New Weak-Key Classes of IDEA

Alex Biryukov*, Jorge Nakahara Jr**, Bart Preneel, Joos Vandewalle

Katholieke Universiteit Leuven, Dept. ESAT/SCD-COSIC, Belgium
{alex.biryukov,jorge.nakahara,bart.preneel,joos.vandewalle}@esat.kuleuven.ac.be

Abstract. This paper presents a large collection of new weak-key classes for the IDEA cipher[1]. The classes presented in this paper contain $2^{53}-2^{64}$ weak keys (as compared with 2^{51} differential weak keys presented by Daemen at CRYPTO'93 and 2^{63} differential-linear weak-keys presented by Hawkes at EUROCRYPT'98). The novelty of our approach is in the use of boomerang distinguishers for the weak-key class membership test. We also show large weak-key classes for reduced-round versions of IDEA.

Key words: IDEA cipher, weak keys, boomerang attack, NESSIE.

1 Introduction

The International Data Encryption Algorithm (IDEA) [8–10] is 64-bit block cipher using a 128-bit secret key. IDEA consists of eight rounds followed by an output transformation. In the last decade considerable cryptanalytic effort was concentrated on IDEA [1, 3–7, 11], however, despite that effort the cryptanalytic progress was very slow. Till now the best attack [1] breaks 4.5 rounds out of 8.5 rounds and it requires the knowledge of all 2^{64} blocks of the codebook and complexity of analysis is 2^{112}. In the same decade some weak-key classes for the full 8.5-round IDEA were found. In [4] a class of 2^{51} weak keys, detectable under differential membership test, was discovered. The membership test uses two chosen plaintexts and runs in at most 2^{12} steps. In [5] a class of 2^{63} weak keys, detectable under differential-linear membership test, was found.

In this paper we describe a series of new classes of weak-keys for the full 8.5-round IDEA. Keys are termed weak in the sense that some multiplicative keys which assume values 0 or 1 turn the modular multiplication into a linear operation. These key classes are detectable with boomerang techniques developed by

* The work described in this paper has been supported in part by the Commission of the European Communities through the IST Programme under Contract IST-1999-12324 and in part by the Concerted Research Action (GOA) project Mefisto 2000/06 of the Flemish Government.

** sponsored in part by the Concerted Research Action (GOA) project Mefisto 2000/06 of the Flemish Government.

[1] The information in this document is provided as is, and no guarantee or warranty is given or implied that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

Appeared in *Information and Communications Security, 4th International Conference, ICICS 2002*, Lecture Notes in Computer Science 2513, F. Bao, R. H. Deng, and S. Qing (eds.), Springer-Verlag, pp. 315–326, 2002.

Wagner ([13]). At least one of the weak-key classes is of size 2^{64} which is larger than the best previously known weak-key class of IDEA that used a differential-linear distinguisher. However, the complexity of the membership test for this class is 2^{16} data and time which is higher than for the Hawkes' class. We also show a collection of other smaller classes all of which are new and not covered by the previously known weak-key classes. In most cases, our membership test can be used to recover up to 16 additional key bits.

Furthermore, for 5-round IDEA (from the third to the seventh round), we found a class of 2^{95} weak keys with a boomerang membership test using only a single boomerang quartet (four chosen texts). This class can be extended four times to the class of size 2^{97} keys (a fraction 2^{-31} of all keys) at the cost of 2^8 quartets for the membership test (80% of success). The best class previously known for the 5-round IDEA contained 2^{13} times less keys [5]. This result also compares favorably to the currently best attack on 4.5-round IDEA mentioned above.

This paper is organized as follows. Section 2 gives a description of the IDEA block cipher, its key schedule and round structure. Section 3 describes the boomerang attack. Section 4 presents our discoveries of the new weak-key classes of IDEA. In Section 5 we show large fractions of weak keys for IDEA reduced to 5 rounds, and compare the complexities of some previously known attacks on IDEA. Section 6 contains a brief discussion and finally section 7 concludes the paper.

2 The IDEA Block Cipher

The International Data Encryption Algorithm (IDEA) is an iterated block cipher designed by Lai, Massey and Murphy in 1991 (see [8–10]).

In November 2000, IDEA was submitted as a candidate block cipher to the NESSIE Project [12], which is a project within the Information Societies Technology (IST) Programme of the European Commission.

The IDEA cipher has a 64-bit block size, 128-bit key size, and iterates eight rounds plus an output transformation. Three algebraic operations are used in IDEA: addition in $\mathbb{Z}_{2^{16}}$ denoted by $\boxplus$, bitwise exclusive-or, denoted by $\oplus$, and multiplication in $GF(2^{16}+1)$, denoted by $\odot$, with 2^{16} interpreted as 0. Encryption and decryption in IDEA use the same framework and differ only in the key schedule.

