IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 4, Issue 6, June 2017

Data Leakage Detection of Video Files using LSB Steganography

Ms. Nilima V. Kayarkar [1], Prof. Ms. Gangotri Nathaney[2]
M.Tech Scholar, CSE Dept., WCEM, Nagpur, India[1]
Assistant Professor, CSE, Dept., WCEM, Nagpur, India[2]

Abstract: Information security is the technique of defending information from unauthorized access, use, disclosure, disruption, modification, inspection, recording or destruction. Data leakage happens whenever a system reveals or disclosed some information to unauthorized parties. We know that in business or in any distribution purpose it is necessary to transfer important data among many business partner and between the numbers of employees. But during this transfer of data information is reach to unauthorized place. So it is challenging and necessary to find leakage and guilty person responsible for information leakage. The goal of this system is to detect guilty person means agent when the distributor's sensitive data have been leaked and if possible to identify the leak data. In this project, we used the methodology for adding fake object into data. Fake object will be added using steganography concept. Steganography is the ancient technique of data hiding used for security. We can say that steganography is the art and science of hiding the existence of information. The ultimate aim of Steganography is to mask the fake data behind the cover file. In this system, we find the guilty person which is responsible for data leakage of video file. Video consist of number frame therefore we add fake data on every frame to increase the chance of detection. Here guilty person means agent is the insider. Insider means the person working in organization. Because in many times, data leakage is happen due to insider.

Keywords: Steganography, fake object, data leakage, distributor, agent.

I. INTRODUCTION

Sometimes data is leaked by someone and it found in unauthorized Places. Nowadays a large amount of data are sold and transmitted on the internet because most of the transactions are carried out on internet. The recently, large growth of the Internet results in wide range of web-based services such as database as a service, digital libraries, e-commerce, online decision support system etc. These applications make the digital assets such as digital images, video, audio, database content etc., easily accessible by ordinary people and easily available for people around the world for sharing, distributing, or many other purposes. In computer based steganography many forms of digital media may be used as cover for hidden information, photos, documents, web pages, images and even MP3 music files. The word Steganography was made from ancient Greek words steganos meaning "covered, concealed, or protected" and graphein meaning "writing". Means it is a technique of hiding of a message within another so that presence of the hidden message is indiscernible. Steganography can be used in a many data formats in the digital world of today. The most common data formats used are .txt, .doc, .bmp, .gif, .jpeg, .mp3, .avi and .wav. We know that now a day data leakage of audio and video file get increases. Recently the copies of many films get leak before the release of film. This is a serious issue. Hence it is necessary to detect the guilty person responsible for this leakage. For the implementation of this system we used two term

1. **Distributor:** It is owner of data who send file information to agent.
2. **Agent:** It is the members of organization, means it is insider and is semi-trusted. They may leak their own data to the outside word.

In this system we design website where distributor and authorized agent will log in. For new registration one has to fill the register form, and after successfully submitting the data he can logged into our system. When distributor wants to send video file they add fake data in original file and then send file to agent. Here the fake data which we added must be unique to detect the leakage. Here we used date and time as a fake data. We used time in millisecond. For adding fake data we used LSB steganography concept. A simple way of steganography is based on modifying the least significant bit layer of images is known as the LSB technique.

II. LITERATURE REVIEW

Peter Shaojui Wang, Feipei Lai, Hsu-Chun Hsiao, And Ja-Ling Wu proposes an insider collusion attack that carried out on data mining systems. In [1] explains how many insiders are sufficient to do this attack. In this system

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 4, Issue 6, June 2017

insiders means person within organizations collude with outsiders. This paper introduced many proposed privacy-preserving schemes to counter the attack.

X. Zhang and S. Wang in their paper [2] efficient steganography embedding by exploiting modification direction deals with the steganography embedding process especially on the Internet and given the large amount of redundant bits present in the digital representation of an image.

H. Zhang and H. Tang proposed in their paper [3] a novel image steganography algorithm against statistical analysis. This is an easiest method for embedding messages in an image with high capacity. It is not detectable by statistical analysis such as RS and Chi-square analysis.

J. Fridrich and P. Lison in their paper efficient steganography embedding by Grid colouring [4] in image steganography discuss a different technique unique to audio steganography is concealing, which exploits the properties of the human ear to hide information unnoticeably. A weak, but audible, sound becomes inaudible in the presence of another louder audible sound.

III. PROPOSED SYSTEM

We developed the system which finds the guilty person responsible for leakage of video. For this we implement two functions.

