Applying Conditional Linear Cryptanalysis to Ciphers with Key-Dependant Operations

REINER DOJEN
TOM COFFEY
Data Communications Security Laboratory
Department of Electronic and Computer Engineering
University of Limerick
IRELAND
reiner.dojen@ul.ie, tom.coffey@ul.ie http://www.ece.ul.ie

Abstract: - Linear cryptanalysis has been proven to be a powerful attack that can be applied to a number of symmetric block ciphers. However, conventional linear cryptanalysis is ineffective in attacking ciphers that use key-dependent operations, such as ICE, Lucifer and SAFER. In this paper conditional linear cryptanalysis, which uses characteristics that depend on some key-bit values, is introduced. This technique and its application to symmetric ciphers are analysed. The consequences of using key-dependent characteristics are explained and a formal notation of conditional linear cryptanalysis is presented. As a case study, conditional linear cryptanalysis is applied to the ICE cipher, which uses key-dependant operations to improve resistance against cryptanalysis. A successful attack on ThinICE using the new technique is presented. Further, experimental work supporting the effectiveness of conditional linear cryptanalysis is also detailed.,

Key-Words: - Cryptography, Cryptanalysis, Linear cryptanalysis, Symmetric ciphers, Attack

1   Introduction

As the volume of data transmitted over the world's expanding communication networks increases so also does the demand for data security services such as confidentiality and integrity. This protection can be achieved by using either a public key cipher or a symmetric cipher. Due to their superior speed, symmetric ciphers are particularly suited for securely handling large amounts of data.

In the early 1990's two statistical attacks on symmetric ciphers, differential [1] and linear cryptanalysis [2][3][4], were invented.

Differential cryptanalysis examines pairs of plaintext with a particular difference and analyses these differences throughout the encipher process. Looking at the resulting differences in the ciphertext, the key, or parts of the key, can be deduced. Differential cryptanalysis is a *chosen-plaintext attack*, as both the plaintext and the ciphertext are known to the attacker and the plaintexts have to satisfy particular properties. A chosen-plaintext attack can be changed to a known-plaintext attack, by increasing the amount of the plaintexts, assuming that within the known plaintexts sufficient pairs with the needed properties are available. Ciphers with weaknesses against differential cryptanalysis include SC2000 [5], SAFER+[6], RC5 [7], KHF [8], ICE[9].

In contrast to differential cryptanalysis, linear cryptanalysis uses linear approximations to describe the F-function of a block cipher. Thus, if some plaintext bits and ciphertext bits are XORed together, then the result equals the XOR of some bits of the key (with a probability p, for the key) and any random plaintext/ciphertext pair.

With this method the adversary can deduce part of the key. The remaining key bits can be obtained by a brute force search. Linear cryptanalysis is a *known-plaintext attack*, as the adversary only has to know, rather than to choose, the plaintext and ciphertext pairs. Ciphers susceptible to linear cryptanalysis include SC2000 [5], Serpent [10], SAFER[11], RC6[12], RC5[13].

This paper presents conditional linear cryptanalysis, which is a derivation on linear cryptanalysis. Conditional linear cryptanalysis is useful for attacks on block ciphers that employ key-dependent operations in their F-function. The process of applying conditional linear cryptanalysis to ciphers with key-dependant operations is detailed. As a case study, conditional linear cryptanalysis is applied to ICE. It is shown that a conditional linear attack can be successfully applied to ThinICE, while standard ICE appears to be secure. The results of an experimental attack, which support the effectiveness of conditional linear cryptanalysis, are also presented.

2 Conditional Linear Cryptanalysis

Symmetric block ciphers sometimes use key-dependant operations (other than simple XOR) to prevent the application of cryptanalysis. Examples for such key-dependant operations include the keyed permutation of ICE [14], the exchange of nibbles in Lucifer [15] and the combination of bytes with sub-keys in SAFER [16]. In Lucifer and ICE certain parts of data are exchanged according to some sub-key bits. As a result, the exact path of the data-bits through the cipher can only be determined if the key is known (if the part of the key that controls the key-dependant operation is known). Conventional linear cryptanalysis cannot be applied to such key-dependant ciphers, as a requirement for this attack is the ability to trace the exact path of bits through the cipher. Here conditional linear cryptanalysis is presented as an enhancement of linear cryptanalysis. Conditional linear cryptanalysis uses key-dependant characteristics to circumvent the problems introduced by the key-dependant operations.

