Adding Event Reconstruction to a Cloud Forensic Readiness Model

Victor R. Kebande*
Department of Computer Science,
University of Pretoria,
Private Bag X20, Hatfield 0028,
Pretoria, South Africa.
Email: vickkebande@gmail.com*

H.S.Venter†
Department of Computer Science,
University of Pretoria,
Private Bag X20, Hatfield 0028,
Pretoria, South Africa.
Email: hventer@cs.up.ac.za†

***Abstract*— During post-event response, proactive forensics is of critical importance in any organisation when conducting digital forensic investigations in cloud environments. However, there exist no reliable event reconstruction processes in the cloud that can help in analysis and examination of Digital Evidence (DE) aspects, during Digital Forensic Readiness (DFR) process, as defined in the standard of ISO/IEC 27043:2015. The problem that this paper addresses is the lack of an easy way of performing digital event reconstruction process when the cloud is forensically ready in preparation of a Digital Forensic Investigation (DFI). During DFR approaches, event reconstruction helps in examination and pre-analysis of the characteristics of potential security incidents. As a result, the authors have proposed an Enhanced Cloud Forensic Readiness (ECFR) process model with event reconstruction process that can support future investigative technologies with a degree of certainty. We also propose an algorithm that shows the methodology that is used to reconstruct events in the ECFR. The main focus of this work is to examine the addition of event reconstruction to the initially proposed Cloud Forensic Readiness (CFR) model, by providing a more enhanced and detailed cloud forensic readiness model.**

Keywords - Cloud; Forensic; Readiness; Model; Event; Reconstruction; Digital

I. INTRODUCTION

Traditionally, the process of digital forensic investigation begins when a digital device has been confiscated. However, whenever a hypothesis about a particular committed digital crime exists, the process of digital forensic investigation is launched to try to prove the existence of admissible digital evidence that can support or refute the occurrence of a digital event in a court of law.

Cloud forensics is best conceptualised as a field of study that is concerned with digital investigations in cloud environments. Reconstructing digital events in a forensically ready environment involves revisiting the characteristics and sequence of digital events in a proactive process and checking whether the collected Potential Digital Evidence (PDE) satisfies admissibility. In spite of that, event reconstruction tends to analyse and examine why evidence portrays certain characteristics [1].

The main contribution of this research is to investigate and propose the possible applicability and acceptability of event reconstruction process in a Cloud Forensic Readiness (CFR) model so that it can enable reconstruction of events and be able to support future investigative technologies.

The paper's main focus is to provide an extension of the initially proposed CFR model that was proposed by Kebande and Venter [2] through the addition of event reconstruction process. The proposed Enhanced Cloud Forensic Readiness (ECFR) is a detailed model that shows systematic occurrence of events and how they are supported by evidence in the cloud environment before Digital Forensic Investigation (DFI) is conducted.

The contribution of this paper is presented in three phases; first we present a scenario which is then followed by a high–level model, thereafter we present the detailed ECFR model. Additionally, an explanation of how event reconstruction is being added to the CFR model is provided too.

The rest of the sections in the paper are structured as follows: The paper begins by describing the background of event reconstruction, Digital Forensic Readiness (DFR), ISO/IEC 27043 readiness process groups is dealt with in section II. Section III presents previous and related work. Thereafter, Section IV discusses the ECFR model. Next, experiments and discussions of the study are presented in Section V. The paper concludes with Section VI stating a conclusion and suggesting future work.

II. BACKGROUND

In this section, the authors present a brief background on event reconstruction, DFR, ISO/IEC 27043 readiness process groups and previous and related work. The goal of discussing event reconstruction is to show its role of characterising digital evidence within a forensically ready cloud, DFR is discussed to show the proactive side of DF. ISO/IEC 27043 which is a standard for security techniques, incident investigation principle and process is discussed to open the forensic readiness spectrum further. Previous and related work is discussed to show initially proposed models and relevant works from other researchers. Finally, cloud computing is incorporated to show how DFR approaches can be adapted in the cloud.

978-1-4799-7755-0/15/$31.00 ©2015 IEEE

A. *Event Reconstruction*

The first step of performing an event reconstruction procedure according to Gardner and Bevel [3] is evidence collection and examination which helps to seek the truth. Afterwards, there is creation and sequencing of event segments from the evidence. However, according to Brian and Carrier [1], a digital event is an occurrence that is meant to change the state of one or more digital objects. Moreover, an object is represented as evidence of an event if the event has changed the object's state. This is basically a discrete collection of digital data which may be examined for possible potential evidence with regard to the occurrence of a security incident. Consequently, event reconstruction tends to question why digital evidence has certain properties and characteristics. Additionally, the existing analogies show that during event reconstruction, evidence analysis and examination is conducted to show the exact causes of the characteristics and properties that digital evidence poses. On the same note, Carrier and Spafford [4] argue that before and after incidents are detected a hypothesis about the events has to be developed and tested to fully determine the cause of the incidents.

B. *Digital Forensic Readiness*

Rowlingson [5] has defined DFR as consisting of two objectives: To maximise the environment's capability of collecting digital forensic information and to minimise the cost of performing a forensic investigation. It is worth noting that this research is inclined towards Rowlingson's objectives.

