On MARS's s-boxes Strength against Linear Cryptanalysis

Conference Paper · May 2003
DOI: 10.1007/3-540-44842-X_9 · Source: DBLP

4 authors, including:

Carlos Javier Hernández-Castro
University of Alcalá
13 PUBLICATIONS **66** CITATIONS
SEE PROFILE

Luis Javier García Villalba
Complutense University of Madrid
162 PUBLICATIONS **638** CITATIONS
SEE PROFILE

Julio Hernandez-Castro
University of Kent
199 PUBLICATIONS **2,244** CITATIONS
SEE PROFILE

On MARS's s-boxes Strength against Linear Cryptanalysis

Carlos Javier Hernández Castro[1], Luis Javier García Villalba[2], Julio César Hernández Castro[3], José María Sierra Cámara[3]

[1] Complutense University, Servicio Informático de Gestión, Madrid, Spain
chcastro@pas.ucm.es
[2] Complutense University, Departamento de Sistemas Informáticos y Programación, Facultad de Informática, Madrid, Spain
javiergv@sip.ucm.es
[3] Carlos III University, Computer Security Group, Computer Science Department, 28911 Leganés, Madrid, Spain
jcesar@inf.ucm.es, sierra@inf.ucm.es

Abstract. MARS's s-boxes were generated using a new algorithm developed by the IBM team, which was supposedly able of producing secure s-boxes against both differential and linear cryptanalysis. In this paper we show this is not the case, because their strength against linear cryptanalysis is not better (in fact, it seems to be worse) that what could be expected if generated randomly.

1 Introduction

MARS's s-boxes were obtained after a week of calculations using a new algorithm developed by IBM [1] which supposedly guaranteed the generated s-boxes had excellent properties against both differential and linear cryptanalysis. The designers of MARS did in fact conjecture that its maximum bias was not greater than 2^{-3}, a value that was, afterwards, proved to be too optimistic in [3] and [4].

In this paper we propose a mathematical model for the number of masks for a given bias value which, in particular, will have shown that the IBM's conjecture was extremely improbable and that the final results presented in [5] are not best than what could have been obtained at random. That simply means the new algorithm for producing s-boxes is not better, at least from the linear cryptanalysis point of view, than using a simple random generation.

In fact, we suggest that the generation procedure, basically a random generation followed by a series of tests and an optimization search between the s-boxes that were found to pass the filters, would probably never produce cryptographically sound s-boxes.

V. Kumar et al. (Eds.): ICCSA 2003, LNCS 2669, pp. 79-83, 2003.

1.1 MARS's s-boxes

As mentioned in [1], MARS's s-boxes were generated "in a pseudorandom fashion" (by using 32 bit words of the output of the SHA-1 algorithm with some fixed constants, an index and another 32 bits as input) and then filtered by testing they have good differential and linear properties. The properties tested were:

1. The S-box does not contain the all-zero or the all-one word.
2. Within each of the two s-boxes S_0, S_1 every two entries differ in at least three of the four bytes
3. S does not contain two entries S[i], S[j] with $i \neq j$ such that S[i]=S[j], S[i]=not(S[j]) or S[i]= - S[j]
4. S has $\binom{512}{2}$ distinct xor-differences and $2\binom{512}{2}$ distinct subtraction differences.
5. Every two entries in S differ by at least four bits.

Additionally, the algorithm tried to minimize the following values, in order to make the resulting s-boxes stronger against linear cryptanalysis:

6. Parity bias $|Pr_x[parity(S[x])=0]-1/2|$, requiring the bias to be at most 1/32
7. Single bit bias. For every j $|Pr_x[(S[x]_j)=0]-1/2|$, being it 1/30 at most
8. Two consecutive bit bias. For every j $|Pr_x[(S[x]_j \text{ XOR } S[x]_{j+1})=0]-1/2|$, requiring the bias to be at most 1/30
9. Single bit correlation: For every generated s-box which satisfies conditions 1-8, minimize $|Pr_x[(S[x]_j)=x_i]-1/2|$

Where the thresholds set above were calculated experimentally. The generation and testing proceeded for "around a week" studying 2^{26} values, and after that, the value which minimized the single bit correlation bias was selected. However, in [2] it its pointed out that the S-box actually fails to meet all the criteria shown above.

