Theoretical Analysis of "Correlations in RC6"

Masahiko Takenaka, Takeshi Shimoyama and Takeshi Koshiba

Secure Computing Lab., Fujitsu Laboratories Ltd.,
4-1-1 Kamikodanaka, Nakahara-ku, Kawasaki 211-8588, Japan
{takenaka,shimo,koshiba}@flab.fujitsu.co.jp

Abstract. In this paper, we give the theoretical analysis of χ^2 attack proposed by Knudsen and Meier on the RC6 block cipher. To this end, we propose the novel method of security evaluation against χ^2 attack precisely including key dependency by introducing a technique "Transition Matrix Computing." On the other hand, the way of security evaluation against χ^2 attack has not been known except the computer experiment. We should note that it is the first results the way of security evaluation against χ^2 attack is shown theoretically. Using this method, we can obtain the "weakest keys" against the attack.
Keyword: RC6, χ^2 attack, Transition Matrix,

1 Introduction

The block cipher RC6 was proposed by Rivest et al. in [8] to meet the requirements of the Advanced Encryption Standard (AES) and is one of the finalists of the AES candidates. It has been admired for its high-level security and high-speed software implementation especially on Intel CPU. RC6 enters also the NESSIE Project selection and it has been nominated to the Phase II evaluation. Moreover, it is also handled as the candidate for standardization in the CRYPTREC Project in Japan, which has been advanced since 2000.

RC6 is designed based on the block cipher RC5 [7] which makes essential use of arithmetic key additions and data-dependent rotations. As additional primitive operations to RC6, the inclusion of arithmetic multiplications and fixed rotations is believed to contribute the strength of the security of RC6. There are some cryptanalyses of RC6: resistance against Differential Attack, Related Key Attack [2, 3], Linear Attack [1, 2, 3], Mod n Attack [5], and Statistical Attack [4]. Shimoyama *et al.* [9] evaluated the resistance of RC6 with 256-bit key against multiple linear attack and showed that the target key of 14-round RC6 can be recovered and also that the target key of 18-round RC6 with weak keys, which exists with probability $1/2^{90}$ at least, can be recovered. One of the most effective attacks is an attack based on χ^2 test. This attack was originally proposed by Vaudenay [10], and was applied to RC6 by Gilbert et al. [4] and Knudsen and Meier [6], independently. In [6], Knudsen and Meier can

Table 1. Previous Attacks on RC6

Attack	Rounds	Data size	Comments
Linear Attack [1]	16	2^{119}	Upper bound of complexity
Differential Attack [2]	12	2^{117}	Upper bound of complexity
Mod n Attack [5]	—	—	—
χ^2 Attack [6]	15	$2^{119.0}$	Lower bound of complexity (estimation)
	17	$\leq 2^{118}$	Lower bound (estimation, $1/2^{80}$ weak keys)
Multiple Linear Attack [9]	14	$2^{119.68}$	Lower bound
	18	$2^{126.936}$	Lower bound ($1/2^{90}$ weak keys)

cryptanalyze up to 15-round RC6 with general keys and 17-round RC6 with weak keys. We call this attack "χ^2 attack" shortly in this paper. We enumerate attacks on RC6 in Table 1.

In this paper, we study χ^2 attack against RC6 more precisely. Knudsen and Meier [6] experimented with 2-, 4- and 6-round RC6 by χ^2 test and estimated the sample complexity necessary to distinguish $(2r+3)$-round RC6 from random permutation at $2^{16.2r+13.8}$. However, the way of security evaluation against χ^2 attack has not been known except the computer experiment. We analyze the sample complexity (in the sense of chosen plaintext) of Knudsen and Meier's χ^2 attack on RC6 more precisely. We introduce a novel technique "Transition Matrix Computing" to evaluate the expected χ^2 value, and to estimate the sample complexity with respect to any fixed key. We show that the sample complexity with respect to the average key to distinguish $(2r+3)$-round RC6 from random permutation is at most $2^{16.0198r+13.1094}$. We note that the sample complexity $2^{16.2r+13.8}$ estimated by Knudsen and Meier is quite close to our results though their value is drawn from the 20 trials. In addition, Knudsen and Meier indicated that the key, by which the least significant five bits of round key is zero, is weaker than keys in average. And, for these weak keys, they estimated 17-round RC6 can be distinguished from a random permutation with less than the sample complexity 2^{118}. Using our method, we can show that such weak keys, mentioned by Knudsen and Meier, are actual "weakest keys."

Moreover, we show that there exist weak keys in 17-round RC6 whose fraction is $1/2^{69.8747}$, which can be distinguished by using less sample complexity than 2^{118}. Therefore, it is said that weak key ratio is about 1024 times larger than the weak key ratio mentioned by Knudsen and Meier ($1/2^{80}$.)

2 Preliminary

In this section, we give notations referred in what follows.

RC6 is a block cipher proposed by Rivest *et al.* [8]. A version of RC6 is more accurately specified as RC6-$w/r/b$ where the word size is w bits, encryption consists of a nonnegative number of rounds r, and b denotes the length of the encryption key in bytes. Currently RC6 with $r = 20$, $4w = 128$ and $8b = 128, 192, 256$ is recommended to give sufficient resistance against known attacks. (See Table 1.)

Let (A, B, C, D) be an input to RC6 and (A', B', C', D') the corresponding output. Let w-bit strings $key[0], key[1], \cdots$ be a user key for RC6. Let $S[0],\cdots, S[2r + 1]$ be w-bit extended keys for RC6.

In this paper, we use a Feistel-like description of RC6 as same as one described in [9]. (See Appendix.)

For $a \in GF(2)^{32}$, we denote by $lsb_5(a)$ the least significant five bits of a. And we denote by $lsb_5(a, b)$ the concatenation $(lsb_5(a)|lsb_5(b))$ of the couple of each least significant five bits of a and b. For $a, b \in GF(2)^8$, we define $lsb_3(a)$ and $lsb_3(a, b)$, similarly.