1. **Distribution of video file:** When distributer sends video to the agent, distributer adds fake data in the original video before send the video. Here we add current date and time as a fake data. The time is considered in millisecond to maintain uniqueness of fake data because we required unique fake data every time. One distributor can send same video to number of agent but for every agent fake data is different. We know that video is made of image and audio. Video also contain many number of frame. We add fake data on every frame to increase the chance of detection.

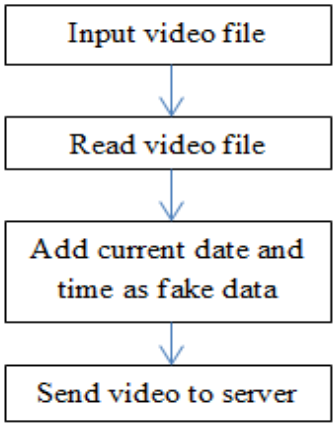

Fig.3.1: Flow diagram of distribution of file

2. **Detection of video file:** To detect guilty agent system compare leak video with the video which is send to the number of agent. If video is match with one of the agent that agent is the guilty one. As we add fake data in every frame it is possible to detect the guilty agent even if some part of video get crop.

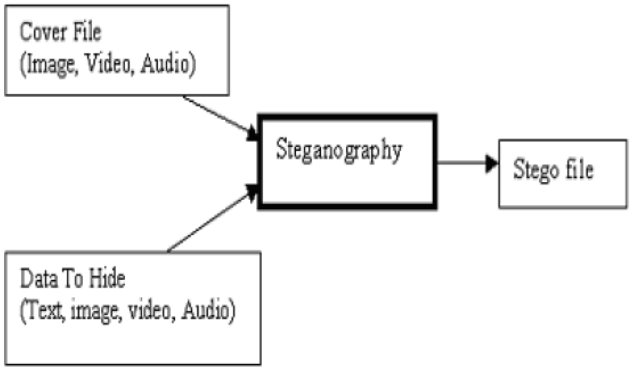

Fig 3.2: The Process of hiding data

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 4, Issue 6, June 2017

In this system we used steganography concept. Steganography system consists of three elements.
i) Cover file which hide secret message.
ii) Data to hide
iii) Stego-file which is cover file with message embedded inside it.

Following figure shows the data hiding in steganography.
In fig 3.2, cover file consist of data and video file which is hidden via. Fake data. We used current date and time as fake data.

IV. METHODOLOGY

To implement this system we used video steganography, we know that video is the combination of image and audio. Hence to understand video steganography first we have to understand image steganography and audio steganography.

1. Image steganography: A simple way of steganography is based on modifying the least significant bit layer of images known as the LSB technique. In the LSB technique, the least significant bits of the pixels is replaced by the message which bits are permuted before embedding. Image steganography is the widely used technique for hiding of secret data into digital image. This steganography technique exploits the faintness of human visual system (HVS). HVS can't notice the variation or changes in luminance of color vectors at high frequency side of the visual band. A picture is represented by a collection of number of pixels. The individual pixel is represented by their optical characteristics such as brightness, Chroma etc. The characteristics of images can be expressed in terms of digital bit likes 1s and 0s. This technique can be applied on both digital image formats i.e. in bitmap format as well as in compressed image format like JPEG. In compressed image format each pixel of the image is digitally coded using discrete cosine transformation (DCT).

2. Audio steganography: Audio steganography is potentially a very powerful. The audio provide users a large amount of choice and make the technology more reachable to everyone. A user can wish to communicate can rank the significance of factors such as data transmission rate, bandwidth and noise audibility and then select the method that best fits their conditions and criterion. For example two persons who want to send rare secret message back and forth might use LSB coding method. On the other hand a large concern wishing to protect its intellectual property from "digital pirates" may consider a more sophisticated technique such as phase coding or echo hiding. Means user used the technique according to his requirement.

