A Distinguish attack on COSvd Cipher

Mohammad Ali Orumiehchi ha, and R. Mirghadri

Abstract—The COSvd Ciphers has been proposed by Filiol and others (2004). It is a strengthened version of COS stream cipher family denoted COSvd that has been adopted for at least one commercial standard. We propose a distinguish attack on this version, and prove that, it is distinguishable from a random stream. In the COSvd Cipher used one S-Box (10×8) on the final part of cipher. We focus on S-Box and use weakness this S-Box for distinguish attack. In addition, found a leak on HNLL that the sub-s-boxes don’t select uniformly. We use this property for an improve distinguish attack.

Keywords—Stream cipher, COSvd cipher, distinguish attack, nonlinear feedback shift registers, chaotic layer.

I. INTRODUCTION

RECENTLY, the COSvd (COS version defense) cipher has been introduced by Filiol, Fontaine and Josse [1]. COSvd Cipher has been chosen as the core of the file system encryption software TURENNE for the security of sensitive data of restricted clearance and also COS family have been used for the IFIC1 project of European PRIAMM call. This cipher is a strengthened version of COS stream cipher family [2] for prevent some proposed attacks [3, 4]. The author’s COS particularly focused on the member of this family producing 128-bit blocks mainly from a 256-bit key. They used some different approaches in cipher design, for example they utilize NLFSR2 and Chaotic layer.

In stream cipher design, one usually use LFSRs3, as building blocks in different ways, and the secret key k is often chosen to be the initial state of the LFSRs. The designers often don’t use NLFSRs because their analysis is not explicit completely. For instance, we can’t determine some very important of characteristics of NLFSRs such as the period of them and etc. But, on the other hand, by using NLFSRs, the more famous attacks on LFSR-based stream cipher become failure. Correlation attacks, fast correlation attacks, algebraic attacks, ... are no longer working and even have no longer significance. Also, by using chaotic layer, increase disordered area defined in 2

Rn

geometric area defined in R2 by the following four points (convergence intervals):

(-1.33, 0.42), (1.32, 0.133), (1.245,-0.14) and (-1.06,-0.5).

Thus Henon [6, 7] proves Zn is always contained in this area for all n ≥ 0. This map is too weak for cryptographic purposes as shown by Erdman [8]. In order to bypass this weakness, two Henon maps have been used in parallel. The additional secret key is denoted σ1 = (x0, y0) and σ2 = (x0′, y0′).

Output blocks shift registers after cross over process are considered byte wise. For each of the byte, the chaotic module produces 10 bits h1, ..., h10.
Denote each output byte by \( b_1, \ldots, b_6 \). Thus HNLL module output one byte denoted \( c_7, \ldots, c_9 \) and computed as follows:

\[
(c_7, c_8, c_9) = F[(b_1 \oplus h_1), (b_2 \oplus h_2), \ldots, (b_6 \oplus h_6)].
\]

The \( F \) function is an optimally nonlinear substitution S-box mapping \( F_2^2 \times F_2^2 \) to \( F_2^8 \) and denoted S-Box 10×8. The overall setting is described in Figure 2.

![Fig. 2 Over all of COSvd Cipher](image)

### III. THEORETICAL BACKGROUND

In this section, we will present statistical theory bases that used in attack. Mainly, a distinguish attack specify that a stream is belong to a particular system or is a random stream. So, a frequently occurring problem in cryptanalysis is the determination of whether a sequence of observation is more likely to be sampled from a device having output distribution \( P \) or from a device having output distribution \( Q \).

Here, we will confine ourselves to a discussion of the following three issues [9]:

- The form of the optimum test.
- The probability of making decision.
- The number of samples needed in order to obtain a certain level of confidence in the decision.

