ENHANCEMENT OF TIMELINE ANALYSIS FOR DIGITAL FORENSICS

[1]SANDEEPAK BHANDARI, [2]VACIUS JUSAS

[1,2]Software Engineering Department Kaunas University of Technology Lithuania
E-mail: [1]sandeepak.bhandari@ktu.edu, [2]vacius.jusas@ktu.lt

Abstract- The field of digital forensics has developed to enable security experts to look at proof from the expanding plenty of computerized gadgets to help figure out what people may have done before. In this paper, we present a methodology based on abstraction of information provided about timeline of events and artefacts at four different levels such as event high level, event low level, artefact high level and artefact low level for enhancement of timeline analysis. In our approach two command based digital forensics tools are also used such as log2timeline and psort to collect data from various files systems found on typical computer for implementing the developed methodology. From results, it is observed that developed methodology can enhance the timeline of events and artefacts by increasing the visualization of timeline, realistic of data and minimize the time utilized for collection and understanding of digital evidences from digital devices.

Keywords- Digital Forensics, Timeline, Events, Artefacts, Digital Evidence and Visualization

I. INTRODUCTION

Digital investigation (sometimes known as digital forensics science) is a branch of forensics science encompassing the recovery and investigation of material found in digital devices, often in relation to computer crime.The term digital forensics was initially utilized as an equivalent word for PC legal sciences yet has extended to cover examination of all gadgets equipped for putting away computerized information. Computerized crime scene investigation examination has an assortment of uses. The most widely recognized is to help or negate a speculation before criminal or common (as a major aspect of the electronic revelation process) courts. The computerized examination process can be arranged into four unique stages as gathering, examination, investigation and reporting. The key terms in digital forensics are timeline, visualization ,events and artefacts.Timeline is defined as a display of list of events in chronological order. It is regularly a visual computerization appearing long bar named with dates nearby itself and for the most part events.In digital forensic,timeline assist examiner with visualizing the occasions more detailed,accurately and complete.Using timetable examination amid examinations can be incredibly helpful yet it some of the time misses critical occasions that are put away inside documents on the speculate framework (log records, OS ancient rarities).In digital forensics, an event can be characterized as an activity performed by clients utilizing digital devices.

The term “artefact” currently does not have a formal definition within the domain of cyber/digital forensics, resulting in a lack of standardized reporting, linguistic understanding between professionals, and efficiency. Generally, artefact can be defined as a piece of data that may or
may not be relevant to the investigation / response. Examples include registry keys, files, time stamps, and event logs and in other words can be defined as something observed in a scientific investigation or experiment that is not naturally present but occurs as a result of the preparation or investigative procedure.

The principle goals of this examination ponder are rightto comprehend the different difficulties in computerized legal sciences to gather and translate the advanced confirmations from advanced gadgets. The primary difficulties are heterogeneity and huge measure of information, touchy development of web and its clients, assembling of advanced gadgets, absence of institutionalization and absence of laws and directions.Secondly, to work with digital forensics tools to get experience and collect timeline of events and artefacts from various file systems which is discussed in detail in section 4 of this research article. The final objective is to develop research methodology which is based on the various challenges faced by digital forensics science and timeline of events and artefacts gathered by using digital forensics tools and how research methodology is capable of enhance the timeline of events and artefacts which is discussed in detail in next sections of the paper.

II. LITERATURE REVIEW

Brady et al.[18] authors propose an ontology the Digital Evidence Semantic Ontology (DESO).It allows an examiner to quickly discover what artefacts may be available on a device before time-consuming processes are commenced. Then it is used for classification of data. DESO ontology use the concept of Gene Ontology (GO), which annotates the work of various genetic researchers using a common vocabulary to allow comparison. In this way, individual researchers can use their own terminology but can query otherareas of research using this common vocabulary to discover findings that may assist them. DESO ontology contains two main

classes namely Artefact Location and Artefact Type.Brady et al. [19] authors used proposed Digital Evidence Semantic Ontology (DESO) from paperBrady et al. [18] for handling with growing variety and volume of digital evidence in a criminal investigation.Buchholz and Falk [20], Zeitline is an open source graphical tool written in Java [20,3] that allows a forensic investigator to import various events from a computing system and then order and classify them into one or more timelines of events. In this paper authors discuss the design, features of Zeitline.

