IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 12, December 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 12, December 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0912038 |

Securing Digital Forensics Data

Santhosh B

Associate Professor, Dept. of MCA, AIMIT, St Aloysius College (Autonomous), Mangalore, India

ABSTRACT: Digital Forensic is a process of using specified methodologies, techniques and tool to identify, extract and analyze data found in digital media . The various data analysis and examination reports can be presented as reliable evidence in the court of law. The security of the data collected is very important because based on which further analysis has been carried out. This research paper identifies key aspects of storing data securely. Here different layers of security has been identified. The proposed model could be usedon hybrid cloud.

KEYWORDS: Cyber Forensic, Digital Forensics; Distributed Database; Digital Evidence , Encryption , Steganography

I. INTRODUCTION

Digital forensic science is a branch of forensic science that focuses on recovering and investigating data from digital devices used in crimes. This is done so that, if necessary, evidence can be presented in a court of law. Digital forensics is becoming an increasingly important part of law enforcement agencies and enterprises as society's reliance on computer systems and cloud computing grows. Digital forensics is concerned with the identification, preservation, examination, and analysis of digital evidence, both inside and outside of a court of law, using scientifically established and proven techniques. [1]

Security Defensive Models

The Lollipop Model and the Onion Model are the two security defensive models. The Lollipop Model is a Defense Model that is inspired by the analogy of a Lollipop. A lollipop has a chocolate center and a layer of crust around it that is largely sugar-flavored syrup. The lollipop is eaten continuously until the chocolate at the center is revealed. Using the Lollipop comparison to the Model, the hacker only must breach one layer of security to have access to the asset, in this example, the Username and Password. The hacker can then get access to the asset. As a result, network security and the Lollipop Model are incompatible.

The Onion Model, a defense concept based on an onion analogy, is used in this study. An onion is a vegetable with several layers. We can only reach the center of the onion by removing each layer. To get access to the asset, the hacker must first breach all levels of security. In this case, breaching each layer should be difficult and time-consuming for the hacker to enter. As a result, the onion model has gained acceptance as a solid network security paradigm. As a result, this study focuses on the onion model of defense. [2]

Cloud Storage Services [15]

Cloud also provides storage as a service which provides superior economies of scale. The categories of cloud storage are managed and un-managed cloud storage. Managed cloud storage is like SaaS and fully managed by the cloud service provider. Unmanaged cloud storage is like IaaS and managed by the users. Since size of the forensics data is very huge, one of the solution is use cloud storage to store the data. These storage has the characteristics like fault tolerance, high reliability and availability. It solves some of the issues of storing digital Evidence. Some of the cloud storage service providers are AWS , iCloud, Google Drive, Dropbox etc

Data Fragmentation [16]

Data Fragmentation is the method to fragment the data into multiple pieces. the different types of data fragmentation methods are horizontal fragmentation, vertical fragmentation and hybrid fragmentation. Fragments created must satisfy three important properties – completeness , disjoint ness and reconstruction. Since the size of digital evidence is very large it can be divided into multiple fragment using these methods according to the requirement.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 12, December 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0912038 |

II. DIGITAL FORENSICS STEPS

The digital forensic investigation process requires systematic and well defined steps to collecting the data from crime scene , transfer of data to investigation agency server and analysis of data. Broadly these steps can be divided as follows

- Identification: the main two phases are identification of crime and digital evidence.
- Collection: In this phase, an investigator collects digital evidence from the crime scene for analysis and examination. It includes collecting the evidence ,secure transfer and store of evidence in investigation agency server. Some times it is hard to collect the evidence . in that case system/device is shifted to the investigation agency office.
- Extraction: this phase deals with extraction of information from various devices
- Analysis: In this phase investigator performs various types of analysis. Report thus generated could be used to n prove or disprove criminal charges.
- Examination: In this phase investigator the investigator extracts and inspects characteristics of the data.
- Report: finally report has been created to present their findings from their forensic analysis.

