A UML-Based Approach for Analysing Potential Digital Forensic Evidence

Victor R. Kebande¹, Nickson M. Karie²
¹Information and Computer Security Architecture (ICSA), Research Lab
¹Department of Computer Science,
¹University of Pretoria, South Africa.
²Department of Computer Science,
²University of Swaziland, Kingdom of eSwatini
vickkebande@gmail.com¹, nickson.karie@gmail.com²

ABSTRACT

This paper introduces the reader to a Unified Modeling Language (UML)-based approach that can be used to model a digital forensic analysis of Potential Digital Evidence (PDE). Additionally, this study has been presented by highlighting a model-based approach of the UML and the methodology that can enable one to visualise the activities in the digital forensic analysis process which is an important process of the digital investigation process. Eventually, the author provides a discussion on the important operations and how the analysis process is executed in a UML-driven approach.

KEYWORDS

1 INTRODUCTION

In the recent years, Digital Forensics (DF) has seen numerous innovative approaches that have been proposed by different researchers in order to acquire and analyse Potential Digital Evidence (PDE) from myriad sources. This is evident from research done by Xiaoyu, Nhien-An and Mark, [1] on evaluation of digital forensic process models with respect to digital forensics as a service. Most of the proposed approaches by most of the previous researchers though, are based on a compilation of the generally accepted practices from the digital forensic domain. The main goal of any previously proposed digital forensic process model, however, is to help digital forensic investigators and Law Enforcement Agencies (LEA) to acquire original digital evidence in such a manner that helps in the protection and preservation of the acquired digital evidence. However, with the changing technological environment, better and efficient ways of dealing with PDE are always sought after. This paper therefore, presents a UML-based approach that can be used to model a digital forensic analysis of Digital Forensic Evidence (DFE) with the primary aim of facilitating the digital forensic analysis process.

This research has been motivated by the fact that, visualisation is increasingly becoming important for understanding information such as investigative digital data that is collected from a digital crime scene as potential evidence. It is also useful for the reconstruction of a crime scene or item, which can further assist the general public (who may have little or no understanding of the subject matter) to perceive and understand what is being analysed [2]. Besides, visualisation can enable digital forensic investigators as well as LEA to rapidly and efficiently locate vital information that is or may be of interest in an investigative perspective. It can also guide investigators towards the next best step in their search for digital events so that potential digital evidence recovery is carried out in a more efficient and effective manner [3]. Visualisation models can also provide an intuitive as well as systematic approach to analyse potential digital evidence. This can further help in evaluating the admissibility of potential evidence in any court of law during criminal, civil or legal proceedings.

As mentioned by Karie & Venter [4], existing approaches in digital forensic have provided
guidelines for identifying and preserving PDE that is captured from a digital crime scene. However, the extent to which such PDE may be admitted as evidence in a court of law remains a challenge to investigators [4]. For this reason, new methodologies and approaches need to be developed in the field of digital forensics with the ability to effectively enhance the digital forensic analysis process.

The remainder of this paper is structured as follows. Section 2, presents a brief background of the existing literature presented as follows: UML Modelling Approaches, Digital Forensics Process, Potential Digital Evidence and lastly Digital Evidence Analysis. Section 3 presents the UML-based approach to digital evidence analysis followed by related work in Section 4. A critical evaluation of the proposed concept is given in Section 5. Finally, conclusions drawn from this work and future work are highlighted in Section 6.

2 BACKGROUND

This section provides background on the following areas: UML Modeling Approaches, Digital Forensics Process, Potential Digital Evidence and lastly Digital Evidence Analysis.

2.1 UML Modeling Approaches

Over the years, the Unified Modeling Language (UML) has been described as a general-purpose and developmental modeling language in the field of software engineering [5]. It is for this reason that, UML has been used by software engineers to provide a standard way to visualise the design of various systems. This has, therefore, made UML become increasingly important as a modeling language across various domains [6]. This is, further, backed up by the fact that, UML is currently being used as the universal technique for modeling object-oriented applications across a myriad of domains [7]. However, this also makes domain specific adapting the UML meta-model increasingly important in the context of model checking and code generation mechanisms across different domains. As a result of this, the authors in this paper are applying UML into the domain of digital forensics primarily to help in modeling the digital forensic analysis of PDE. Besides, UML is also extendable as it offers a profile mechanism for customization. The next sub-section briefly explains the digital forensic process.

