International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 04 Issue: 02 | Feb -2017 www.irjet.net p-ISSN: 2395-0072

# A REVIEW TOWARDS VARIOUS HASH ALGORITHMS AND THEIR COMPARATIVE ANALYSIS

**Umesh Gandhi[1], Mrs. Poonam Sinha[2], Ms. Rachna Kulhare[3]**
*[1]Department of I.I., UIT, Barkatullah University, Bhopal (M.P.), Inda*
*[2]( Professor, Department of Electronics, UIT, Barkatullah University, Bhopal (M.P.), India)*
*[3](Asst. Professor, Department of I.T., UIT, Barkatullah University, Bhopal (M.P.), India)*

---

***Abstract:*** *Today, with the development of modern technologies, there is always need to upgrade the existing algorithm or replaces to cope up with the latest requirement. Hash algorithms are the one which are used to ensure the integrity of transmitted or stored data. Many of the algorithms have been designed to generate hash in which some were rejected and some were become standard. MD and SHA family are the most popular standard used to generate hash. The aim of this paper is to compare these standard integrity algorithms and also gives a focus of light on the problems they are currently facing.*

***Keywords:*** Computer Security, SHA, Hash, Message Digest

## 1. INTRODUCTION

A cryptographic hash function is an algorithm that generates fixed sized bit string of an arbitrary block of data. This fixed size bit string is a hash value is designed in such a way that if any minor change in the data either accidentally or intentionally, change the hash value with a very high probability. The data to be encoded generally called message and the hash value is called message digest or simply a digest. An ideal hash algorithm must have the following four properties:

a. It should be easy to calculate digest for any given message.

b. For a given hash, it should be impractical to generate a message.

c. It should impractical to perform changes in a message without changing hash value.

d. It should be impractical to search two messages with same digest.

Cryptographic hash functions attach the digest with the message and transmit to the other end, at the other end the attached digest is separated and the digest of the received message is calculated again. Now, if the recently generated digested is same as received digest it means no change in the message during transmission process but if digest are not same it means changes happen during transmission process. This change may be cause due to noise or may be done intentionally by intruder.

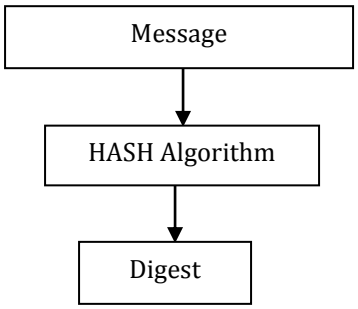


Figure 1 (a) Digest Generation

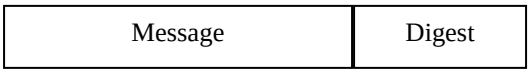


Figure 1 (b) Message preparation for transmission by concatenating message with digest.

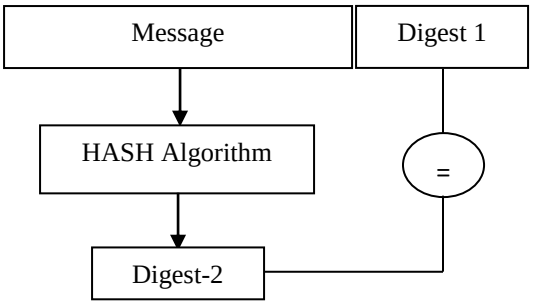


Figure 1 (c) Integrity assurance at receiving end

Figure 1 (a), (b) & (c) shows the above discussed process for the better understanding. Figure 1 (a) shows the

generation of digest of a message by hash algorithm. This process is done at sender end.

Next at the sender end the generated digest is append with the message and transmit to the other end shown in Figure 1 (b) & last Figure 1 (c) shows the process at receiver end. If the digest 1 and digest 2 are same it means receiving message is same that the sender send but if both the digest are different it means message is changed some where during transmission.

Number of algorithms have been proposed to ensure the integrity, among them MD family and SHA family gain most of the popularity.

MD family consist MD-2, MD-4 and MD-5 where as on the other end SHA family includes SHA-0, SHA-1 and SHA-2.

