Square-like Attacks on Reduced Rounds of IDEA

Hüseyin Demirci

Tübitak UEKAE, 41470 Gebze, Kocaeli, Turkey,
huseyind@uekae.tubitak.gov.tr

Abstract. In this paper we develop two new chosen plaintext attacks on reduced rounds of the IDEA block cipher. The attacks exploit the word structure of the algorithm and are based on the observation that suitable chosen plaintexts give rise to some special kind of distributions which provide a way to distinguish reduced round IDEA output from a random permutation with very few plaintexts. As a result, we develop an attack for 3.5 rounds of IDEA which requires only 103 chosen plaintexts. We have reduced the number of required plaintexts significantly up to 4 rounds. We also present some interesting properties of the reduced round variants of the cipher which have not been published before. The properties and the attacks bring a different approach to analyse the cipher.

1 Introduction

The IDEA block cipher is a modified version of the algorithm PES [9], [10]. The main design concept is "mixing operations from different algebraic groups". The authors have developed the idea of Markov ciphers to evaluate the cipher against differential cryptanalysis. IDEA is an original example of a non Feistel cipher with beautiful mathematical ideas and it has been widely used in commercial environment.

Since its description the process of cryptanalysing IDEA has developed slowly. In [13] and [4], differential cryptanalysis was applied to IDEA reduced to 2 and 2.5 rounds. In [3], 3 and 3.5 round IDEA were cryptanalysed using differential-linear and truncated-differential techniques respectively. Finally in [1], Biham, Biryukov and Shamir used impossible differential technique to sieve the key space for 3.5, 4 and 4.5 rounds. These are currently the best known attacks on IDEA, and the 4.5 round attack requires the encryption of the whole plaintext space.

Recently we are aware of a paper [14] which uses the square attack technique to analyse 2.5 rounds of PES and IDEA. The authors have also developed a related key square attack on 2.5 rounds of IDEA using 2 chosen plaintexts and 2^{17} related keys which recovers 32 key bits.

In this paper we describe some distribution properties of the cipher. Some of these properties are "saturation properties "[12], [7]. Also there are properties which are similar to the ones used in square attacks [5]. We preferred the name "square-like attack" rather than square or integral attack since we exploit the word structure of the algorithm in a different sense. Mainly we are interested

K. Nyberg and H. Heys (Eds.): SAC 2002, LNCS 2595, pp. 147–159, 2003.

in the distribution of some variables more than taking sum (or integral) of the variables [5], [8]. Using these distributions we are able to attack the cipher up to 4 rounds. Our main contribution is that we are able to cryptanalyse the cipher with very few chosen plaintexts. This is a result of the powerful eliminating properties of the distributions. As time complexity, our attacks are not better than the known attacks, but we consider that the distribution properties and the reduction of the number of required plaintexts are surprising. We compare our results with the existing ones in Table 1.

Author	**Rounds**	**No of C. Plaintexts**	**Total Complexity**
[13]	2	2^{10}	2^{42}
[13]	2.5	2^{10}	2^{106}
[4]	2.5	2^{10}	2^{32}
[14]	2.5	3.2^{16}	$3.2^{63} + 2^{48}$
[14], Related Key	2.5	2	$2^{37} + 2^{23}$
[3]	3	2^{29}	2^{44}
[3]	3.5	2^{56}	2^{67}
[1]	3.5	$2^{38.5}$	2^{53}
[1]	4	2^{37}	2^{70}
[1]	4.5	2^{64}	2^{112}
This paper Attack 1	2	23	2^{64}
This paper Attack 1	2.5	55	2^{81}
This paper Attack 1	3	71	2^{71}
This paper Attack 1	3.5	103	2^{103}
This paper Attack 2	3	2^{33}	2^{82}
This paper Attack 2	3.5	2^{34}	2^{82}
This paper Attack 2	4	2^{34}	2^{114}

Table 1. Plaintext and time complexity of the attacks on reduced rounds of IDEA

1.1 Notation

Throughout this paper we will use the following notation. The plaintext is denoted by $(P1, P2, P3, P4)$ and ciphertext is denoted by $(C1, C2, C3, C4)$ where the seperated parts show the 16 bit subblocks. The round numbers are denoted by subindices. Therefore $C_2 1$ denotes the first subblock of the ciphertext after 2 rounds. For convenience we denote the subkeys of the MA-box by $K5$ and $K6$, the inputs of the MA-box by p and q and outputs by t and u. We call p as the first input, q the second input, t the first output and u the second output of the MA-box.

Following [12] and [7] we will call a variable "saturated" if it takes every possible value once. For instance if in the plaintext set $\{(P1, P2, P3, P4)\}$ the element $P4$ takes every 16-bit value once, we say $P4$ is saturated. A variable

is said to be "k-saturated" if every possible element of the variable is observed exactly k times.

For the least significant bit of a variable we use the abbreviation lsb. Finally $K_2 1[97...112]$ means that the subkey subblock $K_2 1$ uses the key bits from 97 to 112 of the master key, including the boundaries.

