Asian Journal of Information Technology 5 (9): 1017-1022, 2006

New Approaches for Selective Aes Towards Tackling Error Propagation Effect of Aes

Chandan T. Bhunia
International Centre for Theoretical physics, Trieste, Italy

Abstract: Advanced Encryption Standard (AES) has been developed to replace the Data Encryption Standard(DES). AES suffers from a major limitation of error propagation on its encryption process. To tackle this limitation two approaches are used: Redundancy based Technique and Byte Based Parity Technique. Redundancy based technique has some important advantages over parity based technique but at the cost of higher level of processing and lowering encryption speed. In this study we report the application of selective AES in redundancy based technique. Our technique gives better results in terms of higher throughput and higher speed of encryption over the conventional technique.

Key words: AES, error propagation, redundancy based technique, selective AES, better throughput

INTRODUCTION

Due to several recent past reports of failure[1,2] of security or key of DES (Data Encryption Standard), AES (Advanced Encryption Standard) has been developed to replace DES. The replacement has aimed to provide higher level of security mainly with higher key size. Besides, the higher level of security, AES has aimed to provide higher efficiency and better flexibility by means of encryption at different levels and with different block sizes[3]. But AES suffers from a major limitation of error propagation in the encryption process. The AES encryption is done at several rounds of iteration. Each round of iteration has different input data and different keys. The input data and the keys of different round are all generated from the original source data and the source key respectively.

Thus the input data and the keys at rounds follow a data path and a key path respectively. Any bit error(s) at any round if occurs either at data path or at key path, the effect propagates and results in huge errors. The research[4] reported this limitation of AES in their authoritative work We made a study on the error propagation under AES encryption. Our results[5] are in direct harmony with the study of[4]. It was found that even a single bit error at any stage of encryption process might accumulate as 64-bits error at the output cipher. Thus it is conclusively established that AES must not be applied without addressing the error propagation effect during its encryption process. The limitation of error propagation in AES results in low speed of encryption, more processing and higher complexity, as because until and unless error free encryption is achieved the transmission of the cipher will be meaningless.

For the purpose of tackling error propagation of AES, two techniques, namely the redundancy based technique and the byte based parity technique were studied in literatures[3,4]. The redundancy based technique needs two modules: encryption module and decryption module for producing error free cipher at the transmitter. The output cipher of the encryption module is decrypted by the decryption module. The decrypted output is compared with the plaintext to check for error. If they match, the cipher is error free and it is transmitted. The dual process of encryption and decryption by the technique make the encryption process slow and costly. The byte based parity technique studied in[4] makes use of parity checking at each byte of the plain text to combat error. It is mainly suitable for hardware implementation whereas the redundancy based technique is applicable to both hardware and software based implementations. The redundancy based technique guarantees the error correction for all error vectors that may generate in the AES encryption process. The byte based parity technique guarantees to tackle only one or two bits error vector. Thus there emerges a need to investigate for any modified redundancy based technique that will tackle the error propagation effect of AES for all error vectors but without lowering the speed of encryption to the extent it occurs in normal course of redundancy based technique. If such a modified technique is available, that will be a superior scheme to the byte based parity technique. This paper proposes a modified redundancy based AES scheme.

Basic idea: Recently the selective encryption has become a subject of investigation[6] as it provides a number of advantages in information transportation. In the selective encryption, only a fraction(r) of whole message/plaintext

Corresponding Author: Chandan T. Bhunia, Haldia Institute of Technology, PO: HIT, Pin-721657, East Midnapore, West Bengal, India

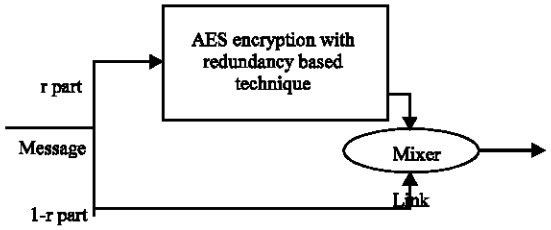

Fig. 1: Selective AES with redundancy technique

is encrypted. Its advantages that are explored so far are: (a) it reduces processing time and complexity (thereby increases speed of encryption) and (b) it provides a new type of system functionality. We will report here the application of selective encryption in the redundancy based AES for the purpose mentioned in study.

We propose an idea portrayed in Fig. 1. This is selective AES with redundancy based technique. The proposed technique will guarantee the error correction for all error vectors that may generate in the AES encryption process by the application of the redundancy based technique but without reducing the speed of encryption to a major extent.