MD-2 was first published in 1989 and was designed by Ronald Rivest. It generates a fixed size message digest of 128 bits. Although, it is not consider enough secure after found of several attack on it. Latest, in 2009 a collision attack was found at time complexity of $2^{63.3}$.

MD-4 was published in 1990 again by Ronald Rivest. It also generates a same fixe size digest of 128 bits. MD-4 again found weak security as collision found on MD-4 in less than 2 hash operations.

MD-5 was published in 1992 again by Ronald Rivest. Again it generates a same fixed 128 bit digest. But again several successful attacks found on MD-5. Latest, in 2010 Tao Xie and Dengguo Feng published MD-5 collision.

SHA-0 on the other end belongs to SHA Family, SHA-0 generates 160 bit digest and was published in 1993. Very soon again SHA-0 was found weak as number of several successful attack found. Latest, in 2008 a boomerang attack was found on SHA-0 at complexity of $2^{33.6}$. It takes hardly one hour on simple personal computer to find collision.

SHA-1 is most popular hash algorithm among all the proposed algorithms. It generates a message digest of fixed length 160 bit. But after gaining lot of popularity SHA-1 also found mathematically weak and a collision attack with time complexity between $2^{60.3}$ and $2^{65.3}$ found.

SHA-2 on the other end is safe till now of getting any successful attack on it, but still it does not gain much popularity due to its inefficiency in terms time as compared to other hash algorithms. SHA -2 is available with different variations of fixed size digest 224, 256, 384, 512. Although many attacks have been proposed on SHA-2 but none of them succeed completely.

From the above discussion, it can be concluded that all the existing standard algorithms are either found breakable or inefficient to use.

This paper has studied above discussed algorithms in detail and many other such algorithms published to discuss the above problem and proposed their unique solutions.

A series of SHA algorithms has been developed by the National Institute of Standards and Technology and published as Federal Information Processing Standards (FIPS). This Standard specifies three secure hash algorithms - SHA-0, SHA-1, SHA-2, for computing a condensed representation of electronic data (message).

## 2. LITERATURE SURVEY

There are lot of algorithms have been proposed since now to ensure the integrity. Many of the them comes out as a standard but none of them stand strong against collision attack. All of them either breakable or those which are not breakable till now are not time efficient.
Many researches have being done afterwards to overcome the above discussed problem. Some of the latest researches have been discussed in this section.

Mirvaziri et al [4], introduced a new hash algorithm in 2007 that generates a message digest of 160 bit. This algorithm generates a hash by combining both MD-5 and SHA-1, to gain the best properties of both the algorithms. Message are first divided into number of sub blocks of size 512 bit and each 512 bits first passes to SHA-1 algorithm and then passes to MD-5 algorithms. The main issue with this algorithm is that the time taken by this algorithm is sum of time taken by SHA-1 and MD-5 individually.

Thulasimani et al [3], present there paper in 2009, and proposed a unique algorithm have fixed size digest of 192 bits and named it SHA -192. They keep the internal structure of SHA -192 same as SHA-1 except it uses six chaining variable of 32 bits instead of five chaining variable of 32 bits. According to birthday attack, it requires $2^{96}$ combinations, which is practically impossible for a super computer to solve in reasonable time.

Gupta et al [2], presented a modification of research discussed in paper [4]. In this the author merges SHA -192 and MD-5 algorithm. Again the problem in this research is same i.e. time required to generate the digest. It required time which is sum of time taken by SHA-192 and MD-5 algorithm.

Wang et al [1], proposed a new concept latest in 2015, which generates a variable size digest in order to meet the requirement of latest technology. They modified MD -5 algorithms in such a way that it generates a variable size digest.

## 3. Comparatively Analysis of SHA-1, MD-5 and Modified MD-5 [1]

This section shows the implementation results of above discussed three algorithms SHA-1, MD-5 and modified MD-5 algorithms and also compare them on the basis of timing, internal robustness and size. Authors have implemented these algorithms in DotNet framework with using of C# language and the configuration used during the comparative analysis is Intel Core I5 2.40 GHz, 4 GB of RAM and Windows 7 Home basic, service pack 2, 64 bit operating system.

