Improved Impossible Polytopic Attacks on Round-reduced DES

Jing Wang\textsuperscript{1}, Qianqiong Wu\textsuperscript{12}, Chaohui Fu\textsuperscript{1}, Guomiao Zhou\textsuperscript{1} and Ming Duan\textsuperscript{1*}

\textsuperscript{1}State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, Henan, 450000, China
\textsuperscript{2}College of electrical engineering and automation, Luoyang Institute of Science and Technology, Luoyang, Henan, 471000, China
\textsuperscript{*}Corresponding author’s e-mail: mdscience@sina.com

Abstract. In Eurocrypt 2016, Tyge Tiessen introduced the $d$-difference which considering the differences between multiple plaintexts, and proposed the impossible polytopic attack on block cipher which effectively reducing the data complexity. In this paper, we improve the impossible polytopic attacks of round-reduced DES by some ideas like truncated differentials. Given the input 3-difference of each S-box in the third round, the number of the output 3-difference is actually smaller than the theoretical upper bound, which helps us reduce the memory complexity of the attack on 5-round DES from $2^{12}$ bytes to $2^{6.9}$ bytes and increase the success rate of the attack. Using the idea of truncated differentials, the time complexity of the attack on 6-round DES is reduced from $2^{22.2}$ encryptions to $2^{25.8}$ encryptions by selecting the output 3-differences of 6 S-boxes for key recovery. We also improve the attack on 7-round DES by using more plaintexts based on our improved attack on 6-round DES.

1. Introduction

Differential cryptanalysis is a kind of chosen plaintext attack firstly proposed by Biham and Shamir in 1990\textsuperscript{1}, which is one of the most effective attack methods for iterative block cipher. Its basic idea is to study the probability propagation characteristics of the given plaintext difference value through the encryption process, and then use this non-random characteristic to distinguish the block cipher from random permutation. From then on, differential cryptanalysis has been applied and generalized in many ways, such as high order differential attacks\textsuperscript{2}, truncated differential attacks\textsuperscript{3}, integral attacks\textsuperscript{4}, impossible differential attacks\textsuperscript{5}, boomerang attacks\textsuperscript{6} and so on.

There are still many unsolved problems in the field of differential cryptanalysis. For example, although the concept of higher order difference has been put forward for a long time, it is difficult to accurately determine the probability of higher order difference without calculating Boolean function. In addition, how to accurately determine the success probability of boomerang attack is also a problem to be solved. In Eurocrypt 2016, Tyge tiessen introduced the polytopic cryptanalysis and proved that the definition and method of differential cryptanalysis can be extended to polytopic cryptanalysis, including the concept of higher order difference and impossible difference\textsuperscript{7}. The impossible polytopic cryptanalysis which can effectively reduce the data complexity was proposed and applied on the round reduced DES and AES.

Impossible polytopic cryptanalysis can be seen as another way using multiple differences different from the higher order differential cryptanalysis. It is interesting that this new method is not limited by
the algebraic degree and has rarely low data complexity. In this paper, we study the details of the attack and improve it. A summary of the improved impossible polytopic attacks on round-reduced DES can be found in Table 1. Data complexity is measured in number of required chosen plaintexts (CP). Time complexity is measured in round-reduced DES encryption equivalents. Memory complexity is measured in plaintexts (8 bytes). Ratio refers to the ratio of possible and impossible d-differences.

Table 1. Impossible polytopic attacks on round-reduced DES

| Rounds | Source | Time | Data | Memory | Ratio |
|--------|--------|------|------|--------|-------|
| 5      | Tiessen| $2^{13.7}$ | 4CP | $2^9$  | $2^{6}$ |
| This paper | $2^{13.2}$ | 4CP | $2^{3.9}$ | $\leq 2^{8.5}$ |
| 6      | Tiessen| $2^{32.2}$ | 4CP | $2^{10}$ | $2^{21.9}$ |
| Tiessen| $2^{18.8}$ | 48CP | $2^{10}$ | $2^6$ | |
| This paper | $2^{25.8}$ | 4CP | $2^9$ | $2^{34.2}$ |
| This paper | $2^{18.3}$ | 48CP | $2^{3.9}$ | $\leq 2^{8.5}$ |
| 7      | Tiessen| $2^{43}$ | 16CP | $2^{43}$ | $2^{54}$ |
| Tiessen| $2^{37.8}$ | 48CP | $2^{10}$ | $2^{71.9}$ |
| This paper | $2^{31.4}$ | 48CP | $2^9$ | $2^{54.2}$ |

