How Fast can be Algebraic Attacks on Block Ciphers ?

Nicolas T. Courtois

Axalto Smart Cards, 36-38 rue de la Princesse
BP 45, 78430 Louveciennes Cedex, France
http://www.nicolascourtois.net
courtois@minrank.org

Abstract. In this paper we give a specification of a new block cipher that can be called the Courtois Toy Cipher (CTC). It is quite simple, and yet very much like any other known block cipher. If the parameters are large enough, it should evidently be secure against all known attack methods.However, we are not proposing a new method for encrypting sensitive data, but rather a research tool that should allow us (and other researchers) to experiment with algebraic attacks on block ciphers and obtain interesting results using a PC with reasonable quantity of RAM. For this reason the S-box of this cipher has only 3-bits, which is quite small.
Ciphers with very small S-boxes are believed quite secure, for example the Serpent S-box has only 4 bits, and in DES all the S-boxes have 4 output bits. The AES S-box is not quite as small but can be described (in many ways) by a very small systems of equations with only a few monomials (and this fact can also be exploited in algebraic cryptanalysis). We believe that results on algebraic cryptanalysis of this cipher will have very deep implications for the security of ciphers in general.

Key Words: algebraic attacks on block ciphers, AES, Rijndael, Serpent, multivariate quadratic equations, MQ problem, overdefined systems of multivariate equations, XL algorithm, XSL algorithm, Gröbner bases, solving systems of sparse multivariate polynomial equations.

1 Introduction

Claude Shannon, the father of information security as a science, has once advised that, breaking a good cipher should require "as much work as solving a system of simultaneous equations in a large number of unknowns", see [9]. This is an important and very explicit recommandation, yet it was ignored and nearly forgotten for more than 50 years.

The public researcher on symmetric cryptography have concentrated on local and statistical aspects of ciphers, and have overlooked the natural "global" approach of the problem. The secret key is defined as a solution of a system of algebraic equations that describes the whole cipher. This system should not be too simple, otherwise somebody might be able solve it...

The United States government encryption standard, AES, which is also expected to become a global de-facto encryption standard, is based on a particularly simple system of algebraic equations. This creates many uneasy feelings as a growing number of symmetric primitives (block ciphers, stream ciphers and hash functions) turn out to be insecure and are being broken by algebraic attacks. If AES is broken, it will have very serious consequences.

2 The Design of the CTC Block Cipher

CTC is an abbreviation for Courtois Toy Cipher. It has been designed in order to have the following properties:

a. It should be very simple, practical, and be implemented with a minimal effort.
b. It should be in general very much like any other known block cipher. If the parameters are large enough it should evidently be secure against all known attacks on block ciphers.
c. For simplicity, the key size should be equal to block size.
d. It should have a variable number of rounds and variable number of S-boxes in each round. However since it is a "research cipher" it is not required that it must encrypt 128-bit blocks. It can use for example 129-bit blocks (in fact it will be any multiple of 3).
e. The S-box should be chosen as a random permutation, and thus have no special structure.
f. Yet this S-box should exhibit an "algebraically vulnerability", by which we mean that it should be described by a small system of multivariate non-linear equations. This is made possible in spite of (e.) because the size of the S-box is quite small.
g. The diffusion should be very good: full avalanche effect should be achieved after about 3-4 rounds.
h. However, at the same time, the diffusion should not be too good, so that the linear parts of the cipher can still be described by (linear) equations that remain quite sparse. (In CTC each bit in the next round is a XOR of two bits from the outputs of two S-boxes from the previous round).
i. Finally and importantly, the cipher should allow to handle complete experimental algebraic attacks on block ciphers using a standard PC, with a reasonable quantity of RAM, and not more than a handful of plaintext / ciphertext pairs.

The simplest way we have found, to design a cipher that satisfies all these criteria, is to:

1. [Easy] Take the toy cipher described by Courtois and Pieprzyk in the appendix of the eprint paper [3] and improve the diffusion (that was excessively poor).
2. [Hard] Then in order to assure (i.), work on algebraic attacks and demonstrate that indeed it can be broken in practice even when parameters are quite large...

3 The Description of the CTC Block Cipher

Here is a short description of Courtois Toy Cipher (CTC) (notations are similar as in [3]).

1. CTC is quite similar to Serpent, except that it is much simpler, and the key schedule is a simple permutation of key bits, like for example in DES.
2. The S-box is the following permutation on $s = 3$ bits that has been chosen as a random non-linear permutation: $\{7, 6, 0, 4, 2, 5, 1, 3\}$. We will number its bits as follows: the input of the S-box is: $4 \cdot x_3 + 2 \cdot x_2 + x_1$, while the output is $4 \cdot y_3 + 2 \cdot y_2 + y_1$.
3. This S-box gives $r = 14$ fully quadratic equations with $t = 22$ terms, i.e. equations of the type:

$$\sum \alpha_{ij} x_i x_j + \sum \beta_{ij} y_i y_j + \sum \gamma_{ij} x_i y_j + \sum \delta_i x_i + \sum \epsilon_i y_i + \eta = 0$$

To be more precise, these equations are exactly:

$$\begin{cases} 0 = x_1x_2 + y_1 + x_3 + x_2 + x_1 + 1 \\ 0 = x_1x_3 + y_2 + x_2 + 1 \\ 0 = x_1y_1 + y_2 + x_2 + 1 \\ 0 = x_1y_2 + y_2 + y_1 + x_3 \\ 0 = x_2x_3 + y_3 + y_2 + y_1 + x_2 + x_1 + 1 \\ 0 = x_2y_1 + y_3 + y_2 + y_1 + x_2 + x_1 + 1 \\ 0 = x_2y_2 + x_1y_3 + x_1 \\ 0 = x_2y_3 + x_1y_3 + y_1 + x_3 + x_2 + 1 \\ 0 = x_3y_1 + x_1y_3 + y_3 + y_1 \\ 0 = x_3y_2 + y_3 + y_1 + x_3 + x_1 \\ 0 = x_3y_3 + x_1y_3 + y_2 + x_2 + x_1 + 1 \\ 0 = y_1y_2 + y_3 + x_1 \\ 0 = y_1y_3 + y_3 + y_2 + x_2 + x_1 + 1 \\ 0 = y_2y_3 + y_3 + y_2 + y_1 + x_3 + x_1 \end{cases} \tag{1}$$

