Linear Cryptanalysis of CTC

Orr Dunkelman*[1] Nathan Keller**[2]

[1]Computer Science Department, Technion.
Haifa 32000, Israel
orrd@cs.technion.ac.il
[2]Einstein Institute of Mathematics, Hebrew University.
Jerusalem 91904, Israel
nkeller@math.huji.ac.il

Abstract. CTC is a toy cipher designed by Courtois in order to prove the strength of algebraic attacks. In this paper we study the differential and the linear behavior of the 85 S-boxes version, which is attacked using algebraic techniques faster than exhaustive key search. We show that an n-round variant of the cipher can be attacked by a linear attack using only 2^{2n+2} known plaintexts, with a negligible time complexity. We conclude that CTC is insecure, even for quite a large number of rounds. We note that our observations can be probably used to devise other attacks that exploit the relatively slow diffusion of CTC.

1 Introduction

Algebraic attacks [2, 3] have stirred quite a debate in the cryptographic community, trying to verify whether these attacks can be used to break ciphers faster than exhaustive key search. For several stream ciphers the answer is yes, as demonstrated in [1, 4].

In a recent paper, Courtois claims to break a relatively "strong" cipher named CTC using algebraic techniques [2]. There are several claims in that paper, including the claim that the algebraic attacks are applicable even for the AES, and for "sanity" reasons, the attack algorithm is omitted.

As the attack itself is unknown, we can only try to analyze CTC using widely known ways. We apply linear cryptanalysis to the CTC variant presented in the paper with 85 S-boxes. We show that there is a 5-round linear approximation with bias of 2^{-6} that ends with one active S-box. Our approximation is based on an iterative approximation that has one active S-box in each round, and has a bias of 2^{-2} per round. This approximation can be used in a linear attack that requires approximately 2^{14} known plaintexts and has a realistic time complexity of less than 2^{8} additional operations.

We also examine the possibility of replacing the S-box by another S-box in order to prevent the iterative linear approximation. We show that for more than

* This work was supported in part by the Israel MOD Research and Technology Unit.

** The research presented in this paper was supported by the Adams fellowship.

half of the possible 3-bit to 3-bit S-boxes, either the linear approximation holds or an iterative differential characteristic with probability of at least 2^{-10} for 5 rounds can be found.

It is important to note that our results do not prove that algebraic attacks are not working. However, they disprove the claim that the CTC cipher can be considered secure against the known attacks.

The rest of the note is organized as follows: Section 2 shortly describes the CTC block cipher with 85 S-boxes that we attack. In Section 3 we present our attacks, and in Section 4 we summarize our findings.

2 Description of CTC

CTC is a toy cipher presented for sake of cryptanalysis using algebraic attacks [2]. The cipher has support for a variable block size, a variable S-box size, and a variable number of rounds. In this paper we are mostly concerned with the version claimed to broken using algebraic attacks. This version has 3-bit to 3-bit S-boxes, and 85 S-boxes in each round out of the six rounds in the cipher. We denote this version by $\mathrm{CTC}_{3,85,6}$.

$\mathrm{CTC}_{3,85,6}$ has a block size and a key size of $3 \cdot 85 = 255$ bits. Each round is composed of XOR with a subkey, parallel application of the same S-box, and a simple linear transformation. After the last round another key is XORed to the output.

The 3-bit to 3-bit S-box used in $\mathrm{CTC}_{3,85,6}$ is $S[i] = \{7, 6, 0, 4, 2, 5, 1, 3\}$. The state is initialized to the plaintext XOR the first subkey, and the bits enter the S-boxes in groups of three consecutive bits, where bit 0 is the least significant bit of the first S-box and bit 254 is the most significant bit of the 85th S-box.

The linear transformation is very simple, and each output bit, denoted by Z_i, depends on one or two of the input bits, denoted by Y_i. We note that in [2] the notations are a bit different (as they include the round number before the number of the bit):

$$\begin{cases} Z_2 = Y_0 \\ Z_{i \cdot 202+2 \bmod 255} = Y_i \oplus Y_{i+137 \bmod 255} \text{ for } i = 1, \ldots 254 \end{cases}$$

3 Linear Attack on CTC

We note that the linear approximation table of the S-box presented for CTC has several one-bit to one-bit approximations. For example, the most significant bit of the input of the S-box is equal to the least significant bit of the output with probability 1/2+1/4. This specific fact will be used in our attack later, but we note that almost any 3-bit to 3-bit S-box may yield similar properties.

