An Event Oriented Approach to Digital Forensics for Tracking Criminals

ONYEMAUCHE U.C.
Department of Computer Science
Nnamdi Azikiwe University Awka
Anambra State, Nigeria.

OKONKWO O.R .
Department Of Computer Science
Nnamdi Azikiwe University Awka
Anambra State, Nigeria.

NWOSU Q.N.
Department Of Physical & Health Education
University Of Nigeria Nsukka
Enugu State, Nigeria.

Abstract: **In this paper, we present a framework for digital forensics that includes an investigation process model based on physical crime scene procedures. In this model, each digital device is considered a digital crime scene, which is included in the physical crime scene where it is located. The investigation includes the preservation of the system, the search for digital evidence, and the reconstruction of digital events. The focus of the investigation is on the reconstruction of events using evidence so that hypotheses can be developed and tested. This paper also includes definitions and descriptions of the basic and core concepts that the framework uses.**

Key words: **Reconstruction, Computer Forensics, heuristics rule.**

INTRODUCTION

Computer forensics emerged in response to the escalation of crimes committed by the use of computer systems and or other related digital devices either as an object of crime, an instrument used to commit a crime or a repository of evidence related to a crime. Digital investigations, or digital forensics, are conducted by law enforcement and corporate investigation teams on a regular basis. Yet, no formal theory exists for the process. A practitioner in the legal field can describe how he recognizes evidence for a specific type of incident, but the recognition process cannot typically be described in a general way.

This study will tend to emphasize highly efficient regimens in cyber threat, cyber related crimes, finger printing analysis, DNA testing and criminal investigative analysis (profiling) that result in capturing serial killers, fraudsters and other perpetrators of homicide.

The primary goal of this study is to examine digital media in a forensically sound manner with the aim of identifying, preserving, recovering, analyzing and presenting facts and opinions about the information.

Computer forensics has become an increasingly important tool in the constant battle against cyber crime, cyber terrorism, cyber stalking. Many national and international law enforcement agencies include specialty computer crime divisions to track down cyber fraudsters, hackers, stalkers, terrorists and pornographers. The agencies deploy computer forensics experts to gather evidence over the Internet and by examining computer hard drives seized during police raids. Many computer forensics organizations, including Expert Data Forensics, provide expert testimony, data recovery, chain of custody, forensic imaging and forensic investigations for the prosecution of civil lawsuits. An example of computer forensics used in civil investigation could be a separated spouse choosing to research the activities of their estranged spouse. Information on online dating, affairs, and other incriminating evidence that could affect divorce proceedings can be analyzed by criminal forensics experts.

Digital forensics has been defined as the use of scientifically derived and proven methods towards the preservation, collection, validation, identification, analysis, interpretation and presentation of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal or helping to anticipate the unauthorized actions shown to be disruptive to planned operations. One important element of digital forensics is the credibility of the digital evidence. Digital evidence includes computer evidence, digital audio, digital video, cell phones, digital fax machines etc.

Digital forensics performs a variety of activities. The most common is to support or refute a hypothesis before criminal or civil (as part of the electronic discovery process) courts. Forensics may also feature in the private sector; such as during internal corporate investigations or intrusion investigation (a specialist probe into the nature and extent of an unauthorized network intrusion).

The technical aspect of this investigation is divided into several sub-branches, relating to the type of digital devices involved; computer forensics, network forensics, forensic data analysis and mobile device forensics. The typical forensic process encompasses the seizure, forensic imaging (acquisition) and analysis of digital media and the production of a report into collected evidence.

RELATED RESEARCH

Event Correlation refer to an array of technique applied to comprehending the dynamic behavior of system, based on events and patterns of events in their history. Garfinkel(2010) uses correlation techniques to identify similar features across entire corpuses of drive, a technique which could prove useful for identifying computers with similar usage pattern. Finally, another useful form of classification is similarity. Fuzzy hashing is a technique which identifies files which are nearly similar.

Abbott et al(2011), have, in their Event Correlation for Forensics (ECF) research, translated textual log events into instances of a generalized data model (canonical form) implemented using a relational database to performing either interactive or automated scenario identification over these events.

