A Review on Data Leakage Prevention using Image Steganography

Mamta Jain

Department of Computer Science and Engineering, Mody University of Science and Technology
Lakshmangarh, Rajasthan, India
Email- mamta11.jain@gmail.com

Saroj Kumar Lenka

Department of Information Technology, Mody University of Science and Technology
Lakshmangarh, Rajasthan, India
Email- lenka.sarojkumar@gmail.com

Abstract-In today's increasingly distributed environment most of the work done by exchanging information from one system to another via network. The confidential data needs to be distributed only between the distributor and the trusted third parties. The data sent by the distributor must be secured, confidential and must not be reproduced as the data shared with the trusted third parties. Sometimes the data distributed by the distributor may be copied by different agents causing a huge damage to the institute and this process of losing the data is known as data leakage. In order to protect the data files being open source the data leakage must be detected in the early stage. This review paper deals with idea of preventing the data from being out sourcing by giving a special inscription to sensitive data from being reproduce using image steganography technique.

Keywords- cryptography; decryption; encryption; embedding; steganography

I. INTRODUCTION

In order to secure data from being out sourced every company focus on security issues with different strategy. The employees are also trained in order to maintain the secrecy of the data and maintain the basic structure of the company. Information security is frequently subjected to metaphors. There is no particular period of data leakage it may happen at any time. Data leakage only depends on the importance of the information distributed by the distributor.

Data Leakage Prevention (DLP) is an important concern for the business organization or IT Company. To prevent the serious consequences of data leakage risk including losing clients and stakeholder confidence, losing market share and goodwill of industry. Organization needs prevention against unwanted activities happening to have control over the flow of data within organization or outside it. DLP not only prevent from the unwanted accidental problem but from the malicious leakage of sensitive information to unauthorized users by providing solutions to organization in successful transmission of confidential information, security policies, blueprint data etc. This data leakage puts the company in ambiguity which results in the downgrade of the business and ultimately failure of the company. To overcome this problem we tried to add concept of image steganography to the distributaries data [1].This aims in providing a step towards the solution of the problem by:

• Presenting current practices and existing security measures that organizations use for preventing internal threats.

• Studying the current technologies that are used in detection of hidden information and methods of steganalysis

• Addressing a specific case for identifying and locating insider data leakage incidents through e-mail messages [2].Steganography is a practice for embedding a file, message and image within another file, message and image. The advantage of steganography over cryptography is that the unauthorized user not able to recognize the data as data is been embedded into image file, that give result as transmission of image file. In cryptography, the plain text data is being encrypted to other format.

For example, if at sender side, user starts with an innocuous image file and adjust the color of every 100th pixel to correspond to a letter in the alphabet, a change that someone not specifically looking for it remains unnoticed In section 2, the existing scenario has been illustrated. In section 3, defining data leakage prevention, information leakage classification and different modules has been illustrated. In section 4, literature survey of DLP using LSB image steganography techniques has been discussed. In section 5, conclusion of the paper is discussed.

II. EXISTING SYSTEM WORK

In the analysis, we are considering a system into which transmission of data in use is taking place at sender and receiver side. Data in use primarily refers to monitoring data movement stemming from actions taken by end users, whether that would detail copying data to a drive, or even cutting and pasting between applications.

Through steganography, we add data object with the image file being sent. This is done in order to improve his effectiveness in preventing the data from getting leaked by third party.

The distributor may be able to add fake objects to the distributed data. Fake objects may impact the correctness of what agents do, so they may not always be allowable. The idea of perturbing data to detect leakage is not new. However, in most cases, individual objects are perturbed, e.g., by adding random noise to sensitive salaries, or adding a watermark to an image. In this case, perturbing the set of distributor objects by embedding the original data with the image is done. Figure 1 shows the architecture of image steganography.

