Turk J Elec Eng & Comp Sci, Vol.20, No.4, 2012, © TÜBİTAK
doi:10.3906/elk-1007-564

Improved square properties of IDEA

Mete AKGÜN*, **Hüseyin DEMİRCİ**, **Mahmut Şamil SAĞIROĞLU**
Pınar KAVAK
TÜBİTAK - UEKAE, POB 74 41470, Gebze, Kocaeli-TURKEY
e-mails: {makgun, huseyind, mahsam, pinar}@uekae.tubitak.gov.tr

Received: 06.07.2010

Abstract

Block cipher encryption algorithms generally process on word structures of fixed length such as 8 or 16-bits. IDEA is one of the most widely used block ciphers and operates on 16-bit words. Square analysis is a method that exploits the word structure of block ciphers. Some square distinguishers of IDEA are given in previous studies. The best attacks against IDEA use square-like techniques. In this paper, we focus on the square properties of the IDEA block cipher. We consider all fixed word combinations of the plaintext to investigate the structural behavior of the algorithm. We observe that the cipher can be distinguished from a random permutation by fixing one, two or three subblocks of the cipher for 2 and 3 rounds. We find out novel 3-round distinguishers that require 2^{16} *chosen plaintexts. Furthermore, this approach enables us to propose the first four and five round square distinguishers of IDEA.*

Key Words: *IDEA, block cipher, cryptanalysis*

1. Introduction

The International Data Encryption Algorithm (IDEA; originally known as IPES, Improved Encryption Standard), proposed in 1991 by Lai and Massey, is a 64-bit, 8.5-round block cipher with 128-bit key size. The main design concept in the design is the “mixing [of] operations from different algebraic groups.” Lai and Massey have developed the idea of Markov ciphers to evaluate the cipher against differential cryptanalysis. IDEA is one of the most popular block ciphers, for it has been widely used in several commercial cryptographic environments such as Pretty Good Privacy (PGP) and Secure Shell (SSH).

The cryptanalysis of IDEA has developed slowly. Several cryptanalytic studies were applied to break IDEA but it resists all of them [1], [2], [3], [4], [5], [6], [7], [8], [9], [10].

In [2], a 2.5-round differential attack on IDEA was introduced. In [3], the authors presented a truncated differential attack on 3.5-round IDEA and a differential linear attack on 3-round IDEA. In [7] and [11], several weak key classes for IDEA were considered. In [4], Biham et al. used impossible differential technique to sieve the key space for 3.5, 4 and 4.5 rounds. In [5], Demirci observed the square properties of the first few

*Corresponding author: TÜBİTAK - UEKAE, POB 74 41470, Gebze, Kocaeli-TURKEY rounds. Nakahara et al. used the Biryukov-Demirci relation to attack on up to four rounds of IDEA with a trade off between time and data [12]. Demirci et al. introduced a chosen-plaintext attack on 5-round IDEA that requires 2^{24} chosen plaintexts and has time complexity of 2^{126} encryptions [6]. A related-key attack on 6.5-round IDEA that requires $2^{57.8}$ chosen plaintexts encrypted under four related keys and has time complexity of $2^{88.1}$ encryptions was developed by Biham et al. [8]. In [9], Biham et al. proposed a linear attack on 5-round IDEA that uses 2^{19} known plaintexts with 2^{103} time complexity. Transforming the relation into a related key one, they applied a 7.5-round attack on IDEA with $2^{43.5}$ known plaintexts and $2^{115.1}$ time complexity. In [13], Junod introduced some new chosen plaintext or chosen ciphertext attacks against reduced round versions of IDEA.Biham et al. also introduced the first known plaintext attack on 6-round IDEA, which exploits the weak-key schedule algorithm of IDEA, a combination of square-like techniques and linear cryptanalysis [10]. In [14], Clavier et al. performed a group of fault attacks against the IDEA block cipher. One of their proposed differential fault analysis of IDEA extracts 93 key bits out of 128 bits exploiting only 10 faults. Finally, Sun and Lai presented a key dependent attack on 5.5 and 6-round IDEA that has the lowest time and data complexity compared to the previous attacks [15].

Consider a block cipher with a word structure. By a “square property” we mean to statistically distinguish the cipher from a random permutation when selected words are fixed as constants while other words vary over a set. The block cipher SQUARE was first cryptanalyzed with the help of square properties [16]. Following this study, such an attack is called a “square attack.” Saturation [17] and integral [18] cryptanalysis also exploit the word structure of the algorithm with different notations. These techniques are applied to some block ciphers, including IDEA and Advanced Encryption Standard (AES).

In this study, we propose new square properties of IDEA by considering different combinations that fix the possible word positions of the algorithm. As an example, we have observed two novel 3-round distinguishers that require 2^{16} chosen plaintexts. We have also obtained a 4-round distinguisher by applying a similar idea that requires 2^{48} chosen plaintexts when 2 subkey blocks are known. It is also possible to carry this observation to the 5-th round for some weak keys.