Stallard et al(2010), employed an anomaly based expert systems approach to identifying semantic inconsistencies in investigation related data. Their approach translated MAC times generated by TCT and the UNIX last log into an XML representation, which was asserted into the HESS expert systems shell. Knowledge is encoded as heuristic rules which specify invariant conditions related to logins and potential file modifications.

Elsaesser et al(2006) employ an AI based approach to automated diagnosis of how an attacker might have compromised a system. Using a model of the topology of a network, the configuration of system, and a set of “action templates”, a class of artificial reasoner called a “planner” generates hypothetical attack sequences which could have led to a particular situation. These hypothetical attack sequences are them run in a simulated environment, and the generated logs compared with the logs of the real world system. The action templates correspond to specifications of how a particular action will transit the state of the world from one state to the next.

Approaches to event correlation in the IDS and network management domains have focused on single domains of interest only, and have employed models of correlation that are very specific in nature. Repurposing these specific existing approaches to the more general task of event correlation in the CF domain is made difficult for a number of reasons. Existing event pattern languages do not necessarily generalize the application in wider domains. For example, while state machine based event pattern languages may work well for events related protocols, they do not work well with patterns where time and duration are uncertain. Most approaches focus exclusively on events, an ignore context related information such as environmental data and configuration information. Furthermore, few approaches have available implementations in a form that is readily modifiable.

Where we have modifiable implementations of event correlation systems, we find that extension is complicated by the software paradigm underlying its implementation, and that the systems are weak on semantics.

Adding new vocabulary to the event language is slowed because of compilation and linkage overheads. Addition of concepts outside of the event pattern language require reengineering of the STATL LANGUAGE compiler and supporting framework.

HEURISTIC EVENT CORRELATION FOR DIGITAL FORENSICS USING ABSTRACT HEURISTIC RULE

The expressiveness of RDF/OWL enables the translation of event log entries into instances of information with fixed and specific semantics. The presence of class/subclass relationship in the event forensics ontology enables the definition of abstract classes of events sharing similar characteristics for example a correlation rule composing a FileReceiveEvent will, in the presence of an ontology describe a WebFileDowloadEvent (an event sourced from web server lods) as a subclass of FileReceiveEvent). Just as happily match the later more specific event.

