The Importance of Hiding Text in Information Security

Saad Nasser Al-Azzam*, Fahad Ali Al-Garni

Faculty of Computing and Information Technology, University of Bisha, Saudi Arabia

Corresponding author E-mail: Snazzam.199@gmail.com

Doi: 10.29072/basjs.20220306

ARTICLE INFO	ABSTRACT

Keywords

Information Hiding, Watermarking, Digital Signature, Cryptography, Steganography

Data or information is a crucial resource for us. Thus, data security is even more crucial. There are alternate methods for securing data, as the communication mediums we use to convey data do not ensure security. Currently, information concealment is essential. It provided methods for encrypting data, rendering it unreadable to unauthorized parties. This article discusses the methods for data concealment and how they can be combined to give an additional layer of security.

Received 19 Aug 2022; Received in revised form 30 Aug 2022; Accepted 27 Oct 2022, Published 31 Dec 2022.

1. Introduction

Any organization or person requires data or information to function. Because the information may be misconstrued, we all detest having our conversations overheard. The same holds true for all organizations' and individuals' data. Two prospective parties must exchange information in a secure manner to prevent manipulation. All information exchange is susceptible to two distinct hazards. Unintentional users who overhear this communication have two options: they can either alter the information to imply something else, or they can listen to the message with the intention of deciphering it and using it to their advantage. Both of these attacks compromised the integrity and confidentiality of the message. It is essential to permit authorized access while restricting unauthorized access. Tough job Long ago, information was concealed. In the past, individuals concealed information by conveying hidden images or writing using invisible ink [1].

2.Importance

Because the means over which the information is sent cannot be trusted, or because the medium is not secure, data masking methods are important. Therefore, various techniques are required to make it impossible for unwanted users to extract the message's contents. Several causes for data concealment include

1. Personal Information

2. private data\

3. trade secrets and private information

4. To prevent data abuse 5. Inadvertent data loss, human error, and data tampering

6. Financial and extortion motives

7. to cover up evidence of crime;

8.As a pleasant bonus

3. Related work

Methods for concealing information have lately emerged as being very significant in a variety of application domains. Digital music, video, and photographs are increasingly being adorned with distinctive markings that are unnoticeable to the naked eye. These marks may conceal a copyright notice or serial number, or they may even actively assist in preventing unlawful reproduction. The employment of traffic security measures in military communications systems is becoming more common. These techniques, as opposed to just encrypting a message's content, strive to disguise the identity of the sender, the recipient, or even the very existence of the message itself. Techniques quite similar to those are employed in some of the systems for mobile phones and the planned plans for digital elections. Criminals try to take advantage of whatever traffic security qualities are supplied deliberately or unintentionally in the many communications systems that are accessible, while law enforcement agencies try to limit criminals' ability to do so. In spite of this, many of the strategies that have been suggested for use in this relatively new but fast developing subject can be traced all the way back to antiquity, and many of them are surprisingly simple to get around. In this post, we will attempt to provide an overview of the area, including what we know about it, what approaches are effective, what approaches are not effective, and what intriguing research issues there are [2]. The protection of sensitive data has emerged as a major concern for academics, as well as for members of the armed forces and officials in government agencies. It is vital to find innovative techniques to conceal information in order to ensure secure communication. Steganography is often employed for this purpose in order to convey secret information to its destination by using a variety of different approaches. The concealment of information via words is going to be the primary topic of this piece. When compared to other steganographic media, text data has very little redundancy, making it difficult to uncover information that has been steganographic ally hidden in text files. For this reason, the technique that we have presented makes use of the Arabic language to conceal sensitive information by using a combination of the zero-width-character and zero-width-joiner Unicode characters. As compared to the methods that were only recently presented, the hidden data capacity per word has been shown to be greatly boosted by the results of the experiments. The great visual resemblance between the cover and the stego-text that our suggested algorithm provides is the primary reason why it is superior to the research that came before it. This similarity may help to divert the attention of potential intruders [3]. Coverless information concealment is a relatively new approach to information concealment that has recently become an important topic of discussion in the world of information security. The present approach of covertly concealing sensitive information can only conceal a single Chinese character inside each piece of natural text. The approach, on the other hand, suffers from the drawback of having an inadequate capacity for concealment. The authors of this work suggest the use of a novel approach that they call the coverless multi-keyword information concealing method based on text. The primary purpose of the strategy is to conceal, within the texts, not only the keywords but also the total number of occurrences of certain terms. The results of the experiments indicate that the suggested approach has the potential to increase the capacity of the already widely used covert information concealing method that is based on text [4].