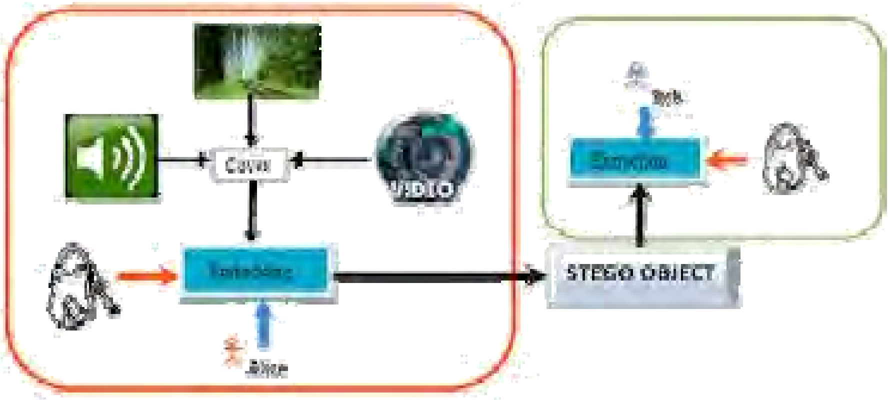

Figure 1. Architecture of Image Steganography

III. DEFINING DATA LEAKAGE PREVENTION

Data leakage prevention is a technique used to hide the confidentiality of data being accessed by unauthorized user

Most DLP solutions include a suite of technologies that facilitates three key objectives:

- Locate and catalog sensitive information stored throughout the enterprise
- Monitor and control the movement of sensitive information across enterprise networks
- Monitor and control the movement of sensitive information on end-user systems [3].

The functional specifications of our model will be:

Input: Corporate outgoing email messages containing attachments, and specifically images.

System’s response: Create an exact copy of the outgoing email.

Output: received the mail without leakage.

A. Classification of information leakage

The data leakage can be classified as:

1) Unintentional Leakage: a. Attach document b. Zip and send c. Copy & Paste.

This type of leakage occur when distributor unintentionally send the sensitive data to the third party. This is done mistakenly.

2) Intentional Leakage: This kind of leakage occurs when agent without knowing company policy sends the sensitive data anyhow. This is done when agent bypass the security rules or device without any personal profit. [4]

Example of intentional leakage is document renames, document type change, partial data copy, remove keyword.

3) Malicious Leakage: This leakage cause when a user intentionally deliberates to sneak the confidential data past the security rules. The problem of vulnerability causes when user tries to sneak confidential data from the company and send them through email. Example of Malicious Leak are character encoding, print screen, Password protected, hide data, policies or product.

B. *Module Information*

There are different kinds of modules:

1) Module1: In this module design a website and database for the application is done. The authorized person will logged in by respective account and new user have to log in by registering himself. The user will request for data in jpeg image format and data will be received by some addition of fake object.

2) Module2: If the particular data get leaked by the agent then the other channel will update those data and at server side the admin will check whether the data leaked is same or not.

3) Module3: The data is being requested by agent in mp3 file. The user will receive the data having fake object which is unique [5].

4) Module4: In this module the two different system get checked whether they both holds the same data or not .If data is same then the agent found as guilty.

IV. STEGANOGRAPHY

Steganography is defined as a technique to hide data into images in such a manner, which is unperceivable. Steganography and Cryptography, both are used for security purposes but with different implementation and approaches.

In cryptography, the text file get converted to other form which provide confidentiality to sensitive data but in steganography we hide the actual data file in image form so that if leakage get occurred the third party fails to recognize the actual data .This provide confidentiality as well as security to the sensitive data. The idea is to hide text in image with the conditions that the image quality is retained along with the size of the image instead we can encrypt the data. So the need is, in cryptography output of an unreadable data files are being send over an internet is easily detectable that some important information is being conveyed. While in steganography hiding message in an image, along with the conditions, it make seem of just an exchange of picture between two user ends.

The steps being followed in steganography are as under:-

1. Firstly the text message is being written, then encryption of the message is done.
2. Later, text is hidden in the selected media like image file and transmitted at the receiver side.
3. At receiver end, reverse method is done to implement and recover the original text message.