2.1 The Operation of Conditional Linear Cryptanalysis

Conditional linear cryptanalysis, like conventional linear cryptanalysis, is a known plaintext attack and is detailed as follows: A characteristic with a sufficiently high bias is selected and those key bits (or sub-key bits) that affect the bits of the chosen characteristic with regard to the key-dependant operation are determined. In the next step, fixed values are assigned to these key bits. The assignment of values to key bits turns the used characteristic into a conditional characteristic. The conditional characteristic is valid, only if the key used for encryption satisfies the condition that the selected bits have the assigned values, otherwise it is void. This assignment effectively divides the key-space in two parts: The first part, also called the covered key-space, contains all the keys for which the characteristic is valid, the second contains all the keys for which the characteristic is void. Under the assumption that the assignment of key bits is correct the attacker knows the exact path of the selected bits through the cipher. This knowledge allows the attacker to continue the attack according to conventional linear cryptanalysis.

2.2 The Consequences of Applying Conditional Characteristics

When using a conditional characteristic, the attacker assumes that some key bits have a fixed value. Hence, the attack can only succeed if this assumption is correct. If the key used is outside the covered key-space, the attack will fail.

At a first glance this seems to render the attack useless, as the size of the covered key-space decreases exponentially with the number of fixed key bits. However, multiple conditional characteristics can be used simultaneously to increase the overall covered key-space significantly.

In order to maximize the effectiveness of the attack, it is advisable to use characteristics that are affected by as few key-bits as possible. Further care must be taken in combining multiple rounds, to ensure that none of the assumptions contradict each other. As most ciphers expand their key into sub-keys, the same key-bit most likely will affect multiple key-dependant operations in several rounds. Hence, the attacker must be careful not to use characteristics that assume opposite values for the same key-bit.

2.3 Formal Notation of Conditional Linear Cryptanalysis

Conditional linear cryptanalysis is formally notated as follows:

P	The plaintext (the data before the first round = L_0R_0).
C	The ciphertext (the data after the last round).
L_r, R_r	The left/right data half after round r.
K	The secret key.
SK_r	The sub-key used in round r.
$F(R_{r-1}, SK_r)$	The F-function as applied in round r.
$A[i]$	The i-th bit of the binary vector A.
$A[i,j,...,k]$	The binary XOR of the named bits, i.e. $A[i] \oplus A[j] \oplus \ldots \oplus A[k]$
$\mid i,j,....,k$	Condition on key bits: 'i' indicates $K[i]=1$, '~i' indicates $K[i]=0$.

A conditional linear characteristic is expressed as:

$$P[i_1,i_2,...,i_a] \oplus C[j_1,j_2,...,j_b] =$$
$$K[k_1,k_2,...,k_c] \mid c_1,c_2,...,c_d$$

where $i_1,i_2,\ldots,i_a$, $j_1,j_2,\ldots,j_b$ and $k_1,k_2,\ldots,k_c$ denote fixed bit locations and the $c_1,c_2, \ldots,c_n$ indicate the conditions on the used key K. The notation for a single-round characteristic and a multi-round characteristic is basically the same. Any intermediate terms in a multi-round characteristic appear twice and hence are cancelled out. The conditions on intermediate terms cannot be neglected and need to be added to the list of conditions. Hence, this list increases with every non-trivial round. In the above equation for a conditional linear characteristic sub-key bits can be

used instead of key bits. However, if the key-schedule is not invertible, sub-key bits must be used, as the sub-key governs the key-dependant operation. The above characteristic can also be expressed as:

$$P[i_1, i_2, .., i_a] \oplus C[j_1, j_2, .., j_b] = SK_1[k_{11}, k_{12}, .., k_{1c}] \oplus$$
$$... \oplus SK_r[k_{k1}, k_{k2}, .., k_{kd}] \mid c_1, c_2, .., c_e$$

Note that in this equation, $c_1, c_2, \ldots, c_e$ express conditions on sub-key bits, rather than on key bits.

3 Case Study: Conditional Linear Cryptanalysis Applied to ICE

In this section we present a case study of the application of conditional linear cryptanalysis. A theoretical analysis of the ICE algorithm is detailed. Experimental attacks, resulting from this analysis, are also discussed.

3.1 The ICE Algorithm

ICE [14] is a standard Feistel cipher similar to DES. The standard version has a 64-bit key and works on 16 rounds. There is also a faster version, known as ThinICE, which has a 64-bit key and works on 8 rounds. The open-ended version ICE-n has a *n* times 64-bit key and works on *n* times 16 rounds. It is more secure, but has slower encryption and decryption rates.