This paper proceeds as follows. In Section 2, we briefly describe the IDEA block cipher. We give some existing square properties of IDEA in Section 3. Our new observations on IDEA are introduced in Section 4. In Section 4.1, 4.2 and 4.3 we give 2, 3 and 4-round distinguishers of IDEA for different fixed word positions. In Section 4.4, a 5-round distinguisher for some weak keys of IDEA is presented. In Section **??**, we conclude the paper.

1.1. Notation

Throughout this paper, we use the following notation. We use the symbol $\oplus$ for the bitwise exclusive-or (XOR), $\boxplus$ for the modular addition and $\odot$ for IDEA multiplication of 16-bit words. The plaintext is denoted by (P^1, P^2, P^3, P^4) and the ciphertext is denoted by (C^1, C^2, C^3, C^4). Each separated part shows one 16-bit subblock. Subscripts denote the round numbers. For example, C_2^1 denotes the first subblock of the ciphertext after second round. The subkeys of the MA-box are denoted by K^5 and K^6. The first input of the MA-box is called as p, the second input is called as q, the first output of the MA-box is called as t and the second output is called as u. The abbreviation lsb denotes the least significant bit of a variable. $K_2^1[97...112]$ denotes that the subkey K_2^1 uses the bits from 97 to 112 of the main key, including the boundaries.

We use the $\bigoplus$ and $\sum$ for XOR and sum of the variables respectively. If these operations are made over a set P, we denote them by

$$\bigoplus_P \text{ and } \sum_P$$

respectively. Therefore, the XOR of the variables $P_k^i, C_k^i, p_k, q_k, u_k, t_k$ is denoted over the plaintext set, P by

$$\bigoplus_P {P_k}^i, \bigoplus_P {C_k}^i, \bigoplus_P p_k, \bigoplus_P q_k, \bigoplus_P u_k, \bigoplus_P t_k$$

and the sum of least significant bits of the variables P_k^i, C_k^i over the plaintext set, P by

$$\sum_P lsb(P_k^i), \sum_P lsb(C_k^i)$$

in the k^{th} round respectively.

2. The IDEA block cipher

The IDEA block cipher is a modified version of the PES block cipher [19], [20]. The main design concept is the mixing operations from different algebraic groups. IDEA is a 8.5-round block cipher encrypting 64-bit data blocks under a 128-bit key. It uses 3 different group operations on 16 bit subblocks: XOR, modular addition and IDEA multiplication. The IDEA multiplication is defined as follows:

$$z = x \odot y$$

$$\text{if } x = 0, \text{ then } x = 2^{16}$$

$$\text{if } y = 0, \text{ then } y = 2^{16}$$

$$z = x.y \bmod (2^{16} + 1)$$

$$\text{if } z = 2^{16}, \text{ then } z = 0.$$

In [21], Lai suggested 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, which is shown in Figure , consists of two parts. The first is a transformation part where each plaintext subblock is operated with the subkey, i.e.,

$$T : (P^1, P^2, P^3, P^4) \rightarrow (P^1 \odot K^1, P^2 \boxplus K^2, P^3 \boxplus K^3, P^4 \odot K^4).$$

The second part is a multiplication-addition layer which is called the MA-box. MA-box has two 16-bit inputs $p = (P^1 \odot K^1) \oplus (P^3 \boxplus K^3)$ and $q = (P^2 \boxplus K^2) \oplus (P^4 \odot K^4)$. Using the inputs p, q and the subkeys K^5, K^6, MA-box produces two output subblocks t and u. The outputs are calculated as follows:

$$t = ((p \odot K^5) \boxplus q) \odot K^6 \text{ and } u = (p \odot K^5) \boxplus t.$$

The outputs of the MA-box are XORed with the outputs of the transformation part and the two middle subblocks are exchanged. After one round the ciphertext is of the form (C^1, C^2, C^3, C^4) where,

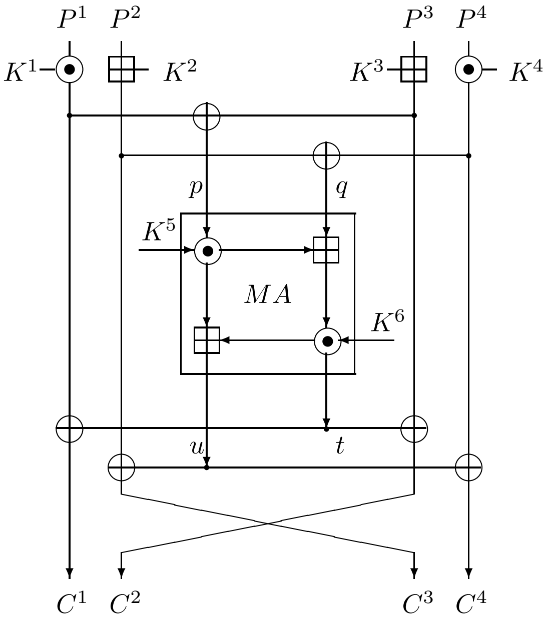

Figure. One Round of IDEA

$$
\begin{aligned}
C^1 &= (P^1 \odot K^1) \oplus t,\\
C^2 &= (P^3 \boxplus K^3) \oplus t,\\
C^3 &= (P^2 \boxplus K^2) \oplus u,\\
C^4 &= (P^4 \odot K^4) \oplus u.
\end{aligned}
$$

