International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 3 Issue 3, March 2014

IDENTIFYING GUILTY AGENT USING MP3-STEGO-FILE

Ashwini Palimkar **Dr. S.H.Patil**
M.Tech (stud of CSE) **Professor & Head Dept of CSE**

Bharati Vidyapeeth Deemed University College of Engineering Pune, Maharashtra, India.

ABSTRACT

This paper contains concept of data leakage, in every organizations, data leakage is very serious problem faced by it. Information within an organizations increases significantly each year. This could be to a point where it becomes unmanageable, resulting in data leakages.
Information leakage (also known as data leakage) is when sensitive data are revealed intentionally or not intentionally to unauthorized parties. The information leaked out can either be private in nature and are deemed confidential, such as credit card numbers or information that could be used by attackers to further exploit the system.
In this paper, We are using Audio Steganography,A perfect audio Steganographic technique aim at embedding data in an imperceptible, robust and secure way and then extracting it by authorized people. Hence, up to date the main challenge in digital audio steganography is to obtain robust high capacity steganographic systems. Leaning towards designing a system that ensures high capacity or robustness and security of embedded data has led to great diversity in the existing steganographic techniques
We study the following problem:
Our goal is to detect when the distributor's sensitive data have been leaked by agents, and if possible to identify the agent that leaked the data. In this project, we are giving the methodology for adding fake object into data. Fake object will be added through steganography concept. Steganography is an ancient technique of data hiding or the steganography is the art and science of hiding the existence of information. Also, we devolved model for assessing the "guilt" of agents. Guilt model are used to improve the probability of identifying guilty third parties.

Keywords: Sensitive data, Fake records, MP3Stego, image, explicit request, sample request.

I. INTRODUCTION

Data leakage is defined as the unintentional distribution of sensitive data to an unauthorized entity. Sensitive data in organizations include intellectual property, financial information, patient information, personal data and other information depending on business. This is the current internet community; secure data transfer is limited due to its attack made on data communication. So more robust methods are chosen so that they ensure secured data transfer. Techniques such as encryption and watermarking are already used in this regard. However, the need for new techniques and new algorithms to counter constantly-changing malicious attempts to the integrity of digital data has become a necessity in today's digital era. One of the solutions which came to the rescue is the audio Steganography. Steganography, which literary means" covered writing" has drawn more attention in the last few years. Its primary goal is to hide the fact that a communication is taking place between two parties. . The sender embeds secret data of any type using a key in a digital cover file to produce a stego file, in such a way that an observer cannot detect the existence of the hidden message. At the other end, the receiver processes the received stego-file to extract the hidden message.

MP3, as a standard for transmission and storage of compressed audio, is a promising carrier format for steganography. First, MP3 is the most popular and widely used audio file format. When audios in MP3 format are taken as cover signals, the stego-audios

ISSN: 2278 – 1323

will be less likely to be noticed by steganalyzers than other audio formats
Basically the data to be hidden is stored as the MP3 file is created, that is during the compression stage. As the sound file is, being compressed during the layer 3 encoding process, data is selectively lost depending on the bit rate the user has specified. The hidden data is encoded in the parity y bi of the information. As
MP3 files are split up into a number of frames each with their own parity bit; reasonable amount of information can be stored. To retrieve the data all u need to do is uncompress the MP3 file and read the parity bits as this process is done.

II. SYSTEM REQUIREMENTS

Hardware Requirements

- Processor: Intel (R) Core(TM) i3 CPU
- Installed RAM: 2 GB
- System type: 32 bit operating systems

Software Requirements

- Java 1.6
- My SQL.
- Editor: Eclipse

III. PROBLEM DEFINITION

The distributor has to distribute data in an efficient manner such that leakage can be detected. There are many cases where alteration to the original data can't be done. In such cases, we can add some realistic records similar to the dataset which don't exist in reality. For example, In Organization the distributor can't alter the personal and contact details of the agent. In order to stegano concept such fake records are acceptable since no real agent matches with this record.