Video file is a combination of both image and audio file. So, video Steganography is nothing but a combination of image and audio steganography. So, the combined evaluations i.e., the evaluations for image and audio steganography can be taken together for evaluation of video steganography. While doing video steganography, the effect on video has to be kept in mind to achieve a secure communicating media. Video files are generally consists of images and sounds so most of the applicable techniques for hiding data into images and audio are also usable for video steganography. In video steganography, sender sends the secret message to the recipient using a video sequence as cover media.
In this system we used two functions

a). Discrete cosine transform (DCT): DCT is the technique for converting a signal into frequency component. DCT work by separating images into parts of differing frequencies working from left to right, top to bottom, the DCT is applied to each block. The DCT is used in JPEG image compression, MPEG, DV, video compression. DCT is a lossy compression transform because the cosine values cannot be calculated exactly, and repeating introduce rounding errors into the final result. Variation between original data values and restored data values depend on the method used for calculate DCT.

b). Discrete wavelet transform (DWT): The discrete wavelet transform is used for transform image. In the numerical analysis and functional analysis a discrete wavelet transform (DWT) is the wavelet transform in which wavelets are discretely sampled. The discrete wavelet transform is an implementation of wavelet transform using discrete set of wavelet scale and translation. This technique is used to transform image pixels into wavelet, which are then used for compression and coding purpose. DWT has both frequency and location information. In this system we used this technique for compression.

V. EXPERIMENTAL RESULT

The following snapshot shows result of this system

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 4, Issue 6, June 2017

Fig.5.1: Home Page Window

Fig 5.1 consists of admin login, distributer login, agent sign in and agent registration. According to requirement we make login. First we make distributor and make login. Inside distributor we make number of agent.

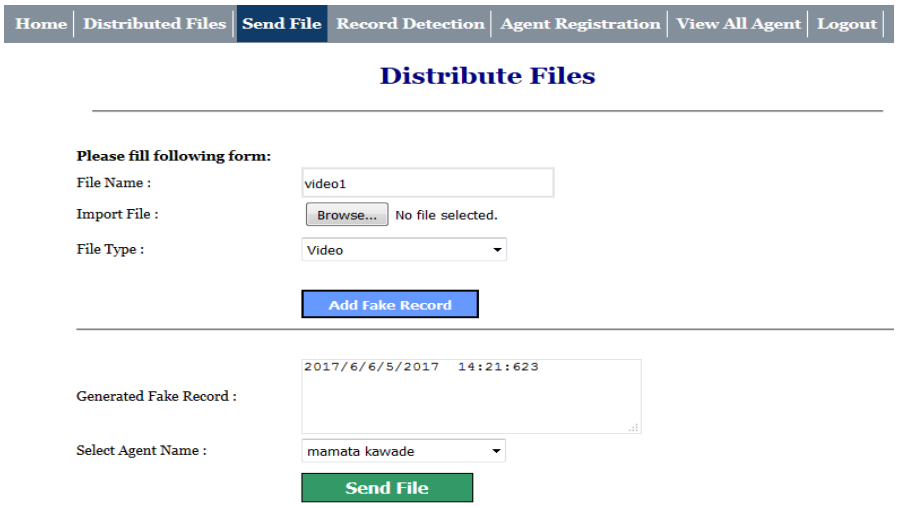

Fig.5.2: Send Window

In Fig 5.2 distributor send file to agent. Here distributor select file name and then add fake data. Finally distributor select agent name to which file is to send and send file. Here we used current date and time as a fake data.

Fig.5.3: Record Detection Window

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 4, Issue 6, June 2017

Import the leakage video file, shows in Fig.5.3 and it gives us the name of agent who is responsible for leakage of file.

Delete	DEMOVIDEOFILE	nilesh.joge	9/5/2017	VIDEO
Delete	DEMONOWVIDEO	nilesh.joge	9/5/2017	VIDEO
Delete	AAAAEE	nilesh.joge	9/5/2017	VIDEO
Delete	FFFFLL	nilesh.joge	10/5/2017	VIDEO
Delete	A1	nilesh.joge	23/5/2017	VIDEO
Delete	A2	mamata kawade	25/5/2017	EXCEL
Delete	IMAGE5	usha kayarkar	25/5/2017	IMAGE
Delete	IMAGE6	usha kayarkar	25/5/2017	IMAGE
Delete	F3	usha kayarkar	25/5/2017	EXCEL
Delete	A3	mamata kawade	26/5/2017	VIDEO
Delete	IMAGE7	usha kayarkar	26/5/2017	IMAGE
Delete	IMAGE8	mamata kawade	26/5/2017	IMAGE
Delete	FL1	mamata kawade	26/5/2017	IMAGE
Delete	VIDEO1	mamata kawade	5/6/2017	VIDEO

Fig.5.4: List of distributed file

Fig 5.4 shows the all file which is to be distributed between the agents.

