Success probability in χ^2-attacks

Takashi Matsunaka*, Atsuko Miyaji**, and Yuuki Takano

Japan Advanced Institute of Science and Technology.
{t-matuna, miyaji, ytakano}@jaist.ac.jp

Abstract. Knudsen and Meier applied the χ^2-attack to RC6. This attack is one of the most effective attacks for RC6. The χ^2-attack can be used for both distinguishing attacks and for key recovery attacks. Up to the present, theoretical analysis of χ^2-attacks, especially the relation between a distinguishing attack and a key recovery attack, has not been discussed. In this paper, we investigate the theoretical relation between the distinguishing attack and the key recovery attack for the first time, and prove the theorem to evaluate the success probability of a key recovery attack by using the results of a distinguishing attack. We also demonstrate the accuracy to χ^2-attack on RC5-64 and RC6 without post-whitening by comparing the implemented results.

Keywords RC6, RC5-64, χ^2 attack, statistical analysis

1 Introduction

The χ^2-attack makes use of correlations between input (plaintext) and output (ciphertext), which is measured by the χ^2-test. The χ^2-attack was originally proposed by Vaudenay as an attack on the Data Encryption Standard (DES) [20], and Handschuh *et al*. applied that to SEAL [6]. The χ^2-attack can be used for both distinguishing attacks and key recovery attacks. Distinguishing attacks have only to handle plaintexts in such a way that the χ^2-value of a part of ciphertexts becomes significantly a high value. On the other hand, key recovery attacks have to rule out all wrong keys, and single out exactly a correct key by using the χ^2-value. Therefore, key recovery attacks often require more work and memory than distinguishing attacks.

In [4, 12], the χ^2-attacks were applied to RC6 [18] or a simplified variant of RC6. They focused on the fact that a specific rotation in RC6 causes the correlations between input and output, and estimated their key recovery attack by using only results of a distinguishing attack [4, 12, 16]. Note that their key recovery attack on RC6 with any round was not implemented because it required too much memory even in the case of small number of rounds. In [5], a key recovery attack on RC5-32 [17] by using the χ^2-attack was proposed. RC5-$w/r/b$ means that two w-bit-word plaintexts are encrypted with r rounds by b-byte keys.

* The author is currently with KDDI.

** Supported by Inamori Foundation.

The χ^2-attack to RC5-32 was further improved by [15]. Their attack can analyze RC5-32 with 10 rounds by a known plaintext attack with negligible memory. They also pointed out the significant difference between the distinguishing attack and the key recovery attack: The distinguishing attack succeeds if and only if it outputs high χ^2-value, but the key recovery attack does not necessarily succeed even if it outputs high χ^2-value. In fact, a key recovery attack to RC5-32 in [5] outputs higher χ^2-value but recovers a correct key with lower probability than that in [15]. This indicates that the security against the key recovery attack cannot be estimated directly from that against the distinguishing attack. The χ^2-attack to a simplified variant of RC6 are further improved in [16, 7], which can work on 4-round simplified variants of RC6.

However, up to the present, any theoretical difference between a distinguishing attack and a key recovery attack in χ^2-attack has not been discussed. Although the theoretical and experimental complexity analysis on the linear cryptanalysis is done by P. Junod in [8], it cannot be applied to the χ^2-attack. His analysis is further generalized by using the normal approximation for order statistics in [19]. However, it is not so sharp or suitable for χ^2-attack.

In this paper, we investigate the theoretical relation between a distinguishing attack and a key recovery attack in χ^2-attack, for the first time, and give the theorem that evaluates the success probability of a key recovery attack by using results of a distinguishing attack. We demonstrate the theorem on a key recovery algorithm against RC5-64, which is given by us, and make sure the accuracy by comparing our approximation to implemented results. We also demonstrate the accuracy to the χ^2-attack against RC6 without post-whitening [7]. As a result, we are able, with our theory, to evaluate the security of key recovery attack in χ^2-attack with less number of plaintexts than expected. We also compare our theory with [19] by applying them on RC5-64 and RC6P, and show our theory is more accurate and more suitable for approximation of χ^2-attack.

This paper is organized as follows. Section 2 summarizes the notation, RC5-64 and RC6 algorithms, the χ^2-test, and statistical facts used in this paper. Section 3 gives the theory of success probability in χ^2-attack and investigates the accuracy by comparing the approximations of success probability to 3-round and 4-round RC5-64 and implemented results. Section 4 applies our theorem to a key recovery algorithm on RC6 without post-whitening. The accuracy of our approximation theorem is compared with that of [19] in Section 5. Conclusion is given in Section 6.

2 Preliminaries

We summarize RC5-64 and RC6 algorithms, the χ^2-test, and statistical facts used in this paper.

2.1 Block cipher RC5-64

Before showing the encryption algorithm of RC5-64, we give some notation.

$\oplus$: bit-wise exclusive OR;
r: number of rounds ;
$a \lll b$: cyclic rotation of a to the left by b-bit;
$a \ggg b$: cyclic rotation of a to the right by b-bit;
(L_i, R_i): input of the i-th round, (L_0, R_0) and (L_{r+1}, R_{r+1}) are a plaintext and a ciphertext after r-round encryption, respectively;
S_i : i-th subkey (S_{2i} and S_{2i+1} are subkeys of the i-th round);
$\text{lsb}_n(X)$: least significant n-bit of X;
$X[i]$: i-th bit of X.

