Modified Blowfish Algorithm

Theda Flare G. Quilala¹, Ariel M. Sison², Ruji P. Medina³
¹,³Technological Institute of the Philippines, 938 Aurora Blvd., Cubao, Quezon City, Philippines
²Emilio Aguinaldo College, 1113-1117 San Marcelino St., Paco, Manila 1000, Philippines

ABSTRACT

Cryptography guarantees security in communication through encryption. This paper proposed a modified Blowfish encryption that uses 128-bit block size and 128-bit key to comply with minimum requirements as an encryption standard. The modification retained the original structure for easy migration but utilized two S-boxes to save memory. A derivation was added to prevent symmetry. The algorithm's performance was evaluated using time, and avalanche. Upon testing, the modified blowfish is slower with key, encryption, and decryption average of 26.99ms, 1651.83ms, and 2765.04ms compared to blowfish with 21.65ms, 1297.76ms and 2176.59ms due to block size difference. Applying 128-bit block size increases security by decreasing the chances of having duplicate blocks that may leak information. The modified Blowfish is faster compared to Twofish with an encryption and decryption average time of 2418.08ms and 4002.70ms. The added derivation improved the avalanche of the modified blowfish. Blowfish achieved 47.14% while modified Blowfish attained 52.86%.

Keywords:
Twofish
Cryptography
Encryption
Security
Cipher

1. INTRODUCTION

Security is involved with the protection of network and data while communicating over the public networks [1]. It is one of the prominent areas of concern in communication and data transmission, particularly in open networks, like the Internet. Examples of sensitive information transmitted through public communication facilities are financial transactions, medical and personal records [2]. As a result, various hackers always try to break into the system to steal critical information or to destroy the integrity of data [3]. With the increasing growth of science and study in the field of the network, we have the responsibility to secure our image, and data from third parties [4].

One way of guaranteeing the protection of information is through the application of cryptography. Cryptography is the practice and study of information hiding and achieving security by encoding messages to make them non-readable [5]. The use of cryptography addresses data privacy preservation and security from modification and unauthorized access during transmission [6]-[10]. Some popular and well respected symmetric-key block ciphers currently available include DES, 3DES, CAST5, RC6, CAST5, Blowfish, Twofish, Serpent, AES (Rijndael), TEA, IDEA, Serpent, and MARS [11], each exhibiting its strength and weakness.

Bruce Schneier designed Blowfish algorithm in 1994; it is a symmetric block cipher that aims to replace the outdated DES. Blowfish is a 64-bit variable length symmetric block cipher [12]. Blowfish is one of the fastest, compact, easy to understand, easy to implement, free alternative to existing encryption algorithms and features variable security level except when changing keys [13]. Several types of research were already conducted to test the security provided by Blowfish and results indicate that it is indeed fast and secure [14]-[16]. An attack developed by John Kelsey could break 3-round Blowfish; however, he was not
able to expand it. Disclosure of F allows performing a differential cryptanalysis which can recover all the rest of the key with 248 chosen plaintexts against the number of rounds reduced to eight [17] but was deemed impossible for round 8 and higher. To date, there is still no known attacks on the complete 16 rounds of blowfish.

Despite the fact that blowfish is a remarkably fast block cipher, extending it to act on 128-bit is the most natural manner [18]. In 1997, a request for candidate algorithm nominations for the Advanced Encryption Standard listed minimum functional requirements and asked for a symmetric block cipher capable of supporting block lengths of 128 bits and a key length of 128 bits [19]. Twofish, an encryption algorithm based on Blowfish, accepts 128-bit block size and was submitted and qualified as one of the finalists for AES [20] that provides a substantial level of security but lacks in encryption speed as compared to blowfish [21]. Twofish has seen less widespread usage than Blowfish [22].

Blowfish algorithm consists of two parts-key expansion and data encryption. In the key expansion, applying XOR to the variable length key and plaintext are used to produce the subkeys and generate the four key-dependent s-boxes. Each round requires around four kB which made the algorithm inapplicable for devices with a small memory like a smart card and phone. Using the algorithm, computation of the subkeys every time results in slower operation which made the algorithm inefficient to use in an application that requires changing secret key frequently [23]. However, three possible simplifications recommended by Schneier aimed at decreasing memory requirements and execution time. Those suggestions include the use of fewer and smaller S-boxes, fewer iterations from 16 to 8, and on-the-fly subkey calculation [12].

