Increasing Cryptographic Security Using Message Authentication Code (MAC)

Seyyed Mehdi Mousavi*[1], Dr.Mohammad Hossein Shakour [2]

[1] Department of Computer Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran.
Email : lord456@gmail.com

[2] Assistant Professor , Department of Computer Engineering , Shiraz Branch,
Islamic Azad University, Shiraz, Iran

Abstract - The study dealt with efficiency and extensive use of the Hash-based Message Authentication Code (HMAC) function in well-known protocols like IP-Sec and SSL and its use in identifying the authenticity of the message using the studies conducted. With the spread of using HMAC, the security of this algorithm becomes important and can greatly help cryptographic community. In cryptography using HMAC, one can ensure the authenticity and validity of a message. HMAC creates the MAC of the message in question using a cryptographic key and a hash functions and adds it to the message intended to be sent to the receiver. The packet is verified if the message authenticity code is identical to the message on the receiver's side. Nowadays, HMAC is widely used in popular security protocols like Ip-Sec, Transport Layer Security (TLS), and Secure Sockets Layer (SSL).

Keywords: Cryptographic security, MAC, HMAC function

Introduction

The ever-increasing growth and expansion of computer networks, especially the Internet, has made drastic changes in the way of life and work of individuals, organizations and institutions. Thus, information security is one of the most important issues in this cycle. By connecting the organization's internal network to the global network, organization data are exposed to the access of individuals and external hosts. Ensuring the non-access of unauthorized people to sensitive information is of the critical security challenges regarding the distribution of information on the Internet. Many solutions have been proposed, such as limited Internet use, data cryptography, and the use of security tools for internal hosts and internal network security. One of the commonest ways to protect information is to encrypt it. Cryptography means converting information into an incomprehensible form and transporting it, and then returning the ciphered data to the original and readable mode. Access to encrypted information is not possible for unauthorized people, and only those who have a password key can open the password and use the information. Cryptography of computer information is based on the cryptographic science. The use of cryptography has a long historic past. Prior to the information age, most users of information cryptography were governments, and especially military users. The history of encrypting information dates back to the Roman Empire’s time. Nowadays, most computer cryptographic methods and models are used in conjunction with computers. The discovery and detection of information normally stored on a computer without any scientific method of cryptography would be easy needless of specialized expertise. Thus, data cryptography has evolved given the recent developments and new algorithms. The equivalent of cryptography in English is the word “Cryptography,” which is derived from the Greek words “kryptos” meaning “confidential” and “writing.” Cryptography is the science that explores the principles and methods of transportation or storing information securely (even if the data transportation path and communication channels or data storage area are insecure) [1].

Literature review

In examining the first users of cryptographic techniques, we see Caesar (the Roman Emperor) and Alkandi, a Muslim scholar, who invented and used very early cryptographic methods. For example, they encrypted the text to a certain degree by moving the alphabet in the full text, and only the one who knew the number of letters’ movement, could extract the original text.

Another primitive cryptography method is wrapping a paper tape on a cylinder with a specified diameter and then writing the message on the wrapped paper. Obviously, without knowing the diameter of the cylinder, reading the message would be very difficult, and only those who have the same copies of the cylinder can read the message.

Lorenz Cipher Machine was being used by Germany in World War II to encrypt military messages. In the twentieth century, this method was used in conjunction with electric motors for high-speed cryptography, whose samples are seen in Lorenz Cipher Machine and Enigma cipher machine.

With the advent of computers and the increase in their computational power, cryptographic knowledge entered the computer science area, which brought about three important changes in cryptographic issues. The high computational power made it possible for more sophisticated and more efficient methods to be encrypted. The cryptography methods that were being used to encrypt the message before, found many new applications. Prior to that, cryptography was mainly being done on text data using the alphabet, but computer introduction enabled cryptography on a variety of bit-based information [2].

Cryptography

Cryptography comes from two words “Crypt” and “Graphy”. The purpose of this science is to examine secret and hidden information. Cryptography is divided into two main categories: classical and modern. The definition of the modern cryptography is slightly different from the classical cryptography presented above. Nowadays, cryptography is one of the branches of mathematics and computer science. This science is also closely related to the information theory sciences, computer security, and engineering [3]. Encryption in digital space is an algorithm that translates your original and clear text into an unreadable text (by the message carrier).