2 Some Distributions in the IDEA Block Cipher

2.1 IDEA Block Cipher

IDEA is a 8.5 round block cipher which uses 3 different group operations on 16 bit subblocks: XOR, modular addition and IDEA multiplication. This multiplication can be described as $z = x \odot y$, where if any of x or y is 0, we convert that element to 2^{16} and calculate $z = (x \times y)$ modulo $2^{16} + 1$. If z is calculated as 2^{16}, we convert z to 0. Since $2^{16} + 1$ is prime, this multiplication is invertible. In [11] Lai suggests that the cipher satisfies "confusion" by using the fact that these operations are incompatible: there are no general commutativity, associativity or distributivity properties when different operations are used respectively. IDEA multiplication provides a strong non-linear component against linear attacks.

The round function of IDEA consists of two parts: first there is a transformation part of each plaintext subblock with the subkey subblocks, i.e. $T : (P1, P2, P3, P4) \rightarrow (P1 \odot K1, P2 \boxplus K2, P3 \boxplus K3, P4 \odot K4)$. In the second part we have the MA-box. MA-box has two inputs $p = (P1 \odot K1) \oplus (P3 \boxplus K3)$ and $q = (P2 \boxplus K2) \oplus (P4 \odot K4)$. Using p, q and the subkey subblocks $K5, K6$ we produce two output subblocks t and u. The outputs are calculated as $t = ((p \odot K5) \boxplus q) \odot K6$ and $u = (p \odot K5) \boxplus t$. The outputs of the MA-box are XORed with the outputs of the transformation part, and the two middle subblocks are exchanged. After 1 round the ciphertext is of the form $(C1, C2, C3, C4)$ where $C1 = (P1 \odot K1) \oplus t$, $C2 = (P3 \boxplus K3) \oplus t$, $C3 = (P2 \boxplus K2) \oplus u$, $C4 = (P4 \odot K4) \oplus u$. The cipher is composed of 8 full rounds and 1 extra transformation round. The 128 bit master key is cyclicly shifted left 25 bits a few times to fill an array. Then we get the bits for subkey subblocks from this array respectively. Since $2^2 + 1, 2^4 + 1$ and $2^8 + 1$ are also prime, it is possible to build smaller variants of IDEA with similiar properties. IDEA with block size 8, 16 and 32 bits can be built with subblock size 2, 4 and 8 respectively.

Remark From [9] we observe that if $x, y \notin \{0, 1\}$, $(x \odot y) + (x \odot (2^{16} + 1 - y)) = 2^{16} + 1$. Therefore for $x, y \notin \{0, 1\}$, $\mathrm{lsb}(x \odot y) = \mathrm{lsb}(x \odot (2^{16} + 1 - y)) \oplus 1$. Also for any z we have $(0 \odot z) \boxplus (1 \odot z) = 1$. As a result for any value of i, there exists a j value, which satisfies $\mathrm{lsb}(i \odot k) = \mathrm{lsb}(j \odot k) \oplus 1$ for all k. This observation will be important during our key elimination process.

2.2 Some Distributions

With the diffusion properties of the MA-box, a single bit change in the plaintext is able to change every bit of the ciphertext after 1 round. Therefore classical

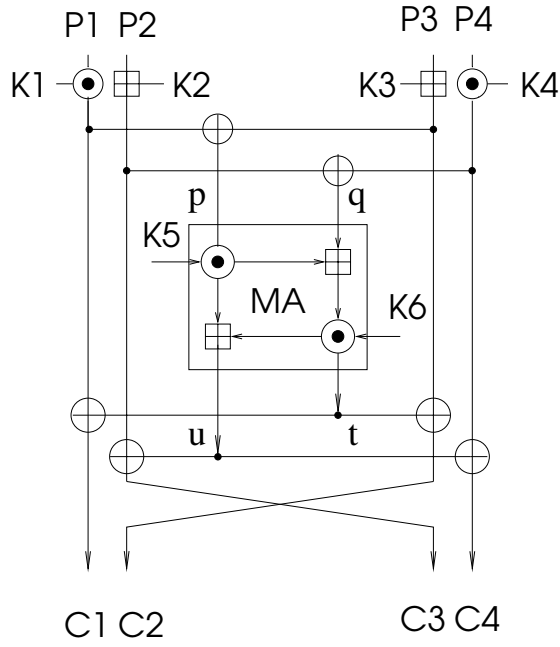

Fig. 1. One Round of IDEA

differential and linear attacks become greatly expensive after a few rounds. But we have observed that the word (16 bit) structure of the cipher result some distribution properties. To observe these, we begin with analysing the MA-box.

In [11], Lai claims that MA-box has complete diffusion: each output subblock depends on every input subblock. But this dependency is exact in the following sense:

Lemma 1. *Let p and q be the inputs of the MA-box respectively. If p is fixed and q is saturated, then both of the outputs t and u of the MA-box are also saturated.*

Proof. If p is fixed and q changes over every value, then $p \odot K5$ is fixed and $t = ((p \odot K5) \boxplus q) \odot K6$ changes over every element in 16 bits. Therefore $u = t \boxplus (p \odot K5)$ changes over every element.

