Using of live response technologies during the blocking of the attacks and investigation of incidents in real time

V I Petrenko, M G Ogur, M V Zubkov, O E Solgalova
North-Caucasus Federal University, Stavropol, Pushkin str., 1, 355017, Russia
E-mail mogur@ncfu.ru

Abstract. Acquisition of a clear idea of the events and artifacts occurring eventually is a difficult task which needs to be solved in digital criminalistics. Recovery of chronology of events and artifacts, allows experts to understand chronology of digital crimes and to interpret a conclusion in the form of digital proofs with use of live-response technologies.

Keywords: forensic; criminalistics; investigation of incidents, live-response technologies

1. Introduction
The digital criminalistics became new area of researches in the last two decades in connection with the explosive growth of the Internet, use of electronic devices, fast innovations in technology and the growing sizes of memory devices and also high growth of number of digital or computer crimes. The purpose of digital criminalistics is search and interpretation of origin of the event or an artifact found in a computer system. Besides, the digital criminalistics is engaged in restoration and investigation of digital proofs; in case of need it is often connected with (cybercrime). Attribution, identification of leak in the organization and assessment of the possible damage caused during violation are the most common causes of conducting digital criminalistic examination. The field of digital criminalistics is subdivided into several divisions depending on type of the digital device which is object of investigation. The computer criminalistics, network criminalistics, criminalistic examination of data and criminalistics of mobile devices belong to them. Digital investigators study data and devices to learn as much as possible about the violation or crime connected with use of digital devices. in the form of digital proofs. The carrier classifies digital criminalistics by three main stages, namely: receiving, analysis and representation [1-3].

The purpose of a stage of data collection is preservation of a state the Digital system allowing to take and analyze subsequently proofs in controlled conditions. Following a phase, i.e., the analysis, consists in studying the obtained data for identification of signs of harmful activity. The stage of the presentation is completely based on precepts of law and rules which change depending on country jurisdiction where there are digital proofs, and, thus, are beyond the present document. As a rule, at an acquisition phase the first step initiated by the investigator is obtaining images of devices of storage of the confiscated computer. The following step after an acquisition phase includes definition of proofs which can be connected with studying the data concerning investigation. In digital criminalistic investigation, as well as at usual restoration on the crime scene, it is important to have an opportunity to make the schedule of work of the file system in a computer system.
The chronology is used for crime reconstruction, and it is favorable in terms of providing to the user access to a target system, implementation of certain applications and also identification of system files and data to which access were modified or removed during the certain periods of time which are of interest to investigation was got. To provide collected proofs from the confiscated computer in court, process by means of which the evidence was obtained has to be accurately shown with observance of the adopted procedures and the practician. In legal proceedings this practice is known as "a chain of deliveries". To prove authenticity of proofs, chains of justification have to be brought into court about the scientific methods accepted in extraction of proofs. Experts face various problems, such as high speed and volumes of data, development of standards, maintaining confidentiality of investigations and growth of anti-criminalistic methods when digital or electronic devices take from them proofs [4,5].

Abstract approach to assistance to experts during the investigation by recovery of chronology of events and artifacts which are clear for investigators is presented in this article and reduce time necessary for collecting digital proofs by means of digital devices [6].

2. Problem research.
Prasad & Satish defined digital criminalistics as the procedure of identification, collecting, preservation, the analysis and submission of digital proofs thus which is legally recognized by court. For effective, effective and during the limited period of time of collecting digital proofs from digital devices the reconstruction of terms is required. In "Review" the time frames are defined as "means of identification or a linking of the sequence of events so that the persons who are responsible for response to incidents could visualize and understand easily". Harrell defined that "the analysis of time frames is an excellent equipment for determination of activity which occurred in a system during a certain period of time in a system". Thus, "creation of a timeline of various events which took place during the incident is one of the key tasks which are carried out by the expert in digital criminalistics". The analysis of time frames has to be a key component of any investigation as time frames of events are almost always relevant. The main source of chronological information are metadata of the file system. File systems trace various temporary tags and have nuances which have to be considered when carrying out the criminalistic analysis [7,8].