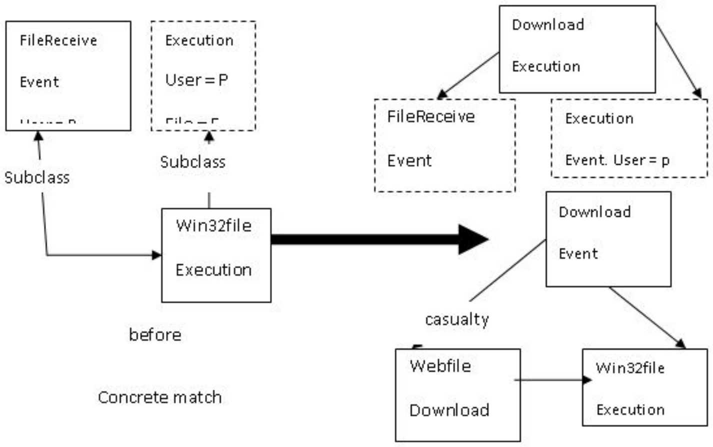

Fig.3.1 Event Correlation Frequency Using Heuristics Rule

3.1 Architecture of the proposed system

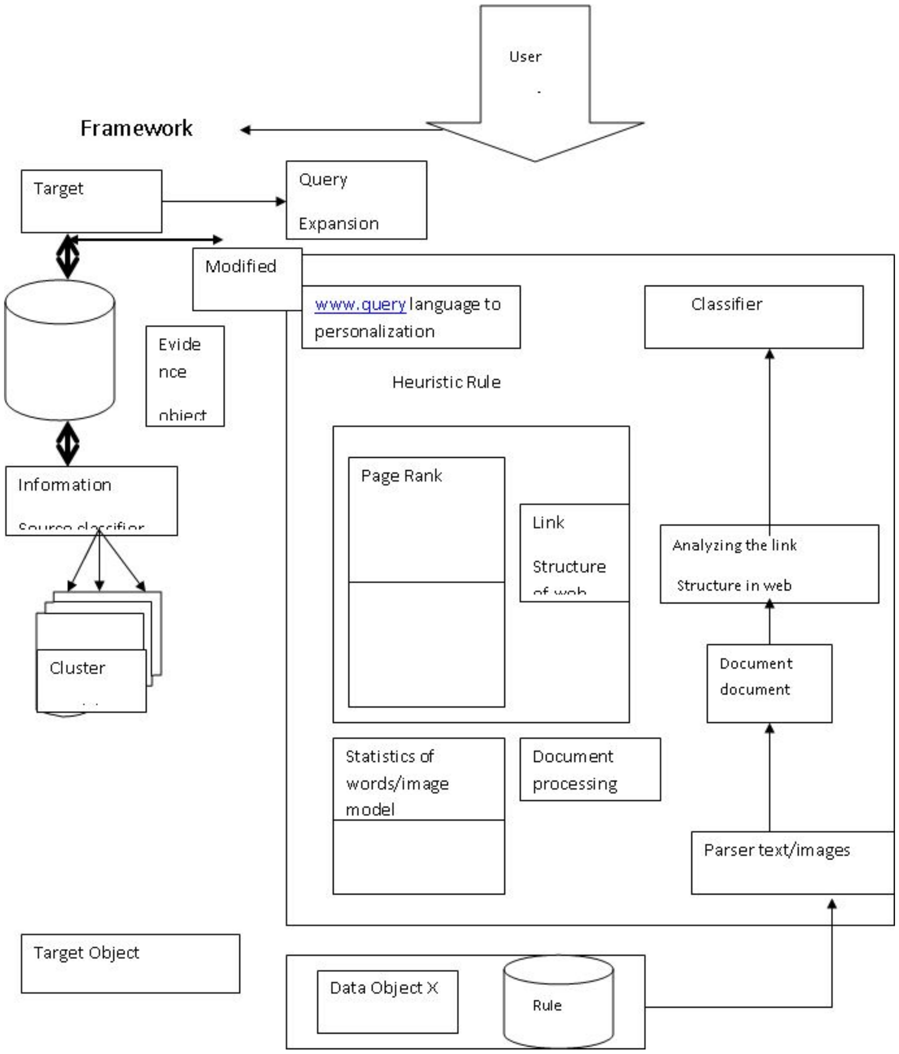

Fig 3.2 High Level Model Of The Proposed Solution

The web is similar to a graph, in that links are like edges and web pages are like nodes. Several approaches have been proposed to overcome the current limitations of ascertaining the integrity of digital evidence. some approaches use employ an AI based approach to automated diagnosis of how an attacker might have compromised a system. Using a model of the topology of a network, the configuration of system, and a set of "action templates", a class of artificial reasoner called a "planner" generates hypothetical attack sequences which could have led to a particular situation. Approaches to event correlation in the IDS and network management domains have focused on single domains of interest only, and have employed models of correlation that are very specific in nature. Repurposing these specific existing approaches to the more general task of event correlation in the CF domain is made difficult for a number of reasons. We shall use the Heuristics rules that will employ the RDF/OWL formalism for representing arbitrary event log related information, and higher order concept such as casual relationship of heterogeneous event logs. Our correlation approach relies on heuristics rules which abstract low level situation of interest into higher level situations. These rules were in this research developed using domain knowledge. From the design, we expect automated means of identifying rules, such as data mining.

3.1.1 Properties Of The High Level Model

The architecture has many components with specific function to track suspicious documents.

i) User Interface: This is the contact point between the user and the agent. It receives user data in the form of query and presents the relevant search results (Halavais, 2008).

ii) Knowledge base: It includes varied sub symbolic representation on event categories. Each of the representations is got through analyzing the co-occurrence chance of the main words in document within a given topic. These would propose vital terms used to reorganize exact term in relation to certain characteristics of a given subject domain Taksa, 2009).

iii) Information Sources: refers to the totality of data sources within the Internet and an example includes Database.

iv) The query Expansion Model: Query expansion is a technique, widely used in information retrieval, for obtaining additional terms relevant to a given query. It is usually used to help information searchers express their intentions more accurately and increase the precision of search results. The relevant terms are extracted from those documents. Since, there is immense corpus on the web, the terms relevant to any kind of search keyword can be obtained, even peculiar events and technical terms.

