Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 6.199

IJCSMC, Vol. 8, Issue. 6, June 2019, pg.147 – 152

A Comparative Study of Hash Algorithms in Cryptography

Prashant P. Pittalia
Department of Computer Science, S. P. University, India
prashantppittalia@yahoo.com

Abstract— In Network Security and Cryptography, hash functions play a very important role. To check the Integrity, Authenticity of information or data transmitting between the Seder and receiver party hash algorithms are used. Digital signature also uses the hash algorithm. Hash function used for key generation in Symmetric and Asymmetric Key Cryptosystems. Different algorithms provide different level of security depending on how difficult is to break them. The most well-known hash algorithms are SHA-1, SHA-2, SHA-3, MD4, MD5 and Whirlpool etc. This paper discusses importance of hash functions, description about various well known hash functions, and comparative analysis of various hash algorithms.

Keywords— SHA 1, MD5, Integrity, Hash Algorithm, Cryptography

I. INTRODUCTION

A hash function is a function that takes an arbitrary amount of input and produces an output of fixed size. The standard hash function serves as a basis for the discussion of Cryptographic Hash Functions. Currently in MD5 and variants of SHA are very commonly used hash algorithms. In a very secure system the data integrity is a crucial part. By cryptographic hash function system users are able to generate a message digest to detect the unauthorized changes in the files. It is especially important when critical system and sensitive databases. To verify the source of data, Hash functions can be combined with other standard cryptographic methods. When hashing algorithms are combined with encryption, they produce special message digests that identify the source of the data; these special digests are called Message Authentication Codes. HMAC is a standard algorithm currently used. The HMAC algorithm provides verification of the source of data, and also prevents against attacks like a replay attack. Hash functions are mathematical computations that take an arbitrary amount of data as input and produce an output of fixed size. The output is always the same when given the same input. The inputs to a hash function are called messages, and the outputs are referred to as message digests. Nearly any piece of data can be defined as a message, including character strings, binary files and TCP packets. All hash functions have the property that it is impossible to convert the output into input. The attacker has no way of

determining what the original message was by being given the digest. This property makes this hash functions a one-way function, it is not possible to deduce the input for a given output. Cryptographic hash functions have another property that it is very difficult to find two different messages that produce the same message digest. To provide the data integrity and data authentication, if a message digest of any information is changes, then the file itself has changed.

II. HASHING ALGORITHMS

In today's Internet environment, MD5 and SHA1 are two primarily cryptographic hash functions. MD5 stands for "Message Digest 5" because it is the fifth revision of a message digest algorithm devised by R.L. Rivest of RSA Laboratories (RSA Laboratories). The early revisions of this algorithm were published prior to 1989, and the most recent revision of the algorithm was published in 1991. SHA1 stands for "Secure Hash Algorithm 1"; it is the first revision of a hash algorithm developed by the National Security Agency. The algorithm was first published in 1995. SHA1 supports messages of any length less than 264 bits as input, and produces a 160-bit digest. If input data for the hash function larger than 264 bits in length the simplest solution would be to divide the large messages into smaller messages. There are other variations of SHA1 which produce longer digests, SHA-256, SHA-512. They produce digests of 256 bits and 512 bits, respectively. The SHA1 and MD5 algorithms are considered secure because there are no known techniques to find collisions, except via brute force. In a brute force attack random inputs are tried, storing the results until a collision is found. SHA1 is more secure than MD5, but it is more costly to compute a message digest using SHA1 than MD5. One-way Hash Function (OWHF) defined by Merkle [1] is a hash function H, that satisfies the following requirements:

1. H can be applied to Block of data of any length. (Any length means size of Block must be greater

Than size of Digest we conclude at the end).

2. H produces a fixed-length output i.e., Message Digest.
3. Given H and x (any given input), it is easy to computer Message Digest H(x).
4. Given H and H(x), it is computationally infeasible to find x.
5. Given H and H(x), it is computationally infeasible to find x and x' such that H(x) = H (x')

for message authentication and digital signature the first three requirements are must for practical applications. The fourth states that it is easy to generate a message code of given message but very difficult to generate a message back from given digest. The fifth requirement guarantees that an alternative message hashing to the same code as a given message cannot be found.