We let the sample space be $\Omega = \{(P, RC6(P)) \mid lsb_5(A, C) = 0, P = (A, B, C, D)\}$ and X_n some sample of n examples. We define the χ^2 value for the sample X_n as follows. We not that the user key is implicitly assumed and we omit the user key unless otherwise stated.

$$\chi^2(X_n) = \frac{m}{n} \sum_{a=0}^{m-1} \left(N_a(X_n) - \frac{n}{m} \right)^2,$$

where $N_a(X_n)$ is the cardinality of the set $\{(P, RC6(P)) \in X_n \mid lsb_5(A', C') = a, RC6(P) = (A', B', C', D')\}$, and the value a in the summation is over 0 to 1023 $(= m - 1)$. We note that $lsb_5(a, b)$ (resp., $lsb_3(a, b)$) is a string of 10 bits (resp,. 6 bits.)

3 Expected χ^2 Value and χ^2 Attack

In this section, we consider the expected χ^2 value $E[\chi^2(X_n)]$ and derive it theoretically.

First, we consider a probability distributing on the least significant five bits of the output with respect to RC6 ($lsb_5(A', C')$.) Let, $p(a) = \frac{N_a(\Omega)}{\sharp\Omega}$, that is occurrence probability. We note that $\sum p(a) = 1$.

Proposition 1. *We have following,*

Table 2. The results of Knudsen-Meier's experiment on RC6-8 and RC6-32

RC6-8

rounds	$\sharp Texts$	χ^2	$\sharp Tests$
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

RC6-32

rounds	$\sharp Texts$	χ^2	$\sharp Tests$
2	2^{13}	1096	20
2	2^{14}	1196	20
2	2^{15}	1332	20
4	2^{29}	1096	20
4	2^{30}	1163	20
4	2^{31}	1314	20

$$E[\chi^2(X_n)] = mn \sum_{a=1}^{m} \left(p(a) - \frac{1}{m} \right)^2 + \frac{(\sharp\Omega - n)}{(\sharp\Omega - 1)} \left(m - m \sum_{a=1}^{m} p(a)^2 \right).$$

Proof. (Refer to Appendix for this proof.)

In addition, when $\sharp\Omega$ is sufficiently large to n and the value of each $p(a)$ is close to $1/m$, the expected χ^2 value can be approximated as follows. (See [5, 10].)

Corollary 2. *If $\sharp\Omega \gg n$ and $\sum p(a)^2 \approx 1/m$, then we have*

$$E[\chi^2(X_n)] \approx mn \sum_{a=1}^{m} \left(p(a) - \frac{1}{m} \right)^2 + m - 1.$$

Especially, the expected value $E[\chi^2(X_n)]$ is almost proportional to n.

In [6], Knudsen and Meier obtained the experimental values of $p(a)$ by the computer experiment, and estimated the value of $E[\chi^2(X_n)]$.

In the next section, we propose a new method to obtain the theoretical value of $p(a)$. And we evaluate the value of $E[\chi^2(X_n)]$ precisely.

4 χ^2 Attack

In this section, we review the attack based on χ^2 test, which is proposed by Knudsen and Meier. In this paper, we call this attack "χ^2 attack" for simplicity. On χ^2 attack, the least significant 5 bits in words B and D of input to RC6 are fixed to zero, and the other bits of input are randomly chosen. And we observe the χ^2 values of the 10-bit integer as obtained by concatenating the lease significant bits in words (A', C') that is the output

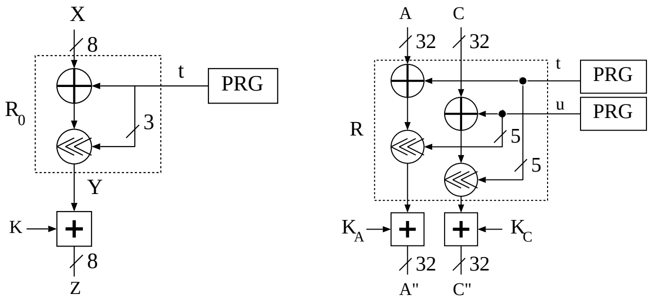


Fig. 1. Basic Models of round function of RC6

of $(2r+1)$-round RC6. If the output of RC6 is perfectly random, the χ^2 value follows the χ^2 distribution with 1023 degree of freedom. When this χ^2 value is larger than 1098, the output of RC6 is able to be distinguished from the uniformly random number with probability of 95%.

Knudsen and Meier experimented χ^2 attack against versions of RC6 with word size $w = 8, 16, 32$ bits. Using the results, they estimated the necessary number of plaintexts to distinguish $(2r+1)$-round RC6-32 from random permutations. And they estimated the necessary number of plaintexts against 15-round RC6-32 is $2^{111.0}$. Their results of the computer experiment with word size $w = 8, 16, 32$ bits is shown in Table 2.

Moreover, expanding this Distinguishing Attack, they proposed the key recovery algorithm and estimated the sample complexity and the computational complexity to derive key. We note that although the necessary number of plaintexts theoretically, we omit discussion of the key recovery. (They are up to [6].)

5 Density Computation using Transition Matrix

5.1 Basic Model

In this section, we briefly sketch our idea of analysis of RC6. First, we consider the following simple model. Let R_0 be a function whose input, denoted by X, t, and output, denoted by Y, are both of the bit size 8.

$$\begin{aligned} Y &= R_0(X,t) = ((X \oplus t) \lll lsb_3(t)) \\ Z &= Add_K(Y) = Y + K \bmod 2^8, \end{aligned}$$

where $a \lll lsb_3(t)$ is τ times rotation of a and τ is the number corresponding to the last significant three bits of t. (See Figure 1.)

