An Approach of Zero Correlation Linear Cryptanalysis

Mehak Khurana,
Dept of Computer Science and Engineering
The NorthCap University, Gurgaon, Haryana, India
mehakkhurana@ncuindi.edu

Meena Kumari
Dept of Computer Science and Engineering
The NorthCap University, Gurgaon, Haryana, India
meenakumari@ncuindia.edu

Abstract: Differential and Linear Cryptanalysis are two most popular techniques that have been widely used to attacks block ciphers to reveal its weakness in substitution and permutation network. Most of the block ciphers which are resistant against Differential and Linear Cryptanalysis may not be immune to their latest extensions such as Impossible Differential Cryptanalysis (IDC) and Zero Correlation Linear Cryptanalysis (ZCLC). These techniques use the statistics of distinguishing the correct key and incorrect key from probability distribution of impossible differentials and linear approximation respectively. Most of IDC and ZCLC are based on miss in middle technique and are independent of non linear components. In this paper, the method used by ZCLC is explained and the block ciphers which have been attacked by ZCLC and multidimensional ZCLC have been stated. The links between IDC and ZCLC have been also projected.

Keywords: Linear Cryptanalysis, Impossible Differential, Zero Correlation, Multidimensional Zero Correlation

I. INTRODUCTION

Most popular cryptanalytic techniques, Differential and Linear Cryptanalysis (LC) have been developed in 1990's and have become most effective against block ciphers. Many block ciphers composed of permutation, substitution, XOR and summation operation were designed to be immune against known cryptanalysis techniques DC and LC. Now a day's latest block ciphers are been developed which are resistant against DC and LC, due to which these techniques have been extended [1] as Truncated Differential, Impossible Differential Cryptanalysis (IDC), Zero Correlation Linear Cryptanalysis (ZC) which may attack these block ciphers. The common way to prove security against DC and LC is to give upper bound on the rounds of differential characteristics/Linear trails that can distinguish a round reduced cipher from round permutation. IDC and ZC are two new approaches which have not been developed to attack the well known block ciphers but also to measure the level of security. The idea behind both the approaches is to exclude wrong key candidates by using miss in middle technique.

Differential cryptanalysis uses the concept where it uses differentials with higher probabilities than those of expected for a randomly drawn permutation. Similarly, for linear cryptanalysis it based on the concept of using linear approximations whose probabilities maximally deviate from $1/2$ i.e. with highly non zero correlation value. Whereas Impossible Differential Cryptanalysis (IDC) and Zero Correlation Linear Cryptanalysis (ZCLC) are based on the concept of miss in middle technique where differentials with zero probability are targeted and exploited in IDC and in ZCLC, linear approximations with probability of exactly $1/2$ i.e. with correlation zero are targeted. They are counterpart of each other in their own domains. In rest of the paper, section II will introduce linear cryptanalysis and its method of construction. Section III comprises of the Zero correlation linear cryptanalysis technique, the basic conditions to find its distinguisher, its construction method and list of block ciphers on which it has been applied. Section IV shows the links between IDC and ZC.

II. LINEAR CRYPTANALYSIS

In 1993, Matsui proposed Linear Cryptanalysis (LC). In LC, the Linear Expression of this form is found such that expression has high or low probability of occurrence. Deviation or bias ($\in$) from probability $1/2$ is exploited [2].

$$P_{i_1} \oplus P_{i_2} \oplus \dots P_{i_j} \oplus C_{i_1} \oplus C_{i_2} \oplus .. C_{i_k} = K_{i_1} \oplus K_{i_2} .. \oplus K_{i_l}$$

P is the plaintext and C is its corresponding Ciphertext under key K respectively $C = f_K(P)$ for n bit block cipher. If this type of expression exits it states cipher is trivially weak.

The linear approximation $M_P \rightarrow M_C$ is denoted by $M_p^T\ P \oplus M_C^T\ C = 0$ has probability:

$$p_{M_P,M_C} = Pr_{P \in F_n^2}(M_p^T\ P \oplus M_C^T\ C = 0)$$

M_P, M_C are linear mask of plaintext and ciphertext respectively of n bit each. M_p^T denotes the multiplication of the transposed bit vector M_P by column bit vector of P over F_n^2. This probability p_{M_P,M_C} can be written in term of correlation C_{M_P,M_C}:

$$C_{M_P,M_C} = 2p_{M_P,M_C} - 1$$

More the probability (p_{M_P,M_C}) deviates from value $1/2$, better the attack can be applied by the cryptnalyst. The high bias probability for holding this attack for $(r-1)$ rounds becomes the distinguisher for the attack.