Based on this research it is the authors' opinion that the collection, preservation, presentation of digital evidence from digital sources for purposes of furthering reconstruction of events are the most important relevant aspects that Palmers [6] defined during the first Digital Forensic Research Workshop (DFRWS) in 2001 at Utica, New York. In this context, collection and preservation can be mapped with DFR as presented by Rowlingson [5], while reconstruction of events is mapped with examining the properties and characteristics of digital events during event reconstruction as highlighted by Carrier and Spafford [1].

Nevertheless, an environment that is forensically ready makes the effort that is needed to conduct a Digital Forensic Investigation (DFI) to be minimised. This can only be achieved through collection, validation, preservation of critical information that is related to crimes as per ISO/IEC 27043 [13]. Furthermore, information gathered from digital sources is then used as part of a DFI. The next section discusses the ISO/IEC 27043 readiness process groups.

C. *ISO/IEC 27043 Readiness Process Groups*

The ISO/IEC 27043 [11] is a standard that deals with security techniques, incident investigation principles and processes that have been proposed by Valjarevic and Venter [7]. ISO/IEC 27043 clearly defines readiness through its classes of processes as a process that occurs before incident detection. This has been presented as a proactive process that precedes incident detection. At the time when this paper was written ISO/IEC 27043 was published as a standard for IT, security techniques and incident investigation principles and processes [11].

Additionally, ISO/IEC 27043 has defined readiness process groups that are able to maximise the potential use of digital evidence in order to reduce the cost of a DFI. Moreover, this has been done as a measure that can improve organisations' level of information security of systems. The readiness process groups are grouped into three: planning process group, implementation process group and assessment process group. This is shown in Figure 1.

Planning process group defines the scenario, identifies PDE sources, and plans pre-incident collection, storage of PDE, pre-incident analysis planning, planning incident detection and defining the system architecture. Next the implementation process group implements the planning process group activities and finally the assessment process group performs the following activities; assessing the implemented process and implementation of assessment result processes. The concurrent process shown by the arrow pointing downwards allows the processes to be executed as on-going [11].

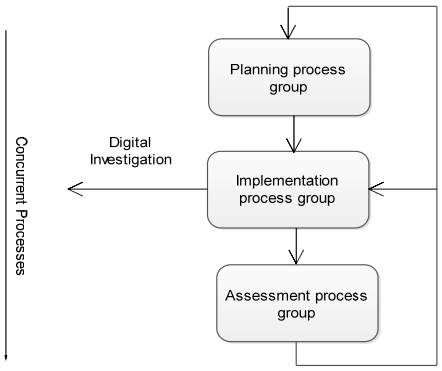

Figure 1.Readiness process groups, (Source, [11])

Having looked at the ISO/IEC 27043 readiness process groups, it is evident that there is no event reconstruction of security events throughout the readiness process groups and the readiness spectrum has been opened further. The next subsection discusses previous and related work.

III. PREVIOUS AND RELATED WORK

The Cloud Forensic Readiness (CFR) model shown in Figure 2 depicts a DFR approach in the cloud environment. The model presents a proactive forensic approach that is based on active monitoring, gathering and retaining digital information within the cloud environment [2]. The main role of CFR is to collect relevant digital information existing as PDE that can be used to support or refute a hypothesis based on the occurrence of a digital event. The relevant information collected may include: access logs, hypervisor logs, networks and activity logs [2]. Figure 2 shows a high-level view of the CFR model.

The CFR model consists of Cloud Service Provider (CSP) and the Non-Malicious Botnet (NMB) "infection".

It is worth noting that "infection" in this context involves modifying the originally considered malicious botnet to collect digital information with a positive connotation. The CSPs provide virtual services to cloud clients where proactively the NMB harvests digital information that can be used as potential evidence. This has been shown by the arrow pointing downwards named proactive process.

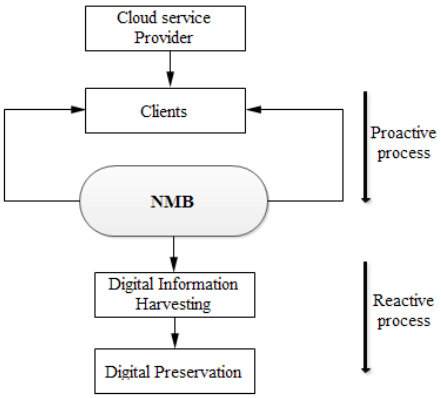

Figure 2.High-level view of CFR model (source; [2])

The evidence retained through the CFR is digitally preserved and retained for DFR purposes. Additionally, evidence is collected according to the guidelines and readiness processes explained in the ISO/IEC 27043 which are also highlighted in Figure 1. Additionally, in previous work [10],[12], the NMB is obfuscated to avoid deterrence based on the botnet detection and infiltration strategies.

According to existing literature [1],[3],[4], event reconstruction processes at digital crime scenes have been proposed before, however more focus has been on physical crime scene but not on digital forensic readiness perspective in the cloud. A well-documented research by Carrier and Spafford [1] highlights that evidence should have been recognised and collected at the crime scene when the process of event reconstruction begins. Further these authors present the five phases in reconstruction process as: Evidence examination, role classification, event construction and testing, event sequencing and hypothesis testing. Through these phases, the authors are able to identify an object as being the initiator that causes a particular event to occur using a role-based event reconstruction model.

