Statistical Analysis of the MARS Block Cipher

Andrey Pestunov

Abstract

The work contains a statistical investigation of the MARS block cipher — one of the AES finalists. It is shown that 8 MARS round ciphertext can be recognized from a uniform distribution with the help of the "Book Stack" test providing that 2^{18} blocks of plaintexts and 2^{20} bytes of memory are avaliable. The previous published attacks on this cipher were only theoretical with unrealistic resource requirements.

1 Introduction

Contemporary computer science widely uses different schemes for information security. One of the most popular ways is to use block ciphers. Any block cipher is a determined function that transforms plaintext to ciphertext with the help of a secret key. If one wants to encode a plaintext he needs to divide it into the blocks of the same length and then to encode them. The majority of the block ciphers include three algorithms: expanding the user's key into a key array, encryption and decryption. The structure of key expanding procedure is unique for different ciphers but encryption and decryption usually consists of several rather simple functions called rounds. They are repeated for R times and every round uses keys from the key array. Decryption is common with encryption but rounds go in reverse order. The security grows with the round number, but the speed goes down, that is why the designer recommends such number of rounds that would provide both security and speed.

There is no other way to study security than trying to attack the cipher i.e. to look for its weaknesses. Not always it is possible to attack all R rounds, so it is interesting even to attack $R_1 < R$ rounds, later the attack could spread over all rounds. Such investigations are really important and several competitions passed during the last years: American AES [1], Japaneese CRYPTREC [2], European

NESSIE [3]. Researchers all over the world worked for the goal of finding the best ciphers, that would be used as standard.

MARS block cipher [4] was one of the finalists of the AES competition. This cipher is not well studied because of its complicated structure: it consists of 32 rounds of four types. Published attacks on this cipher require unrealisic resources and so they are only theoretical, but not practical.

One of the requirements for block ciphers is a uniform distribution of ciphertext if all plaintexts blocks are different. It means that any bit takes a value 0 or 1 with the probability 1/2 independently from the others. In the current work we will show, that it is possible to carry out a distinguishing attack on this cipher using the "Book Stack" test [5]. The 8 rounds MARS ciphertext could be recognized from a uniform distribution using 2^{18} plaintexts and 2^{20} bytes of memory.

2 MARS description and previous results

MARS cipher deals with 128 bit blocks and uses's key can be of 128, 192 or 256 bits. Before encryption the user's key should be expanded into an array of forty 32 bit keys. For encryption we need to

1. Divide 128 bit block into four 32 bits subblocks;
2. Find sum between subblocks and the last four keys from the array;
3. Transform the block by eight FORWARD MIXING rounds;
4. Transform the block by eight FORWARD CORE rounds and the key array;
5. Transform the block by eight BACKWARD CORE rounds and the key array;
6. Transform the block by eight BACKWARD MIXING rounds;
7. Find difference between subblocks and the last four keys from the array.

Here is an encoding pseudocode:

```
Y[0]=X[0]+KEYS[0]; Y[1]=X[1]+KEYS[1];
Y[2]=X[2]+KEYS[2]; Y[3]=X[3]+KEYS[3];
For i=0 To 7 Do Y=Fmix(i, Y);
For i=0 To 7 Do Y=Fcore(i, Y, KEYS[4+4*i]);
For i=0 To 7 Do Y=Bcore(i, Y, KEYS[20+4*i]);
For i=0 To 7 Do Y=Bmix(i, Y);
Y[0]=Y[0]-KEYS[36]; Y[1]=Y[1]-KEYS[37];
Y[2]=Y[2]-KEYS[38]; Y[3]=Y[3]-KEYS[39];
```

After all these operations the block Y would be the result of encoding of the block X. We need to execute almost the same operations for decryption.

