Implementation Analysis of Simplified AES (S-AES) Algorithm on Matyas-Meyer-Oseas (MMO), Davies-Meyer (DM), and Miyaguchi-Preneel (MP) Schemes using Yuval's Birthday Attack

Elena Sabarina, Bety Hayat Susanti, Agus Winarno
Sekolah Tinggi Sandi Negara

Keywords:

Hash functions
Simplified AES
Matyas-Meyer-Oseas scheme
Davies-Meyer scheme
Miyaguchi-Preneel scheme
Yuval's birthday attack

ABSTRACT

Matyas-Meyer-Oseas (MMO), Davies-Meyer (DM), and Miyaguchi Preneel (MP) schemes are block cipher based hash functions that used to provide data integrity mechanism. These schemes should be able to fulfill the collision resistance properties. In this paper, we analyze the implementation of Simplified AES (S-AES) algorithm as a compression function on MMO, DM, and MP schemes. We use Yuval's birthday attack to test the collision resistance of these schemes by conducting 120 experiments of extreme input and pseudo random input. The test results show that the collision occurred. Yuval's birthday attack for extreme input show that MP has a least number of collisions where MP, DM, and MMO have 102, 112, and 140 collisions, respectively. Otherwise, Yuval's birthday attack for pseudo random input show that MP has a greatest number of collisions, where MP, DM, and MMO 157, 133, and 117 collisions, respectively. Thus, these three schemes using either extreme or pseudo random input are not resistant to the collision.

Corresponding Author:

Bety Hayat Susanti,
Sekolah Tinggi Sandi Negara,
Jalan H.Usa Putat Nutug, Ciseeng, Bogor, 16330.
Email: bety.hayat@lemsaneg.go.id

1. INTRODUCTION

Hash functions maps an input of arbitrary length to an output with a fixed length [1]. On security application, hash functions are often used to maintain the integrity and authentication of data. For the last two decades, most widely used of hash functions in applications of cryptography are hash functions based on block ciphers and the dedicated hash functions [2]. The security of a hash functions based on block cipher is closely related to the security of the underlying block cipher. Not only for the security of block ciphers is well studied and their weaknesses are exploited, but also because they have good implementations both in hardware and software, such as hash functions have been popular. Although there are many possibilities, most designers concentrate on the following constructions: *single-block length compression functions* and *double-block length compression functions*. For single-block length hashing, the three most known modes are Matyas-Meyer-Oseas (MMO), Davies-Meyer (DM), and Miyaguchi-Preneel (MP) [3].

In the 1980's, several hash functions schemes are constructed using Data Encryption Standard (DES) block cipher algorithm. At present, the designers prefer the Advanced Encryption Standard (AES) block cipher algorithm as the basic construction of a hash functions [4]. This happens because the DES algorithm itself has been found a collision [5]. In addition, the AES algorithm is a block cipher algorithm standard adopted by the National Institute of Standards and Technology (NIST) in 2001. Simplified AES (S-AES) is a simplification of the AES algorithm that has the same structure but with a smaller key size and round [6].

A hash functions algorithm must satisfy a certain cryptographic properties. One of the basic criteria that must be fulfilled in the design of the hash functions was collision resistance [7]. Meanwhile, one of the attacks which can be used to find collisions on hash functions was Yuval's birthday attack. In this paper, we

analyzed the implementation of S-AES on MMO, DM, and MP schemes. Then, we will apply Yuval's birthday attack to find the output collision of each scheme. The purpose of this study was to determine the implementation analysis of S-AES algorithm on an MMO, DM, and MP schemes using Yuval's birthday attack.

2. THEORETICAL BACKGROUND

a. Unkeyed Hash Functions

A hash functions h maps bitstrings of arbitrary finite length to strings of fixed length, say n bits. For a domain D and range R with h : D→R and |D| > |R|, the function is many-to-one (input could be greater than the output), implying that the existence of *collisions* (pairs of inputs with identical output) is unavoidable. A hash functions must fulfill at least two properties, namely [8]:

1. *compression* — h maps an input x of arbitrary finite bitlength, to an output h(x) of fixed bitlength n.
2. *ease of computation* — given h and an input x, h(x) is easy to compute an output h(x).