- Entities and Agents

Let the distributor database owns a set S={t1….tm} which consists of data objects. Let
the no of agents be A1, A2,…..An,. The distributor distributes a set of records S to any
agents based on their request such as sample or Explicit.

-Algorithms

A *Allocation for Explicit Data Requests.*--In this request the agent will send the request with appropriate condition. Agent gives the input as request with input as well as the condition for the request after processing the data after processing on the data the gives the data to agent by adding fake object with an encrypted format.
Explicit request Ri = EXPLICIT (S, Condi)

B. *Allocation for Sample Data Requests*--In this request agent request does not have condition. The agent sends the request without condition as per his query he will get the data. With sample data requests agents are not interested in particular objects. Hence, object sharing is not explicitly defined by their requests.
Sample request Ri = SAMPLE(S, ui)

IV. SYSTEM ANALYSIS & IMPLEMENTATION

In this paper, the main important point of this paper is to analyze the guilt of every agent which can be responsible the leakage. The main concept deals with agents' valuable information for an Organization. If data leakage happens, this uncontrolled data leakage puts Organization in a backward position.
For implementing this system, we used an Organization, pune.In this system, we consider data distributor is main channel and other employee is called agents.
The distributor maintains the entire database. The distributor registers the details of all agents**.** All Entities must select "New" when they enter and register for first time. The new register will enter details. The distributor validates the request and if he finds the agent is guilty, he adds fake objects. Choose "Add/Update" to make changes to an existing registration entered using the new Registration process or to add "tagging" information. Use "View" to query the contents of the registration database or download the database to local computer. It will be asked to login a logon username and password to validate login process. And it verifies the username and password with database. Once verified, it allows continuing the requesting process. The objects are serialized to prevent the data leakage. Only the valid user can unserialize the objects. In the implementation of the system we maintain,

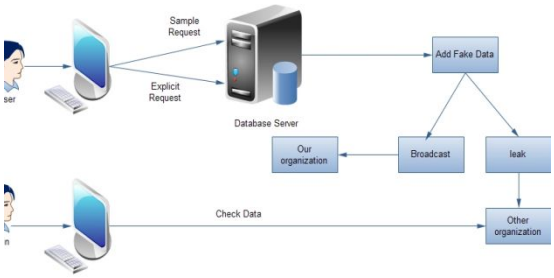

Figure1.System Architecture

ISSN: 2278 – 1323

A. Database Maintenance

Here the agent registration details are maintained and the sensitive data which are provided to agents are specified. The designing of the whole database is done.

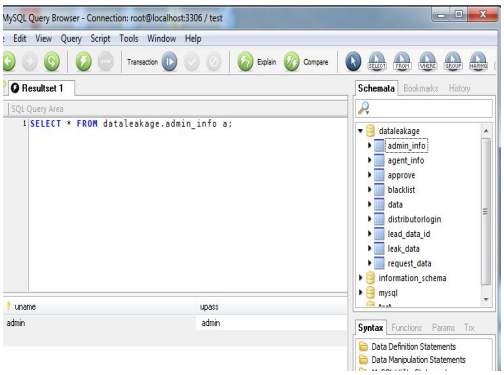

Figure 2: Tables used in this system

B. Guilty Agent Detection System

i. Distributor selects the agents to send the data according to agent request.
ii. Distributor creates fake data and allocates it to the agent. The distributor can create
iii. fake data and distribute with agent data In such cases, we can add some realistic
iv. records similar to the dataset which don't exist in reality.
v. For example In an Organization
   a. Distributor can't change the personal and contact details of the agent.
   b. To use stegano concept such data, some fake agent records are acceptable since no real agent matches with this record.
vi. Distributor checks the number of agents, who have already received data.
vii. Distributor chooses the remaining agents to send the data. Distributor can increase the number of possible allocations by adding fake record
viii. Each and every agent has its unique data.