III. PROBLEM DEFINITION

One of the main problem in the investigation agency is how to store the data securely. In the literature most of the work has been carried out collection and analysis of the digital evidence. once the digital evidence has been collected from the crime scene securing evidence also been equally important. To fill this research gap this paper deals with how to store the data securely in the investigation agency server.

IV. CHALLENGES IN STORING DIGITAL EVIDENCE

One of the main challenge here is size of the digital evidence. Diskfile or the log record collected usually a huge file (50+ GB)[12]. To make it high available and file should be stored in multiple servers. It makes the system more fault tolerant. In India day by day no of cyber crimes has been increasing exponentially. Instead of personal server ,using of cloud storage would be more economical. Another main challenge is isolation of digital evidence from one group of investigator to another.

V.THE PROPOSED METHOD

Since the size of the digital evidence is very large the proposed method uses data fragmentation and cloud storage .Since evidences are stored in the cloud storage – the third party service provider provides one more level of security- which includes data protection, high data availability and trusted access.

Different types of data collected for the forensic analysis consists of Email datasets, different types of files (PEG, ZIP, HTML, Text, Microsoft Office, MP3, MPG, WMV, PDF, and EXE) , Chat logs, Android Application Packets (APK) , Disk images, Media (pictures\video) or network traffic logs. Since size of the data sets are very large it can be fragmented and stored in distributed database , cloud storage like AWS S3,AWS EFS , AWS EBS or in managed cloud database like awsdynamoDB, Aurora etc.

Proposed Method

Step 1: fragment the forensics data
Step 2: store fragmented data in Distributed DBMS ,cloud storage or in managed cloud database
Step 3: collect information about each fragment like case no, case supervisor, date & time, case description
Step 4: Encrypt the data using symmetric key algorithm.
Step 5: compress the data using data compression technique
Step 6: Encode the compressed data in the cover image using Steganography.
Step 7: store stegno image in the local server

Reconstructing the data is the reverse procedure of proposed method. Since the metadata about the digital evidence is encoded in the image, it hides the existence of the message. It gives one layer of security. Even if the intruder guesses existence of the information inside the image, compression layer makes more difficult to retrieve the information. The proposed method uses one more encryption level of security which provides higher level of security.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 12, December 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0912038 |

VI. IMPLIMENTATION & EXPERIMENTAL RESULTS

Experimental Setup:

The Fragmentation technique used is Horizontal Fragmentation. Since data s are collected in the table format and disk image file it is easier to fragment than Vertical Fragmentation technique. Every fragment thus created must satisfy the conditions like completeness, reconstruction and disjointness [7]. In horizontal fragmentation it is easier to create fragments which satisfy the above conditions. The encryption algorithm used is AES [8]. The AES algorithm is a Symmetric cipher model which uses the same key for encryption and decryption. The algorithm can use various cryptographic keys like 128, 192, and 256 bits for encryption and decryption of data in blocks. The data compression method methods used is loss less data compression [10] and algorithm used is LZW [11]. LZW data compression first it reads a set of symbols , create groups the symbols , covert symbols into strings, and finally convert the strings into codes. The steganography encoding algorithm used is LSB[9]. To embed message in the cover image ,LSB is one the very efficient algorithm. The data set used is Emaldataset[12]. It is the data set created by cyber forensics lab, university of new Haven,US

Experimental Results:

The emaldata set is fragmented horizontally into four data sets. These files are named as dataset1.xls , dataset2.xls, dataset3.xls, dataset4.xls. The sample image of dataset1 is as shown in fig [1]. These emails were periodically sent to the joe (reciver) . it contain parts of the logfile generated by the keylogger on the MacBook Air. Data sent consist of 12 emails and the its source is digital corpora.

```
X-Account-Key: account4
X-UIDL: GmailId13834add2ec42977
X-Mozilla-Status: 0001
X-Mozilla-Status2: 00000000
X-Mozilla-Keys:
Delivered-To: joe.sum.twelve@gmail.com
Received: by 10.114.63.177 with SMTP id h17csp36946lds;
        Thu, 28 Jun 2012 13:00:04 -0700 (PDT)
Received: by 10.101.165.37 with SMTP id s37mr1641075ano.71.1340913603499;
        Thu, 28 Jun 2012 13:00:03 -0700 (PDT)
Return-Path: <root@tracys-macbook-air.local>
Received: from Tracys-MacBook-Air.local ([2600:1003:b016:ceaf:dca8:ff95:5012:998a])
        by mx.google.com with ESMTP id f18si156733ani.108.2012.06.28.13.00.02;
        Thu, 28 Jun 2012 13:00:03 -0700 (PDT)
Received-SPF: neutral (google.com: 2600:1003:b016:ceaf:dca8:ff95:5012:998a is neither
Authentication-Results: mx.google.com; spf=neutral (google.com: 2600:1003:b016:ceaf:dc
Received: by Tracys-MacBook-Air.local (Postfix, from userid 0)
    id 085D364123; Thu, 28 Jun 2012 16:00:00 -0400 (EDT)
To: joe.sum.twelve@gmail.com
Subject: Logfile
Message-Id: <20120628200001.085D364123@Tracys-MacBook-Air.local>
Date: Thu, 28 Jun 2012 16:00:00 -0400 (EDT)
From: root@Tracys-MacBook-Air.local (System Administrator)

![LogKext Daemon starting up : Thu Jun 28 15:41:39 2012]

![User 'tracysumtwelve' has logged in : Thu Jun 28 15:42:13 2012]
legalBee
otou<del><del><del>utlothunderbird.com
```

Fig.1:Dataset1 Sample

To store the datasets instances of AWS S3 is being used. Separate instances of S3 is been created to store each datasets.

Sample file about the fragment is as shown in fig[2]

```
case no:123
case supervisor emp no and name: ps121kar Mr Ramesh
case desciption: email data set has been collected from the crime scene
case fragment information: s1_dataset1,s2_dataset2,s3_dataset3,s4__dataset4
```

Fig 2: Fragment Information

The information about the fragment is encrypted using AES algorithm .The encrypted text is shown in fig [3]

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 12, December 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0912038 |

```
be5EmC6U/vFOQePf8a40l8pUCOSRQzJa7H1EUOfajyDNZtPH4UF3t+
4aKm/v9m8u28zSyqqvgdXws0LE9d28R/7nHfP4/Gzzrm8SSUbnQdO48vAAfSYdhvaSl9hCVP2X43Lu8LJm6lObHgezGIElhe4q+
6HpQ0wNjmpBS8zeG0qJrtp2MsIEYovc9kaUY33+RceP79b/DnJviXMxqeV6ZsBnktjl47KWsAD+
nVcW2SMAqLcOwZxwdZnipdf0LPSVG6t6tCxCzMDSHVaOVeA4S394q/uKY4ES1qjqFPetOHA=
```

Fig 3: Encrypted Message of the Fragment

Next the encrypted text is Compressed using LZW data lossless compression. The encrypted text is shown in fig [4]

```
01100100 01000011 01101111 01100100 01100101 00100000
10000011 01100011 10000010 01100101 01110011 00100000
01001100 01011010 01010111
```

Fig 4: Compressed message using LZW

The above data is encoded in the cover image using LSB algorithm. The fig [5] and Fig [6] represents the image before and after encoding.

Fig 5 : Before Encoding

Fig 6 : After Encoding

The encoded image is stored in the local server. Even if multiple admins allowed to access the server, without knowing key information it is not possible to get unauthorized case information.