Table 3. Output $Y = R_0(X, t)$ and the distribution on $lsb_3(Y)$ (in the case $lsb_3(X) = 0$)

t	Y		$lsb_3(Y)$ 000	001	010	011	100	101	110	111
*****000	*****	000	1/8	0	0	0	0	0	0	0
*****001	****0	01*	0	0	1/16	1/16	0	0	0	0
*****010	***01	0**	1/32	1/32	1/32	1/32	0	0	0	0
*****011	**011	***	1/64	1/64	1/64	1/64	1/64	1/64	1/64	1/64
*****100	*100*	***	1/64	1/64	1/64	1/64	1/64	1/64	1/64	1/64
*****101	101**	***	1/64	1/64	1/64	1/64	1/64	1/64	1/64	1/64
*****110	10***	**1	0	1/32	0	1/32	0	1/32	0	1/32
*****111	1****	*11	0	0	0	1/16	0	0	0	1/16
Prob.			13/64	7/64	9/64	15/64	3/64	5/64	3/64	9/64

Matrix Representation of the Input-Output Transition of R_0 For the simplicity, we consider the case $lsb_3(X) = 0$. Since $lsb_3(X \oplus t) = lsb_3(t)$, we have an input-output transition of R_0 (See Table 3). In [10], Vaudenay mentioned the similar transition matrix for consideration of the generalized linear test.

The symbols "*" corresponds with the most significant five bits of $X \oplus t$. If we assume that t are randomly and independently given from X, the most significant five bits of $X \oplus t$ is uniformly random for any X. Under the assumption, we can treat the symbol "*" as either 0 or 1 with the equal probability 1/2. Thus, we can calculate the distribution on Y that depends on the value $lsb_3(t)$. And, we can say that $Y = R(X, t)$ is biased if and only if $lsb_3(X)$ is biased. Moreover, since we take into account that the output Y from R_0 flows to the next R_0, we may consider the least significant three bits of Y.

In case that $lsb_3(t) = 0 = (0, 0, 0)$, which occurs with probability 1/8, $lsb_3(Y)$ is always equal to $0 = (0, 0, 0)$. In case that $lsb_3(t) = 1 = (0, 0, 1)$, which occurs with probability 1/8, $lsb_3(Y)$ is distributed on $2 = (010)$ and $3 = (0, 1, 1)$ with each probability 1/2.

It is easy to compute the distribution on $lsb_3(Y)$ in any other cases. (See Table 3).

Now, by the Table 3, we can easily see that $lsb_3(Y)$ is biased. Similarly, we consider the other cases than $lsb_3(X) = 0$ and have the transition probability matrix between the $lsb_3(X)$ of input to R_0 and $lsb_3(Y)$ of output from R_0 in Table 4. We note that the distribution on $lsb_3(Y)$, which is seen in Table 4, can be easily calculated using the value of $lsb_3(X)$ and the independent uniform randomness of t, and the most significant

Table 4. Transition probability matrix between the lsb_3 of input to R_0 and lsb_3 of output from R_0 ($\times 1/64$)

	$lsb_3(X)$							
$lsb_3(Y)$	000	001	010	011	100	101	110	111
000	13	7	5	3	7	9	11	9
001	7	17	7	5	9	11	5	3
010	9	3	17	7	11	5	7	5
011	15	9	7	13	9	3	5	3
100	3	5	3	9	13	7	9	15
101	5	7	5	11	7	17	3	9
110	3	5	11	9	5	7	17	7
111	9	11	9	7	3	5	7	13

five bits of X does not affect the distribution of $lsb_3(Y)$.

Let M_{R_0} be the matrix given in Table 4 and $p_{R0}(x)$ be the probability that the least significant three bits $lsb_3(X)$ of X which is input of R_0 is equal to x. We denote, by ϕ_x, the column vector consisting of $p_{R0}(0), ..., p_{R0}(7)$. (For example, if the least significant three bits of X is always equal to 0 then $\phi_x = {}^t(1, 0, ...0)$.) Similarly, we denote, by ϕ_y, the column vector of the occurrence probabilities of 0,...,7 in $lsb_3(Y)$. Then, we have the following equation.

$$\phi_y = M_{R_0}\phi_x$$

Matrix Representation of Key Addition Next, we consider the matrix representation of key addition $Z = Add_K(Y) = Y + K \bmod 2^8$ when restricted on the least significant three bits in the simple model. (See Figure 1.) Since the least significant three bits of output Z from Add_K is just an additive result of the least significant three bits of Y and key K. For example, if the key satisfies that $lsb_3(K) = 1$ then the probability that $lsb_3(Z) = 1$ is the one that $lsb_3(Y) = 0$. Let ϕ_z be the column vector of the occurrence probability of 0,...,7 in $z = lsb_3(Z)$. Then it is easy to see that ϕ_z is the vector whose elements coincide with a rotation of elements in ϕ_y. Thus, the key addition in case that the least significant three bits of the key K is equal to 1 is represented by the following matrix

T_0 and the equation $\phi_z = T_0 \phi_y$ holds.

$$T_0 = \begin{pmatrix} 0&0&0&0&0&0&0&1 \\ 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \end{pmatrix}$$

Moreover, if the least significant three bits of the key K is equal to k then the transition by the key addition is represented as follows.

$$\phi_z = {T_0}^k \phi_y$$

5.2 Matrix Computation of Input-Output Transition of RC6

In this subsection, we expand the Basic Model, which has been considered in the previous subsection, into a generalized model in which we can handle the functions in RC6. For simplicity, we consider RC6-8, which is a variant of RC6 of the block size 32-bit (and of the word size 8-bit). We note that the following discussion is applicable to RC6-32, which is of block size 128-bit. Let (A, C, t, u) be an input to RC6-8 and (K_A, K_C) be a key of RC6-8, where A, C, t, u, K_A, K_C are of the bit length 8. Then, a function R in RC6-8 and the key addition $Add_{(K_A,K_C)}$ in RC6-8 are defined as follows.

$$\begin{aligned} (A', C') &= R(A, C, t, u) = ((A \oplus t) \lll lsb_3(u), (C \oplus u) \lll lsb_3(t)) \\ (A'', C'') &= Add_{(K_A,K_C)}(A', C') = (A' + K_A \bmod 2^8, C' + K_C \bmod 2^8) \end{aligned}$$

