Using DEMF in Process of Collecting Volatile Digital Evidence

Miroslav Bača*, Jasmin Ćosić ** , Petra Grd*
* Faculty of organization and informatics, Center for Forensics, Biometrics and Privacy, Varaždin, Croatia
** Independent researcher, Bihać, Bosnia and Herzegovina
miroslav.baca@foi.hr, jasmin.cosic.ba@gmail.com, petra.grd@foi.hr

Abstract – Acquisition of volatile data for further forensic analysis still represents a challenge to both practitioners and researchers. The current tools used for acquisition of such data are focused exclusively on a way to capture content. However, the development of forensic science, in particular in the area of digital evidence in terms of the admissibility in court, has introduced additional elements to be evaluated. Mainly, the integrity of the collected digital evidence, authenticity and other elements of the digital chain of evidence to be presented in court. This paper describes a framework for capturing volatile data using Digital Evidence Management Framework (DEMF) with regards to integrity of captured data.

I. INTRODUCTION

The problems researchers and practitioners face in the context of respecting digital chain of evidence[1] for the acquisition of volatile digital evidence lies in the fact that with existing tools it is not possible to present evidence in court and confirm their integrity and completeness. It is relatively easy, and so far it is accepted by the courts, to confirm the integrity of the file as well as the media using hash functions. But even "standard" static data analysis leads to different results from the same values, which can be seen in the analysis conducted by Zimmerman [1]. Exploring a variety of tools for media acquisition as well as different methods these tools use, results were quite different. For example, using the X-Ways Imager tool using DD format, obtained size was 1,000,204,886,016 bytes, whereas using the E01 method size was 1,000,449,415,330 bytes. The difference between formats can also be observed by using other tools like FTK where DD image was 1,000,204,886,016 bytes, while E01 image was 1,000,571,704,661 bytes. In instances where EnCase Imager tool was used for image format E01 value obtained was 1,000,571,589,032 bytes. Although it was the same media, the results indicate that different tools in the same format have very different values of the media image size. As the size of the media is directly related to the hash function calculation, conclusion is that neither the hash function for the same medium with different tools is the same.

Even greater difference occurs with the acquisition of volatile data, such as data collected from RAM or data collected from computer networks. Other that it is unknown how large the data is initially, it is rather difficult to calculate the hash function and confirm the integrity of the data. To solve this problem to some extent, researchers have begun to develop algorithms that would be able to calculate the hash function of data whose length is not known. Such functions usually take parts of the same size and sign them with hash functions.

The question posed here is whether it is possible, only on the basis of such data, to guarantee the integrity of the data.

II. DIGITAL EVIDENCE MANAGEMENT FRAMEWORK

In order to solve the problem described in the introduction, some solutions are suggested to deal with the digital chain of evidence of which the DEMF is one of the last and most comprehensive in terms of confirming the integrity of the collected data. Elements of DEMF are described in [2], [3] and they use hash functions as one of the mechanisms to confirm the integrity of the data. DEMF with its concept answers the questions who, what, when, how, where and why, which fully ensures the integrity of digital evidence, but also its confidentiality and availability. Figure 1 shows a rough diagram of the DEMF concept where two elements that are most important for the process of proving integrity are marked red and that are *report* and *judge*.

The concept of DEMF is universally applicable and allows different variations according to the area of

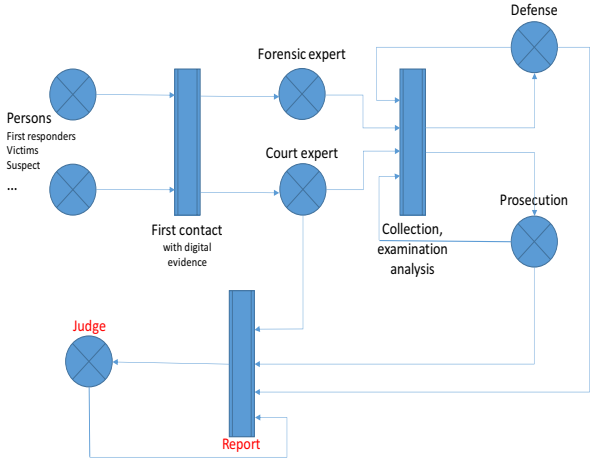

Figure 1 Processes in DEMF model

[1] Chain of Custody is sometimes called Chain of Evidence in literature and scientific papers

application. Further in this paper, the application of the DEMF concept on files whose content and metadata are not known will be described, which is especially important for collection, testing and analysis modules.