The viability of the idea revolves around the choice of r. As r increases the security level increases. The choice of r must have to guarantee the desired level of security. The choice of r is dependant from message to message/application to application. Any message or application is having some keywords that are the main issues for security. In general any block of the plain text that contains a key word must be encrypted. The distribution of the keywords in whole of the message and their frequency of occurrence in whole of the message and in the blocks of the message (blocks are fragmented parts of message made for encryption) will be the main factors for choice of r. We propose the following algorithm I for selectio n of those blocks of the message, that require encryption.

Algorithm I:

- Input Key Words for the message of N words (say, n number of keywords are given). Find frequency of occurrence of each keyword in whole of the message, f_i (i=1 to n)
- [Say, the AES encrypts blocks each of M words (M<N). So there are k blocks where k=N/M]. Find frequency of occurrence of each keyword in each block, F_{ij} (i=1 to n, j=1 to k)
- If $F_{ij} \geq \{f_i.(1/k)\}$ for any one, more or all keywords, i.e, i=1 to n, the jth block will be encrypted in AES with redundancy based technique, otherwise not.
- Repeat (3) for all blocks, j=1 to k. When j=k, the proposed scheme of encryption is complete.

Analysis: If P is the probability of the failure of encryption of a block due to error, the average number of times the block requires encryption for a success is 1/(1-P) under AES with redundancy-based technique. This means that the speed of encryption is lowered down by (P/1-P)%.

When the plain text is made of k blocks, the total number of encryption required in full encryption of message with redundant based technique = k/1-P.

With selective encryption the number of blocks to be encrypted = kr. For these k.r blocks, the number of encryptions required =kr/1-P. Thus total number of encryption required under selective AES with redundancy based technique = kr/1-P + k(1-r).

Advantages of proposed scheme

Speed up factor: Thus the selective AES with redundancy based technique has a speed up factor over the redundant based AES as:

$$\text{speedupfactor (s)} = 1/r+(1-r)\,(1-P)$$

Throughput: The throughput (average number of encryption that a block requires to generate its cipher without error) of the AES with redundancy based technique is:

$$(1-P)$$

whereas that with proposed selective encryption with redundancy based technique is

$$1-P/r+(1-r)\,(1-P) = s(1-P)$$

This provides a throughput gain for the proposed scheme over the conventional scheme as:

$$\text{gain} = (s-1) \times 100\% \qquad (1)$$

As the expression of gain includes the speed up factor, in subsequent paragraphs the parameter gain has been taken for comparing different techniques.

An estimation of the value of r in case of random distribution: We assume that the key words are randomly distributed with α_i as the probability of occurrence of ith key word in any application or service. Assume that the same application generates a message of N words. Then

$$f_i = N\alpha_i$$

The appropriate probability model for the analysis will be binomial probability distribution when the probability that the jth block will be encrypted is

$$P_{ji}=\left[\sum_{l=M.\alpha_i}^{M}\binom{M}{l}\alpha_i^{\,l}(1-\alpha_i)^{M-l}\right]$$

where M is the block size; and number of blocks, k=N/M. Thus the probability of jth block being selected for encryption is P_{hj} where:

P_{hj} = highest value of all P_{ji} for all key words (i=1 to n).

Let us assume that P_{hj} (=p) is same for all the key words. Then the probability that r number of blocks will be selected out of k blocks is

$$P_{rs}=\binom{k}{r}P^r(1-P)^{k-r}$$

This suggests a value of r as r_{random} where:

$$r_{random}=\left[\sum_{1}^{k}\left(P_{rs}.k\right)\right]\times(1/k) \qquad (2)$$

Disadvantages of proposed scheme: The lower value of r will speed up the AES encryption but at the cost of lowering security level. The selective encryption opens a scope of attack by guessing. The security level will not be a linear relation with decreasing value of r, as because as r decreases larger part of the message will not be encrypted. This may result in more effect of guessing. Thus the effect of the guessing attack may be related to r in some exponential form. A reasonable empirical formula may be:

Level of security = $b\,r^t$ where $t \geq 0$.

The loss of the security level in the proposed scheme over the full encryption scheme = $b(1-r^t)$ where b and t are system constants. The percentage loss of security level will be:

$$\text{Loss} = (1-r^t)\times 100\% \qquad (3)$$

Lower the value of t, higher is the percentage loss. Thus for analysis purposes t is assumed as 1.1, 2 and 3 in the subsequent paragraphs.