The encryption operation is composed of 8 full rounds and an extra transformation round. The 128 bit main key is used in 16-bit round subkeys and it is cyclically shifted 25 bits left to fill an array. Then the bits for subkeys are taken from this array. Building smaller versions of IDEA is also possible, since $2^2+1, 2^4+1$ and 2^8+1 are also prime. IDEA with block sizes 8, 16 and 32 bits can be built with subblock sizes 2, 4 and 8 respectively. This fact enables us to evaluate analytical properties of the cipher easily. The square properties mentioned in this study are verified using the smaller versions of IDEA.

3. Some distributions of the IDEA block cipher

In some previous studies, it is shown that the word structure of IDEA gives rise to some square properties. The important square distinguishers mentioned in [5] are given below.

Corollary 1 *Let $P = \{(P^1, P^2, P^3, P^4)\}$, where P^1 and P^3 are fixed and P^2 and P^4 take every possible combination, 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 $r_3 = r_4 = r_6$.*

Theorem 1 *Let $P = \{(P^1, P^2, P^3, P^4)\}$ and $P' = \{(P'^1, P2^2, P'^3, P^4)\}$ denote the sets of plaintexts where P^1, P^3, P'^1, P'^3 are fixed, and P^2 and P^4 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 $lsb(C_3^2 \oplus C_3^3 \oplus (K_3^5 \odot (C_3^1 \oplus C_3^2)))$ for the set E_3. Then, the number of 0's of the variable $lsb(C_3^2 \oplus C_3^3 \oplus (K_3^5 \odot (C_3^1 \oplus C_3^2)))$ for E_3' is either n_0 or $2^{32} - n_0$.

Theorem 2 *Fix one of the subblocks P^1 or P^3 in the plaintexts (P^1, P^2, P^3, P^4) 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 $lsb(C_3^2 \oplus C_3^3 \oplus (K_3^5 \odot (C_3^1 \oplus C_3^2)))$ takes the values 0 and 1 equal 2^{47} times.*

These results investigate square properties of IDEA for 2 and 3 rounds. At this point a natural question arises: What are the square properties when different combinations are considered for the position of the fixed words? We give the answer of this question in the following sections. We search for similar square properties by going through all combinations produced by fixing the subblocks at different positions. We observe that the previously mentioned properties are not complete. For instance, the variable n_0 is an even number for some combinations.

4. New distributions of the IDEA block cipher

4.1. Two round distinguishers

In this section, we present square distinguishers of 2-round IDEA that are found by fixing one, two and three of the subblocks. Hence, we are able to see that previous square distinguishers are incomplete.

The following theorem investigates the effect of fixing three subblocks and changing the remaining subblock over all possible values.

Theorem 3 *Let $P = \{(P^1, P^2, P^3, P^4)\}$ denote the set of plaintexts that satisfies the following cases:*

a) P^1, P^2, P^3 are fixed and P^4 takes every possible value.

b) P^1, P^2, P^4 are fixed and P^3 takes every possible value.

c) P^1, P^3, P^4 are fixed and P^2 takes every possible value.

d) P^2, P^3, P^4 are fixed and P^1 takes every possible value.

Encrypt this set with 2 rounds of IDEA. Then, for these cases, respectively, we have:

a)

$$\bigoplus C_2^1 = \bigoplus C_2^2 = \bigoplus t_2 \text{ and } \bigoplus C_2^3 = \bigoplus u_2 \text{ and } \bigoplus p_2 = 0 \text{ and}$$

$$\sum lsb(C_2^2 \oplus C_2^3 \oplus (K_2^5 \odot (C_2^1 \oplus C_2^2))) \text{ is either 0 or } 2^{16}.$$

b)

$$\bigoplus C_2^1 = \bigoplus t_2 \text{ and } \sum lsb(C_2^2 \oplus C_2^3 \oplus (K_2^5 \odot (C_2^1 \oplus C_2^2))) \text{ is even.}$$

c)

$$\bigoplus C_2^1 = \bigoplus t_2\ ,\ \bigoplus C_2^3 = \bigoplus C_2^4\ ,\ \bigoplus q_2 = 0 \text{ and}$$

$$\sum lsb(C_2^2 \oplus C_2^3 \oplus (K_2^5 \odot (C_2^1 \oplus C_2^2))) = 2^{15}.$$

d)

$$\bigoplus C_2^3 = \bigoplus u_2 \text{ and } \sum lsb(C_2^2 \oplus C_2^3 \oplus (K_2^5 \odot (C_2^1 \oplus C_2^2))) = 2^{15}.$$

