# A Review Paper In Steganography Method that Using Message Integrity Techniques

**Article** *in* International Journal of Applied Engineering Research · October 2015

CITATIONS
0

READS
219

**1 author:**

Mohammed Salem Atoum
Al al-Bayt University
**25** PUBLICATIONS **141** CITATIONS

SEE PROFILE

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 20 (2015) pp 41267-41270


# A Review Paper In Steganography Method that Using Message Integrity Techniques

**Mohammed Salem Atoum**

*Assistant Professor, Computer Science Department Irbid National University Irbid, Jordan Moh_atoom1979@yahoo.com*

**Ahmad Khader Habboush**

*Associate Professor, Computer Science Department Jerash University Jerash, Jordan Ahmad_ram2001@yahoo.com*

**Abstract**
Steganography is the techniques can be concealed secret messages inside the cover. The aim of steganography is to keep a secret message within no modified. Message integrity is the scheme that can validate the secret message when it arrived in order to ensure if it is correctly delivered or has been altered. This paper contributes to the information security community by reviewed message integrity methods that using steganography technique**.**

**Keywords:** Steganography, Scrambling, Hash.

## Introduction

In this century, the Global Network Internet is a key technology in which the information is gathered, processed and distributed among people in different places in the world. Internet enables faster communication among people in offices in the world which make the source of information more valuable. One of the subjects that are of concern to many users of the internet is information security. The three security goals for information security are confidentiality, integrity and availability [1].

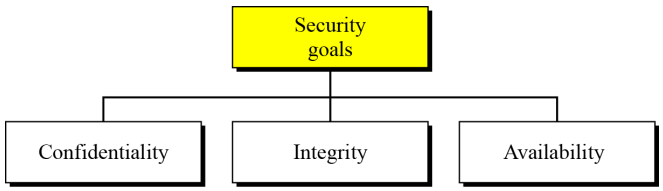


**Fig.1. Taxonomy of security goals**

Confidentiality refers to means of concealing information to from being access by authorized parties. This is regarded as the most important aspect of information security and provides privacy of information for those that own the information. Therefore, confidentiality is the responsibility of an organization to prevent access from any malicious event that may threaten the confidentiality of its information [1]. Integrity on the other hand dwells on the degree of trustworthiness of any information coming from a different party. Information could only be trusted if it is ensured that the content send has not been altered during transmission or in anywhere. This could be traced only if an organization provides a measure of validating any information received. Thus, only the authorize subject is responsible for modifying information received [1-2].

Availability is another information security element that ensures information that is generated by organization is available on demand at the right time in a right place. Only authorized access should have that information, if it is available for them, otherwise no any other party should have access to that information. Furthermore, the storage of information should be made available to only application that process it; otherwise, information might be vulnerable when unauthorized applications use it [1]. Information security in general requires some help from mathematical application, hence there are two basic ways in which information can be secured; these include steganography and cryptography [2].

This paper primarily focuses on steganography technology those using message integrity methods such as hashing and scrambling techniques

## Steganography and Watermarking

Much of the research that has been done on information hiding adopts steganography and watermarking methods. Steganography refers to methods that are used to transmit the embedded message without the observer being able to notice the embedded message in the cover message. In steganography method, transparency is a critical issue but not robustness. Watermarking ensures security of data by embedding a "watermark" for authentication purpose, which is an important step for copyright protection and tamper proofing. The embedded watermark is usually perceivable and cannot easily be removed from the stego message. Watermarks are extended information and are considered as attributes of the cover image but more so they are usually required to be semi-fragile or robust [3].

## Steganography and Cryptography

In the digital world, although steganography and cryptography operate in different ways they have the same goal of securing data from unwanted parts [4]. The two methods are often lumped together even though they are very dissimilar. Steganography is operated by hiding the message transfer process in such a manner that the observer is not able to see and observe the exchange of messages. This method secures the overall communication process because it obscures the message exchange process. However in some steganographic