The importance of this is shown through the UML

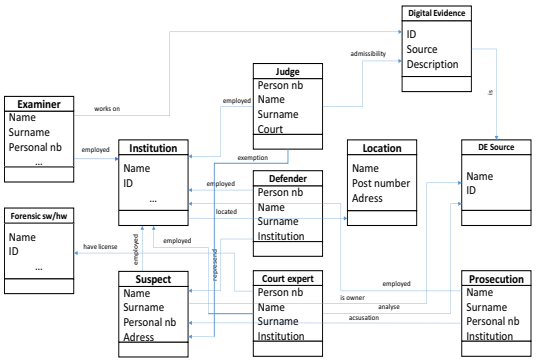

Figure 2 DEMF UML model

concept in Figure 2.

As it can be seen from the Figure 2, the system of integrity proofing in the court as well as the chain of evidence is based solely on the digital evidence and therefore the way to collect digital evidence is the most important factor in this process.

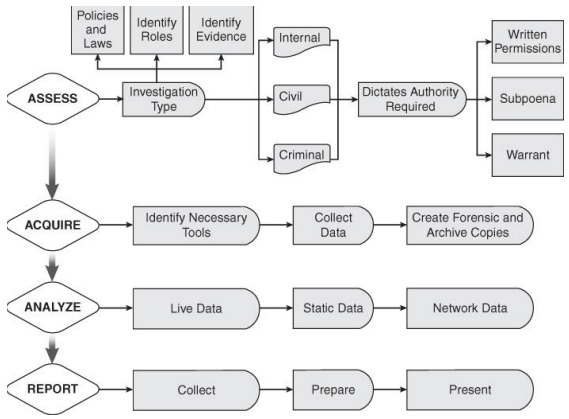

Figure 3 Forensic examination steps [6]

III. MEMORY DATA

Collecting memory content, along with the collection of the hard disk content and other holders of information, has become mandatory in the process of working with digital evidence. In order to collect the contents of RAM, it is necessary to run an application that will allow the collection of the data and that will affect the content of what is collected. Tools used to collect the contents of RAM are mainly open source based tools that do not necessarily comply with the technical standards regulating the collection of memory contents. A small number of commercial companies develop tools for the acquisition of RAM and are advertised as a standard for that purpose. Encouraged by the results obtained by Zimmerman[2] [1] and which are mentioned in the introduction and refer to the established methods, it can be concluded that the difference resulting from the use of various tools for the acquisition of RAM, is much higher than is the case with the acquisition of hard drives.

The volatility of the RAM content is not constant and depends on a number of different parameters like the amount of RAM and the number of running processes as well as to the characteristics of other peripherals that are in direct interaction with RAM memory. Therefore, the framework proposes, for working with the RAM memory, that the content is divided into blocks of fixed memory. This blocks will, during storage, be signed with all the elements contained in the DEMF.

This proposal has its negative aspects also. First of all, this refers to the fact that during memory acquisition a larger number of elements than is currently the case will be changed. The reason for this is the longer acquisition time required by DEMF in comparison with other state-of-the-art methods. Furthermore, an additional problem can be later work with a larger amount of elements. Today's tools save RAM memory as a single file which can facilitate the work if the files are of smaller memory range. However, in the case of files that have higher memory range (256 GB and larger), this can generate big problems.

In addition, there is the question, if the file is divided into smaller units, how to ensure the integrity of all the individual parts, and whether the process itself affects the quality of captured RAM images.

The methods used today for RAM acquisition have their advantages and disadvantages. There are two types of RAM acquisition: software and hardware. Software acquisition is widespread today, which can be used to acquire the content of RAM through tools such as dd. The problem with them is that one piece of content (possible evidence) will be rewritten because the tool itself must be installed in RAM. The amount that will be lost depends of course on the size of the memory itself, and the smaller it is, the larger part of content will be deleted. Hardware methods have different constraints and are less widespread than software. They rely on the fire wire for acquisition because it is possible to make a Direct Memory Access via IEEE 1349 [4]. This method is much better than software because it does not cause changes in RAM but it requires certain prerequisites that must be met and which are not common in the daily work (for example, the user should not shut down the computer, the computer has to be unlocked and the computer should have a specific type of fire wire adapter).

IV. COMPUTER NETWORK DATA

Network forensics deals with the capture, recording and analysis of network events in order to discover

[2] Results can be downloaded from https://docs.google.com/spreadsheets/d/1wXX5zYql7KIPgrsDdt6S5bTuGt_WRjWaBde1D0fhG5k/edit?type=view&gid=0&f=true&sortcolid=11&sortasc=true&rowsperpage=250&pref=2&pli=1#gid=0

evidential information about the source of security attacks in a court of law [5].