Assume that we have a sequence of \( n \) independent and identically distributed (i.i.d) random variables \( X_1, X_2, \ldots, X_n \) over an alphabet \( N \). The distribution is denoted, \( Q(x) = \Pr(X_i = x) \) \( 0 \leq i \leq n \) and the sampled values are denoted \( x = x_1, x_2, \ldots, x_n \) where \( x_i \in N \), \( 0 \leq i \leq n \). We consider two hypotheses:

\[
H_0 : Q = P_0
\]

\[
H_1 : Q = P_1
\]

Let \( \phi(x) \) be a decision function where \( \phi(x) = 0 \) implies that \( H_0 \) is accepted and \( \phi(x) = 1 \) implies that \( H_1 \) is accepted. Furthermore let \( P_0^n(\phi) \) denote the simultaneous probability \( \prod_{i=1}^{n} P_0(x_i) \) and similarly we have \( P_1^n(\phi) = \prod_{i=1}^{n} P_1(x_i) \). Since \( \phi(x) \) only takes two values, we can specify a set \( A \in \{N\}^n \) over which \( \phi(A) = 0 \) and the complementary set \( A^c \in \{N\}^n \) over which \( \phi(A^c) = 1 \).

We can now specify the two types of error that can occur:

\[
P_F = \Pr(\phi(x) = 1 \mid H_0 \text{ is true}) = P_0^n(A^c)
\]

and,

\[
P_M = \Pr(\phi(x) = 1 \mid H_1 \text{ is true}) = P_1^n(A)
\]

Ideally, we would like to minimize both \( P_F \) and \( P_M \) but normally there is trade off. The optimum test between the two hypotheses is by given the Neyman-Pearson lemma, given here without proof.

**Lemma 1 (Neyman-Pearson [10]):** Let \( X_1, X_2, \ldots, X_n \) be drawn i.i.d according to the mass function \( Q \). Consider the decision problem corresponding to the hypotheses \( Q = P_0 \) vs. \( Q = P_1 \).

For \( T \geq 0 \) define a region

\[
A_n(T) = \left\{ \frac{P_0^n(x_1, x_2, \ldots, x_n)}{P_1^n(x_1, x_2, \ldots, x_n)} > T \right\}
\]

Let \( P_F = P_0^n(A_n(T)) \) and \( P_M = P_1^n(A_n(T)) \) be the probabilities of error corresponding to the decision region \( A_n(T) \).

Let \( B_n \) be any other decision region with associated probabilities of error \( P_F^B \), \( P_M^B \).

If \( P_F^B \leq P_F \) then \( P_M^B \geq P_M \).

The Neyman-Pearson lemma tells us that the region \( A_n(T) \), determined by the likelihood ratio \( \frac{P_0^n(x)}{P_1^n(x)} > T \), is the one that (jointly) minimizes \( P_F \) and \( P_M \). If we have symmetrical distributions of equal shape and would like to have the probabilities of error \( P_F \) and \( P_M \) equally large, we should choose \( T = 1 \). When computing the likelihood ratio for a large sample, both the numerator and the denominator tend to become very small and if a computer is used we could run into serious numerical problems. So, we can rewrite the test using a 2-logarithmic measure and \( T = 1 \) as

\[
\sum_{i=1}^{n} \log_2 \left( \frac{P_0^n(x_i)}{P_1^n(x_i)} \right) > 0 \quad (3.1)
\]

In (3.1), we have a simple, computationally robust test, which is easy to implement and tells us which of the two hypotheses \( H_0, H_1 \) is the most likely. The ratio is called a log-likelihood ratio, and the test is called a log-likelihood test. Assigning a priori probabilities to the two hypotheses, we can write the overall probabilities of error as

\[
P_\pi = \pi_0 P_F + \pi_1 P_M
\]

Where \( \pi_0 \) is the prior probability of \( H_0 \) and \( \pi_1 \) is the prior probability of \( H_1 \) and \( \pi_0 + \pi_1 = 1 \).
It can then be shown that $P_e$ is essentially equal to the larger of $P_F$ and $P_M$ and the total error is given by

$$P_e = 2^{-nC(P_0, P_I)}$$

(3.2)

Where $n$ is the number of samples, and $C(P_0, P_I)$ is the chernoff information [10].

$$C(P_0, P_I) = -\min_{\lambda \in [0, 1]} \log\left(\sum_{i=0}^{255} P_0(x)^i P_I(x)^{256-i}\right)$$

(3.3)

So, firstly we must calculate the chernoff information between $P_0$ and $P_F$ by using (3.3), and then we settle for an appropriate error probability $P_e$ and from (3.2).