A. MD5

The MD5 message-digest algorithm is a widely used hash function producing a 128-bit hash value. Although MD5 was initially designed to be used as a cryptographic hash function, it has been found to suffer from extensive vulnerabilities. It can still be used as a checksum to verify data integrity, but only against unintentional corruption. It remains suitable for other non-cryptographic purposes, for example for determining the partition for a particular key in a partitioned database.[2] MD5 digests have been used in the software to provide some assurance that a transferred file has arrived without any changes during transmission. File servers provide a pre-computed MD5 checksum for the files, so that a user can compare the checksum of the downloaded file to it. MD5 processes a variable-length message into a fixed-length output of 128 bits. The input message is broken up into chunks of 512-bit blocks; the message is padded so that its length is divisible by 512. This is followed by as many zeros as are required to bring the length of the message up to 64 bits less than a multiple of 512. The remaining bits are filled up with 64 bits representing the length of the original message,

modulo 264. The main MD5 algorithm operates on a 128-bit state, divided into four 32-bit words, denoted A, B, C, and D. These are initialized to certain fixed constants. The main algorithm then uses each 512-bit message block in turn to modify the state. The processing of a message block consists of four similar stages, termed rounds; each round is composed of 16 similar operations based on a non-linear function F, modular addition, and left rotation.

B. Whirlpool

It was designed by Vincent Rijmen (co-creator of the Advanced Encryption Standard) and Paulo S. L. M. Barreto, who first described it in 2000. The hash has been recommended by the NESSIE project. It has also been adopted by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) as part of the joint ISO/IEC 10118-3 international standard. The Sub Bytes operation applies a non-linear permutation (the S-box) to each byte of the state independently. The 8-bit S-box is composed of 3 smaller 4-bit S-boxes. The Shift Columns operation cyclically shifts each byte in each column of the state. Column j has its bytes shifted downwards by j positions. The Mix Rows operation is a right-multiplication of each row by an 8×8 matrix over {\display style GF ({2^ {8}})} {\display style GF ({2^ {8}})}. The AddRoundKey operation uses bitwise xor to add a key calculated by the key schedule to the current state. The key schedule is identical to the encryption itself, except the AddRoundKey function is replaced by an AddRoundConstant function that adds a predetermined constant in each round. The block cipher W consists of an 8×8 state matrix {\displaystyle S} S of bytes, for a total of 512 bits. The encryption process consists of updating the state with four round functions over 10 rounds. The four round functions are SubBytes (SB), ShiftColumns (SC), MixRows (MR) and AddRoundKey (AK). During each round the new state is computed as {\displaystyle S=AK\circ MR\circ SC\circ SB(S)} S=AK\circ MR\circ SC\circ SB(S).

C. SHA- 1

In cryptography, SHA-1 (Secure Hash Algorithm 1) is a cryptographic hash function which takes an input and produces a 160-bit (20-byte) hash value known as a message digest. All major web browser vendors ceased acceptance of SHA-1 SSL certificates in 2017. [3][4][5] Several widely used security applications and protocols, TLS and SSL, PGP, SSH, S/MIME, and IPSec may use SHA-1. It is also used to identify that during transmission of information from sender to receiver is any changes occur or not.

