A Forensics Activity Logger to Extract User Activity from Mobile Devices

Priscila Cedillo
Dept. of Computer Science
Universidad de Cuenca
Cuenca, Ecuador
priscila.cedillo@ucuenca.edu.ec

Jessica Camacho
Faculty of Engineering
Universidad de Cuenca
Cuenca, Ecuador
jessica.camachoc@ucuenca.edu.ec

Karina Campos
Electronical, Electronics and Telecommunications Dept.
Universidad de Cuenca
Cuenca, Ecuador
karina.campos@ucuenca.edu.ec

Alexandra Bermeo
Faculty of Engineering
Universidad de Cuenca
Cuenca, Ecuador
alexandra.bermeo@ucuenca.edu.ec

***Abstract*—Nowadays, mobile devices have become one of the most popular instruments used by a person on its regular life, mainly due to the importance of their applications. In that context, mobile devices store user's personal information and even more data, becoming a personal tracker for daily activities that provides important information about the user. Derived from this gathering of information, many tools are available to use on mobile devices, with the restrain that each tool only provides isolated information about a specific application or activity. Therefore, the present work proposes a tool that allows investigators to obtain a complete report and timeline of the activities that were performed on the device. This report incorporates the information provided by many sources into a unique set of data. Also, by means of an example, it is presented the operation of the solution, which shows the feasibility in the use of this tool and shows the way in which investigators have to apply the tool.**

Keywords—forensics; tool; register; activity; mobile; smartphone; Android

I. INTRODUCTION

Nowadays, mobile devices are used for a wide spread of tasks (e.g., entertainment, education, communication, socialization, research, commercial transactions). As a result of said use, the devices store information related to the user´s behavior. Therefore, they constitute an important source of evidence for forensics analysis[1].

Also, the forensics analysis uses a set of techniques that allow the collection and extraction of information from different devices without altering their original state [2]. For example, it can recover deleted files, browsing history, instant messaging information, login data, among others, all these types of information are known as digital evidence. According to Iorio et al., [3], there are three aspects that should be considered during the forensics analysis: i) avoid contamination of the evidence to prevent misinterpretations; ii) act methodically, that is, all the results of the forensics process must be well documented; and iii) control the chain of custody through the use of a protocol. Also, there are legal aspects to take into consideration when performing a forensics investigation, that do not comply always, these leads to the misuse of applications, fraud, theft, dissemination of copyrighted materials, etc. Thus, according to Taylor et al., [4] it is necessary to follow all the legal guidelines corresponding to the jurisdiction where the conflict is generated, to avoid undue exposure of personal information.

Also, there are a variety of applications (e.g., Encase, DFF, FTK, Helix, Oxygen, MOBILEdit, UFED), which are used for forensic analysis and allow the inspection of various elements of mobile devices (e.g., internal memory, applications, messages). Now, the so-called suites take all the previous points and join them in a single analysis creating a powerful and useful tool [5]. Also, it is important to take into account that there are advantages of using open source tools for forensics analysis during an investigation (e. g., no-cost, easy to examine in court, allows verification) [6]. But, commercial tools are also used because they provide a great variety of alternatives for analysis [6]. In Yadav et al., [7] it is presented a comparison among six commercial and open source applications. Those tools perform processes such as: recovering, performing keyword searches, recovering cookies, creating forensic images and locating partitions of the digital devices. Also, Shortall and Azhar [8] and Tajuddin and Manaf [9] present several popular forensic tools, such as Cellebrite UFED, MOBILedit Forensic, Forensic Toolkit, XRY, Oxygen Forensic Suite, EnCASE Forensic, and Paraben's device seizure. Each one of them has different capabilities, effectiveness and options to acquire information, but also, they offer similar services, analysis techniques and ways to present retrieved data. For example, UFED looks for physical data on the hard drive in order to recover deleted data, while the Oxygen Forensic Suite has a variety of options to perform a deep forensics analysis. By the analysis of the indicated studies, and as far as we know, there are not solutions that provide a complete log of the users' actions when using a mobile device, therefore the investigator needs to use more than one tool in order to recover all the data. Thus, this paper presents a tool, which has been implemented in Python [10], that generates a unique report with all the information about the mobile device user's behavior, by means of the collection of information from different applications that are installed on the it, which runs on Android OS. This information is then used to obtain a track of the users' activities while using the mobile device.