The input mask we choose is in bit 2, the most significant bit of the first S-box. This bit is equal to bit 0, the least significant bit of the output of the S-box, with bias 1/4. Note that by the linear transformation $Z_2 = Y_0$, and hence, bit 2 of the round output is equal to bit 2 of the round input with bias 1/4. Thus, this linear approximation is an iterative one, and can be concatenated to itself as many rounds as needed.

To obtain an r-round approximation from bit 2 of the input to bit 2 of the output, the basic approximation is concatenated r times, resulting in a linear approximation with a bias of $2^{-(r+1)}$. This approximation can be used to attack $r+1$ rounds with about 2^{2r+4} known plaintexts, and time complexity of about $2^{2r+4} \cdot 2^3$ partial decryptions of one S-box (about $2^{2r+4}/10r$ full r-round encryptions). The attack retrieves the equivalent of 3 key bits and the parity of another r key bits. Thus, the attack on $\mathrm{CTC}_{3,85,6}$ requires about 2^{14} known plaintexts, and has a running time of about 2^8 encryptions.

We note that if the difference distribution table of the S-box used in $\mathrm{CTC}_{3,85,6}$ had a non-zero probability in the entry corresponding to input difference in the middle bit and output difference in the most significant bit, an iterative differential characteristic could be constructed. This characteristic would be based on having an input difference in bit 136, that becomes a difference in bit 137 after the S-box, and returns to a difference in bit 136 after the linear transformation.

We note that constructing a 3-bit permutation that is not affected by the differential or by the linear attack, or has some other undesired properties, is highly unlikely. Out of the 40320 possible 3-bit permutations, 12288 have an iterative differential characteristic with probability 2^{-2} per round, 4608 have an iterative differential characteristic with probability 2^{-1} per round, and 384 have an iterative differential characteristic with probability 1. There are also 18432 permutations that have an iterative linear approximation with bias of $\pm 2^{-2}$ per round, and 1152 permutations with bias of 2^{-1} per round (i.e., where the least significant bit of the output is always equal to the most significant bit of the input, or is always its complement). Only 14336 permutations do not possess any of these iterative statistical properties, but they also can have other weaknesses.

4 Summary and Conclusions

The $\mathrm{CTC}_{3,85,6}$ cipher can be attacked using regular linear attacks, using a relatively low number of known plaintexts. This has no effect on the algebraic attacks on the cipher, but it does show that the mixing of CTC is relatively weak. As this is the case, it is expected that the relevant equations have a very sparse form, even for the entire cipher, resulting in the (probable) existence of efficient algorithms for solving the equation set.

We note that even if the number of rounds is increased, then the linear approximation can be extended, exploiting the fact that the same linear transformation is used in each round, and the "fixed" approximation of bit 2.

We currently collaborate with Nicolas Courtois to devise a cipher that is secure against differential and linear attacks, while algebraic attacks still succeed in attacking the cipher.

References

1. Nicolas T. Courtois, *Fast Algebraic Attacks on Stream Ciphers with Linear Feedback*, Advances in Cryptology, proceedings of CRYPTO 2003, Lecture Notes in Computer Science 2729, pp. 176–194, Springer-Verlag, 2003.
2. Nicolas T. Courtois, *How Fast can be Algebraic Attacks on Block Ciphers?*, IACR eprint paper 2006/168, 2006.
3. Nicolas T. Courtois, Josef Pieprzyk, *Cryptanalysis of Block Ciphers with Overdefined Systems of Equations*, Advances in Cryptology, proceedings of ASIACRYPT 2002, Lecture Notes in Computer Science 2501, pp. 267–287, Springer-Verlag, 2002.
4. Nicolas T. Courtois, Willi Meier, *Algebraic Attacks on Stream Ciphers with Linear Feedback*, Advances in Cryptology, proceedings of EUROCRYPT 2003, Lecture Notes in Computer Science 2656, pp. 345-359, Springer-Verlag, 2003.