C. Taking the action on Agent

i. Distributor selects the agents to send the data according to agent request.
ii. Distributor creates fake data and allocates it to the agent if the distributor is identify who leaked data using fake object
iii. Distributor should take the action on agent &put down this agent into Blacklist.
iv. Finding Probability of Leak data

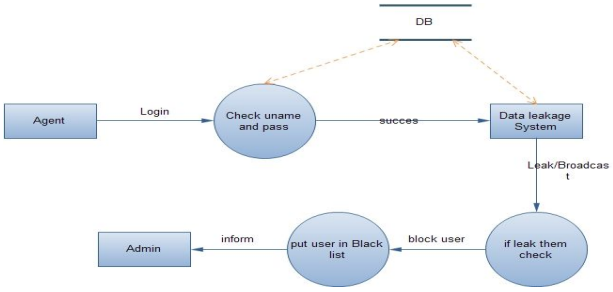

Figure 3: Action on agent

D. Data Allocation

This module checks whether this system data as well as another system data is same or not. If he found the same data then the agent will be called guilty agent.
we describe allocation strategies that categorize the normal and authorized agents based on their requests given to the distributor. We deal with explicit data requests of the agents.

A. Explicit Data Requests

The authorized agents send the request for available records which contain both sensitive and nonsensitive data in the distributor owed set i.e. Re=EXPLICIT ({t1,t2,…,tn}, cond1), then the request is said to be explicit data request to the distributor. The distributor cannot remove or alter Re data to decrease the overlap between requests from all other agents. So the distributor adds fake objects along with the requested data which do not come under the condition mentioned in agent's request ie. R= {t1, t2… tn, f}. If the distributor is able to create more fake objects, he could further improve the objective. We present the algorithms for explicit data requests allocation, agent selection for e-random and e-optimal as follows. In Algorithms 1 a strategy for randomly allocating fake objects.

Algorithm1. Allocation for Explicit Data Requests (EF)

Input: R1. Rn, cond1. . . condn, b1, . . . ,bn, B
Output: R1. Rn, F1. Fn
1: R←Ø //Agents that can receive fake objects
2: for i=1…….n do
3: if bi > 0 then
4: R ←R U {i}
5: Fi ←Ø
6: while B > 0 do

ISSN: 2278 – 1323

7: i ←SELECT AGENT(R; R1 . . . Rn)
8: f ←CREATE FAKE OBJECT (Ri; Fi; condi)
9: Ri ←Ri U {f}
10: Fi ←Fi U {f}
11: bi ←bi – 1
12: if bi = 0 then
13: R←R / {Ri}
14: B ←B – 1

Algorithm explanation

a. R set contains the resources
b. "cond " are used as the condition for steganography
c. B set contains the already created fake objects
d. SELECTAGENT function used to selecting the agent with their needed resources R1…

B. Sample data request:

An object allocation that satisfies requests and ignores the distributor's objective is to give each agent *Ui* a randomly selected subset of *T* of size *mi* [5].

Algorithm2 . Allocation for Sample Data Requests (SF)

Input: m1….., mn, |T| // Assuming mi _ jTj
Output: R1……., Rn

1: a ← 0 |T | //a[k]:number of agents who have received object tk
2 :R ← Ø……., Rn ← Ø
3: remaining ← S i=1 mi
4: while remaining › 0 do
5: for all i=1,…., n : | Ri | < mi do
6
: k ← SELECT OBJECT(i, Ri) // May also use additional parameters
7:Ri ← Ri U { tk }
8:a[k] ← a[k] + 1
9: remaining ← remaining – 1

V.EXPERIMENTAL RESULTS

In this, we have taken a set of 100 objects and requests from every agent are accepted. There is no limit on number of agents, as we are considering here their trust values.

The flow of our system is given as below:

- Agent's Request: Either Explicit or Implicit.
- Leaked dataset given as an input to the system.
- The list of all agents having common records as that of leaked records is found.
- It shows that as the overlap with the leaked dataset minimizes the chances of finding guilty agent increases.