The present work is organized as follows: Section 2 presents the related work, then Section 3 introduces to the making of the solution and Section 4 presents the operation of the activity logger. Section 5 discusses the implementation of forensics tools, while Section 6 presents the application of a proof of concepts to analyze the digital information generated by a mobile device and describes the results obtained with the tool. Finally, the conclusions and future work are presented.

II. RELATED WORK

Recent studies on forensics analysis for mobile devices are mostly focused on Android and iOS operating systems [11], which also are only oriented to the study of specific applications. Anglano et al, [12] study the artifacts generated by WhatsApp when it is deployed on devices running Android, and explain how those artifacts are correlated to extract several types of data. The tools that they use are: FTK Imager, SqliteMan and SQLite v.3 databases [12]. On another study by the same authors, they analyze data obtained from Telegram; as a result,

978-1-7281-1704-1/19/$31.00

it presents the way to show the contact list, the chronology, the messages that have been exchanged, and the contents of the files that have been sent or received, all these with the use of the tools: SQLite database, UFED and Oxygen Forensic SQLite Viewer [11]. Moreover, Alyahya and Kausar [13] analyze Snapchat application on an Android platform by using two forensics analysis tools, Autopsy and AXIOM Examine. On the same context, Walnycky et al., [14] analyze 20 Android applications (e.g., WhatsApp, Viber, Instagram, Facebook Messenger, Tango), in which the digital evidence that could be used for forensics analysis, is examined, and also they evaluate the security involved in sending/receiving data and application privacy.

Furthermore, Walnycky et al., [14] developed an application named Datapp that captures packages that transit in a wireless network, and display the results on real time. Also, Rahaditya et al., [15] designed an application that helps in the search of text messages as evidence. This application makes copies and filters evidence of SMS; and generates an output file in PDF, Word or Excel format.

In summary, there are several studies focused on the extraction of evidence from individual applications. It leads to using forensics tools that focus on extracting specific information. Therefore, the forensics tools are wasting their extraction potential as a whole. The solution to this problem is the implementation of a tool that collects all the evidence into a single final report.

III. Making of the Solution

Before handling each file, it is necessary to know its structure, how they are distributed, which kind of information it contains, the path where they are stored, among other aspects. According to UNE 71505 [16], it is important to be careful in not altering the files that can be considered as evidence.

A. File structure

It is necessary to choose the appropriate software package (forensic tool) suitable for the extraction of digital evidence in a mobile device with Android OS. The tools that are going to be used provide reports in several formats. They contain information of the data that has been extracted from the device. Moreover, applications such as Andriller [17], MOBILedit [18] and Oxygen Forensic [19] could be installed and obtain a good performance on Windows, while Kali Linux is preferred on Linux platforms. The former tools for Windows generate .xls, .pdf, and htlm documents, among other formats, while Kali provides results in .txt format.

Each generated report has information about the mobile device (i.e., investigator data, characteristics of the device, date on which the extraction is carried out, scheduled events on calendar, passwords of wi-fi, web browsing history, multimedia, applications either installed or deleted, application use, download history, cookies, storage, call list, SMS messages, deleted files, data on SD external memory). The report includes the following information about each of the extracted data: name, label, URL, directory, folder that contains it, last visit, date that was accessed, date on which it was modified, size, created date, type of file, description. On the other hand, in the case of calls, messages, calendar and passwords the information presented on the report are: name, date on which the activity was made, duration of the activity and a brief description.

To summarize, each forensic software, despite its characteristics or brand, returns a report with the details mentioned above.