2.2 Digital Forensic Process

According to Carrier [8], a digital forensic investigation process is one special case of an investigation, where the procedures and methods used will allow the outcome to stand up in any court of law. However, Karie and Venter, [4] states that to convince the court that the digital evidence presented is worthy of inclusion into the criminal process, the methods and procedures used during investigation must possess scientific validity grounded in scientific methods and procedures. This implies that the investigation processes should be compatible with the relevant policies and/or laws in various jurisdictions, since evidence may not be admissible in court if it was not properly or legally acquired [4].

In a bid to help collect potential digital evidence in a forensically sound manner, numerous models, frameworks and methodologies have been proposed. These models are able to help in the gathering or specifying different phases in the digital forensic investigation process [9]. However, according to Kohn et al., [10], these vast numbers of proposed models and frameworks have added to the complexity of the field of digital forensics. For this reason, the authors in this paper propose a UML-based approach for analysing PDE with the primary aim of enabling the visualisation of the digital forensic analysis process. The next sub section explains the concepts of potential digital evidence.

2.3 Potential Digital Evidence

According to Kozushko [11] Digital Evidence (DE) comprises of digital data that is able to establish whether a digital crime has occurred or not. DE helps to establish the existence or absence of a link between a suspect and the crime or the victim of a particular crime. In addition, DE can help forensic analysts and the Law Enforcement Agencies (LEAs) in coming
up with a conclusion during a digital forensic investigation process.

However, Carrier and Spafford [12] present Digital Evidence (DE) as an object. Being an object DE has unique features and characteristics based on their functionality and their creator. Besides, the notion that digital data has a physical form implies that physical evidence may contain digital evidence. Moreover, DE is an aspect that cannot be overlooked because it may not be obvious when a computer-based crime occurs. This is backed up by Karie and Venter [13] who stated that the sources of Potential Digital Evidence (PDE) have also grown exponentially with the advances in digital technology. This has made it even hard for investigators and LEAs to establish reliable PDE sources for different types of digital evidence captured during the investigation process.

While digital evidence remains to be a very important aspect in digital forensics, the way it is acquired is of much importance. In that respect Karie, Kebande and Venter [14] have highlighted that the knowledge, skills, tools and techniques used to collect, organise and analyse digital evidence has become indispensable. This is important because such techniques are important during presentation and analysis of digital evidence in a court of law.

According to Carrier and Spafford [12], the focus in digital investigations is mostly on how DE can be recovered as well as examining the properties that the collected evidence has. However, this current paper presents a UML-based approach for analysing PDE with the primary aim of helping in modeling the digital forensic analysis process. The next sub section explains the digital evidence analysis.

2.4 Digital Evidence Analysis (DEA)

Analysis is the process of interpreting any extracted digital evidence data in order to determine the significance they hold to any case at hand. According to [15], there are three main tasks that investigators perform during the analysis of evidentiary digital media:

a. Creation of a forensic image;

b. Creation of a forensic archive from the forensic image; and

c. Exporting potential probative digital data related to the investigation.

All the above-mentioned tasks are critical to the overall success of the investigation process and eventual prosecution of any case. From these concepts therefore, DEA is used in this paper to portray the process of interpreting any recovered digital evidence data and putting it in a logical and useful format that can be used to support or refute a hypothesis. Knowing that, digital evidence, by its very nature, is fragile and can be altered, damaged, or destroyed by improper handling; examination and analysis is, thus, best conducted on a copy of the original evidence. This also implies that, the original evidence must be acquired in a manner that protects and preserves the integrity of the evidence [16].

The authors in this paper, however, argue that the process can be enhanced using visualisation by the use of the UML and object diagrams. That is the main motivation behind this research paper; to propose a UML-based approach for analysing PDE with the aim of enabling the visualisation of the digital forensic analysis process. The UML-based approach to digital evidence analysis is explained in the section to follow.

3 A UML-BASED APPROACH TO POTENTIAL DIGITALE EVIDENCE ANALYSIS

This section presents a UML-based approach as a contribution that shows how potential digital Evidence (PDE) can be analysed. This approach has been designed in such a manner that digital evidence that has been acquired can be analysed based on their characteristics. This may eventually help one to identify Digital Forensic Evidence (DFE) sources and ultimately increase the chances of attribution during a digital forensic investigation.

The UML-based analysis approach has primarily been designed with a view that shows that irrespective of the source of digital evidence, one can easily provide facts by extracting real digital
evidence from the accumulated digital evidence from a potential digital crime scene.