Other studies improved key generation of Blowfish to minimize the time required to produce subkeys [23]-[26]. For all studies involving the production of subkeys, the generation of elements achieved reduced time complexity though approaches used was entirely different from that of the original algorithm. Some researchers focused on the security aspect by modifying the f-function [27]-[29], but the latter concluded that the original blowfish algorithm was still more compact and more secure. Some optimizations on blowfish concentrated on the modification of the number of rounds to increase speed and ultimately enhance security [30] Although there was a mention of a minimum of five rounds, there was no minimum number set [31]. However, the recommended number of rounds is 16-8. Optimization on the reduction of S-boxes from four to two was also applied to increase the speed [32]. Several researchers have attempted to extend the block size of blowfish to 128-bit [18], [33] [34], but results indicate an increase in time complexity and need for higher memory.

From here, a modified Blowfish algorithm is proposed accepting block size of 128-bits and a key size of 128 bit to meet AES requirements that would exhibit speed and simplicity comparable to original blowfish. Iteration count and the number of S-boxes was reduced to achieve speed during key generation and replacement. For on-the-fly subkey calculation, there was an added S-box derivation technique.

This research aims to propose a modified blowfish algorithm that uses 128-bit block size and 128-bit key with a reduced number of iterations while maintaining the original structure of blowfish for a smooth migration. This study sought answers to the following objectives: to compare the execution of the modified Blowfish, Blowfish and Twofish algorithm regarding speed based on encryption, and decryption; and to compare the performance of the modified algorithm and blowfish regarding security using avalanche effect.

This study will modify Blowfish to make use of 128-bit block size and key, a design criterion set during the AES competition. The change in block size would allow encryption of file with lesser chances of having duplicate blocks. The original structure of Blowfish will still be used but will reduce the number of S-boxes from four to two to provide less memory consumption. The modification made on the algorithm ensures compatibility to the original version of Blowfish. A derivation technique will be introduced to reduce time in the generation of the key and to reduce symmetry in the s-boxes.

2. RESEARCH METHOD

2.1. Research Design

The key expansion will still convert the 128-bit key length into several subkey arrays. The modified Blowfish reduced the size from the previous 4168 bytes to 2128 bytes. These keys may also be generated separately and stored before any data encryption or decryption occurs. The P-array will now consist of 20 (P1, P2…P20) 32-bit subkeys. The four S-Boxes will still consist of 256 individual entries comprising 32-bits each (S1 - 0…255, S2 - 0…255). In the modified key expansion scheme, the total number of iterations will be reduced to 266 to generate all required subkeys.

Calculation of the subkeys are done using the same Blowfish algorithm, but the algorithm reduced the size to two S-boxes:

1. Initialization of the P-array followed by the four S-boxes was executed using constant strings that consist of the hexadecimal digits of pi.

Modified Blowfish Algorithm (Theda Flare G. Quilala)
2. P1 is XORed with the first 32 bits of the key, P2 is XORed with the second 32-bits, continuously until all bits of the key is exhausted up to P20. Repeat the cycle until the whole P-array has been XORed against the key bits.

3. The Blowfish algorithm is used to encrypt an all-zero string using the subkeys described in the previous steps (1 and 2).

4. The outcome of step 3 substituted P1 and P2.

5. Using the Blowfish algorithm, encrypt the output of step 3 using the revised subkeys.

6. Results obtained in step 5 replaced P3 and P4.

This process is continuously repeated replacing all entries of the P array, followed by the two S-boxes with the output of the continually varying Blowfish algorithm. Figure 2 shows the new process of encryption of the modified blowfish algorithm. The structure of the original blowfish algorithm is still adopted, but the modified Blowfish reduce the number of iterations to 8.