B. Programming language

The programming language used to build our tool is Python, used over Linux. The forensics software packages for Linux operate by commands on a console, even so this does not diminish their efficiency, let alone their potential.

Besides, in the forensics area of mobile devices, there is a great inequality between commercial and open source software, due to the fact that most of the solutions are commercial. According to Limodio et al. [6], the open source forensics tools allow experts, judges or litigants to verify that the evidence has not been manipulated, due to its easiness of access. Finally, it must be considered that the software is only a tool and the success of the process depends on professionals that operate it.

IV. Operation of Activity Logger

The entire process for using the tool is illustrated in Fig. 1, it is divided into four main activities: i) identification, ii) collection, iii) analysis, and iv) preservation. In each of these activities, there are tasks and artifacts involved. It is important that the solution presented is based on the general methodology of the treatment of digital evidence [20].

The developed tool gathers the data to be included on the report which will be generated by the various forensics software; all reports must be stored in a determined folder, so that the application can later query them.

Also, the tool accomplishes the following objectives:

- Detects and counts the number of files according to the type.
- Determines the number of sheets, columns and rows in a Microsoft Excel file and the number of lines in text files. This activity is performed to indicate the length of each file.
- Gets the column that contains the date and time of the users' activity.
- Compares the date entered by the forensics investigator with the date of the evidence.
- Saves the filtered data.
- Merges the data in a single file.
- Organizes the data in a descending order, so it is chronologically order.
- Deletes repeated data.
- Assigns a code to each activity.
- Saves the report.

V. Implementation

This section describes the implementation of the activities shown in Fig. 1. Once the program is executed, as shown on Fig. 2, the first action is to enter and validate the data required for the investigation (date and time). These items require a certain format, otherwise a validation message will be shown.

A. Identification

In the identification stage, the investigator works with different types of files, which should be stored in a folder called *Report*. The first step is to determine the number of files that the folder contains according to their type; this because each type should be processed in a different way. For example, if there are Microsoft Excel files obtained from the forensics software package running in Windows, the number of files is identified; also, within each file the number of pages, rows and columns are stored. On the other hand, in a Linux operating system, the number of files is also identified, with the difference that within each file the number of lines is also obtained.