Categorization of cryptography algorithms

Cryptographic algorithms can be categorized into three categories:

1. Classical cryptographic algorithm

In the past, given the lack of computational tools, it was natural for the cryptographic process to be done using classic methods including displacement and replacement. Although the operation was done manually, it was impossible to access the content of the message in that period with the tools of that time. Among the classic ciphers, the following cryptographic algorithms can be mentioned.

1. Caesar cipher
2. Wigner cipher
3. Skatel cipher
4. Hill Cipher
5. Playfair cipher

If it is detected that with which of the classic algorithms the text is encrypted, the initial text extraction will be very simple. Thus, in this class of algorithms, an attempt was made to hide the algorithm used [4].

2. Symmetric encryption algorithms

In symmetric encryption algorithms, unlike classical algorithms, no attempt is made to hide the algorithm used. In this class of algorithms, special numbers are used as keys for cryptography. In these algorithms, attempts are made to hide the cryptography keybetween origin and destination. Among the symmetric algorithms, one can cite the following cryptographic algorithms:

1 - DES
2 - 3DES
3 - AES
4- MD5

This class of algorithms is used for data cryptography because of displacement and replacement operations and / or data XOR. One of the problems with this class of algorithms is the distribution of the key between the source and the destination. These algorithms use a common key for encryption and decryption of data [4].

3. Asymmetric algorithms

In Rivest Shamir Adlemen (RSA) (asymmetric) algorithms, unlike symmetric algorithms, that use a key for cryptography and decryption, there are two public and dedicated keys to this process. The source code is encrypted with the public key and decrypted using the dedicated key in the destination. Asymmetric algorithms can be used for electronic signature, data cryptography, cryptography and key transport, smart cards, and so on. Among the asymmetric algorithms, the following cryptographic algorithms can be mentioned:

1 - RSA
2 - ECC
3 - DSA
4 - Diffie-Hellman
5 - ELGamal

Asymmetric algorithms are slower than symmetric methods given the application of numerical factorization or discrete logarithms to mathematical processes. These algorithms are less commonly used in data cryptography [4].

Classification of symmetric cryptography algorithms

1. One-way symmetric cryptography

This kind of cryptography is called one-way as the existing algorithms can only encrypt information and there is no way to decrypt information, such as MD5algorithm. This type of cryptography is usually used to encrypt passwords in databases and there is no way to access the password. For the correctness of a password, one should only encrypt the existing algorithm of the password entered by the user and compare it with the password stored in the database [4].

2. Two-way symmetric cryptography

In two-way symmetric cryptography, this is usually done with one or more public or dedicated keys that enable us to encrypt and decrypt using the cryptography keys, such as the Blowfish and Twofish algorithms [4].

Introducing some famous encryption protocols

1-Internet Protocol Security (IPsec)

IPsec is a collection of several protocols used to secure the Internet protocol in communications by authenticating and encrypting each packet in a single data stream. This is a joint product of Microsoft and Cisco Systems, which is so interesting [5].

IPsec, unlike other security protocols such as Secure Sockets Layer (SSL), Transport Layer Security (TLS), and Secure Shell (SSH) that are in the transport layer (layer 4),works on the network layer or the same layer of OSI reference model, the layer where the IP is located, which makes the protocol more flexible to protect Layer 4 protocols such as TCP and UDP. The next advantage of IPsec compared to other security protocols such as SSL is that the program does not have to be designed in accordance with the protocol. The IPSec protocol family includes two protocols: authentication header (AH) and ESP: both these protocols will be independent of IPSec.

2. AH protocol

In sum, AH protocol will actually provide the following security services:

1. Integrity of the sent data

2. Authentication of the origin of the sent data

3. Rejecting the re-sent packets

This protocol uses HMAC for the integrity of the data sent, and to do this, it will base its work on the secret key, which will payload the packet and the unchanging parts of the IP header will be similar to the IP address. After this, the protocol adds its header to it in the form of subfields and AH fields.

4. Security parameter index (SPI) field is composed of 32 bits. This field is composed of SA used to open encapsulated packets. Finally, 96 bits are used for HMAC.

5. HMAC protects the integrity of the data sent, as only peer-to-peer points know the secret key generated by the HMAC and checked by the same. As HA protocol protects the IP datagram including irreplaceable parts, such as IP addresses, AH protocol does not allow the translation of the network address. Network address translation (NAT) is placed in the IP field of another address (usually the later IP address), so the next HMAC change will not be valid.

3. Encapsulation Security Payload (ESP) protocol

ESP protocol provides the following security services:

1. Confidentiality

2. Authentication of the origin of sent data

3. Rejection of the re-sent packets

Indeed, ESP protocol both secures the integrity of the data (the health of the data transported), the packets that use the HMAC, and the confidentiality of the encryption principle. After packet cryptography and HMAC calculations, the ESP header is calculated and added to the packet.