Remark 1 *Let* $n_1 = \sum lsb(C_2^2 \oplus C_2^3 \oplus (K_2^5 \odot (C_2^1 \oplus C_2^2)))$. *Then,* n_1 *either remains the same or becomes* $2^{16} - n_1$ *according to changes of the least significant bits of the fixed words.*

Proof

a) Since P^1 and P^3 are fixed, $p_1 = (P^1 \odot K_1^1) \oplus (P^3 \boxplus K_1^3)$ is also fixed and $q_1 = (P^2 \boxplus K_1^2) \oplus (P^4 \odot K_1^4)$ takes every possible value once. Therefore, t_1 and u_1 take every possible value once. Then,

$$C_1^1 = (P^1 \odot K_1^1) \oplus t_1,\ \ C_1^2 = (P^3 \boxplus K_1^3) \oplus t_1 \text{ and } C_1^3 = (P^2 \boxplus K_1^2) \oplus u_1$$

take every possible value once in the first round. For the second round, $P_2^1 \odot K_2^1$, $P_2^2 \boxplus K_2^2$ and $P_2^3 \boxplus K_2^3$ take every possible value once. Then,

$$\bigoplus (P_2^1 \odot K_2^1) = 0 \text{ and } C_2^1 = (P_2^1 \odot K_2^1) \oplus t_2 \Rightarrow$$

$$\bigoplus C_2^1 = \bigoplus (P_2^1 \odot K_2^1) \oplus \bigoplus t_2 = \bigoplus t_2.$$

$$\bigoplus (P_2^3 \boxplus K_2^3) = 0 \text{ and } C_2^2 = (P_2^3 \boxplus K_2^3) \oplus t_2 \Rightarrow$$

$$\bigoplus C_2^2 = \bigoplus (P_2^3 \boxplus K_2^3) \oplus \bigoplus t_2 = \bigoplus t_2.$$

$$\bigoplus (P_2^2 \boxplus K_2^2) = 0 \text{ and } C_2^3 = (P_2^2 \boxplus K_2^2) \oplus u_2 \Rightarrow$$

$$\bigoplus C_2^3 = \bigoplus (P_2^2 \boxplus K_2^2) \oplus \bigoplus u_2 = \bigoplus u_2.$$

As a result,

$$p_2 = (P_2^1 \odot K_2^1) \oplus (P_2^3 \boxplus K_2^3) \Rightarrow \bigoplus p_2 = \bigoplus (P_2^1 \boxplus K_2^1) \oplus (P_2^3 \boxplus K_2^3) = 0.$$

Therefore, we have

$$lsb(C_2^2 \oplus C_2^3 \oplus (K_2^5 \odot (C_2^1 \oplus C_2^2))) = lsb(C_2^2 \oplus C_2^3 \oplus u_2 \oplus t_2) = lsb((C_1^2 \boxplus K_2^2) \oplus (C_1^3 \boxplus K_2^3))$$

$$= lsb(((P^3 \boxplus K_1^3) \oplus t_1) \boxplus K_2^2) \oplus (((P^2 \boxplus K_1^2) \oplus u_1) \boxplus K_2^3). \quad (1)$$

Since p_1 is fixed, $u_1 = t_1 \boxplus c_0$, where c_0 is a constant, and

$$lsb(C_2^2 \oplus C_2^3 \oplus (K_2^5 \odot (C_2^1 \oplus C_2^2)))$$

$$= lsb(((P^3 \boxplus K_1^3) \oplus t_1) \boxplus K_2^2) \oplus (((P^2 \boxplus K_1^2) \oplus (t_1 \boxplus c_0)) \boxplus K_2^3). \tag{2}$$

From the fact that $lsb(a \oplus b) = lsb(a \boxplus b)$, the lsb of the above expression is

$$lsb(C_2^2 \oplus C_2^3 \oplus (K_2^5 \odot (C_2^1 \oplus C_2^2)))$$

$$= lsb(P^3 \oplus K_1^3 \oplus t_1 \oplus K_2^2 \oplus P^2 \oplus K_1^2 \oplus t_1 \oplus c_0 \oplus K_2^3)$$

$$= lsb(P^3 \oplus P^2 \oplus c_1), \tag{3}$$

where c_1 is the new constant that depends on c_0 and the keys. Here P^3, P^2 and c_1 are all constants; then,

$$\sum_{P^4} lsb(P^3 \oplus P^2 \oplus c_1) \text{ is either 0 or } 2^{16}.$$

b) Since P_2, P_4 are fixed, q_1 is also fixed and t_1 takes every possible value once. Therefore, $(P^1 \odot K_1^1) \oplus t_1 = C_1^1$ takes every possible value once. Then,

$$C_2^1 = (C_1^1 \odot K_1^1) \oplus t_2 \Rightarrow \bigoplus C_2^1 = \bigoplus (C_1^1 \odot K_1^1) \oplus \bigoplus t_2 = \bigoplus t_2.$$

