Linearly Weak Keys of RC5

H. M. Heys

Faculty of Engineering and Applied Science
Memorial University of Newfoundland
St.John's, Newfoundland, Canada A1B 3X5
Email: howard@engr.mun.ca

Abstract: In this Letter, we examine the application of linear cryptanalysis to the RC5 private-key cipher and show that there are expected to be weak keys for which the attack is applicable to many rounds. It is demonstrated that, for the 12-round nominal RC5 version with a 64-bit block size and a 128-bit key, there are 2^{28} weak keys for which only about 2^{17} known plaintexts are required to break the cipher and there are 2^{68} keys for which the cipher is theoretically breakable requiring about 2^{57} known plaintexts. The analysis highlights the sensitivity of RC5 security to its key scheduling algorithm.

Key Words: cryptography, private-key ciphers, RC5, linear cryptanalysis, weak keys

Introduction: RC5 [1] is a class of private-key block ciphers designed to be efficiently implemented in software by utilizing three basic operations: exclusive-OR, addition, and data-dependent rotations. An RC5 cipher is designated as RC5-$w/r/b$ where w represents the word size of the target processor in bits, r is the number of rounds of the cipher, and b is the number of bytes of key. The cipher block size is fixed at $2w$ bits. The nominal version of the cipher is RC5-32/12/16.

The application of the two powerful attacks of linear and differential cryptanalysis to RC5 is considered by Kaliski and Yin [2] who show that only 7 rounds of the nominal version of RC5 are required to thwart a linear attack; the full 12 rounds are required to provide security against the differential attack. In [3], Knudsen and Meier extend the analysis of differential attacks to RC5 and show that, by searching for the appropriate plaintexts to use, the complexity of the attack can be reduced by a factor of up to 512 for a typical key of the nominal RC5. As well, it is shown that keys exist which make RC5 even weaker against differential cryptanalysis.

In this Letter, we investigate the existence of weak keys for RC5 with respect to linear cryptanalysis. It is shown that, based on the randomizing assumptions of the key scheduling algorithm, keys exist for which the application of a linear attack is much more likely to be successful.

Description of the Cipher: In our description of the cipher we use the notation of [2]. All words are represented with the most significant bit at the left and the i-th least significant bit of a word X is represented by $X[i-1]$. The algorithm consists of r rounds involving the application of $2r+2$ subkeys. Alternatively, the cipher can be viewed as the mixing of 2 subkeys with the plaintext, followed by $2r$ half-rounds. Let L_0 and R_0 represent the left and right half of the plaintext input, respectively, each half consisting of w bits. We use the notation L_k and R_k to represent the left and right half, respectively, of the cipher data after the $(k-1)$-th half-round. Also S_k represents the k-th subkey consisting of w bits associated with the $(k-1)$-th half-round and generated by the key scheduling algorithm of [1]. Letting L_{2r+1} and R_{2r+1} represent the left and right half of the ciphertext, respectively, the RC5 encryption algorithm is given by:

$$\begin{array}{l} L_1 = L_0 + S_0 \\ R_1 = R_0 + S_1 \\ \text{for } k = 2 \text{ to } 2r+1 \text{ do} \\ \quad L_k = R_{k-1} \\ \quad R_k = ((L_{k-1} \oplus R_{k-1}) \leftarrow R_{k-1}) + S_k \end{array} \tag{1}$$

where "+" represents addition modulo-2^w, "$\oplus$" represents bit-wise exclusive-OR, and "$X \leftarrow Y$" is the rotation of X by Y^{rot} positions to the left with Y^{rot} representing the $\log_2 w$ least significant bits of Y.

Review of Linear Cryptanalysis Applied to RC5: The basic principle of linear cryptanalysis [4] is to find expressions consisting of a mod-2 linear combination of plaintext, ciphertext, and key bits which holds with a probability $p \neq 1/2$. In [2], using the assumption that each round in the cipher is independent, Kaliski and Yin determine the most likely linear approximation by concatenating individual round approximations of the form $R_1[0] \oplus R_0[0] = S_1[0]$, $L_k[0] = R_{k-1}[0]$ for even k, $2 \leq k \leq 2r$, and $R_k[0] \oplus L_{k-1}[0] = S_k[0]$ for odd k, $3 \leq k \leq 2r-1$. The resulting $(r-1)$-round approximation $L_{2r}[0] \oplus R_0[0] = S_1[0] \oplus S_3[0] \oplus \ldots \oplus S_{2r-1}[0]$ can be used to attack an r-round cipher. Using the piling-up lemma [4], the bias of $L_{2r}[0] \oplus R_0[0]$ is given by