2. Notions
In this section, we briefly introduce the notions of DES and impossible polytopic attack. For details, please refer to the original reference.

Data Encryption Standard (DES) was defined as the federal data processing standard (FIPS) by the National Institute of Standards and Technology of the United States in 1977[8], and authorized to be used in non secret level government communications. Subsequently, the algorithm has been widely spread in the world.

An s-polytope in $\mathbb{F}_2^n$ is an s-tuple of values in $\mathbb{F}_2^n$. A d-difference over $\mathbb{F}_2^n$ is a d-tuple of values in $\mathbb{F}_2^n$ describing the relative position of the texts of a (d+1)-polytope from one point of reference.

An impossible (d+1)-polytopic attack on n-bits block cipher can proceed in three steps as follows. Firstly, choose a number d and a d-difference such that the ratio of possible to impossible (d+1)-polytopic transitions is lower than $2^{-k-1}$, where k is the number of bits in the last round key. Secondly, get the r-round encryption of d+1 plaintexts chosen such that they adhere to the input d-difference. At last, keep the last round guess key as a candidate if the obtained d-difference after the (r−1)th round is possible, otherwise discard it.

3. Impossible polytopic attacks on DES

3.1 Impossible 4-polytopic attack on 5-round DES
We start with an impossible 4-polytopic attack on 5-round DES. We split our input 3-difference into two parts, one for the left 32 state bits and one for the right 32 state bits. We denote the left 3-difference as $(\alpha, \beta, \gamma)$. For the right half we choose the 3-difference $(0, 0, 0)$. This allows us to pass the first round for free (as can be seen in Fig. 1).
The number of possible 3-differences after the second round depends now on our choice of $\alpha$, $\beta$ and $\gamma$. To keep this number low, clearly it is good to choose the differences to activate as few S-boxes as possible. We tried to activate one S-box and the results are listed in Table 2.

Table 2. The number of possible 3-differences after round 2

| $(\alpha, \beta, \gamma)$          | S-box | The number of possible 3-differences |
|-----------------------------------|-------|-------------------------------------|
| (20000000, 40000000, 60000000)     | $S_1$ | 48                                  |
| (02000000, 04000000, 06000000)     | $S_2$ | 35                                  |
| (00200000, 00400000, 00600000)     | $S_3$ | 55                                  |
| (00020000, 00040000, 00060000)     | $S_4$ | 52                                  |
| (00002000, 00004000, 00006000)     | $S_5$ | 64                                  |
| (00000200, 00000400, 00000600)     | $S_6$ | 48                                  |
| (00000020, 00000040, 00000060)     | $S_7$ | 47                                  |
| (00000002, 00000004, 00000006)     | $S_8$ | 42                                  |

We choose the values with the fewest possible 3-differences and all of these three 3-differences only activate S-box 2 in round 2. With this choice we get 35 possible 3-differences after round 2. Note that the left 3-difference is still $(\alpha, \beta, \gamma)$ after round 2 while the 35 variations only appear in the right half.

As discussed earlier in Lemma 1, the maximal number of output d-differences for a fixed input d-difference is inherently limited by the size of the domain of the function. A consequence of this is that for any of the 35 3-differences after round 2 the possible number of output 3-differences of any S-box in round 3 is limited to $2^6$ as shown in Fig.1. To verify this, we do an experiment using algorithm 1.

Algorithm 1. Calculation of the possible 3-difference numbers of each S-box.