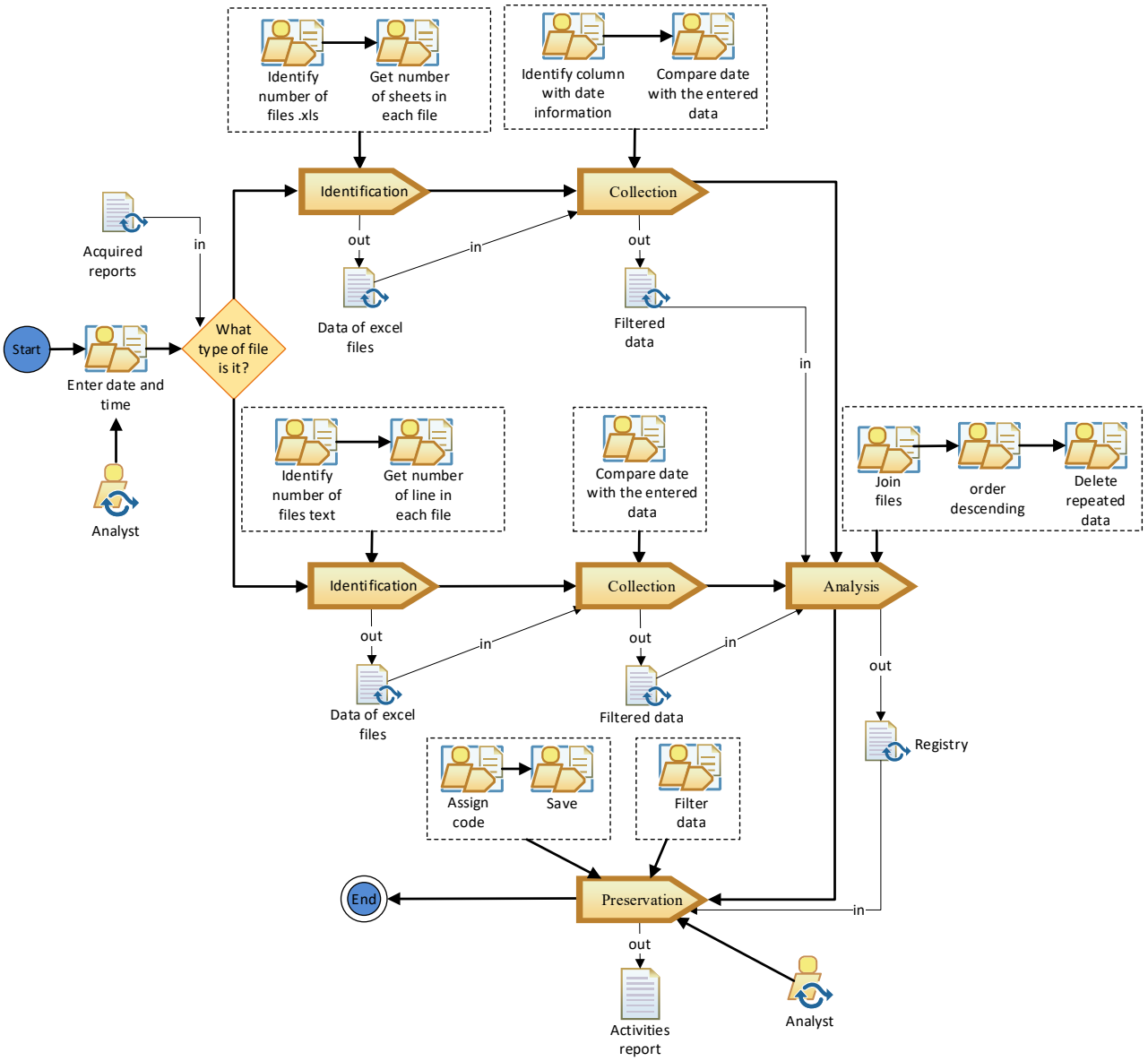

Fig. 1. Tasks and processes of the proposed tool.

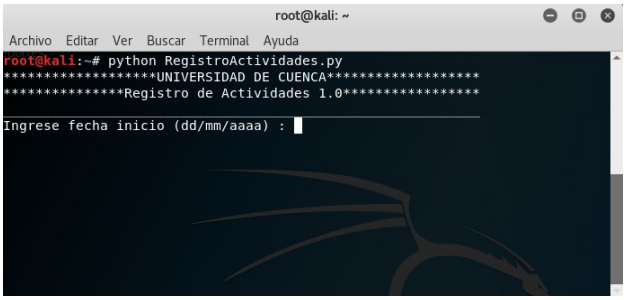

Fig. 2. Activity registration tool.

B. *Collection*

Once the data are identified, the evidence contained in the report is collected. In the case of Microsoft Excel files, each page contains information about the device and also, there is a column that indicates the date and hour in which the activity was executed. Once the date and hour have been identified, the program starts comparing the dates, when finding a match, the whole row is saved in a text file. Whereas, in the case of text files, the date entered into each row is compared, and later stored in a text file. Headings are included in order to guide the reader through the report. There are two types: component heads and text heads.

C. *Analysis*

In this stage, the reports are merged into a single file; here, no writing tasks should be performed because this would alter the evidence. An algorithm is used to order chronologically the data using the date and time previously obtained, Finally, the data with the same date and label are deleted in order to eliminate duplicated information.

D. *Preservation*

Finally, once all the clean records have been obtained, an "axxx" code is assigned to each activity and it is saved in a text

file called *Final Report*. Moreover, there is an option to filter the information that the investigator considers relevant for the case.

Additionally, the program prints a summary with the number of each file, the amount of activities collected in the established slot of time, the date and time and the filtered report. Finally, in the case that the entered code is incorrect or it does not exist, it is printed in the summary.