The encryption algorithm of RC5-64 is reviewed as follows: a plaintext (L_0, R_0) is added with (S_0, S_1) and set to (L_1, R_1); and (L_1, R_1) is encrypted to (L_{r+1}, R_{r+1}) by r iterations of a main loop. The detailed algorithm is given:

Algorithm 1 (RC5-64 Encryption Algorithm)
1. $L_1 = L_0 + S_0$; $R_1 = R_0 + S_1$;
2. for $i = 1$ to r do: $L_{i+1} = ((L_i \oplus R_i) \lll R_i) + S_{2i})$.
$R_{i+1} = ((R_i \oplus L_{i+1}) \lll L_{i+1}) + S_{2i+1})$.

Two rotations by R_i or L_{i+1} in i-th round are called by first rotation or second rotation, respectively.

2.2 Block cipher RC6

In addition to notation used in RC5-64, we use the following notation.

(A_i, B_i, C_i, D_i) : input of the i-th round;
(A_0, B_0, C_0, D_0) : plaintext;
$\text{msb}_n(X)$: most significant n-bit of X;
$f(x)$: $x \times (2x + 1)$;
$F(x)$: $f(x) \pmod{2^{32}} \lll 5$;
$x||y$: concatenated value of x and y.

The detailed algorithm of RC6 is given:

Algorithm 2 (RC6 Encryption Algorithm)
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: $t = F(B_i)$; $u = F(D_i)$; $A_{i+1} = B_i$;
$B_{i+1} = ((C_i \oplus u) \lll t) + S_{2i+1}$; $C_{i+1} = D_i$; $D_{i+1} = ((A_i \oplus t) \lll u) + S_{2i}$;
3. $A_{r+2} = A_{r+1} + S_{2r+2}$; $B_{r+2} = B_{r+1}$; $C_{r+2} = C_{r+1} + S_{2r+3}$; $D_{r+2} = D_{r+1}$.

Parts 1 and 3 of Algorithm 2 are called pre-whitening and post-whitening, respectively. We call the version of RC6 without post-whitening to, simply, RC6P.

2.3 χ^2-Test

We make use of the χ^2-tests to distinguish a non-uniformly random distribution from uniformly random distribution [10, 12, 13]. Let $X = X_0, ..., X_{n-1}$ be sets of $\{a_0, ..., a_{m-1}\}$, and $N_{a_j}(X)$ be the number of X which takes on the value a_j.

The χ^2-statistic of X which estimates the difference between X and the uniform distribution is defined as follows:

$$\chi^2(X) = \frac{m}{n} \sum_{i=0}^{m-1} \left(N_{a_i}(X) - \frac{n}{m} \right)^2 .$$

Table 1 presents each threshold for 63 degrees of freedom. For example, (level, χ^2) = (0.95, 82.53) in Table 1 means that the value of the χ^2-statistic exceeds 82.53 in the probability of 5%, if the observation X is uniform.

Table 1. χ^2-distribution with 63 degrees of freedom

Level	0.50	0.60	0.70	0.80	0.90	0.95	0.99
χ^2	62.33	65.20	68.37	72.20	77.75	82.53	92.01

2.4 Statistical facts

Let us describe statistical facts together with the notation.

Theorem 1 (Distribution of the Means [3]). *Let μ and σ^2 be the mean and the variance of a population, respectively. Then the mean and the variance of the distribution of the mean of a random sample with the size n drawn from the population are μ and σ^2/n, respectively.*

Theorem 2 (Central Limit Theorem [3]). *Choose a random sample from a population. If the sample size n is large, then the sampling distribution of the mean is closely approximated by the normal distribution, regardless of the population.*

Theorem 3 (Law of large numbers [3]). *The larger the sample size, the more probable it is that the sample mean comes arbitrarily close to the population mean.*

The probability density function of the normal distribution with the mean μ and the variance σ^2 is given by the following equation,

$$\phi_{(\mu,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right].$$

We also follow commonly used notation: the probability density and the cumulative distribution functions of the standard normal distribution are denoted by $\phi(x)$ and $\Phi(x)$; the probability of distribution X in the range $X \leq I$ is denoted by $\Pr(X \leq I)$; and $\mathcal{N}$ is used for the normal distributions.

3 Theoretical analysis on χ^2-attacks

This section presents the theorem of success probability in χ^2-attack, where we use a key recovery algorithm to RC5-64 based on [15].

3.1 Key recovery algorithm of RC5-64

The following algorithm recovers the least significant five bits of S_{2r+1}. Let us set $(x, y) = (\mathrm{lsb}_6(L_{r+1}), \mathrm{lsb}_6(R_{r+1}))$, $s = \mathrm{lsb}_5(S_{2r+1})(s = 0, 1, \cdots, 2^5 - 1)$, and $S_{2r+1}[6] = 0$, where x corresponds to the rotation amount in the r-th round.

Algorithm 3

1. Choose a plaintext (L_0, R_0) with $\mathrm{lsb}_6(R_0)$=0, and encrypt it.
2. For each s, decrypt a 6-bit y with a key $S_{2r+1}[6]||s$ by 1 round to a 6-bit z.
3. For each value s, x, and z, update each array by incrementing count$[s][x][z]$.
4. For each s and x, compute $\chi^2[s][x]$.
5. Compute the average $ave[s]$ of $\{\chi^2[s][x]\}_x$ for each s and output s with the highest $ave[s]$ as $\mathrm{lsb}_5(S_{r+1})$.

Figure 1 shows the outline of Algorithm 3. Algorithm 3 averages the χ^2-values

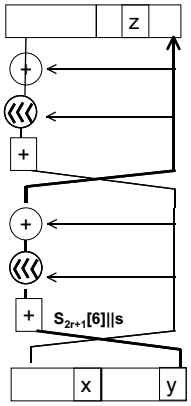

Fig. 1. Algorithm 3

$\chi^2[s][x]$ by second rotation amount x in the r-th round, in which there are 2^6 rotations.