Research by Liao and Langweg [8] has proposed a resource-based event reconstruction of digital crimes prototype which includes a readiness phase that helps to ensure that evidence is admissible. The prototype corresponds to the DF framework and it has the following phases: Readiness for collecting system call traces, deployment phase for receiving detection alerts, investigation phase for preserving and recognising evidence and reconstructing events. Basically the prototype's feasibility is assessed based on the applicability of the existing application. However, the cloud is hardly mentioned in this research.

An automated timeline reconstruction approach for digital forensic investigations is a Python-based framework that enables automatic reconstruction of high-level events. This approach consists of a framework that allows forensic analysts to visualise high-level events using existing tools while preserving the provenance of the high-level events. Further this approach uses a pattern matching method to automatically reconstruct high-level and humanly understandable events [9].

While our study is inclined towards the proactive approaches, the aforementioned approach is integrated into the digital investigation process that mainly focuses on the analysis phase. We highly acknowledge research presented above as related work by other researchers. Additionally this has given more insight on the relevance of this area, in the next section the reader is introduced to the enhanced cloud forensic readiness model.

IV. ENHANCED CLOUD FORENSIC READINESS MODEL

Before we present the process for adding event reconstruction, we give a brief introduction to an Enhanced Cloud Forensic Readiness model (ECFR), event reconstruction is implemented in this model. Figure 3 shows a block diagram ECFR. It is composed of the following modules: Cloud Service Providers (CSPs), clients, Non-malicious Botnet (NMB) "infection", information harvesting, hashing, digital preservation and event reconstruction. Details of the composition of the ECFR have been explained in the following section.

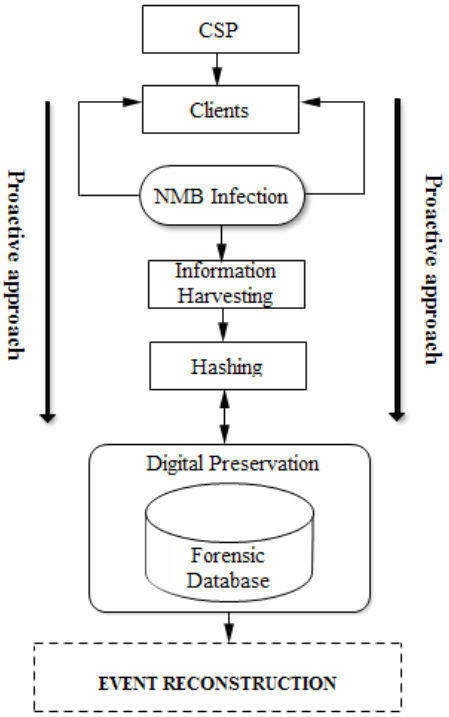

Figure 3.Block diagram of ECFR

The ECFR has *N* number of Virtual machines (VMs) in the cloud that are provided by the cloud environment. The *N*

number of VMs is “infected” by the Non-Malicious Botnet (NMB) that captures possible attack logs as potential digital evidence in the logging process through an NMB “infection” process. As discussed in section III infection carries a positive connotation whereas the NMB is used as an agent.

Next, the captured evidence is then hashed to preserve the integrity of the PDE, this is done using blocks of hashes before digitally preserving digital evidence. We are adding event reconstruction to the CFR model. This is shown using a dotted rectangle at the bottom of Figure 3. The dotted rectangle constantly represents event reconstruction throughout the paper.

Event reconstruction is used to analyse and examine the PDE in order to identify why it holds certain characteristics. This helps in identifying the causality of the possible event and also helps to build a hypothesis before a DFI is done. In the later section (section 4.1), we have discussed how the reconstruction of events has been added to this model.

Hypothetical Scenario:

P is a Cloud Service Provider (CSP) offering cloud services to organisation SX. Furthermore SX has enforced digital forensic readiness approaches for its organisation as shown in Figure 4. For a given period of time, there has been no suspicious incident within organisation SX until client X complains of possible Internet Protocol (IP) flooding. X experiences this for a couple of hours and finally S decides to seek the services of a Digital Forensic Investigator (DFI), D. We can see this is a problem and we want to investigate it. Therefore, the authors have come up with a process to do this investigation that is basically going to happen during event reconstruction process as shown in Figure 5.

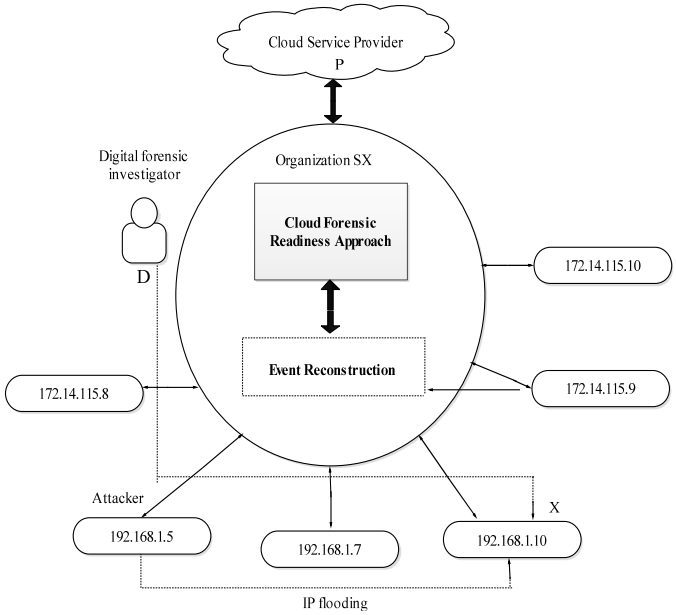

Figure 4.Event reconstruction scenario

A. Adding Event Reconstruction to a CFR Model

This section proposes a process for adding event reconstruction in a CFR model as a contribution for the enhanced model. The next section briefly highlights the high-level overview of event reconstruction process while in a later section a detailed event reconstruction process is discussed.