Zeitline is a graphical timeline editor that allows a forensic investigator to create a timeline of events that were gathered from di erent sources. The main design objectives of Zeitline are:

- Generating events from arbitrary data sources
- Grouping events together into logical groups recursively •
- Filtering of the data that is displayed •
- Locating specific events •
- Intuitive interface •
- Platform-independent implementation

Forensic analysis is the science of collecting, examining, and presenting evidence to support or refute a hypothesis Inglot and Liu [17]. There are two approaches for timeline analysis.The first one is to use applications that have been specifically created to accelerate the analysis of a timeline (or any sort of textual log file), and the other is more labor intensive, that is, a combination of command line tools and spreadsheet application. One of the biggest challenges in Digital investigation is volume of data that need to be analyzed [17-20,11,7]. Hargreaves and Patterson [11], authors propose a framework for extract low level events to an SQLite backing store which is automatically analyzed for pattern to generate high level events. TimeLine analysis find out when event took place and quickly lead you to the evidences that needs further analysis. Log2timeline is framework design for problem analysis, extract timestamps from various solutions and output them in various formats Guðjónsson[2].

During past two decades to assist computer forensic investigators variety of digital forensics tools have been developed but most of the tools are not able to handle vast volume of data efficientlyKhan et al.[3]. In this paper authors use the neural network methodology for digital forensic analysis by formulating a timeline of relevant activities take place on computing machine by tracing the previous file system activities. Traditional analysis is immensely manual process which is not suitable for growing vast amount of data James et al. [7]. In this, authors propose a signature-based approach for high level user event reconstruction from a collection of low-level traces found during a post-mortem forensic analysis of a system.

Cho [15], authors propose an intuitive computer forensic method by timestamp changing patterns of operations on file in windows NTFS file system. Chabot et al. [14], authors combine the knowledge of experts from field of digital forensics and software development to allow a semantically rich representation of events related to the incident to develop a new methodology. This methodology also meets the requirements of the courts such as reproducibility, verifiability, validation etc. Marrington et al. [10], authors work with detecting inconsistency in timeline. For this author developed a software tool i.e. Computer Activity Timeline Detection (CAT) for the detection of inconsistency within timelines of computer activity..

Bang et al. [9], authors work with different file systems such as NTFS and FAT and analyse the changes in the time information of files and folders for different operations and attempts to reconstruct the user's actions. Prasad and Satish [5], authors work on reconstruction of events based on crime event and event characteristics in digital forensics. Authors Kalber et al.[4], presented a prototype of a forensic fingerprinting framework called Py3xF.It is possible to generate profiles of user activities in software applications by using Py3xF.Py3xF can automatically derive unique fingerprints – if such exist – based on timestamp information stored in the file system metadata. Moreover, Py3xF can extract the file system metadata of given disk image files or parse XML representations of already extracted metadata information generated by popular forensic toolkits like EnCase or The Sleuthkit.

Guðjónsson [2], author address the problem of traditional timeline analysis such as easy to manipulate, sensitive to changes and not updated always and required manual labour. To address these problems author developed Log2timeline a framework designed for timeline analysis. This can be used on Mac OS X, Linux and Windows OS X. In this paper explain in detail about Log2timeline framework including designing, developing and installation. James and Gladyshev [1], authors work on the timely investigation of requests. For this author, proposes the use of signature-based methods for automated action instance approximation to automatically reconstruct past user activities within a compromised or suspect system. Cho [15,16], authors develop a computer forensic method for detecting timestamp forgeries in Windows NTFS file system to address the problem of classic timestamp in digital forensic science. The method is based on past and present timestamp change patterns by file operations. Authors also develop logical rules for detecting file time forgery. Khan et al. [3], authors develop a methodology for post event re-construction based on artificial neural networks. The main aim of this

methodology is to handle the large volumes of data efficiently. This approach based on the monitoring the file system manipulation, capturing file systems snapshots at discrete interval of time to characterise the use of different software applications and then using this capture data to train a neural network to recognise the execution patterns of application programs.

Olajide et al. [6], authors reviewed about forensic live analysis and event re-construction methods in Linux systems. Khan et al. [8], authors proposed a methodology based on machine learning for post event timeline re-construction, the focus of this methodology is to outline the effectiveness of employing machine learning methodology for computer forensic analysis by tracing previous file-system activities and preparing a timeline of the events Awawdehet al. [12], authors show preliminary results for a real time computer forensics agent that logs computer activity on a Windows computer system for subsequent investigation and developed an agent using .NET 2010 framework having six modules. Among these, two modules are implemented and tested with carefully designed scenarios using Windows XP and Windows 7 operating systems.Battistoni et al. [13], authors work on integrity of timeline. Firstly, authors reviewed on issue related to the cloud machine time reliability then introduce novel architecture (CURE) aimed at providing timeline resilience to cloud nodes.Lillis et al.[21], authors explore the current challenges contributing to the backlog in digital forensic such as complexity of problem, diversity of problem, consistency and correlation problem, volume problem and unified time lining problem. Authors also outlines several future research topics that could greatly contribute to a more e□cient digital forensic process.