The Inglot & Liu company specified that in principle there are two approaches to the analysis of a timeline. First, there are applications which were specially created for the analysis of a timeline, and they to focus on visual representation of a timeline, for example, Cyber Forensics Time Lab (CFTL), Zeitline, Encase, tool kit for forensic medicine and many other things. Secondly, there is a combination of tools of a command line and applications for work with spreadsheets which are labor-consuming, for example, Log2timeline and excel together. Guðjónsson developed the tool for extraction of temporary tags from various files found on the ordinary computer. and to aggregate them. The Log2timeline tool is a part of the plazo with the back engine on the basis of Python. Purpose of these systems is to receive temporary tags in one place for computer criminalistic examination. Such timeline sometimes is called "Super-timeline" [5].

Sitompul, Handoko, and Rahmat claim that to have a clear idea of events which took place during the certain period of time it is difficult to reach in digital investigation. The restoration of events allowing the expert to understand chronology of commission of crime is one of the most important stages of digital investigation. This difficult task demands a research of a large number of events that is connected with frequent technological innovations, heterogeneity, huge number of data and also manual process of reconstruction of events which is inefficient and expensive [9-11].

Bang, etc. discussed how creation time, time of the last record and time of the last access to the file or the folder are important factors which can indicate the events which affected a computer system. They analyzed changes in time of information on files and folders for various operations of the FAT and NTFS file systems and tried to restore actions of the user. Besides, they show use of information on time for the digital analysis of proofs on the example of a specific case. Shabo, etc. was defined by two main problems (heterogeneity and volume of data) with reconstruction of events. In them the approach
supported with theoretical concepts which can help investigators throughout all process, including construction and interpretation of the events describing business is stated [12,13].

Hargreaves and Patterson considered the possibility of use of the Log2timeline tool for reconstruction of a timeline, but was decided to construct an own prototype on Python. The basis allowing to write the analysis which can lead to the event of high level based on presence of one or several events of low level became other contribution of research. A lack of this approach was the fact that analyzers which were focused on certain events have to be written in advance [14].

Brady, Overill and Keppens suggested to use digital semantic ontology of proofs (DESO) which allows the expert to find out quickly what artifacts can be available on the device. DESO is constructed on the ideas of Gene ontology (GO). The general principle which is the cornerstone of DESO is bidirectional [15,16]:

1. Experts use a certain form of classification or the system of tags which allow experts to estimate easily what artifacts are available;
2. after artifacts are taken from various sources, DESO provides means for their comparison.

The main idea is in having an opportunity to compare the artifacts taken from different sources. Brady & Overill continued development of ontology of DESO. A main objective - performance of functions of a repository and the qualifier of artifacts of the digital evidence for a possibility of comparison of the data obtained from heterogeneous sources. Research objectives are established and carried out by statement of simple questions, based on the one who that when where why and how? ("5WH"). Only a subclass "That" it was detailed. Classes "why" and "as" were not discussed at all. Realization of DESO in general was not opened.

Debinski, Breytinger and Mokhan claim that reconstruction of events is a fundamental step for experts to understand the main point, and the important tool for generation of time frames is Log2timeline. In spite of the fact that these schedules give big proofs and help to understand the event, they are difficult and demand additional tools and also educational scenarios. Moreover, some of the main restrictions of Log2timeline are also specified in them, for example, the fact that there is no simple tool in use which beginners/investigators could use for the analysis of the generated time frames, and is not present a free training material which would also allow the practicing experts to study and improve the acquaintance to Log2timeline and also to instruments of visualization.

Authors developed Timeline2GUI for support of researchers - the independent tool written in the Python language which supports the analysis of a timeline of CSV (result of Log2timeline). The tool is similar to widely used sheet Excel that provides easy transition for the practicing experts. Besides, three educational cases which are in free access were developed and can be used for improvement of skills of the analysis of time frames of the investigator or by means of the graphic Timeline2GUI interface, the sheet Excel or any other tool. Performance of Timeline2GUI can be increased due to process acceleration, for example, due to removal of irrelevant fields or association of fields in future work.

Soltani, Seno & Yazdi offered the scheme of reconstruction of events which defines whether the application was started on the compromised system. The offered structure constructed a signature or TPFSM-A (temporary template of modification of the file system of the application) and TPFSM-D (temporary template of modification of the file system of the hard drive). Besides, the framework presented a distance metrics which is used for calculation of distance between a signature of the application and TPFSM-D of the hard drive. The instrument of decision-making of a framework used the calculated distance for making decision on whether the application was started on the compromised system. Zhao and Kopliyen said that success of applications of symmetry in various scientific disciplines induced them to study the concept of symmetry in the software. Besides, authors explained that in object-oriented designs of language the concept of symmetry is used, and studied symmetry in object-oriented language and also gave some examples of symmetry outside object-oriented programming to show that reasons of symmetry are beyond object-oriented programming and extend to other fields of development of the software [17].