Among all works devoted to MARS cryptanalysis we want to emphasize the paper [6], where several MARS attacks are described along with the memory and computing resources required for their implementation. Specifically, for an attack on 8 rounds (2 of each type) of MARS 2^{25} plaintexts, 2^{29} memory bytes and 2^{68} encryptions are required. An attack on 21 rounds (16 MIXING 5 CORE) requires 8 plaintexts, 2^{232} encryptions and 2^{236} bytes of memory. Biham [7, 8] claims that 12 MARS rounds are not secure, but no such attacks were shown.

3 The Brief "Book Stack" test description

We will study statistical features of MARS using the "Book Stack" test, so let us give its brief description. The "Book Stack" test is used for checking hypotesys which claims, that elements of a sample $Z = (z_1, z_2, \ldots, z_N)$ from the alphabet $A = \{a_1, a_2, ..., a_S\}$ have uniform distribution, i.e. they are independent and

$$\mathbf{P}(z_n = a_i) = 1/S;\ n = 1, \ldots, N;\ i = 1, \ldots, S.$$

We should fix some order in the alphabet before testing. This order changes after analysis of every element z_n as follows: letter z_n gets number 1; numbers of all letters that were less than the number of z_n, increase for 1; all other letters numbers stay unamended. More formally: let $\omega^n(a)$ be number of the letter $a \in A$ after analysis of $z_1, z_2, \ldots, z_{n-1}$, then

$$\omega^{n+1}(a) = \begin{cases} 1, & z_n = a\,; \\ \omega^n(a) + 1, & \omega^n(a) < \omega^n(z_n); \\ \omega^n(a), & \omega^n(a) > \omega^n(z_n)\,. \end{cases}$$

This structure is common for a book stack if it is supposed that books number concurs with its position in the stack. The book is extracted from the stack and is put on the top so that its number becomes 1. All the books that were above it are shifted lower and all other stay on their places. Before testing the set of all numbers $\{1, \ldots, S\}$ is divided into two non-intersected parts $A_1 = \{1, 2, \ldots, K\}$ and $A_2 = \{K + 1, \ldots, S\}$, where $K \in \{1, ..., S\}$ is a parameter. Then with the help of the sample $(\ z_1, z_2, \ldots, z_N)$ we should count ν_N — quantity of numbers $\omega^n(z_n)$, that belong to the first part A_1, i.e. quantity of hittings in the topside part of the book stack. $(N - \nu_N)$ — it is, obviously, quantity of hittings in the lower part. Then we estimate

$$x^2 = \frac{(\nu_N - NP_1)^2}{NP_1} + \frac{((N - \nu_N) - N(1 - P_1)^2)}{N(1 - P_1)},\ P_1 = |A_1|/S,$$

and if x^2 is greater than the critical level $\chi^2_{1,1-\alpha}$, then uniform distribution is accepted otherwise rejected. The value $\chi^2_{1,1-\alpha}$ is quantile of chi square distribution with the level of significance equal to $(1-\alpha)$ with one degree of freedom. The degree of freedom is equal to one because the set of numbers is divided into two parts.

4 Investigation of MARS statistical features with the "Book Stack" test

Now let us describe how the sample for our experiments were constructed. Consider the block sequence X_i^u, $i \in \{1, ..., N\}$, $u \in \{0,1,2,3\}$, these blocks have three zero subblocks and subblock number u is equal to i. For example, $X_7^0 = (7,0,0,0)$, $X_5^3 = (0,0,0,5)$.

Let $y_i^{u,v}$ be a 32 bit subblock with the number v of ciphertext block after encryption of the block X_i^u. For example, after encoding of X_7^0, we will get the block ($y_7^{0,0}$, $y_7^{0,1}$, $y_7^{0,2}$, $y_7^{0,3}$). With the help of 100 random keys we will get 100 different samples of the view ($y_0^{u,v}$, ..., $y_{N-1}^{u,v}$) with the fixed parameters u and v. For some experiments we will take only s bits from the 32 bit words $y_i^{u,v}(i = 0, ..., N-1)$ and 2^s is an alphabet size. For each sample we estimate x^2 and calculate $U_{95\%}$ and $U_{99\%}$ that mean how many times of 100 this value overrided quantiles of chi square distribution of the level of significance 95% (equal to 3.84) 99% (equal to 6.64) accordingly. Speaking differently it means how many samples did not pass the test.