4. The number of rounds is N_r.

5. Let $B = 1, 2, \dots, 128$ be the number of S-boxes in each round. There are $B * s$ bits in each round. We number them $0..Bs-1$, and we have in order 0 being x_1 of the first S-box, then we have x_2, x_3 of the first S-box, then x_1, x_2, x_3 of the second S-box (if any), etc.
6. The key size is equal to the block size and has $H_k = B * s$ bits, so that one known plaintext should be (on average) sufficient to determine (more or less uniquely) the secret key $K_0 = (K_{0\ 1}, \dots, K_{0\ Bs})$.
7. Each round i consists of the XOR with the derived key K_{i-1}, a parallel application of the B S-boxes, and then of a linear diffusion layer D is applied (this replaces the simple permutation of wires used in [3]). For the last round an additional derived key K_{N_r} is XORed (as in AES).
8. We denote $X_{i\ j}$, for $i = 1..N_r$, $j = 0..Bs-1$, the *inputs* of the $i-th$ round after the XOR with the derived key.
9. We denote $Z_{i\ j}$, for $i = 1..N_r$, $j = 0..Bs-1$, the *outputs* of the $i-th$ round before the XOR with the next derived key.

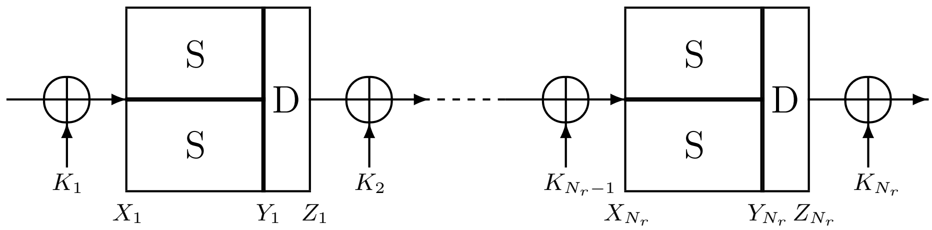

Fig. 1. A toy cipher with $B = 2$ S-boxes per round

10. I order to have uniform notations, we may also denote the plaintext by Z_0 and the ciphertext by X_{N_r+1}. These should not considered as variable names, but as abbreviations that denote (known) constant values.
11. There is no S-boxes in the key schedule and the derived key in round i, K_i is obtained from the secret key K_0, by a very simple permutation of wires:

$$K_{i\ j} \stackrel{def}{=} K_{0\ (j+i \bmod Bs)}. \tag{2}$$

12. With all these notations, the linear equations from the key schedule are as follows:

$$X_{i+1\ j} = Z_{i\ j} \oplus K_{i\ j} \quad \text{for all } i = 0..N_r. \tag{3}$$

13. The diffusion part D of the cipher is defined as follows:

$$\begin{cases} Z_{i\ (257 \bmod Bs)} = Y_{i\ 0} & \text{for all } i = 1..N_r \\ Z_{i\ (j\cdot 1987+257 \bmod Bs)} = Y_{i\ j} \oplus Y_{i\ (j+137 \bmod Bs)} & \text{for } j \neq 0 \ \text{ and all } i. \end{cases} \tag{4}$$

4 Cryptanalytic Results on CTC

Algebraic attacks on block ciphers appear to be (at least for CTC) much easier and much faster than it was ever expected. More details will be published soon.

References

1. Ross Anderson, Eli Biham and Lars Knudsen: *Serpent: A Proposal for the Advanced Encryption Standard.* Available from `http://www.cl.cam.ac.uk/~rja14/serpent.html`
2. Nicolas Courtois and Josef Pieprzyk: *Cryptanalysis of Block Ciphers with Overdefined Systems of Equations,* Asiacrypt 2002, LNCS 2501, pp.267-287, Springer.
3. Nicolas Courtois and Josef Pieprzyk: *Cryptanalysis of Block Ciphers with Overdefined Systems of Equations,* Available at `http://eprint.iacr.org/2002/044/`.
4. Joan Daemen, Vincent Rijmen: *AES proposal: Rijndael,* The latest revised version of the proposal is available on the internet, `http://csrc.nist.gov/encryption/aes/rijndael/Rijndael.pdf`
5. Joan Daemen, Vincent Rijmen: *The Design of Rijndael. AES - The Advanced Encryption Standard,* Springer-Verlag, Berlin 2002. ISBN 3-540-42580-2.
6. Nicolas Courtois: *The security of Hidden Field Equations (HFE)*; Cryptographers' Track Rsa Conference 2001, LNCS 2020, Springer, pp. 266-281.
7. Jacques Patarin: *Cryptanalysis of the Matsumoto and Imai Public Key Scheme of Eurocrypt'88*; Crypto'95, Springer, LNCS 963, pp. 248-261, 1995.
8. Adi Shamir, Jacques Patarin, Nicolas Courtois, Alexander Klimov, *Efficient Algorithms for solving Overdefined Systems of Multivariate Polynomial Equations*, Eurocrypt'2000, LNCS 1807, Springer, pp. 392-407.
9. Claude Elwood Shannon: *Communication theory of secrecy systems,* Bell System Technical Journal 28 (1949), see in particular page 704.