3.2 Statistical analysis of χ^2-attacks

We show the theorem on the success probability of Algorithm 3 by investigating the distribution of χ^2-values for a correct key and wrong keys.

Notation Let us use the following notation.

- e : recovered-key bit size (There are one correct key and $2^e - 1$ wrong keys.);
- P_S : success probability of a key recovery attack;
- $X_{d[r,n]}$: distributions of χ^2-values on $\mathrm{lsb}_6(R_{r+1})$ of RC5-64 with $\mathrm{lsb}_6(R_0) = 0$ by using 2^n plaintexts;
- $\mu_{d[r,n]}$ ($\sigma^2_{d[r,n]}$): mean (variance) of distribution of χ^2-values on $\mathrm{lsb}_6(R_{r+1})$ of RC5-64 with $\mathrm{lsb}_6(R_0) = 0$ by using 2^n plaintexts;
- $X_{c[r,n]}$ ($X_{w[r,n]}$) : distributions of χ^2-values of a key recovery attack to r-round RC5-64 by using a correct key (a wrong key);
- $\mu_{c[r,n]}$ ($\sigma^2_{c[r,n]}$) : mean (variance) of distribution of mean of χ^2-values of a key recovery attack in r-round RC5-64 with a correct key by using 2^n plaintexts;
- $f_{c[r,n]}(x)$: probability density function of distribution of χ^2-values with a correct key in r-round RC5-64;
- μ_w (σ^2_w) : mean (variance) of distribution of χ^2-values in a key recovery attack r-round RC5-64 with a wrong key;
- $f_{w[r,n]}(x)$: probability density function of distribution of χ^2-values with a wrong key in r-round RC5-64.

Distributions of χ^2-values In this section, we put forward three hypotheses on distribution of χ^2-values.

Hypothesis 1 *If the number of plaintexts to compute the χ^2-values is enough large, then the sample of χ^2-values on each key candidate approximately follows a normal distribution.*

Hypothesis 2 (Wrong-Key Randomization Hypothesis 1) *Each distribution of χ^2-values of key-recovery attack on i-th wrong key $X_{w(1)[r,n]}$, $X_{w(2)[r,n]}$, $\cdots$, $X_{w(2^e-1)[r,n]}$ is independent and approximately equal to each other.*

Hypothesis 3 (Wrong-Key Randomization Hypothesis 2) *Distribution of χ^2-values of key-recovery attack on a wrong key $X_{w[r,n]}$ is approximately equal to that of $X_{d[r,n']}$, where n' is the real number of plaintexts that is used for computing χ^2-value of $X_{w[r,n]}$.*

Hypotheses 1 and 2 are simple and natural, which are often used in a statistical analysis of the security, including the differential and linear attack as in [8, 19]. On the other hand, Hypothesis 3 means that the distribution of χ^2-values recovered by using a wrong key is approximately equal to that before recovering. This is considered as a variant of Hypothesis 2, which means that a wrong key randomizes data. We note here that Hypothesis 3 is the ideal case for an attacker, and, thus, the results can be seen as an upper bound for the actual success probability. It also reflects experimental results in [7].

Success probability of χ^2-attacks We show the theorem on the success probability of Algorithm 3 by investigating the distribution of χ^2-values for a correct key and wrong keys. We may note that χ^2-attacks compute the χ^2-value on a part for every key candidate and output a key with the highest χ^2-value as a correct key.

Lemma 1. *Let $n \geq 6$ and $r \geq 4$. The distribution of χ^2-values on a correct key in Algorithm 3, $X_{c[r,n]}$, follows a normal distribution of $\mathcal{N}(\mu_{d[r-1,n-6]}, \sigma^2_{d[r-1,n-6]}/2^6)$. Therefore, the probability density function of distribution of χ^2-values on a correct key in Algorithm 3, $f_{c[r,n]}(x)$, is given by*

$$f_{c[r,n]}(x) = \phi_{(\mu_{d[r-1,n-6]}, \sigma^2_{d[r-1,n-6]}/2^6)}(x).$$

Proof. The distribution $X_{c[r,n]}$ follows a normal distribution from Hypothesis 1. When a correct key is used in Algorithms 3, six-bit data $\mathrm{lsb}_6(R_{r+1})$ is decrypted correctly by 1 round. χ^2-values are computed for every second rotation in r-th round, where each rotation amount is uniformly distributed on 2^n plaintexts. As a result, the χ^2-values in Algorithm 3 is computed by using roughly 2^{n-6} plaintexts. Putting together the facts and Theorem 1, the distribution $X_{c[r,n]}$ follows a normal distribution $\mathcal{N}(\mu_{d[r-1,n-6}, \sigma^2_{d[r-1,n-6]}/2^6)$. Thus, we get

$$f_{c[r,n]}(x) = \phi_{(\mu_{d[r-1,n-6]}, \sigma^2_{d[r-1,n-6]}/2^6)}(x).$$

Lemma 2. *Let $n \geq 6$ and $r \geq 4$. The distribution of χ^2-values on a wrong key in Algorithm 3, $X_{w[r,n]}$, follows a normal distribution of $\mathcal{N}(\mu_{d[r,n-6]}, \sigma^2_{d[r,n-6]}/2^6)$. Therefore, the probability density function of distribution of χ^2-values on a wrong key in Algorithm 3, $f_{w[r,n]}(x)$, is given by*