From Equation 1 and $lsb((p_1 \odot K_1^5) \boxplus t_1) = lsb(u_1)$:

$$lsb(C_2^2 \oplus C_2^3 \oplus (K_2^5 \odot (C_2^1 \oplus C_2^2)))$$

$$= lsb(((P^3 \boxplus K_1^3) \oplus t_1) \boxplus K_2^2) \oplus (((P^2 \boxplus K_1^2) \oplus ((p_1 \odot K_1^5) \boxplus t_1)) \boxplus K_2^3).$$

Since p_1 takes every possible value once, $p_1^{'} = p_1 \odot K_1^5$ takes every possible value once while P^3 changes. Consider the sum

$$\sum lsb(P^3 \oplus P^2 \oplus c_1 \oplus p_1^{'}). \tag{4}$$

Here, c_1 and P^2 are fixed and every 16 bit number occurs once for P^3 and once for p_1'. Therefore, we have

$$\sum lsb(P^3 \oplus c_2 \oplus p_1^{'}) \text{ is even.}$$

c) The first part can be proved as in Proof (a). For the second part, since P^1 and P^3 are constants and p_1 is fixed, Equation 3 is also valid:

$$lsb(C_2^2 \oplus C_2^3 \oplus (K_2^5 \odot (C_2^1 \oplus C_2^2))) = lsb(P^3 \oplus P^2 \oplus c_1).$$

Here, P^3 and c_1 are constants and P^2 takes every possible value then

$$\sum lsb(P^3 \oplus P^2 \oplus c_1) = 2^{15}.$$

d) The first part can be proved as in Proof (b). For the second part, using Expression 4 we have:

$$lsb(C_2^2 \oplus C_2^3 \oplus (K_2^5 \odot (C_2^1 \oplus C_2^2))) = lsb(P^3 \oplus c_1 \oplus P^2 \oplus p_1^{'}).$$

Here P^2, P^3 and c_1 are constants and $p_1^{'}$ takes every possible value once as P^1 changes. Therefore,

$$\sum lsb(P^3 \oplus c_1 \oplus P^2 \oplus p_1^{'}) = 2^{15}.$$

□

The following theorem is on all possible combinations that fix two subblocks.

Theorem 4 *Let $P = \{(P^1, P^2, P^3, P^4)\}$ denote the set of plaintexts that satisfies the following cases:*

a) P^1, P^2 are fixed and P^3, P^4 visit each of 2^{32} possible combinations once.

b) P^1, P^3 are fixed and P^2, P^4 visit each of 2^{32} possible combinations once.

c) P^1, P^4 are fixed and P^2, P^3 visit each of 2^{32} possible combinations once.

d) P^2, P^3 are fixed and P^1, P^4 visit each of 2^{32} possible combinations once.

e) P^2, P^4 are fixed and P^1, P^3 visit each of 2^{32} possible combinations once.

f) P^3, P^4 are fixed and P^1, P^2 visit each of 2^{32} possible combinations once.

Encrypt this set with 2 rounds of IDEA. Then, we have the following conditions:

(i) The cases (a), (b), (c), (d), (e) and (f) satisfy

$$\bigoplus C_2^1 = \bigoplus C_2^2 = \bigoplus t_2,$$

$$\bigoplus C_2^3 = \bigoplus C_2^4 = \bigoplus u_2$$

and

$$\bigoplus p_2 = \bigoplus q_2 = 0.$$

(ii) The cases c, d, e and f satisfy

$$\sum lsb(C_2^2 \oplus C_2^3 \oplus (K_2^5 \odot (C_2^1 \oplus C_2^2))) = 2^{31}.$$

Proof
(i) Consider part (a). From Theorem 3a,

$$\bigoplus_{P^4} C_2^1 = \bigoplus_{P^4} C_2^2 = \bigoplus_{P^4} t_2 \Rightarrow \bigoplus_{P^3}\bigoplus_{P^4} C_2^1 = \bigoplus_{P^3}\bigoplus_{P^4} C_2^2 = \bigoplus_{P^3}\bigoplus_{P^4} t_2$$

and

$$\bigoplus_{P^4} C_2^3 = \bigoplus_{P^4} u_2 \Rightarrow \bigoplus_{P^3}\bigoplus_{P^4} C_2^3 = \bigoplus_{P^3}\bigoplus_{P^4} u_2.$$