Data Dependent Rotation in RC6 As in the Basic Model, we assume that t, u and A, C are distributed uniformly and independently and given to R. Let a (resp., c) be the least significant three bits $lsb_3(A)$ (resp., $lsb_3(C)$) of A (resp., C). We consider the case of $a = c = 0$. Then, we have $lsb_3(A \oplus t) = lsb_3(t)$, $lsb_3(C \oplus u) = lsb_3(u)$. Let $(a'|c') = (lsb_3(A')|lsb_3(C'))$, where $(A', C') = R(A, C, t, u)$. We denote by $lsb_3(t, u)$ the concatenation $(lsb_3(t)|lsb_3(u))$ of the least significant three bits of t and the least significant three bits of u. Then, the values of $(a'|c')$ are calculated as in Table 5.

The symbols in "*" in Table 5 correspond to some bits in the most significant 27 bits of either $A \oplus t$ or $C \oplus u$. Since we assume that t, u are

Table 5. Output of R-function (in the case that each of the least significant three bits of A and C is 0)

$lsb_3(t,u)$	$(a'\|c')$	$lsb_3(t,u)$	$(a'\|c')$	$lsb_3(t,u)$	$(a'\|c')$	$lsb_3(t,u)$	$(a'\|c')$
000000	000000	010000	0100**	100000	100***	110000	110**0
000001	00*001	010001	10*1**	100001	00****	110001	10***0
000010	0**010	010010	0**0**	100010	0*****	110010	0****0
000011	***011	010011	***1**	100011	******	110011	*****0
000100	***100	010100	***0**	100100	******	110100	*****1
000101	***101	010101	***1**	100101	******	110101	*****1
000110	**0110	010110	**00**	100110	**1***	110110	**1**1
000111	*00111	010111	*011**	100111	*10***	110111	*11**1
001000	00100*	011000	011***	101000	101***	111000	111*00
001001	01*01*	011001	11****	101001	01****	111001	11**00
001010	1**10*	011010	1*****	101010	1*****	111010	1***01
001011	***11*	011011	******	101011	******	111011	****01
001100	***00*	011100	******	101100	******	111100	****10
001101	***01*	011101	******	101101	******	111101	****10
001110	**010*	011110	**0***	101110	**1***	111110	**1*11
001111	*0011*	011111	*01***	101111	*10***	111111	*11*11

uniformly random values and chosen independently from A, C, we can treat the symbol "*" as either 0 or 1 with the equal probability 1/2. As in Basic Model, we can calculate the distribution on $(a'|c')$. (See Table 6).

As in the case of $(a|c) = (0|0)$, it is not hard to calculate the distribution of $(a'|c')$ in any other cases than $(a|c) = (0|0)$. Let M_R be the transition probability 64×64 matrix from $(a|c)$ to $(a'|c')$. That is, the element which is in the ith row and in the jth column of M_R represents the probability that $(a'|c') = i$ occurs when $(a|c) = j$. As in Basic Model, ϕ denotes the column vector of occurrence probabilities of values on $(a|c)$ and ψ denotes the column vector of occurrence probabilities of values on $(a'|c')$. Then we have the following equation.

$$\psi = M_R \phi$$

Key Addition in RC6 We consider the matrix representation of the key addition $Add_{(K_A,K_C)}$ in $RC6$. For example, if the least significant three bits k_1 of the key K_A is equal to 1 and the least significant three bits k_2 of the key K_C is equal to 0, then the input-output transition matrix T_1 by the key addition is the following, where E is the 8×8 identity matrix. We note that the transition matrix T_2 in the case $k_1 = 0$ and $k_2 = 1$ is

Table 6. Distribution on $(a'|c')$ (in the case that each of the least significant three bits of A and C is 0)

(0|0)	$1616/2^{16}$	(2|0)	$976/2^{16}$	(4|0)	$688/2^{16}$	(6|0)	$1136/2^{16}$
(0|1)	$1136/2^{16}$	(2|1)	$1008/2^{16}$	(4|1)	$784/2^{16}$	(6|1)	$848/2^{16}$
(0|2)	$976/2^{16}$	(2|2)	$1616/2^{16}$	(4|2)	$816/2^{16}$	(6|2)	$1008/2^{16}$
(0|3)	$688/2^{16}$	(2|3)	$1328/2^{16}$	(4|3)	$720/2^{16}$	(6|3)	$784/2^{16}$
(0|4)	$688/2^{16}$	(2|4)	$816/2^{16}$	(4|4)	$1104/2^{16}$	(6|4)	$1424/2^{16}$
(0|5)	$720/2^{16}$	(2|5)	$848/2^{16}$	(4|5)	$1200/2^{16}$	(6|5)	$1136/2^{16}$
(0|6)	$1136/2^{16}$	(2|6)	$1008/2^{16}$	(4|6)	$1424/2^{16}$	(6|6)	$1232/2^{16}$
(0|7)	$1232/2^{16}$	(2|7)	$592/2^{16}$	(4|7)	$1456/2^{16}$	(6|7)	$624/2^{16}$
(1|0)	$1136/2^{16}$	(3|0)	$688/2^{16}$	(5|0)	$720/2^{16}$	(7|0)	$1232/2^{16}$
(1|1)	$1744/2^{16}$	(3|1)	$912/2^{16}$	(5|1)	$880/2^{16}$	(7|1)	$880/2^{16}$
(1|2)	$1008/2^{16}$	(3|2)	$1328/2^{16}$	(5|2)	$848/2^{16}$	(7|2)	$592/2^{16}$
(1|3)	$912/2^{16}$	(3|3)	$1616/2^{16}$	(5|3)	$944/2^{16}$	(7|3)	$1200/2^{16}$
(1|4)	$784/2^{16}$	(3|4)	$720/2^{16}$	(5|4)	$1200/2^{16}$	(7|4)	$1456/2^{16}$
(1|5)	$880/2^{16}$	(3|5)	$944/2^{16}$	(5|5)	$1360/2^{16}$	(7|5)	$1104/2^{16}$
(1|6)	$848/2^{16}$	(3|6)	$784/2^{16}$	(5|6)	$1136/2^{16}$	(7|6)	$624/2^{16}$
(1|7)	$880/2^{16}$	(3|7)	$1200/2^{16}$	(5|7)	$1104/2^{16}$	(7|7)	$1104/2^{16}$

represented in the following, where T_0 is the matrix considered in Basic Model.