**Timing Analysis:** Timing is one of the important factors in evaluation of performance of any algorithm. An algorithm that take more time to generate the message digest will considered less preferable than other which generate fast message digest. Authors have implemented both the algorithms and evaluated the time taken by these algorithms to generate the message digest and after testing on more than 50 files of each size the average time of the experimental results is shown in Table 1.

Table 1 Timing Comparison between SHA-1, MD-5 and Modified MD-5 [1] algorithm

| File Size in KB | Algorithms (Time in Seconds) | | |
|---|---|---|---|
| | SHA-1 | MD-5 | Modified MD-5[1] |
| 5 KB | 0.174 | 0.128 | 0.140 |
| 10 KB | 0.525 | 0.423 | 0.492 |
| 15 KB | 1.156 | 1.054 | 1.121 |
| 20 KB | 1.982 | 1.921 | 1.935 |

The graphical analysis of Table 1 is shown in Figure 2. In this blue bar represent MD-5, red bar represent Modified MD-5[1] and green bar represent SHA-1.

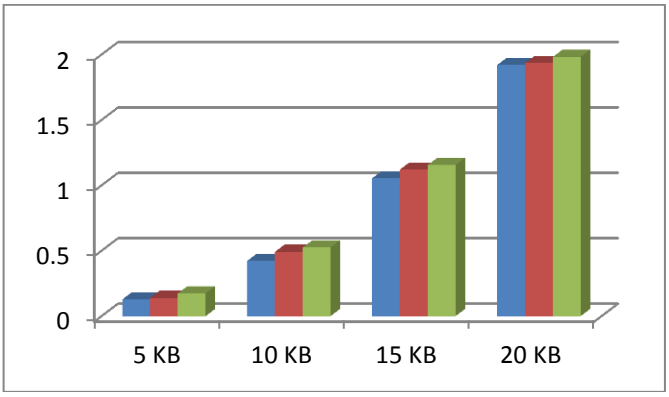


Figure 2 Timing Comparison between SHA-1, MD-5 and Modified MD-5 [1] algorithm

From the Figure 2, it is clearly seen that MD-5 takes less time as compared to other two algorithms and SHA-1 takes maximum time compare to other two algorithms.

**Security Analysis:** Another important factor in designing an algorithm is security. Whether the algorithm is secure or not is always a question. It is always a point of discussion that how to measure a security of any algorithm. As such no cryptanalysis attack has been found on Modified MD-5 [1], but because of only this reason nobody can say that Modified MD-5 is secure. So to check the security Modified MD-5 algorithm avalanche effect of all three algorithms is calculated. Avalanche effect is one parameter which can be used to check the internal strength of any cryptographic algorithm. According to avalanche effect, change in a single bit closer to avalanche value is considered more preferable. After testing on more than 50 different files authors have concluded the result shown in Table 2.

Table 2. Avalanche effect of SHA-1, MD-5 and Modified MD-5 [1] algorithm

| Algorithm | Avalanche Effect | |
|---|---|---|
| | Bits Changed | Percentage |
| MD-5 | 58/128 | 45.31% |
| Modified MD-5 [1] | 52/128 | 40.63% |
| SHA-1 | 73/160 | 45.63% |

The graphical analysis of avalanche effect of SHA-1, MD-5 and Modified MD-5 [1] algorithm is shown in Figure 3. From this it can be easily concluded that the avalanche effect of SHA-1 is more than other two algorithms. Hence the internal structure of SHA-1 is more secured than other 2.