Now using this lemma, we have the following result on 1 round distribution of the cipher.

Corollary 1. *Consider the set of plaintexts $(P1, P2, P3, P4)$ where we fix $P1$, $P2, P3$ and $P4$ is saturated. Encrypt this set with 1 or 1.5 round IDEA. Then in the ciphertexts $(C1, C2, C3, C4)$, each of the subblocks $C1, C2$ and $C3$ are also saturated.*

Proof. After the first transformation part, $P1 \odot K1, P2 \boxplus K2, P3 \boxplus K3$ are fixed and since $\odot$ is invertible, $P4 \odot K4$ varies on every element. Therefore the first input to the MA-box is fixed, whereas the other varies on every element. By Lemma 1, the outputs of the MA-box take every value once. XOR of a fixed value with all the possible values gives the result.

This distribution can be selected as a distinguisher of 1 or 1.5 round cipher output from a random permutation. The probability of such an event in a random permutation is:

$$\left(2^{16}!/(2^{16})^{(2^{16})}\right)^3.$$

This number is approximately $2^{-281720}$. This is a strong indicator for one round. We note that such kind of distributions were used in structural cryptanalysis [2], the square attack [5], [6] applied on the ciphers Square and Rijndael by the designers, and the saturation attacks[12], [7].

The following property of the MA-box is crucial for us in the development of our attacks.

Lemma 2. $\mathrm{lsb}(t \oplus u) = \mathrm{lsb}(p \odot K5)$.

Proof. Since $u = t \boxplus (p \odot K5)$ and for the least significant bit XOR is the same as addition, we have $\mathrm{lsb}(t \oplus u) = \mathrm{lsb}(p \odot K5)$.

This property is useful for us because 1 bit of information related with MA-box outputs can be got using only one input and one subkey subblock. Therefore in our attacks we consider only the key bits of $K5$ and the ones acting on $p = C1 \oplus C2$.

As a result of Lemma 2, we observe the following fact:

Corollary 2. *Consider the set of plaintexts obtained by fixing the first 3 subblocks, and letting the last subblock take distinct values. Apply 1 round of IDEA to this set. Then in the ciphertexts* $(C_1 1, C_1 2, C_1 3, C_1 4)$ *the variables* $C_1 1 \oplus C_1 2$ *and* $\mathrm{lsb}(C_1 2 \oplus C_1 3)$ *are constant. Therefore as the last subblock takes every* 2^{16} *value, the first input to the MA-box and the last bit of the XOR of middle subblocks are constant.*

Proof. We have that $p = C_1 1 \oplus C_1 2 = (P1 \odot K_1 1) \oplus (P3 \boxplus K_1 3)$, therefore the first input of the MA-box is fixed. By Lemma 2, this gives that the last bit of XORs of the MA-box outputs is fixed. But $C_1 2 \oplus C_1 3 = (P3 \boxplus K_1 3) \oplus t \oplus (P2 \boxplus K_1 2) \oplus u$. Since for the last bit addition is the same as XOR and the last bit of $K_1 2 \oplus K_1 3$ is constant, we have that $\mathrm{lsb}(C_1 2 \oplus C_1 3)$ is the same for all ciphertexts.

We now extend this result to the second round with the use of Lemma 2. This observation is the basis of our first attack on IDEA block cipher.

Lemma 3. *Consider the set of plaintexts obtained by fixing the first 3 subblocks, and letting the last subblock take distinct values. Apply 2 rounds of IDEA to this set. Then, in the ciphertexts* $(C_2 1, C_2 2, C_2 3, C_2 4)$ *the variable* $\mathrm{lsb}(C_2 2 \oplus C_2 3 \oplus K_5 \odot (C_2 1 \oplus C_2 2))$ *takes the same value for all ciphertexts.*

Proof. In the second round we have that $C_2 2 = (C_1 3 \boxplus K_1 3) \oplus t_2$ and $C_2 3 = (C_1 2 \boxplus K_1 2) \oplus u_2$. Then $\mathrm{lsb}(C_2 2 \oplus t_2 \oplus C_2 3 \oplus u_2) = \mathrm{lsb}((C_1 3 \boxplus K_1 3) \oplus (C_1 2 \boxplus K_1 2))$. By Lemma 2, $\mathrm{lsb}(t_2 \oplus u_2) = \mathrm{lsb}(K_2 5 \odot (C_2 1 \oplus C_2 2))$ and we have $\mathrm{lsb}(C_2 2 \oplus C_2 3 \oplus (K_2 5 \odot (C_2 1 \oplus C_2 2))) = \mathrm{lsb}(C_1 2 \oplus C_1 3 \oplus K_1 2 \oplus K_1 3)$. Since the first 3 subblocks are fixed, by Corollary 2 $\mathrm{lsb}(C_1 2 \oplus C_1 3)$ is constant, and we have the result.

