**DIGITAL FORENSICS & CYBER CRIME DATAMINING**

**Guided By**

**Dr. Rakesh Kumar**

**Assistant Professor, Department- CEA**

**Year of Submission- 2022-2023**

**Submitted to**

**Department of Computer Engineering and Applications**

**GLA University, Mathura**

Submitted by-

**Aditi Saraswat, 20153002, 6th Semester**

**Kamini Sengar , 201530015, 6th Semester**

**Priyanka Singh, 201530021, 6th Semester**

**Chapter Index Page No**

Abstract 1

Certificate 2

Declaration 3

List of Abbreviations, Tables and Figures 4

1. Introduction 5
2. Literature Reviews with Table Summary 6
3. Research Gap Motivation 7
4. Proposed Methodology with Diagram 8
5. Results & Experiments 9
6. Conclusion & Future Work 10

References

**ABSTRACT**

Digital forensics is the science of identifying, extracting, analyzing and presenting the digital evidence that has been stored in the digital devices. Various digital tools and techniques are being used to achieve this. Our paper explains forensic analysis steps in the storage media, hidden data analysis in the file system, network forensic methods and cyber crime data mining. This paper proposes a new tool which is the combination of digital forensic investigation and crime data mining. The proposed system is designed for finding motive, pattern of cyber attacks and counts of attacks types happened during a period. Hence the proposed tool enables the system administrators to minimize the system vulnerability.

#### BONAFIDE CERTIFICATE

Certified that this project report **“DIGITAL FORENSICS & CYBER CRIME DATAMINING”** is the bonafide work of “**ADITI SARASWAT, KAMINI SENGAR, PRIYANKA SINGH”** who carried out the project work under my/our

|  |  |
| --- | --- |
| supervision.    <<Signature of the HOD>>  **SIGNATURE** : ROHIT AGRAWAL  **HEAD OF THE DEPARTMENT**  <<Department>> Computer Science Engineering (CSE) | <<Signature of the Supervisor>>  **SIGNATURE** : Mr. Rakesh Kumar    <<Name>>  **SUPERVISOR**  <<Academic Designation>>  **ASSISTANT PROFESSOR**    <<Department>> Computer Science Engineering (CSE) |

**DECLARATION**

I hereby declare that the project work entitled

**“DIGITAL FORENSICS & CYBER CRIME DATAMINING”**

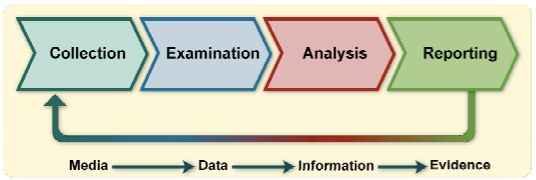
submitted to the Department of Computer Engineering and Application

GLA University, Mathura is a record of an original work done by me & my team mates. It gives us a great sense of pleasure to present the synopsis of the BTech mini project undertaken during BTech IIIrd Year. This project is going to be an acknowledgement to the inspiration, drive and technical assistance will be contributed to it by many individuals. We owe special debt of gratitude to Mr. Rakesh Kumar (our mentor) for providing us with an encouraging platform to develop this project, which thus helped us in shaping our abilities towards a constructive goal and for his constant support and guidance to our work.

1. **Introduction**

Computer forensics is the process that applies computer science and technology to collect and analyze evidence which is crucial and admissible to cyber investigations. Network forensics is used to find out attackers’ behaviours and trace them by collecting and analyzing log and status information. A digital forensic investigation is an inquiry into the unfamiliar or questionable activities in the Cyber space or digital world. The investigation process is as follows (As per National Institute of Standards and Technology) [1]. Figure 1 shows the complete phases of Digital Forensic investigation processes.

* Collection phase: The first step in the forensic process is to identify potential sources of data and acquire forensic data from them. Major sources of data are desktops, storage media, Routers, Cell Phones, Digital Camera etc. A plan is developed to acquire data according to their importance, volatility and amount of effort to collect [2]. Examination: Once data has been collected, the next phase is to examine it, which involves assessing & extracting the relevant pieces of information from the collected data [2].
* Analysis: Extracted and relevant data has been analyzed to draw conclusions. If additional data is sought for detail investigation will call for in depth data collection.
* Reporting: This is the process of preparing and presenting the outcome of the Analysis phase.
* Digital Forensic Science covers Computer forensics, Disk forensics, Network forensics, Firewall forensics, Device forensics, Database forensics, Mobile device forensics, Software forensics, live systems forensics etc.