The F-function consists of an expansion function, a keyed permutation, XOR with the sub-key, substitution boxes and the permutation function. The keyed permutation, which swaps bits according to sub-key SK3, is regarded as the key-dependent operation for this analysis. Each round uses 60 key bits, which are divided into the three 20-bit values SK1, SK2 and SK3:

The ICE algorithm has already shown weaknesses against differential cryptanalysis. So far the best attack on ICE was presented by Rompay et.al. with their differential attack on ICE [9]. The best multi-round characteristic in their attack on ThinICE has a bias of 2^{-19}.

3.2 Details of an Attack on ICE

Examining the keyed permutation, it can be seen that in this operation each bit of a single S-box is affected by a different sub-key bit. To minimise the number of fixed bits per characteristic (and thereby maximising the size of the covered key-space) only single-bit characteristics are considered. Reviewing all single-bit conventional linear characteristics, it can be seen, that there are no single-bit characteristics that can be combined directly with

the trivial characteristic. However, it will be shown, that some characteristics can be combined with the trivial characteristic to form three-round characteristics. Even though this decreases the overall bias, as only one in three rounds is a trivial round (rather than every second round when multiple-bit characteristics are employed), this approach is preferred as the decrease of the bias is linear. In contrast to this, using multiple-bit characteristics would result in an exponential decrease in the covered key-space. As mentioned before, care must be taken to ensure that the conditions of the used characteristics do not contradict each other. The linear equation for 3-round conditional linear characteristics becomes:

$$L_0[\alpha] \oplus L_3[\beta] = SK_1[\gamma] \oplus SK_2[\delta] \mid v_\alpha, v_\beta$$

where α and β denote single bits in the data halves, v_α and v_β denote the conditions on the sub-key bits and γ and δ denote the bits in the sub-keys that affect α and β in the sub-key XOR, respectively. Figure 1 shows the structure of these 3-round characteristics.
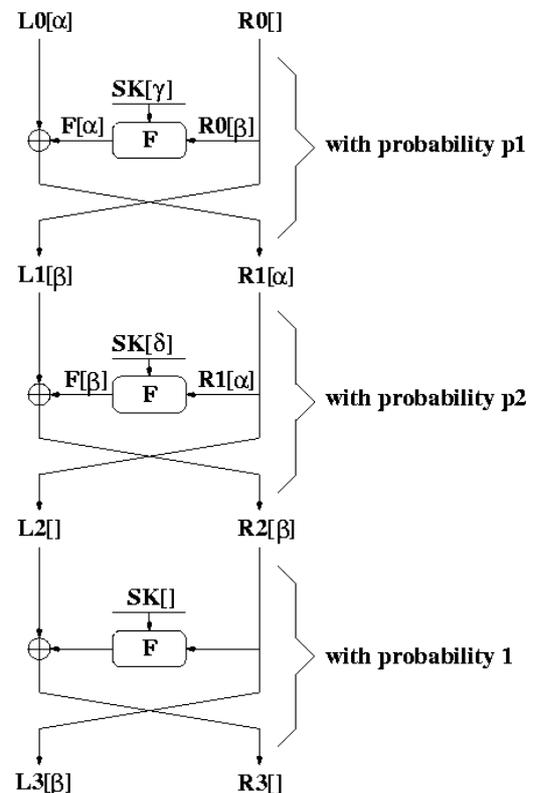

Fig. 1: 3-round conditional linear characteristic

Such a 3-round characteristic is build as follows: Assume a single bit α is used on the left

half (and hence on the result of the F-function) in the first round. This requires the use of a single bit β before the F-function on the right half. As a result, in the next round β is used on the left half and α on the right half. To ensure that the trivial round is used, β must be such, that it requires the use of α before the F-function on the right half. When this condition holds, it leads to the use of an empty pattern on the left half in the next round, as $\alpha \oplus \alpha = \emptyset$. This is exactly the property of the trivial round.

The best 3-round characteristic uses the bit at position 16 in both of the non-trivial rounds, resulting in a bias $\varepsilon = 2^{-9}$. As this characteristic has two fixed key values, it covers 2^{-2} of the key space. There are 89 different conditional characteristics with a bias $\varepsilon \geq 2^{-13}$. Table 1 shows the six characteristics with the highest bias. Each of these characteristics can be used in multiple ways, i.e. the pattern can be applied in several orders. However, the selection of the used characteristic and the order of the pattern must be considered carefully, as some conditions for key-bit values mutually exclude each other.