$$T_1 = \begin{pmatrix} 0&0&0&0&0&0&0&E \\ E&0&0&0&0&0&0&0 \\ 0&E&0&0&0&0&0&0 \\ 0&0&E&0&0&0&0&0 \\ 0&0&0&E&0&0&0&0 \\ 0&0&0&0&E&0&0&0 \\ 0&0&0&0&0&E&0&0 \\ 0&0&0&0&0&0&E&0 \end{pmatrix} \qquad T_2 = \begin{pmatrix} T_0&0&0&0&0&0&0&0 \\ 0&T_0&0&0&0&0&0&0 \\ 0&0&T_0&0&0&0&0&0 \\ 0&0&0&T_0&0&0&0&0 \\ 0&0&0&0&T_0&0&0&0 \\ 0&0&0&0&0&T_0&0&0 \\ 0&0&0&0&0&0&T_0&0 \\ 0&0&0&0&0&0&0&T_0 \end{pmatrix}$$

Let ψ be the column vector of occurrence probabilities on $(a'|c')$, the least significant three bits of inputs to the key addition $Add_{(K_A,K_C)}$, and ω be the column vector of probabilities on $(a''|c'')$, the least significant three bits of outputs from the key addition. Then, we have the following equation.

$$\omega = {T_1}^{k_1} {T_2}^{k_2} \psi$$

2*r*-Round RC6 Case In discussion so far, we have assumed that t and u that are inputs to R-function are uniformly distributed. We comments on the validity of this assumption. The values on t and u in the original RC6 are calculated as follows: $t = (B(2B+1)) \lll 5$ and $u = (D(2D +$

1))⋘5. It is easy to see that these functions are one-to-one w.r.t. B and D respectively. Furthermore, each least significant three bits of t and u is a value obtained by a multiplication. In general, the most significant bits of a value obtained by a multiplication are highly disturbed. Thus, we regard our assumption as being appropriate.

Let $(A^{(r)}, C^{(r)}, B^{(r)}, D^{(r)})$ be output from $2r$-round RC6-8. For convenience, let $(A^{(0)}, C^{(0)}, B^{(0)}, D^{(0)})$ be input (i.e., plaintexts) to $2r$-round RC6-8. Let $\phi^{(0)}$ be the column vector of occurrence probabilities on $(a^{(0)}, c^{(0)})$, pair of the least significant three bits of $A^{(0)}$ and $C^{(0)}$. For example, if $(a^{(0)}, c^{(0)}) = (0, 0)$ then $\phi^{(0)} = {}^{t}(1, 0, ..., 0) \in \mathcal{R}^{64}$.

Using the discussion in this section, we have the following equation w.r.t. $\phi^{(0)}$ and $\phi^{(1)}$.

$$\phi^{(1)} = {T_1}^{lsb_3(S[2])} {T_2}^{lsb_3(S[3])} M_R \phi^{(0)}$$

Using similar discussion, we can calculate the output distribution on $2r$-round RC6 (w.r.t. any fixed extended key $S[i]$) as follows.

$$\phi^{(r)} = \prod_{i=1}^{r} (T_1^{lsb_3(S[4i+2])} T_2^{lsb_3(S[4i+3])} M_R) \phi^{(0)}$$

6 χ^2 Statistic with Transition Matrix Computation

In the previous section, we consider RC6-8 of the block size 32-bit. The similar discussion is applicable to RC6-32 of the block size 128-bit. A major difference is the size of transition matrices. The size of matrices T_1, T_2 and M_R is 1024×1024 in case of RC6-32.[1] In this section, we calculate the distribution on the least significant five bits of output from each round in RC6-32 by using the way shown in the previous section. We also exactly calculate the expected χ^2 value from the resulting distribution and compare with experimental results in Knudsen and Meier [6].

Using the fixed user key, we calculate round keys $S[i]$ that is generated from the key schedule function of RC6. Then, we calculate the value of $\phi^{(r)}$ with the below matrix equation, which is obtained in the previous section.

$$\phi^{(r)} = \prod_{i=1}^{r} (T_1^{lsb_5(S[4i+2])} T_2^{lsb_5(S[4i+3])} M_R) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

[1] The (0,0)-element of M_R is $1813/2^{20}$, the whole matrix (1024×1024) is too large to show in this paper, though.

Table 7. The average value for $\theta(\phi^{(r)})$ and the sample complexity to distinguish $2r$-round RC6 from random permutation (log_2)

round	$\theta(\phi^{(r)})$	♯ text (log_2)		
		average	deviation	weakest
3	-6.7566	12.9854	0.0	12.9854
5	-22.9979	29.2067	0.3397	27.8251
7	-39.0473	45.2761	0.4667	42.3149
9	-55.0994	61.3282	0.5577	56.3422
11	-71.1414	77.3702	0.6330	70.4079
13	-87.1729	93.4017	0.6781	84.3690
15	-103.2145	109.4433	0.7324	98.2064
17	-119.2522	125.4810	0.7865	112.0031
19	-135.2883	141.5171	0.8416	125.8010