The approach in discussion has been presented using a UML activity diagram and an object diagram as is shown in Figures 1 and 2 respectively. Figure 1 (activity diagram) has been divided into five process and activities namely: (1) pre-incident analysis, (2) evidence analysis design, (3) evidence analysis implementation, (4) evidence examination and (5) evidence reporting. Nevertheless, Figure 2 (object diagram) has been divided into two processes namely digital forensic investigator and digital forensic crime scene. More discussions on the aforementioned divisions will become apparent in the later sections of this paper; however, the next section presents a discussion on the UML based approach.

3.1 Construction of UML-Based Approach

In this section, the reader is introduced to the construction of the UML-Based approach to the potential digital evidence analysis. The following has been discussed in this section: Pre-incident Analysis, Evidence Analysis and Design, Analysis Implementation, Evidence Examination and lastly, Reporting.

3.1.1 Pre-Incident Analysis

This approach is faceted via Potential Digital Evidence (PDE) that is collected from different sources. The sources are realised by studying the potential attack scenarios that are capable of yielding digital evidence that can be used to develop a hypothesis to prove or disprove a fact in a court of law. These sources in the context of the attack scenarios include targets that an adversary focuses on and digital evidence that is able to be extracted from such sources.

During pre-incident analysis, there is need to check where evidence is being extracted from and the media is defined as a potential source (physical or logical) where an object can originate from. This allows one to have an idea of what the analysis is based on as highlighted by Carrier [17]. Carrier has also highlighted this as the use of the notion of abstraction layers. Additionally, another important factor is the tool requirement which allows the evaluation of the tools that have to be used during the analysis process. As a result of increased technology and instrumentation, NIST in its project on Computer Forensic Tool Testing (CGTT) allows evaluation of these tools; also the scientific Working Group on Digital Evidence (SWGDE) allows assessing of these tools [18].

Another important aspect that has been employed in this context is the methodologies employed. It is important to note that the methodologies that are supposed to be employed should be based on scientific discoveries and proven methods. Lastly, verification is needed to ascertain whether the pre-incident process followed the prescribed methodology before analysis design is initiated. A summary of these aspects that have been highlighted in Figure 1 have also been shown in Table 1.

<table>
<thead>
<tr>
<th>Table 1: Pre-Incident Analysis Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-Incident Analysis</strong></td>
</tr>
<tr>
<td>1 Define media, file system, network and memory to be analysed</td>
</tr>
<tr>
<td>2 Define the tool requirements to be used</td>
</tr>
<tr>
<td>3 Provide the methodology to be used</td>
</tr>
<tr>
<td>4 Verify pre-</td>
</tr>
</tbody>
</table>
If the analysis process follows the prescribed methodology.

The process of sifting accumulated evidence in order to extract facts is initiated.

## 3.1.2 Evidence Analysis Design and Implementation

The capacity to design the analysis process and to implement it ensures that one is able to outline specific functions that allow a representation on how digital evidence can be analysed. It is worth noting that the aim of this is to be able to create a hypothesis that can be used to prove/disprove a fact in a court of law during criminal or civil proceedings. Moreover, it is also vital to note that analysis in the context of this paper has been presented as a significant phase of the digital forensic investigation model.

The successful implementation of analysis process ensures that activities are able to be assessed in detail while performing examination. The reason of doing this is to be able to bring out the importance of PDE that is collected from a digital crime scene. This ultimately enables the development of an action plan that can help digital forensic investigators and the law enforcement agencies to attain the required readiness. A summary has also been given in Table 2 that is shown next.

<table>
<thead>
<tr>
<th>Incident Analysis Process</th>
<th>if the analysis process follows the prescribed methodology.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialize digital forensic analysis process</td>
<td>The process of sifting accumulated evidence in order to extract facts is initiated.</td>
</tr>
</tbody>
</table>

### Table 2: Evidence Analysis Design and Implementation Activities

<table>
<thead>
<tr>
<th>Evidence Analysis Design and Implementation</th>
<th>Analysis and Tool Requirements</th>
<th>While designating analysis process it is important to incorporate the requirements of the digital forensic tools so that examination can be done with ease.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Analysis and Tool Requirements</td>
<td>Initiate Analysis Design</td>
<td>The design of potential digital evidence analysis strategies is initiated.</td>
</tr>
<tr>
<td>2 Implement Evidence Analysis process</td>
<td>Provide updated Analysis implementation process</td>
<td>The process is implemented by assessing the user-based activities in order to extract digital events (evidence).</td>
</tr>
<tr>
<td>2 Provide updated Analysis implementation process</td>
<td>Updated events are highlighted as analysis implementation is carried on.</td>
<td></td>
</tr>
</tbody>
</table>
3.1.3 Evidence Examination

During examination, it is important to apply the digital forensic principles that have been mentioned previously in the digital forensic investigation processes. By using accepted methodologies (scientifically proven techniques), the extracted evidence that is treated as real evidence is able to be examined. As shown in Table 1, these pieces of evidence are normally recovered from a given media, i.e. network, file system, memory etc. In the context of this research paper, analysis has been used to portray the interpretation done from the recovered digital data that is arranged in a logical manner.