1) High-level overview of Event Reconstruction Process

Figure 5 shows the high-level view of the reconstruction process that is divided into five sections. The first section is the retrieval of initially preserved digital data (see fig 3). Next, is clustering of data. Clusters are able to be detected using the algorithm that has been described in (section IV part d), for example log files containing usernames and passwords may each represent a category of clusters respectively. Next is a search function that checks the availability of events in clusters. Thereafter, a similarity measure follows that matches events based on the distance function. The final part of the high-level view gives an event report on the reconstructed events.

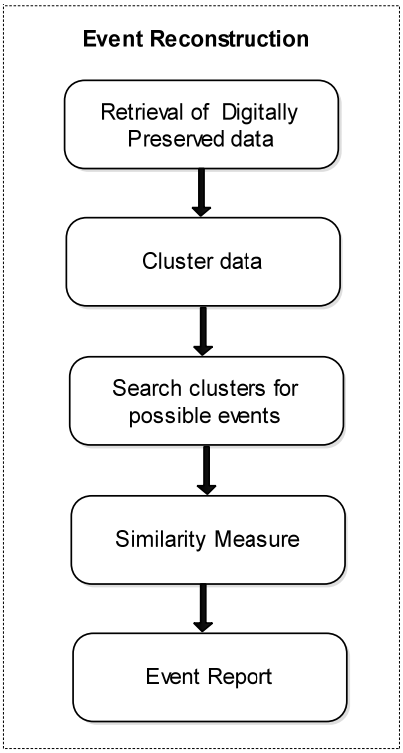

Figure 5. High-level view of event reconstruction process

2) Detailed Event Reconstruction Process

This section provides a detailed discussion on reconstruction process. The goal of the detailed reconstruction process is to show the systematic approaches that have been used to add event reconstruction in the ECFR model. Each of the processes shown in Figure 5 is described as a subsection in this section with respect to the hypothetical scenario described

previously, however Figure 6 shows the UML sequence diagram that shows the procedural flows for the process shown in Figure 5. An explanation of each procedure has been given below.

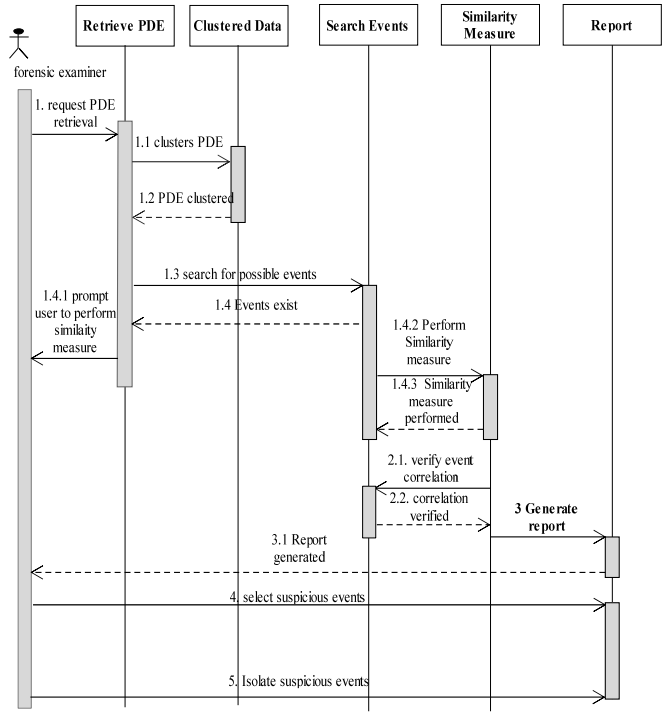

Figure 6.procedural flows of reconstruction process

2.1 Procedural Flows for Reconstruction Process

Figure 6 provides a description of the procedural flows for event reconstruction process as shown previously in Figure 5. It consists of five parts: Retrieval of PDE, clustering data, searching events, similarity measure and event reporting.

Firstly, in step 1 of Figure 6 the Forensic Examiner (FE) forwards the request to retrieve PDE from the initially preserved digital data in the forensic database that was previously shown in Figure 3. Next in step 1.1 the FE clusters the data and a response is given back on the availability of clusters in step 1.2. This is followed by the search for potential events in the clusters in step 1.3. Apart from that, the clusters are formed based on data of the same class. Example: IP addresses, usernames, passwords form a single cluster respectively. When the existence of an event is detected in step 1.4, the FE is prompted to perform a similarity measure on the event in step 1.4.1. A measure on whether particular events are similar is done in step 1.4.2 and control is returned in step 1.4.3 which is then followed by a verification to show if the events correlate in step 2.1 and step 2.2. Event report consisting of reconstructed events is generated in step 3 and 3.1. Eventually the FE has options of selecting and isolating suspicious events in step 4 and 5.