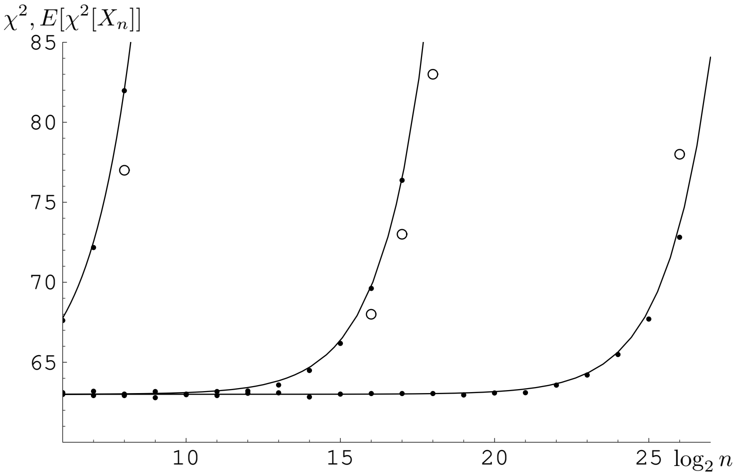
Proposition in Section 3 implies that the expected value $E[\chi^2(X_n)]$ of $\chi^2(X_n)$ can be computed from m and probability $\phi_x^{(r)}$ for each x and that the expected value is almost proportional to n. Where, $\phi_x^{(r)}$ is the x-th element of vector $\phi^{(r)}$.

$$\theta(\phi^{(r)}) = m \sum_{x=0}^{1023} (\phi_x^{(r)} - 1/1024)^2$$
$$E[\chi^2(X_n)] = \theta(\phi^{(r)}) \cdot n + 1023$$

We set $m = 1024$ and $\#\Omega = 2^{118}$ and compute $\theta(\phi^{(r)})$ from randomly chosen 1000 user keys. We also compute the averages and the derivations from the experimental data. By using these averages, we calculate the value of n satisfying that $E[\chi^2(X_n)] = \theta(\phi^{(r)}) \cdot n + 1023 = 1098$, that means the average sample complexity which the χ^2 value exceeds 1098. (See Table 7.) The logarithm of the sample complexity to distinguish $2r$-round RC6-32 from random permutation is almost linear in r. Using the least square method, we obtain that the sample complexity to distinguish $(2r+3)$-round RC6-32 from random permutation is $2^{16.0198r+13.1094}$.

We note that the sample complexity $2^{16.2r+13.8}$ estimated by Knudsen and Meier is quite close to the theoretical value though their value is drawn from the 20 trials for 2- and 4-round RC6-32.

In [6], Knudsen and Meier estimated the sample complexity in case of 2-round RC6-32 at 2^{13}. They also experimented with 4-round RC6-



Theoretical χ^2 values obtained from the matrix computation.(-)
Experimental χ^2 values from our 1000 tests with RC6-8.(•)
Knudsen and Meier's χ^2 value with 20 experiments.(∘)

Fig. 2. Theoretical χ^2 values, our experimental χ^2 values, and Knudsen-Meier's experimental χ^2 values on RC6-8

32 and claimed that $2^{16.2}$ times sample complexity should be required to distinguish two more rounds of RC6-32 from random permutation. From the estimation they had, they calculated the sample complexity to distinguish $2r$-round RC6-32 from random permutation by the linear interpolation.

In Figure 2, we illustrate the theoretical χ^2 values, our experimental χ^2 values, and Knudsen-Meier's experimental χ^2 values on RC6-8. (w.r.t. RC6-8 as a simple case in which we explain our theory.) We note that these experimental results endorse our theory.

7 Distinguishing Attack and Weakest Key

In this section, we consider the distribution of sample complexity with respect to user key. We call keys which are whose sample complexity is relatively large (resp., small) "strong keys" (resp., "weak keys"). Then, we can regard the strength of keys as the sample complexity ($\log_2(\sharp text)$). Table 7 shows the average and the standard deviation of the sample complexity to distinguish RC6-32 from random permutation, where $\sharp text = \min\{n|E[\chi^2(X_n)] \geq 1098\}$. They are calculated from 1000 instances of

$E[\chi^2(X_n)]$ values randomly chosen from user keys. We note that $E[\chi^2(X_n)]$ depends on implicitly assumed user key. In this section, we consider the distribution of weak keys by using the results in Table 7.

Here we assume that the distribution of $\log_2(\sharp text)$ for random keys can be approximated to the normal distribution. (Our 1,000 computer experiments endorse this assumption.) This implies the weak keys are symmetrically distributed as strong key. The ratio of (weak) keys where 15-round RC6-32 is distinguishable from the random permutation with the same complexity $2^{109.4433}$ is 50 %. Moreover, with the sample complexity $2^{110.1757}(= 2^{109.4433+0.7324})$ (resp., $2^{110.9081}(= 2^{109.4433+2\cdot 0.7324})$) the ratio of weak keys augment to be 84.1 % (resp., (97.77 %)). In the same way, we consider 17 round RC6-32. From Table 7 and the normal density function, we have $(125.4810 - 118)/0.7865 = 9.5113$ and $\mathit{erfc}(9.5113) = 1 - 2^{-69.8474}$. Then, we can show that there exist weak keys in 17-round RC6-32 whose fraction is one over $2^{69.8474}$ which can be distinguished by using less sample complexity than 2^{118}. Therefore, it is said that the weak key ratio is about 1024 times larger than the weak key ratio (one in 2^{80} keys) mentioned by Knudsen and Meier [6]. On the other hand, for 19-round RC6-32, the ratio of such weak keys is one in $2^{569.3812}$, that is much fewer than whole size of user keys. So such weak keys do not exist.

Now we study about the weak keys mentioned by Knudsen and Meier such that the least significant five bits of extended keys are zero for every 2 rounds. In our method, the value of $E[\chi^2(X_n)]$ depends only on each of the input value $lsb_5(A, C)$ and $lsb_5(S[4i+2], S[4i+3])$ for i from 0 to $\lfloor r/4 \rfloor$. Therefore, it is possible to evaluate the security of RC6-32 against the χ^2 attack for any keys. Our results of the "alert" security evaluation for almost all keys imply that the weak keys mentioned by Knudsen and Meier are actual "weakest keys". The column of "weakest" in Table 7 shows its security level for each round. Table 7 shows that the distinguishing attack can not be applicable to RC6-32 more than 18 rounds even in the case of the weakest key. Therefore, we are able to prove that the 20-round RC6-32 is secure against χ^2 attack if it is shown the 3-round elimination attack (including the key-recovery algorithm) can not be applicable.