Sitompulet al. [22], authors developed Aho-Corasick parsing technique to read file attributes from the master file table (MFT) to examine the file condition. If the file was deleted, then the parser searches the file content to reconstruct the file.

III. METHODOLOGY

From the literature studies, it is studied and found that the immense amount and heterogeneity of data are the two main challenges in digital investigation for digital investigator to collect digital evidences from digital devices with in time with high understandability and visualization of timeline of events and artefacts. The research methodology is based on and to deal with immense amount and heterogeneity of data. Basically, the research methodology is split into four different sub parts with the different level of abstraction of data provided at each sub parts namely Events: High Level (New entries and Web Surfing), Events: Low Level (Web Surfing and actions of modifying), Artefact Location: High Level and Artefact Location: Low Level to provide highly recognizable, compact and visualization of timeline of events and artefacts.

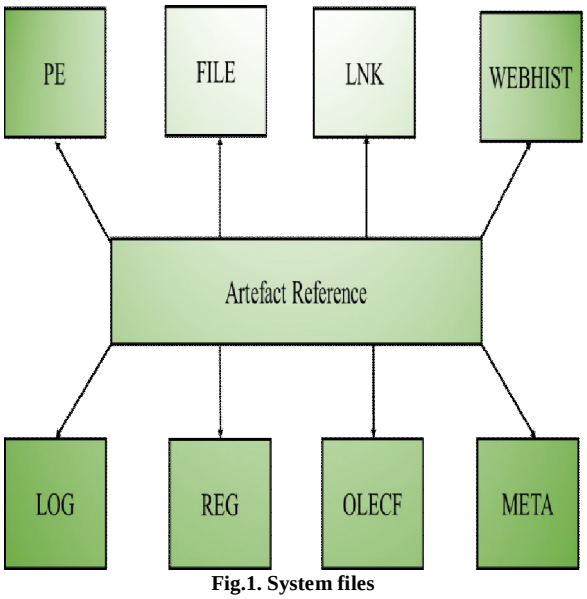


Fig.1. System files

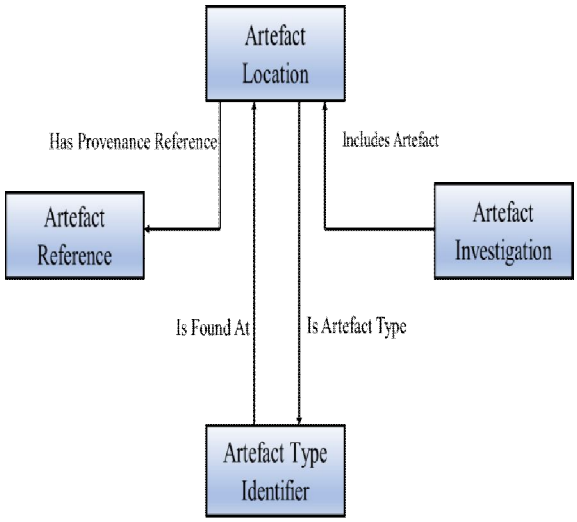


Fig.2. Different Modules of Methodology

IV. RESULTS AND DISCUSSION

To extract the timeline of artefacts and events from various file systems found on a typical computer, two open source and command based digital forensics tools are used Log2timeline and Psort as shown in below Fig.3. These two tools are used together to generate timeline. In our case, we used 40 GB and 110 GB of Hard disk image files to generate timeline by using these two tools it includes 17 different fields ,9 different sources of data and 3,13,596 number of events and artefacts, which is challenging, time consuming and labor intense work to collect and analysis of digital evidences. To overcome all these challenges, we developed methodology which is based on different level of abstraction of timeline provided at four different levels of events and artefacts. Not only abstraction, we also used Data

Deduplication (Redundancy Elimination) technique and alteration and eliminate unneeded information from timeline of artefacts and events to compose highly visulaization of timeline and collectable and analysis of digital evidences from digital devices. To implement the research methodology Core Java Language is used and four different modules are developed to implement the research methodology.