In the next section we give a detailed explanation of reconstruction process highlighted in Figure 5 with respect to the hypothetical scenario presented previously in section IV.

a) Retrieval of Digitally Preserved Data

Based on the scenario highlighted in Figure 4, *D* has no idea from where to begin his investigation , but organisation *SX* tells *D* that the organization is forensically ready for security incidents. In this process the first step that *D* has to take is to retrieve, digitally preserved PDE from forensic databases for reconstruction purposes as shown in Figure 3. Forensic database in the cloud as described by Kebande and Venter [13] consists of collected, retained and digitally preserved data as evidence. These may include; forensic logs, monitored data, service artifacts, VM images, activity logs, hypervisor logs and system logs.

b) Cluster Data

As soon as *D* has access to the PDE that he requires to perform an investigation he forms a category of clusters to help in grouping the events based on their occurrence and later their similarity. In order to cope with the need of clustering, we use the following technique;-we first cluster data based on categories and then check for the occurrence of possible events. The occurrence (*X*) of categorical clusters A_i of events e_i with timestamp (*t*) are reconstructed based on cluster name (*Clu-N*) and the interval between the events that is denoted as the distance (*d*). When a distance exists between the events e_i in A_i, it shows that the events exhibit different timestamps hence an event may have occurred. The pattern of occurrence of e_i can be computed using the distance function $d(e_i)$.

$$
\begin{aligned}
&A_{1(Clu\_N)} \rightarrow \{\, w_{1t} \rightarrow w_{2t} \rightarrow w_{3t} \ldots\ldots\ldots\ldots w_{nt} \,\} \\
&A_{2(Clu\_N)} \rightarrow \{\, x_{1t} \rightarrow x_{2t} \rightarrow x_{3t} \ldots\ldots\ldots\ldots x_{nt} \,\} \\
&A_{3(Clu\_N)} \rightarrow \{\, y_{1t} \rightarrow y_{2t} \rightarrow y_{3t} \ldots\ldots\ldots\ldots y_{nt} \,\} \\
&\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots \\
&A_{n(Clu\_N)} \rightarrow \{\, z_{1t} \rightarrow z_{2t} \rightarrow z_{3t} \ldots\ldots\ldots\ldots z_{nt} \,\}
\end{aligned}
\qquad \Big\} \; ei \qquad (1)
$$

The assumptions in deriving the technique shown in *Eqn (1)* are that a cluster A_i is tagged with respective cluster name (*Clu_N*). Thereafter, e_i represents the following possible events; w_{1t}, x_{1t}, y_{1t} and z_{1t} represent first events for A_1, A_2, A_3, A_n clusters respectively. This is followed by w_{2t}, x_{2t}, y_{2t} and z_{2t} as the second events for cluster A_1, A_2, A_3 and A_n , finally the third and last events are represented using w_{3t}, x_{3t}, y_{3t} and w_{nt}, x_{nt}, y_{nt} and z_{nt} for cluster A_1, A_2, A_3, A_n respectively. Investigator *D* finally clusters the data based on IPs, usernames and access logs as A_1(*IP*), A_2(*User_name*), A_3(*Access_logs*) to represent clusters under different categories respectively and thereafter he finds the need for

searching for possible events. Next *D* performs a search as shown in the next subsection.

c) Searching Clusters for Possible Events

In this sub-process, *D* performs a look-up using a search function (*S*) while trying to detect the possible events that are present within the clusters of *IPs*, usernames or access logs in the retrieved original data (*P*) presented as A_1(*IP*), A_2(*User_name*), A_3(*Access_logs*) respectively. *P* is assumed to contain clusters through which possible events might occur. When *D* goes through the sequence of events in the IP cluster he discovers something very alarming, he discovers that a particular cluster of IPs containing address *192.168.1.5* flooded *client X* with *IP 192.168.1.10*, 48 times in a span of 10 seconds. Moreover, D discovers that the flooding originated from a single source.

Figure 6 shows a search function S that performs a look-up in *P*. *P* is a set of collected potential digital evidence. *S* returns the cluster A_1 A_2, $A_3 \ldots A_n$ with respective cluster names (*Clu_N*), based on step 2 of the algorithm in the next sub-section. Further, e_i represents the possible events in A_i while w_1, x_1, $y_1 \ldots z_n$ represents the possible events in A_1 A_2, $A_3 \ldots A_n$ respectively. Search function(S) should not be confused with organisation *S* in the hypothetical scenario.

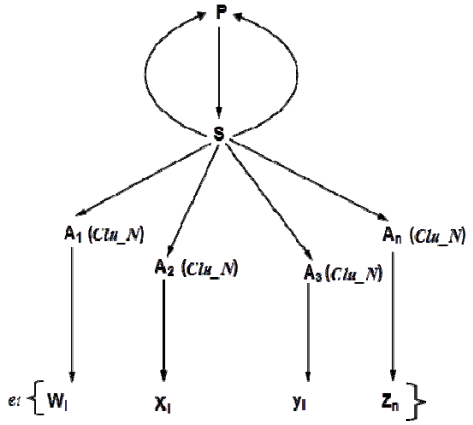

Figure 7. Search function(S)

d) Similarity Measure

In this sub-process *D* performs a similarity measure on the events in the cluster A_1(*IP*) of IP's that contains address *192.168.1.5* and checks the pattern of occurrence and a reflection of the timestamps. *P* is assumed to be a set of collected potential evidence in a digital forensic readiness approach while $\{A_1, A_2, A_3 \ldots A_n\}$ are different clusters as shown in Figure 7. If $\{A_1, A_2, A_3 \ldots A_n\}$ are distinct clusters with events e_i, (say a flood from *192.168.1.5*) that happened 48 times after the interval of 10 seconds, then the pattern of occurrence of two given events say w_{1t} and w_{2t} can be computed using the following distance function d^{MD} (w_{1t}, w_{2t}) where w_{1t}, is the first event IP flood and w_{2t} is the second event IP flood and *d* is the distance between the events. The general distance function equation is given by *Eqn 2*.