Input: 35 possible 3-differences, Indiff[35]; S-box N
Output: The number of possible 3-differences of S-box N after round 3, Np[35]

1. for i←0 to 34 do
2. num←0
3. for k←1 to 64 do
4. Anchor←Calc(N,k) //Go through all the anchor
5. Input←Xor(Anchor,Indiff[i])
6. Output←DESf(Input),OutAnchor3←DESf(Anchor) //Do the transformation
7. Outdiff←Xor(OutAnchor,Output) //Calculate the output 3-differences
8. if num=0 then
9. OutdiffList[0]←Extract(Outdiff,N) //Add the 4-bit differences to the list
10. num←num+1
11. else
12. for j←0 to num-1 do
13. if Match(Outdiff,OutdiffList[j])=0 then
14. Insert(Outdiff,OutdiffList) //Add the 3-differences without repetition
15. num←num+1
16. end if
17. end for
18. Np[i]←num
19. end for
20. return Np

The results of the experiment show that, due to the specificity of the input, for any of the 35 3-differences after round 2 the possible numbers of output 3-differences of all the S-boxes in round 3 are the same and in fact are less than $2^6$. The ratio of possible to impossible 3-differences of each S-box is under $2^{-7}$. For each S-box after round 3, different input 3-differences can correspond to one output 3-difference. Thus we can remove duplicates further (as can be seen in Table 3). This allows us to improved the 5-round attack on DES. In this case, the data complexity and the time complexity of the attack remain the same. Similarly, the memory requirement in the attack is storing the list of possible 3-differences for each key guess in each S-box. This should be no more than $77\cdot3\cdot2^4=2^{12.9}$ bytes. We have a total of $77+1+30+27+33+1+63+56=2^8.2$ steps each of which should be easier than one round of DES encryption. This leaves us with a time complexity of about $2^{5.8}$ 5-round DES encryption equivalents. Thus the total time complexity is $2^{5.8}+2^{13.2}=2^{19.2}$.

For the 4 output bits of each S-box, the ratio of possible to impossible 3-differences is reduced, and the success rate of the attack is higher.

### Table 3. The number and the ratio of output 3-differences of each S-box after round 3.

| S-box | S_1 | S_2 | S_3 | S_4 | S_5 | S_6 | S_7 | S_8 |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|
| The number of possible 3-differences (for each 3-difference after round 2) | 11 | 1 | 10 | 9 | 11 | 1 | 9 | 8 |
| The sum of possible 3-differences | 77 | 1 | 30 | 27 | 33 | 1 | 63 | 56 |
| Ratio | $2^{-8.5}$ | $2^{-12}$ | $2^{-8.7}$ | $2^{-8.8}$ | $2^{-8.5}$ | $2^{-12}$ | $2^{-8.8}$ | $2^{-9}$ |

### 3.2 Impossible 4-polytopic attack on 6-round DES
The 6-round attack on DES is based on the 5-round attack, using the same input 3-difference. We get 35 possible 3-differences after round 2, and after round 3 we can get possible 4-bit 3-differences of each S-box. We choose S_2, S_3, S_4, S_6, S_7, and S_8 to obtain 35\cdot10\cdot9\cdot8=2^{17.8} 24-bit output 3-differences. Thus the 24-bit 3-differences at the corresponding position of the actual output after round 3 should be included. The number should be less than $2^{17.8}$. The ratio of possible to impossible 3-differences is now $2^{13.8}/(2^{14.3}\cdot2^{17.8})=2^{5.2}\cdot2^{0.8}$. By guessing the round keys of round 5 and round 6, we can determine
the 3-difference in the same 24 output bits of round 3 now coming from the ciphertexts. For the right guess of the keys, the determined 3-difference will be possible.

This allows us to improve the 6-round attack on DES. In this case, the data complexity of the attack remains the same. The time complexity of the attack is \(2^{17.8} \cdot 2^8 \approx 2^{25.8}\) (ignoring the time complexity of the matching process). The memory complexity is the same as the 5-round attack i.e. no more than \(35 \cdot 2^9 \cdot 3 \cdot 2^{11} \approx 2^{12}\).