$$f_{w[r,n]}(x) = \phi_{(\mu_{d[r,n-6}, \sigma^2_{d[r,n-6]}/2^6)}(x).$$

Proof. The distribution $X_{w[r,n]}$ follows a normal distribution $\mathcal{N}(\mu_{d[r,n']}, \sigma^2_{d[r,n']})$ from Hypotheses 1 and 3. Here, n' is the real number of plaintexts that is used for computing χ^2-value of $X_{w[r,n]}$. In the same discussion as Lemma 1, χ^2-values are computed for every second rotation amount in r-th round, which is uniformly distributed on 2^n plaintexts. As a result, the χ^2-values in Algorithm 3 is computed by using roughly 2^{n-6} plaintexts. Putting together the facts and Theorem 1, the distribution $X_{w[r,n]}$ follows a normal distribution $\mathcal{N}(\mu_{d[r,n-6]}, \sigma^2_{d[r,n-6]}/2^6)$. Thus, we get

$$f_{w[r,n]}(x) = \phi_{(\mu_{d[r,n-6]}, \sigma^2_{d[r,n-6]}/2^6)}(x).$$

Using the above preparations, the success probability of the key recovery attack on χ^2-attack is evaluated as follows.

Theorem 4. *The success probability P_S of e-bit key recovery algorithm to r-round RC5-64 with 2^n plaintexts can be evaluated by using $f_{c[r,n]}(x)$ and $f_{w[r,n]}(x)$ as follows,*

$$P_S = \int_{-\infty}^{\infty} f_{c[r,n]}(x) * \left(\int_{-\infty}^{x} f_{w[r,n]}(u) du \right)^{2^e - 1} dx.$$

Proof. The e-bit key can be recovered correctly if and only if the χ^2-value of a correct key is higher than that of all $2^e - 1$ wrong keys. This means that the key recovery algorithm to r-round RC5-64 with 2^n plaintexts succeeds if and only if

$$X_{c[r,n]} > X_{w[r,n]} \ (\forall w).$$

From Hypothesis 2, any distribution on wrong keys is independent and approximately equal to each other, which is denoted by $X_{w[r,n]}$. Thus, the success probability P_S can be evaluated by

$$\begin{aligned} P_S &= \Pr(X_{c[r,n]} > X_{w[r,n]})^{2^e-1} \\ &= \int_{-\infty}^{\infty} f_c(x) * \left(\int_{-\infty}^{x} f_w(u)du \right)^{2^e-1} dx. \end{aligned}$$

Theorem 5. *The success probability P_S of e-bit key recovery algorithm to r-round RC5-64 with 2^n plaintexts can be evaluated by using the distributions of χ^2-values in the distinguishing algorithm as follows,*

$$P_S = \int_{-\infty}^{\infty} \phi_{(\mu_{d[r-1,n-6]}, \sigma^2_{d[r-1,n-6]}/2^6)}(x) * \left(\int_{-\infty}^{x} \phi_{(\mu_{d[r,n-6]}, \sigma^2_{d[r,n-6]}/2^6)}(u)du \right)^{2^e-1} dx.$$

Proof. Theorem 5 follows immediately from Lemmas 1 and 2 and Theorem 4.

Theorem 5 indicates the following two factors for high success probability.

- **(Factor 1)** Maximize the average of χ^2-values computed by a correct key;
- **(Factor 2)** Minimize the variances (the error) of each distribution of χ^2-values computed by each key.

3.3 Accuracy of the Approximations of the security on RC5-64

We estimate the success probability of Algorithm 3 by using Theorem 5. In the beginning, we conduct the following distinguishing test on 2 - 4 rounds and get the distribution of χ^2-values on $\mathrm{lsb}_6(R_{h+1})$, $X_{d[r,n]}$. Our experiments use 100 kinds of plaintexts and 100 keys and, thus, conduct 10000 trials in total.

Distinguishing Test: The χ^2-test on $\mathrm{lsb}_6(R_{h+1})$ with $\mathrm{lsb}_6(R_0) = 0$.

The experimental results are shown in Table 2.

The success probability of Algorithm 3 to RC5-64, based on Theorem 5, is computed on Table 3. To evaluate the estimation, we also implement Algorithm 3 on 2-round and 3-round RC5-64. Our implementations generate all plaintexts by using M-sequence: Algorithm 3 uses 122-bit random numbers generated by M-sequence, whose primitive polynomial of M-sequence is $x^{122} + x^{108} + x^8 + x + 1$. The platform is IBM RS/6000 SP (PPC 604e/332MHz $\times$ 256) with memory of 32 GB. Table 3 shows the implemented results among 100 keys for RC5-64 with 3 - 4 rounds. Comparing the estimation with the implemented results, we see that our theory can evaluate the success probability of key recovery algorithm of χ^2-attack. Furthermore, the necessary number of plaintexts for this evaluation is reduced by 2^6 from that of Table 3. In summary, our theory can evaluate the success probability in χ^2-attack by using less number of plaintexts.

Table 2. mean and variance for $X_{d[r,n]}$ ($r = 2, 3, 4$, 10000 trials)

mean $\mu_{d[r,n]}$ (variance $\sigma_{d[r,n]}$)			
#texts	2 rounds	3 rounds	4 rounds
2^{12}	63.41 (125.83)	63.02 (126.76)	–
2^{13}	63.40 (130.56)	62.91 (125.31)	–
2^{14}	64.03 (132.04)	62.96 (125.09)	–
2^{15}	64.80 (139.10)	62.91 (123.37)	–
2^{16}	66.63 (149.53)	62.97 (125.71)	–
2^{17}	70.02 (178.01)	62.83 (124.80)	–
2^{20}	–	63.09 (124.31)	62.84 (123.69)
2^{21}	–	63.41 (126.33)	62.99 (123.79)
2^{22}	–	63.68 (125.68)	63.01 (122.39)
2^{23}	–	64.39 (130.16)	63.18 (124.95)
2^{24}	–	66.11 (135.60)	63.06 (122.94)
2^{25}	–	69.22 (152.86)	63.05 (124.57)
2^{26}	–	75.43 (174.06)	63.23 (124.58)

4 Theoretical analysis on χ^2-attacks to RC6 without post-whitening

We apply Theorem 5 to a key recovery algorithm on RC6P [7] and investigate the accuracy of approximations by comparing it with implemented results.