α	β	$\varepsilon(\log_2)$	Round α	Round β
16	16	-9	SK3[18] = 0	SK3[18] = 0
12	11	-9.68	SK3[14] = 1	SK3[13] = 1
31	31	-9.83	SK3[17] = 0	SK3[17] = 0
29	29	-9.83	SK3[15] = 1	SK3[15] = 1
25	14	-9.83	SK3[11] = 1	SK3[16] = 1
22	19	-9.83	SK3[06] = 1	SK3[03] = 0

Table 1: The best conditional linear characteristics of ICE

3.3 Attack on 4-round ICE
Each of these 3-round characteristics can be directly applied to ICE with 4 rounds in a 1R-attack. There are 20 active key bits, 10 from the XOR with the sub-key and 10 from the keyed permutation. Hence a total of 2^{20} counters are needed to perform an attack. However, the number of active key bits can be reduced, as some of the fixed key bits may be part of the sub-key in the final round. As there are two fixed key bits, a single attack covers a quarter of the key space.

Using the best characteristic, with a bias $\varepsilon = 2^{-9}$, would require typically $\varepsilon^{-2} = 2^{18}$ plaintext/ciphertext pairs. Assuming that the patterns are applied to the left half in an attack on the encryption process, in the order $[\alpha\beta-]$ (where '-' denotes the trivial round), then 19 bits of key

information can be obtained: 18 bits from the adopted key candidate (two fixed key bits are used in the 4th round within the active sub-key bits) and one bit from the equation:

$$P[16] \oplus C[16] \oplus F[16] = K[16,33] \oplus 1 \mid \sim 34, 17$$

If the 4 characteristics ([31,31,-],[29,31,-], [14,31,-], [8,15,-]) are used in an attack, then 75% of the key-space is covered. The attack requires typically $2^{20.84}$ plaintext/ciphertext pairs.

3.4 Attack on ThinICE
For an attack on ThinICE (8-rounds) a 3-round characteristic can be concatenated with itself. Using the best 3-round characteristic and applying the pattern in the order $[-\alpha\beta - \beta\alpha -]$ to the left half, results in a 7 round characteristic with bias $\varepsilon = 2^{-17}$, which covers 2^{-4} of the key space. This attack would typically require 2^{34} plaintexts. The attack on the encryption process reveals 19 bits of key information. The attack can also be optimized for coverage of the key space. This has a bias of $\varepsilon = 2^{-18.66}$ and covers 2^{-3} of the key-space.

The conditional linear attack outlined above is a known plaintext attack. Therefore, an attacker doesn't need the ability to select plaintext/ciphertext pairs with particular properties. The used characteristic has a bias of 2^{-17}, indicating that on average 2^{34} known plaintexts/ciphertext pairs are required. To date, the best published attack on ICE is a chosen plaintext attack by Rompay et.al. [9] with a bias of 2^{-19}. This differential attack requires on average 2^{19} chosen plaintext/ciphertext pairs with a particular difference, which equates to a requirement of 2^{40} known plaintext/ciphertext pairs in a known plaintext attack.

3.5 Attack on Standard ICE
The conditional linear 3-round characteristics can be concatenated five times with themselves, resulting in a 15-round approximation. Using the characteristic with the highest bias leads to an overall bias of $\varepsilon = 2^{-41}$. As it would typically require about 2^{82} plaintexts/ciphertext pairs, standard ICE with 16 rounds appears secure against such an attack.

3.6 An Experimental Attack on ICE
Experimental work was undertaken on reduced versions of the ICE algorithm, to demonstrate the operation of conditional linear cryptanalysis. These attacks on ICE use large numbers of

plaintext/ciphertext pairs and up to 20 active key bits. This is a computationally intensive task and its execution takes a long time. However, as the computation on each plaintext/ciphertext pair is independent of the other pairs, performing the computations in parallel can significantly reduce the execution time.

The following results were computed using a distributed computer system with up to 17 processors. One of the machines served as a master, responsible only for distributing the workload. The other sixteen processors served as slaves, which received part of the workload from the master and performed the calculation on this part. When the computations were completed by the slaves, all partial results were merged by the master.