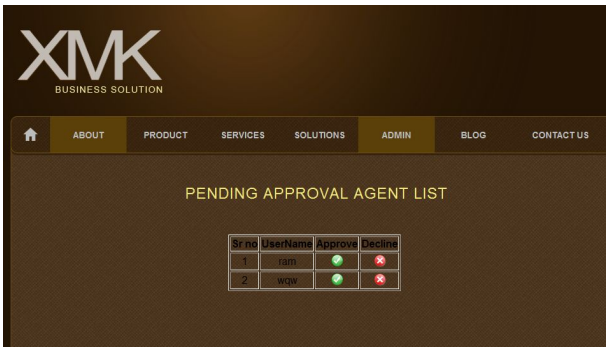

Figure 4: Approval permission By Admin

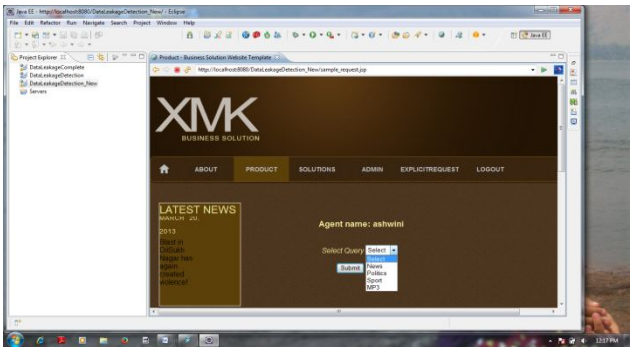

Figure 5: Agent requesting sample request

Here agent can request sample or explicit. Agents are selected as sample request as MP3 file.
Path as "E://data//song2.mp3"[before embedded data]
Path as "E://data//song2ashwini.mp3"[after embedded data]

ISSN: 2278 – 1323

The proposed technique is simulated in java with the data and audio files. The data is encrypted and is embedded into the audio file then that audio file is encrypted. By using this technique we find no difference in the size and the quality of the audio file before embedding data and after embedding data.

