REVIEW PAPER ON STEGANOGRAPHY USING KMEANS CLUSTERING ALGORITHM

[1]Ambe Nelsa Flora, MCA-ISMS Student,Jain University, Bangalore, India

[2]Umarani C, Assistant Professor,Jain University, Bangalore, India

Abstract - Steganography is a method of hiding away classified information inside any media. In this day and age, the correspondence is the essential need of each developing region. Everybody needs the mystery and security of their imparting information. In our day to day life, we utilize many secure pathways like web or phone for exchanging and sharing information; however it's not 100 percent protected. So to share the information in a hidden way, two procedures could be used. One of such procedures includes steganography. This method has the ability to conceal the information behind an image, video, audio or text file. As in the up and coming years the need of copyright security, information concealingand secrecy builds, Steganography comes up as one of the best methods in achieving this on account of its unique features. In this paper, we propose a method which clusters a selected dataset and groups them into various centroids and then hides that output behind an image with a process or method called image steganography. Various clustering algorithms can be used for this but in this project, we would be using the K-means clustering technique which is the most common clustering algorithm but also it is known to get accurate results.

***Keywords*:** Steganography, Kmeans, Clustering, Algorithm, Centroids, Secure.

1. INTRODUCTION

The Internet has emerged as the most convenient and efficient medium for communication. The recent growth in computational power and technology has propelled internet users to seek higher and better security techniques. Thus, throughout the ages,people have devised a means of hiding secrets from plain sight called steganography. Basically, the information hiding process of steganography starts by identifying a cover medium especially those that can be modified without destroying that medium's integrity. Steganography has emerged as a

glowing research area in which various methods have been proposed. The word steganography is built up of two words of ancient Greek origin “steganos” meaning “concealed” and “graphie” meaning “writing”. In this paper, we are going to discuss a brief overview of some already proposed methods with regards to steganography and kmeans clustering algorithm as separate entities.

2. BODY

2.1: A NEW METHOD OF IMAGE STEGANOGRAPHY USING 7TH BIT OF A PIXEL AS INDICATOR BY INTRODUCING THE SUCCESSIVE TEMPORARY PIXEL IN THE GRAY SCALE IMAGE

Steganography has emerged as a glowing research area in which various methods have been proposed in several carrier media. While the Internet has as the most convenient and efficient medium for communication with messages going from one place to another in a fast and cheap way in various fields like offices, medical areas, military, private sector, and military, thus, confidentiality of the transferred message needs to be maintained. To ensure that the message is transferred safely and securely over the network, a worthy and suitable method is needed. Steganography has proven over time to be a trusted method for achieving this goal. The proposed method on this paper is performed on gray images. A mathematical function is applied to the 7th bit of the pixels. This method provides a range of advantages. For example, two bits of message storage in each pixel and nondependency of the technique on the 8th bit. This particular method tackles the limitations of steganography to a higher extent.

2.2: AN OVERVIEW OF IMAGE STEGANOGRAPHY

Due to the rise of the Internet one of the most important concerns of information technology and communication has been on the security of information. Steganography is the origin and grounding of invisible communication. This is done by hiding information in other information, thus hiding the existence of the original communicated information. Almost all digital file formats can be used for steganography. Even though Image and audio files especially comply with this requirement, research and researchers have also uncovered other file formats that can be used for information hiding which includes protocol and text steganography. The term protocol steganography refers to the technique of inserting or embedding

information within messages and network control protocols used in network transmission. Images are the most popular cover objects used for steganography. An image a computer is a collection of numbers that constitute different light intensities. When working with larger images of greater bit depth, the images tend to become too large to transmit over a standard Internet connection. The process of compression is employed to analyse and condense image data, resulting in smaller file sizes. Least significant bit insertion is a common, simple approach to embedding information in a cover image. Least significant sit makes use of BMP images, since they use lossless compression.

There exists a large variety of stenographic techniques of which some are more complex than other, but yet, they all have their respective strong and weak points.

2.3: APPLICATION OF K-MEANS ALGORITHM IN IMAGE COMPRESSION

K-means algorithm is an iterative algorithm that divides the unlabelled dataset into k different clusters in such a way that each dataset belongs only to one group that has similar properties. Clustering is a type of unsupervised learning that can be used to probe data structures. Clustering is the process of dividing data into multiple clusters, each of which consists of one or more similar data. The clustering algorithm necessitates thegreatest similarity between the data of the same cluster. The clustering algorithm does not need to label the categories of the samples, but divides the data set into several clusters according to the similarity of the samples. The clusters of data are not predefined, but are defined according to the similarity of the characteristics of the samples. K-means clustering is closely related to a number of other clustering and location problems. The K-means algorithm solves the clustering problem of data in multidimensional space. The goal of the K-means algorithm is to find a group, assign each data to the cluster with the centroid closest to itself, and find minimum accumulated distance of each data point with the cluster center. K-means algorithm can be used to compress an image. Below are the principles of K-means clustering algorithm for compressing images:

- Preferred number of selected clusters K is very import, K must be less than the number of image pixels N.

- Using each pixel of the image as a data point, clustering it with the K-means algorithm to obtain the centroid.

- Storing the centroid and the index of the centroid of each pixel, so it not important to keep all the original data.

3. CONCLUSION

Although one can see that there exists a large selection of approaches to hiding information in images. All the major image file formats have different methods of hiding messages, with different strong and weak points respectively. Where one technique lacks in payload capacity, the other lacks in robustness.

ACKNOWLEDGEMENT

I would like to acknowledge God almighty for the strength and knowledge to write this paper. I would also like the acknowledge my above mentioned supervisor for her encouragement and guidance.

REFERENCES

1. Swati Gill,[1] and Rajkumar Yadav[1] "A New Method of Image Steganography Using 7th Bit of a Pixel as Indicator by Introducing the Successive Temporary Pixel in the Gray Scale Image". 2018

2. An efficient k-means clustering algorithm: analysis and implementation - Pattern Analysis and Machine Intelligence, IEEE Transactions on (umd.edu)

3. https://digifors.cs.up.ac.za/issa/2005/Proceedings/Full/098_Article.pdf

4. https://www.researchgate.net/publication/335080736_Application_of_K-means_Algorithm_in_Image_Compression