It should be noted that each file was checked by a function that returns a single hash for each file. When a file is modified, it returns a completely different hash, which helps verify the integrity of the data. This must be done in the preservation part of the evidence proposed by the general methodology of the treatment of digital evidence [21].

VI. Results

For the proof of concepts, a process aligned with international standards is followed (i.e., ISO / IEC 27037 (2012), RFC 3227 [22] and UNE 71505 [16]) focused on mobile devices. It also considers guides such as: Forensics Guide in Cell Phones of the National Institute of Standards and Technology (NIST) [23] and the Best Practice Guide for Forensic Analysis in Mobile Phones of the Scientific Working Group on Digital Evidence (SWGDE) [24]. Hence, a guide to develop a forensics investigation on mobile devices is presented, where the main activities are: i) identification and preservation, ii) acquisition, iii) analysis, and iv) documentation.

On the following section, the most important results from the proof of concepts are presented. In addition, the most relevant activities of the case are analyzed for the presentation of the expert report.

A. Identification and preservation of evidence

1) Proof of concepts: The proposed scenario for this case is a common one for a University, where one of the main concerns is the unauthorized use of electronic mobile devices during an academic assessment. This rules are followed to prevent academic dishonesty, which are punishable by law according to the internal protocols of every academic institute and is mandated by the Organic Law of Intercultural Education in Ecuador [25].

The context in which this example has been carried out, is the use of a mobile device by a student during an assesment. The student is taking a test inside the University, which was applied at 11h00 of July 9th, 2018.

2) Identification: The first phase of the process is very important because is the starting point for the investigation. In this phase, the strategy to be followed during the case is determined, as well as the personnel that will intervene and the role that each person will take. The identification allows an ordered process and enables involved people to act timely and efficiently. Also, another informatio is obtained at this time: the devices that will be analyzed, the characteristics of each one and the materials that will intervene during the investigation.

B. Acquisition of evidence

In this phase, the greatest amount of information coming from the mobile device is obtained. Forensics tools are used for extraction in both Windows and Linux, and the obtained reports are shown in Fig 3. Finally, a total of 26 files were obtained.

1) Forensics extraction: The tools that will be useful to perform this step are chosen. As mentioned previously, there are several tools, but only the most relevant ones for this case are chosen, such as: Andriller [17], Kali Linux [26], Oxigen forensic [19], and MOBILedit. In addition, more than one application has been used in order to obtain data from various sources and correlate the events

2) Acquisition of registration activities: The developed tool allows the collection of information for each report obtained by entering the date (i.e., July 9, 2018 between 11h00 to 13h00). Fig. 3 and Fig. 4 shows the files, and the 283 activities that were obtained.

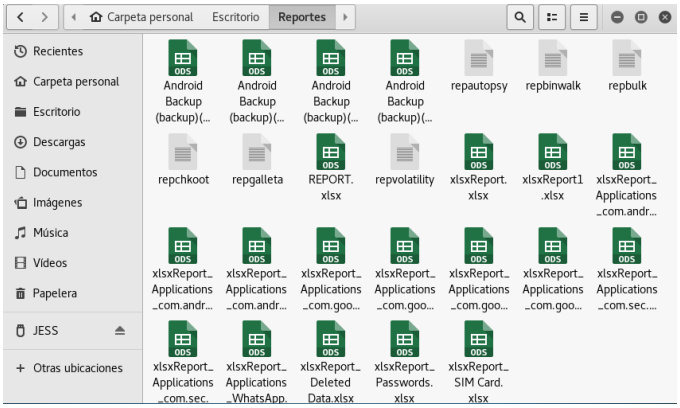

Fig. 3. Reports obtained from forensic tools.

With the activity report, the next step is to obtain the most relevant activities. The report is studied and the data related to the subject of the exam is obtained.