4. Features for Information Hiding Techniques

Any method of information concealment must have the following qualities:

1. Capacity: The quantity of data that may be concealed in a cover medium is referred to as capacity [1]. The quantity of information that may be concealed is limited by the need to not significantly change the original message in order to avoid drawing the attention of an unwanted user.
2. Security: Data should be protected using the information concealing approach so that only the intended user may access it. In other terms, it relates to a user's incapacity to recognize concealed information. The confidentiality and sensitivity of the information being conveyed must be protected in this way [1, 5].
3. Robustness is a measure of how much information can be hidden without causing problems or hurting the information that is hidden [1].
4. How much information can be hidden without making problems or hurting the information that is hidden is a measure of robustness [1].

5. Data Hiding Techniques

The three most often used methods of data concealment are steganography, cryptography, and watermarking.

5.1 Watermarking

A recognized picture or pattern that is imprinted on paper as a watermark serves as proof of authenticity [3]. When seen in transmitted light, a watermark appears as a variety of bright and dark tones. Banknotes, passports, postage stamps, and other security documents often have watermarks added as security elements. This idea is expanded upon in the digital sphere by digital watermarking. We now have to utilize procedures that can secure the ownership of digital material since there is so much data available online. Digital content, including photos, text, music, and video, is often pirated. These are fairly simple to create and distribute. Therefore, it becomes crucial to identify who is the document's owner. Digital watermarking offers a solution to the longstanding issues with digital data copyright [3]. A kind of marking subtly incorporated into any digital data, such as audio or picture data, is called a digital watermark. It may then be extracted or found to support a claim about the data. This data may include author, copyright, or image-specific information [3, 5]. Because the digital watermark is unaffected by transmission or change, we may safeguard our ownership rights in digital form. Digital watermarks are unnoticeable at all other times and only visible after using a certain algorithm. Digital watermarks are useless if they distort the carrier signal in a way that can be used to find them. The main objective of a watermarking system is to achieve resilience, which means that the watermark cannot be removed without altering the original data. Digital watermarking serves as a passive security measure. The data is only marked; it is not changed in any way, nor is access to it limited. Source tracking is one use for digital watermarking. At every stage of dissemination, a watermark is included in a digital signal. If a duplicate of the work is subsequently discovered, the watermark may be extracted from the copy, revealing where it was distributed. According to reports, this method has been used to identify the origin of illegally pirated movies. Another use is in broadcast monitoring, where watermarked footage from foreign agencies is often used in the news on television.

5.2 Cryptography

Both crypt and graphein refer to writing as well as the hidden or secret. The phrase is Greek in origin language. The technique of converting data into the unintelligible cypher text format is known as cryptography. The communication is deciphered or converted into plain text by the recipient on the other end. Data secrecy, data integrity, authentication, and non-repudiation are all provided by cryptography. Access to some kinds of information is restricted or limited due to confidentiality. Integrity is the upkeep and assurance of the veracity of the data being transmitted, i.e., the absence of any alterations, deletions, etc. Authentication confirms the source and recipient of the information's identities. Non-repudiation is the ability to keep the sender or the recipient from denying that the information they sent was signed by them. Today's cryptography and encryption are interchangeable terms. In this case, the unencrypted information is referred to as "cypher text," whereas the original information is known as "plain text. Three stages make up cryptography:

1. Encryption: converting plain text into an unintelligible format. Cryphertext is the result of this process.

2. Message transfer entails delivering the recipient's encryption text.

3. Decryption: To get the original plain text, the receiver on the opposite end of the communication decrypts the encrypted text.

Symmetric key cryptography and asymmetric key cryptography are two main categories for cryptography.

1. Symmetric Key Cryptography is the name given to encryption techniques in which the sender and the recipient use the same key. This type of encryption is used by several encryption algorithms, including AES, DES, RC5, etc.