The numerical results (with i=1 and $\alpha_i = \alpha$) based on Eq. 1-3 are shown in Fig. 2 and 3. It is found that:

- As expected when α=1, r becomes 1 (100%);
- As M increases, r decreases. As M increases, block size increases resulting in higher probability of the block to be encrypted. Again as M increases, k decreases. Thus less number of blocks will meet the criterion of being encrypted. These are the reasons behind such results.
- As expected with lower value of α, the value of r goes down. This is due to the fact that with lower α, the probability of any block to be encrypted goes down
- As gain depends on both r and P, the variation of gain has been shown in Fig. 3 for different sets of r and P. As expected gain decreases with increase of r and decrease of P. As expected, both gain and loss are 0, when r=1
- The important observations arrived from Fig. 3 are: (a) gain overshoots loss only when P is high and r is low, (b) at the low value of r, loss is more than gain at low P, (c) at high r, loss overshoots gain for wide sets of P and t.

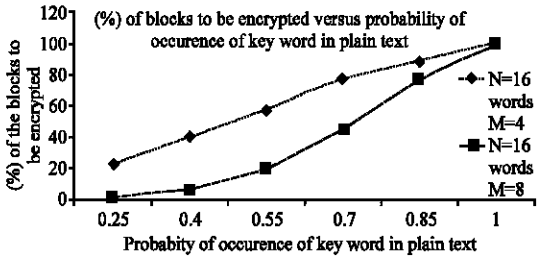

Fig. 2: Variation of r with α [Note: the results are given only for N=16 as because if N is high, the computation using binomial formula becomes unmanageable. In that study, Laplace approximation has to be used]

Table 1: Optimal t for different P

r	P	Optimal t
0.2	0.2	0.1312
	0.4	0.3950
	0.6	1.59
	0.8	Does not meet the criterion (i) and/or (ii)
	1.0	
0.4	0.2	0.1599
	0.4	0.4141
	0.6	0.9022
	0.8	2.799
	1.0	Does not meet the criterion (i) and/or (ii)
0.6	0.2	0.178
	0.4	0.4136
	0.6	0.7428
	0.8	1.2450
	1.0	2.150
0.8	0.2	0.190
	0.4	0.407
	0.6	0.656
	0.8	0.9469
	1.0	1.289

An optimal solution is when gain becomes loss. This provides the following condition of t for a set of r and P:

$$t=\frac{\log\left(\dfrac{1-2P(1-r)}{r+(1-r)(1-P)}\right)}{\log(r)} \qquad (4)$$

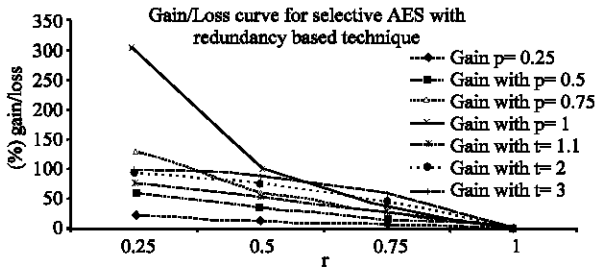

Fig. 3: Variation of gain and loss with r for a set of P and that of t, respectively

This means that t is real optimal value when r< 1 provided i) 2P (1-r)<1 and/or ii). Numerical results of equ 4 are given in Table I.

An estimation of the value of r in case of burst distribution: The value of r as given in Eq. 2 is for randomly distributed key words with binomial probability distribution. The distribution of key in any application/message may be

- randomly distributed like in normal text, or
- burst distributed. When the key words occur in consecutive positions or in clustered in message/blocks, burst occurs. One example of burst application is quotation that must have a key word unit price. The proposed technique of selective AES will provide better result for burst distribution in which case the value of r will be lower than that of random distribution as in Eq 1. If the average length of burst in the blocks of a given message is

$(\bar{L})$, $\bar{L} \geq (1/k)f_i$.

