IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Volume-5, Issue-3 E-ISSN: 2347-2693

Imminent accession of Artificial Intelligence based Forensic Exploratory with Data Mining Analysis

S. Umar [1*], A. Praveen[2], S. Gouse[3], N. Deepthi[4]

[1*] Department Of Computer Science Engineering, MLRIT, Hyderabad
[2] Department Of Computer Science Engineering, IARE, Hyderabad
[3] Department Of Computer Science Engineering, MLRIT, Hyderabad
[4] Department Of Computer Science Engineering, CMR ENGG & TECH, Hyderabad

[*]*Corresponding Author: umar332@gmail.com*

Available online at: www.ijcseonline.org

Received:21/Feb/2017 Revised: 02/Mar/2017 Accepted: 14/Mar/2017 Published: 31/Mar/2017

Abstract— Data mining is part of the interdisciplinary field of knowledge discovery in databases. Data Mining research began in 1980 and has grown rapidly in 1990s.Specific methods developed in disciplines such as artificial intelligence, machine learning and pattern recognition used in data mining. Data mining has been introduced in various sectors. key functional area of mining technology of the World Wide Web Recently mining techniques applied to the data in the field of criminal law, but that digital forensics. Examples can be found misleading to establish criminal identity criminal groups involved in illegal activities and much more. the politics of data mining technology typically generate a summary of large amounts of data. Digital Forensics is the area of research discoveries and advanced tip. Canvass search field, and digital forensic applications are developing rapidly with the economic giant digital information. law enforcement and military agencies have confidence in digital forensics today. Since the age of information is the speed of thought and data stored in digital form, the need for an accurate intellectual interception, and timely decision errors close to zero digital data processing cores issue. This article will research focusing on the role of data mining techniques for digital forensics. also identifies how data-mining techniques can be applied in the field, a digital forensic forensic examiner to take the next step in the process, which is cost-effective digital program is a crime.

Keywords— ***Data Recovery, Forensic exploratory, Digital Forensic***

I. INTRODUCTION

Expert in the use of forensic identification, storage, retrieval and documentation of digital sources of successful persecution of digital evidence [1], the scientific method. Digital Forensics, the true sense of the answer when, what, digital divide, where, how and why [2]. If the search on the computer, for example, the word "when" the time interval occurs during the activity. "We" for the activity on the computer system. Responsible "for the" concern "about" the expression, in which the evidence is to go into the "how" the manner in which the activity takes place and the "why" to determine the motive of the crime.

Forensic expert technology audits can be used to create digital crime or an accident. The purpose of research, and truth, which often leads to prosecutions and convictions. Significant increase in the number of digital development has been caused by a range of legal instruments. The device provides a numerical proof found and maintain, and that the results support numerical proofs of precision processing [3]. Can occur have been developed in the form of software, and such devices to help carry out a digital study, digital tester.

The integration of useful data mining techniques described in the digital forensics process. This helps to improve the performance and reliability of the test object. Including the formal methodology for data extraction are the basic steps [4] as follows:

Define the type of representation and the structure of the data sets. Determine how the measurement data; Compare the different appropriate data representations
Select algorithmic process to use the evaluation function
Determine the type of algorithm for executing the actual data management principles.

The data mining functions are also used for different types of samples. The first part of this work is to study and the

existing information in the field of forensic numerical review. After examining the thread of traditional forensic tools. The analysis results are used for classification codes of basic legal instruments. This ranking is based on recognition of the limitations of the tools and recommendations of computer evidence at the time by the development of advanced data mining techniques to present the latest tools proposed for digital forensics.

II. EXISTING SYSTEM USED IN FORENSIC

Literature, computer evidence tools are essentially classified

1. Judicial equipment
2. Legal software tool

The equipment can be used for legal means of territorial parts or complete systems and computer servers. The software that is used in forensic applications of command line and graphical applications 90s.These powerful advertising and development tools that began in 1980 and is generally divided into three categories, as follows.

1. Common legal tools: Tools such as a series of tests, especially the search for keywords in digital media.
2. Specialized legal instrument: It puts the tested focus to the forensic material every image, illustration or Internet. One of the most important tools is often based on production.
3. Management tools: these are used to monitor, control and report cases.