Every full round of IDEA can be split into two halves: a key-mixing layer and a multiplication-addition (MA) structure (see Fig. 1). Let $X^i = (X_1^i, X_2^i, X_3^i, X_4^i)$ be the input block to the i-th round of IDEA, where $1 \le i \le 8$, and $X_j^i \in \mathbb{Z}_2^{16}$, for $1 \le j \le 4$. Let $Z^i = (Z_1^{(i)}, Z_2^{(i)}, Z_3^{(i)}, Z_4^{(i)}, Z_5^{(i)}, Z_6^{(i)})$, with $Z_j^{(i)} \in \mathbb{Z}_2^{16}$, for $1 \le j \le 6$ represent the six subkey words used in the i-th round of IDEA. The first operation in a round is a key-mixing half-round that combines the four 16-bit input words with the subkey words $Z_1^{(i)}$, $Z_2^{(i)}$, $Z_3^{(i)}$, $Z_4^{(i)}$, in parallel, by either modular addition or multiplication. The result is input to the MA structure (or

Appeared in *Information and Communications Security, 4th International Conference, ICICS 2002*, Lecture Notes in Computer Science 2513, F. Bao, R. H. Deng, and S. Qing (eds.), Springer-Verlag, pp. 315–326, 2002.

half-round), together with $Z_5^{(i)}$ and $Z_6^{(i)}$. At the end of the MA half-round there is a swap of the two middle words.

The output transformation (OT) is composed of a swap of the two middle input words and a key-mixing half-round.

2.1 Key Schedule of IDEA

The key schedule of IDEA processes the initial 128-bit master key into fifty-two 16-bit subkeys. Each one of the eight rounds uses six subkeys, and the output transformation (OT) uses four subkeys. The initial 128-bit key is partitioned into eight 16-bit words, and is used as the first eight subkeys. Successive sets of eight subkeys are generated by: rotating left by 25 bits the 128-bit block containing the previous eight 16-bit subkey words. Partitioning the resulting block into eight 16-bit words.

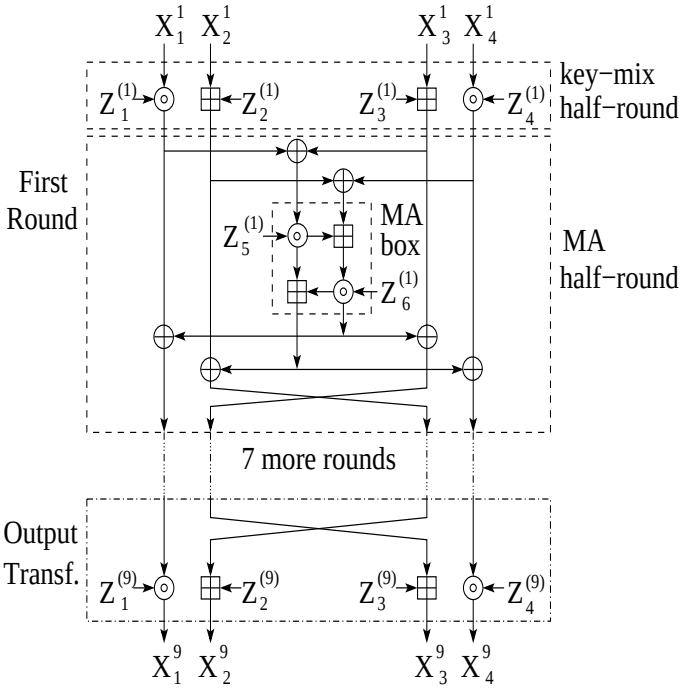

Fig. 1. Encryption scheme of IDEA block cipher.

Table 1 shows the dependency of subkey bits on the master key bits, which is indexed from 0 (MSB: most significant bit) to 127 (LSB: least significant bit). Bit numbering is taken modulo 128, that is, in a circular fashion, due to the rotation operation.

Appeared in *Information and Communications Security, 4th International Conference, ICICS 2002*, Lecture Notes in Computer Science 2513, F. Bao, R. H. Deng, and S. Qing (eds.), Springer-Verlag, pp. 315–326, 2002.

Table 1. Dependency of subkey bits on the master key bits of IDEA.

i-th round	$Z_1^{(i)}$	$Z_2^{(i)}$	$Z_3^{(i)}$	$Z_4^{(i)}$	$Z_5^{(i)}$	$Z_6^{(i)}$
1	0–15	16–31	32–47	48–63	64–79	80–95
2	96–111	112–127	25–40	41–56	57–72	73–88
3	89–104	105–120	121–8	9–24	50–65	66–81
4	82–97	98–113	114–1	2–17	18–33	34–49
5	75–90	91–106	107–122	123–10	11–26	27–42
6	43–58	59–74	100–115	116–3	4–19	20–35
7	36–51	52–67	68–83	84–99	125–12	13–28
8	29–44	45–60	61–76	77–92	93–108	109–124
OT	22–37	38–53	54–69	70–85	—	—

3 The Boomerang Attack

In this section we describe the cryptanalytic technique called the *boomerang attack* developed by Wagner [13].