At the highest level, a hash functions is divided into two categories namely keyed hash functions or MACs (Message Authentication Codes) which requires a secret key and a message as input and unkeyed hash functions or MDCS (Modification Detection Codes) which only requires a message as input. At the MDC's classification, in addition must satisfy two properties mentioned earlier, there are three additional properties, namely [4]:

1. *preimage resistance* : it is computationally infeasible to find any input which hashes to that output, i.e., to find any preimage x' such that h(x') = y when given any y for which a corresponding input is not known.
2. *2nd-preimage resistance* (weak collision resistant) : it is computationally infeasible to find any second input which has the same output as any specified input, i.e., given x, to find a 2nd-preimage x'≠ x such that h(x) = h(x').
3. *collision resistance* (strong collision resistant) : it is computationally infeasible to find any two distinct inputs x, x' which hash to the same output, i.e., such that h(x) = h(x').

Unkeyed hash functions (MDCS) can be classified into three categories based on the operation of the compression functions, i.e: hash functions based on block cipher, a dedicated hash functions, and hash functions based on modular arithmetic. Matyas-Meyer-Oseas, Davies-Meyer, and Miyaguchi Preneel are a single-length (n-bits) hash functions scheme of MDCS that have different characteristics.

2.1.1 Matyas-Meyer-Oseas (MMO) Scheme

In the MMO scheme, each block of plaintext message is encrypted using the key of block cipher which was a hash value generated from the previous iteration. The MMO scheme will be explained as follows [8]:

INPUT : bitstring x.
OUPUT : n-bit hash-code of x.

1. Input x is divided into n-bit blocks and padded, if necessary, to complete last block. Denote the padded message consisting of t n-bit blocks: $x_1x_2 \dots x_t$. A constant n-bit initial value IV must be pre-specified.
2. The output is H_t defined by: H_0 = IV ; $H_i = E_{g(H_{i-1})}(x_i) \oplus x_i$; $1 \le i \le t$

Descriptions of MMO scheme can be seen in Figure 1.

2.1.2. Davies-Meyer (DM) Scheme

DM scheme is a dual of MMO scheme. In this scheme, which is the key of block cipher is a message while the hash output of the previous iteration act as plaintext. The DM scheme will be explained as follows [8] :

INPUT : bitstring x.
OUPUT : n-bit hash-code of x.

1. Input x is divided into n-bit blocks where k is the key size, and padded, if necessary, to complete last block. Denote the padded message consisting of t blocks with each block of n-bit : $x_1x_2 \dots x_t$. A constant n-bit initial value IV must be pre-specified.
2. The output is H_t defined by: H_0 = IV ; $H_i = E_{x_i}(H_{i-1}) \oplus H_{i-1}$, $1 \le i \le t$.

Descriptions of DM scheme can be seen in Figure 2.

2.1.3. Miyaguchi-Preneel (MP) Scheme

MP scheme is the development of MMO scheme. The difference is the hash output from the previous iteration XORed with the ciphertext. The MP scheme will be explained as follows [8]:

INPUT : bitstring x.
OUPUT : n-bit hash-code of x.

1. Input x is divided into n-bit blocks where k is the key size, and padded, if necessary, to complete last block. Denote the padded message consisting of t blocks with each block of n-bit : $x_1x_2 \ldots x_t$. A constant n-bit initial value IV must be pre-specified.
2. The output is H_t defined by: $H_0 = IV$; $H_i = E_{g(H_{i-1})}(x_i) \oplus x_i \oplus H_{i-1}$, $1 \le i \le t$.

Descriptions of MP scheme can be seen in Figure 3.

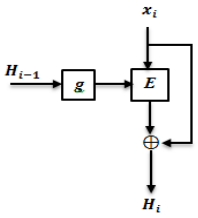

Figure 1. MMO Scheme [8]

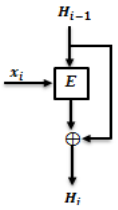

Figure 2. DM Scheme [8]

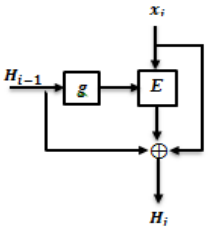

Figure 3. MP Scheme [8]

b. Simplified AES (S-AES)

Simplified AES (S-AES) was developed in 2002 as a teaching tool to help students understand AES [6]. Structure of S-AES algorithm similar to the structure of the AES algorithm. The S-AES algorithm operates on 16-bit plaintexts and generates 16-bit ciphertexts, using the expanded key. The key length is 16-bits and the number of rounds is 2.