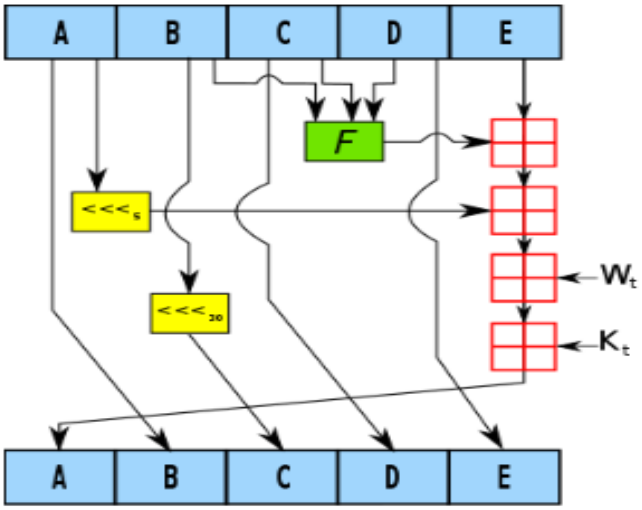

Fig. 1 One SHA-1 Iteration[7]

One iteration within the SHA-1 compression function:
A, B, C, D and E are 32-bit words of the state; F is a nonlinear function that varies; left shiftn denotes a left bit rotation by n places; n varies for each operation; Wt is the expanded message word of round t; Kt is the round constant of round t; Addition denotes addition modulo 232.

D. SHA - 2

SHA-2 variants are SHA-224, SHA-256, SHA-384, SHA-512, SHA-512/224, and SHA-512/256. SHA-2 includes
A significant number of changes from its predecessor, SHA-1. SHA-2 consists of a set of six hash functions with digests that are 224, 256, 384 or 512 bits. SHA-256 and SHA-512 are novel hash functions computed with 32-bit and 64-bit words, respectively. They use different shift amounts and additive constants, but their structures are otherwise virtually identical, differing only in the number of rounds. The SHA-2 functions were not easily adopted in compare to SHA-1. Reasons was that the lack of support for SHA-2 on systems running Windows XP SP2 or older and a lack of perceived urgency since SHA-1 collisions had not yet been found. From 2017, Microsoft announced that Internet Explorer and Edge would stop honouring public SHA-1-signed TLS certificates. From 2014-15 Google Chrome team announced a plan to make their web browser gradually stop honouring SHA-1-dependent TLS certificates. In January 2016, Mozilla disabled SHA-1 but had to re-enable it temporarily via a Firefox update, after problems with web-based user interfaces of some router models and security appliances.

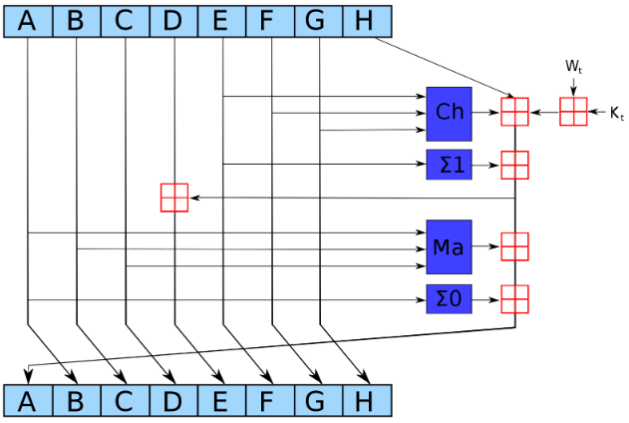

Fig. 2 One SHA-2 Iteration [8]

E. SHA - 3

It is the latest standard released by NIST on August 5, 2015. SHA-3 is internally different from the MD5-like structure of SHA-1 and SHA-2. NIST does not currently plan to withdraw SHA-2 or remove it from the revised Secure Hash Standard. The purpose of SHA-3 is that it can be directly substituted for SHA-2 in current applications if necessary, and to significantly improve the robustness of NIST's overall hash algorithm toolkit. [6]

To ensure the message can be evenly divided into r-bit blocks, padding is required. SHA-3 uses the pattern 10*1 in its padding function: a 1 bit, followed by zero or more 0 bits (maximum r − 1) and a final 1 bit.
The block transformation f, which is Keccak-f for SHA-3, is a permutation that uses XOR, AND and NOT operations, and is designed for easy implementation in both software and hardware. In 2016 the same team that made the SHA-3 functions and the Keccak algorithm introduced faster reduced-rounds (reduced to 12 and 14 rounds, from the 24 in SHA-3) alternatives which can exploit the availability of parallel execution because of using tree hashing: Kangaroo Twelve and Marsupilami Fourteen.