List of Available Agents

LoginID	First Name	Last Name	City	DOB	Email Id
nilesh.joge	Nilesh	Joge	wardha	3/1/2013	nilesh.joge@gmail.com
mamata kawade	mamata	kawade	wardha	3/18/2000	mamata@gmail.com
usha kayarkar	usha	kayarkar	wardha	10/1/1998	usha@gmail.com

Fig.5.5. Available Agents

Fig 5.5 shows list of all available agents.

List of Available Files

Distributed By : megha shinde

File Type : VIDEO

Posted Date : 5/6/2017

File Name : VIDEO1

View this File

Fig.5.6.Available file window

Fig 5.6 shows list of available file in agent. From this window agent can view the video file.

VI. CONCLUSION

In this system,we find the guilty person responsible for the data leakage of video. This technique is used in many fields. For example in film industry, because now a days copies of many film get leak before the release of film. This causes financial losses. Therefore this system is very useful and effective in today era.In this paper, different techniques are discussed for embedding data in audio and video files as cover media. Here video before send and video after send are seem to be same. User cannot differentiate these two video by only watching them because fake data is hide from user.

Detectability, capacity, and robustness are three main factor of steganography. We can use any steganography methods but it is necessary to satisfy these factors. In this respect the research to device strong Steganography and Steganalysis technique is a continuous process. Our main aim is to increase the security level of the system and also maintain integrity and quality of the data.

ACKNOWLEDGMENT

Initially, we would like to thank our almighty in the success of completing this work. We would extend our gratitude to all the experts for their critical comments.