The review of literature shows that the set of tools and techniques was developed for support of the criminalists who are engaged in the analysis of time frames, but any of them is not able to solve a
problem of difficult creation of chronology of events and artifacts in real time. The following belongs
to number of the main causes of inefficient adjustment of terms:

1. huge number of data,
2. heterogeneity of data,
3. fast innovations of technologies,
4. rapid growth of the Internet and integrity of the schedule.

The purpose of work is development and introduction of methodology of recovery of chronology of
events and artifacts which is capable completely and it is clear to restore the timeline based on
abstraction level with use of technologies of real time and the so-called "live-response technologies"
[18].

3. Technique of creation of a timeline in real time
In the section of the review of literature various problems connected with the huge volume of data when
the timeline is taken from an image of a disk, in particular, when approaching "Super-chronometrical"
are covered. This research is based on the data provided to Log2timeline in the form of the file which is
sorted by means of the Psort tool, as shown in Figure 1.

![Figure 1. Formation of a timeline.](image)

Besides, Psort allows to transform the file to usual formats, such as CSV. The timeline contains 17
fixed fields as it is specified in table 1. Then data from the CSV file are imported to the sheet Excel. One of the most often found problems on the timeline taken by means of Log2timeline is repeated
repetition of the same unit of time. For example, unit of time "11:49:53" in the field "time". Repeated
repetitions correspond to various data sources, such as "WEBHIST" and "LNK" in the field "source"
showing that emergence of the same event caused reflection in various magazines and various artifacts,
as shown in Table 2. Such scenarios formulate chronology big, difficult and difficult analyzed for
experts. So, the corresponding scenarios were taken from the schedule and replaced with the
Corresponding methods (for example, on one event per unit of time for duplication elimination) to
reconstruct the simple, compact, recognizable and structured timeline for experts.

| Field             | Description                                   |
|-------------------|----------------------------------------------|
| Date              | Date of approach of an event                  |
| Time              | Time when there was an event                  |
| Time zone         | The time zone which was used for a tool call  |
| MACB              | Modification, access, creation and origin     |
| Source            | Short name of a source, for example, records of the register - REG |
| Source type       | Description of a source                       |
| Type              | Time tag type, for example, last access or last record |
| User              | What user name is associated with an event if that is available |
| Host              | what name of a host is associated with record  |
| Short description | contains the field of the short description in which the text is stored |
| Description       | here the most part of the sorted information is stored |
The methodology by means of which it is expedient to represent chronology of events and artifacts is presented using the concept of abstraction for expansion of a timeline of events and artifacts on four various levels of abstraction, i.e., events: High level; events: Low level; location of an artifact: High level; and location of an artifact: Low level, as shown in figure 2, for creation of a timeline, readable and clear for criminalists. The ideas which are the cornerstone of breakdown of the terms presented to Log2Timeline at four levels of abstraction give different idea to information. For each level the structure and also separate levels of specification of information has to be specified to reduce complexity of the schedule, to pass undesirable details, to provide correctness of the schedule, and submission only of that information which will help to distinguish and understand specific actions. It is carried out by users by the analysis of various sources and fields.

Table 2. A timeline with repetition (duplication) of the same unit of time

| Date          | Time     | MACB | Source       | Sourcetype                | Type                        |
|---------------|----------|------|--------------|---------------------------|-----------------------------|
| 19 November 2017 | 11:49:53 | .A.  | WEBHIST      | Firefox History           | Last Visited Time           |
| 19 November 2017 | 11:49:53 | .A.  | LNK          | Windows Shortcut          | Creation; Last Access Time  |
| 19 November 2017 | 11:49:53 | MA.B | LNK          | Windows Shortcut          | Content Modification; Creation; Last Access Time |
| 19 November 2017 | 11:49:53 | .A.  | LNK          | Windows Shortcut          | Creation; Last Access Time  |
| 19 November 2017 | 11:49:53 | .A.  | WEBHIST      | Firefox History           | Last Visited Time           |
| 19 November 2017 | 11:49:53 | MA.B | LNK          | Windows Shortcut          | Content Modification; Creation; Last Access Time |
| 19 November 2017 | 11:49:53 | MA.B | LNK          | Windows Shortcut          | Creation; Last Access Time  |

In a forensic the event can be characterized as the activity which is carried out by users with use of digital devices. On the other hand, "artifact" has no formal definition in limits area a cyber / digital criminalistics now that results in lack of the standardized reporting and linguistics for mutual understanding between professionals.