4.1 Key recovery algorithm of RC6P

Intuitively, a key recovery algorithm [7] fixes some bits out of $\mathrm{lsb}_n(B_0)||\mathrm{lsb}_n(D_0)$, check the χ^2-value of $\mathrm{lsb}_3(A_r)||\mathrm{lsb}_3(C_r)$, and recover $\mathrm{lsb}_2(S_{2r})||\mathrm{lsb}_2(S_{2r+1})$ of r-round RC6P. Here we use the following notation: $(y_b, y_d) = (\mathrm{lsb}_3(B_{r+1}), \mathrm{lsb}_3(D_{r+1}))$, $(x_c, x_a) = (\mathrm{lsb}_5(F(A_{r+1})), \mathrm{lsb}_5(F(C_{r+1})))$, $(s_a, s_c) = (\mathrm{lsb}_2(S_{2r}), \mathrm{lsb}_2(S_{2r+1}))$, $s = s_a||s_c$, and $(S_{2r}[3], S_{2r+1}[3]) = (0, 0)$, where x_a (resp. x_c) is the rotation amounts on A_r (resp. C_r) in the r-th round.

Algorithm 4 ([7])

1. Choose a plaintext (A_0, B_0, C_0, D_0) with $(\mathtt{lsb}_5(B_0), \mathtt{lsb}_5(D_0)) = (0, 0)$ and encrypt it.
2. For each (s_a, s_c), decrypt $y_d||y_b$ with a key $(S_{2r}[3]||s_a, S_{2r+1}[3]||s_c)$ by 1 round to $z_a||z_c$, which are denoted by a 6-bit integer $z = z_a||z_c$.
3. For each s, x_a, x_c, and z, update each array by incrementing $\mathtt{count}[s][x_a][x_c][z]$.
4. For each s, x_a, and x_c, compute $\chi^2[s][x_a][x_c]$.
5. Compute the average $ave[s]$ of $\{\chi^2[s][x_a][x_c]\}_{x_a,x_c}$ for each s and output s with the highest $ave[s]$ as $\mathtt{lsb}_2(S_{2r})||\mathtt{lsb}_2(S_{2r+1})$.

Figure 2 shows the outline of Algorithm 4.

Table 3. Comparison of theoretical and implemented results in Algorithm 3

3 rounds				
theoretical results: mean (variance)				implemented results
#texts	correct key	wrong key	P_S	SUC
2^{20}	64.03 (2.06)	62.96 (1.95)	0.128	0.15
2^{21}	64.80 (2.17)	62.91 (1.93)	0.277	0.36
2^{22}	66.63 (2.34)	62.97 (1.96)	0.683	0.62
2^{23}	70.02 (2.78)	62.83 (1.95)	0.991	0.92
4 rounds				
theoretical results: mean (variance)				implemented results
#texts	correct key	wrong key	P_S	SUC
2^{28}	64.68 (1.96)	63.01 (1.91)	0.080	0.09
2^{30}	66.11 (2.12)	63.06 (1.92)	0.552	0.53
2^{31}	69.22 (2.39)	63.05 (1.95)	0.973	0.89
2^{32}	75.43 (2.72)	63.23 (1.95)	1.000	1.000

SUC: the probability of recovered keys in 100 keys

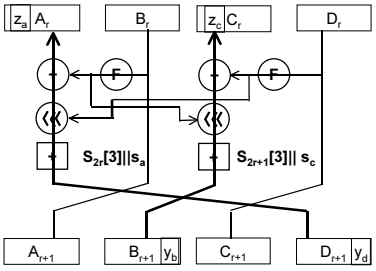

Fig. 2. Algorithm 4

4.2 Success probability of Algorithm 4

By applying Lemmas 1 and 2 and Theorem 4 to Algorithm 4, we get the theorem of success probability on RC6P. Before showing the theorem, we give some notation, which has the same meaning as that in Section 3.2.

- e : recovered-key bit size (There are one correct key and $2^e - 1$ wrong keys.);
- P_S : success probability of a key recovery attack;
- $X_{d[r,n]}$: distributions of χ^2-values on $\mathrm{lsb}_3(A_{r+1})||\mathrm{lsb}_3(C_{r+1})$ of RC6P with $\mathrm{lsb}_5(B_0)||\mathrm{lsb}_5(D_0) = 0$ by using 2^n plaintexts;
- $\mu_{d[r,n]}$ ($\sigma^2_{d[r,n]}$): mean (variance) of distribution of χ^2-values on $\mathrm{lsb}_3(A_{r+1})||\mathrm{lsb}_3(C_{r+1})$ of RC6 with $\mathrm{lsb}_5(B_0)||\mathrm{lsb}_5(D_0) = 0$ by using 2^n plaintexts;
- $X_{c[r,n]}$ ($X_{w[r,n]}$) : distributions of χ^2-values of a key recovery attack to r-round RC6P by using a correct key (a wrong key);
- $\mu_{c[r,n]}$ ($\sigma^2_{c[r,n]}$) : mean (variance) of distribution of mean of χ^2-values of a key recovery attack to r-round RC6P with a correct key by using 2^n plaintexts;
- $f_{c[r,n]}(x)$: probability density function of distribution of χ^2-values with a correct key in r-round RC6P;

– μ_w (σ_w^2) : mean (variance) of distribution of χ^2-values in a key recovery attack to r-round RC6P with a wrong key;
– $f_{w[r,n]}(x)$: probability density function of distribution of χ^2-values with a wrong key in r-round RC6P.