technique, cryptographic layer is included which enables the encryption of the data before embedding. According to the combination of traditional cryptography and steganographic increases the strength of securing an embedded data while on transmission, because even if an attacker could be able to break a steganographic system, there is also an added cryptographic layer that will make it difficult to be broken [5].
Cryptography operates by modifying the contents of a file or message in a way that data can only be accessed by the right people. In this method, the intended recipient is given a decryption key that can enable him to access the data. It is difficult to figure out the means of encryption, and decryption keys used to secure the information. But, with the growing use of networks to send and receive data on the global information network, it has become very difficult to maintain this data. Because cryptography method is not being able to conceal the message transfer processes [5].
A few methods have been developed in image steganography that provide message integrity [6-10]. Generating codes from coefficient in DCT domain is the techniques can apply to achieve message integrity in an image cover [11]. These techniques not concern for the security, which ensure that since the content of the message, is neither altered nor modified or destroyed [12]. In image steganography, this technique is not suitable, because there is probability that an attacker can detect the present of message from stego object and will easily change the validation code and eventually create a new validation code matching with the new embedded message, which tries to embed after modification. However, this could drag the suspicion of the presence of message embedded. The tendency with which the message could be extracted by the random bits is minimal, thus this could be vulnerable to the steganographic technique [13].

**Message Integrity Methods**

Message integrity is an approach to check whether a secret message is altered or not. Three techniques: watermarking, cryptography and steganography used message integrity approach to achieve unity. In fragile watermarking, the technique does not provide firm embedding of watermark, it relies on the fact the any alteration with the watermark, and then the mark gets destroyed. This process ensures that if an adversary try to temper with watermarked file, then the file get destroyed. This technique only observed changes and response to the changes [14].
While though there are some certain allowable levels of modification in image yet, image authentication, to some degree might affect the visual appearance. However, to a lower degree of image modification will not alter those visual appearances. Thus a semi fragile watermarking can be more resistance and robust to any image modification and significantly better than the fragile watermark [15]. A closely related to semi-fragile watermark is seen in [16], where the technique relies on robustness as the major goal. This technique utilized a decomposed image size in medium size, blocks. In each block a watermark is embedded, thus it is made easier to compare the block, if some part of it is removing or altered with the original image. On the other hand if some part of the image is added it will be easier to detect as well [17].
Self-embedding is proposed by [17] in order to tract and corrects any possible changes that are made to the image. It could be that the changes comes as result of addition of bits, hence this will be tract and corrected. It may also be that the changes come from the fact that some part of the image is removing or modified. In this technique the entire embedding of the image is within itself. Unfortunately, due to the high embedded data, the ratio of the underlying quality metrics will be in the form that it will experience low visibility and the quality of the extracted image will be poor [17].
In steganography literature little research uses the integrity methods. The methods are developed in image steganography. In [12] proposed method to generate verification code by using the AC coefficients of the (DCT) domain to detect unintentional changes to embedded information during communication. The weakness for using same technique is the security of message integrity. However, in this research the security is improved by applying secure hash function. The advantage of using hash function is that the size of the output of the hash function is constant, no matter the size of the input.

**Hashing Algorithms**

Hash algorithms are also known as message digests or one-way transformations. A cryptographic hash function is a mathematical transformation that carries a message of arbitrary length and computes from it a fixed-length number. The hash of a message M is h (M). It has the following properties [18]:

i. For any message M, it is relatively easy to compute h(M). This just means that in order to be practical it cannot hold a lot of processing time to calculate the hash.
ii. Given h (M), there is no way to obtain an M that hashes to h (M) in a way that is substantially easier than going through all possible values of m and computing h (M) for each one.
iii. Even though it is obvious that many different values of M will be transformed to the same value h(M) because there are many more possible values of M, it is computationally infeasible to find two values that hash to the same thing.

There are two families of modification detection codes: message digits family (MD) and secure hash Algorithm family (SHA). A message digest MD is a cryptographic hash function containing a string of digits created by a one-way hashing formula. Message digests are designed to protect the integrity of a piece of data or media to observe changes and alterations to any part of a message. They are a type of cryptography utilizing hash values that can warn the copyright owner of any modifications applied to their work [19].
Message digest hash numbers represent specific files containing the protected works. One message digest is assigned to particular data content. It can reference a change made deliberately or accidentally, but it prompts the owner to identify the modification as well as the individual(s) making

the change. Message digests are algorithmic numbers. This condition is also known as a hash value and sometimes as a checksum [19].

There are four series for MD family will be explained below [19]:

**MD2**: is a cryptographic hash function developed by Ronald Rivest in 1989. The algorithm is optimized for 8-bit computers. MD2 is described in RFC 1319. Although MD2 is no longer considered secure, even as of 2010, it remains in use in public key infrastructures as part of certificates generated with MD2 and RSA.

**MD4** is a cryptographic hash function developed by Ronald Rivest in 1990. The digest length is 128 bits. The algorithm has influenced later designs, such as the MD5 and SHA-1 algorithms. The security of MD4 has been severely compromised. The first full collision attack against MD4 was published in 1995 and several newer attacks have been published since then. As of 2007, an attack can generate collisions in less than MD4 hash operations. A theoretical preimage attack also exists.