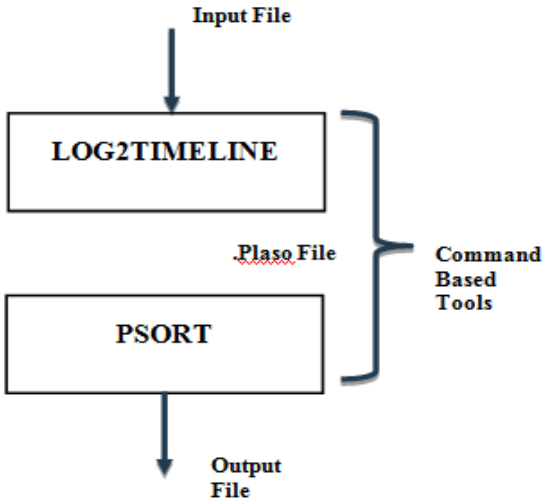


Fig.3. Command Based Digital Forensics Tools

In this paper, we consider a case source "WEBHIST"and source type "Chrome History"to demonstrate the outcome of methodology at four different levels of abstraction.Although there are distict number of
sources and source types which is not possible due to limitation of number of pages for paper.This case depicts about web history which is demonstrate by field Source "WEBHIST" from Google Chrome browser shown by field Source Type "Chrome History " at the artefact high level and artefact low level.

Date:11/13/2017
Time:19:01:19
Source:WEBHIST
Short:https://www.magnetforensics.com
Visit:LINK
Event Number:255696

Fig.4. High Level Event

At each level different level of abstraction of information provided from general to specific. In the above figure 4 the higher level of abstraction of information of timeline is provided among all four level.

Date:11/13/2017
Time:19:01:19
Source:WEBHIST
Short:https://www.magnetforensics.com
Visit: LINK
Extra:https://www.magnetforensics.com/computer-forensics
Event Number: 255696

Fig.5. Low Level Event

In the event low level of timeline more detail is provided as compared to event high level of timeline. From the event high level, it is shown that when and what kind of event is performed by the user at the typical computer system demonstrate by three fields namely "Date", "Time " and "Short"but low-level event provide more detail and information about the event by field "Extra"which present the title of the webpage accessed by the user and the sub web page as shown in above figure 5.To provide more detail and information the number of fields are increased at and add needed information at each level of abstraction. At the artefact high level, number of fields increased up to 10 to provide more detail and relevant information. At this level not only, what type of action is performed

Date:11/13/2017
Time:19:01:19
MACB:.A..
Source:WEBHIST
Source Type: Chrome History
Short:https://www.magnetforensics.com/computer-forensics/how-to-analyze-usb-device-...
Visit:LINK
Extra:https://www.magnetforensics.com/computer-forensics
Desc.:https://www.magnetforensics.com/computer-forensics/how-to-analyze-usb-device-history-in-windows/ (How to Analyze USB Device History in Windows - Magnet Forensics Inc.) [count: 0] Host: www.magnetforensics.com Type: [LINK - User clicked a link] (URL not typed directly - no typed count)
Event Number: 255696

Fig.6. Artefact High Level

by user is shown but also shows how the user perform the perform action and how many numbers of time

user perform the action demonstrated by fields "Short"and "Desc."as shown in figure 6.It is note that at each level field "Event Number"is used for the reference purpose to the event.The last level of abstraction and developed methodology is artefact low level. More detail is provided as compared to previous level about action such as last visited time of the web page accessed by the user as shown infigure 7.The developed methodology not only enhance the timeline of events and artefacts but also increase the visualization, reduce labor intense work, noise reduction and reduce time to collect the digital evidences from digital devices.



Fig.7. Artefact Low Level

CONCLUSION

This research aimed to firstly analysis the problems with collection of digital evidences from digital devices for digital forensics such as heterogeneity of data, amount of data due to immense growth of internet and its users and manufacturing of electronic devices, and rapid changes in technologies which increase labor intense work , reduced visualization of timeline of events and artefacts and increase time. Secondly, based on these problems formulation from literature studied, a research methodology is developed. The research methodology is developed into four different modules namely "high level event, low level event, artefact high level and artefact low level" based on the level of abstraction of information about timeline of events and artefacts (from general level to specific level) provided. At last, from result it is found that developed methodology increases the visualization of timeline, reduce noise and create more realistic and effective timeline which help digital investigator to understand and work with digital evidence. In the future, an ontology will be develop based on the outcome from the research methodology.

REFERENCES

[1] J. I. James and P. Gladyshev, "Automated inference of past action instances in digital investigations," Int. J. Inf. Secur., vol. 14, no. 3, pp. 249–261, 2015.