Since P^3 and P^4 take every possible value, (p_1, q_1) and (t_1, u_1) take every possible 2^{32} combination once and $\bigoplus_{P^3}\bigoplus_{P^4} u_1 = 0$. Then, $\bigoplus_{P^3}\bigoplus_{P^4}(P_2^4 \odot K_2^4) = 0$ since $P_2^4 = C_1^4 = (P^4 \odot K_1^4) \oplus u_1$ and P^4 takes every possible value once. Therefore,

$$\bigoplus_{P^3}\bigoplus_{P^4} C_2^4 = \bigoplus_{P^3}\bigoplus_{P^4}(P_2^4 \odot K_2^4) \oplus \bigoplus_{P^3}\bigoplus_{P^4} u_2 = \bigoplus_{P^3}\bigoplus_{P^4} u_2.$$

The fact

$$\bigoplus_{P^3}\bigoplus_{P^4} p_2 = 0$$

again follows from Theorem 3a. On the other hand, since

$$C_1^2 = (P^3 \boxplus K_1^3) \oplus t_1$$

visits every value 2^{16} times as P^3 and P^4 changes, we get

$$\bigoplus_{P^3}\bigoplus_{P^4}(P_2^2 \boxplus K_2^2) = 0.$$

This gives

$$\bigoplus_{P^3}\bigoplus_{P^4} q_2 = \bigoplus_{P^3}\bigoplus_{P^4}(P_2^2 \boxplus K_2^2) \oplus \bigoplus_{P^3}\bigoplus_{P^4}(P_2^4 \odot K_2^4) = 0.$$

The parts (b), (c), (d), (e) and (f) can be shown similarly.

(ii) As an example, consider part (c). From the result of Theorem 3c,

$$\sum_{P^2} lsb(P^3 \oplus P^2 \oplus c_1) = 2^{15} \Rightarrow \sum_{P^3}\sum_{P^2} lsb(P^3 \oplus P^2 \oplus c_1) = 2^{16}.2^{15} = 2^{31}.$$

The parts (d), (e) and (f) can be shown similarly. □

Finally, the following theorem investigates the effect of fixing one subblock.

Theorem 5 *Let $P = \{(P^1, P^2, P^3, P^4)\}$ denote the set of plaintexts where one of the subblocks P^i is fixed and the remaining three subblocks visit each of 2^{48} possible combinations once. Encrypt this set with 2 rounds of IDEA. Then, we have*

(i)

$$\bigoplus C_2^1 = \bigoplus C_2^2 = \bigoplus t_2,$$

$$\bigoplus C_2^3 = \bigoplus C_2^4 = \bigoplus u_2$$

and

$$\bigoplus p_2 = \bigoplus q_2 = 0.$$

(ii)

$$\sum lsb(C_2^2 \oplus C_2^3 \oplus (K_2^5 \odot (C_2^1 \oplus C_2^2))) = 2^{47}.$$

Proof
(i) The result follows from Theorem 4.
,
(ii) Consider the case where P^1 is fixed and P^2, P^3 and P^4 vary. From Theorem 4,

$$\sum_{P^3}\sum_{P^2} lsb(P^3 \oplus P^2 \oplus c_1) = 2^{31} \Rightarrow \sum_{P^4}\sum_{P^3}\sum_{P^2} lsb(P^3 \oplus P^2 \oplus c_1) = 2^{16}.2^{31} = 2^{47}.$$

Other cases can be shown by using similar arguments.

□

4.2. Three round distinguishers

In this section, we investigate if there exists similar square properties for 3 rounds of IDEA. We observe that we could not get a property by fixing three subblocks. The following theorem summarizes the results for fixing two subblocks.

The following theorems summarize the results obtained by fixing two and one subblocks respectively.

Theorem 6 *Let $P = \{(P^1, P^2, P^3, P^4)\}$ denote the set of plaintexts that satisfies the following cases:*

a) P^1, P^2 are fixed and P^3, P^4 visit each of 2^{32} possible combinations once.

b) P^1, P^3 are fixed and P^2, P^4 visit each of 2^{32} possible combinations once.

c) P^1, P^4 are fixed and P^2, P^3 visit each of 2^{32} possible combinations once.

d) P^2, P^4 are fixed and P^1, P^3 visit each of 2^{32} possible combinations once.

Encrypt this set with 3 rounds of IDEA. Then, for these cases we have:

$$\sum lsb(C_3^2 \oplus C_3^3 \oplus (K_3^5 \odot (C_3^1 \oplus C_3^2))) \text{ is an even number.}$$

For the remaining two cases where $P^2 - P^3$ and $P^3 - P^4$ are fixed, there is no such relation.

Theorem 7 *Let $P = \{(P^1, P^2, P^3, P^4)\}$ denote the set of plaintexts that satisfies the following cases:*

a) P^1 is fixed and P^2, P^3, P^4 visit each of 2^{48} possible combinations once.

b) P^2 is fixed and P^1, P^3, P^4 visit each of 2^{48} possible combinations once.

c) P^3 is fixed and P^1, P^2, P^4 visit each of 2^{48} possible combinations once.