For the end of this section, we comment the randomness of extended keys of RC6-32. In this paper, we use the number 2^{-10} for the ratio of the user keys such that the related extended keys $S[4i+2], S[4i+3]$ satisfy the equation $lsb_5(S[4i+2], S[4i+3]) = 0$ holds. We avoid the difficulty of theoretical analysis about the distribution of extended keys, since the key schedule part of RC6-32 is very complex. Instead, we adopt computer-experimental ratio. Our results show that the user keys, for which $lsb_5(S[2i], S[2i+1]) = 0$ hold, exist one in $2^{9.9994}$ on average, and its variance is $2^{-34.296}$. These results are obtained from the experiment with sampling on the 2^{30} user keys of RC6-32 using 128-bit key such that $(key[0], key[1], key[2], key[3]) = (0,0,0,0), \cdots, (0,0,0,0,2^{30}-1)$. From this, we can say that each of the least significant five bits are uniformly distributed. This implies validity of our assumption.

8 Conclusion

In this paper we have given the theoretical analysis of χ^2 attack by Knudsen and Meier on the RC6 block cipher. For this, we have proposed the novel method of security evaluation against the χ^2 attack precisely including key dependency by introducing the technique "Transition Matrix Computing." On the other hand, the way of security evaluation against χ^2 attack had not been known except the computer experiment. We have shown the way of security evaluation theoretically by this paper. For 17-round RC6, it has be shown that the user keys which has been able to be distinguished from a random permutation with less than the sample complexity 2^{118} exist with probability $1/2^{69.8474}$. Moreover, using this method, we have found the "weakest key" against χ^2 attack on RC6.

References

1. J. Borst, B. Preneel, and J. Vandewalle. Linear cryptanalysis of RC5 and RC6. FSE'99, LNCS 1636, pp.16–30, 1999.
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 `http://www.rsasecurity.com/rsalabs/rc6/`.
3. S. Contini, R.L. Rivest, M.J.B. Robshaw, and Y.L. Yin. Improved analysis of some simplified variants of RC6. FSE'99, LNCS 1636, pp.1–15, 1999.
4. H. Gilbert, H. Handschuh, A. Joux and S. Vaudenay, A Statistical Attack on RC6. FSE 2000, LNCS 1978, pp.64–74, 2001.
5. J. Kelsey, B. Schneier, and D. Wagner. Mod n cryptanalysis, with applications against RC5P and M6. FSE'99, LNCS 1363, pp.139–155, 1999.
6. L.R. Knudsen and W. Meier. Correlations in RC6 with a reduced number of rounds. FSE 2000, LNCS 1978, pp.94–108, 2001.
7. R.L. Rivest. The RC5 encryption algorithm. FSE'94, LNCS 1008, pp.86–96, 1995.
8. R.L. Rivest, M.J.B. Robshaw, R. Sidney and Y.L. Yin. The RC6 block cipher. v1.1, August 20, 1998. Available at `http://www.rsasecurity.com/rsalabs/rc6/`.
9. T. Shimoyama, M. Takenaka and T. Koshiba. Multiple linear cryptanalysis of a reduced round RC6. FSE 2002.
10. S. Vaudenay. An Experiment on DES Statistical Cryptanalysis. *3rd ACM Conference on Computer and Communications Security*, ACM Press, pp. 139-147, 1996.