REFERENCES

[1] Peter Shaojui Wang, Feipei Lai, Hsu-Chun Hsiao, And Ja-Ling Wu, "Insider Collusion Attack on Privacy-Preserving Kernel-Based Data Mining Systems", IEEE Trans, Apr.2016.
[2] X. Zhang and S. Wang, "Efficient steganographic embedding by exploiting modification direction", IEEE Communication Lett., vol. 10, no. 11, (2006) November, pp. 781-783.
[3] H. Zhang and H. Tang, "A Novel image steganography algorithm against statistical analysis", Proceeding of the IEEE, vol. 19, no. 22, (2007), pp. 3884-3888.
[4] J. Fridrich and P. Lison, "Grid coloring in steganography", IEEE Transaction on Information theory, vol. 53, no. 4, (2007**)** April, pp. 1547-1549.
[5] P. Johri, and A. Kumar, "Review paper on text and audio steganography using GA", International Conference on Computing, Communication & Automation (ICCCA), Uttar Pradesh, India, (2015), May 15-16, pp. 190-192.
[6] P.P. Dandavate, S.S. Dhotre, "Data Leakage Detection using Image and Audio Files", International Journal of Computer Applications (0975 – 8887) Volume 115 – No. 8, April 2015
[7] Ratul Chowdhury , Debnath Bhattacharyya, Samir Kumar Bandyopadhyay and Tai-hoon Kim," A View on LSB Based Audio Steganography " International Journal of Security and Its Applications Vol. 10, No. 2 (2016
[8] Digital video steganalysis exploiting collusion sensitivity- Udit Budhiaa and Deepa Kundur Sensors, Command Control, Communications and intelligence (C3I) Technologies for Homeland Security and Homeland Defense, Edward M. Carapezza, ed., Proc. SPIE, Orlando, Florida, http://www.ece.tamu.edu/~deepa/pdf/BudKun04.pdf, vol. 5403, April (2004).
[9] Methods of Audio Steganography, Internet Publication on http://www.snotmonkey.com/work/ school/405/methods.html.
[10] Direct-sequence spread spectrum (DSSS), Frequency-hopping spread spectrum (FHSS) Wikipedia, the free encyclopedia, GNU Free Documentation license. http://en.wikipedia.org/wiki/Direct-sequence_spread_ spectrum, http://en.wikipedia.org/wiki/Frequency-hopping_spread_spectrum.
[11] M. Kharrazi Husrev, T. Sencar and N. Memon, "Performance study of common image steganography and steganalysis techniques", SPIE Proceedings, vol. 5681.
[12] W. Bender, D. Gruhl, N. Morimoto and A. Lu, "Techniques for data hiding", IBM Systems Journal, vol. 35, (1996).
[13] N. F. Johnson and S. Jajodia, "Steganalysis of Images Created Using Current Steganography Software", Proceedings of the 2nd Information Hiding Workshop, **(1998)** April.
[14] J. C. Joo, H. Y. Lee, C. N. Bui, W. Y. Yoo and H. K. Lee, "Steganalytic measures for the steganography using pixel-value differencing and modulus function", Proceedings of the 9th Pacific Rim Conference on Multimedia, Lecture Notes in Computer Science, vol. 5353, (2008), pp. 476-485.
[15] C. F. Lee and H. L. Chen, "A novel data hiding scheme based on modulus function", Journal of Systems and Software, vol. 83, no. 1, (2009) December, pp. 832-843.
[16] J. C. Joo, H. Y. Lee and H. Y. Lee, "Improved Steganographic Method Preserving Pixel-Value Differencing Histogram with Modulus Function", EURASIP Journal on Advances in Signal Processing, vol. (2010).
[17] C. M. Wang, N. I. Wu, C. S. Tsai and M. S. Hwang, "A high quality steganography method with pixel-value differencing and modulus function", Journal of System Software, vol. 81, no. 1, (2008), pp. 150-158.
[18] K. C. Chang, C. P Chang, P. S. Huang and T. M. Tu, "A Novel Image Steganographic Method Using Tri-way Pixel-Value Differencing", Journal of Multimedia, vol. 3, no. 2, (2008), pp. 37-44.
[19] J. K. Mandal and D. Das "Steganography Using Adaptive Pixel Value Differencing (APVD) for Gray Images through Exclusion of Underflow/Overflow", Computer Science & Information Series, ISBN: 978-1-921987-03-8, (2012), pp. 93-102.
[20] K. Gulve Avinash and M. S. Joshi, "A Image Steganography Method with Five Pixel Pair Differencing and Modulus Function", International Journal of Computer Applications (0975 – 8887) ,vol. 68, no. 1, (2013) April.
[21] V. Jithu, and A. Mary Alex. "Audio steganography using dual randomness LSB method." International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), IEEE, (2014), Jul 10-11, Tamilnadu, India, pp. 941-944.
[22] M. Zamani, A. Manaf, RB Ahmad, F. Jaryani , H. Taherdoost, AM Zeki, "A secure audio steganography approach". International Conference on Internet Technology and Secured Transactions, (ICITST) (2009) Nov 9, London, UK, pp. 1-6. IEEE.
[23] Banerjee, Sean, Sandip Roy, M. S. Chakraborty, and Simpita Das. "A variable higher bit approach to audio steganography." International Conference on In Recent Trends in Information Technology (ICRTIT), (2013) Jul 25-27, Chennai, India, pp. 46-49. IEEE.
[24] R. Din, H. Shaker Hussain, and S. Shuib, —"Hiding secret messages in images: suitability of different image file types", WSEAS TIONSRANSAC on COMPUTERS, vol. 6, no. 1, January 1 (2006), pp. 127 -132.
[25] K. Bhowal, D. Bhattacharyya, AJ Pal, TH Kim, "A GA based audio steganography with enhanced security." Telecommunication Systems. (2013) Apr 1, vol. 52, no. 4, pp. 2197-2204.
[26] LB Rahim, S. Bhattacharje and IB Aziz, "An Audio Steganography Technique to Maximize Data Hiding Capacity along with Least Modification of Host". In Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013) (2014) Jan 1, pp. 277-289, Springer Singapore.
[27] Balgurgi, P. Pooja, and S. K. Jagtap. "Audio steganography used for secure data transmission." In Proceedings of International Conference on Advances in Computing. Springer India, (2012), pp. 699-706
[28] R. J. Anderson (ed.), "Information hiding", 1st international workshop, volume 1174 of Lecture Notes in Computer Science, Isaac Newton Institute, Springer-Verlag, Berlin, Germany, (1996).

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 4, Issue 6, June 2017

[29] Nathan, Mark, N. Parab and K. T. Talele. "Audio Steganography Using Spectrum Manipulation." In Technology Systems and Management, Springer Berlin Heidelberg, (2011), pp. 152-159.
[30] S. Malviya, M. Saxena, A. Khare, "Audio Steganography by Different Methods", International Journal of Emerging Technology and Advanced Engineering [20] (ISSN 2250-2459, vol. 2, issue 7. (2012).

BIOGRAPHIES

Nilima Kayarkar has received her B.E. degree in Information Technology in 2012. She is pursuing Master in Technology in Computer Science and Engineering from Wainganga College of Engineering and Management, Nagpur. Her areas of interest include Security and Image processing.

Gangotri Nathaney has received her B.E. degree in Information technology in 2011. She has completed her Masters in Computer Science and Engineering from Shri Ramdeobaba College of Engineering and Management, Nagpur-440013. Her areas of interest include Image Processing, Pattern Recognition and Artificial Intelligence.