ALGORITHM

Event Reconstruction Module

Weight-Bound-Search (WBS)

```
WBS(Graph,Source,Destination,GapWeight)
getNieghbours(Source)  nei flist of neighbours
for i = 1 to i _ size of(nei) do
    if nei[i] = \exhausted" then
        Remove(nei[i]) remove nei[i] if it is tagged as exhausted"
        Continue floop to the next neighbor nei[i + 1]
    end if
    get Edges(nei[i]) ! nei[i].edges fget all nei[i]'s outgoing edges
    for j = 1 to j _ sizeof (nei[i].edges) do
        if nei[i].edges[j] 6= \visited" then
            Counter++ fHow many \unvisited" edges nei[i] has
            nei[i].edges[j] ! nei[i].current Edge
            nei[i].edges[j] = \visited" ftag as visited"
            nei[i].untraversed Edge = true
        end if
    end for
    if nei[i].untraversed Edge = true then
        if Counter = 1 then fif there was only 1 \unvisited" edge
            nei[i] = \exhausted" tag nei[i] as exhausted"
        end if
        break  out of the if statement and traverse nei[i + 1]
    end if
end for
if nei = ? then
    Terminate WBS if all neighbors are \exhausted"
end if
for i = 1 to i _ size of(nei) do
```

```
get Weight(nei[i].current Edge) = wei[i] call mobility model to assign a weight
_ = _ + wei[i] update the time variable counter
push Route(nei[i].current Edge, nei[i]) update route list
if nei[i] = Destination then
if Gap Weight  Gap Weight + _ then
return _; _
Back-up(nei[i].current Edge, nei[i], wei[i], _ , _) back 1 level up to Source
Continue iterating the next neighbor
else if GapWeight + _ then
Back-up(nei[i].current Edge, wei[i], _ , _); Continue
end if
else if nei[i] G or nei[i] 2 _ then fif childless/End-tail vertex is reached
Back-up(nei[i].currentEdge, wei[i], _ , _); Continue
else if    GapWeight   then
WBS(Graph,nei[i],    Destination, Gap Weight) recursion
end if
end for
```

CONCLUSION

As agents gain a wider acceptance and become more sophisticated, they will become a major factor in the future of the Internet. Through a thorough review of the literature that is available on the topic, an analysis of autonomous tracking system search agents are created, guiding the manner through search engines retrieval process and effectively reducing search engine time spent. The result regarding temporal provenance indicate the event correlation processes would benefit from richer models of temporal progress including time scale derivation, event time uncertainty, and orthogonal to this, assumption about these. It would appear likely that their effects on algorithmic complexity of a correlation approach would be adverse to a high degree.

REFERENCES

[1]. Abbott, M. A, Gbendo and K.A.(2006). A Comprehensive Approach To Digital Incidence Investigation. Elsevier Information Security Technical report. PP 10 – 15.

[2]. Elsaesser, M.N., Artificial Intelligence in Digital Forensics. Explanation in Knowledge Systems. ELBS publishers New Jersey, 5th Edition, Pp 21 – 23.

[3]. Halavais, K.F.(2008). Criminal Shadows. Journals of forensics identification. Vol 3, pp 42.

[4]. Garfintel, S.L., Milan, D, J.(2011). Disk imaging with the advanced forensics format, library and Tools. Advances in Digital Forensics(2nd annual International conference on Digital forensics. Pp 10 – 11.

[5]. Stallard, M.K., Hannam, M.P.(2010). Electronic Crime Scene Investigation. A guide for first Responders. National Institute of Justice, Washington DC, pp5.

[6]. Taska, R.V.(2008). Application of Theorems Proving to solving problems. Proceedings from the 1st International joint conference on Artificial Intelligence. Stanford Research Institute, Artificial Intelligence group. Pp 50.