$$d^{MD}(w_{1t}, w_{2t}) = \sqrt[p]{\sum_{i=1}^{n} | w_{1t} - w_{2t} |^p} \quad (2)$$

where*[small(p)=1, 2….n]* for the distance function of an event w_{1t} and w_{2t}, d^{MD} (w_{1t}, w_{2t}) is the distance between event w_{1t} and event w_{2t}. The similarity, behaviour and pattern of the events e_i can be calculated using the distance function d^{MD} (w_{1t}, w_{2t}) given *small(p)*, this is shown in *Eqn 2*. The similarity between the events should be noted so as to ascertain the initiator of the event if the process is not continuous. Table 1 shows the notations used and the algorithm represents the procedure followed when performing a similarity measure.

ALGORITHM

1. *Let P be the set of all collected potential digital evidence, Clu_N be the cluster name, t be the timestamp, X the occurrence of events, A_i be a distinct cluster of events and S be the search function.*
2. *Search P for possible cluster A_i.*
3. *For each $A_i \in P$, step number 4,otherwise step number 8*
4. *Search A_i in P having e_{it}.*
5. *While e_i>1, determine d i.e how close (similar) the events e_i are, using the distance function $d(e_i)$,otherwise step no 2.*
6. *If t for events e_i is the same consider the events to be identical, match those events .*
7. *Return e_i.*
8. *End.*

Table 1. Procedure Notation

Notation		Meaning
X	:	Occurrence
Ai	:	Cluster
ei	:	Number of events
t	:	Timestamps
Clu_N	:	Cluster Name
d	:	Distance metric
d(ei)	:	Distance between events
S	:	Search
P	:	Potential digital evidence

e) Event Report

Reporting is an integral part of the event reconstruction process and typically reporting contains the information and descriptions of all the steps taken towards potential evidence examination, classification and how event reconstruction hypothesis is formulated.

Moreover, ISO/IEC 27043 describes this as results from digital evidence interpretation process, at this stage the roadmap of the readiness process and the event properties is interpreted and presented to DFI and law enforcement agencies. Reporting is aimed at generating an audit record that shows the scope, occurrence and the characteristics of the

events [12]. The following are the possible exhibited characteristics of events that might be presented during reporting.

- **Event type**: Checks the event type whether it has inbound threats or outside threats.
- **Timestamps**: Gives the exact time when the event happened.
- **Sender**: Shows who was sending.
- **Receiver**: Shows who was to be the receiver.
- **Content:** Shows the content that the messages carry.
- **Attack Details**: Nature of the attack.
- **Category**: Shows how the event be classified.
- **Where**: The session where the event happened.

Having looked at event reconstruction process, in the next section, the reader is introduced to the experiments and discussions of this study.

V. EXPERIMENTS AND DISCUSSIONS

In this section, we present the experimental findings and discussions on the suitability and applicability of the proposed method and processes.

A. Experimental Data

Based on the similarity measure described in (sub-section 2 of sec IV), we have conducted a comparative analysis on the distance function with the values of *small(p)* using the attributes that e_i exhibits. By revisiting *Eqn 2*, *small(p)=1, 2,.n*, we take *small(p)* as the number of occurrence of possible events during event reconstruction using *Eqn 2*. We assume that the attributes of the given two events w_{1t} and w_{2t} are represented as a set respectively. We therefore take random non-negative values for size, time and occurrence attributes as w_{1t} (*0, 5, 3*) and w_{2t} (*6, 7, 2*). Table 2 shows the attributes. Afterwards we test the attributes when *[p=1, p=2 and when p>2 to ∞]* using *Eqn 2*.

Table 2.w_{1t} and w_{2t} events with attributes

Event (e_i)	size	t	X
W_{1t}	1	5	3
W_{2t}	6	7	5

When *p=1* in *Eqn 2*, we find the absolute difference between the pair of event attributes by examining the absolute value distance. This is given by

$$d(w_{ij}) = \sum_{k=1}^{n} |w_{ik} - w_{jk}| \tag{3}$$

When *p=2* in *Eqn 2*, we find the root of square differences between the set of event attributes by examining the distance metric. This is given by

$$d(w_{1t}, w_{2t}) = \sqrt{\sum_{i=1}^{n} (w_{it} - w_{2t})^2} \tag{4}$$

When *p>2 to ∞* in *Eqn 2* the maximum value distance is checked by examining the absolute difference in magnitude between the set of event attributes. The distance metric is given by

$$d(w_{1t}, w_{2t}) = \max_i |w_{1ti} - w_{2ti}| \tag{5}$$

B. Experimental Findings

To evaluate this approach, we have selected random non-negative attributes based on two sets shown in Table 2, then we evaluated them using the proposed *Eqn 3, Eqn 4 and Eqn 5* as shown in (3),(4) and (5) respectively, to find the inferences in Table 3. Furthermore, these attributes represent the possible events e_i. On the other hand, Figures 8, 9 and 10 represent our experimental findings that are depicted in Table 2, which shows the convergence of the distance metric *d* when *small(p)* is tested with *[1, 2 and >2 to ∞]* in equation (3),(4) and (5) respectively.