Traditional differential attacks [2] are powerful methods of cryptanalysis in which the attacker considers pairs of plaintexts (P_1, P_2) with a fixed difference $\triangle = P_1 \oplus P_2$, and studies the propagation of differential patterns throughout the cipher. The aim of the attacker is to predict the resulting ciphertext difference $C_1 \oplus C_2$ with non-negligible probability. If this can be done then the cipher can be distinguished from a random permutation, and in many cases a key-recovery attack can be mounted on the cipher.

The boomerang attack is a differential-style attack in which the attacker does not try to cover the whole cipher with a single highly-probable differential pattern. Instead, the attacker tries to find several high-probability patterns that are not necessarily related to each other but together cover the whole cipher. The boomerang attack requires the ability to make both chosen-plaintext and chosen-ciphertext queries.[2]

Let's denote the encryption operation by E and its decomposition into two parts (not necessarily dividing the cipher into halves) as $E = E_1 \circ E_0$. Suppose that we start with two plaintexts P_1, P_2, such that $P_1 \oplus P_2 = \triangle$. Suppose that we have a differential pattern $\triangle \rightarrow \triangle^*$ propagating through the E_0 part of the cipher with probability p. Now consider the corresponding ciphertexts C_1, C_2 and their "shift" by the difference ∇ as follows: $C_3 = C_1 \oplus \nabla$, $C_4 = C_2 \oplus \nabla$. As ∇ we use a pattern that goes up through E_1^{-1} with high probability q, i.e. $\nabla \rightarrow \nabla^*$. We decrypt the new ciphertexts C_3, C_4 to obtain their corresponding plaintexts P_3 and P_4. If the previous three difference patterns happened as predicted, between

[2] The chosen-ciphertext queries in the boomerang attack are *adaptive* in the sense that one first obtains ciphertexts which are the results of the chosen-plaintext queries to the encryption oracle, then one performs appropriate modifications to these ciphertexts and finally feeds them to the decryption oracle.

Appeared in *Information and Communications Security, 4th International Conference, ICICS 2002*, Lecture Notes in Computer Science 2513, F. Bao, R. H. Deng, and S. Qing (eds.), Springer-Verlag, pp. 315–326, 2002.

E_0 and E_1, we obtain:

$$E_0(P_3)\oplus E_0(P_4) = E_0(P_1)\oplus E_0(P_2)\oplus E_1^{-1}(C_1)\oplus E_1^{-1}(C_3)\oplus E_1^{-1}(C_2)\oplus E_1^{-1}(C_4) =$$

$$= \triangle^* \oplus \nabla^* \oplus \nabla^* = \triangle^*.$$

Thus we can decrypt backwards through E_0 using the pattern $\triangle^* \rightarrow \triangle$. The claim is that with probability p^2q^2, the difference $P_3 \oplus P_4 = \triangle$ holds, which can be readily checked. See Figure 2 for a graphical representation of a boomerang quartet. This is an example of a **top-down boomerang**.

There are several refinements to the technique described above: we may send boomerangs from the decryption direction, that is, starting from the ciphertext and then performing adaptive plaintext queries (**bottom-up boomerangs**); we may guess the keys of the first or last rounds and send boomerangs (top-down or bottom-up) based on the key guesses. We can use the careful choice of $\triangle$ and ∇ in order to obtain additional half-round(s) in the middle (for the IDEA cipher it is the MA half-round) for free (this observation was used in [13] to attack the Khufu-16 cipher). We can use truncated differentials whenever the two faces of the boomerang produce the same difference patterns, the boomerang goes through, no matter what these difference patterns are. If the E_0 part is short enough we may use more analysis to check how differences propagate through it, without waiting for the perfect match of the difference in the second pair of plaintexts to $\triangle$.

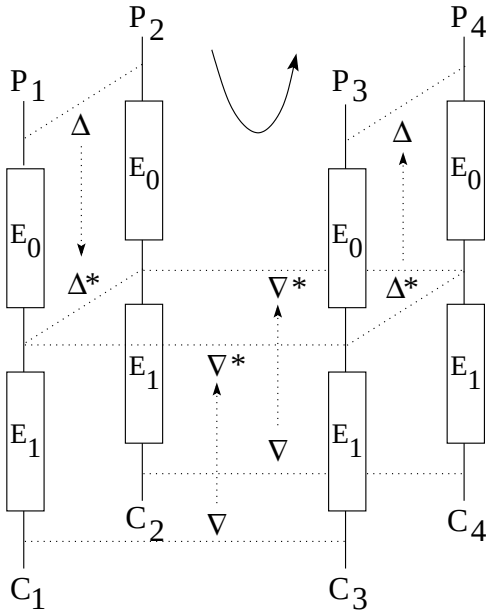

Fig. 2. A (top-down) boomerang quartet (P_1, P_2, C_3, C_4).