Another consequence of Lemma 2 about the behaviour of 1 round cipher is the following corollary.

Corollary 3. *Fix the plaintext subblocks* $P1, P3, P4$ *and let* $P2$ *take different values. If we apply 1 round IDEA to these plaintexts, then the variable* $\mathrm{lsb}(C_1 2 \oplus C_1 3)$ *takes the same value for all plaintexts such that* $\mathrm{lsb}(P2) = 0$*, and takes the complement of that value for the plaintexts where* $\mathrm{lsb}(P2) = 1$.

Proof. Again by Lemma 2, $t \oplus u$ is constant. But $C_1 2 \oplus C_1 3 = (P3 \boxplus K3) \oplus t \oplus (P2 \boxplus K2) \oplus u$. Since for the last bit XOR is the same as addition, we have the result.

2.3 An Attack on IDEA

Lemma 3 leads to an attack for 2 rounds. We know that the correct value of $K_2 5$ satisfies the condition that $\mathrm{lsb}(C_2 2 \oplus C_2 3 \oplus K_2 5 \odot (C_2 1 \oplus C_2 2))$ is constant for all ciphertexts when we fix the first 3 subblocks of the plaintexts, where as the wrong key values will behave randomly. To eliminate the key candidates for $K_2 5$ continue the following steps.

1. Take a set of plaintexts by fixing the first three subblocks and changing the last subblock. Encrypt these plaintexts with 2 rounds of IDEA.
2. For any value of $K_2 5$, calculate the value of $\mathrm{lsb}(C_2 2 \oplus C_2 3 \oplus K_2 5 \odot (C_2 1 \oplus C_2 2))$ for all ciphertexts.
3. Eliminate the keys where the variable $\mathrm{lsb}(C_2 2 \oplus C_2 3 \oplus K_2 5 \odot (C_2 1 \oplus C_2 2))$ does not give the same value for all ciphertexts.
4. If more than 2 keys stay after elimination, take another plaintext where the first three subblocks are the same as previous ones, and the last subblock is different. Repeat step 3 for this ciphertext.

Repeat step 4 until only two key values stay. Recall from Remark in Section 2.1 that, for any value of $K_2 5$, there exists a K' which satisfies $\mathrm{lsb}(K_2 5 \odot x) = \mathrm{lsb}(K' \odot x) \oplus 1$ for all x. If $K_2 5 \notin \{0, 1\}$, this attack eliminates all keys except the correct subkey value $K_2 5$ and $2^{16} + 1 - K_2 5$. If $K_2 5 \in \{0, 1\}$, this attack eliminates all keys except 0 and 1. Therefore it is enough to search half of the key space, but at the end we will have 2 candidates for the subkey subblock $K_2 5$.

The probability that a wrong key has the property that $\mathrm{lsb}(C_2 2 \oplus C_2 3 \oplus K_2 5 \odot (C_2 1 \oplus C_2 2))$ is constant for m ciphertexts is $1/2^{m-1}$. Therefore with probability $(1 - 1/2^{m-1})^{N_k/2}$, all but 2 of the candidates for the key from a key space of N_k elements would be eliminated. Then 2 candidates for $K_2 5$ may be decided with only 23 chosen plaintexts with a probability about 0.99.

We may use Corollary 2 to decide the subkey values K_26, K_21, K_22 and the correct choice of K_25 from the two candidates. Since the first inputs of the MA-box for all the chosen ciphertexts should be equal in the first round, we can easily decide the values of K_26, K_21 and K_22 with the ciphertexts that we used to decide K_25. We may decide the remaining 64 bits of the key by exhaustive search. Therefore the total complexity of this attack is about 2^{64}.

Consider 2.5 rounds of IDEA. The least significant bit of XOR of the middle blocks is $\mathrm{lsb}(C_22 \oplus C_23 \oplus K_32 \oplus K_33)$, therefore for any ciphertext set, there are two possible sequences for the variable $\mathrm{lsb}(C_23 \oplus C_23)$ where one is the complement of the other. Also we may calculate the values of C_21 and C_22 by trying every possible value for K_31 and K_32. Then we may check $\mathrm{lsb}(C_22 \oplus C_23 \oplus K_25 \odot (C_21 \oplus C_22))$. We continue the steps above and eliminate the subkey values where $\mathrm{lsb}(C_22 \oplus C_23 \oplus K_25 \odot (C_21 \oplus C_22))$ is not constant for all chosen plaintexts. To decide the correct 48 bit subkey subblock, about 55 chosen plaintexts will be enough to eliminate all the keys with a probability near to 1. As above two keys will survive after the elimination process.

For the 3 round version of this attack, we have to search the subkey subblocks of the MA-box, K_35 and K_36 also. This requires 2^{64} trials for the key and needs about 71 chosen plaintexts. Total complexity of the elimination process is about 2^{71} decryptions.