A. Construction Method

i. Steps for finding distinguisher and recovery are as follows:

1. Generate the Linear Approximation (LA) table for an S-Box $F_2^n \rightarrow F_2^m$ (nxm) of $2^n\ x 2^m$ entries.
   (a) Find each element of table by calculating the number of coincidence between linear relation of input and output for an input
   (b) Then for each element calculate probability p_L by dividing it by 2^n
   (c) calculate bias probability $\in = \left|p_L - \frac{1}{2}\right|$
2. Concatenate the linear approximations of S-Boxes by marking the linear trail from input of first round to input of last round of cipher where linear approximation $M_P \rightarrow M_C$ for S-Box holds true for high probability i.e. Bias probability is high.
3. Using pilling up lemma considering approximations of S-Boxes are independent of each other, calculate the expected bias probability p_D till second last round$(r-1)$ [2].

Once $(r-1)$ rounds of linear approximation with high linear probability bias for r rounds of cipher is obtained, it will help in extracting last round subkey bits (associated bits of S-Boxes in the last round that are influenced by the data bits involved in the linear approximation) [3].

ii. Steps for Key Recovery

The single path linear trail gives a straight forward way for key recovery.

1. Given N Plaintext Ciphertext (PC) pairs to find right key
2. For all PC pairs guess all possible last round (r) subkey
3. For all possible subkey guess of last round (r), the Ciphertext is partially decrypted by one round for each PC pairs
4. Count the number of times the linear approximation $M_P \rightarrow M_C$ is satisfied.
5. The key which has maximum bias is the actual key i.e. the key that differs maximum from half the number of plaintexts is the actual key.

Matsui proved the complexity of the attack in [3], that if the bias is $\in$ of the linear approximation then the number of plaintext required for attack is nearly equal to $1/\in^2$.

III. ZERO CORRELATION LINEAR CRYPTANALYSIS

In ZCLC proposed by Bogdanov and Rijmen [4], instead of exploiting linear approximation with high probability,

$$p_{M_P,M_C} = Pr_{P \in F_n^2}\left(M_p^T\ P \oplus M_C^T\ C = 0\right) = \frac{1}{2}$$

probability of exactly $1/2$ is exploited which amounts to correlation C_{M_P,M_C} is zero [4]

$$C_{M_P,M_C} = 2p_{M_P,M_C} - 1$$

$$C_{M_P,M_C} = 2 * [Pr_{P \in F_n^2}\left(M_p^T\ P \oplus M_C^T\ C = 0\right)] - 1$$

(If only one of $M_p^T,\ M_C^T$ is zero)

$$Pr_{P \in F_n^2}\left(M_p^T\ P \oplus M_C^T\ C = 0\right) = \frac{1}{2}$$

$\therefore\ M_P \rightarrow 0\ and\ 0 \rightarrow M_C$ have correlation C_{M_P,M_C}

$$C_{M_P,M_C} = 2 * \frac{1}{2} - 1$$

$$C_{M_P,M_C} = 0$$

where $M_P, M_C \neq 0$. Then

$$C_{0\to 0} = 1, C_{M_P\to 0} = C_{0\to M_C} = 0$$

i.e. B. Rijmen [4] uses linear approximations with correlation exactly zero for all possible keys for exploitation.

A. Sufficient conditions for distinguisher

Each iterative block cipher has r rounds and for a linear trail, linear approximation is concatenated for each round such that output mask of ith round equals the input mask of $i^{th}+1$ round. Correlation value for each individual round is calculated where i^{th} round has input mask u_i and output mask u_{i+1}.