Appeared in *Information and Communications Security, 4th International Conference, ICICS 2002*, Lecture Notes in Computer Science 2513, F. Bao, R. H. Deng, and S. Qing (eds.), Springer-Verlag, pp. 315–326, 2002.

4 Boomerang Attack under Weak-Key Assumptions

In this section we show a variety of attacks on the full 8.5-round IDEA block cipher under some weak-key assumptions. These attacks provide new weak-key classes, larger than the ones discovered by Daemen [4], and not covered by the class discovered by Hawkes [5]. Some of these classes are the largest found so far for this cipher, but they require more effort for their membership test compared to Hawkes' differential-linear classes.

In order to build our new weak-key classes we use the boomerang-style distinguisher. The benefit of using boomerang distinguishers is two-fold: first, boomerang distinguishers pose different constraints on the key schedule than the previous differential or differential-linear distinguishers, thus we are likely to find new weak-key classes; second, we can pick unrelated patterns to cover both the E_0 (top) and E_1 (bottom) parts of the cipher and optimize the number of key-bit constraints to be minimal. As in the previous attacks we consider input xor-differences that only differ in the most significant bit (8000_x). Such differences have the advantage of propagating across the modular addition for free (i.e. with probability one). We are thus concerned only with the multiplicative keys.

4.1 Advanced Boomerang Techniques

In this section we describe our general method of search for the weak-key classes of the IDEA cipher. The method includes several refinements to the basic boomerang attack. These refinements help us to increase the key-class sizes.

We have written a program that searches through all possible plaintext/ciphertext differences in order to find the largest boomerang weak-key classes. We also considered gaps[3] of one, two and three half-rounds in the middle of the cipher, in order to increase the sizes of the key-classes at the cost of higher data and time complexities of the membership test. Another relaxation was not to cover either the top-most or bottom-most key-mixing half-round, assuming that the attacker can guess the required top or bottom keys or use special structures to construct appropriate input (or output) difference after the key-mixing. In these cases, given a correct boomerang quartet, the attacker can find up to 16 bits of multiplicative subkeys of the first or last key-mixing half-rounds, in addition to the zero key bits of the weak-key class. In the following subsections we describe several examples of our weak key classes together with their membership tests. In Table 2 we summarize the findings of this paper and compare them with the previously best classes.

4.2 A Weak-Key Class of Size 2^{53}

Consider a boomerang distinguisher which consists of two differentials: one with plaintext xor-difference $\triangle = (8000_x, 0000_x, 0000_x, 0000_x)$ that causes the xor-difference $\triangle^* = (8000_x, 8000_x, 0000_x, 8000_x)$, after 2.5 (encryption) rounds with

[3] Half-rounds with no constraints on the key bits.

Appeared in *Information and Communications Security, 4th International Conference, ICICS 2002*, Lecture Notes in Computer Science 2513, F. Bao, R. H. Deng, and S. Qing (eds.), Springer-Verlag, pp. 315–326, 2002.

probability one, provided that the 64 key bits 0–23, 64–103 are zero. The other differential has ciphertext xor-difference $\nabla = (0000_x, 8000_x, 8000_x, 0000_x)$, and causes the xor-difference $\nabla^* = (0000_x, 8000_x, 0000_x, 8000_x)$ after 5.5 (decryption) rounds with probability one, provided the 63 key bits numbered 0–25, 77–107, 123–127 are zero. These two differentials together require that the 75 key bits numbered 0–25, 64–107, 123–127 be zero. One MA half-round, with subkeys $Z_5^{(3)}$ and $Z_6^{(3)}$, is not included in the boomerang. However, due to the proper choice of the differences coming from top and the bottom end of the cipher, we gain this MA half-round for free (a similar trick was used by Wagner in his attack on Khufu [13]). This boomerang can be used to identify a weak-key class of size $2^{128-75} = 2^{53}$ using a single quartet: two chosen-plaintext and two chosen-ciphertext queries.