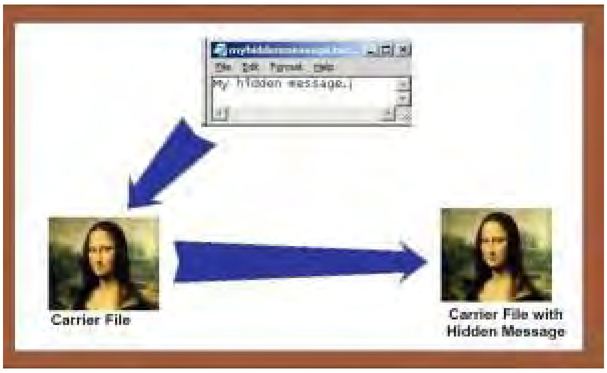

Figure 2. Process of steganography

Various techniques are used in the field of steganography by arranging the different bits of the character of the text message in the image file and other media. In order to encrypt the data two files are needed: (i) image file and (ii) the text file containing the data.

Our algorithm is simple and flexible using LSB (Least Significant Bit) technique. We have selected the formats that commonly use lossless compression that is BMP, PNG, TIFF and GIF. When data is streamed, it is captured after the header and chopped into 8 bits.

In 24-bit BMP, using RGB color model, with header size of 54 bytes, each pixel value contains the value of the color and is represented in bits (0 & 1).Similarly, text are also hidden in bit form. Therefore comparing bit values byte by byte both of text and image. The technique we are using is LSB i.e. storing in LSB of a byte (pixel). As the RGB model is used, we first stream an Image file and read the file in bits and then obtain the position ahead the header bits.

LSB affects the smallest changes of the 8 bits therefore it alters the image to minimum. The most common method used is called LSB mechanism that is hiding the data object in the LSB of the message. LSB is extremely vulnerable to attacks. LSB techniques implemented of 24 bit formats are difficult to detect as compare to 8 bit format. The other techniques include Filtering and Masking. This is normally associated with JPEG. In this, image data is extended by masking secret data over it therefore, experts do not include this as a form of Steganography.

V. CONCLUSION

In some cases, to improve the chances of detecting leakage and identifying the guilty party “realistic but fake” data records are injected. Using cover messages (container) to embed secret messages is the most popular use of steganography today. The method of steganography is very useful when a party need to send a secret, private or highly sensitive document over an open systems environment like the Internet. By embedding the hidden data into the cover message no one knows you have sent more than a harmless message other than the intended recipients. Steganography is in the nascent stage of development which allows users to hide files of larger sizes while also preserve the appearance of data by any cover image used. It is concluded that the original image and the final embedded image appear to be identical to the human eye. LSB makes use of BMP image, to be able to hide a secret message inside a BMP file; one would require a very large cover image. For this reason, LSB Steganography has also been developed for use with other image file formats. All the major image file formats have different methods of hiding messages, with different strong and weak points respectively. LSB in GIF images has the potential of hiding a large message, but only when the most suitable cover image has been selected

REFERENCES

[1] V. Shobana and M. Shanmugasundaram. "Data Leakage Detection Using Cloud Computing", International Conference on Information Systems and Computing (ICISC-2013), INDIA, 2013.

[2] V. Stamati-Koromina, et al. "Insider threats in corporate environments: a case study for data leakage prevention." Proceedings of the Fifth Balkan Conference in Informatics, ACM, 2012.

[3] Raman, Preeti, H. G. Kayacık and A. Somayaji. "Understanding Data Leak Prevention." 6th Annual Symposium on Information Assurance (ASIA’11). 2011.

[4] B. Purohit and P. P. Singh. "Data leakage analysis on cloud computing." International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622, 2013.

[5] Palimkar, Ashwini, and S. H. Patil. "Using Stegnography Technique for Data Leakage Problems Detect." International Journal of Engineering Research and Applications, vol 3, no.2, pp. 379-384, 2013.