Finally, for the attack on 3.5 rounds, we have to search for K_41, K_42, K_43 and K_44 additionally. As a result of the key schedule this attack requires 96 bit key search and needs about 103 chosen plaintexts. The work load of the elimination process is about 2^{103} decryptions.

This is a divide and conquer attack, after finding the key bits by the elimination process, we find the remaining key bits by exhaustive search. For the 2.5 round attack, after finding two candidates for the subkey subblocks K_25, K_31 and K_32, we decide the remaining subblocks by trying every possible combination of 80 bits. Then total complexity of the attack is about 2^{81}. For the 3 and 3.5 round attacks, the work done to decide the remaining key bits is negligible near the work done during the elimination process. The results are summarised in Table 1.

For this type of attacks, we have considered the following subkey subblocks:

$$K_21[97...112], K_22[113...128], K_25[58...73], K_26[74...89],$$

$$K_31[90...105], K_32[106...121], K_35[51...66], K_36[67...82],$$

$$K_41[83...98], K_42[99...114], K_43[115...2], K_44[3...18].$$

2.4 More Distributions

Using the word structure of the algorithm, we want to extend our observations to the third round. For this reason we need the following lemma and corollary.

Lemma 4. *Consider the two plaintext sets P and P' defined as: $P = \{(P1, P2, P3, P4)\}$ and $P' = \{(P1, P2', P3, P4)\}$ where $P1, P2, P2', P3$ are fixed, $P2 \neq$*

$P2'$ and $P4$ is saturated. Let E and E' denote the sets obtained by encrypting P and P' with 1 round IDEA respectively. Then if (x, y, z, s) is an element in E, there is exactly one element in E' of the form (x, y, z', s') where $z \neq z'$, $s \neq s'$ and $z \oplus s = z' \oplus s'$.

Proof. Recall from Corollary 1 that, if we fix the first 3 subblocks of the plaintext and change the last subblock over every possible 16 bit value, then after 1 round, the variables $C_1 1, C_1 2, C_1 3, t_1, u_1$ are all saturated. On the other hand if we change $P2$ to $P2'$ and repeat this procedure again, we will obtain the same set of $(t_1, u_1)'s$. Therefore, if (x, y, z, s) is a ciphertext in E, then there will be exactly one element in E' with the first two subblocks x and y, respectively. To produce the same t_1, u_1, we should have $z \oplus s = z' \oplus s'$.

Observe that although there is a great similarity between the sets E and E', it is not possible to see this using classical differential analysis. Because where the similiar ciphertext pairs, i.e. (x, y, z, s) and (x, y, z', s') occur is not certain directly from the plaintext differences. This relation can be seen if we compare a set of ciphertexts with respect to another one.

For different values of the second subblock we obtain distinct z' and s' values. As a result we have the following:

Corollary 4. *Let $P = \{(P1, P2, P3, P4)\}$ where $P1$ and $P3$ are fixed and $P2$ and $P4$ take every possible combination. Encrypt P with 1 round IDEA and denote the resulting set by E. Then the sets $M_x = \{(x, y, z, s)\}$ where x and y are fixed and z and s are saturated, form a partition for the set E.*

Corollary follows from the Lemma 4, and Corollary 1.

The outputs of 2 round IDEA seem to be randomly distributed, but the following interesting properties are again result of the word structure of the algorithm.

Corollary 5. *Let the set P be defined as in Corollary 4 and let E_2 denote the set obtained when P is encrypted with 2 round IDEA. Let r_i denote the XOR of the i-th subblocks of the ciphertexts, r_5 denote the XOR of the first outputs, and r_6 denote the XOR of the second outputs of the MA-box of all the elements of E_2. Then we have $r_1 = r_2 = r_5$ and $r_3 = r_4 = r_6$.*

Proof. When P is encrypted with 1 round IDEA, for each value of x, there is a set M_x defined as in Corollary 4. In each of these sets, the third and fourth subblocks visit each 16 bit value once. Therefore after 1 round each subblock of the ciphertext will take every value 2^{16} times. In the second round we have $C_2 1 = (C_1 1 \odot K_2 1) \oplus t_2$, $C_2 2 = (C_1 3 \boxplus K_2 3) \oplus t_2$, $C_2 3 = (C_1 2 \boxplus K_2 2) \oplus u_2$ and $C_2 4 = (C_1 4 \odot K_2 4) \oplus u_2$. The terms in paranthesis repeat equal times so their XOR is 0 and we have the result.

This is very similiar to the result used in the square attack [5]. It is trivial that an attack can also be developed using this property, but we skip this as it neither decreases the number of plaintexts nor the total complexity. Instead, we use this property as a part of our second attack on 3 rounds.