By assuming three hypotheses on wrong-key distribution in Section 3.2, we get the following lemmas and a theorem in the same way as those of RC5-64. The detailed proof will be given in the final version.

Lemma 3. *Let $n \geq 10$ and $r \geq 4$. The distribution of χ^2-values on a correct key in Algorithm 4, $X_{c[r,n]}$, follows a normal distribution of $\mathcal{N}(\mu_{d[r-1,n-10]}, \sigma^2_{d[r-1,n-10]}/2^{10})$. Therefore, the probability density function of distribution of χ^2-values with a correct key in Algorithm 4, $f_{c[r,n]}(x)$, is given by*

$$f_{c[r,n]}(x) = \phi_{(\mu_{d[r-1,n-10]}, \sigma^2_{d[r-1,n-10]}/2^{10})}(x).$$

Lemma 4. *Let $n \geq 10$ and $r \geq 4$. The distribution of χ^2-values on a wrong key in Algorithm 4, $X_{w[r,n]}$, follows a normal distribution of $\mathcal{N}(\mu_{d[r+1,n-10]}, \sigma^2_{d[r+1,n-10]}/2^{10})$. Therefore, the probability density function of distribution of χ^2-values with a correct key in Algorithm 4, $f_{w[r,n]}(x)$, is given by*

$$f_{w[r,n]}(x) = \phi_{(\mu_{d[r+1,n-10]}, \sigma^2_{d[r+1,n-10]}/2^{10})}(x).$$

Lemma 4 is derived from Hypothesis 3. In the case of Algorithm 4, $X_{d[r,n']}$ defined in Algorithm 3 of Hypothesis 3 corresponds to the distributions of χ^2-values on $\mathrm{lsb}_3(B_{r+1})||\mathrm{lsb}_3(D_{r+1}))$, which are equal to that on $\mathrm{lsb}_3(A_{r+2})||\mathrm{lsb}_3(C_{r+2})$ and thus it corresponds to $X_{d[r+1,n']}$ defined in Algorithm 4.

Theorem 6. *The success probability P_S of e-bit key recovery algorithm to r-round RC6P with 2^n plaintexts can be evaluated by using the distributions of χ^2-values in the distinguishing algorithm as follows,*

$$P_S = \int_{-\infty}^{\infty} \phi_{(\mu_{d[r-1,n-10]}, \sigma^2_{d[r-1,n-10]}/2^{10})}(x) * \left(\int_{-\infty}^{x} \phi_{(\mu_{d[r+1,n-10]}, \sigma^2_{d[r+1,n-10]}/2^{10})}(u)du \right)^{2^e-1} dx.$$

4.3 Accuracy of the Approximations of the security on RC6P

We estimate the success probability of Algorithm 4 by using Theorem 6. In the beginning, we conduct the following distinguishing test on 3 and 5 rounds and get the distribution of χ^2-values on $\mathrm{lsb}_3(A_{r+1})||\mathrm{lsb}_3(C_{r+1})$, $X_{d[r,n]}$. Our experiments use 100 kinds of plaintexts and 100 keys and thus conduct 10000 trials in total.

Distinguishing Test:
The χ^2-test on $\mathrm{lsb}_3(A_{r+1})||\mathrm{lsb}_3(C_{r+1})$ with $\mathrm{lsb}_3(B_0)||\mathrm{lsb}_3(D_0) = 0$.

The experimental results are shown in Table 4.

The success probability of Algorithm 4 to RC6P, based on Theorem 6, is computed on Table 5. To evaluate the estimation, we implement Algorithm 4

Table 4. mean and variance for $X_{d[r,n]}$ ($r = 3, 5$, 10000 trials)

mean $\mu_{d[r,n]}$ (variance $\sigma_{d[r,n]}$)		
#texts	3 rounds	5 rounds
2^6	63.03(124.25)	63.02(123.82)
2^7	63.05(125.25)	63.02(125.06)
2^8	63.12(125.92)	62.99(125.47)
2^9	63.26(126.89)	63.01(125.82)
2^{10}	63.51(128.07)	63.02(125.97)
2^{11}	64.08(130.62)	62.99(125.97)
2^{12}	65.17(135.41)	63.00(125.98)

on 4-round RC6P. Our implementations generate all plaintexts by using M-sequence: Algorithm 4 uses 118-bit random numbers generated by M-sequence, whose primitive polynomial of M-sequence is $x^{118}+x^{36}+x^8+x+1$. The platform is the same as that in Section 3.3. Table 5 also shows implemented results among 100 keys for 4-round RC6P. Comparing the estimation with implemented results, we see that our theorem can evaluate the success probability of key recovery algorithm of χ^2-attack. Furthermore, the necessary number of plaintexts for this evaluation is reduced by 2^{10} from that of Table 5. In summary, our theory can also evaluate the success probability in χ^2-attack by using less number of plaintexts.