$$\epsilon_{r-1} = |P(L_{2r}[0] = R_0[0]) - 1/2| = 1/(2w^{r-1}). \tag{2}$$

To derive the overall cipher key, Kaliski and Yin [2] demonstrate that the bias $\epsilon_{r-1} \neq 0$ can be exploited to determine the cipher subkeys using about N_P known plaintexts where

$$N_P = w \cdot \epsilon_{r-1}^{-2}. \tag{3}$$

Modification to the Linear Attack: The implementation of the linear attack in [2] can be modified by considering that one of w^2 possible values for the $\log_2 w$ least significant bits of L_0 and R_0, i.e. L_0^{rot} and R_0^{rot}, will result in $R_1^{rot} = R_0^{rot} + S_1^{rot} = 0$ and $R_2^{rot} = L_0^{rot} + S_0^{rot} + S_2^{rot} = 0$. When this is true, $R_3^{rot} = S_3^{rot}$ and a linear attack for an r-round cipher can now be based on an $(r-2)$-round approximation of $L_{2r}[0] = R_3[0] = S_3[0]$. The resulting bias of the $(r-2)$-round approximation is significantly larger than the bias of an $(r-1)$-round approximation which does not use fixed values for L_0^{rot} and R_0^{rot}.

For a cipher of r rounds, the number of known plaintexts required for cryptanalysis is based on the number of plaintexts required for an $(r-2)$-round approximation multiplied by w^2. Hence, using this modification to the attack, based on a linear approximation bias of $\epsilon_4 = 1/(2w^4)$, the linear attack requires about $2^{10} \cdot 2^{47} = 2^{57}$ known plaintexts to break a 6-round cipher with the nominal word and key sizes. Note that, although this modified approach to the attack results in the same complexity as the implementation outlined by Kaliski and Yin [2], we shall use this modified approach as the basis for the analysis of the next section.

Keys That Trivialize Rounds in the Attack: In this section, we investigate the existence of weak keys with respect to linear cryptanalysis. We begin by examining a scenario of extremely weak keys. Similarly to the modified implementation of the attack described in the previous section, weak keys can be identified when the inputs to L_0^{rot} and R_0^{rot} are such that $R_1^{rot} = 0$ and $R_2^{rot} = 0$, implying $R_3^{rot} = S_3^{rot}$. Now assume that $S_3^{rot} = 0$. The probability of this being true for the selected key is $1/w$ if the key schedule generates reasonably pseudo-random subkeys. If $S_3^{rot} = 0$, then $R_3^{rot} = 0$, $L_3^{rot} = 0$, and

$R_4^{rot} = S_4^{rot}$. If we extend this argument so that $S_i^{rot} = 0$ for $3 \leq i \leq 2(r-1)$, then $L_{2r}^{rot} = R_{2r-1}^{rot} = S_{2r-1}^{rot}$ for all input plaintexts with the correct values of L_0^{rot} and R_0^{rot}. In such cases, the bias of the approximation $L_{2r}[0] = 0$ is $1/2$ and the subkey values are easily determined using the techniques of [2]. We label this class of extemely weak keys as WK_{r-1} since it trivializes the first $r-1$ rounds of the cipher and requires $2(r-1)-2$ subkeys to have the least significant $\log_2 w$ bits equal to 0. Assuming that the subkey values are random and independent, the probability of this weak key occurring is given by the product of the probabilities that the partial subkey at each round is 0. Hence, $P(WK_{r-1}) = 1/(w^{2r-4})$. To discover the weakness and determine the key bits, about $4w$ known plaintexts corresponding to each of the w^2 possible values for L_0^{rot} and R_0^{rot} must be examined. In the case of the 12-round nominal version of RC5, the probability that the selected cipher key is in the WK_{r-1} class is 2^{-100}. Hence, out of 2^{128} cipher keys, we expect there to be 2^{28} weak keys in class WK_{r-1} that make the cipher easily breakable using about 2^{17} known plaintexts.