4. Transport Layer Security

The Transport layer security protocol is based on a secure socket layer, which is one of the cryptographic protocols designed to secure Internet communications. It uses X.509 certificates and asymmetric cryptography to ensure the identity of the other party and symmetric exchange. This security protocol is used to transfer data to the Internet for purposes such as working with web sites, e-mail, internet faxes, and instant internet messaging. Although TLS and SSL are slightly different, the bulk of this protocol remains more or less the same. TLS and SSL in the TCP / IP model perform cryptography in the lower layers of the application layer, but are executed in OSI model in the session layer and work on the display layer: first, the session layer performs asymmetric cryptography settings necessary for cryptography and then display layer performs the cryptography of the connection. In both models, TLS and SSL work on the transportation layer.

Secure Sockets Layer (SSL) is a protocol developed by Netscape to reject private documents over the Internet. SSL uses a private key to encrypt information transported over an SSL connection. Both Netscape Navigator and Internet Explorer (and today's most modern browsers) support this protocol. Moreover, many websites use this protocol to provide an appropriate context for maintaining user confidential information (such as a credit card number) as stated in the standard. URLs that need an SSL-type connection start with https instead of http. SSL is an application-independent protocol layer. Thus, protocols such as FTP, HTTP and remote network can use it. However, SSL is optimized for FTP, HTTP and IP addresses. In SSL, public and private keys are used. Moreover, in SSL, symmetric and asymmetric states can be cited that can be like public and private key discussion, so that the symmetric representation of two key domains is used by the server and the client, in which case the encrypted contents will not be secured. This is because the shared key (server and the client) can be eavesdropped or hacked by a third party. Thus, asymmetric state is used. B and A keys are used in asymmetric encryption; i.e., if the contents are encrypted by key A, they will not be decrypted by the same key and can be decrypted only with key B, which corresponds to key A.

