each array is used to find the probability distribution of the 10 target bits. Repeat the attack sufficiently many times, until one array has a significantly higher value in the χ^2-test.

For an estimate of the complexity to recover subkey $S[1]$, consider r-round versions of RC6 with r even. For each trial key $S[1]$ we perform a χ^2 test with

$$2^{13} \times (2^{16.2})^{\frac{r-2}{2}} \times 2^{-8.1} \tag{4}$$

plaintexts as described. Then for the correct choice of $S[1]$ the χ^2-value is expected to be around 1100, that is, significantly higher than 1023. For each key which produces an expected χ^2-value, repeat the attack with additional plaintexts.

To rule out all false values of the key, we increase the number of texts by up to a factor of 2^3. Enlarging the amount of plaintexts by this factor has the effect of a substantial increase of the χ^2-value, as observed in our experiments (see the tables in Section 3). Thus, to single out the correct key out of suggested key values we would need about $2^{16} \times (2^{16.2})^{\frac{r-2}{2}} \times 2^{-8.1}$ texts. And since only one in every 2^5 texts gives the desired zero rotation at word D, the total number of plaintexts needed is

$$2^5 \times 2^{16} \times (2^{16.2})^{\frac{r-2}{2}} \times 2^{-8.1} = 2^{r \times 8.1 - 3.3}.$$

The amount of work is estimated as follows. For each plaintext in the attack, we update the counters of at most 2^{27} keys. If we assume that after the first two iterations of the attack, the number of remaining keys are reduced by a factor of 4 or more, we obtain a complexity of

$$2^{27 + r \times 8.1 - 5.3} = 2^{21.7 + r \times 8.1},$$

where one unit is the time to update one entry of one array of size 2^{10} of totally 2^{32} arrays.

After $S[1]$ is correctly found, subkey $S[0]$ can be determined with a reduced amount of texts and work. Knowing $S[0]$ and $S[1]$, the data dependent rotations in the first round can be fixed to zero without effort. Thus the χ^2

tests can now be applied by controlling inputs to the second round. This enables finding subkeys $S[2]$ and $S[3]$ in much the same way as we did for $S[0]$ and $S[1]$. After this we peel of the first round and proceed to determine the other subkeys.

This attack is faster than an exhaustive key search for the 128-bit key version of RC6 with up to 12 rounds, and for the 192-bit and 256-bit versions of RC6 with up to 14 rounds. Table 10 lists the complexity for 12, and 14 rounds of RC6. For 16 rounds the number of texts needed is $2^{126.3}$ and thus exceeds the number of available texts of 2^{118}.

For key sizes 192 bits and 256 bits the computational effort for searching subkeys can be larger. Thus for a 192-bit key we can do a simultaneous search over $S[0]$ and $S[1]$, thereby improving the χ^2 statistic by two rounds. In addition, we increase the factor 2^3 to 2^4 in order to single out the correct pair $S[0]$, $S[1]$ among the remaining pairs. Here only one in every 2^{10} plaintexts give zero rotations at words B and D. The number of plaintexts needed for this version of the attack is

$$2^{10} \times 2^{17} \times (2^{16.2})^{\frac{r-2}{2}-1} = 2^{r \times 8.1 - 5.4},$$

and the time complexity is

$$2^{54 + r \times 8.1 - 7.4} = 2^{46.6 + r \times 8.1},$$

where one unit is the time to update one entry of one array of size 2^{10} of totally 2^{64} arrays. Table 10 lists the complexities of this attack for 14 rounds of RC6. The number of texts needed in the attack on 16 rounds is about 2^{124} and thus still exceeds 2^{118}. However, as reported earlier there are keys for which the complexities improve. We estimate that the attack is possible for at least one in 2^{60} keys with the complexity as stated in the table.

Finally, for the 256-bit key version of RC6 it is possible to further extend the attack. In a 15-round version, one can search over the keys $S[0], S[1], S[32]$, and $S[33]$. The latter two keys are used to decrypt the ciphertexts one round. In the updating of the probability-arrays, one only uses ciphertexts for which there are zero rotations in the last round. The number of texts needed is approximately 2^{10} times that of 14 rounds, and the time complexity increases with a factor of about 2^{54}. To rule out all false values of the keys, we estimate that the number of plaintexts needed increases by yet a factor of 2.