---

**Figure 2. Blowfish proposed modification using 128-bit**

The difference lies in the input block. The input block changes to 128-bit and will be split into two 64-bit equal segments LE0, RE0. Second, the first segment 64-bit block (LE0) is XORed to the first entry in the P-array (P1, P11) with two 32-bit entries. Third, input the two 32-bit data obtained to the F-function. The output from the F-function will then be XORed with the second segment (RE0) of the plaintext. Then, swap LE0 and RE0. This cycle will continue up to the eighth round. After the eighth round, exchange LE8 and RE8 reversing the last swap. Then, RE8 is XORed to P-array (P9, P19) and LE8 is XORed to P-array (P10, P20). Finally, we recombine LE9 and RE9 to get the ciphertext. The decryption process is the reverse of the encryption process.

Figure 3 also shows the details of the construction of the new F-function in the modified blowfish. The F-function now accepts a 64-bit data stream and will be divided into eight 8-bits where a is the first 8 bits, b is the second 8 bits, up to the last 8 bits. Transform each 8-bit data bits into a 32-bit data. The first four 8-bit data stream utilizes the first S-box while the next four 8-bit data stream uses the second S-box. The output from the S-boxes are then XORed or added to obtain the final 32-bit value per S-box and then concatenated to obtain the 64-bit output as shown in the equation:
The S-boxes are derived at runtime from S-box 1 by a simple rotation by one position of either the input or the output or either by left or right. Below defined the details of the derivation process:

\[
F(LE0) = ((SI(a) + SI(b) \ll 1 \bmod 2^{32}) \oplus SI(c) \gg 1) + SI(d \ll 1) \bmod 2^{32} \mid ((S2(e) + S2(f) \ll 1 \bmod 2^{32}) \oplus S2(g) \gg 1) + S2(a \ll 1) \bmod 2^{32}
\]

Figure 3. Modified F-function

The researcher changed the structure of the F-function as can be seen from the equation above.

2.2. Metrics
Performance metrics for the analysis of the performance of the algorithm is time (milliseconds), throughput (Mb/sec), and avalanche effect (%). Below is the description of the evaluation parameters:
1. Key Generation time: The amount of time needed to generate the subkeys
2. Encryption time: The amount of time necessary to change the plaintext to equivalent ciphertext.
3. Decryption time: The amount of time needed to change ciphertext to plaintext.
4. Throughput: Throughput indicates the speed of encryption. The size of plaintext divided by the total time is the calculation for throughput.
5. Avalanche effect: Refers to the characteristic where a minimal change in the input text results in a significant change of the output sometimes referred to as diffusion, reflecting the cryptographic strength of a cryptographic algorithm. Avalanche effect is calculated using hamming distance which is a measure of difference. It is the XOR calculation bit by bit of the ASCII value.

2.3. Research Procedure
Experimentation was done using different file sizes ranging from 10kb to 1000kb. The average time is computed using 20 trials of each file size. During the experimentation, the researcher used Intel® Core ™2 Quad CPU Q6600 @2.40 GHz with 4G RAM. The file and key used for all testing done were the same.

3. RESULTS AND ANALYSIS
As seen in Table 1, the average key generation time for Blowfish is 21.65ms while the modified Blowfish algorithm is 26.99ms. The number of rounds was reduced to 8 rounds to compensate for the time difference. Although the original algorithm is still faster than the modified version, the modified algorithm already uses 128-bit block size. Extending the block size to 128-bit lessen the chances of having duplicate blocks that may lead to the leak of information thus increases security. For a block cipher which uses 64-bit blocks, the threshold is about 32 gigabytes (232 blocks of 8 bytes). If a 1TB drive is encrypted, there exist 32 duplicated cipher blocks.
Table 1. Key generation time comparison in milliseconds using different file sizes