We can generalize the argument by defining weak keys notated WK_m for $1 \leq m \leq r-1$, where $S_i^{rot} = 0$ for all i, $3 \leq i \leq 2m$. For a key in the class WK_m, when L_0^{rot} and R_0^{rot} are selected so that $R_1^{rot} = 0$ and $R_2^{rot} = 0$, the bias of approximation $L_{2r}[0] = 0$ is given by the bias for an $(r-m-1)$-round linear approximation. The number of known plaintexts required is given by w^2 multiplied by the number of known plaintexts required in an attack using an $(r-m-1)$-round linear approximation, resulting in

$$N_P = 4w^{2(r-m)+1}. \tag{4}$$

The probability that a key in the WK_m class is selected is given by

$$P(WK_m) = \frac{1}{w^{2m-2}}. \tag{5}$$

Hence, there is a tradeoff between the probability that a key selected is weak and the amount of weakness of the key. The two extreme cases are class WK_1 (the class containing all keys) and WK_{r-1}.

The compromise cases are particularly interesting. For example, considering the nominal RC5 of 12 rounds, the likelihood of selecting a key that trivializes the first 7 rounds (i.e., WK_7) is 2^{-60} and the resulting linear approximation used for the remaining 4 rounds implies that the number of known plaintexts required is about $2^{10} \cdot 2^{47} = 2^{57}$, which is well below the complexity of the approximately 2^{128} encryptions required in an exhaustive key search. For a cipher of a reduced number of rounds, the linear attack can be practically applicable for a potentially large set of keys. For example, for an 8-round cipher, there are 2^{88} keys which require only 2^{37} known plaintexts in the attack. A summary of some sample results of the analysis for the nominal word and key sizes is given in Table 1.

In order to establish the validity of the randomizing assumptions for the key scheduling algorithm, 2^{24} keys where randomly selected and the number of keys satisfying

$$S_3^{rot} = ... = S_i^{rot} = 0 \tag{6}$$

was found to be 523748, 16351, 494, and 14, for i = 3, 4, 5, and 6, respectively. These results are consistent with the assumption of randomness for the subkey values which implies theoretical values of 524288, 16384, 512, and 16, respectively.

Conclusion: In this Letter, we have shown that, for some keys, RC5 can be significantly more vulnerable to linear cryptanalysis than previously implied. Although the analysis presented here does not seem to pose a practical threat to the security of the nominal RC5 implementation - either the number of known plaintexts required is too large or the likelihood of selecting a weak key is too small - it does highlight the importance of the key scheduling algorithm and the non-equivalence of RC5 keys. The next step in the work should be to establish that there are not weaknesses in the pseudo-random properties of the RC5 key scheduling algorithm which might more easily lead to the generation of weak keys than is expected by the randomness assumptions of this Letter.

References

[1] R.L. Rivest, "The RC5 Encryption Algorithm", *Proceedings of Fast Software Encryption - Second International Workshop at K.U. Leuven, Dec. 1994*, Springer-Verlag, pp. 86-96, 1995.

[2] B.S. Kaliski and Y.L. Yin, "On Differential and Linear Cryptanalysis of the RC5 Encryption Algorithm", *Advances in Cryptology - CRYPTO '95*, Springer-Verlag, pp. 171-184, 1995.

[3] L.R. Knudsen and W. Meier, "Improved Differential Attacks on RC5", *Advances in Cryptology - CRYPTO '96*, Springer-Verlag, pp. 216-228, 1996.

[4] M. Matsui, "Linear Cryptanalysis Method for DES Cipher", *Advances in Cryptology - EUROCRYPT '93*, Springer-Verlag, pp. 386-397, 1994.

Rounds	*Kaliski and Yin*		*Weak Keys*		
r	N_P	$\#keys$	N_P	m	$\#keys$
6	2^{57}	2^{128}	2^{37}	3	2^{108}
7	–	–	2^{37}	4	2^{98}
8	–	–	2^{37}	5	2^{88}
10	–	–	2^{47}	6	2^{78}
12	–	–	2^{57}	7	2^{68}
12	–	–	2^{37}	9	2^{48}
12	–	–	2^{17}	11	2^{28}

Table 1: Sample Complexities of the Linear Attack on RC5-32/r/16