3.3 Impossible 4-polytopic attack on 7-round DES

The attacks for 5 and 6 rounds can be extended by one round at the cost of a higher data complexity. The extension can be made at the beginning of the cipher in the following way.

Suppose we start with a 3-difference \((\delta_1, \delta_2, \delta_3)\) in the left half and the 3-difference \((\alpha, \beta, \gamma)\) in the right half. If we knew the output 3-difference of the round function in the first round, we could choose \((\delta_1,\delta_2,\delta_3)\) accordingly to make sure that we end up at the starting position of the original attack. Thus by guessing this value and repeating the attack for each guessed value of this 3-difference we can make sure we still retrieve the key.

Since the values of \((\alpha, \beta, \gamma)\) are already chosen to give a minimal number of possible 3-difference in the round function, the time complexity only increases by 35 and 48. The data complexity increases even less. As it turns out, 12 different values for the left half of the input text are enough to generate all of the 35 or 48 3-differences. Thus the data complexity only increases to 48 chosen plaintexts.

Using Tiessen’s impossible polytopic attack, the 5-round attack can be extended to 6-round attack with a time complexity of \(35 \cdot 2^{13.7} \cdot 2^{18.8}\) and a memory complexity of \(2^{12}\) which is the same as that in 5-round attack. The 6-round attack can be extended to 7-round attack with a time complexity of \(48 \cdot 2^{32.2} \cdot 2^{37.8}\) and a memory complexity of \(2^{13}\) which is the same as that in 6-round attack.

Using the improved impossible polytopic attack, the 5-round attack can be extended to 6-round attack with a time complexity of \(35 \cdot 2^{13.2} \cdot 2^{18.3}\) and a memory complexity of \(2^{6.9}\) which is the same as that in 5-round attack. The 6-round attack can be extended to 7-round attack with a time complexity of \(48 \cdot 2^{25.8} \cdot 2^{31.4}\) and a memory complexity of \(2^{12}\) which is the same as that in 6-round attack.

4. Conclusions

In this paper we improved the impossible polytopic attack of round reduced DES by some ideas like truncated differentials. In our opinion, the next worth studying is the optimization of impossible polytopic differential attack, the mining of new polytopic differential attack patterns and the application of deep learning in this area.

References

[1] Biham E., Shamir, A. (1991) Differential cryanalysis of DES-like crytosystems. J. Cryptology 4(1), 3-72.
[2] Lai, X. (1994) Higher order derivatives and differential cryptanalysis. In: Blahut, R.E., Daniel J. Costello, J., Maurer, U., Mittelholzer, T. (eds.) Communications and Cryptography, Two Sides of One Tapestry, pp. 227–233.
[3] Knudsen, L.R. (1995) Truncated and higher order differentials. In: Preneel, B. (ed.) Fast Software Encryption, FSE ’94. Lecture Notes in Computer Science, vol. 1008, pp. 196–211.
[4] Knudsen, L.R., Wagner, D. (2002) Integral cryptanalysis. In: Daemen, J., Rijmen, V. (eds.) Fast Software Encryption, FSE 2002. Lecture Notes in Computer Science, vol. 2365, pp. 112–127.
[5] Biham E., Biryukov A., and Shamir, A. (1999) Cryptanalysis of Skipjack Reduced to 31 Rounds Using Impossible Differentials. Lecture Notes in Computer Science, vol. 1592, pp. 12–23.
[6] Wagner, D. (1999) The boomerang attack. In: Knudsen, L.R. (ed.) Fast Software Encryption, FSE ’99. Lecture Notes in Computer Science, vol. 1636, pp. 156–170.
[7] Tiessen, T. (2016) Polytopic cryptanalysis. In: Fischlin, M., Coron, J (eds.) Advances in Cryptology - EUROCRYPT 2016. Lecture Notes in Computer Science, vol. 9665, pp. 214-239.

[8] National Institute of Standards and Technology. (1977) Data Encryption Standard. Federal Information Processing Standard (FIPS), Publication 46, U.S. Department of Commerce, Washington D.C.