<table>
<thead>
<tr>
<th>Input Size (kb)</th>
<th>BA</th>
<th>MBA</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>21.00</td>
<td>28.55</td>
</tr>
<tr>
<td>20</td>
<td>23.30</td>
<td>25.60</td>
</tr>
<tr>
<td>50</td>
<td>22.10</td>
<td>25.05</td>
</tr>
<tr>
<td>100</td>
<td>22.30</td>
<td>26.85</td>
</tr>
<tr>
<td>200</td>
<td>22.15</td>
<td>27.60</td>
</tr>
<tr>
<td>500</td>
<td>20.30</td>
<td>26.00</td>
</tr>
<tr>
<td>1000</td>
<td>21.00</td>
<td>27.70</td>
</tr>
<tr>
<td>Average (ms)</td>
<td>21.65</td>
<td>26.99</td>
</tr>
</tbody>
</table>

Comparing the key expansion of Blowfish and modified blowfish, Blowfish uses 4168 bytes while the modified Blowfish uses 2128 bytes. Bytes used for the modified blowfish is lower which would make it suitable for small devices with limited memory size. The difference is due to the differences in P-array and S-boxes used. Table 2 shows the comparison.

Table 2. Key generation comparison of bytes used

<table>
<thead>
<tr>
<th>BA</th>
<th>MBA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uses 4168 bytes, (computed as 4 bytes * 4 s-boxes * 256 entries each plus 4 bytes * 18 Parray entries)</td>
<td>Uses 2128 bytes (computed as 4 bytes * 2 s-boxes * 256 entries plus 4 bytes * 20 Parray entries)</td>
</tr>
<tr>
<td>The P-array consist of 18 (P1, P2...P18) 32-bit subkeys.</td>
<td>The P-array consist of 20 (P1, P2...P20) 32-bit subkeys</td>
</tr>
<tr>
<td>Four S-Boxes consists of 256 individual entries comprised of 32-bits each (S1 - 0...255, S2 - 0...255, S3 - 0...255, S4 - 0...255)</td>
<td>Two S-Boxes consists of 256 individual entries consisting of 32-bits each (S1 - 0...255, S2 - 0...255)</td>
</tr>
<tr>
<td>Number of iterations is 521 to generate all required subkeys</td>
<td>Number of iterations is 266 to make all the necessary subkeys</td>
</tr>
</tbody>
</table>

For the encryption and decryption time, as can be observed in Table 3 and Table 4, BA is still faster among the three algorithms having 1297.76ms and 2176.59ms average encryption and decryption time respectively and an encryption throughput value of 234.91Mb/sec and decryption throughput of 120.20Mb/sec. The difference in time is due to the difference in block size. The modified Blowfish is compared to Twofish, an encryption algorithm based on Blowfish because it also accepts 128-bit block size. Comparing modified blowfish and Twofish, each accepting 128-bit block size and key, the modified blowfish gained faster average encryption and decryption time at 1651.83ms and 2765.04ms compared to 2418.08ms and 4002.70ms for Twofish. Encryption throughput for modified Blowfish is at 182.92Mb/sec and decryption throughput is 89.65Mb/sec. For Twofish, encryption throughput is 124.32Mb/sec, and decryption throughput is 63.56. From the results, the modified Blowfish has higher throughput value meaning the modified algorithm has greater efficiency than Twofish.

Table 3. Encryption time comparison in milliseconds using different file sizes

<table>
<thead>
<tr>
<th>Input Size (kb)</th>
<th>BA</th>
<th>MBA</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>51.10</td>
<td>69.40</td>
<td>93.30</td>
</tr>
<tr>
<td>20</td>
<td>93.60</td>
<td>119.75</td>
<td>174.55</td>
</tr>
<tr>
<td>50</td>
<td>212.65</td>
<td>272.30</td>
<td>438.85</td>
</tr>
<tr>
<td>100</td>
<td>413.55</td>
<td>527.15</td>
<td>779.95</td>
</tr>
<tr>
<td>200</td>
<td>809.00</td>
<td>1030.45</td>
<td>1517.15</td>
</tr>
<tr>
<td>500</td>
<td>2007.55</td>
<td>2556.85</td>
<td>3737.90</td>
</tr>
<tr>
<td>1000</td>
<td>3999.40</td>
<td>5085.40</td>
<td>7423.25</td>
</tr>
<tr>
<td>Average</td>
<td>1297.76</td>
<td>1651.83</td>
<td>2418.08</td>
</tr>
<tr>
<td>Throughput</td>
<td>234.91</td>
<td>182.92</td>
<td>124.32</td>
</tr>
</tbody>
</table>