Synchronized key Plain text, an encryption method, a secret key, cypher text, and a decoding algorithm are the five elements of cryptography. Using a secret key, an encryption method manipulates plain text in numerous ways. The secret key, which is selected by one of the communication parties, is independent of plain text. This results in cypher text as the output. The result of a decryption algorithm is plaintext, which is produced from the inputs of cypher text and a secret key. A major disadvantage of the symmetric key cypher is that each pair of communicating parties must share the secret key, and the key itself must be communicated in a protected medium. The text can be ciphered by any unauthorized user who has access to the secret key.

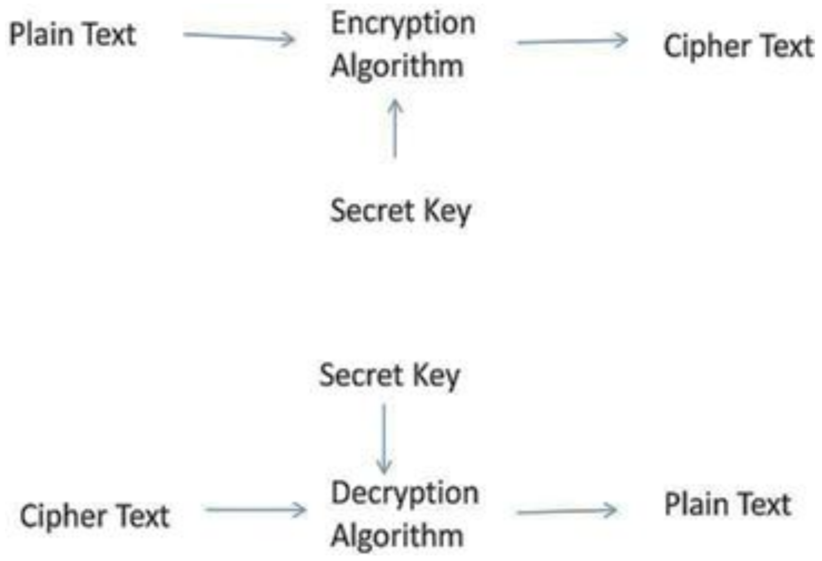

Figure 1: Symmetric Key Cryptography

Public key cryptography is another name for asymmetric key cryptography. Public key and private key are the two keys it employs. While the pairing private key must be kept a secret, the public key may be freely released. For encryption, the public key is used. The recipient is then provided the encryption text. To get the plain text at the receiver end, a secret key and decryption algorithm are used.

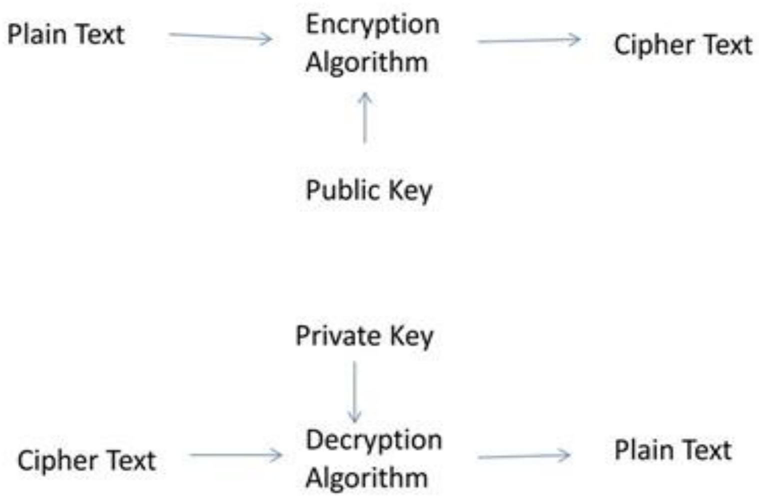

Figure 2: Asymmetric Key Cryptography

Digital signatures may also employ public key cryptography. Digital signatures allow the content of the communication being signed to be permanently connected. The contents are signed with a secret key, and the validity of the signature is checked with the public key that goes with it.

5.3 Steganography

Steganography is the process of obscuring or hiding a message, file, or picture within another message, file, or image. Greek in origin, the term steganography means "covered writing" or "concealed writing". Or to put it another way, it is the art and science of communicating in a manner that conceals the communication itself. The objective is to conceal messages within other benign communications in a manner that prevents the opponent from even realizing that a second message is there [6].