4.3 A Weak-Key Class of Size 2^{56}

Consider a boomerang distinguisher which consists of two differentials: one with plaintext xor-difference $\triangle = (8000_x, 0000_x, 8000_x, 0000_x)$ that causes the xor-difference $\triangle^* = (8000_x, 8000_x, 0000_x, 0000_x)$ after 1.5 (encryption) rounds with probability one, provided that the 30 key bits 0–14, 96–110 are zero. The other differential has xor-difference $\nabla = (0000_x, 8000_x, 0000_x, 8000_x)$ at the end of the 8th round (without including the last key mix half-round) and causes the xor-difference $\nabla^* = (0000_x, 8000_x, 0000_x, 8000_x)$ at the beginning of the third round with probability one, provided the 69 key bits 0–32, 84–107, 116–127 are zero. These two differentials jointly require that the 72 key bits 0–32, 84–110, 116–127 be zero. One MA half-round is not included in the boomerang. If the fourteen MSBs of $Z_4^{(9)}$ can be guessed, then this **bottom-up boomerang** can be used to identify a weak-key class of size $2^{128-72} = 2^{56}$. However we can do better than guessing the last subkey. Instead, we will prepare two pools of ciphertexts which have appropriate differences in all words and random 2^9 values in the ciphertext word after the unknown key $Z_4^{(9)}$. The pools contain 2^{18} pairs and we assume that due to the birthday paradox we'll have several pairs that have the required difference 8000_x after the decryption by the unknown key. If this event happens, then the boomerang runs for the rest of the rounds with probability one (the MA half-round in the middle is bypassed for free). In order to detect this class we thus need 2^{11} queries.

4.4 The Largest Weak Key Class of Size 2^{64}

In this section we describe the largest weak-key class of the IDEA cipher which we discovered. It is twice larger than that described by Hawkes, although its membership test is more complex. We consider Daemen's weak-key class (see Table 2) in which we do not restrict the subkey $Z_4^{(7)}$ thus increasing the key class size from 51 to 66 bits. Restricted key bits are 0-25, 41-71, and 123-127. We use the boomerang distinguisher **bottom-up** and create special structures

Appeared in *Information and Communications Security, 4th International Conference, ICICS 2002*, Lecture Notes in Computer Science 2513, F. Bao, R. H. Deng, and S. Qing (eds.), Springer-Verlag, pp. 315–326, 2002.

of ciphertexts in order to bypass the bottom round without the need to guess the multiplicative key $Z_1^{(8)}$.

In more detail: we produce two pools of chosen ciphertexts of size 2^{14} texts each, in order to generate many pairs with the difference $\triangle = (0000_x, 8000_x, 0000_x, 8000_x)$ just above the bottom key-mixing half-round. This pattern covers the next MA half-round with probability one, since it causes zero difference in the inputs to the MA-box. We create a pool C_1 of 2^{16} chosen ciphertexts, in which the 4th word takes 2^{14} random values and the other three words are arbitrary but fixed for all the texts in the pool. The second pool C_2 is created using the difference $(0000_x, 8000_x, 0000_x)$ from the texts of the first pool in the 1st, 2nd and 3rd words respectively, and the 4th word runs through random 2^{14} values. Thus, between the pools we have 2^{28} pairs with the difference $(0000_x, 8000_x, 0000_x, *)$ and among these we have 2^{12} pairs with the required difference $(0000_x, 8000_x, 0000_x, 8000_x)$ after decrypting a single key-mixing half-round and thus with a difference $(0000_x, 8000_x, 0000_x, 8000_x)$ after decrypting the last round.

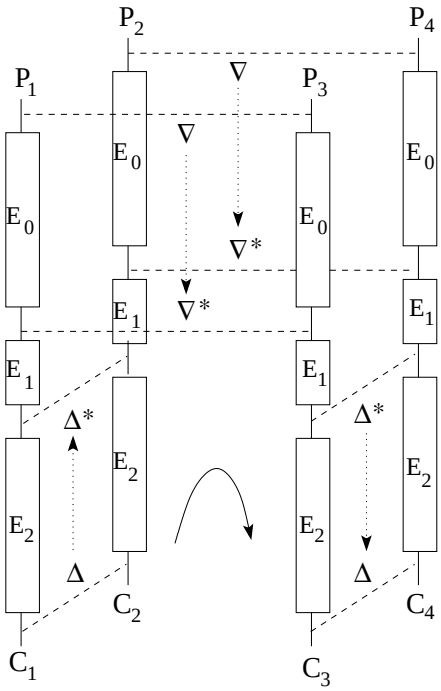

Fig. 3. Boomerang quartet for our largest key-class. E_0: 6 rounds, E_1: 1.5 rounds, E_2: 1 round.

We decrypt the pools C_1, C_2 producing plaintext pools P_1, P_2, from which we create two new pools P_3, P_4 by using the difference $\nabla = (0000_x, 8000_x,$

Appeared in *Information and Communications Security, 4th International Conference, ICICS 2002*, Lecture Notes in Computer Science 2513, F. Bao, R. H. Deng, and S. Qing (eds.), Springer-Verlag, pp. 315–326, 2002.

0000_x, 8000_x) and encrypt these pools to obtain new ciphertext pools C_3, C_4. We sort these pools by the highest 48 bits (1st, 2nd and 3rd words) and check if there is a pair with difference: $0000_x, 8000_x, 0000_x$ in the highest-order 48 bits, between the pools. If so, we proclaim that the boomerang has returned and the key belongs to the weak-key class.