C. Analysis of the evidence

In this section, the most relevant evidence of the proof of concepts is presented; that is, the activities that the owner of the device performed during the exam. It has been divided into five different types of evidence: web browsing history, cookies, images, downloads and instant messaging applications. Finally, the process followed by the system are also described, in order to establish that there was telephone activity. As for browsing history, 11 web pages were found which, by their label (title of the web page), are related to the subject of the course. As for cookies, there are 8 activities that left a trace on the pages that the user tried to access. On the other hand, the 4 images that were recovered, by the use of the address that the registry provided, gave information pertinent to the case. Finally, an attempt was made to download a PDF file during the exam, and an audio was sent using the WhatsApp messaging application. In total there were 25 unauthorized activities that the researcher determined to be important evidence of academic dishonesty during the exam.

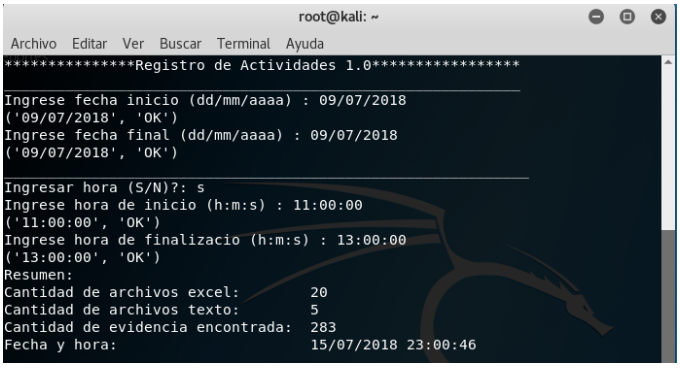

Fig. 4. Results obtained from registration activities.

When performing the general report manually, the researcher or person in charge would have to unify each individual report in a single file, bearing in mind that there are two formats of files. The final result will be a messy file because, as explained in section III, each report provided by the forensic tools has a different structure. The final file will have repeated information and will not be ordered chronologically, this proves that the search for a certain activity is tedious. While, the proposed tool simplifies all this work. The program unifies each file in a single general report facilitating the work of ordering chronologically, in this way an event can be identified in a certain period of time.

V. Conclusions and Further Work

Based on several tests performed with different brands of Android mobile devices; it can be concluded that the activity registration tool is stable and complies with the requested examinations.

The tool automates and reduces the time of evidence analysis. Selecting the right tools for the acquisition of evidence that serves as input to the application represents a crucial piece of research; however, none of them possess the ability to acquire all the information of a mobile device. Therefore, it is necessary to use several of them to improve the desired result. Finally, the advantage of using Python programming language, is that it allows to verify the source code and thus, validate that it does not alter the digital evidence.

The main advantage found while using this tool is that it reduces the time used on an investigation and saves resources. This because each installed software returns large volumes of information that must be analyzed step by step by the researcher in charge. Thus, this tool avoids the manual use of more than one software to get all the information that is required for the case.

The evidence has to be carefully manipulated, because if the information is altered in any way, this will not be valid for the investigation.

Finally, the presented study, gives a first view on the handling of digital evidence in mobile devices with Android OS, this later can be developed for other operating systems such as iOS and Windows Phone. For further work, it is necessary to increase the interoperability to gather the information from third party solutions and propose connectors and generic ways to extract evidence. Also, it is important to measure and perform future improvements in certain non-functional characteristics of this tool (e.g., efficiency, latency, usability).

References

[1] H. K. S. Tse, K. P. Chow, and M. Y. K. Kwan, "The next generation for the forensic extraction of electronic evidence from mobile telephones," *Int. Work. Syst. Approaches Digit. Forensics Eng., SADFE*, 2014.

[2] K. Barmpatsalou, D. Damopoulos, G. Kambourakis, and V. Katos, "A critical review of 7 years of Mobile Device Forensics," *Digit. Investig.*, vol. 10, no. 4, pp. 323–349, 2013.

[3] A. Di Iorio, R. Sansevero, and M. Castellote, "La recuperación de la información y la informática forense: Una propuesta de proceso unificado," no. March, 2013.