Depending on what network technology is applied to transfer data, different technologies for acquisition can be applied. For the purposes of this paper, the example of the most common method in Ethernet infrastructure will be given (for example, IEEE 802.11x WiFi). Traditional tool for data collection with this type of network technology is tcpdump. The output of this tool is PCAP file. Tcpdump can be used to filter out collected network traffic.

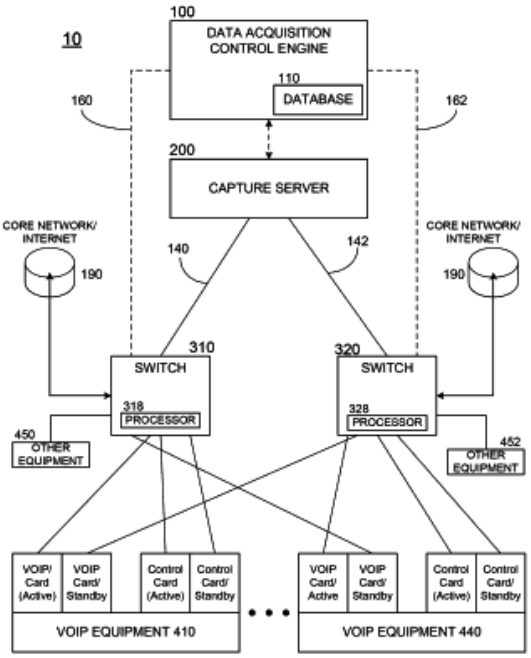

Figure 4 Patent network acquisition model [7]

Problems with data acquisition from the network lies in the fact that it is not known how much data can be collected. An example of such model is given in Figure 7.

„According to one aspect, the invention is directed to a communication network monitoring system that may include at least one switch serving as an intermediary to a plurality of data input streams and a plurality of data output streams; a capture server in communication with the at least one switch; and a data acquisition control engine operable to receive data acquisition instructions from a user and cause the received instructions to be implemented at the at least one switch. According to another aspect, the invention is directed to a method that may include presenting a graphical user interface (GUI) to a user by a data acquisition control engine, in a communications network; receiving data acquisition instructions from the user that specify a data acquisition plan; deriving commands to issue to one or more switches based on the data acquisition plan; and transmitting the derived commands to the one or more switches.“ [7]

In addition, there is a need for data acquisition from different services, for example IaaS (Infrastructure as a service). The paper [8] described a comparison of data acquisition with Amazon EC2, and it is shown how with present forensic tools data acquisition of that type is done. The differences between individual tools are pointed, from the amount of captured content to duration of data acquisition. Five different products showed similar acquisition results but still different enough that it can cause some doubt in court. On the side of the forensic process especially in an amount of acquisition time, the difference between the fastest and the slowest tool is 1:6.

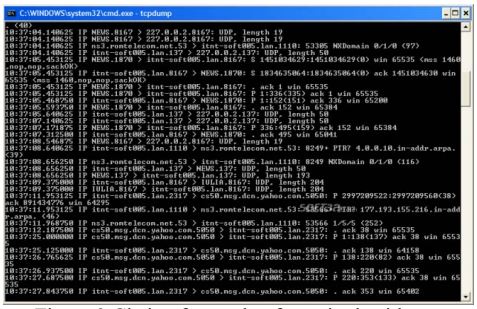

Figure 3 Chain of custody of acquired evidence

V. MOBILE DATA

The data collected from mobile phones and mobile devices are particularly sensitive to change. In the context of this study this passage refers to data that is collected during the "live" acquisition or with a mobile device that is operational and that cannot be turned off. Today's models are mainly based on two large commercial producers Cellebrite [9] and XRY [10]. In one part of mobile phones it is possible to make data acquisition using these tools without changing the content of mobile phones. It should also be noted that for certain mobile phones and operating systems on them, these commercial systems are unusable.

In order to conduct the acquisition from such devices it is necessary to use noncommercial tools, which implies the lack of standardization models as well as the testing of these tools, which can be a problem in court. One such system is Firefox OS [12], for which there is no commercial tool for the acquisition. For the purpose of the acquisition it is necessary to use alternative tools, such as [12]. However, when using such tools there is still a question as to what changed during installation of those tools on the mobile phone, whether it is possible to document and ultimately confirm the integrity of the data in court.

VI. PROPOSED MODEL

Regardless of whether the data is collected from RAM or from another source, difficulties that arise can be observed. In the first place is that the forensic scientist does not know the length of the memory to be collected, and that the length is not permanent. On the other hand, the installation of the tool itself causes changes in memory and proving the integrity of collected evidence is necessary.