Corollary 6. *If we fix the first and third subblocks of the plaintexts and range the second and fourth ones over all 2^{32} values, then in the second round we have that the first input of the MA-box takes is 2^{16}-saturated.*

Proof. By Lemma 4, the ciphertext set after 1 round may be decomposed into sets $M_x = (x, y, z, s)$ where x and y are fixed, and z and s change over every possible element. For a ciphertext (x, y, z, t) in M_x, the first input to the MA-box in the second round is $p = (x \odot K_2 1) \oplus (z \boxplus K_2 3)$. Since x is fixed and z changes over every element, for one set, p visits every element exactly once. For every different value of x, we have such a set, therefore we have the result.

This distribution may also be used as a distinguisher of the cipher from a random permutation. The probability of such an event in a random permutation is:

$$\frac{2^{32}!}{(2^{16}!)^{2^{16}}(2^{16})^{2^{32}}}.$$

This is approximately $\frac{2^{16}\sqrt{2\pi}}{\left((2^{2^{19}+2^{15}})\pi^{2^{15}}\right)} \approx 2^{-611154}$.

Our main result which is used in the attack for 4 rounds is the following:

Theorem 1. *Let $P = \{(P1, P2, P3, P4)\}$ and $P' = \{(P1', P2, P3', P4)\}$ denote the sets of plaintexts where $P1, P3, P1', P3'$ are fixed, and $P2$ and $P4$ take every possible value. Encrypt these sets with 3 rounds of IDEA. Denote the resulting sets by E_3 and E_3' respectively. Let n_0 denote the number of 0's of the variable $\mathrm{lsb}(C_3 2 \oplus C_3 3 \oplus K_3 5 \odot (C_3 1 \oplus C_3 2))$ for the set E_3. Then the number of 0's of the variable $\mathrm{lsb}(C_3 2 \oplus C_3 3 \oplus K_3 5 \odot (C_3 1 \oplus C_3 2))$ for E_3' is either n_0 or $2^{32} - n_0$.*

Proof. By Lemmma 2, $\mathrm{lsb}(C_3 2 \oplus C_3 3 \oplus K_3 5 \odot (C_3 1 \oplus C_3 2)) = \mathrm{lsb}(C_2 2 \oplus C_2 3 \oplus K_3 2 \oplus K_3 3)$. Since $\mathrm{lsb}(K_3 2 \oplus K_3 3)$ is constant, it is enough to consider the variable $\mathrm{lsb}(C_2 2 \oplus C_2 3)$ in the second round. Now let E and E' denote the resulting ciphertext sets when P and P' are encrypted with 1 round IDEA, respectively. By Corollary 4, E and E' can be written as a union of sets $M_x = \{(x, y, z, s)\}$ and $M_x' = \{(x, y', z', s')\}$, where x, y, y' are fixed and z, z', s, s' change over every 16 bit value. Therefore if (x, y, z, s) is an element in E, then there exists an element in E' of the form (x, y', z, s'). The variable $(C_2 2 \oplus C_2 3)$ is of the form $(K_2 3 \boxplus z) \oplus t_2 \oplus (K_2 2 \boxplus y) \oplus u_2$ for (x, y, z, s) and $(K_2 3 \boxplus z) \oplus t_2' \oplus (K_2 2 \boxplus y') \oplus u_2'$ for (x, y', z, s') for some t_2, u_2, t_2', u_2'. But since the first input of the MA-box in the second round is $p = (x \odot K_2 1) \oplus (z \boxplus K_2 3)$ for both (x, y, z, s) and (x, y', z, s'), Lemma 2 implies $\mathrm{lsb}(t_2 \oplus u_2) = \mathrm{lsb}(t_2' \oplus u_2')$. Therefore it is enough to show that $\mathrm{lsb}(y \oplus y')$ is constant for every value of y and y'. But $y = (P3 \boxplus K_1 3) \oplus t_1$ and $y' = (P3' \boxplus K_1 3) \oplus t_1'$ for some t_1, t_1'. The first subblocks are equal in the first round, so we have $x = (P1 \odot K_1 1) \oplus t_1 = (P1' \odot K_1 1) \oplus t_1'$ which implies $t_1 \oplus t_1' = (P1 \odot K_1 1) \oplus (P1' \odot K_1 1)$. Since $P3$ and $P3'$ are constant and $t_1 \oplus t_1'$ depends only on $P1$ and $P1'$, we have the result.

After observing the distributions as a result of the word structure, it is natural to ask the question what happens if we fix one subblock only, and change the remaining three subblocks over every possible element. We would like to conclude this section with the answer of this question.

Theorem 2. *Let us fix one of the subblocks* $P1$ *or* $P3$ *in the plaintexts* $(P1, P2, P3, P4)$ *and change the other three subblocks over all possible values. Encrypt these plaintexts with 3 round IDEA. Then in the ciphertexts* $(C_3 1, C_3 2, C_3 3, C_3 4)$*, the variable* $\mathrm{lsb}(C_3 2 \oplus C_3 3 \oplus K_3 5 \odot (C_3 1 \oplus C_3 2))$ *takes the value 0 and 1, exactly equal, i.e.* 2^{47} *times.*

This follows from the fact that in the proof of Theorem 1, $\mathrm{lsb}(t \oplus t')$ and $\mathrm{lsb}(P3 \oplus P3')$ take the value of 0 and 1 equal times when ranging over one of $P1$ or $P3$ and keeping the other constant.