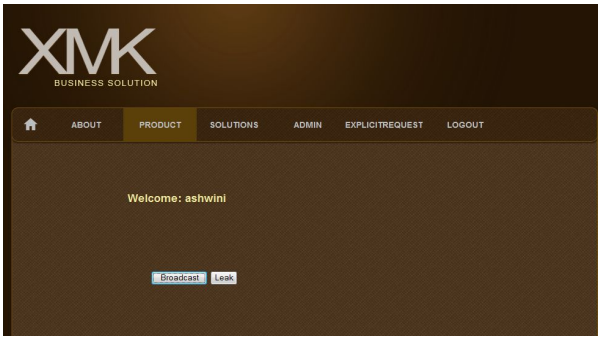

Figure 6: Broadcast /Leak MP3 files

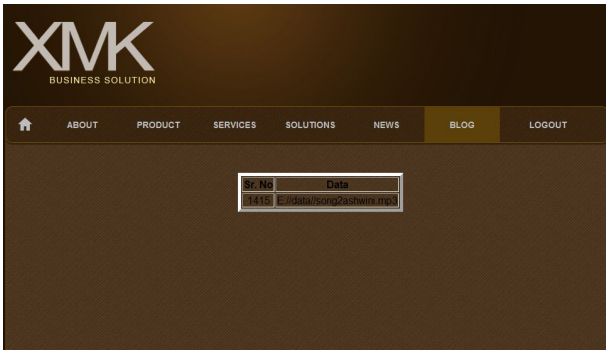

Figure 7: Leaked MP3 file by Agent

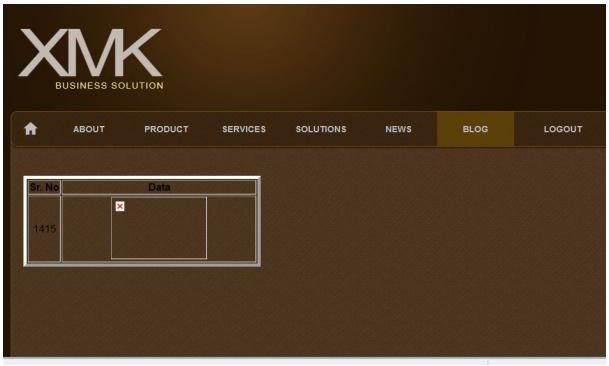

Figure 8: Broadcast MP3 files by Agent

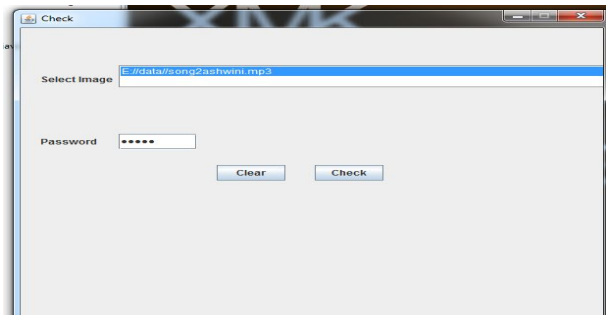

Figure 9: selection path of leakage MP3 files

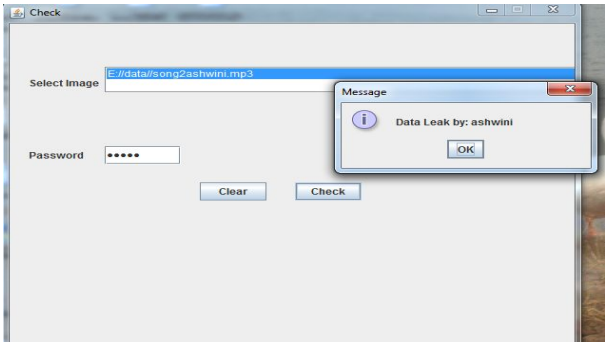

Figure 10: Data leakage can seen agent Guilt

VI. CONCLUSION

This system is to provide a good, efficient method for hiding the data in MP3 files .The algorithms we have presented implement a variety of data distribution that can improve the distributor's chances of identifying a leaker, for detecting guilty agent we have used proposed steganographic algorithm (SBR). The Encryption and Decryption techniques have been used to make the security system robust.

This proposed system will not change the size of the file even after embedding and also suitable

ISSN: 2278 – 1323

for any type of audio file format.

VII. REFERENCES……

[1] Panagiotis Papadimitriou, Hector Garcia- Molina (2010)'Data Leakage Detection', IEEE Transactions on knowledge and data engineering, Vol.22h, No.3.

[2] "A proposed algorithm for steganography in digital image based on LSB"by A.E. Mustata, AhmedBD, A.M.F.Elnamal &M.E.Elalm,Research journal specific education.

[3] "Stegnography Implementation", Beenish, Mehboob and Rashid Aziz Faruqui.

[4] Moreland, T "Steganography and Steganalysis", Leiden Institute of Advanced Computing Science.

[5] Mr. Rudragouda G Patil, "Development of Data leakage Detection Using Data Allocation Strategies," IJ of Computer Applications in Engineering Sciences **[**VOL I, ISSUE II, JUNE 2011**]**

[6] L. Sweeney, "Achieving K-Anonymity Privacy Protection Using Generalization and Suppression,"

[7] Archana Vaidya, Prakash Lahange, Kiran More, Shefali Kachroo & Nivedita Pandey, "data leakage detection," vol. 3, issue 1, pp. 315-321.

[8] Jagtap N.P., Patil S.S. And Adhiya K. P., "Implementation Of Guilt Model With Data Watcher For Data Leakage Detection System," Volume 4, Issue 1, 2012.

[9] Rohit Pol, Vishwajeet Thakur, Ruturaj Bhise, Prof. Akash Kate , " Data leakage Detection," International Journal of Engineering Research and Applications (IJERA) ISSN: 2248- 9622 ,Vol. 2, Issue 3, May-Jun 2012, pp. 404-410.