Note that in the above attacks the number of available texts is bounded by 2^{118}, since we need to fix 10 bits of each plaintext. The probability distributions for each such fixed 10-bit value will be different, but their distance to the uniform distribution can be expected to be similar. As an extension of the above attacks consider the following. Run the attack with x texts for one fixed value of the 10 bits in the plaintexts. Record the χ^2-value for each key in the attack, and rank the keys. Reset the arrays. Repeat the attack x texts for another fixed value of the 10 bits. Record again the χ^2-value for each key in the attack, and rank the keys. Repeat this a number of times. If the χ^2-values for the correctly guessed keys will be larger than

r	#Texts	Work	χ^2	Memory	Comment
12	$2^{93.9}$	$2^{118.9}$	1500	2^{42}	
14	$2^{110.1}$	$2^{135.1}$	1500	2^{42}	
14	$2^{108.0}$	$2^{160.0}$	2000	2^{74}	
16	$2^{118.0}$	$2^{171.0}$	2000	2^{74}	1 in 2^{60} keys
15	$2^{119.0}$	$2^{215.0}$	> 2000	2^{138}	

Table 10: Complexities for key-recovery attacks on RC6 with 128-bit blocks and r rounds. One unit in "Work" is the time to increment one counter. The χ^2 value is the expected value for the correct key.

for random values, one can expect that the correct key will be high in the rankings, and it can be detected after sufficiently many iterations. Thus this variant would make available all 2^{128} texts. We conjecture that this attack is applicable to 15 rounds of RC6 with a complexity as given in the last entry of Table 10.

We leave it as an open question whether the attack and its variants can be used to attack RC6 with 16 or more rounds.

Finally, note that the reported attacks are chosen plaintext attacks. However, it follows that the basic attack reported earlier can be easily transformed into a known plaintext attack with an increase in the needed texts of a factor of at most 2^{10}, leaving the total time complexity unaltered.

5 Conclusion

In this paper we have presented an attack on RC6 which is based on a strong relationship between the effects of data dependent rotations in the round function and statistical input-output dependencies.

Estimates which are based on systematic experimental results show that versions of RC6 with up to 15 rounds can be distinguished from a random permutation. A class of weak keys has been identified for which this nonrandomness is estimated to persist up to 17 rounds. Finally, we have derived a method for key-recovery for RC6 with up to 15 rounds which is faster than exhautive key search. It is open whether our analysis can be used to attack RC6 with 16 or more rounds.

We remark that similar attacks are applicable to reduced-round versions of RC5. Work is in progress.

6 Acknowledgments

The authors would like to thank Vincent Rijmen for helpful comments and discussions.

References

[1] O. Baudron, H. Gilbert, L. Granboulan, H. Handschuh, A. Joux, P. Nguyen, F. Noilhan, D. Pointcheval, T. Pornin, G. Poupard, J. Stern, and S. Vaudenay Report on the AES candidates. Available at `http://csrc.nist.gov/encryption/aes/round1/conf2/papers/baudron1.pdf`.

[2] S. Contini, R.L. Rivest, M.J.B. Robshaw and Y.L. Yin. The Security of the RC6 Block Cipher. v.1.0, August 20, 1998. Available at `www.rsa.com/rsalabs/aes/`.

[3] S. Contini, R.L. Rivest, M.J.B. Robshaw, and Y.L. Yin. Improved analysis of some simplified variants of RC6. In L. Knudsen, editor, *Fast Software Encryption, Sixth International Workshop, Rome, Italy, March 1999, LNCS 1636*, pages 1–15. Springer Verlag, 1999.

[4] S. Contini, R.L. Rivest, M.J.B. Robshaw, and Y.L. Yin. Some Comments on the First Round AES Evaluation of RC6. Available at `http://csrc.nist.gov/encryption/aes/round1/pubcmnts.htm`.

[5] J. Kelsey, B. Schneier, and D. Wagner. Mod n cryptanalysis, with applications against RC5P and M6. In L. Knudsen, editor, *Fast Software Encryption, Sixth International Workshop, Rome, Italy, March 1999, LNCS 1636*, pages 139–155. Springer Verlag, 1999.

[6] A.G. Konheim. Cryptography: A Primer. John Wiley & Sons, 1981.

[7] L.R. Knudsen, and W. Meier. Correlations in RC6. Technical Report 177, Department of Informatics,University of Bergen, Norway, July 29, 1999.

[8] D.E. Knuth. The Art of Computer Programming, Vol. 2. Addison-Wesley, 1981.

[9] M. Matsui. Linear cryptanalysis method for DES cipher. In T. Helleseth, editor, *Advances in Cryptology - EUROCRYPT'93, LNCS 765*, pages 386–397. Springer Verlag, 1993.

[10] R.L. Rivest, M.J.B. Robshaw, R. Sidney and Y.L. Yin. The RC6 Block Cipher. v1.1, August 20, 1998. Available at `www.rsa.com/rsalabs/aes/`.

[11] S. Vaudenay. An Experiment on DES Statistical Cryptanalysis. 3rd ACM Conference on Computer and Communications Security, ACM Press, 1996, pp. 139-147.