Appendix A: Figure for Feistes-like description of RC6

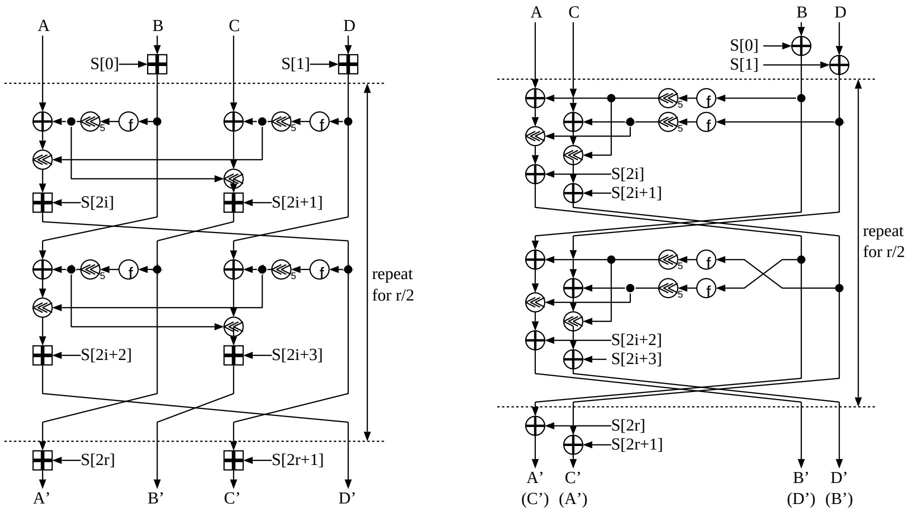


Fig. 3. RC6 structure and its transformation

Appendix B: Proof of Proposition 1

$$
\begin{aligned}
Y &= \frac{m}{n} \sum_{a=1}^{m} \left(N_a(X) - \frac{n}{m} \right)^2 \\
E[Y] &= \frac{m}{n} \sum_{a=1}^{m} \left(E[N_a(X)^2] - \frac{2n}{m} E[N_a(X)] + \frac{n^2}{m^2} \right) \\
&= \frac{m}{n} \sum_{a=1}^{m} \left(\frac{\sum_{X \subset \Omega} N_a(X)^2}{\binom{\sharp\Omega}{n}} - \frac{2n}{m} \frac{\sum_{X \subset \Omega} N_a(X)}{\binom{\sharp\Omega}{n}} + \frac{n^2}{m^2} \right) \\
&= \frac{m}{n} \sum_{a=1}^{m} \left(\frac{\sum_{i=1}^{n} \binom{n}{i} \binom{\sharp\Omega - P_a}{n-i} i^2}{\binom{\sharp\Omega}{n}} - \frac{2n}{m} \frac{\sum_{i=1}^{n} \binom{n}{i} \binom{\sharp\Omega - P_a}{n-i} i}{\binom{\sharp\Omega}{n}} + \frac{n^2}{m^2} \right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{m}{n}\sum_{a=1}^{m}\left(\frac{\sum_{i=1}^{n}\binom{n}{i}\binom{\sharp\Omega-n}{P_a-i}i^2}{\binom{\sharp\Omega}{P_a}} - \frac{2n}{m}\frac{\sum_{i=1}^{n}\binom{n}{i}\binom{\sharp\Omega-n}{P_a-i}i}{\binom{\sharp\Omega}{P_a}} + \frac{n^2}{m^2}\right)\\
&= \frac{m}{n}\sum_{a=1}^{m}\left(\frac{n\sum_{i=1}^{n}\binom{n-1}{i-1}\binom{\sharp\Omega-n}{P_a-i}i}{\binom{\sharp\Omega}{P_a}} - \frac{2n^2}{m}\frac{\sum_{i=1}^{n}\binom{n-1}{i-1}\binom{\sharp\Omega-n}{P_a-i}}{\binom{\sharp\Omega}{P_a}} + \frac{n^2}{m^2}\right)\\
&= \frac{m}{n}\sum_{a=1}^{m}\left(\frac{n\sum_{j=1}^{n-1}\binom{n-1}{j}\binom{\sharp\Omega-n}{P_a-j-1}(j+1)}{\binom{\sharp\Omega}{P_a}} - \frac{2n^2}{m}\frac{\binom{\sharp\Omega-1}{P_a-1}}{\binom{\sharp\Omega}{P_a}} + \frac{n^2}{m^2}\right) \quad \dagger\\
&= \frac{m}{n}\sum_{a=1}^{m}\left(\frac{n(n-1)\sum_{j=1}^{n-1}\binom{n-2}{j-1}\binom{\sharp\Omega-n}{P_a-j-1}}{\binom{\sharp\Omega}{P_a}} + n\frac{\binom{\sharp\Omega-1}{P_a-1}}{\binom{\sharp\Omega}{P_a}} - \frac{2n^2}{m}\frac{\binom{\sharp\Omega-1}{P_a-1}}{\binom{\sharp\Omega}{P_a}} + \frac{n^2}{m^2}\right)\\
&= \frac{m}{n}\sum_{a=1}^{m}\left(\frac{n(n-1)\binom{\sharp\Omega-2}{P_a-2}}{\binom{\sharp\Omega}{P_a}} + \left(n-\frac{2n^2}{m}\right)\frac{\binom{\sharp\Omega-1}{P_a-1}}{\binom{\sharp\Omega}{P_a}} + \frac{n^2}{m^2}\right)\\
&= \frac{m}{n}\sum_{a=1}^{m}\left(\frac{n(n-1)(\sharp\Omega-2)!P_a!(\sharp\Omega-P_a)!}{\sharp\Omega!(P_a-2)!(\sharp\Omega-P_a)!} + \left(n-\frac{2n^2}{m}\right)\frac{P_a!(\sharp\Omega-1)!(\sharp\Omega-P_a)!}{\sharp\Omega!(\sharp\Omega-P_a)!(P_a-1)!} + \frac{n^2}{m^2}\right)\\
&= \frac{m}{n}\sum_{a=1}^{m}\left(n(n-1)\frac{P_a(P_a-1)}{\sharp\Omega(\sharp\Omega-1)} + \left(n-\frac{2n^2}{m}\right)\frac{P_a}{\sharp\Omega} + \frac{n^2}{m^2}\right)\\
&= \frac{m}{n}\sum_{a=1}^{m}\left(n(n-1)p(a)^2 + \left(n-\frac{2n^2}{m}\right)p(a) + \frac{n^2}{m^2}\right) + n(n-1)\frac{m}{n}\sum_{a=1}^{m}\left(\frac{P_a(P_a-1)}{\sharp\Omega(\sharp\Omega-1)} - \frac{P_a^2}{\sharp\Omega^2}\right) \quad \ddagger\\
&= mn\sum_{a=1}^{m}\left(\frac{n-1}{n}p(a)^2 + \left(\frac{1}{n}-\frac{2}{m}\right)p(a) + \frac{1}{m^2}\right) + \frac{(n-1)}{(\sharp\Omega-1)}\left(m\sum_{a=1}^{m}p(a)^2 - m\right)\\
&= mn\sum_{a=1}^{m}\left(p(a)-\frac{1}{m}\right)^2 + mn\sum_{a=1}^{m}\frac{p(a)}{n} - mn\sum_{a=1}^{m}\frac{p(a)^2}{n} + \frac{(n-1)}{(\sharp\Omega-1)}\left(m\sum_{a=1}^{m}p(a)^2 - m\right)\\
&= mn\sum_{a=1}^{m}\left(p(a)-\frac{1}{m}\right)^2 + \frac{(\sharp\Omega-n)}{(\sharp\Omega-1)}\left(m - m\sum_{a=1}^{m}p(a)^2\right)\\
&\simeq mn\sum_{a=1}^{m}\left(p(a)-\frac{1}{m}\right)^2 + m - 1
\end{aligned}
$$

$$\dagger : \sum_k \binom{r}{n} \binom{s}{n-k} = \binom{r+s}{n}$$

$$\ddagger : P_a = p(a) \cdot \sharp\Omega$$