Table 5. Comparison of theoretical and implemented results in Algorithm 4 to 4-round RC6P

theoretical results: mean (variance)				implemented results
#texts	correct key	wrong key	P_S	SUC
2^{18}	63.12(0.123)	62.99(0.123)	0.111	0.11
2^{19}	63.26(0.124)	63.01(0.123)	0.185	0.15
2^{20}	63.51(0.125)	63.02(0.123)	0.388	0.40
2^{21}	64.08(0.128)	62.99(0.123)	0.882	0.75
2^{22}	65.17(0.132)	63.00(0.123)	1.000	1.00

SUC: the probability of recovered keys in 100 keys

4.4 Approximations of the security on 6-round RC6P

By using Theorem 6, we can estimate the security on 6-round RC6P theoretically although it is not easy to compute experimentally. The experimental results of distinguishing test on 5- and 7-round RC6P are shown in Table 6. The approximation of the security on 6-round RC6P is shown in Table 7. The results indicate that a correct key on 6-round RC6P can be recovered by using 2^{16} times as many

texts as those on 4-round RC6P, which reflects the estimation of security of RC6 or RC6P [11, 7].

Table 6. mean and variance for $X_{d[r,n]}$ ($r = 5, 7$, 10000 trials)

mean $\mu_{d[r,n]}$ (variance $\sigma_{d[r,n]}$)		
#texts	5 rounds	7 rounds
2^{24}	63.3060(128.796)	63.0142(126.729)
2^{25}	63.3729(126.632)	63.0180(126.843)
2^{26}	63.7322(129.106)	62.8563(126.375)
2^{27}	64.4361(132.361)	62.8587(126.070)
2^{28}	66.0068(141.824)	63.1763(124.279)

Table 7. Approximations of the security on 6-round RC6P

#texts	correct key	wrong key	P_S
2^{34}	63.31(0.126)	63.01(0.124)	0.215
2^{35}	63.37(0.124)	63.02(0.124)	0.263
2^{36}	63.73(0.126)	62.86(0.123)	0.747
2^{37}	64.44(0.129)	62.86(0.123)	0.990
2^{38}	66.01(0.139)	62.86(0.123)	1.000

SUC: the probability of recovered keys in 100 keys

5 Comparison of approximation theorems of χ^2-attack

Another approximation of success probability was proposed in [19]. It is based on order statistics and applied to differential and linear attack. Although it is also applicable to χ^2-attack, the accuracy has not been reported yet. From the point of view of accuracy of success probability in χ^2-attack, we compare our theory to [19].

5.1 Success probability based on order statistic

The main idea of an analysis based on order statistics is as follows:

1. distributions of a correct key follows a normal distribution;
2. distributions of wrong keys are sorted in increasing order;
3. the highest distribution of wrong keys follows a normal distribution;
4. the success probability is computed as the probability that the distribution of correct key is greater than the highest distribution of wrong keys.

We may note that assumptions on distributions of a correct key and a wrong key is the same as those in Section 3.2. When we apply an analysis of order statistics to e-bit key recovery on RC5-64 or RC6P, the success probability is computed as follows:

1. distributions of a correct key follows a normal distribution $\mathcal{N}(\mu_{c[r,n]}, \sigma^2_{c[r,n]})$;
2. distributions of wrong keys are sorted in increasing order, $X_{w(1)[r,n]}, \cdots, X_{w(2^e-1)[r,n]}$;
3. the highest distribution $X_{w(2^e-1)[r,n]}$ are assumed to follow a normal distribution $\mathcal{N}(\mu_{[r,n]}, \sigma^2_{[r,n]})$, where the average and the variance are given as:

$$\mu_{[r,n]} = \mu_{w[r,n]} + \sigma_{w[r,n]}\Phi^{-1}(1-2^{-e}) \text{ and } \sigma_{[r,n]} = \frac{\sigma_{w[r,n]}}{\phi(\Phi^{-1}(1-2^{-e}))}2^{-e}.$$

Then, the success probability is computed as the probability that the distribution of correct key is greater than the highest distribution of wrong keys as follows:

Theorem 7 ([19]). *The success probability P_{Sel} of e-bit key recovery algorithm can be evaluated by using the distributions of χ^2-values in the distinguishing algorithm as follows*

$$P_{Sel} = \int_{-\frac{\mu_{c[r,n]}-\mu_{[r,n]}}{\sqrt{\sigma^2_{c}[r,n]+\sigma^2_{[r,n]}}}}^{\infty} \phi(x)dx,$$

where

$$\mu_{[r,n]} = \mu_{w[r,n]} + \sigma_{w[r,n]}\Phi^{-1}(1-2^{-e}) \text{ and } \sigma_{[r,n]} = \frac{\sigma_{w[r,n]}}{\phi(\Phi^{-1}(1-2^{-e}))}2^{-e}.$$

5.2 Accuracy of approximations of success probability in χ^2-attack

We compare approximations of the success probability of 3-round and 4-round RC5-64 and 4-round RC6P based on our theorems to those of Theorem 7, especially. Table 8 or 9 shows results of 3-round and 4-round RC5-64 or 4-round RC6P, respectively. These results indicate that our approximation is more accurate than Theorem 7. Theorem 7 gives rather loose upper bounds. On the other hand, our theorem approximates the success probability more accurately. Especialy when $P_S > 0.8$, our estimation gives a lower upper bound.

Our theorem deals with distributions of all wrong keys. On the other hand, Theorem 7 deals with only the highest distribution of wrong keys. This is one reason that our theorem can estimate strictly. Furthermore, Theorem 7 aims at dealing with differential or linear attack rather than χ^2-attack. This is why our theorems are more suitable for computing the success probability in χ^2-attack.

6 Conclusion

In this paper, we have proved the theorems that evaluate the success probability in χ^2-attack by using the distinguishing test. The derived formulae can be computed efficiently and provide a practical analysis for the estimation of the success probability in χ^2-attack. We have also demonstrated that our theorems can estimate success probability in χ^2-attacks against RC5-64 and RC6P.