[10] Naresh Bollam, Mr. V. Malsoru, "REVIEW ON DATA LEAKAGE DETECTION," International Journal of Engineering Research and Applications (IJERA),Vol. 1, Issue 3, pp.1088-1091.

[11]Unnati kavali, tejal abhang, mr. vaibhav narawade / international journal of engineering research and applications (ijera) " data allocation strategies in data leakage detection"

[12] Ingemar J. Cox, Ton Kalker, Georg Pakura and Mathias Scheel,"Information Transmission and Steganography", Springer, Vol.3710/2005, pp. 15-29.

[13] Lobo Guerrero, A., Marques, F., Lie nard, P.B.J., "Enhanced audio data hiding synchronization using nonlinear filters",

[14] K. Gopalan, et al, "Covert Speech Communication Via Cover Speech ByTone Insertion", Proceeding of IEEE Aerospace Conference, Big,Sky,MT, March 2003.

[15] K.Gopalan and S.Wenndt,"udio Steganography for Covert Data Transmission by Imperceptible Tone Insertion", WOC 2004, Banff, Canada July 8 10, 2004.

[16] Gang. L, A.N. Akansu, M. Ramkumar, "MP3 resistant oblivious steganography", Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, UT, Vol. 3, pp.1365-1368, 7-11 May 2001.

[17] X. Dong, M. Bocko, Z. Ignjatovic, "Data hiding via phase manipulation of audio signals", IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 5, pp. 377-380, 17-21 May 2004.

[18] Deshpande Neeta, Kamalapur Snehal, Daisy Jacobs, "Implementation of LSB Steganography and Its Evaluation for Various Bits", 2004.

[19] K.B.Raja, C.R.Chowdary, Venugopal K R, L.M.Patnaik, "A Secure Image Steganography using LSB, DCT and Compression Techniques on Raw Images", IEEE-0-7803-9588-3/05/$20.00 ©2005.

[20] Vijay Kumar Sharma, Vishal shrivastava, "A Steganography Algorithm for Hiding Images by improved LSB substitution by inize detection."Journal of Theoretical and Applied Information Technology, Vol. 36 No.1, ISSN: 1992-8645, 15th February 2012.

[21] Po-Yueh Chen and Hung-Ju Lin, "A DWT Based Approach for Image Steganography", International Journal of Applied Science and Engineering 4, 3: 275-290, 2006.

[22] Hassan Mathkour, Batool Al-Sadoon, AmeurTouir, "A New Image Steganography Technique",IEEE-978-1-4244-2108-4/08/$25.00 © 2008.

[23] NageswaraRaoThota, Srinivasa Kumar Devi eddy, "Image Compression Using Discrete Cosine Transform", Georgian Electronic Scientific Journal: Computer Science and Telecommunications, No.3 (17), 2008.

[24] MamtaJuneja,Parvinder Singh Sandhu, "Designing of Robust Image Steganography Technique Based on LSB Insertion and Encryption", International Conference on Advances in Recent Technologies in Communication and Computing, 2009.

[25] Dr. EktaWalia, Payal Jain, Navdeep, "An Analysis of LSB & DCT based Steganography" , global Journal of Computer science & technology, Vol. 10 Issue 1 (Ver 1.0), April 2010.

ISSN: 2278 – 1323

AUTHOR'S PROFILE

Ashwini Palimkar---Received her B.E.degree in computer engineering from MGM'S college of engineering Nanded, India
Currently pursuing MTECH Computer engineering from bharati vidyapeeth college of engineering pune.Her area of interests include steganography and network security.

Dr. S. H. Patil working as a Professor and Head of Department in Computer engineering, Bharati
Vidyapeeth Deemed University college of Engg, Pune-43. He is having total 22 years of teaching experience & working as HOD from last ten years.

ISSN: 2278 – 1323