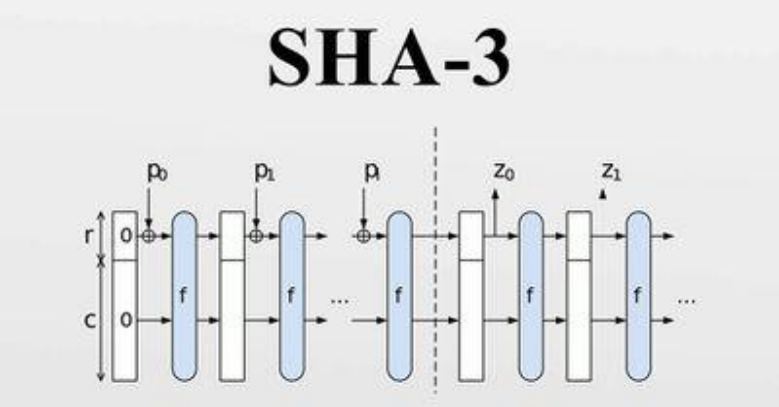

Fig. 3 One SHA-3 Iteration [9]

The sponge construction for hash functions. Pi are input, Zi are hashed output. The unused "capacity" c should be twice the desired resistance to collision or preimage attacks.

III. COMPARISON OF HASH ALGORITHMS WITH PARAMETERS

In following table the comparison between the various cryptographic hash algorithms like MD5, SHA-1, SHA-2, SHA-3 and Whirlpool shows with various parameters.

TABLE-I Hash Algorithms Comparisons

Parameters	MD5	SHA-1	SHA-2	SHA-3	Whirlpool
Block size (bits)	512	512	512,1024	1600-2*bits	512
Digest size (bits)	128	160	160,224,256,384,512	160,224,256,384,512	512
Word size (bits)	32	32	32,64	64	8
Rounds	4	80	80	24	10
Collision found	Yes	Theoretical attack	None	None	Yes
Operations	and,or,xor,rot	and,or,xor,rot	and,or,xor,rot,shr	and,or,xor,rot,shr	and,or,xor,rot

CONCLUSIONS

Cryptographic integrity and security should be provided with the proper implementation of the proper hash algorithm. The detail design of each hash function is provided. We discuss the MD5, variants of SHA and Whirlpool algorithms by considering the various parameters. In current scenario in digitization it is very important that when the digital certificates are issued to the particular organization, people or device that certificate must use the proper has function to provide the integrity of the data as well as to authenticate the legitimate user.

REFERENCES

[1] R.C. Merkle, "One Way Hash Functions and DES", in CRYPTO, 1989, pp.428-446.
[2] Kleppmann, Martin (2 April 2017). Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems (1 ed.). O'Reilly Media. p. 203. ISBN 978-1449373320.
[3] Goodin, Dan (2016-05-04). "Microsoft to retire support for SHA1 certificates in the next 4 months". Ars Technica. Retrieved 2019-05-29.

[4] "Google will drop SHA-1 encryption from Chrome by January 1, 2017". VentureBeat. 2015-12-18. Retrieved 2019-05-29.
[5] "The end of SHA-1 on the Public Web". Mozilla Security Blog. Retrieved 2019-05-29.
[6] "Announcing Request for Candidate Algorithm Nominations for a New Cryptographic Hash Algorithm (SHA-3) Family [U.S. Federal Register Vol. 72 No. 212)]" (PDF). November 2, 2007. Archived (PDF) from the original on March 31, 2011. Retrieved July 18, 2017.
[7] https://en.wikipedia.org/wiki/File:SHA-1.svg
[8] https://en.wikipedia.org/wiki/File:SHA-2.svg
[9] https://en.bitcoinwiki.org/wiki/SHA-3