The computed values of *p=1, p=2 and p>2 to ∞* are shown in Table 3. Table 3 lists the distance metric based on the occurrences(x) of *small(p)* that has been shown in *Eqn 2*.

Table 3.w_{1t} and w_{2t} events with attributes and values when *small (p)=1,2*>to ∞ in *Eqn 3, 4 and 5*

Event (e_i)	size	t	X	P=1	P=2	p>2to ∞
W_{1t}	1	5	3	8	5.744	5
W_{2t}	6	7	2	8	5.744	5

Based on the value of *small(p),* we have shown the distance metric for attributes that represent events. We will give a discussion on event reconstruction based on *small(p)* value however, Figures 8, 9 and 10 show the corresponding experimental findings respectively.

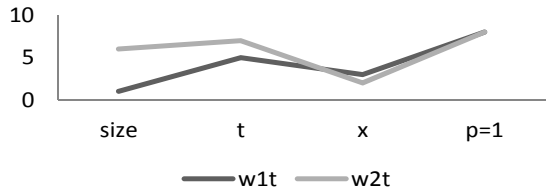

Figure 8. Similarity measure for *small(p)=1.*

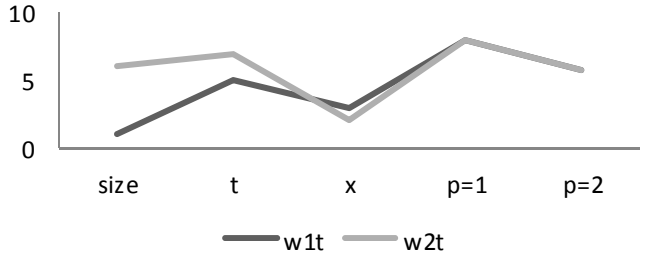

Figure 9. Similarity measure for *small(p)=2.*

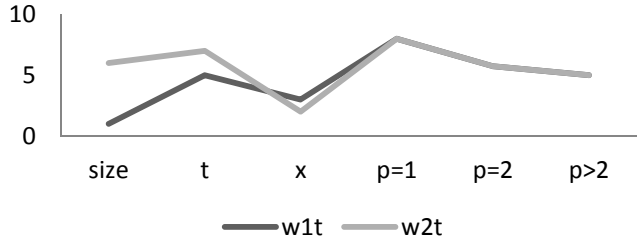

Figure 10.Similarity measure of events for *small(p)>2* to ∞.

C. Further Discussions

Judging from the hypothetical scenario, the observation made by *D* is that there is a close match between the events that occurred to c*lient X*, which eventually showed that there was a possibility that these events had some similarity. *D* was able to observe this based on the event timestamps that were closely related with the distance metric d^{MD} (w_{1t}, w_{2t}) between the two events(w_{1t}, w_{2t}) being the same after checking the similarity measure. This happened after making 48 observations. Through these parameters, *D* is able to correlate the causal relationship between the sequence of events after fragmenting the A_1(*IP*), *IP* cluster of data with the address *192.168.1.5* that flooded *client X* within the same time interval of 10s.

Consequently, the investigation conducted by *D* was timeous because he naturally did this through performing a comparative analysis of the events using the potential evidence that was at his disposal.

Nevertheless, for each of the tasks shown in Figure 5, the procedure proceeds as follows; firstly, potential digital evidence that is to be used for DFR purposes is captured in a proactive approach from the cloud environment as shown in Figure 2 and stored in the forensic databases as described in [14]. This approach allows the cloud to be forensically ready through implementation of the readiness as highlighted in the standard of ISO 27043[12]. Thereafter event reconstruction is added to this model as shown with the dotted square.

The idea in *Eqn 2* is that d^{MD} (w_{1t}, w_{2t}) represents the Minkowski distance metric between two possible events w_{1t} and w_{2t}. On the one hand *Eqn 3, 4 and 5* represent the distance metric based on the occurrence of *small(p).* We first observe Table 3; when the value of *small(p) =1*, the distance metric for w_{1t} and w_{2t} is computed as 8 while when *small(p)=2* the distance metric is computed as 5.744, finally when *small(p)>2 to* ∞ it is computed as 5. Besides, this has been demonstrated in Figure *8, 9* and *10* respectively. In our opinion, the outcome of Figure 8 on the similarity measure between w_{1t} and w_{2t} shows a close match when *small(p=1),* therefore this outcome is sufficient enough to prove the concept of similarity measure in this context.

On the other hand, during event reconstruction, the purpose of computing the distance metric *d* between possible events e_i is for the following reasons:

- To be able to distinguish one event from the other during event reconstruction.
- Enables a discovery of the structure of the events.
- Helps to distinguish one event from the other by focusing on the relationship between the events.
- Prediction of the behaviour of events.