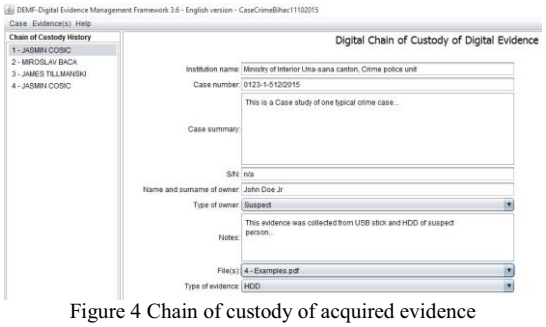

Figure 4 Chain of custody of acquired evidence

DEMF model can be used to prove the integrity of the collected evidence in court. In this sense, it is necessary to define a specific memory size of the package that will be collected. This step is necessary to reduce the impact of installed applications from the complete contents of the memory, to only part of it. As the memory changes, the overall size which is to be collected in this process should be defined.

Figure 5 Metadata acquired by using DEMF

When the individual package is collected, metadata from DEMF model need to be collected:

- specific task on the basis of which these data are collected, this may be a court order in case of official investigations by law enforcement agencies or management decisions in case of corporate investigations,
- first or/and last name of the person, or some other designation of the person who collects this data - identification can be done by a biometric characteristic or even digital signature stored on a personal documents.
- the geographic location of computer, or geo-data (latitude and longitude), and unique data on the computer to which the data is collected,
- a time stamp - the exact date and time obtained as a confirmation from an external TSA, as well as a public IP address.
- calculation of a hash function, selection from offered MD/SHA functions (to the level of 512 bytes)
- complete metadata files (author, time of creation, modification, last access, etc.)

Each individual package made is stored in a container. The containers are, at the end of the process, protected by a certain key - AES256 encryption. Key management procedures must be defined in institutions which are using DEMF, through a document called "Security policy". During the encryption process all metadata for the container, as well as for each individual package, are written (Figure 5).

In the process of proving the integrity it is possible, by analysis of each package as well as container, to read all metadata that are confirmation that the process is done in accordance with the preservation of the chain of evidence, which can guarantee the integrity of all the packages contained in a container.

Today, DEMF is initially tested in institutions in Bosnia and Herzegovina, and Croatia and some court expert witnesses for ICT are using DEMF tool to improve integrity of digital evidence acquired from seized devices.

VII. CONCLUSION

The idea of this paper was to give an overview of DEMF as a framework for managing digital evidence. The problems of collecting volatile digital evidence have been addressed, as were their possible solutions by using this model and software - DEMF. Memory and computer network data collection has been described. The main part of the paper was showing the usage of DEMF in collecting volatile digital evidence.

REFERENCES

[1] Zimmerman Eric, Speed testing of different forensic imaging tools, available at www.hecfblog.com, (accessed 01.02.2016.), 2013

[2] Ćosić, Jasmin; Bača, Miroslav; ,Do we have full control over integrity in digital evidence life cycle?,"Information Technology Interfaces (ITI), 2010 32nd International Conference on",,,429-434,2010,IEEE

[3] Ćosić, Jasmin; Bača, Miroslav; Leveraging DEMF to Ensure and Represent 5ws&1h in Digital Forensic Domain,International Journal of Computer Science and Information Security, Vol.13, No.2, pp. 7-10,2015

[4] Adam Boileau, Hit By A Bus: Physical Access Attacks With Firewire. Ruxcon, 2006

[5] Natarajan Meghanathan, Sumanth Reddy Allam and Loretta A. Moore: Tools and Techniques for Network Forensics, International Journal of Network Security & Its Applications (IJNSA), Vol. 1, No. 1, April 2009

[6] Graves, Michael Digital Archeology, The Art and Science of Digital Forensics, London, UK; O'Reilly-Media

[7] US Patent 9178791 B2, System and method for data acquisition in ana internet protocol network, 2015

[8] Dykstra, Josiah, Sherman, Alan Acquiring forensic evidence from infrastructure-as-a-service cloud computing: Exploring and evaluating tools, trust, and techniques, Digital Investigation 9 (2012) S90-S98

[9] Cellebrite forensic, available at www.cellebrite.com (accessed 30.01.2016.)

[10] MSAB Inc. availbable at www.msab.com (accessed 30.01.2016.)

[11] Mozilla, available at www.mozzila.org (accessed 15.01.2016.)

[12] GitHub repository, availbable at http://github.com/elninosi (accessed 30.01.2016.)