[2] K.Guðjónsson, "Mastering the Super Timeline Who am I ?," 2010.

[3] M. N. A. Khan, C. R. Chatwin, and R. C. D. Young, "A framework for post-event timeline reconstruction using neural networks," Digit. Investig., vol. 4, no. 3–4, pp. 146–157, 2007.

[4] S. Kalber, A. Dewald, and F. C. Freiling, "Forensic application-fingerprinting based on file system metadata," Proc. - 7th Int. Conf. IT Secur. Incid. Manag. IT Forensics, IMF 2013, pp. 98–112, 2013.

[5] M. A. R. Prasad and Y. N. Satish, "Reconstruction of Events in Digital Forensics," vol. 4, no. 8, pp. 3460–3467, 2013.

[6] F. Olajide, N. Savage, D. Ndzi, and H. Al-Sinani, "Forensic live response and event reconstruction methods in Linux systems," Proc. Converg. Telecommun. Netw. Broadcast., pp. 141–146, 2009.

[7] J. I. James, P. Gladyshev, and Y. Zhu, "Signature based detection of user events for post-mortem forensic analysis," Lect. Notes Inst. Comput. Sci. Soc. Telecommun. Eng. LNICST, vol. 53, pp. 96–109, 2011.

[8] M. N. Khan, E. Mnakhansussexacuk, and I. Wakeman, "Machine Learning for Post-Event Timeline Reconstruction," PGnet, no. January 2006, pp. 1–4, 2006.

[9] J. Bang, B. Yoo, J. Kim, and S. Lee, "Analysis of time information for digital investigation," NCM 2009 - 5th Int. Jt. Conf. INC, IMS, IDC, pp. 1858–1864, 2009.

[10] A. Marrington, I. Baggili, G. Mohay, and A. Clark, "CAT detect (computer activity timeline detection): A tool for detecting inconsistency in computer activity timelines," Digit. Investig., vol. 8, no. SUPPL., 2011.

[11] C. Hargreaves and J. Patterson, "An automated timeline reconstruction approach for digital forensic investigations," Digit. Investig., vol. 9, no. SUPPL., pp. S69–S79, 2012.

[12] S. Al Awawdeh, I. Baggili, A. Marrington, and F. Iqbal, "Towards a unified agent-based approach for real time computer forensic evidence collection," Proc. 2013 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Min. - ASONAM '13, no. August, pp. 1492–1493, 2013.

[13] R. Battistoni, R. Di Pietro, and F. Lombardi, "CURE—Towards enforcing a reliable timeline for cloud forensics: Model, architecture, and experiments," Comput. Commun., vol. 91–92, pp. 29–43, 2016.

[14] Y. Chabot, A. Bertaux, C. Nicolle, and M. T. Kechadi, "A complete formalized knowledge representation model for advanced digital forensics timeline analysis," Digit. Investig., vol. 11, no. SUPPL. 2, pp. 95–105, 2014.

[15] G. S. Cho, "A computer forensic method for detecting timestamp forgery in NTFS," Comput. Secur., vol. 34, pp. 36–46, 2013.

[16] G. S. Cho, "An Intuitive Computer Forensic Method by Timestamp Changing Patterns," 2014 Eighth Int. Conf. Innov. Mob. Internet Serv. Ubiquitous Comput., pp. 542–548, 2014.

[17] B. Inglot and L. Liu, "Enhanced Timeline Analysis for Digital Forensic Investigations," Inf. Secur. J., vol. 23, no. 1–2, pp. 32–44, 2014.

[18] O. Brady, R. Overill, and J. Keppens, "Addressing the increasing volume and variety of digital evidence using an ontology," Proc. - 2014 IEEE Jt. Intell. Secur. Informatics Conf. JISIC 2014, no. December, pp. 176–183, 2014.

[19] O. Brady, R. Overill, "DESO: Addressing volume and variety in large scale criminal cases," Elsevier, vol. 88, no. 4, pp. 72–82, 2015.
[20] F. Buchholz and C. Falk, "Design and Implementation of Zeitline: a Forensic Timeline Editor.," Dfrws, pp. 1–7, 2005.
[21] D. Lillis, B. A. Becker, and M. Scanlon, "Current Challenges and Future Research Areas for Digital Forensic Investigation," 2016.
[22] S. Sitompul, A. Handoko, and R. F. Rahmat, "File Reconstruction in Digital Forensic," TELKOMNIKA (Telecommunication Comput. Electron. Control., vol. 16, no. 2, p. 776, 2018.

★★★