Encrypt this set with 3 rounds of IDEA. Then, we have:

$$\sum lsb(C_3^2 \oplus C_3^3 \oplus (K_3^5 \odot (C_3^1 \oplus C_3^2))) = 2^{47}.$$

Proof

Parts (a) and (c) have been shown in [5] as mentioned in Theorem 2. Hence it is enough to show part (b).

b)

$$P_2^2 = (P^3 \boxplus K_1^3) \oplus t_1,$$

$$P_2^1 = (P^1 \odot K_1^1) \oplus t_1 = P_2^2 \oplus (P^1 \odot K_1^1) \oplus (P^3 \boxplus K_1^3)$$

and

$$P_2^3 = (P^2 \boxplus K_1^2) \oplus u_1$$

Assume that P^3 is fixed and P^1 and P^4 take every possible 2^{32} value once. Then, u_1 and t_1 take every possible 2^{32} value. Therefore P_2^2 and P_2^3 also take 2^{32} values once. In this situation, if the value of P^3 is changed then, P_2^1, P_2^2 and P_2^3 take every possible 2^{48} value. From the 2^{48} values, if we choose the values of $P_2'^1$ and $P_2'^3$ as constants then, p_2 will be fixed. For this situation, Equation 3 can be written for the second round as:

$$\sum_{P_2^2} lsb(P_2'^3 \oplus P_2^2 \oplus c_1) = 2^{15}.$$

Since P_2^2 takes every possible value, for all $P_2'^1$ and $P_2'^3$ values, we have

$$\sum_{P_2'^1} \sum_{P_2'^3} \sum_{P_2^2} lsb(P_2'^3 \oplus P_2^2 \oplus c_1) = 2^{32}.2^{15} = 2^{47}.$$

Other cases can be proved similarly. □

We would like to remark that the case where P^4 is fixed has no such a deterministic property.

The following theorem provides two 3-round distinguishers which require 2^{16} chosen plaintexts.

Theorem 8 *Let $P = \{(P^1, P^2, P^3, P^4)\}$ denote the set of plaintexts such that*

a) P^1, P^3 are fixed and P^2 and P^4 take every possible value in such a way that q_1 is fixed. Encrypt this set with 3 rounds of IDEA. Then

$$\sum_{(P^2 \boxplus K_1^2) \oplus (P^4 \odot K_1^4) = q_1} lsb(C_3^2 \oplus C_3^3 \oplus (K_3^5 \odot (C_3^1 \oplus C_3^2))) \qquad (5)$$

is an even number.

b) P^2, P^4 *are fixed and* P^1 *and* P^3 *take every possible value in such a way that* p_1 *is fixed. Encrypt this set with 3 rounds of IDEA. Then*

$$\sum_{(P^1 \odot K_1^1) \oplus (P^3 \boxplus K_1^3) = p_1} lsb(C_3^2 \oplus C_3^3 \oplus (K_3^5 \odot (C_3^1 \oplus C_3^2))) \tag{6}$$

is an even number.

4.3. A four round distinguisher of IDEA

If we fix the first input of the MA-box in the first round, p_1, these square properties are valid for one more round. Note that this condition can be satisfied with a guess of the subkey blocks K_1^1 and K_1^3.

Theorem 9 *Let* $P = \{(P^1, P^2, P^3, P^4)\}$ *denote the set of* 2^{48} *plaintexts such that* P^2 *and* P^4 *take every possible value and* P^1 *and* P^3 *satisfy*

$$((P^1 \odot K_1^1) \oplus (P^3 \boxplus K_1^3)) = p_1 \text{ for a given } p_1.$$

Consider the ciphertext set obtained after 4 rounds of IDEA. Then

$$\sum_{((P^1 \odot K_1^1) \oplus (P^3 \boxplus K_1^3)) = p_1} lsb((C_4^2 \oplus C_4^3 \oplus (K_4^5 \odot (C_4^1 \oplus C_4^2))) \tag{7}$$

is an even number.

This theorem can be used to distinguish the correct values of K_1^1, K_1^3 and K_4^5. If these values are correct, this count is always even whereas wrong guesses of the subkey blocks will behave randomly. To the best of our knowledge, this is the first four round distinguisher of IDEA which is based on square properties. On the other hand, we need the whole plaintext set to use this distinguisher in an attack. Therefore, such an attack would not be more advantageous than the existing attacks.

4.4. A five round distinguisher of IDEA

In this section, we carry the square based distinguishers to the fifth round under some assumptions. For some weak keys, a 5-round distinguisher is observed.

Theorem 10 *Let* $P = \{(P^1, P^2, P^3, P^4)\}$ *denote the set of* 2^{48} *plaintexts such that* P^2 *and* P^4 *take every possible value and* P^1 *and* P^3 *satisfy*

$$((P^1 \odot K_1^1) \oplus (P^3 \boxplus K_1^3)) = p_1 \text{ for a given } p_1.$$

Assume also that the subkey values K_2^5 *and* K_3^5 *take the values* 0 *or* 1. *Consider the ciphertext set obtained after 5 rounds of IDEA. Then*

$$\sum_{((P^1 \odot K_1^1) \oplus (P^3 \boxplus K_1^3)) = p_1} lsb(C_5^2 \oplus C_5^3 \oplus (K_5^5 \odot (C_5^1 \oplus C_5^2)) \tag{8}$$

is an even number.