**MD5** is one in a series of message digest algorithms designed by Professor Ronald Rivest of Massachusetts Institute of Technology (MIT) in 1992. When analytic work indicated that MD5 predecessor MD4 was likely to be insecure, MD5 was designed in 1991 to be a secure replacement. Weaknesses were indeed later found in MD4 by Hans Dobbertin. In 1993, Den Boer and Bosselaers gave an early, although limited, the result of finding a "pseudo-collision" of the MD5 compression function; that is, two different initialization vectors which produce an identical digest.

**MD6** is the final series of MD family. It uses a Merkle treelike structure to allow for immense parallel computation of hashes for very long inputs. The authors claim a performance of 28 cycles per byte for MD6-256 on an Intel Core 2 Duo and provable resistance against differential cryptanalysis

The SHA is a family of cryptographic hash functions published by the National Institute of Standards and Technology (NIST) as a U.S. Federal Information Processing Standard (FIPS), and may refer to [20]:-

**SHA-1**: A 160-bit hash function which resembles the earlier MD5 algorithm. This was designed by the National Security Agency (NSA) to be part of the Digital Signature Algorithm. Cryptographic weaknesses were discovered in SHA-1, and the standard was no longer approved for most cryptographic uses after 2010.

**SHA-2**: A family of two similar hash functions, with different block sizes, known as SHA-256 and SHA-512. They differ in the word size; SHA-256 uses 32-bit words where SHA-512 uses 64-bit words. There are also truncated versions of each standardized, known as SHA-224 and SHA-384. These were also designed by the NSA.

**SHA-3**: A hash function formerly called Keccak, chosen in 2012 after a public competition among non-NSA designers. It suffers the same hash lengths as SHA-2, and its internal structure differs significantly from the remainder of the SHA family.

**Scrambling Methods**

Scrambling method is a procedure that changes the positions of symbols in a message, thus it is difficult for attackers to find out the true of a message. Cryptography, encoding and permutation are three methods can be implemented to generate scrambled message. In cryptography method that encryption is aimed to encrypt the secret message to be scrambled.

Encoding methods such as LZ-77 and Huffman encoding are aimed to represent the data and decrease the size of secret message. LZ-77 is a simpler technique, effective in text compression [21]. It employs the concept that words and phrases in a text stream can be repeated and encodes this as a point to earlier occurrence with the pointer having the number of characters to be matched. Pointers and decompressed characters are differentiated by leading flag bit with a 0 and 1 denoting a pointer and decompressed character respectively. The capability of this technique is to reduce the file size and increase compression ratio making it useful in text compression application. But it is not good for BMP, GIF, TIF image file format and requires longer compression data processing which can cause some loss of character

Huffman coding is a popular technique used by researchers of steganography. Characters in a data file are converted to a binary code where the most common characters in the file have the shortest binary codes, and the least common have the longest. The assignment of code words to source message is on probability bases which the source messages appear in the message ensembles. Short code words are used to represent more messages that appear more frequently while longer code words for messages with smaller probabilities map and probabilities are confirmed before transmission begins. A dynamic code is a mapping from a set of messages to set of code words which changes over time. Huffman coding is used to compress and decompress files of different format such as BMP, JPG, TIFF, TIF and GIF. The scheme is also used to compress binary file by representing every single binary digit with 2 bit sequence which leads to reduction of the compression ratio and file size after the compression process. Its limitation is a bigger sized image after compression when compared with the original image and cannot capture higher order relationship between words and phrases making it imperfect for text encoding [22]. The disadvantage of using cryptography and encoding methods is the computational complexity. Moreover, the encoding method maybe lost the data after encode and decode procedure.

Finally [23-24] proposed a Permutation method which is an ordered arrangement of numbers, terms, or string. The number of permutations on a set of S elements is given by S!. This factorial shows the possibilities that can be produced from this re-ordering.

**Conclusion**:

Several techniques have been discussed in this paper for methods that using message integrity within steganography techniques. The challenges faced for steganographic technique is how to scrambling secret message which embedded in cover media that achievement the attackers cannot read the content as well as ensure the secret message is arrived

correctly. Finally we expect the scrambling method is better technique to use it in message integrity.

**References**

[1] Schneier, B. (1963). Secrets and Lies-Digital Security in a Networked World, John Willey & Sons.