A good estimate of r for burst distribution may then be

$$r_{burst} = r_{random} \times \left(\frac{fi}{k\bar{L}} \right)$$

In Fig. 3, we have shown the variation of gain and loss with r for a set of P and that of t . As P depends on the probability of bit error, β of the system running the AES, a better form of investigation will be to see the variation of gain and loss with respect to β. This has been done in Fig. 4. We find that

- high β provides higher gain. This is what was found in Fig. 3 as high β amounts to high probability (P) that the generated cipher under AES is erroneous.
- gain is lower at lowβ, but the gain is lower than all the losses for all ts in the Fig. 4. This further confirm that the proposed scheme provides better results at high P

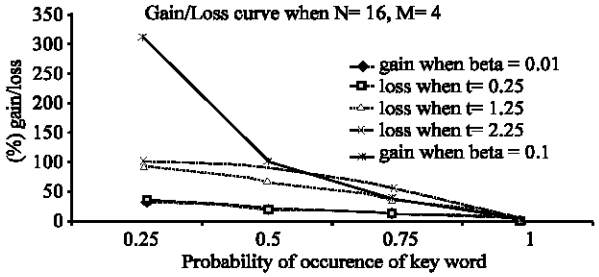

Fig. 4: Variation of gain/loss for different probability of bit error on AES process

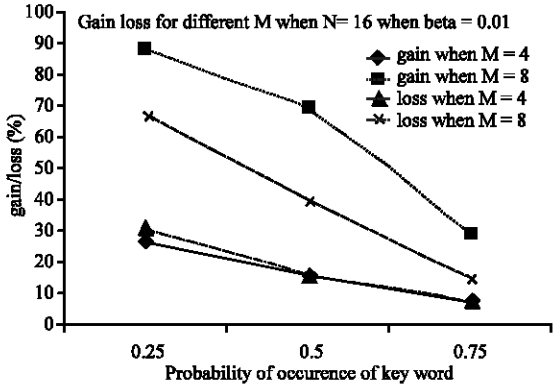

Fig. 5: Variation of gain/loss for different M

In Fig. 5, the variation of gail and loss with different block sizes (M determines the block size) has been shown. We have seen in Fig. 2 that the value of r changes with M and this suggests to study the present variation. We find in Fig. 5:

- As M increases, gain as well loss increases. But we need more gain and less loss. At high M gain is higher than the loss. This is not the case for low M. This suggests that the choice of higher size of block is desirable for the proposed scheme.

The observations on Fig. 4 and 5 thus together gives a rule for the proposed technique that: when β is high, proposed selective AES with redundancy based technique may be applied forthwith, but when β is low, the application of the proposed technique may be made with higher block size (higher value of M).

Algorithm II: The study made above on the proposed technique shows it has to compromise with security level. The algorithm I proposed for the selection of blocks to be encrypted has inherently compromise on the level of security. The algorithm proposes to encrypt when the block contains a minimum of $\{f_i.(1/k)\}$ key words. If the

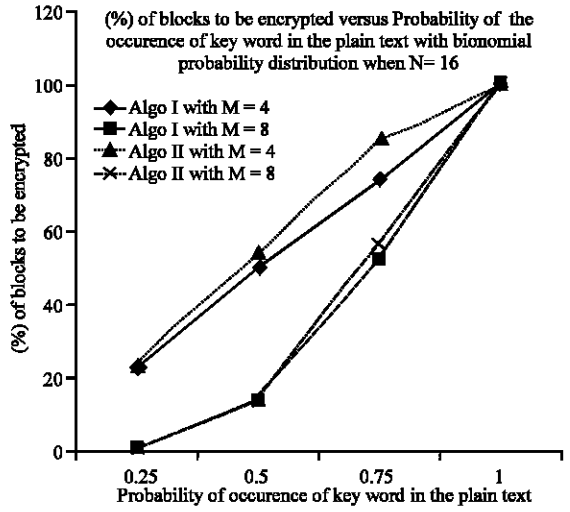

Fig. 6: Comparison of variation of r with probability of occurrence of key word for algorithm I and II

keyword is the parameter of secrecy, logically a block should be encrypted if the block has even a single keyword. Secondly algorithm I is applicable only when the whole of message is available prior to encryption. When the message is being generated and being processed for encryption on stream, the proposed algorithm I will not work. Thus the algorithm I may be modified as in algorithm II to provide a higher level of security confidence and to make it applicable to any message coming as a stream.

- Input key words for the message of N words (say, n number of keywords are given).
- [Say, the AES encrypts blocks each of N words (M<N). So there are k blocks where k=N/M].. If it occurs, encrypt the block otherwise not.
- Repeat (2) for all blocks, j=1 to k. When j=k, the proposed scheme of encryption is complete.