Using this system configuration, the execution of the conditional linear attack on ThinICE using the characteristic with the highest bias can be interpolated to take 600 days approximately. Note that this attack has only 18 active key-bits, as two of the sub-key bits in the final round are assumed to have fixed values due to the conditional characteristics used. This time can be significantly reduced by using a system with a much higher degree of parallelism. Figure 2 shows how the execution time increases, for 1, 4, 8 and 16 slave processors, as the number of active key bits increases (The dashed line indicates the time of the attack on ThinICE).
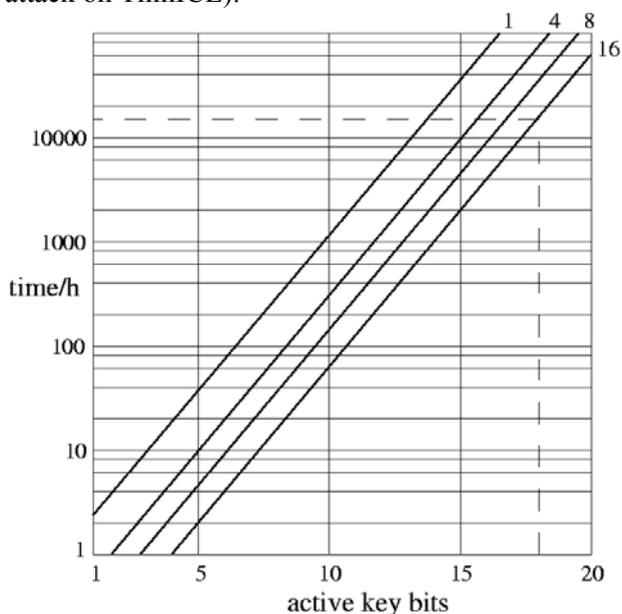

Fig. 2: Execution times for attacks on ThinICE

A reduced attack was used for this experimental work, to secure realistic execution times. The reduced attack works only on parts of either SK2 or SK3. Knowledge of some sub-key bits was assumed and with this information the number of active key-bits was reduced. For example, the attack with 3 active key bits assumes that the other 17 sub-key bits are known or fixed due to the used conditional characteristics.

Figure 3 shows the results of several attacks on ICE with 4-rounds. The number N denotes ε^{-2} (which equals 2^{18} for 4-round ICE) and k denotes the number of active key bits. It can be seen that, for small numbers of pairs, the success rate of the full attack is lower than success rates of the reduced attacks. However, with increasing numbers of pairs the success rates converge. Success is defined as finding the correct key, whereas failure results in finding a wrong key.
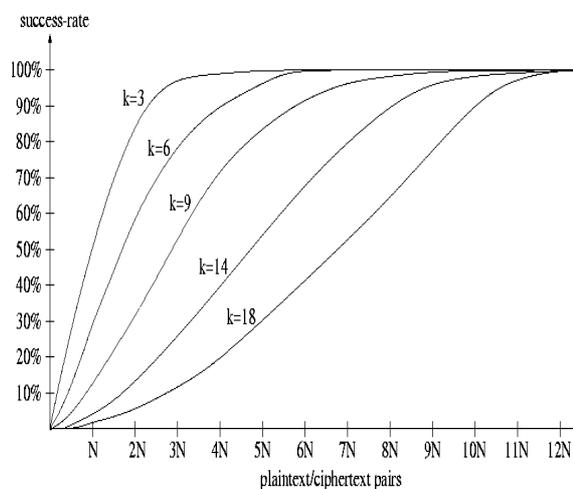

Fig. 3: Success-rates of attacks on 4-round ICE

Table 2 shows the results of the attack on ThinICE (8-rounds) with 3 active key bits. The number N again denotes ε^{-2} (which equals 2^{34} for ThinICE and 2^{18} for 4-round ICE). The results are similar to the attack on 4-round ICE with 3 active key bits. This shows that the full attack on ThinICE performs as expected, even though the reduced attacks only have been undertaken.

Pairs	Hit rate 8-rounds	Hit rate 4-rounds
N/4	17.1%	19%
N/2	35.6%	34%
N	58%	50%
2N	94.1%	84%
4N	98.2%	99%

Table 2: Experimental results on ThinICE

4 Conclusion

Symmetric block ciphers sometimes use key-dependant operations to prevent the application of differential and linear cryptanalysis. Examples of such ciphers include ICE, Lucifer and SAFER.