VII.CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

In this research paper, we propose a novel methodology to store the data securely. The onion model is used to implement the security. Three levels of security has been implemented – Encryption, Compression and Steganography. To access the data all the three layers has to be compromised .since these standard algorithms already discussed in the literature ,used widely in different scientific applications, it is not easy to compromise. But practically no system provides 100 % secure still higher security could be achieved by using advanced encryption techniques , by creating different authorization permission to access central server or by isolating servers for each cases by using VMs

ACKNOWLEDGEMENTS

The authors would like to acknowledge the funding support from **MJES Minor Research Projects**grant , St Aloysius college ,Mangalore. Thanks also go to the dedicated research group in the area of cloud computing & digital forensics at the Dept of MCA, AIMIT, St Aloysius College (Autonomous), India, for many stimulating discussions. Lastly but not least the author would like to thank everyone, including the anonymous reviewers.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 12, December 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0912038 |

REFERENCES

1. M. A. Caloyannides, N. Memon and W. Venema, "Digital Forensics," in *IEEE Security & Privacy*, vol. 7, no. 2, pp. 16-17, March-April 2009, doi: 10.1109/MSP.2009.34..
2. Basant Kumar, 2019, Effective Approach Toward Intrusion Detection and Pretention Systems in Implementing Defense in Depth, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) CICTAB – 2019 (Volume 7 – Issue 04)
3. AES page available via http://www.nist.gov/CryptoToolkit. 4
4. P. Dipak, Bhagwan and Anurag, "Strengthen Data Concealing Volume Utilizing LSB Based Image Steganography Method," vol. 4, 4, April 2014
5. S. Rahimi and F.S. Haug, Distributed database management systems: A Practical Approach, IEEE, Computer Society, Hoboken, N. J: Wiley, 2010
6. H. Q. Beyers, "Database forensics: Investigating compromised database management systems", 2014.
7. Iacob, N.-M. (2011). Fragmentation and data allocation in the distributed environments. Annals of the University of Craiova, Mathematics and Computer Science Series. 38. 76-83.
8. Dworkin, M. , Barker, E. , Nechvatal, J. , Foti, J. , Bassham, L. , Roback, E. and Dray, J. (2001), Advanced Encryption Standard (AES), Federal Inf. Process. Stds. (NIST FIPS), National Institute of Standards and Technology, Gaithersburg, MD, [online], https://doi.org/10.6028/NIST.FIPS.197
9. K. Thangadurai and G. Sudha Devi, "An analysis of LSB based image steganography techniques," *2014 International Conference on Computer Communication and Informatics*, 2014, pp. 1-4, doi: 10.1109/ICCCI.2014.6921751.
10. K. Holtz and E. Holtz, "Lossless data compression techniques," *Proceedings of WESCON '94*, 1994, pp. 392-397, doi: 10.1109/WESCON.1994.403566.
11. Zhou Yan-li, Fan Xiao-ping, Liu Shao-qiang and XiongZhe-yuan, "Improved LZW algorithm of lossless data compression for WSN," *2010 3rd International Conference on Computer Science and Information Technology*, 2010, pp. 523-527, doi: 10.1109/ICCSIT.2010.5563620.
12. CinthyaGrajeda, Frank Breitinger, and Ibrahim Baggili. “Availability of Datasets for digital forensics – and what is missing”. In: Digital Investigation (2017). (Presented at DFRWS 2017, Austin, TX)
13. A. Al-Dhaqm et al., "Digital Forensics Subdomains: The State of the Art and Future Directions," in IEEE Access, vol. 9, pp. 152476-152502, 2021, doi: 10.1109/ACCESS.2021.3124262.
14. D. Kim, Y. Pan and J. H. Park, "A Study on the Digital Forensic Investigation Method of Clever Malware in IoT Devices," in IEEE Access, vol. 8, pp. 224487-224499, 2020, doi: 10.1109/ACCESS.2020.3043939.
15. Zenuni, Xhemal&Ajdari, Jaumin& Ismaili, Florie&Raufi, Bujar. (2014). Cloud storage providers: A comparison review and evaluation. 883. 272-277. 10.1145/2659532.2659609
16. A. H. Al-Sanhani, A. Hamdan, A. B. Al-Thaher and A. Al-Dahoud, "A comparative analysis of data fragmentation in distributed database," *2017 8th International Conference on Information Technology (ICIT)*, 2017, pp. 724-729, doi: 10.1109/ICITECH.2017.8079934

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details