Within the S-AES encryption process, there are 4 main components, namely *NibbleSub*, *ShiftRow*, *MixColumn* and *KeyAddition*. The application of these 4 components in sequence constitutes a *round* of S-AES. Each function on these components operate at state. Each state represented as a matrix of 4-bits (nibble) with 2 rows and 2 columns. The encryption processes of S-AES algorithm can be seen in Figure 4.

c. Yuval's birthday attack

Yuval's birthday attack is application of the birthday paradox derived from the classical selection problem (classical occupancy problem) which if selected elements of the population N at random, there will be a recurrence after the election of the O ($\sqrt{N}$). The concept is still the same as the birthday attack in general, which makes finding collision in hash functions. This attack is relevant to hash functions because it is easier to find collision on a one-way hash functions than finding preimage or second preimage. It can be applied to all the unkeyed hash functions that has a running time O ($2^{m/2}$) with output length of m-bit hash. The algorithm of Yuval's birthday attack will be explained as follows [8]:

INPUT: legitimate message x_1; fraudulent message x_2; m-bit one-way hash functions h.

OUTPUT: x_1', x_2' resulting from minor modifications of x_1, x_2 with $h(x_1') = h(x_2')$.

1. Generate $t = 2^{m/2}$ minor modification x_1' of x_1.
2. Hash each such modified message, and store the hash-values (grouped with corresponding message) such that they can be subsequently searched on hash-value.
3. Generates minor modifications x_2' of x_2, computing $h(x_2')$ for each and checking for matches with any x_1' above; continue until a match is found.

Birthday attack is basically used to facilitate collision finding in hash functions. According to [9], Yuval's birthday attack is used not merely to find collision but also for knowing the performance of the collision resistance or resilience to the birthday attack of the hash functions.

d. Mode

The observation cluster mode is the value that occurs most often or have the highest frequency. Mode does not always exist [10]. This occurs when all observations have the same frequency. For specific data, there may be some value with the highest frequency and thus the mode of observation is more than one.

e. Quartile

The value of a distribution for which some fraction or a certain percentage of the sample lies below is called fractiles or quantiles [10]. Fractiles are numbers that partition, or divide, an ordered dataset into equal parts. For instance, the median is a fractile because it divides an ordered data set into two equal parts. One of the interesting and important of fractiles is called quartiles. Quartiles are values that divide a cluster observations into four equal parts. Quartiles can be obtained by sorting the data from the smallest to the largest. We calculate the quartiles of an individual data by the formula (1) [11]

$$Q_i = \frac{i(u+1)}{4} \qquad \text{.....(1)}$$

where Q_i is *i*-th quartile and *u* is the number of observation sample.

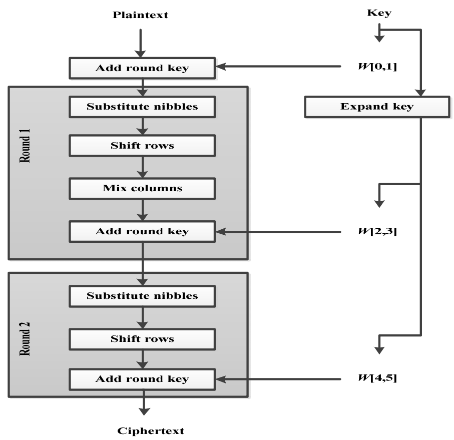

Figure 4. Encryption Processes of S-AES [12]

3. RESEARCH METHOD

Experiments performed by implementing S-AES algorithm into MMO, DM, and MP schemes using the C + + Dev programming language. Then, Yuval's birthday attack against the MMO, DM, and MP schemes. Collision searches performed by trying 16 extreme inputs (with certain characteristics) and 16 pseudo random inputs with a total trials for each were C_2^{16} = 120 experiments.

The size of the hash input used for both extreme and pseudo random inputs is 512-bits or equivalent to 128-digit hex. We take plaintext size 512 bits since it is large enough to analyze the collisions. With large domain (512 bits) and small codomain (16 bits) then collision may occur. Moreover, 512 bits can be operated on MMO, DM, and MP schemes without using padding bits (512 bits divisible by 16). On Yuval's birthday attack, it takes input as much as $2^{n/2}$ in the form of minor modifications of the original messages and fake messages, n denotes the bitlength of output of hash functions. Selection of the number of inputs taken as much as half of all possible input of hash functions in order to find the collision with a probability approximately 0.5. Due to the bitlength of the output of MMO, DM, and MP schemes are 16 bits, Yuval's birthday attack can be applied to all schemes by generating 2^8 minor modifications of input. Minor modification performed by changing the 8 final bits (2 hex) of messages as much as 256 possibilities. Hash values of the minor modification of the original message and the fake message will be counted and recorded. Furthermore, the hash value of minor modifications of the original message will be compared with the fake message to get the number of collisions happened. Analysis performed by looking at the value of quartiles and mode of collision of each scheme. If one collision is found, then all the hash functions schemes are considered broken which means that the properties of collision resistance are not fulfilled [13].