In this boomerang attack we bypass the bottom key-mixing half-round using the birthday paradox, and we cover the IDEA cipher with one round from the bottom and six rounds from the top. In between we have a gap of three half-rounds: key-mixing, MA and another key-mixing which our differences have to bridge in order for the boomerang to work. Some of the multiplicative keys in the gap are already restricted by our key class ($Z_6^{(7)}$ has the 13 most significant bits set to 0, and $Z_5^{(7)} = 0$), which helps for the boomerang differences to bridge the gap. Although the key mask for this class has 66 bits we observed that boomerangs returned for about 25% of all such keys, for 2^{14} quartets, which reduces the key class to 64 bits.[4] Data complexity of this membership test is 2^{14} quartets or 2^{16} texts. Figure 3 shows one quartet used for our largest key class.

5 Attacks on Round-Reduced IDEA

In this section we show boomerang attacks on round-reduced IDEA under weak-key assumptions.

5.1 Attack on 5-round IDEA

Using the program described above we discovered that 5-round IDEA (from the 3rd to the 7th round) can be attacked for a fraction of 2^{-33} keys (2^{95} weak keys) with just one quartet (4 text queries). The plaintext difference to the boomerang is $\triangle = (0000_x, 8000_x, 0000_x, 8000_x)$ and the ciphertext difference after 5 rounds is $\nabla = (8000_x, 8000_x, 0000_x, 8000_x)$. The key-mixing at the 10th half-round is covered for free. There is another smaller class with 2^{92} weak keys (fraction of 2^{-36}) also from the 3rd to the 7th round which can be detected with a single boomerang quartet. Compare these results to the class of size 2^{84} (fraction 2^{-44}) found previously by Hawkes. Recall also that the best attack on IDEA [1] covers only 4.5 half-rounds, uses all the 2^{64} blocks of the codebook and has 2^{112} complexity.

Furthermore there is a larger class of size 2^{97} (fraction 2^{-31}) which requires 2^8 quartets for 80% success probability of the boomerang attack. The increased data requirements are due to the two half-round gap in the middle of the boomerang. Also note that this class includes the previous class of size 2^{95}.

In Table 2 we summarize our results for round-reduced versions of IDEA (from four to six rounds) and compare them to the best-previously known results [5]. In this table "Flow" indicates the top-down (↓), or bottom-up (↑)

[4] By increasing the number of quartets more keys would be covered but the amount of additional data required are larger than the gain in the key bits.

Appeared in *Information and Communications Security, 4th International Conference, ICICS 2002*, Lecture Notes in Computer Science 2513, F. Bao, R. H. Deng, and S. Qing (eds.), Springer-Verlag, pp. 315–326, 2002.

direction of the boomerang attack, $|WKC|$ denotes the size of the weak key class. The data complexity is measured in the number of texts (divide by four to get the number of quartets), time complexity is measured in the number of reduced-round encryptions.

Table 2. Summary of Weak-key Boomerang Distinguishers.

#Rnds	Hawkes $\|WKC\|$	Our $\|WKC\|$	Half-Rnds♯	Flow	Weak-Key Bit Positions	Input‡ Difference	Output‡ Difference	Complexity Data	Complexity Time
4	2^{99}	2^{104} ♭	6-13	↓	11-32	(0 0 1 *)	(1 1 0 1)	2^{14}	2^{14}
4.5	2^{97}	2^{101}	6-14	↑	0-18, 123-127	(0 1 0 1)	(0 1 0 *)	2^{18}	2^{18}
4.5	2^{97}	2^{101}	4-12	↑	2-25	(0 1 0 1)	(0 1 0 *)	2^{18}	2^{18}
5	2^{84}	2^{97}	2-11	↓	0-25, 123-127	(0 1 0 *)	(0 1 0 1)	2^{10}	2^{10}
5	2^{84}	2^{97}	4-13	↓	0-18, 116-127	(0 1 0 *)	(0 1 0 1)	2^{10}	2^{10}
5	2^{84}	2^{95}	4-13	↓	2-34	(0 1 0 1)	(1 1 0 1)	4	4
5	2^{84}	2^{97}	4-13	↓	2-32	(0 1 0 1)	(1 1 0 1)	2^{10}	2^{10}
5.5	2^{82}	2^{95}	4-14	↑	2-34	(0 1 0 1)	(* 1 0 *)	2^{18}	2^{18}
5.5	2^{82}	2^{97}	4-14	↑	2-32	(0 1 0 1)	(* 1 0 *)	2^{23}	2^{23}
6	2^{82}	2^{83}	2-13	↓	0-32, 116-127	(0 1 0 *)	(0 1 0 1)	2^{10}	2^{10}
8.5	2^{63}	2^{53}	0-16	↓	0-25,64-107,123-127	(1 0 0 0)	(0 1 1 0)	4	4
8.5	2^{63}	2^{56}	0-16	↑	0-32,84-110,116-127	(1 0 1 0)	(0 0 1 *)	2^{11}	2^{11}
8.5	2^{63}	2^{57}	0-16	↓	0-23,57-91,116-127	(0 0 1 *)	(0 1 0 1)	2^{11}	2^{11}
8.5	2^{63}	2^{57}	0-16	↓	0-32,57-91,125-127	(0 0 1 *)	(0 1 0 1)	2^{13}	2^{13}
8.5	2^{63}	2^{58}	0-16	↑	0-18,41-71,77-91,123-127	(0 1 0 1)	(0 1 0 *)	2^{25}	2^{25}
8.5	2^{63}	2^{59}	0-16	↑	0-32,84-107,116-127	(1 0 1 0)	(0 0 1 *)	2^{24}	2^{24}
8.5	2^{63}	2^{59}	0-16	↓	0-25,73-110,123-127	(0 0 0 *)	(0 1 1 0)	2^{16}	2^{16}
8.5	2^{63}	2^{60}	0-16	↓	4-25,66-110	(0 0 0 *)	(0 1 1 0)	2^{25}	2^{25}
8.5	2^{63}	2^{64}	0-16	↑	0-25,41-71,123-127	(0 1 0 1)	(0 1 0 *)	2^{16}	2^{16}