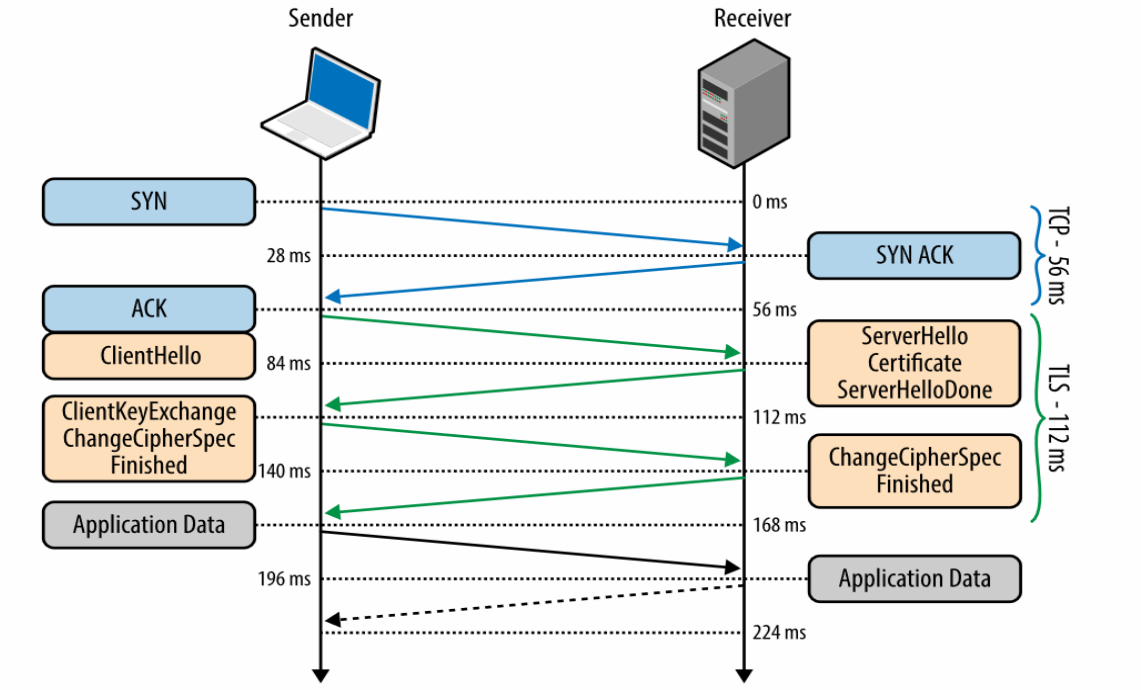

Figure 1: Schematic image of TLS protocol

HMAC

HMAC was originally developed by Bolary, Kantian and Crouches in 1996. In HMAC cryptography, there is a definite structure for calculating MAC, including a cryptography hash function in combination with a password key. Like every MAC, HMAC can examine the integrity of the message and validate a message simultaneously. Each cryptographic hash function, such as MD5 or SHA-5, can be used to compute HMAC. Thus, the resulting MAC algorithm is called HMAC-MD5 or HMAC-SHA1. The HMAC cryptography power depends on the ciphering power of the hash function used, the bit size of its intruder output length and the size and quality of the cryptography key.

A duplicate hash function divides the message into blocks of a given size and repeats the compression function on them. For instance, MD5 and SHA-1 work on 512-bit blocks. The size of HMAC output is the same as the size of the hash function used (in MD5 or SHA-1 mode, 128 or 160 bit). However, this size can be shortened if necessary.

The definition and analysis of an HMAC structure were first published in 1996 by Mayer Blair, Ron Centi and Hugo Crozic, who wrote RFC 2104. The paper also defined a type called NMAC rarely used so far. The Federal Information Processing Standard has promoted the use of HMACs and standardized them. HMAC-SHA-1 and HMAC-MD5 are used in IPsec and TLS protocols.

Explaining MAC functioning based on hashing

While sending the packet, the attacker can create the desired changes on the packet and then attempt to generate a hash for this data, and then replace the hash with the original hash; thus, in the hash destination generated by the receiver for the data will be the same with the sent hash and will not notice the change. HMACmechanism is used to avoid this problem, where besides the closed data; a hash is used to create a secret key. This secret key is only possessed by the two parties, so the attacker does not have this secret key, and if the attacker makes a change to the packet and generates a hash for it, the hashes will certainly not be the same in the destination as the attacker does not possess the secret key to produce the hash.

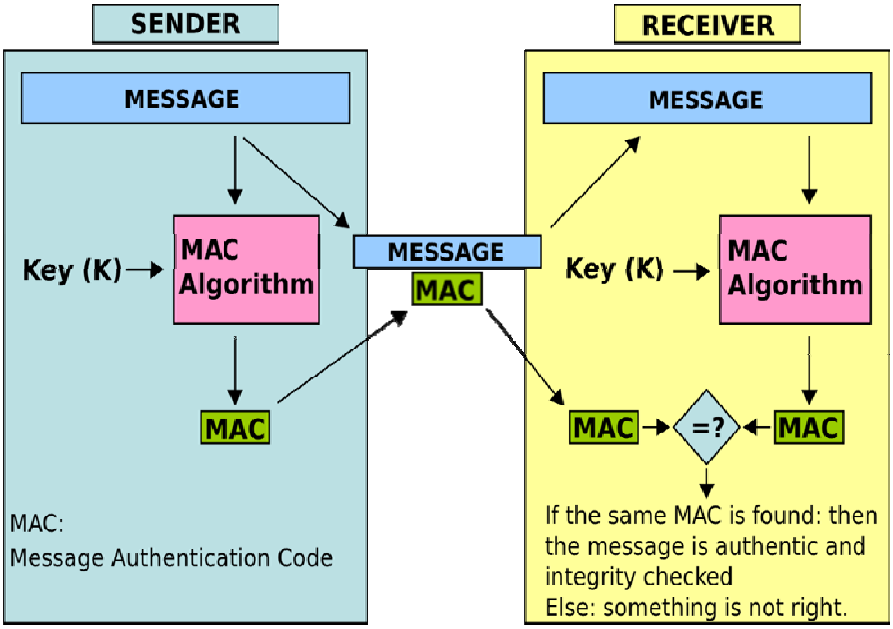

Figure 2: HMAC structure

Sending packet after implementation of HMAC algorithm on the message (the message can be encrypted or not) is as follows:

1 HMAC of the message is calculated using the private key and the hash function using the following formula:

$$HMAC(K,m) = H((K \oplus opad) \parallel H((K \oplus ipad) \parallel m)) \quad (1)$$

2. The calculated HMAC is sent along with the message packet to the receiver (HMAC is added to the beginning or the end of the message and is sent along with the package or in another packet)

3. The sent packet reaches the receiver and the authenticity of the message is checked through the synchronization of the message with HMAC using the private key HMAC and the desired hash function.