This is also a strong distinguisher from random. The probability of having equal number of 0's and 1's in a random binary sequence of 2^{48} elements is:

$$\frac{\binom{2^{48}}{2^{47}}}{2^{2^{48}}} = \frac{2^{48}!}{\left((2^{47}!)^2 2^{2^{48}}\right)}.$$

Using Stirling's Approximation, this number is approximately 2^{-24}.

Theorem 1 and Theorem 2 both consider the same variable. Theorem 2 is a much stronger distinguisher than Theorem 1, but we prefer to use the first one in an attack since it requires less number of plaintexts.

2.5 Another Attack on IDEA

Using Theorem 1, we may develop an attack on 3 rounds of the cipher. The attack proceeds as follows.

1. Take two plaintext sets of 2^{32} elements where the first and third subbloks are different fixed values and the second and fourth subblocks change over every possible element. Encrypt these sets with 3 rounds of IDEA.
2. For every possible value of the subkey $K_3 5$, count the number of 0's and 1's of the variable $\mathrm{lsb}(C_3 2 \oplus C_3 3 \oplus K_3 5 \odot (C_3 1 \oplus C_3 2))$ for both sets.
3. Let us denote the number of 0's and 1's of the first set by n_0 and n_1 and the second set r_0 and r_1 respectively. Eliminate the keys where the sets $\{n_0, n_1\}$ and $\{r_0, r_1\}$ do not coincide.
4. If more than 2 key values stay, then change the fixed part and take another set of 2^{32} plaintexts by ranging the second and fourth subblocks over every possible element. Continue the elimination by the same way.

Consider two random binary sequences of 2^{32} elements. The probability that the number of 0's of the first sequence is equal to either the number of 0's or the number of 1's of the second sequence is:

$$\frac{\left(4 \sum_{i=0}^{i=2^{31}-1} \binom{2^{32}}{i}^2 + \binom{2^{32}}{2^{31}}^2\right)}{2^{2^{33}}}.$$

We may approximate this probability as follows. Since we are counting the number of 0's, this is a binomial distribution. We have the parameters $N = 2^{32}$ and

$p = 1/2$. Therefore this distribution has mean $\mu = 2^{31}$ and variance $\sigma^2 = 2^{30}$. Since N is large and $p = 1/2$, we may assume that the number of 0's are normally distributed for both sets. Define a new variable as the difference of the number of 0's of two sets, i.e. $X_0 = n_0 - r_0$. Then X_0 is normally distributed with mean $\mu = 0$ and variance $\sigma^2 = 2^{31}$. Now for any value of ϵ we may calculate the probability that X_0 lies between $-\epsilon$ and ϵ.

$$P(-\epsilon \leq X_0 \leq \epsilon).$$

For instance for $\epsilon = 512$ this probability is about 2^{-25}. Since $P(n_0 = c \text{ or } n_0 = 2^{32} - c) = 2P(n_0 = c)$, our probability is 2 times this probability. Therefore only 2 plaintext sets will be enough to decide the two candidates of $K_3 5$. By Remark in Section 2.1, there is a K' value which satisfies $\text{lsb}(K35 \odot x) = \text{lsb}(K' \odot x) \oplus 1$ and as in the previous one, this attack eliminates all but 2 keys. This attack requires about 2^{33} chosen plaintexts and 2^{16} key search, and the work load in the elimination process is about 2^{49}. We may use Corollary 5 or Corollary 6 to decide the subkey blocks $K_3 6, K_3 1, K_3 2$ and the correct choice of $K_3 5$. Using these subkey values, we decrypt the ciphertexts and find the values of $C_2 1$ and $C_2 2$ to check if the conditions of the corollaries are satisfied. To decide the correct combination of $K_3 3$ and $K_3 4$ we may use Corollary 5 similarly with the values of $C_2 3$ and $C_2 4$. After finding these subkey subblocks, the remaining bits can be found by exhaustive search. Therefore the total work done in this attack is about 2^{49} decryptions for 2^{33} plaintexts, 2^{82}.

To extend the attack on 3.5 rounds, we shall search for the keys $K_3 5$, $K_4 1$ and $K_4 2$ since $K_4 1$ and $K_4 2$ are the subkey subblocks which affect the first two subblocks in the transformation part of the fourth round. Therefore we have to do 2^{48} trials for key search, and on the average $48/24 = 2$ comparisons will be enough to find 2 candidates for the correct key combination. About 3 plaintexts sets are required for the elimination process. The remaining 80 bits can be found by trying every possible combination. The total complexity of this attack is about 2^{82}.

We continue this way. For a 4 round attack we have to search for $K_4 5$ and $K_4 6$ also. This brings extra 32 bits search. Therefore totally we will search for 80 bits of the key. About 4 sets will be enough to eliminate all the keys in this case. The remaining subkey values are found by an exhaustive search.