Table 1. Extreme Input on MMO, DM, and MP schemes

Symbol	*Hex input* (128 hex/512 bits)
0	000000000000000000000000000000000000000000000…000
1	1111111111111111111111111111111111111111111111…111
2	222222222222222222222222222222222222222222222…222
⋮	⋮
d	ddddddddddddddddddddddddddddddddddddddddddddd…ddd
e	eeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeee…eee
f	ffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffff…fff

4. RESULTS AND ANALYSIS

4.1. Collision Analysis Results of the Implementation of S-AES algorithm on MMO, DM, and MP Schemes with Extreme Input

Yuval's birthday attack results on MMO, DM, and MP schemes with extreme input indicates that the collision happened in the three schemes as shown in Table 2. The input columns in the table 1 contains a minor modification of the input x_1 and x_2 as much as 256 pieces. While the column of number of collision $h(x'_1) = h(x'_2)$ describes the number of Yuval's collision that happened to MMO, DM, and MP schemes.

Table 2. Result of Yuval's Collision of MMO, DM, and MP Schemes for Extreme Inputs

Input (128 hex)		Number of collisions ($h(x_1') = h(x_2')$)		
x_1	x_2	MMO	DM	MP
0	1	2	2	0
0	2	2	2	0
0	3	0	0	0
…	…	…	…	…
7	e	16	0	1
7	f	0	5	2
…	…	…	…	…
a	d	0	3	9
…	…	…	…	…
d	e	1	1	0
d	f	2	0	1
e	f	0	1	4
TOTAL		140	112	102

Table 2 shows that MP has a least number of collisions where MP, DM, and MMO have 102, 112, and 140 collisions, respectively for a total of 120 trials. The largest collision on the MMO, DM, and MP schemes were 16, 5, and 9 respectively, which occurred on the input pair 7-e, 7-f, and a-d. Yuval's collision results can be classified according to the number of collisions happened so that it can be presented in the frequency table as in Table 3.

Table 3. Collision Frequency Table of MMO, DM, and MP schemes with extreme input

Number of collisions		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	MMO	71	16	12	9	7	-	2	-	1	1	-	-	-	-	-	-	1
Frequency	DM	45	49	18	6	1	1	-	-	-	-	-	-	-	-	-	-	-
	MP	72	21	13	8	4	-	1	-	-	1	-	-	-	-	-	-	-

From Table 3, we can calculate the value of Q_1, Q_2 (median), and Q_3, and the mode of collision that occurs in MMO, DM, and MP schemes. Using equation (1), it was found that the Q_1, Q_2, Q_3 values are 30.25, 60.5 and 90.75, respectively for a total of 120 trials.

In the MMO scheme Q_1, Q_2 (median), Q_3, and collision modes that occurred is 0, 0, 2, and 0 respectively. In the DM scheme Q_1, Q_2 (median), Q_3, and collision mode that occured is 0, 1, 1, and 1 respectively. Whilst, in the MP scheme Q_1, Q_2 (median), Q_3, and collision modes that occurred is 0, 0, 1, and 0 respectively. Based on Yuval's collision results data that have been described, it can be concluded that the MMO, DM, and MP schemes with extreme input are not resistant to the collision that happened even though the amount is very small.

4.2. Collision Analysis Results of the Implementation of S-AES algorithm on MMO, DM, and MP Schemes with Pseudo random Input

Yuval's birthday attack results on MMO, DM, and MP scheme with pseudo random input indicates that the collision happened in the three schemes as shown in Table 4.

Table 4. Result of Yuval's Collision of MMO, DM, and MP Schemes for Pseudo random Inputs

Pseudo random Input		Number of collisions ($h(x_1') = h(x_2')$)		
x_1 (128 hex)	x_2 (128 hex)	MMO	DM	MP
4b576abd2e…8e529	71cd83e315…8210b	-	1	-
4b576abd2e…8e529	26df5c7bd9…5a6cf	-	-	2
4b576abd2e…8e529	c380184e8f…db992	-	2	-
…	…	…	…	…
…	…	…	…	…
684f04ecd8…61250	53f3fb02cd...94563	-	3	2
684f04ecd8…61250	1fc0299778…b0900	-	1	-
…	…	…	…	…
…	…	…	…	…
beafa8ebe4…bf68b	0c0d48e636…53073	-	-	-
beafa8ebe4…bf68b	a4c96537dc…ad17f	-	1	2
0c0d48e636…53073	a4c96537dc…ad17f	-	-	1
TOTAL		117	133	157

Table 4 shows that MMO has a least number of collisions where MMO, DM, and MP have 117, 133, and 157 collisions, respectively for total of 120 trials. The largest collision on the MMO, MP, and DM schemes were 12, 12, and 4 respectively. Yuval's collision results can be classified according to the number of collisions happened so that it can be presented in the frequency table as in Table 5.