For a cipher correlation contribution C_U for each linear trail U is calculated by multiplying correlation values of each individual round.

$$C_U = \prod_{i=0}^{r-1} C_{r_i}(u_i, u_{i+1})$$

For a linear hull $M_P \to M_C$ i.e. for all possible linear trails U in a cipher, correlation contributions C_U are summed to compute total correlation C.

$$C = \sum_U C_U$$

The distinguisher $M_P \to M_C$ for zero correlation linear cryptanalysis is constructed by finding zero correlation contribution in each linear trail. It is only possible if for at-least any i^{th} round in each linear trail correlation value turns out to be zero,

$$C_i(u_i, u_{i+1}) = 0$$

$\therefore$ for all trails correlation contribution C_U becomes zero

$$C_U = \prod_{i=0}^{r-1} C_{r_i}(u_i, u_{i+1}) = 0$$

thus correlation for a linear hull $M_P \to M_C$ becomes zero.

$$C = \sum_U C_U = 0$$

To describe linear approximation for XOR, Branching and permutation in the encryption or decryption round for a fiestel ciphers, we use the following conditions [5]:

Let $M_X = (M_{X_1}, M_{X_2}, .., M_{X_n})$, $M_Y = (M_{Y_1}, M_{Y_2}, .., M_{Y_n})$ and $M_Z = (M_{Z_1}, M_{Z_2}, .., M_{Z_n})$

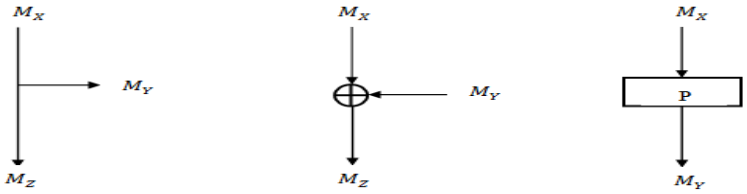

Figure 1: Branching, XOR, Permutation

- *Branching operation*: Summation of all masks (input or output masks) at a branching point is zero i.e. $M_{X_i} \oplus M_{Y_i} = 0$ and $M_{X_i} \oplus M_{Z_i} = 0$ or correlation is zero.
- *XOR operation:* All the masks at XOR operation are equal i.e. $M_{X_i} \oplus M_{Y_i} \oplus M_{Z_i} = 0$ or correlation is exactly zero.
- *Bijective function*: Either the input mask or output mask of permutation is zero or nonzero $M_Y^T = (P^T)^{-1}.M_X^T$, then the correlation is zero.

B. Construction Method

To construct the ZCLC, one adopts the miss in middle technique just like in IDC extension of differential cryptanalysis. In IDC, miss in middle technique constructs the impossible differential characteristic by two truncated differential paths (forward and backward) with probability one and which lead to a contradiction at the middle [6].

In ZCLC, any linear approximation with nonzero bias is concatenated to any linear approximations with nonzero bias in the inverse direction, where the intermediate masks states contradict with each other. To state more simply, linear approximation patterns of input mask and output masks in the intermediate rounds in both

forward and backward directions respectively are followed and inquired if no linear characteristics with nonzero correlation exists i.e. check the incoherence of events [7].

i. Steps to find linear approximation and recover is as follows

For a block cipher f_k, identify all the zero correlation linear approximations till it has compact description with a distinct structure and linear approximations with zero correlation for any key K.

1. Find input mask M_I and output mask M_J for a part of a cipher and continue encryption and decryption by using above stated 3 lemmas till some intermediate level of the design.
2. If at the intermediate level of the design, the output of the i^{th} round and input to $(i^{th}+1)$ are not equal i.e. non equivalence exists, then this linear hull $M_I \rightarrow M_J$ gives zero correlation distinguisher.
3. If more no. of rounds can be attacked, it is more efficient distinguisher.

ii. Steps for key recovery

According the zero correlation linear approximations $M_I \rightarrow M_J$ obtained for a part of cipher from the above steps, similar approach to linear cryptanalysis key recovery approach can be used here for key recovery. Let N are the number of known Plaintext-Ciphertext pairs (P-C pairs) with an adversary and s be the number of zero correlation linear approximations $M_I \rightarrow M_J$. The linear approximations $M_I \rightarrow M_J$ for a part of cipher are placed in the middle of the attacked cipher so as to attack more number of rounds. Let I and J be the partial intermediate states of the data transform at the boundaries of the linear approximations [8]. Following is the approach

1. Guess the bits of the subkeys k_1 and k_2 that are needed to compute I and J. For each guess of subkey k_1 and k_2 :

(a) Partially encrypt and partially decrypt the plaintexts P and its corresponding ciphertexts C respectively upto the boundaries I and J of the zero correlation linear approximation $M_I \rightarrow M_J$.