Steganography is primarily concerned with capacity and good security. The meaning of a stego medium may be altered even slightly. In any cover material, including photos, music, and video on the Internet, steganography hides the sensitive data. Four stages are involved in steganography:

1. Choosing the cover medium on which to conceal the data

2. The covert message or data that will be concealed in the cover material.

3. A function that hides data in the cover medium and its opposite, which shows the hidden data.

4. A password or key that can be used to verify information or hide and show it [5].

Carefully consider the cover you want. Given that steganography replaces redundant data with the secret message to be sent, the cover selected should have enough redundant information to be utilized to conceal the data. The three most common steganographic protocols are as follows: 1. Pure steganography: This kind of steganography doesn't involve the sending and receiving of cyphers like stego-keys, but both parties must have access to embedding and extraction algorithms. This method's cover is designed to minimize modifications brought on by the embedding procedure. Since the security of these systems relies on the assumption that no other party is aware of this secret message, they are not particularly safe 2. Secret key steganography: This technique embeds the hidden message into the cover using a key. The sender and the recipient are the only parties who are aware of the key before contact. Additionally, a secure means should be used to exchange the key. This strategy's drawback is that it is vulnerable to interception. 3. Public key steganography: it employs two keys, the secret key being known only to the people involved in communication and being used to reconstruct the original message [7]. The public key is maintained in a public database and is used for the embedding process.

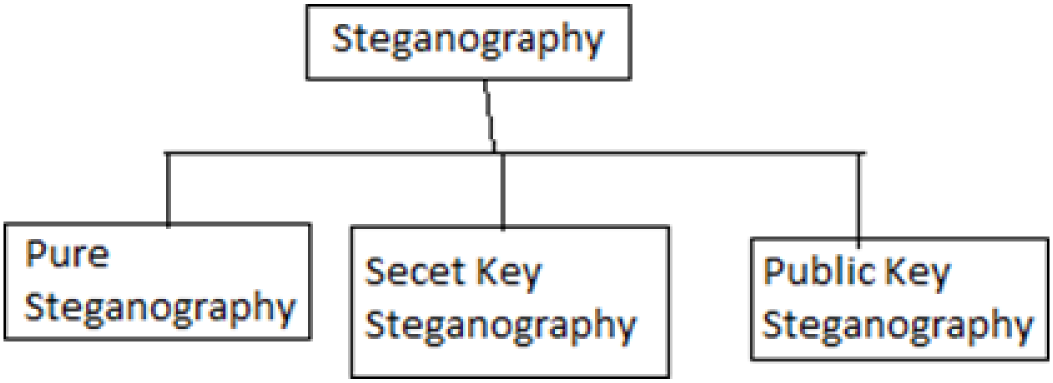

Figure 3: Steganographic Protocols

Different steganographic methods include:

1. Steganography for text
2. Steganography in audio and video
3. Steganography in images
4. IP steganography

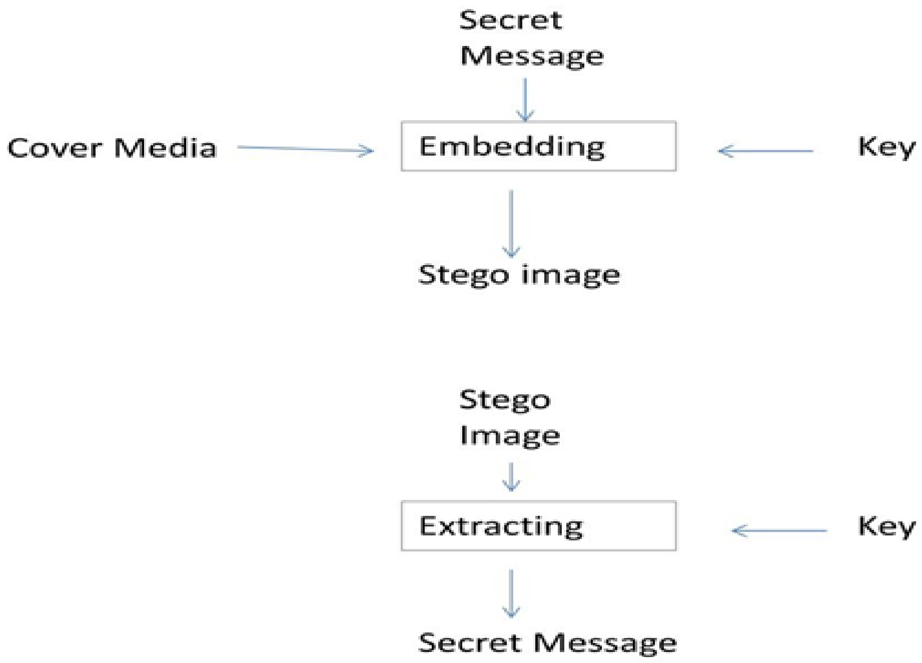

Figure 4: Steganography

6. Steganography Vs. Cryptography

Cryptography refers to "secret writing," while steganography refers to "cover writing" [6]. Steganography and cryptography are often mixed together; however, there are important differences between the two. The former sends the information to the network while concealing it behind a cover. Any accidental user will have a tough time figuring out if any secret information is contained or not. Steganography's key feature is that the cover should be selected with enough redundant information that, even after the message has been embedded, it is difficult to identify the message after seeing the message.

Using cryptography, on the other hand, the message is encrypted in a way that either makes it impossible to read or completely changes its original meaning. While cryptography changes the structure of the secret message, steganography leaves it unchanged. While the latter prevents an unauthorized user from learning the contents of the communication, the former prevents the existence of the communication from being detected.

7. Combined Cryptography and Steganography

Combining the two methods will provide an additional layer of security. The message may first be converted to a cypher text using cryptography. Steganography may then be used to insert this encrypted text into a cover material. Security, capacity, and robustness are the three objectives of data concealing that will be met by this integrated strategy.

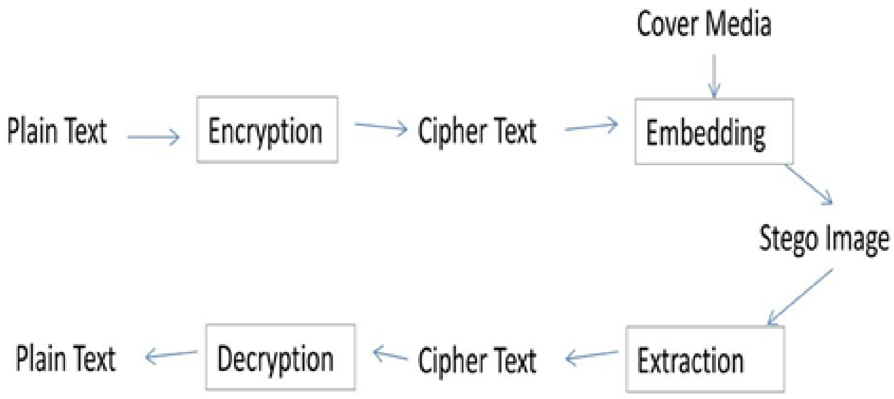

Figure 5: Cryptography with Steganography

8. Combined Watermarking and Steganography

Watermarking may be used on a document to ensure its legitimacy. Using a stego-key, this watermarked document may be conveyed through the communication channel by being included in the cover picture. The information may be decrypted via the reverse process at the receiving end, and then watermarked to verify its validity. This integrated strategy will meet the four goals of data hiding, which are security, capacity, robustness, and visibility.