Individual needs led to the creation of forensic computer forensics computer science, in the form of computer software. Make sure that the tools necessary to achieve and maintain properly maintained the integrity of the digital evidence, digital evidence. For example, copying and pasting data cannot by other means of communication be recognized by the court as reliable forensic evidence. This is because the process cannot copy and paste data to modify, for example, to change the data time-stamp. Thus, a typical digital precision test set (or bit-stream) a copy of all data storage media. This is exactly the copy to each of the image called, and the process images are often referred to as images. Forensic computing tool especially for the retrieval of digital evidence, in other words, to come back after the media data. The device is usually the ability to assist in the analysis of limited data recovery. All information provided by misleading computer forensics tools, sometimes. The reason for this is that the size, complexity and amount of data to computer forensics tools to researchers. Not the data and the conclusions presented appear. At present, computer forensics tool is not suitable solution performs the following operations:

- The relationship between the Association identification data.
- Classification: the recognition and classification of data into groups on the basis of the same data.
- Before locating and displaying groups of strangers or left unnoticed; cluster
- The forecast data to find reasons and cause a reasonable forecast.

III. FORENSIC TECHNIQUES INTERNAL METHODOLOGY

A. Imaging

One of the first techniques used in digital image forensic examination, or copy, should be included in the media. Although this seems simple first steps file system runs the modern operating system (OS), many related functions, such as sorting or magazines solutions. Without treatment, the material cannot be changed, but slightly, and may risk the integrity of the proof. [5]

B. Hashing

To quickly identify the files, and that the image file's reliability or change, judge the encryption approved. To use a one-way hash function modern cryptographic hash function to find. The unique character depends on the cryptographic hash function used. MD5 hash value when the decision was founded in 1991, Ron Rivest and quickly took the legal community. NIST will soon decide SHA-1 in which federal regulations [2].

C. Sculpture

A category of tools for forensic digital tool called a sculptor file. These tool analyzes allow disk blocks that do not include the existing file to access the deleted data. The use is known as the sculptor's head and foot signed the connection node "unused" cancellation of the original file [6]. Carving can recover deleted files, but files are not overwritten and temporary media. Analysis of Mikus sculpture techniques in 2005 [6]. Recent advances in the incision, allowing files to recover to a more fragmented. Garfunkel sculpture file real confirmation procedure [7].

IV. ROLE OF DATA MINING IN FORENSIC TECHNIQUES

Data Mining and Soft Computing for many applications in forensic. This can be the right data correlation (Association) data models to track records and because the right information that the group composition number (classification), the hidden positioning of the fact (optional), and Trace profit forecasts can (forecasts) [8]. Even though the ideal method association, classification, grouping and unpredictable, it is very helpful to look. [9]

The display allows the digital quest for basic information of interest to be found quickly and efficiently. In addition, the best digital step researchers in the search, so that digital evidence of effective and effective recovery from stroke. [10]

In 2003, the Artificial Intelligence Laboratory at the University of Arizona, case studies are reported in the COPLAN-related task. This project is particularly interested in information overload an effective analysis of crime and terrorism, law enforcement and prevent national security personnel. Her work using the mining industry to help solve the problem. In its report, determine the context of data mining and analysis of criminal intelligence units, including mining, clustering techniques, error detection, classification, and finally with respect to the rope .

Four case studies in the report shows how useful it was extraction extract descriptive report of the identity fraud investigation of the police information analysis, analysis of copyright cyber-crime. And finally, to analyze the criminal network. Today the program is implemented COPLAN and works to meet, share and search for information. Any Hewlett create online databases and save [11] Mining Packard data to solve the problem to find the same folder in 2005, using major library materials [12].

After the analysis of the different groups of files associated, and further improved by a pair of graph partitioning algorithm [13].

In 2006, Galloway and Sim off try more network approach Case Study Data Mining. Their work to extract specific information from their networks, for example between unique data warehouse fluid networks. [14]. Shatz, Mohay and Clark in 2006 developed a discovery method for the majority of digital proofs of time-stamping data relationships. Work on the time-stamping of the important and important impact of digital forensic investigations on several occasions to the complexity of the problem of time zone differences, drift, prejudice and possible human intervention to administer. [15].