Therefore, specific values of two subkey blocks K_2^5 and K_3^5 provide a five round square distinguisher of IDEA.

5. Conclusion

Recent analysis of IDEA block cipher depend on square properties of the algorithm. Some distinguishers were already observed as a result of the word structure. In this paper, we have analyzed the security of IDEA by investigating different square scenarios that have a potential to yield distinguishing properties. Therefore, we have considered all one, two and three fixed word combinations. We have showed that there are still square properties which were not remarked before for 2 and 3 rounds. We have discovered two 3-round distinguishers that require the information of two subkey blocks and 2^{16} chosen plaintexts. The observations on 2 and 3-round IDEA conduct us to find out the first 4-round distinguisher of IDEA. This distinguisher requires 2^{48} chosen plaintexts. Finally, we have revealed a 5-round distinguisher for a specific subset of the key space.

We have observed that careful analytical investigation of a block cipher is required during a square attack. Different fixed position combinations may lead to different properties of the cipher some of which may be more advantageous than the others. In the case of IDEA block cipher, we are able to present the first 4 and 5-round distinguishers of the cipher. However, the attacks in which these properties are directly used have more complexity than the best existing attacks. The main reason is that these properties require large amount of data. It is a new research direction to carry out the square properties to more rounds and to develop better attacks on IDEA with less complexity.

References

[1] W. Meier, On the security of the IDEA block cipher, in: EUROCRYPT, 1993, pp. 371–385.

[2] J. Daemen, R. Govaerts, J. Vandewalle, Cryptanalysis of 2.5 rounds of IDEA (extended abstract), Tech. rep. (1993).

[3] J. Borst, L. R. Knudsen, V. Rijmen, Two attacks on reduced IDEA, in: EUROCRYPT, 1997, pp. 1–13.

[4] E. Biham, A. Biryukov, A. Shamir, Miss in the middle attacks on IDEA and Khufu, in: FSE, 1999, pp. 124–138.

[5] H. Demirci, Square-like attacks on reduced rounds of IDEA, in: Selected Areas in Cryptography, 2002, pp. 147–159.

[6] H. Demirci, A. A. Selçuk, E. Türe, A new meet-in-the-middle attack on the IDEA block cipher, in: Selected Areas in Cryptography, 2003, pp. 117–129.

[7] A. Biryukov, J. N. Jr., B. Preneel, J. Vandewalle, New weak-key classes of IDEA, in: ICICS, 2002, pp. 315–326.

[8] E. Biham, O. Dunkelman, N. Keller, Related-key boomerang and rectangle attacks, in: EUROCRYPT, 2005, pp. 507–525.

[9] E. Biham, O. Dunkelman, N. Keller, New cryptanalytic results on IDEA, in: ASIACRYPT, 2006, pp. 412–427.

[10] E. Biham, O. Dunkelman, N. Keller, A new attack on 6-round IDEA, in: FSE, 2007, pp. 211–224.

[11] J. Daemen, R. Govaerts, J. Vandewalle, Weak keys for IDEA, in: CRYPTO, 1993, pp. 224–231.

[12] J. J. Nakahara, B. Preneel, J. Vandewalle, The Biryukov-Demirci attack on reduced-round versions of IDEA and MESH ciphers, in: ACISP, 2004, pp. 98–109.

[13] P. Junod, New attacks against reduced-round versions of IDEA, in: FSE, 2005, pp. 384–397.

[14] C. Clavier, B. Gierlichs, I. Verbauwhede, Fault analysis study of IDEA, in: CT-RSA, 2008, pp. 274–287.

[15] X. Sun, X. Lai, The key-dependent attack on block ciphers, in: ASIACRYPT, 2009, pp. 19–36.

[16] J. Daemen, L. R. Knudsen, V. Rijmen, The block cipher Square, in: FSE, 1997, pp. 149–165.

[17] S. Lucks, The saturation attack - a bait for Twofish, in: FSE, 2001, pp. 1–15.

[18] L. R. Knudsen, D. Wagner, Integral cryptanalysis, in: FSE, 2002, pp. 112–127.

[19] X. Lai, J. L. Massey, A proposal for a new block encryption standard, in: EUROCRYPT, 1990, pp. 389–404.

[20] X. Lai, J. L. Massey, S. Murphy, Markov ciphers and differential cryptanalysis, in: EUROCRYPT, 1991, pp. 17–38.

[21] X. Lai, On the design and security of the block ciphers, in: ETH Series in Information Processing, Vol. 1, 1995.