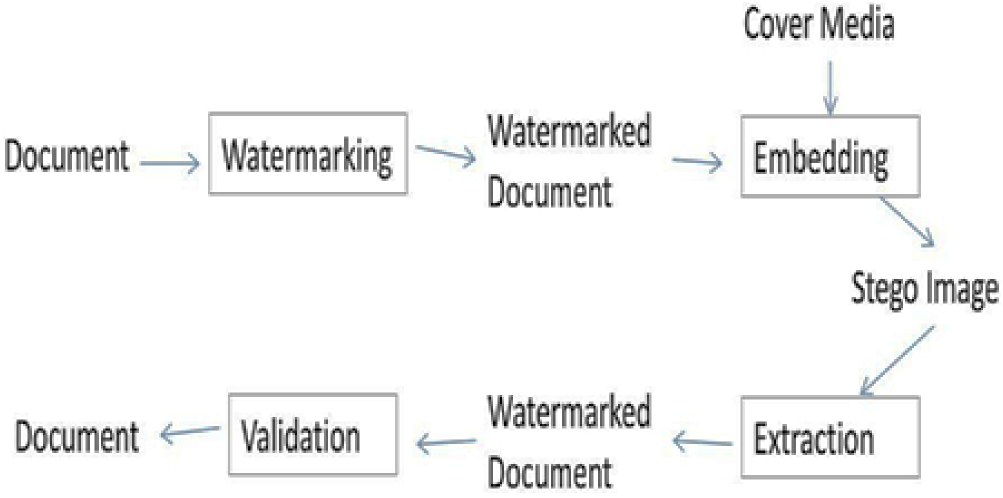

Figure 6: Watermarking with Steganography

9. Conclusions

In this study, we examined current data masking strategies, along with their advantages and disadvantages. This essay also examines the reasons why data concealment is becoming increasingly crucial and the objectives that any data concealment strategy must satisfy. In addition, we have attempted to demonstrate how combining one or more data-concealing strategies can help achieve the underlying data-concealing goals.

References

[1] S.N Wawale, A Dasgupta, Review of Data Hiding Techniques, Int. J Adv Res Eng. Technolo.2(2014)260-265

[2] F.A Petitcolas, R.J Anderson, M.G Kuhn, Information hiding-a survey, Proceedings of the IEEE, 87(2009)1062-1078, https://doi.org/10.1109/IEEC.2009.32

[3] A. Ditta, C. Yongquan, M. Azeem, K.G Rana, H. Yu, M.Q Memon, Information hiding: Arabic text steganography by using Unicode characters to hide secret data, Int. J Electronic Secur Digit Forensics, 10(2018) 61-78, https://doi.org/10.1504/IJESDF.2018.089214

[4] Z. Zhou, Y. Mu, C.N Yang, N. Zhao, Coverless multi-keywords information hiding method based on text. Int. J Secur Appl., 10(2016) 309-320.

http://dx.doi.org/10.14257/2016.10.9.30

[5] H. Kayarkar, S. Sanyal, A Survey of Data Hiding Techniques and their Comparative Analysis, ACTA Technica Corviniensis, 5 (2012)35-40, https://doi.org/10.48550/arXiv.1206.1957

[6] S.M Thampi, Information Hiding Techniques: A Tutorial Review, ISTE-STTP on Network Security & Cryptography, LBSCE (2004) https://doi.org/10.48550/arXiv.0802.3746

[7] M. Ghonge, A. Dhawale, A. Tonge, Review of Steganography Techniques, Int. J Adv Res Comp Electron, 1(2014) 260-265, doi 10.5120/ijca2016911203

أهمية إخفاء النص في أمن المعلومات

سعد بن ناصر آل عزام[1,2,3] فهد علي القرني[4]

1.كلية الأعمال - قسم إدارة الأعمال - جامعة الملك خالد- المملكة العربية السعودية

2.كلية الحاسبات ونظم المعلومات- قسم الأمن السيبراني- وجامعة بيشة- المملكة العربية السعودية

3. كلية القانون والدراسات القضائية – قسم القانون- وجامعه جدة- المملكة العربية السعودية

4.كلية الحاسبات وتقنية المعلومات- بجامعة بيشة- المملكة العربية السعودية

المستخلص

بالنسبة لنا ، فإن المعلومات أو البيانات مورد مهم للغاية. لذلك ، من الأهمية بمكان تأمين البيانات. هناك طرق بديلة لتأمين البيانات لأن وسائل الاتصال التي نستخدمها لنقل البيانات لا توفر أمان البيانات. اليوم ، إخفاء المعلومات ضروري. لقد عرضت طرقًا لتشفير البيانات ، مما يجعلها غير قابلة للقراءة لأي مستخدمين غير مصرح لهم. تتم مناقشة مناهج إخفاء البيانات في هذه المقالة ، إلى جانب كيفية اقترانها لتوفير طبقة إضافية من الحماية.