This paper presented conditional linear cryptanalysis as an enhancement of linear cryptanalysis. Conditional linear cryptanalysis uses key-dependant characteristics to circumvent the problems of applying linear cryptanalysis, which are introduced by the key-dependant operations. The main disadvantage of using conditional characteristics is, that the attack is only valid for a fraction of the overall key-space: The attack can only succeed, if the used encryption key meets the conditions implied by the conditional characteristics. However, since linear cryptanalysis is a known-plaintext attack, multiple attacks with different characteristics can be executed with the same set of plaintext/ciphertext pairs. Thus, the size of the covered key-space can be increased significantly. The overall number of pairs required to successfully execute multiple attacks equals the number of plaintext/ciphertext pairs required by the characteristic with the lowest bias.

As a case study, conditional linear cryptanalysis was applied to ICE. The paper showed how conditional linear cryptanalysis can be successfully applied to ThinICE. This attack requires typically $N=2^{34}$ plaintext/ciphertext pairs and covers 2^{-4} of the key space. The attack can also be optimized for coverage of the key space, covering 2^{-3} of the key-space using $N=2^{37.3}$ plaintext/ciphertext pairs. In an experimental attack on ICE, using a distributed system, it has been shown that by using 12N pairs the attack has a hit rate approaching 100%.

In this paper it has been shown that conditional linear cryptanalysis is effective against symmetric ciphers with key-dependant operations.

5 Acknowledgement

References:
[1] Biham, E. and Shamir, A.: Differential Cryptanalysis of the full 16-round DES, *CRYPTO'92, LNCS740*, Springer Verlag, 1993, pp.487-496
[2] Matsui, M. and Yamagishi, A.: A New Method for Known Plaintext Attack of FEAL Cipher, *EUROCRYPT'92, LNCS658*, Springer Verlag, 1993, pp.81-91
[3] Matsui, M.: Linear Cryptanalysis Method for DES cipher. *EUROCRYPT'93, LNCS765,* Springer Verlag, 1994, pp.386-397
[4] Matsui, M.: The First Experimental cryptanalysis of the Data Encryption Standard. *CRYPTO'94, LNCS839*, Springer Verlag, 1994, pp.1-11
[5] Yanami, H., Shimoyama, T. and Dunkelman, O.: Differential and Linear Cryptanalysis of a Reduced-Round SC2000, *FSE02, LNCS 2365*, Springer Verlag, 2002, pp.34-48
[6] Hu, Y., Zhang Y., and Xiao, G.: Integral Cryptanalysis of SAFER+, *IEE Electronic Letters, Vol.35, No. 17*, 1999, pp.1458-1459
[7] Biryukov, A. and Kushilevitz, E.: From Differential Cryptanalysis to Ciphertext-Only Attacks, *CRYPTO '98, LNCS 1462,* Springer Verlag, 1998, pp.72-88
[8] Wagner, D.: Differential Cryptanalysis of KHF, *FSE98, LNCS 1372*, Springer Verlag, 1998, pp.203-296
[9] Rompay, B. van, Knudsen, L.R. and Rijmen, V.: Differential Cryptanalysis of the ICE Encryption Algorithm, *FSE98, LNCS 1372*, Springer Verlag, 1998, pp.270-283
[10] Biham, E., Dunkelman, O. And Keller, N.: Linear Cryptanalysis of Reduced Round Serpent, *FSE01, LNCS 2355*, Springer Verlag, 2002, pp.16-27
[11] Nakahara Jr., J., Preneel, B. and Vandewalle, J.: Linear Cryptanalysis of Reduced-Round Versions of the SAFER Block Cipher Family, *FSE00, LNCS 1978,* Springer Verlag, 2000, pp.244-261
[12] Borst, J., Preneel B. and Vandewalle, J.: Linear Cryptanalysis of RC5 and RC6, *FSE99, LNCS 1636,* Springer Verlag, 1999, pp.16-30
[13] Heys, H.M.: Linearly Weak Keys of RC5, *IEE Electronic Letters, Vol. 33, No. 10*, 1997, pp.836-838
[14] Kwan, M.: The Design of the ICE Encryption Algorithm, *FSE97, LNCS1267*. Springer Verlag, 1997, pp.69-82
[15] Sorkin, A.: Lucifer, a Cryptographic Algorithm, *Cryptology, Vol. 8, No.1,* 1987, pp.22-41
[16] Massey, J.L.: SAFER K-64: A Byte-Oriented Block-Ciphering Algorithm, *FSE93, LNCS 809,* Springer Verlag, 1994, pp.1-17