As a rule, the artifact can be defined as a part of data which can have or not have relations to investigation/reaction.

For example: register keys, files, temporary tags and logs.

In other words, artifacts can be defined as something observable in scientific research or an experiment which is made as a result of the preparatory or investigative procedure.
The chosen methodology of a research assumes development of four various modules, each of which corresponds to one of levels of abstraction of chronology of events and artifacts. At each level of methodology various types of sources, parameters and fields for drawing up advanced structure and representation of a timeline are considered. On actions: information of high level connected with the activity which is carried out by users such as creation of files, access to web pages and loading of information in the form of files of various formats is provided (jpeg, pdf and docx). Such information with higher level of abstraction is taken by the analysis of six various sources - "LNK", "LOG", "META", "OLECF", "PE" and "WEBHIST" - and five certain areas, namely: "date", "time", "source", "visit" and "exile" that allows to obtain necessary information for recognition of concrete types of activity. Besides, the quantity of the sources and areas considered for the analysis depends on the type of information and the level of its specification demanded at the concrete level of a time scale.

For obtaining detailed information the different types of operations connected with files are carried out namely: modification, access, change and origin and also actions on web surfing which concern concrete web pages to which the user got access, and loading of files by the user [12].

The second level of abstract events: Low level, is reconstructed with additional details in comparison with the previous events: High level, for the analysis is required additional quantity of sources and fields. At this level nine various sources and seven area with expansion of three new sources, including "FILE", "REG" and "RECBIN" and also a new field, "additional", respectively, are studied for reconstruction of a time scale. The first two levels of abstraction of methodology can provide the smallest volume of information, related to what types of operations are performed by the user. To get a clear idea of a timeline, additional information is required. Here the list of all .exe-files started by the user, applications, and files which often address authors of files, the list of files enters. Namely, automaticDestinations-ms provides information on the used application which is attached to the taskbar of the user, temporary tags and also ways to elements (documents, web pages, images, etc.) which the program addressed recently while .msp provides updating of information relating to the Windows operating system and other Microsoft .msi programs containing information on installation for the specific program, such as files expecting installation. Besides, the detailed information concerning various events held on the Internet is provided. For example, to what application program (web browser) and as (the LINK user clicked on the link, GENERATED - chose record from the list, RELOAD user rebooted the page and gathered TYPED user. The URL address in a line of the URL address) each concrete web page is available. This detailed information is provided on an artifact: High level: "sourcetype" and "desc" (description) with lower value abstraction level is reconstructed by studying nine various sources and ten fields with addition of three new fields of "MACB". At various levels of abstraction, the relevant information joins and undesirable details for reconstruction of the corresponding timeline taking into account various number of sources and the field are removed. At the end the timeline with the smallest level of abstraction from all levels is reconstructed: Low level - by the analysis of the available all nine sources.
and seventeen fields by means of a method and also addition of 9 fields, namely: "time zone", "type", "user", "host", "version", "file name", "code", "notes", and "format" for providing the full detailed information concerning all operations performed by the user.

The offered methodology consists of four various levels of abstraction of events and artifacts. Each level provides various information connected with various actions, carried out by the user, with different levels of abstraction of information, since minimum and finishing with the maximum details to reconstruct a relevant and recognizable timeline for the best understanding by the practicing digital expert.