3.1.4 Reporting

Reporting provides the details of the findings that emanate from the analysis process. Details regarding the examination results are submitted as potential digital evidence that is analysed by presenting examination notes.

3.2 Object Flow of Digital Evidence Analysis Diagram

This section provides an interaction of the diagrams that provides the flow of activity between the digital forensic investigator and the digital forensic crime scene. The main objective of using an object diagram is to be able to show how analysis of potential digital evidence is achieved in a step by step approach. Firstly, a DFI requests to access the crime scene which is then seized and the details of seizure are then highlighted. Acquisition follows, which in this context allows evidence analysis process to begin. Lastly, the desired evidence is extracted which is then examined for possible incidents. This is shown in Figure 2.
Request to access crime scene

Provide seized crime scene
crime scene details
Access
Process evidence acquisition
Start evidence analysis
Provide evidence
Extract desired evidence
Examine and analyse the evidence
Provide access
Digital Forensic Investigator
Digital Forensic crime scene

Figure 2. Object flow diagram for digital evidence analysis process

In the next section, work that has been used as related work is given.

4 RELATED WORK

The authors presents works that has somewhat been used as related work in this paper. To begin with, there has been a number of researches that has applied the UML concepts on forensic processes, but not much concentration has focused on potential digital evidence. Kohn, Eloff and Olivier [19] have used a UML paradigm to model digital forensic processes where they have applied activity diagrams and use-cases from the behavioural of UML specifications. It is important to note that this work was entirely focused on the Digital Forensic Process Models (DFPMs), however, work presented in this research paper is focused on digital evidence. Additionally, formalism and structured approaches methodologies for computer forensics by Ruan and Huebner [20] has shown how to use UML diagrams to model and visualize different aspects of computer forensics. Also, the Advanced Data Acquisition Model (Adam), is a generic process model for digital forensic practice that provides a step towards the development of an accepted standard of the activity of digital forensic evidence [21]. Other relevant work include a Cloud Forensic Readiness as a Service (CFRaaS) by Ke bande and Venter [22] that has focused on modelling forensic readiness processes using UML [22].

5 CRITICAL EVALUATION OF THE PROPOSITION

Based on the contributions that have been highlighted in this research paper, the authors have been able to identify a research gap that has allowed the presentation of a systematic UML-based approach that caters for the design of digital evidence analysis process. Furthermore, the authors have introduced different UML interaction processes with sub activities. The processes that have been introduced includes the media being accessed, design analysis, implementation of analysis, examination and reporting. It is worth noting that the comprehensibility that the UML-based approach presented helps one to be able to relate to evidence analysis from a UML point of view.

The authors have demonstrated that realisation of a UML-based approach in the analysis of potential digital evidence is of vital importance. Moreover, through the UML activity and object diagram it has become apparent that the activities employed in this process can be used by digital forensic experts. These activities are defined and executed holistically in a very dynamic approach and the digital forensic experts are able to successfully examine potential evidence in a more realistic approach.

It is imperative for one to realise that future digital forensic investigation techniques will need to have a proper design and description of the case that needs to be investigated and this research paper tries to provide this relationship. Ultimately, this helps investigators to be able to understand the tool requirements that are needed during analysis process. Additionally, the authors have explored different analysis techniques that are based on the UML standpoint that stipulates the systematic process of evidence analysis.
6 CONCLUSION AND FUTURE WORK

The research work that has been presented in this paper was aimed at checking if a UML-based approach can be used during analysis of digital evidence. The authors have proposed a UML-based approach as a contribution that allows digital forensic investigator to be able to relate with potential evidence based on the activities defined in the UML and the object diagram.

The authors defined a UML-based approach with the following processes: Pre-incident analysis, evidence analysis design, analysis implementation, evidence examination and reporting. Nevertheless, the authors also defined an object diagram with the following activities: Digital forensic investigator abilities and a digital forensic crime scene.

In future, the authors will design a scenario that is able to allow the incorporation of tool requirement in order to create prototype as a proof of concept of the propositions that have been mentioned in this research paper.

REFERENCES