Finally, we can distinguish a COSvd cipher stream from Random stream corresponding the following:

$$I = \sum_{i=0}^{255} f_i (\log_2 \left(\frac{F_i}{2^8}\right))$$

(3.4)

That $F_i$ is the distribution probability of numbers of S-Box, $f_i$ is the distribution probability of tested stream and $0 \leq i \leq 256$.

If $I > 0$ then the tested stream has been generated by COSvd Cipher. Also, if $I < 0$, we suppose that it’s a random stream.

### IV. DESCRIPTION OF ATTACK

As mention, this attack is applied on weakness of S-Box. So, we explain analysis of S-Box in two ways as that input of S-Box is random or generated by COSvd Ciphers.

#### A. Analysis of S-Box

The final part of cipher is one S-Box that must be increase nonlinearity of over all cipher. This S-Box is a mapping $F_2^8 \times F_2^8$ to $F_2^8$ but the probability for every output numbers is not equal. We know that each 8-bit happens with probability equal $\frac{1}{2^8}$ in a random stream but on S-Box the probability some 8-bit is different with $\frac{1}{2^8}$. For example, the number 0 repeats 14 times on S-Box (i.e. the probability of number 0 is $\frac{9}{1024} = \frac{9}{256}$) or number 7 repeats 2 times on S-Box (i.e. the probability of number 7 is $\frac{2}{1024} = \frac{2}{256}$).

We use this weakness and design a distinguish attack on cipher. Now, by using the relations (3.2), (3.3) and (3.4), we can calculate according as following:

For calculate I:

- Calculate the distribution probability of numbers of S-Box ($F_i$), $0 \leq i \leq 256$
- Calculate the distribution probability of tested stream ($f_i$), $0 \leq i \leq 256$
- Calculate ‘I’ according to (3.4).

If $I > 0$ then the output stream is COSvd otherwise output random.

Also, for $n=1024$ (1 kb) we have the error probability is:

$$P_e = 2^{-nC(P_0, P_I)} = 2^{-1024(0.01339)} = 0.0003196$$

We can choose the error probability firstly and then determine the number of samples.

For instance, for $P_e = 10^{-10}$, determine $n$ equal to 10 kb approximately.

#### B. The Improved Distinguish Attack

In section 4.1, we supposed that the input stream to S-Box is random. In other word, the probability of choose sub S-Boxes is uniform (Equal to $\frac{1}{2^8}$).

But, it isn’t true completely. We know that, by bits of $h_1, h_0$, which don’t have uniform distribution probability, the sub S-Boxes are selected. We understand that 00, 11 are more probable than 10, 01. Namely, the probabilities of 00, 11 are approximately equal to $\frac{1.14}{4}$. Therefore, the probabilities of 10, 01 are approximately equal to $\frac{0.86}{4}$. We determine these values by simulation. Then we have,

$$Pr(A_i) = Pr(00 | A_i(s = box1)) + Pr(01 | A_i(s = box2)) + Pr(10 | A_i(s = box3)) + Pr(11 | A_i(s = box4))$$

A_i’s are amounts of sub s-boxes and $0 \leq i \leq 256$.

Using these results, the error probability of decision for $n=1$ kb is:

$$P_e = 2^{-nC(P_0, P_I)} = 2^{-1024(0.0135737)} = 0.000550$$

### V. RESULTS AND SIMULATIONS

We wrote a program for testing this method. In other word, this program generates a lot of COSvd streams with different lengths ($2^{10} \leq n \leq 2^{14}$) and then decides whether it is a COSvd stream or not. We applied this method on $2^{20}$ streams that and all decisions were all right.

### VI. CONCLUSIONS

We presented that the stream of COSvd Cipher isn’t like a random sequence and is distinguishable from a random stream with negligible probability. In order to distinguishing, we need only a short length ($n_1 = 1$ kb or $n_2 = 10$ kb) of COSvd Ciphers with the negligible error probability ($P_e = 0.0005, P_{2e} = 10^{-10}$ respectively).
ACKNOWLEDGMENT

Many thanks to S. Khazaei for an initial idea and his helpful comments.

REFERENCES