2006 Abraham showed the exploitation of proletarians' event research to develop forensic computer science purposes. Abraham analyzed using the computer to search for information based on the owners or users of the profiles at the sequence of events that occur to examine the system. Abraham profile owners are divided into four different categories. Materials, Objects, Actions and Time Stamps [1]. In 2007, Beebe and Clark invited to work for the restoration and research chain fights the digital forensics team. Although their work focuses on text exploration, data collection has shown that the search for information and effectiveness of the algorithm effectiveness of their efforts. [16]

V. CONCLUSION

More and computer forensic tools on the market, it is important to know the different features in this area. The goal is to provide an overview of the current capacity of the computer forensic tools. It takes into account the limitations of registered funds and recommendations. The device is usually the ability to help in the analysis of the limited data recovery. Data mining techniques that are currently used in order to come up with a variety of devices. That may be the various aspects of digital forensics solution mining techniques to study existing data. The analysis of the revision of digital forensics tools and techniques available.

REFERENCES

[1] A.J. Marcella, R.S. Greenfield, *"Cyber forensics: a field manual for collecting, examining and preserving evidence of computer crimes*", Auerbach, 2002.

[2] J. Guynes, C. Nicole, L. Beebe, "*Digital forensics text string searching: Improving Information retrieval effectiveness by thematically clustering search results*", In 6th Annual Digital Forensic Research Workshop, vol. 4, pages 49–54, 2007.

[3] J.G. Heiser, "Computer forensics: incident response essentials", Addison- Wesley, 2002.

[4] DFRWS, "*A road map for digital forensic research*", DTR - T001-01 FINAL - DFRWS Technical Report, 1(1), August 2001. PDF.

[5] M. Usama, *"Summary from the kdd-03 panel: data mining: the next 10 years*", SIGKDD Explor. Newsl., 5(2): 191–196,. ISSN 1931- 0145, 2003.

[6] H. Jan, Kroeze, C. Machdel, Matthee, J.D. Theo, "*Differentiating data- and Text- mining terminology*", In SAICSIT '03: Proceedings of the annual research conference of the South African institute of computer scientists and information technologists on Enablement through technology, South African Institute for Computer Scientists and Information Technologists, Republic of South Africa, pp.93–101, 2003., 2003.ISBN 1-58113-774-5.

[7] L. Simson, Garfinkel, "*The advanced forensic format*", 2008

[8] A. Clark, B. Schatz, G. Mohay, "*A correlation method for establishing Provenance of timestamps in digital evidence*", 6th Annual Digital Forensic Research Workshop, In Digital Investigation, volume 3, supplement 1, pages 98–107, 2006.

[9] C. Apt´e, S. Weiss (1997), "*Data mining with decision trees and decision rules*", Future Generation Computer System., 13(2-3):197–210, 1997.

[10] H. Atabakhsh, "C*rime data mining: an overview and case studies*", Proceedings of the 2003 annual national conference on Digital government research, pages 1–5, 2003

[11] P. Smyth, D. Hand, H. Mannila, "*Principles of Data Mining*", The MIT Press, 2001.

[12] O. de Vel, A. Anderson, M. Corney, G. Mohay,"*Mining e-mail content for author identification forensics*", SIGMOD Rec., 30(4):55–64, ISSN 0163-5808, 2001.

[13] T. Abraham, "*Event sequence mining to develop profiles for computer forensic investigation purposes*", In ACSW Frontiers '06: Proceedings of the 2006 Australasian workshops on Grid computing and e-research, Australian Computer Society, Inc., Darlinghurst, Australia, Australia, pp.145–53, 2006.

[14] G. Forman, K. Eshghi, S. Chiocchetti, "*Finding similar files in*

large document repositories.", In KDD '05: Proceeding of the eleventh ACM SIGKDD international Conference on Knowledge discovery in data mining, ACM, New York, NY, USA, pp.394–400, 2005., ISBN 1-59593-135-X.

[15] J. Galloway, J. Simeon, Simoff ,"*Network data mining: methods and techniques for discovering deep linkage between attributes*", In APCCM '06: Proceedings of the 3rd Asia-Pacific conference on Conceptual modelling, pages 21–32. Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 2006. ISBN 1-920-68235-X.

[16] L. Simson, Garfinkel, "*Forensic feature extraction and cross-drive analysis*", Digital Investigation, 3(Supplement-1):71–81, 2006.