(b) Evaluate the correlations C_{M_I,M_J} of all linear approximations in $M_I \rightarrow M_J$ for the subkey guess using the partially encrypted I and partially decrypted values J.

(c) Check if correlation is zero or not or by counting how many times $M_I^T I \oplus M_J^T J = 0$ over N input/output pairs. **Note:** For every linear trail U in a linear hull, if at least one pair of adjacent

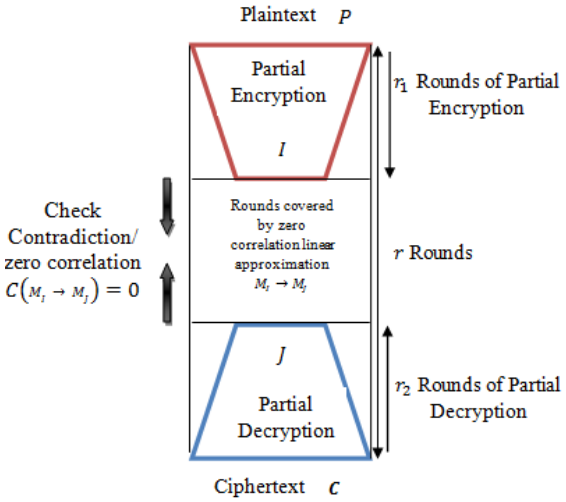

Figure 2: Zero Correlation Linear Cryptanalysis

linear selection patterns is incompatible $C_i(u_i, u_{i+1}) = 0$, the correlation for the linear hull $M_I \rightarrow M_J$ will be exactly 0, $C_{M_I,M_J} = \sum_U C_U = 0$, since $C_U = \prod_{i=0}^{r-1} C_{r_i}(u_i, u_{i+1}) = 0$

(d) If the correlation is zero, subkey bits are correctly guessed. Otherwise repeat the steps.

$\mathcal{U}$ method [4] and matrix method [11] are the two ways to search Zero Correlation Linear Hulls (ZCLH) for block ciphers.

AES [6][8], TEA [8], XTEA [8], CLEFIA-type generalized feistel networks [9][16], Skipjack [6], CAST256 [6], Simon [7], Camellia [8][9], LBLOCK [11], ARIA [8], E2 [12], KASUMI [13], HIGHT [13], are some of the block ciphers which have been attacked by zero correlation linear cryptanalysis [9][10].

C. Multidimensional Zero Correlation Linear Cryptanalysis

Bogdanov et al. in [11][12][13] proposed a multidimensional zero correlation linear distinguisher using zero correlation linear approximations. He proved that for n bits of block cipher, if there are s independent zero correlation linear approximations then the required number of plaintexts are $O(2n/\sqrt{l})$ such that $l = 2^s$ nonzero linear combinations of them have zero correlation. Multidimensional zero correlation attack on HIGHT lightweight block cipher on 27 rounds has been shown in [14].

D. Links between Impossible and Zero Correlation Cryptanalysis

IDC that follows the differential cryptanalysis mechanism is known to be useful and popular technique for attacking some block ciphers which may be immune to DC. ZCLC that follows the mechanism of linear cryptanalysis is the latest technique developed similar to IDC in technical terms but has its theoretical foundation in a different mathematical theory and has advantage over IDC because it covers more number of rounds when applied on a block cipher, or else it perform the attack in less time on the same number of rounds [15]. When IDC and ZCLC search distinguishers using matrix-method then the number of differentials and linear approximations involved in the attack are approximately same and in case of ZCLC attack, the data complexity is large as compared to IDC attack.

The other similarity involved between two is that they both uses miss in middle technique. To distinguish between a small correlation value and a correlation with exact zero value becomes difficult. This corresponds to the situation where LC usually tends to break slightly smaller number of rounds than DC, except in the case of DES where LC was more efficient and broke the cipher with more number of rounds as compared to DC.

Some of the example of block ciphers

TEA,XTEA, HIGHT, CLEFIA, AES, ARIA, Camellia, E2, MIBS, LBlock, Piccolo are the block ciphers that have been attacked by the Impossible Differential [16] [17].