The results of obtaining 'r' for algorithm I and algorithm II are compared in Fig. 6. It is found that:

- As expected the value of r is higher in study of algorithm II thereby ensuring higher level of security that was the aim for algorithm II
- There is nominal difference in the values of 'r' in two studies The difference is higher when M is lower. This suggests that for low M, algorithm II may be used and for high M, algorithm I may be employed. The better response of algorithm I for higher M was established in Fig. 5 above.

Algorithm III: To further increase the level of security and to make the scheme applicable to a message

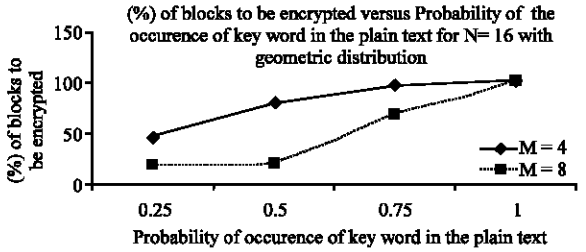

Fig. 7: Variation of r with probability of occurrence of keyword for algorithm III

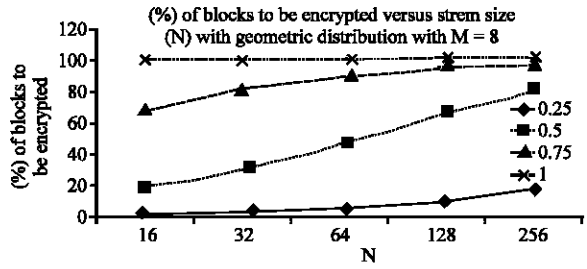

Fig. 8: Variation of r with respect to N under algorithm III

generating as a number of streams, we propose an algorithm III. The algorithm III fits to geometric distribution model.

- Input key words for the message of N words Message is divided into K parts each of k blocks. Each block is of M words
- Find occurrence of any keyword in the blocks, starting with first block in the first part. If it occurs, encrypt that and all the blocks thereafter in the part
- Repeat (1-2) for all parts, j=1 to K. When j=K, the proposed scheme of encryption is complete.

We assume a single keyword that a probability of occurrence in the message as β. Thus the probability, Q that a block of M words be selected for encryption is given as:

$$Q = \left[\sum_{i=1}^{M}\binom{M}{I}\alpha_i^{\ I}(1-\alpha_i)^{M-I}\right]$$

The probability, Q_i that the out of i sequential stream blocks, the last ith block will be encrypted is then:

$$Q_i = Q(1-Q)^{i-1}$$

Thus the value of r be obtained as:

$$r = (1/k)\left(\sum_{i=1}^{k}\left[(k-i+1)Q_i\right]\right) \quad (5)$$

The value of r under algorithm III is given in Fig. 7. We find that the variation is like that of binomial distribution, but r is higher in geometric distribution confirming better security.

We have studied the variation of r under algorithm III with message size, N. It was not done in case of algorithm I and II for computational problem with binomial co-efficient. We find in Fig. 8:

- With N, r increases as expected, thereby increasing the security level.
- As expected with probability of of occurrence of keyword in the message, r increases.

CONCLUSION

We have illustrated a selective AES with redundancy based technique that will tackle the error propagation effect of AES for all error vectors without degrading the performance (speed of encryption, throughput) of AES as happened in case of conventional AES with redundancy based technique. Three algorithms proposed for the selection of a few blocks out of all the blocks of the plain text for the purpose of encryption are presented. The selection of blocks is the key factor for maintaining thedesired security level of the technique. We have compared the algorithms. The observation for comparison are presented and explained. We find that the proposed technique will be effective at high value of P and low value of r , i.e when the system running the AES has high probability of bit error and the plain text has a low probability of occurrence of key word irrespective of algorithm used for selection of blocks to be encrypted.

REFERENCES

1. Allen Householde *et al*, 2002. Computer attack trends and challenges, internet security, Security and Privacy, IEEE Computer Soc., pp: 5-7.
2. NIST, 2001. Announcing the Advanced Encryption Standard (AES), Federal Information Processing Standards Pub., pp: 197.
3. Chandan T. Bhunia, 2005. Information Technology, Networks and Internet", New Age International Publishers, New Delhi.
4. Guido Bertoni *et al*, 2004. Error analysis and detection procedures for a hardware implementation of the advanced encryption standard, IEEE Trans on Computers, 52: 492-504.
5. Chandan T. Bhunia *et al*, 2004. Project Work on AES Error Propagation, ISM, Deemed University, India.
6. Tom Lookabaugh *et al*, 2004. Selective Encryption for Consumer Applications, IEEE Communication Magazine, 42: 124-129.