2 MARS's s-boxes Linear Probabilities Complete Distribution and Model

In [5] the author presented the complete distribution of the linear probabilities of MARS's s-boxes, calculated after "using about 2 months idle time of our processors" (5 processors ranging from 500 to 266 MHz.). In Table 1, we show both the real number of masks for each bias and the expected number of masks, as calculated by the Formula in **(1)** below:

$$\#\{m \in Z_2^{9+32} \mid bias(m) = n\} \approx 2 \frac{\binom{512}{256+n}}{2^{512}} 2^{41} \qquad \textbf{(1)}$$

Table 1. Number of masks for a given bias, with the expected number given by the formula in **(1)**, for some selected values

Bias * 2^9	Frequency	Expected
0	77498737588	77503773066,8126
1	154403399557	154404403930,770
2	152613000039	152609003885,064
3	149664380489	149662884118,943
4	145634858332	145633498777,279
5	140608578512	140611653991,855
…	…	…
64	14136	15183,8400327846
…	…	…
77	6	9,889082203093
78	4	5,299837468124
79	3	2,816033042764
80	0	1,483445977885
81	1	0,774737365304
82	2	0,401121416947
83	1	0,205885329052
84	1	0,104759299782

As shown in Table 1, the goodness of the formula given in (1) is quite good, except for the higher biases. However, the real values will not pass a chi-square goodness of fit test In particular, we can observe that the conjecture of the developers of MARS

about the higher bias being at most 2^{-3}, which translates to a value of 64, was very optimistic, as one could expect around 20,000 masks above this value. The discoveries in [3], which first pointed out this conjecture was too optimistic and exhibited biases over 80, and later the findings in [5], which showed biases over 82, would have been easily predictable by simply using the proposed formula.

Recalling that this formula reflects the distribution that could be expected if the s-boxes were generated randomly, one is tempted to conclude that not only the algorithm proposed by the developers of MARS is not clearly better (with respect to linear cryptanalysis, at least) than the much simpler random generation, but indeed worst for high biases, which are precisely those that matter to find the best linear approximations for the s-boxes. So it seems that, instead of spending a hole week or more in generating and testing 2^{26} s-boxes, the MARS team would have done better (on average) just by generating a single s-box at random.

3 Conclusions

The claim of a bias not higher than 2^{-3} by the MARS's team was spectacular and rather improbable, as far as there should be more than 18,000 masks with biases higher than 64 if the s-boxes were generated completely at random. Their conjecture would only hold if, by means of using a brand new s-box generation algorithm, they were able of obtaining much better s-boxes than those one could obtain at random, which would have been a very important achievement in the field but is obviously not the case, as the final s-box was not better with respect to its linear approximations than what could be expected of a random s-box.

Additionally, we do not believe the general procedure for obtaining s-boxes proposed in [1] nor the supposedly better method shown in [2] have much future in cryptology. Although there are previous works on generating Boolean functions by using heuristic optimization methods such as hill-climbing [6] and genetic algorithms [7], we believe that the best cryptographic primitives could not be obtained by maximizing/minimizing any simple value as in those proposals. Improving the strength against linear or differential cryptanalysis will surely imply weaknesses against new kinds of cryptanalytic methods, some perhaps yet to be discovered.

References

1. Burwick, C., Coppersmith, D., D'Avignon, E., Gennaro, R., Halevi, S., Jutla, C., Matyas,S., O'Connor, L., Peyravian, M., Safford, D., Zunic, N.: MARS- a candidate cipher for AES. Proceedings of the First AES Conference (1999). Revised September 22, 1999
2. Burnett, L., Carter, G., Dawson, E., Millan, W.: Efficient Methods for generating MARS like S-boxes. Proceedings of the Fast Software Encryption 2000 (FSE'2000)
3. Robshaw, M., Yin, Y.L.: Potential flaws in the conjectured resistance of MARS to linear cryptanalysis: Proceedings of the 3rd AES Conference

4. Knudsen, L., Raddum, H.: Linear Approximation to the MARS S-box. NESSIE Deliverable, April 2000
5. Aoki, K.: The Complete Distribution of Linear Probabilites of MARS's s-box. Cryptology e-print nº 33, June 30, 2000
6. Millan, W., Clark, A., Dawson, E.: Boolean Function Desing using Hill-Climbing Methods. Proceedings of the Symposium on Applied Cryptography (SAC'97). LNCS 1587
7. Millan, W., Burnett, G., Carter, G., Clark, A., Dawson, E.: Evolutionary Heuristics for finding Cryptographically strong s-boxes. Proceedings of the Information and Communication Security, Second International Conference, ICICS'99, Sydney, Australia, November 9-11, 1999. Lecture Notes in Computer Science 1726