IV. CONCLUSION

This paper describes the approach of zero correlation linear cryptanalysis to find distinguisher and to recover key through which designer of block ciphers can evaluate and analyse the security of their block ciphers against ZCLC. Some of the block ciphers which have been attacked by IDC and ZCLC are also mentioned. Multidimensional zero correlation and links between IDC and ZCLC have also been projected in this paper.

REFERENCES

[1] Mehak Khurana, Meena Kumari, Variants of Differential and Linear Cryptanalysis, in International Journal of Computer Applications (0975 – 8887) Volume 131 – No.18, December 2015

[2] Howard M. Heys, A Tutorial on Linear and Differential Cryptanalysis pp 1-11.

[3] E. Biham. On Matsui's Linear Cryptanalysis. In EUROCRYPT'94, volume 950 of Lecture Notes in Computer Science, pages 341–355. Springer, 1995.

[4] A. Bogdanov, V. Rijmen, Linear Hulls with Correlation Zero and Linear Cryptanalysis of Block Ciphers. Designs, Codes and Cryptography March 2014, Volume 70, Issue 3, pp. 369-383, 2014.

[5] Bogdanov, V. Rijmen: Zero Correlation Linear Cryptanalysis of Block Ciphers, IACR Eprint Archive Report 2011/123, March 2011.

[6] Andrey Bogdanov, Gregor Leander, Kaisa Nyberg, and Meiqin Wang. Integral and Multidimensional Linear Distinguishers with Correlation Zero. In ASIACRYPT, in Lect. Notes Computer Science, vol. 7658, Springer, Heidelberg, pages 244-261, 2012.

[7] C. Boura, M. Naya-Plasencia, and V. Suder. Scrutinizing and Improving Impossible Differential attacks: Applications to CLEFIA, Camellia, LBlock and Simon. In ASIACRYPT 2014, LNCS 8873, pages 179-199, 2014.

[8] A. Bogdanov, M. Wang: Zero Correlation Linear Cryptanalysis with Reduced Data Complexity, FSE'12, LNCS, Anne Canteaut (ed.), Springer-Verlag, 2012.

[9] Andrey Bogdanov, Huizheng Geng, Meiqin Wang, Long Wen, and Baudoin Collard. Zero-Correlation Linear Cryptanalysis with FFT and Improved Attacks on ISO Standards Camellia and CLEFIA. In Selected Areas in Cryptography - SAC 2013, LNCS 8282, pages 306-323, 2013.

[10] A.Bogdanov, V.Rijmen, Linear Hulls with Correlation Zero and Linear Cryptanalysis of Block Ciphers. Designs, Codes and Cryptography, Springer, US, 2012, pp. 1–15.

[11] H. Soleimany and K. Nyberg. Zero-Correlation Linear Cryptanalysis of Reduced round LBlock. Des. Codes Cryptography, 73(2):683-698, 2014.

[12] L. Wen, M. Wang, and A. Bogdanov. Multidimensional Zero-Correlation Linear Cryptanalysis of E2. In AFRICACRYPT'14, LNCS 8469, Springer-Verlag, pages 147-164, 2014.

[13] Wentan Yi and Shaozhen Chen, Multidimensional Zero-Correlation Linear Cryptanalysis of the Block Cipher KASUMI, [cs.CR] 14 Oct 2014.

[14] L. Wen, M. Wang, A. Bogdanov, and H. Chena. Multidimensional Zero-Correlation Attacks on Lightweight Block Cipher HIGHT: Improved Cryptanalysis of an ISO Standard. Information Processing Letters, 114(6):322-330, 2014.

[15] Bing Sun, Zhiqiang Liu, Vincent Rijmen, Ruilin Li, Lei Cheng, Qingju Wang, Hoda Alkhzaimi, Chao Li, Links among Impossible Differential, Integral and Zero Correlation Linear Cryptanalysis, Cryptology ePrint Archive, Report 2015/181, http://eprint.iacr.org/, 2015

[16] Wentan Yi and Shaozhen Chen, Improved Integral and Zero-correlation Linear Cryptanalysis of Reduced-round CLEFIA Block Cipher, Cryptology ePrint Archive, Report 2016/149, http://eprint.iacr.org/, 2016

[17] J. Chen, M.Wang, B. Preneel: Impossible Differential Cryptanalysis of Lightweight Block Ciphers TEA, XTEA and HIGHT. IACR Eprint Archive Report 2011/616, 2011.