[4] M. Taylor, G. Hughes, J. Haggerty, D. Gresty, and P. Almond, "Digital evidence from mobile telephone applications," *Comput. Law Secur. Rev.*, vol. 28, no. 3, pp. 335–339, 2012.

[5] B. B. Carrier, "Open Source Digital Forensics Tools : The Legal Argument.," *@Stake*, no. October, p. 11, 2002.

[6] G. F. Limodio and P. A. Palazzi, "El uso de software abierto para el análisis de la evidencia digital," 2016.

[7] S. Yadav, K. Ahmad, and J. Shekhar, "Analysis of Digital Forensic Tools and Investigation Process," *High Perform. Archit. Grid …*, pp. 435–441, 2011.

[8] A. Shortall and M. A. H. Bin Azhar, "Forensic Acquisitions of WhatsApp Data on Popular Mobile Platforms," *Proc. - 2015 6th Int. Conf. Emerg. Secur. Technol. EST 2015*, pp. 13–17, 2016.

[9] T. B. Tajuddin and A. A. Manaf, "Forensic investigation and analysis on digital evidence discovery through physical acquisition on smartphone," *2015 World Congr. Internet Secur. WorldCIS 2015*, pp. 132–138, 2015.

[10] "Welcome to Python.org." [Online]. Available: https://www.python.org/. [Accessed: 21-Aug-2018].

[11] C. Anglano, M. Canonico, and M. Guazzone, "Forensic analysis of Telegram Messenger on Android smartphones," *Digit. Investig.*, vol. 23, pp. 31–49, 2017.

[12] C. Anglano, "Forensic analysis of whats app messenger on Android smartphones," *Digit. Investig.*, vol. 11, no. 3, pp. 201–213, 2014.

[13] T. Alyahya and F. Kausar, "Snapchat Analysis to Discover Digital Forensic Artifacts on Android Smartphone," *Procedia Comput. Sci.*, vol. 109, pp. 1035–1040, 2017.

[14] D. Walnycky, I. Baggili, A. Marrington, J. Moore, and F. Breitinger, "Network and device forensic analysis of Android social-messaging applications," *Digit. Investig.*, vol. 14, no. S1, pp. S77–S84, 2015.

[15] I. P. Agus, "Prototyping SMS Forensic Tool Application Based On Digital Forensic Research Workshop 2001 (DFRWS) Investigation Model," 2016.

[16] "Norma UNE 71505-1:2013." [Online]. Available: https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma/?c=N0051411. [Accessed: 21-Aug-2018].

[17] "Andriller | Android Forensic Tools." [Online]. Available: https://www.andriller.com/. [Accessed: 21-Aug-2018].

[18] "MOBILedit." [Online]. Available: https://www.mobiledit.com/. [Accessed: 21-Aug-2018].

[19] "Oxygen Forensics - Mobile forensics solutions: software and hardware." [Online]. Available: https://www.oxygen-forensic.com/en/. [Accessed: 21-Aug-2018].

[20] ISO/IEC, "Information technology — Security techniques — Guidelines for identification, collection, acquisition and preservation of digital evidence." 202AD.

[21] "ISO/IEC 27037:2012 - Information technology -- Security techniques -- Guidelines for identification, collection, acquisition and preservation of digital evidence." [Online]. Available: https://www.iso.org/standard/44381.html. [Accessed: 30-Aug-2018].

[22] T. Killalea and D. Brezinski, "Guidelines for Evidence Collection and Archiving."

[23] "National Institute of Standards and Technology | NIST." [Online]. Available: https://www.nist.gov/. [Accessed: 30-Aug-2018].

[24] "SWGDE." [Online]. Available: https://www.swgde.org/. [Accessed: 30-Aug-2018].

[25] Gobierno del Ecuador, "Ley Orgánica de Educación Intercultural." 2012.

[26] "Kali Linux | Penetration Testing and Ethical Hacking Linux Distribution." [Online]. Available: https://www.kali.org/. [Accessed: 21-Aug-2018].