For a 4.5 round attack, we have to search $K_5 1, K_5 2, K_5 3$ and $K_5 4$ additionally. But as a result of the key schedule, some of the bits we are searching are common and indeed we have to search for 114 bits. Again about 4 sets will be enough to find out the correct 114 bit combination. It is trivial that the attack for 4.5 rounds is slower than exhaustive search, but it is interesting to see that the number of required plaintexts do not change significantly as the number of rounds increase. This is a result of the strong distinguishing properties of the distributions.

The subkey subblocks we have used for this version of the attack are the following:

$$K_3 1[90...105], K_3 2[106...121], K_3 3[122...9], K_3 4[10...25], K_3 5[51...66],$$

$$K_3 6[67...82], K_4 1[83...98], K_4 2[99...114], K_4 5[19...34], K_4 6[35...50],$$

$$K_5 1[76...91], K_5 2[92...107], K_5 3[108...123], K_5 4[124...11].$$

3 Conclusion

We have observed some interesting distribution properties of the IDEA block cipher reduced to 1, 2 and 3 rounds as a result of the word structure of the algorithm. With the use of these properties, we have developed two chosen plaintext attacks. Up to 4 rounds, we are able to decrease the number of required plaintexts for an attack, but our total time complexities are not smaller than the known attacks. We consider that the distribution properties bring a different view to analyse the cipher, and they can be useful in the future attacks with more rounds. As an open question we remark that, if any distinguishing property related with the distribution of the last bits of the middle blocks of ciphertexts in the third round is found, this will immediately give rise to attacks for 4.5 or more rounds using the ideas in this paper.

4 Acknowledgements

We are very thankful to Ali Aydın Selçuk for his valuable comments and suggestions. We would also like to thank to Erkan Türe and Ali Doğanaksoy for their help on calculating and approximating the probabilities. We also thank to reviewers of SAC 2002 for pointing out integral cryptanalysis and saturation attacks.

References

1. E. Biham, A. Biryukov, A. Shamir, *Miss in the Middle Attacks on IDEA and Khufu*, LNCS 1636, Proceedings of Fast Software Encryption - 6th International Workshop, FSE' 99, pp. 124-138, Springer-Verlag, 1999.
2. A. Biryukov, A. Shamir, *Structural Cryptanalysis of SASAS,* LNCS 2045, Advances in Cryptology - Proceedings of EUROCRYPT'2001, pp. 394-405, Springer-Verlag, 2001.
3. J. Borst, L. R. Knudsen, V. Rijmen, *Two Attacks on Reduced IDEA (extended abstract),* LNCS 1223, Advances in Cryptology - Proceedings of EUROCRYPT'97, pp. 1-13, Springer-Verlag, 1997.
4. J. Daemen, R. Govaerts, J. Vandewalle, *Cryptanalysis of 2.5 round of IDEA (extended abstract),* Technical Report ESAC-COSIC Technical Report 93/1, Department Of Electrical Engineering, Katholieke Universiteit Leuven, March 1993.

5. J. Daemen, L. Knudsen and V. Rijmen, *The Block Cipher SQUARE,* LNCS 1267, FSE'97, pp. 149-165, Springer-Verlag, 1997.
6. FIPS PUB 197, NIST.
7. K. Hwang, W. Lee, S. Lee, S. Lee, J. Lim, *Saturation Attacks on Reduced Round Skipjack,* FSE'2002, Pre-Proceedings.
8. L. Knudsen, D. Wagner, *Integral Cryptanalysis,* FSE'2002, Pre-Proceedings.
9. X. Lai, J. L. Massey, *A Proposal for a New Block Encryption Standard,* LNCS 473, Advances in Cryptology - Proceedings of EUROCRYPT'90, pp. 389-404, Springer-Verlag, 1991.
10. X. Lai, J. L. Massey and S. Murphy, *Markov Ciphers and Differential Cryptanalysis,* LNCS 547, Advances in Cryptology - Proceedings of EUROCRYPT'91, pp. 17-38,Springer-Verlag, 1991.
11. X. Lai, *On the Design and Security of the Block Ciphers,* ETH Series in Information Processing, Volume 1, Hartung-Gorre Verlag Konstanz, 1995.
12. S. Lucks, *The Saturation Attack - a Bait for Twofish,* LNCS 1039, FSE'2001, pp. 189-203, Springer-Verlag, 2001.
13. W. Meier, *On the Security of the IDEA Block Cipher,* LNCS 765, Advances in Cryptology - Proceedings of EUROCRYPT'93, pp. 371-385, Springer-Verlag, 1994.
14. J. Nakahara Jr., P.S.L.M. Barreto, B. Preneel, J. Vandewalle, H.Y. Kim, *SQUARE Attacks Against Reduced-Round PES and IDEA Block Ciphers,* IACR Cryptology ePrint Archive, Report 2001/068, 2001.