In spite of that, the possible characteristics of any event during event reconstruction process might not be the solution to any potential security incident rather it will be used to develop a hypothesis in preparation for a digital forensic investigation. Carrier and Spafford [4] highlight that a hypothesis should always be created showing how a suspect's file got into the web browser's cache if it exists. In the next section we give a conclusion of the study and state future work.

D. Threat Model

In this section a threat model that addresses the issues that may arise based on reconstruction of events using the ECFR as depicted in the scenario that has been presented in part III. A number of challenges that may face the ECFR model have been documented on draft NISTIR 8006 by the National Institute of Standards and Technology (NIST) cloud computing forensic science working group (NCC FSWG) [14]. NCCFSWG has documented 65 cloud forensics challenges, however, Kebande and Venter [15] have also identified a number of challenges that are experienced as a result of the implementation of the CFR model that is shown in Figure 2. The threats identified in [15] have been mapped with the challenges in the draft NISTIR 8006. The following section gives a conclusion and future work.

VI. CONCLUSION AND FUTURE WORK

The Enhanced Cloud Forensic Readiness (ECFR) model is an enhanced version of Cloud Forensic Readiness that is aimed at observing the characteristics of potential evidence at crime scenes to help digital forensic investigators in identifying the causality. If we revisit the statement of the problem, "there exist no reliable event reconstruction processes in the cloud that can help in analysis and examination of Digital Evidence (DE) aspects, during Digital Forensic Readiness". The authors have proposed the addition of digital event reconstruction

process to the initially proposed cloud forensic readiness model.

The proposed ECFR can be used by digital forensic practitioners to aid in analysis of PDE at the crime scenes in the cloud during DFR. One advantage of the ECFR is the ability to show the characteristics and properties exhibited the potential evidence in the form of a hypothesis which might help to identify the causality.

Further the ECFR will use the standardised guidelines of ISO 27043 which means if a potential security incident camouflages itself within the forensic tools, the ECFR can help to outline its properties. Verification and validation of all the proposed study in this paper will be implemented using an applicable prototype as future work.

ACKNOWLEDGMENT

This work is based on research supported by the National Research Foundation (NRF) of South Africa (Grant-specific unique reference number UID85794). The Grant holder acknowledges that opinions, findings and conclusions or recommendations expressed in any publication generated by the NRF-supported research are those of the author(s) and the NRF accepts no liability whatsoever in this regard.

REFERENCES

[1] B. D Carrier, and E.H Spafford," Defining event reconstruction of digital crime scenes". Journal of forensic sciences, 49(6), 1291-1298, 2004.

[2] V.R.Kebande, & H.S.,Venter "A Cloud Forensic Readiness Model Using a Botnet as a Service. In *The International Conference on Digital Security and Forensics (DigitalSec2014)* (pp. 23-32). The Society of Digital Information and Wireless Communication, (2014).

[3] T.Bevel and RM Gardner."Bloodstain pattern analysis: with an introduction to crime scene reconstruction". 2nd ed. Boca Raton, FL:CRC Press, 2002.

[4] B. D Carrier, and E.H Spafford," An event-based digital forensic investigation framework". In Digital forensic research workshop (pp. 11-13), July 2004.

[5] R.Rowlingson, "A ten-step process for forensic readiness". International Journal of Digital Evidence, 2(3), 1-28,2004.

[6] G.Palmer, "A road map for digital forensic research". In First Digital Forensic Research Workshop, Utica, New York (pp. 27-30),August 2001.

[7] A.Valjarevic and H.S. Venter, "Implementation guidelines for a harmonised digital forensic investigation readiness process model". In ISSA (pp. 1-9),August 2013.

[8] Yi-Ching Liao and H.Langweg"Resource-Based Event Reconstruction of Digital Crime Scenes," Intelligence and Security Informatics Conference (JISIC), 2014 IEEE Joint , vol., no.,pp.129,136,24-26Sept.2014.

[9] C. Hargreaves and J. Patterson, "An automated timeline reconstruction approach for digital forensic investigations". Digital Investigation, 9, S69-S79, 2012.

[10] V.R. Kebande and H.S. Venter, "Obfuscating a Cloud-Based Botnet Towards Digital Forensic Readiness". In Iccws 2015-The Proceedings of the 10th International Conference on Cyber Warfare and Security (p. 434). Academic Conferences Limited, (2015, February).

[11] ISO/IEC 27043:2015,"Information technology -- Security techniques -- Incident investigation principles and processes".[online], Available at http://www.iso.org/iso/catalogue_detail.htm?csnumber=44407

[12] V.R. Kebande and H.S. Venter, "A Cognitive Approach for Botnet Detection Using Artificial Immune System in the Cloud",2014.

[13] V.R. Kebande and H.S. Venter, "A Functional Architecture for a Cloud Forensic Readiness Large-scale Potential Digital Evidence Analysis". In 13th European Conference on Cyber Warfare and Security ICCWS - 2015. Hatfield, United Kingdom,2015.

[14] Draft NISTIR 8006 NIST Cloud Computing Forensic Science Challenges. Accessed at http://csrc.nist.gov/publications/drafts/nistir-8006/draft_nistir_8006.pdf

[15] V.R. Kebande and H.S. Venter, "On Digital Forensic Readiness in the Cloud Using a Non-Malicious Botnet: Issues and Challenges". Journal of forensic sciences,-To appear.