4. Conclusion

In digital criminalistics of one of the main tasks the reconstruction and the analysis of terms, events and artifacts for interpretation of digital proofs in connection with a huge number and a variety of data is. Researches of scientific and technical literature show that various approaches were developed to show it, but most of them are not able to resolve an issue finally. In this work the flexible approach based on four various levels of abstraction, chronology of events and artifacts, such as events was considered: High level (new records and web surfing), events: Low level (web surfing, actions for modification), arrangement of artifacts: High level (including all .exe-files), and arrangement of artifacts: Low level. At each level of the schedule the unique structure which gives information relating to the concrete operation performed by the user is observed. For a timeline various parameters, fields and criteria (for example, on one event in unit of time), and various relevant structures of data excluding irrelevant for reconstruction of a compact and effective time scale with the accurate level of abstraction from details are considered to show the structured timeline which helps the investigator to understand and quickly to analyze information on incident. For realization of the offered approach and check of results various experiments are made. Results show that abstract approach is capable to recreate the chronological sequence of events and artifacts effectively, irrespective of quantity and heterogeneity of data, in comparison with the existing methods.

References

[1] Carrier, B. 2019 Open Source Digital Forensics Tools
[2] Olajide, F.; Savage, N.; Ndzi, D.; Al-Sinani, H. 2018 Forensic Live Response and Event Reconstruction Methods in Linux Systems
[3] Thomas, D.S.; Forcht, K.A. 2004 Legal methods of using computer forensics techniques for computer crime analysis and investigation (Issues Inf. Syst. 2004, 5) pp 692-698.
[4] Harrell, C. 2011 What’s a Timeline
[5] Esposito, S.; Peterso, G. 2013 Creating Super Timelines in Windows Investigations (Proceedings of the 9th International Conference on Digital Forensics, Orlando, FL, USA, 28-30 January 2013) pp. 135-144.
[6] James, J.I.; Gladyshev, P. 2015 Automated inference of past action instances in digital investigations (Int. J. Inf. Secur. 2015, 14) pp 249-261. [CrossRef]
[7] Inglo, B.; Liu, L. 2014 Enhanced Timeline Analysis for Digital Forensic Investigations. (Inf. Secur. J. Glob. Perspect. 2014, 23) pp 32-44. [CrossRef]
[8] Guðjónsson, K. 2010 Mastering the Super Timeline Who Am I?
[9] Sitompul, O.S.; Handoko, A.; Rahmat, R.F. 2018 File Reconstruction in Digital Forensic (TELKOMNIKA Indones. J. Electr. Eng. 2018, 16) pp 776-794. [CrossRef]
[10] Cho, G.S. 2013 A computer forensic method for detecting timestamp forgery in NTFS (Comput. Secur. 2013, 34) pp 36-46. [CrossRef]
[11] Kalber, S.; Dewald, A.; Freiling, F.C. 2013 Forensic application-fingerprinting based on file system metadata (Proceedings of the Seventh International Conference on IT Security Incident Management and IT Forensics, Nuremberg, Germany, 12-14 March 2013) pp 98-112.
[12] Bang, J.; Yoo, B.; Kim, J.; Lee, S. 2009 Analysis of time information for digital investigation (Proceedings of the Fifth International Joint Conference on INC, IMS and IDC, Seoul, Korea, 25-
27 August 2009) pp 1858-1864.

[13] Chabot, Y.; Bertaux, A.; Nicolle, C.; Kechadi, M.T. 2014 *A complete formalized knowledge representation model for advanced digital forensics timeline analysis* (Digit. Investig. 2014, 11) pp 95-105. [CrossRef]

[14] Hargreaves, C.; Patterson, J. 2012 *An automated timeline reconstruction approach for digital forensic investigations* (Digit. Investig. 2012, 9) pp 69-79. [CrossRef]

[15] Brady, O.; Overill, R. 2015 *DESO: Addressing volume and variety in large scale criminal cases* (Digit. Investig. 2015, 88) pp 72-825. [CrossRef]

[16] Brady, O.; Overill, R.; Keppens, J. 2014 *Addressing the increasing volume and variety of digital evidence using an ontology* (Proceedings of the 2014 IEEE Joint Intelligence and Security Informatics Conference, Hague, The Netherlands, 24-26 September 2014) pp 176-183.

[17] Debinski, M.; Breitinger, F.; Mohan, P. 2019 *Timeline2GUI: A Log2timeline CSV parser and training scenarios* (Digit. Investig. 2019, 28) pp 34-43. [CrossRef]

[18] Soltani, S.; Seno, S.A.H.; Yazdi, H.S. 2019 *Event reconstruction using temporal pattern of file system modification* (IET Inf. Secur. 2019, 13) pp 201-212. [CrossRef]