Table 5. Collision Frequency Table of MMO, DM, and MP schemes with pseudo random input

Number of collisions		0	1	2	3	4	5	6	7	8	9	10	11	12
Frequency	MMO	76	14	12	9	4	-	4	-	-	-	-	-	1
	DM	41	45	18	12	4	-	-	-	-	-	-	-	-
	MP	66	17	14	11	4	1	4	-	-	1	-	-	2

From Table 5, we can calculate the value of Q_1, Q_2 (median), and Q_3, and the mode of collision that occurs in MMO, DM, and MP scheme. In the MMO scheme, the value of Q_1, Q_2 (median), Q_3, and collision modes that occurred is 0, 0, 1.75, and 0 respectively. In the DM scheme Q_1, Q_2 (median), Q_3, and collision modes that occurred is 0, 1, 2, and 1 respectively. Whilst, in the MP scheme Q_1, Q_2 (median), Q_3, and collision modes that occurred is 0, 0, 2, and 0 respectively. Based on Yuval's collision results data that have been described, it can be concluded that the MMO, DM, and MP schemes with pseudo random input are not resistant to the collision that happened even though the amount is very small.

5. CONCLUSION

In this paper, we analyze the implementation of Simplified AES (S-AES) block cipher algorithm on MMO, DM, and MP schemes using Yuval's birthday attack. The test result shows that the MMO, DM, and MP schemes with both extreme and pseudo random inputs are not resistant to the collision that happened even though the amount is very small. In the future, we will implement other block ciphers with different input to MMO, DM, and MP schemes so we can analyze the effect of their implementation to hash functions.

REFERENCES

[1] Bart Preneel, "Analysis and Design of Cryptographic Hash Functions," PhD thesis, Katholieke Universiteit Leuven, 2003.

[2] M.K. Reddy Danda, "Design and Analysis of Hash Functions," A Thesis in Network Security and Cryptography, Victoria University, 2007.

[3] D. Toz, "Cryptanalysis of Hash Function", Dissertation presented in partial fulfillment of the requirements for the degree of Doctor in Engineering, Katholieke Universiteit Leuven, 2013.

[4] Bos, et al., "Efficient Hashing using the AES Instruction Set," Ecole Polytechnique Federale de Lausanne, University of Bristol, 2011.

[5] J.J. Quisquater and J.P.Delescaille, "How easy is collision search ? Application to DES," Advances in Cryptology-EUROCRYPT '89: Springer-Verlag, 1989.

[6] M. Musa, et al., " A Simplified Rijndael Algorithm and Its Linear and Differential Cryptanalysis", 2002.

[7] Doganaksoy, et al., "Cryptographic Randomness Testing of Block Ciphers and Hash Functions," Institute of Applied Mathematics, Middle East Technical University, 2010.

[8] A. J. Menezes, et al., "Handbook of Applied Cryptography," Boca Raton : CRC Press LLC, 1997.

[9] Kocarev, et al., "Chaos-Based Cryptography Theory, Algorithms and Applications", Studies in Computational Intelligence, Volume 354, Springer.

[10] R. Walpole, "Pengantar Statistika (Edisi Ketiga)", Jakarta: Gramedia Pustaka Utama, 1997.

[11] A.Supangat, "Statistika Untuk Ekonomi dan Bisnis", Bandung: Pustaka, 2006.

[12] J. Holden, et al., "A Simplified AES Algorithm," Institute of Technology, 2010.

[13] Stamp, et al., "Applied Cryptanalysis Breaking Ciphers in the Real World," John Wiley & Sons, Inc. Hoboken, New Jersey, 2007.

BIBLIOGRAPHY OF AUTHORS

	Elena Sabarina is a student at Sekolah Tinggi Sandi Negara, Bogor. Her interests are in mathematics, cryptography, and its related.
	Bety Hayat Susanti is a lecturer at Sekolah Tinggi Sandi Negara, Bogor. She obtained undergraduate degree in Math and master's degree in Planning and Public Policy at University of Indonesia. Her interests are in mathematics, cryptography, and its related.
	Agus Winarno is a student at Sekolah Tinggi Sandi Negara, Bogor. His interests are in mathematics, cryptography, and its related.