‡: the symbol '0' denotes 32-bit difference $\texttt{0000}_\texttt{x}$, and '1' denotes $\texttt{8000}_\texttt{x}$; '*' denotes arbitrary difference, used to produce $\texttt{8000}_\texttt{x}$ difference after multiplication by an unrestricted key.
♯: half-round numbering starts from 0, and ends at 16.
♭: out of a class of size 2^{106}, about 1/4 of the keys allows boomerangs to return, for the given amount of data.
†: in the upper-half of the table the last round does not unswap the middle two words; in the lower-half, there is an unswap of the two middle words, as in the full 8.5-round IDEA.

6 Discussion

We have also discovered that not only keys with certain subkeys equal to zero or one are weak (as was known before) but keys with few runs of ones are also weak and contribute to a very slow avalanche inside the IDEA cipher. While the

Appeared in *Information and Communications Security, 4th International Conference, ICICS 2002*, Lecture Notes in Computer Science 2513, F. Bao, R. H. Deng, and S. Qing (eds.), Springer-Verlag, pp. 315–326, 2002.

zero-one weak keys problem of IDEA can be corrected just by XORing a fixed constant to all the keys (one such constant may be $\texttt{0DAE}_\texttt{x}$ as suggested in [4]) the problem with the runs of ones may still remain and will require complete redesign of the IDEA key schedule.

7 Conclusions

IDEA is a strong block cipher that for more than a decade has evaded attempts of cryptanalysis. However during the same period of time large weak key classes for this cipher were found. This is due to the fact that the main non-linear part of the cipher is based on multiplication with a chosen master key and due to linearity of the key schedule. In this paper we have shown new large weak-key classes of IDEA. We have used the boomerang distinguisher as a membership test for these classes. These results strengthen the need for the redesign of the key schedule of IDEA.

References

1. Biham, E., Biryukov, A., Shamir, A.: Miss-in-the-Middle Attacks on IDEA, Khufu and Khafre, *6th Fast Software Encryption Workshop, LNCS 1636*, L.R. Knudsen, Ed., Springer-Verlag, 1999, 124–138.
2. Biham, E., Shamir, A.: Differential Cryptanalysis of the Data Encryption Standard, Springer-Verlag, 1993.
3. Borst, J., Knudsen, L.R., Rijmen, V.: Two Attacks on Reduced IDEA (extended abstract), *Advances in Cryptology, Eurocrypt'97, LNCS 1233*, W. Fumy, Ed., Springer-Verlag, 1997, 1–13.
4. Daemen, J., Govaerts, R., Vandewalle, J.: Weak Keys for IDEA, *Advances in Cryptology, Crypto'93, LNCS 773*, D.R. Stinson, Ed., Springer-Verlag, 1994, 224–231.
5. Hawkes, P.: Differential-Linear Weak Key Classes of IDEA, *Advances in Cryptology, Eurocrypt'98, LNCS 1403*, K. Nyberg, Ed., Springer-Verlag, 1998, 112–126.
6. Hawkes, P., O'Connor, L.: On Applying Linear Cryptanalysis to IDEA, *Advances in Cryptology, Asiacrypt'96, LNCS 1163*, K. Kim and T. Matsumoto, Eds., Springer-Verlag, 1996, 105–115.
7. Kelsey, J., Schneier, B., Wagner, D.: Key-Schedule Cryptanalysis of IDEA, G-DES, GOST, SAFER and Triple-DES, *Advances in Cryptology, Crypto'96, LNCS 1109*, N. Koblitz, Ed., Springer-Verlag, 1996, 237–251.
8. Lai, X.: On the Design and Security of Block Ciphers, Hartung-Gorre Verlag, Konstanz, 1992.
9. Lai, X., Massey,J.L.: A Proposal for a New Block Encryption Standard, *Advances in Cryptology, Eurocrypt'90, LNCS 473*, I.B. Damgård, Ed., Springer-Verlag, 1990, 389–404.
10. Lai, X., Massey, J.L., Murphy, S.: Markov Ciphers and Differential Cryptanalysis, *Advances in Cryptology, Eurocrypt'91, LNCS 547*, D.W. Davies, Ed., Springer-Verlag, 1991, 17–38.
11. Meier, W.: On the Security of the IDEA Block Cipher, *Advances in Cryptology, Eurocrypt'93, LNCS 765*, T. Helleseth, Ed., Springer-Verlag, 1994, 371–385.