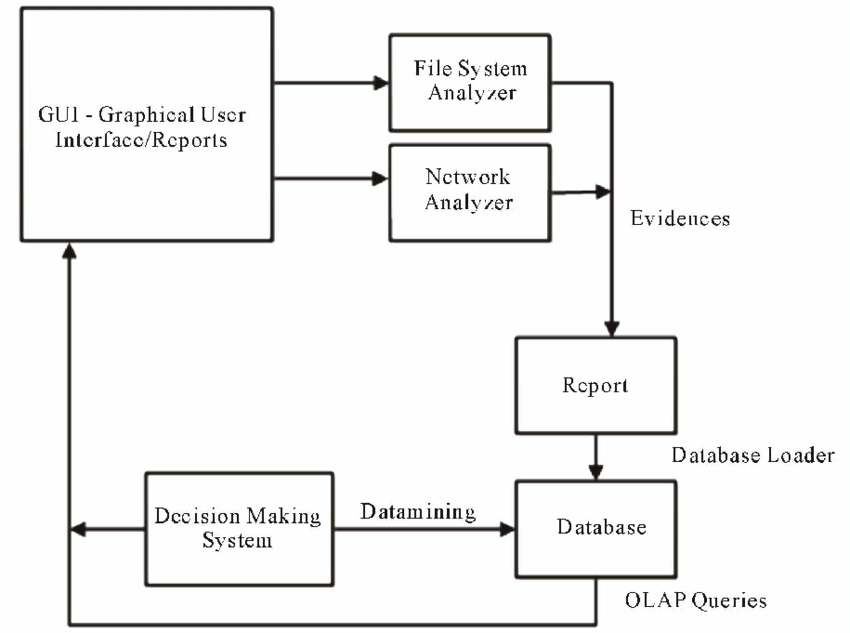


1. **Literature Reviews** (Minimum 8 research papers**) with Table Summary**
2. **Research Gap Motivation**
3. **Proposed Methodology with Diagram**

Our proposed model is the combination of digital forensics and data mining. Our proposed system helps to increase the security of the organization. When an incident reported, it investigates and report is saved in the database. Using crime data mining tool the nature of the attack is identified and alert administrator about similar attacks in future. Proactive measures can be initiated to prevent future cyber attacks. Figure shows the Block diagram of our proposed tool.

**Block Diagram of the Proposed System Graphical User Interface**: It is used by the forensic investigator to enter case details and apply tools (File system, Network) to collect evidences. Investigators can input their queries in the system. This also displays the result of the query in the form of Bar chart or report. It is the presentation layer of our three tier architecture.

* **File System Analyser:** This tool Collects evidence from the file system, it recovers all files, searches data in the free space, slack spaces and deleted spaces.
* **Network Analyser**: This tool collects data from the network traffics and server log files.
* **Database:** Database loader collects evidences from the above tools and loader loads into the database as attributes of the tables. OLTP (Online Transaction Processing): Set relations between the tables of the detected crime attributes. This applies data mining and extracts of required data. OLAP (Online Analytical processing) apply analytical queries and retrieves the output/decisions. Database server helps to store and retrieve crime attributes and results.
* **Decision Making System:** This module applies data mining algorithm and also SQL queries into the database and generates reports.

km

**Block diagram of proposed system**

1. **Results & Experiments**

analyzeMFT can produce output in CSV or bodyfile format.

CSV output

---------

The output is currently written in CSV format. Due to the fact that Excel

automatically determines the type of data in a column, it is recommended that

you write the output to a file without the .csv extension, open it in Excel, and

set all the columns to "Text" rather than "General" when the import wizard

starts. Failure to do so will result in Excel formatting the columns in a way

that misrepresents the data.

We could pad the data in such a way that forces Excel to set the column type correctly

but this might break other tools.

GUI:

You can turn off all the GUI dependencies by setting the noGUI flag to 'True'. This is for installations that don't want to install the tk/tcl libraries.

1. **Conclusion and Future Work**

This paper explains the hidden evidence acquisition from file system. Second section explains investigation on the Network. There are two types of investigation in network, live data acquisition (Packet capturing and analysis) and log file analysis. Third section explains crime data mining. On the basis we propose a new system with Digital forensic tool for decision making in the computer security domain.

**FUTURE WORK:**

1) Figure out how to write the CSV file in a manner that forces Excel to interpret the date/time

fields as text. If you add the .csv extension Excel will open the file without invoking the import

wizard and the date fields are treated as "General" and the date is chopped leaving just the time.

2) Add version switch

3) Add "mftr" switch - produce MFT Ripper compatible output

4) Add "extract" switch - extract or work on live MFT file

5) Finish parsing all possible attributes

6) Look into doing more timeline analysis with the information

7) Improve the documentation so I can use the structures as a reference and reuse the code more effectively

8) Clean up the code and, in particular, follow standard naming conventions

9) There are two MFT entry flags that appear that I can't determine the significance of. These appear in the output as Unknown1 and Unknown2

10) Parse filename based on 'nspace' value in FN structure

11) Test it and ensure that it works on all major Windows OS versions

12) Output HTML as well as CS

13) If you specify a bad input filename and a good output filename, you get an error about the output filename.

**References**

[1] K. Kent, S. Chevaller, T. Grance and H. Dang, “Guide to Integrating Forensic Techniques into Incident Response,” NIST SP800-86 Notes, 2006.

[2] S. K. Brannon and T. Song, “Computer Forensics: Digital Forensic Analysis Methodology,” Computer Forensics Journal, Vol. 56, No. 1, 2008, pp. 1-8.

[3] D. Klieiman, K. Timothy and M. Cross, “The Official CHFI Study Guide for Forensic Investigators,” 2007.