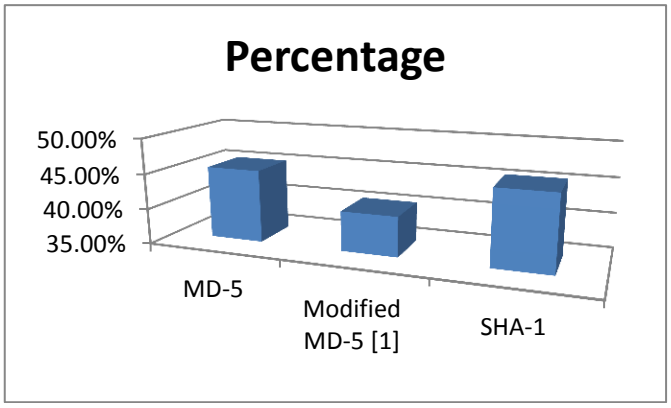

Figure 2 Avalanche effects of SHA-1, MD-5 and Modified MD-5 [1] algorithm

**Space Analysis:** Another parameter to evaluate the performance of all the three algorithms is space. As discussed SHA -1 uses five chaining variable of 32 bit which actually store the hash value, but on the other end MD-5 uses four chaining variable of 32 bit while modified MD-5 [1] of n bit digest uses n/32 chaining variable. If the value of n is more chaining variable is more. Therefore, SHA needs more space than MD-5 whereas Modified MD-5[1] generates variable size digest hence if value of n is more than it required more space and if value of n is low than it required low space.

**Analysis of Hash Code:** If there are n bit digest than it has $2^n$ distinct hash value exist for the n bit digest. But there is no constraint on message, messages are more than $2^n$. So, it is impossible to design algorithms which have no collision. Now the question comes out, how to make an algorithm that makes difficult to find that collision. So if n is more than number of combination of digest become more hence difficult to generate two messages having same digest. According to birthday attack, it requires $2^{n/2}$ combination to find collision in any hash hence SHA- required $2^{80}$ combinations and MD-5 required $2^{64}$ combinations. But this paper already discussed that these algorithms are proven breakable far before this combinations.

## 4. CONCLUSION

In this paper, authors have discussed details of various existing hash integrity algorithms i.e. SHA-1, MD-5 and other latest algorithms. All the existing standard algorithms are either proven breakable or the rest are proven inefficient in term of timing. Also this paper have implemented latest modified MD-5[1], MD-5 and SHA-1 algorithm and found that SHA-1 has more secure internal structure tan other two but inefficient in terms of timing as compared to other two algorithms. So it can be concluded that there is a need of development of hash algorithm which should be enough secure and efficient than the existing algorithms as they are not sufficient to meet the requirement of latest technologies and security concern.

## REFERENCES

[1] Meng-jiao WANG, Yong-zhen LI, "Hash Function with Variable Output Length" 2015 International Conference on Network and Information Systems for Computers, IEEE-2015
[2] Gupta G., Sharma, S. "Enhanced SHA-192 Algorithm with Larger Bit Difference" Published in IEEE International Conference on Communication Systems and Network Technologies (CSNT), 6-8 April 2013 Page(s):152 - 156 Print ISBN:978-1-4673-5603-9
[3] L.Thulasimani and M.Madheswaran "Security and Robustness Enhancement of Existing Hash Algorithm" IEEE International Conference on Signal Processing Systems 2009.
[4] A new Hash Function Based on Combination of Existing Digest Algorithms pub 2007.
[5] The Collision Rate Tests of Two Known Message Digest Algorithms 2009.
[6] Harshvardhan Tiwari. A Secure Hash Function MD-192 with Modified Message Expansion" Vol. 7 No. 2 February 2010 International Journal of Computer Science and Information Security.
[7] Marc Stevens hash clash - Framework for MD5 & SHA-1 Differential Path Construction and Chosen-Prefix Collisions for MD5
[8] X. Wang, H. Yu and Y.L. Yin, "Efficient Collision Search Attacks on SHA-0",(Pub 2005)
[9] K. Matusiewicz and J. Pieprzyk, "Finding good differential patterns attacks on SHA-1", (Pub 2004), Available: http://eprint.iacr.org/2004/364.pdf
[10] William Stallings, "Cryptography and Network Security: Principles and Practice. Third edition, Prentice Hall.2003.
[11] Florent Chabaud, Antoine Joux, "Differential collisions in SHA-0," Advances in Cryptology-CRYPTO'98, LNCS 1462, Springer-Verlag, 1998.
[12] http://en.wikipedia.org/wiki/Secure_Hash_Algorithm
[13] Ricardo Chaves, Georgi Kuzmanov, Leonel Sousa, and Stamatis Vassiliadis " Cost-Efficient SHA Hardware Accelerators" IEEE transactions on very large scale integration (VLSI)Systems, VOL. 16, NO. 8, AUGUST 2008