Appeared in *Information and Communications Security, 4th International Conference, ICICS 2002*, Lecture Notes in Computer Science 2513, F. Bao, R. H. Deng, and S. Qing (eds.), Springer-Verlag, pp. 315–326, 2002.

12. *NESSIE Project – New European Schemes for Signatures, Integrity and Encryption* – available at `http://cryptonessie.org`.
13. Wagner, D.: The Boomerang Attack, *6th Fast Software Encryption Workshop, LNCS 1636*, L.R. Knudsen, Ed., Springer-Verlag, 1999, 156–170.

A Example of 5-round Boomerang

In this section we show an example printout of our program for the 5-round boomerang from the 3rd to the 7th rounds (4th – 13th half rounds). The differences are the input differences to the half-rounds, the keys printed are those that need to be restricted and the masks show the effect of these restrictions on the weak-key class size (denoted by $|WKC|$). The key-class below has size 2^{95}. The gap in half-round 10 is covered for free.

```
 HR     xor-difference     weak subkeys   key masks                  |WKC|
----|-------------------|---------------|-------------------------|-------
  4  0000 8000 0000 8000  Z_4^3 ff8000ff ffffffff ffffffff ffffffff  2^113
  5  0000 8000 0000 8000        ff8000ff ffffffff ffffffff ffffffff  2^113
  6  0000 0000 8000 8000  Z_4^4 c00000ff ffffffff ffffffff ffffffff  2^106
  7  0000 0000 8000 8000  Z_5^4 c0000000 7fffffff ffffffff ffffffff  2^97
  8  0000 8000 8000 0000        c0000000 7fffffff ffffffff ffffffff  2^97
  9  0000 8000 8000 0000  Z_5^5 c0000000 7fffffff ffffffff ffffffff  2^97
 10  ---- ---- ---- ----
 11  8000 0000 0000 0000  Z_5^6 Z_6^6 c0000000 1fffffff ffffffff ffffffff 2^95
 12  0000 8000 0000 0000        c0000000 1fffffff ffffffff ffffffff  2^95
 13  0000 8000 0000 0000  Z_6^7 c0000000 1fffffff ffffffff ffffffff  2^95
```

Below we show another example of a weak-key class. This class is of size 2^{97} and includes the class shown above. However due to the gap of 2 half-rounds (10th and 11th) the membership test for this class requires 2^8 quartets for 80% success probability.

```
 HR     xor-difference     weak subkeys   key masks                  |WKC|
----|-------------------|---------------|-------------------------|-------
  4  0000 8000 0000 8000  Z_4^3 ff8000ff ffffffff ffffffff ffffffff  2^113
  5  0000 8000 0000 8000        ff8000ff ffffffff ffffffff ffffffff  2^113
  6  0000 0000 8000 8000  Z_4^4 c00000ff ffffffff ffffffff ffffffff  2^106
  7  0000 0000 8000 8000  Z_5^4 c0000000 7fffffff ffffffff ffffffff  2^97
  8  0000 8000 8000 0000        c0000000 7fffffff ffffffff ffffffff  2^97
  9  0000 8000 8000 0000  Z_5^5 c0000000 7fffffff ffffffff ffffffff  2^97
 10  ---- ---- ---- ----
 11  ---- ---- ---- ----
 12  0000 8000 0000 0000        c0000000 7fffffff ffffffff ffffffff  2^97
 13  0000 8000 0000 0000  Z_6^7 c0000000 7fffffff ffffffff ffffffff  2^97
```

Appeared in *Information and Communications Security, 4th International Conference, ICICS 2002*, Lecture Notes in Computer Science 2513, F. Bao, R. H. Deng, and S. Qing (eds.), Springer-Verlag, pp. 315–326, 2002.