[2] Lenti J. (2000). Steganographic Methods, Department Of Control Engineering and Information Technology, Budapest University. Periodica Poltechnica Ser. El. Eng. Vol.44, No. 3-4, 249-258.

[3] Cvejic, N and Seppanen, T. (2002). Increasing the Capacity of LSB Based Audio Steganography. IEEE Transactions on Computers,.336-338

[4] Dunbar B. (2002). A Detailed Look at Steganographic Techniques and Their Use in an Open-Systems Environment, Sans Institute

[5] Katzenbeisser S., and Petitcolas F. (2000). Information Hiding Techniques for Steganography and Digital Watermarking, Artech House Inc

[6] Atoum, M. S. (2015). A Comparative Study of Combination with Different LSB Techniques in MP3 Steganography. In Information Science and Applications(pp. 551-560). Springer Berlin Heidelberg.

[7] Atoum, M. S., Ibrahim, S., Sulong,G. and Ahmed, A. (2012). MP3 Steganography: Review. Journal of Computer Science issues, 9(6).

[8] Atoum, M. S., Suleiman, M., Rababaa, A., Ibrahim, S., and Ahmed, A. (2011). A steganography Method Based on Hiding secretes data in MPEG / Audio Layer III. International Journal of Computer Science and Network Security, 11(5), 184-188.

[9] Atoum, M. S., Rababah, A. and Al-attili, A. I. (2011).New Technique for Hiding Data in Audio Files. International Journal of Computer Science and Network Security, 11(4), 173-177.

[10] Atoum, M. S., Ibrahim, S., Sulong, G., Zeki, A and Abubakar, A. (2013). Exploring the Challenges of MP3 Audio Steganography. Proceding IEEE from 2nd International Conference on Advanced Computer Science Applications and Technologies (ACSAT),Sarawak, Malaysia.

[11] Potdar, V.M., Han, S. and Chang, E. (2005). Fingerprinted Secret Sharing Steganography for Robustness Against Image Cropping Attacks. Proceedings of the IEEE International Conference on Industrial Informatics, 717-724.

[12] Park, Y., Kang, H., Yamaguchi, K., and Kobayashi, K. (2006). Integrity Verification of Secret Information in Image Steganography. In Symposium on Information Theory and its Applications, Hakodate, Hokkaido, Japan,1-4.

[13] Morkel, T. (2012). Image Steganography Applications for Secure Communication, Doctoral dissertation, University of Pretoria.

[14] Walton, S. (1995). Information Authentication for a Slippery New Age. Dr. Dobbs Journal, 20(4):18-26.

[15] Fei, C., Kundur, D. and Kwong, R.H. (2006). Analysis and Design of Secure Watermark-Based Authentication Systems. IEEE Transactions on Information Forensics and Security, 1(1):43-55.

[16] Fridrich, J. and Goljan, M. (1999). Protection of Digital Images Using Self-Embedding. Proceedings of the Symposium on Content Security and Data Hiding in Digital Media.

[17] Fridrich, J. (1998). Methods for detecting changes in digital images. Proceedings of the IEEE International Workshop on Intelligent Signal Processing and Communication Systems.

[18] Paar, C. (2000). Applied Cryptography and Data Security. Lecture Notes), Ruhr-Universität Bochum.

[19] Deng, K., Zhang, R., Tian, Y., Yu, X., Niu, X., and Yang, Y. (2010). Steganalysis of the MP3 Steganographic Algorithm Based On Huffman Coding. In Intelligent Computing and Integrated Systems (ICISS), International Conference on IEEE, 79-82.

[20] Rogaway, P., and Shrimpton, T. (2004). Cryptographic Hash-Function Basics: Definitions, Implications, and Separations for Pre-Image Resistance, Second-Pre-Image Resistance, and Collision Resistance. In Fast Software Encryption Springer Berlin Heidelberg. 371-388.

[21] Al-Laham, M., and El Emary, I. M. (2007). Comparative Study Between Various Algorithms of Data Compression Techniques. IJCSNS, 7(4), 281.

[22] Connell, J. B. (1973). A Huffman-Shannon-Fano Code. Proceedings of the IEEE, 61(7), 1046-1047.

[23] Atoum, M. S., Ibrahim, S., Sulong, G., and Zamani, M. (2013). A New Method for Audio Steganography Using Message Integrity, Journal of Convergence Information Technology,8(September), 35-44.

[24] Atoum, M. S., Ibrahim, S., Sulong, G., and Ahmed, A. (2013).New Secure Scheme in Audio Steganography (SSAS). Australian Journal of Basic and Applied Sciences, 7(6), 250-256.