4. If the authenticity is confirmed, the receiver of the message will check the message.

5. If the message is encrypted, the message decryption stage is completed after the authentication is confirmed.

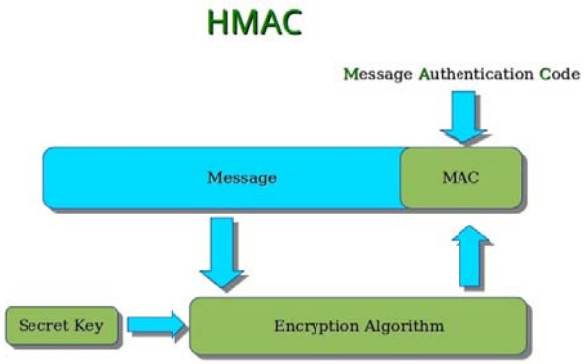

Figure 3: the overall scheme o HMAC

The length of HMAC input is smaller than 2^{128}.

The length of HMAC output is 512 bit.

Given the significance of HMAC in identifying the authenticity of the message and the current use of this algorithm in popular protocols such as IPC and TSL, increasing the security of this algorithm is an effective step in increasing the security of today's protocols. However, on the other hand, the security attacks on the HMAC algorithm in the recent years may affect the security of the protocols that use this algorithm in the future. Nonetheless, this does not mean that HMAC is obsolete and its use does not have any effect on increasing security, but its algorithm trend has to be modified or improved.

Considering the examination of the attacks on HMAC, generally, due to the weakness of the private HMAC key or the hash function used in this algorithm [6, 7], some papers have been written to increase the security of HMAC key or using valid hash algorithms to increase HMAC security [8 and 9].

Conclusion

Nowadays, with the development of computers and the speed of sending data on the network, a lot of information is received and sent in a unit of time. Thus, the need for information security in the network and speeding the exchange of information are of the needs of a two-way parallel cryptographic algorithm to encrypt and decrypt bulk information.

In cryptography using HMAC, authenticity and authenticity of a message can be verified. HMAC creates the message authentication code and attaches it to the end of the message to be sent to the receiver using a password key and a hash function. The packet is confirmed if the recipient message code is identical to the authenticity of the message. Nowadays, HMAC is widely used in popular security protocols such as IPsec, TLS, and SSL.

According to the studies conducted in recent years, there have been some attacks on HMAC, which could jeopardize the security of this widely used function. However, most of the attacks have been due to the weakness of the key used in this function or the weakness of the hash function used in HMAC. In recent years, some studies have been conducted using securing the key or more secure hash functions, but unfortunately, they have not been able to function properly and provide security.

As stated, in the sent packet, message and HMAC are identifiable and using this feature, the attacker may be able to start attacks as a birthday attack, fast attack, or a collision attack, reach the key, falsify the message, make HMAC using the encrypted key, and send the message to the recipient, and the recipient does not notice the attack and authenticates the message.

References

[1] Bencsáth, B., Pék, G., Buttyán, L., & Felegyhazi, M. (2012). The cousins of stuxnet: Duqu, flame, and gauss. Future Internet, 4(4), 971-1003.
[2] Stallings, W. (2003). Cryptography and network security: principles and practice. Pearson Education India.
[3] Buchmann, J. (2013). Introduction to cryptography. Springer Science & Business Media.
[4] Robling Denning, D. E. (1982). Cryptography and data security. Addison-Wesley Longman Publishing Co., Inc.
[5] Ansari, S., Rajeev, S. G., & Chandrashekar, H. S. (2002). Packet sniffing: a brief introduction. IEEE potentials, 21(5), 17-19.
[6] Contini, S., & Yin, Y. L. (2006, December). Forgery and partial key-recovery attacks on HMAC and NMAC using hash collisions. In International Conference on the Theory and Application of Cryptology and Information Security (pp. 37-53). Springer, Berlin, Heidelberg.
[7] Guo, L., Wang, L., Li, Q., Zhang, Z., Liu, D., & Shan, W. (2015, March). A first-order differential power analysis attack on HMAC-SM3. In First International Conference on Information Science and Electronic Technology (ISET 2015). Atlantis Press.
[8] Jeong, K., Lee, Y., Sung, J., & Hong, S. (2013). Security analysis of HMAC/NMAC by using fault injection. Journal of Applied Mathematics, 2013.
[9] Naqvi, S. I., & Akram, A. (2011, May). Pseudo-random key generation for secure HMAC-MD5. In Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on (pp. 573-577). IEEE.