Table 4. Decryption time comparison in milliseconds using different file sizes

<table>
<thead>
<tr>
<th>Input Size (kb)</th>
<th>BA</th>
<th>MBA</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>91.80</td>
<td>153.75</td>
<td>179.85</td>
</tr>
<tr>
<td>20</td>
<td>174.00</td>
<td>255.30</td>
<td>336.45</td>
</tr>
<tr>
<td>50</td>
<td>402.90</td>
<td>562.05</td>
<td>833.45</td>
</tr>
<tr>
<td>100</td>
<td>827.00</td>
<td>1038.80</td>
<td>1538.50</td>
</tr>
<tr>
<td>200</td>
<td>1624.15</td>
<td>2047.55</td>
<td>3007.95</td>
</tr>
<tr>
<td>500</td>
<td>4045.35</td>
<td>5114.10</td>
<td>7413.50</td>
</tr>
<tr>
<td>1000</td>
<td>8070.90</td>
<td>10183.75</td>
<td>14709.20</td>
</tr>
<tr>
<td>Average</td>
<td>2176.59</td>
<td>2765.04</td>
<td>4002.70</td>
</tr>
<tr>
<td>Throughput</td>
<td>120.20</td>
<td>89.65</td>
<td>63.56</td>
</tr>
</tbody>
</table>
A suitable feature of every encryption algorithm is that a minimal change in the key should produce a significant difference in the ciphertext. The avalanche effect of the modified Blowfish is compared to Blowfish to ensure that the diffusion of the algorithm was not affected by the removal of the two S-boxes and to ensure that the derivation technique has removed the symmetry between the S-boxes.

The plaintext message used was “The quick brown fox jumps over the lazy dog” and varying 1 bit of the key. The following keys were used: FEDCBA9876543210, FEDCBA9876543211, FEDCBA9876543212, FEDCBA9876543213, and FEDCBA9876543214. Figure 4 shows the avalanche percentage. Blowfish has 47.14% avalanche, and the modified Blowfish is at 52.86%. The higher the avalanche percentage, the higher will be the security [35], this means that the modified algorithm even had a better avalanche, thus better security.

The efficiency of the modified algorithm was also evaluated using CrypTool 1.4.40 [36], a free and comprehensive e-learning program including cryptography and cryptanalysis. For this test, a file of 104Kb was encrypted using a 128-bit key. Blowfish accepts variable key sizes ranging from 32-448 bits. The modified Blowfish algorithm set the smallest key size to 128-bit. A 128-bit key has a complexity of \(2^{128}\) or \(3.40 \times 10^{38}\). For an encryption algorithm with a 128-bit key, a brute force attack will take \(5 \times 10^{25}\) years which makes brute force utterly impossible. See Figure 5.

Entropy is a measure of randomness or uncertainty in the information usually referred to as confusion. Based from CrypTool, the document contains all 256 possible byte values. The entropy of the text is 7.99 with a maximum entropy value of 8.0. Encryption desires high randomness so that there is less or no dependency between key and ciphertext to make it difficult to guess by an attacker. Frequency test is a frequency analysis which studies the frequency of characters in a ciphertext.

4. CONCLUSION

This paper proposed a modified blowfish algorithm that uses 128-bit block size and a 128-bit key, maintaining the original structure of blowfish for a smooth migration with a reduced number of S-boxes to provide less memory consumption. Results show that the modified algorithm design continues to offer sufficient avalanche effect as the original with less storage requirement for the P-array and S-boxes.
Although the modified algorithm is a little slow compared to Blowfish, the increase in the input block size is the reason. Compared to Twofish, a related algorithm to Blowfish, the modified Blowfish is faster and has a better throughput efficiency. Based from Cryptool, the modified Blowfish algorithm also passed the entropy and frequency test. In a 128-bit key, a brute force attack will take 5×10^25 years which makes brute force utterly impossible. Studies involving any data encryption or file encryption process can use the algorithm. For future works, other researchers may study hardware optimization implementation of the modified algorithm.

REFERENCES