Table 8. Comparison between Theorems 5 and 7 in the accuracy on to RC5-64

#texts	Theorem 5		Theorem 7		SUC
	P_S	error rates	P_{Sel}	error rates	
2^{20}	0.128	15%	0.252	68%	0.15
2^{21}	0.277	23%	0.441	23%	0.36
2^{22}	0.683	10%	0.815	31%	0.62
2^{23}	0.991	8%	0.997	8%	0.92
#texts	Theorem 5		Theorem 7		SUC
	P_S	error rates	P_{Sel}	error rates	
2^{28}	0.080	11%	0.387	330%	0.09
2^{30}	0.552	4%	0.716	35%	0.53
2^{31}	0.973	9%	0.991	11%	0.89
2^{32}	1.000	0%	1.000	0%	1.000

SUC: the probability of recovered keys in 100 keys (implemented results)

Table 9. Comparison between Theorems 5 and 7 in the accuracy on to 4-round RC6P

#texts	Theorem 5		Theorem 7		SUC
	P_S	error rates	P_{Sel}	error rates	
2^{18}	0.111	0.9%	0.150	36%	0.11
2^{19}	0.185	23%	0.233	55%	0.15
2^{20}	0.388	3%	0.452	13%	0.40
2^{21}	0.882	18%	0.916	22%	0.75
2^{22}	1.000	0%	1.000	0%	1.00
$2^{21.1}$	0.889	16%	0.922	20%	0.768
$2^{21.2}$	0.921	12%	0.947	16%	0.817
$2^{21.3}$	0.946	12%	0.966	14%	0.846
$2^{21.4}$	0.966	8%	0.980	9%	0.896
$2^{21.5}$	0.979	7%	0.989	8%	0.919

SUC: the probability of recovered keys in 100 or 1000 keys (implemented results)

References

1. A. Biryukov and E. Kushilevitz, "Improved Cryptanalysis of RC5", *Proc. EUROCRYPT'98*, **1403**(1998), pp. 85–99, Springer-Verlag.
2. J. Borst, B. Preneel, and J. Vandewalle, "Linear Cryptanalysis of RC5 and RC6", *Proc. Fast Software Encryption'99*, **1636**(1999), pp. 16–30, Springer-Verlag.
3. R.J. Freund and W.J. Wilson, *Statistical Method*, Academic Press, San Diego, 1993.
4. H. Gilbert, H. Handschuh, A. Joux, and S. Vaudenay, "A Statistical Attack on RC6", *Proc. Fast Software Encryption'2000*, **1978**(2000), pp. 64–74, Springer-Verlag.
5. J. Hayakawa, and T. Shimoyama, "Correlation Attack to the Block Cipher RC5 and the Simplified Variants of RC6", *Third AES Candidate Conf.*, April 2000.
6. H. Handschuh and H. Gilbert, "χ^2 Cryptanalysis of the SEAL Encryption Algorithm", *Proc. Fast Software Encryption*, **1267**(1997), pp.1–12.
7. N. Isogai, T. Matsunaka, and A. Miyaji, "Optimized χ^2-attack against RC6", *Proc. ANCS 2003*, **2846**(2003), to appear in Springer-Verlag.

8. Pascal Junod, "On the Complexity of Matsui's Attack", *Proc. Selected Areas in Cryptography'01*, **2259**(2001), pp. 199–211, Springer-Verlag.
9. B. Kaliski and Y. Lin, "On Differential and Linear Cryptanalysis of the RC5 Encryption Algorithm", *Proc. CRYPTO'95*, **963**(1995), pp. 171–184, Springer-Verlag.
10. J. Kelsey, B. Schneier, and D. Wagner, "Mod n Cryptanalysis, with applications against RC5P and M6", *Proc. Fast Software Encryption'99*, **1636**(1999), pp. 139–155, Springer-Verlag.
11. L. Knudsen and W. Meier, "Improved Differential Attacks on RC5", *Proc. CRYPTO'96*, **1109**(1996), pp. 216–228.
12. L. Knudsen and W. Meier, "Correlations in RC6 with a reduced number of rounds", *Proc. Fast Software Encryption'2000*, **1978**(2000), pp. 94–108, Springer-Verlag.
13. D. Knuth, *The art of computer programming*, vol.2, Seminumerical Algorithms, 2nd ed., Addison-Wesley, Reading, Mass. 1981.
14. A. Menezes, P.C. van Oorschot, and S. Vanstone, *Handbook of applied cryptography*, CRC Press, Inc., Boca Raton, 1996.
15. A. Miyaji, M. Nonaka, and Y. Takii, "Known plaintext correlation attack against RC5", *Proc. RSA'2002 Conf.*, **2271**(2002), pp. 131–148, Springer-Verlag.
16. A. Miyaji and M. Nonaka, "Cryptanalysis of the Reduced-Round RC6", *Proc. ICICS 2002*, **2513**(2002), pp.480-494.
17. R. Rivest, "The RC5 Encryption Algorithm", *Proc. Fast Software Encryption'95*, **1008**(1995), pp. 86–96, Springer-Verlag.
18. R. Rivest, M. Robshaw, R. Sidney, and Y. Yin, "The RC6 Block Cipher. v1.1," August 20, 1998. Available at http://www.rsasecurity.com/rsalabs/rc6/.
19. A. A. Selcuk and A. Bicak, "On probability of success in differential and linear cryptanalysis", *Security in Communication Networks SCN 2002*, Lecture notes in Computer Science, **2576**(2003), 1751-185, Springer-Verlag. Available at http://www.rsasecurity.com/rsalabs/rc6/.
20. S. Vaudenay, "An Experiment on DES Statistical Cryptanalysis", *Proc. 3rd ACM Conference on Computer and Communications Security*, ACM Press, pp.139–147, 1996.