When the cipher consists of the same rounds reduction of their quantity is done in a usual way: we investigate the cipher consisted of one round, two rounds, three rounds and so on. MARS consists of the rounds of different types, so we can reduce them in different ways:

1. Encryption with only FORWARD MIXING rounds;
2. Encryption with only BACKWARD MIXING rounds;
3. Encryption with only FORWARD CORE rounds;
4. Encryption with only BACKWARD CORE rounds;
5. Symmetric round reduction (one or two of each type, this variant has been used in [6]).

Experiments have shown that the highest biasses from uniform distribution were when $u = 3$. The value v was different for different cipher modifications. The table contains the results of the experiments.

R	N	$U_{95\%}$	$U_{99\%}$	x^2_{av}	K	s
FORWARD MIXING rounds $u = 3, v = 2$						
2	$\mathbf{2^6}$	100	100	139.5	2^6	8
4	$\mathbf{2^6}$	100	100	139.5	2^6	8
6	$\mathbf{2^{14}}$	84	67	10.8	2^{14}	24
8	$\mathbf{2^{18}}$	70	43	7.8	2^{18}	32
BACKWARD MIXING rounds $u = 3, v = 1$						
2	$\mathbf{2^6}$	100	100	139.5	2^6	8
4	$\mathbf{2^6}$	100	100	139.5	2^6	8
6	$\mathbf{2^8}$	100	99	29.5	2^6	8
8	$\mathbf{2^{14}}$	41	25	4.4	2^{18}	24
FORWARD CORE rounds $u = 3, v = 3$						
1	$\mathbf{2^6}$	100	100	139.5	2^6	8
3	$\mathbf{2^6}$	100	100	139.5	2^6	8
5	$\mathbf{2^{20}}$	28	17	892.1	2^{18}	32
BACKWARD CORE rounds $u = 3, v = 0$						
1	$\mathbf{2^6}$	100	100	139.5	2^6	8
3	$\mathbf{2^6}$	100	100	139.5	2^6	8
5	$\mathbf{2^6}$	100	100	85.1	2^6	8
Symmetric reduction $u = 3, v = 0$						
1+1+1+1	$\mathbf{2^6}$	100	100	139.5	2^6	8
2+2+2+2	$\mathbf{2^{18}}$	51	29	6.1	2^{18}	32

here x^2_{av} is an average value of x^2 calculated out of 100 samples and $4K$ bytes of memory are required for the test implementation.

References

[1] Advanced encryption algorithm (AES) development effort. http://csrs.nist.gov/encryption/aes. 1997-2000.

[2] CRYPTREC project. http://www.ipa.go.jp/security/enc/CRYPTREC. 2000-2002.

[3] New European Schemes for Signatures, Integrity, and Encryption, Deliverables of the NESSIE project. http://www.cosic.esat.kuleuven.ac.be/nessie. 2003.

[4] IBM Corporation. MARS—a candidate cipher for AES. 1999.

[5] B. Ryabko, A. Pestunov. "Book stack" as a new statistical test for random numbers, Probl. Imform. Transmission, 40, 1, pp. 66-71. 2004.

[6] J. Kelsey, B. Schneier. MARS Attacks! Preliminary Cryptanalysis of Reduced-Round MARS Variants. In proceedings of the Third AES Candidate Conference. 2000.

[7] E. Biham, V. Furham. Impossible differentials on 8-Round MARS Core. In proceedings of the Third AES Candidate Conference. 2000.

[8] E. Biham. A Note on Comparing AES Candidates. In proceedings of the Second AES Candidate Conference. 1999.