HoneyNetCloud Investigation Model, A Preventive Process Model for IoT Forensics

Jayakrishnan Anilakkad Raman1*, Vasanthi Varadharajan2

1 Department of Computer Science, Ratham College of Arts and Science, Bharathiar University, Coimbatore 641021, India
2 Department of IT, SriKrishna Adithya College of Arts and Science, Coimbatore 641042, India

Corresponding Author Email: jk.anilakkad@gmail.com

https://doi.org/10.18280/isi.260309

Received: 2 April 2021
Accepted: 3 June 2021

Keywords:
network forensics, honeypots, IoT attacks, preventive model, forensics process model, HoneyNetCloud

ABSTRACT

With the pervasive usage of sensing systems and IoT things, the importance of security has increased. Attempts towards breaching IoT security systems by attackers are on upsurge. Many intrusions in embedded systems, sensing equipment and IoT things have occurred in the past. Though there are cyber security tools like Antivirus, Intrusion detection and prevention systems available for securing the digital devices and its networks. However, a forensic methodology to be followed for the analysis and investigation to detect origin cause of network incidents is lacking. This paper derives a comprehensive preventive cyber forensic process model with honeypots for the digital IoT investigation process which is formal, that can assist in the court of law in defining the reliability of the investigative process. One year data of various attacks to the IoT network has been recorded by the honeypots for this study. The newly derived model HIM has been validated using various methods and instead of converging on a particular aspect of investigation, it details the entire lifecycle of IoT forensic investigation. The model is targeted to address the forensic analysts’ requirements and the need of legal fraternity for a forensic model. The process model follows a preventive method which reduce further attacks on network.

1. INTRODUCTION

The discernment of extensive computing and sensing has been made general through the IoT, which integrates the benefits of intelligent embedded arrangements and the power of linked Internet-driven computing, service provision, and management [1]. Varied classes of IoT gadgets are developed having abilities to acquire data automatically, enable control, and conduct networking. Thus, it becomes obvious that soon, and in the long term, variety of computing gadgets will be pervasively deployed that would be linked by varied communicating methods. Usage of advanced security tools and firewall protections have their limitations in real scenario, such loopholes are being taken as an advantage by hackers to penetrate the network [2]. Once any IoT attack happens in a network system, it becomes then the forensic investigators responsibility in finding the root attack source. Nevertheless, thousands of tools, applications and methods are currently available to help and support the investigation, a specific process model ready and handy to follow the investigation for IoT forensics doesn’t exist [3]. This research derives a comprehensive preventive cyber forensics process model which uses honeypots for the digital IoT investigation process that is formal and can assist in court for defining the reliability of investigation process.

2. BACKGROUND

Literature study on a few of the generic digital cyber forensics model reveals that there exist weaknesses in various stages. Though, many of them were robust models at their times, however since technology keeps changing rapidly, they require significant updates to come up to the par of current worlds technology [4]. Moreover, forensics for IoT network in specific doesn’t exist. Stephenson proposed FORZA (Forensics Zach Man) model which could solve complex problems by integrating reactions to the questions, however the model is human dependent. The forensic investigation of IoT applications were performed in standard acceptable models, so that evidences collected which are relevant are acceptable at the court [5]. Standard phases of forensic investigations which are establishing context, collection, investigation and analysis of case and reporting are being followed by every standard model among others.

Digital Forensic Investigation Model (DFIM) is a four-phase model that primarily aims to uncover hidden evidence in the data collected. Conversely, it is not concerned on actual evidence, i.e., physical evidence, which is unfavorable in IoT’s case. Various existing models of cyber forensic such as DFRW, Enhanced Digital Investigative Process Model (EDIPIM), Abstract Digital Forensics Model (ADFM), Digital Forensics Model for Digital Forensic Investigation (DFMDFI), Integrated Digital Investigation Model (IDIP), Xtended Model of Cybercrime Investigation (EMCI), Enhanced Systematic Digital Forensics Investigation Model (ESDFIM) and Systematic Digital Forensic Investigation Model (SDFIM). DFRW is established on seven phases (e.g., identification, preservation, collection, examination, analysis, presentation, and decision). ADFM has added three new components (e.g., preparation, approach strategy, and return of evidence), which were missing in DFRW. IDIP is a five-phase model (e.g., readiness, deployment, physical investigation of crime scene, digital investigation of crime scene, and review). EDIPM aims to enhance IDIP model by including two further steps (e.g.,
tracing, and dynamite). EMCI model contains thirteen steps which include awareness, authorization, identify evidence, planning, notification, collection and examination of evidences, logistics, storage, hypothesis, proof and presentation of hypothesis and finally storage archival. DFMDFI is a four layers’ iterative approach model. First tier handles preparation, authorization & identification, and communication; the second layer handles the rules related with stages collection, preservation, and documentation. Whereas, the third layer supports rules relating to the examination, exploratory testing, the analysis, and finally the fourth layer will be responsible for giving results, reviews and reports. SDFIM manages forensic process in eleven phases. Whereas, ESDFIM model deals the investigation procedures in a detailed six phase method [6].

Currently, IoT devices are exposed to major risks by malware. At present, more than 8 billion smart gadgets are linked worldwide. When the Mirai botnet struck 5 years ago, the world realized the dangers such equipment pose in the hands of hackers [7]. The IoT is being used in areas such as traffic management, automating homes, industrial process management, etc. [8]. Nowadays, typically, most IoT devices run on Linux due to ease of programming and availability (DD-WRT, 2019). Connected ‘things’, however, constantly emit streams of data as part of their communication in the network. More and more Linux-based IoT devices are being targeted by attackers. Attacks on IoT devices may be classified as follows: malware and file-less attacks [9]. In file-less intrusions, also known as non-malware intrusions on IoT devices, as against malware-based intrusions, downloading and running malware files for infecting the devices does not occur. They exploit the vulnerabilities existing in the victim’s devices and silently enter their network, making them more dangerous. In the recent past, progressively more file-less attacks have been reported [10].

There have been studies concerning the use of honeypots for defense against intrusions. For example, McCarthy et al. [11] were among the first to conduct analysis on integrating HP and honeynet methodologies. Pomsathit [12] proposed an HP system for ensuring security in communication in distributed networks. Li and Liu [13] proposed a model for automatically constructing signatures by using honeypots utilizing the data captured in them. Researchers da Silva Vargas and Kleinschmidt [14] developed a honeypot based system that detects malicious and unauthorized entry in VoIP networks. Selvaraj et al. [15] observed that the distributed denial of service (DDS) is effective, but this wastes bandwidth. Therefore, they proposed an advanced ant-based DDS method combined with honeypots for providing a solution for this wastage. Liu and Zhang [16] proposed a honeypot-based method for securely communicating data. Limited research exists on the features of attacking patterns in IoT systems, which is a necessity for preparing defense strategies. In addition, there is an urgent compulsion for a methodology for IoT forensic.

3. METHODOLOGY

Experiments were conducted to analyse IoT attacks on IP/Wi-Fi camera and SCA attack for getting AES key. The objective was to demonstrate intrusions by a hacker into unprotected Things of IoT connected to network. This experiment is the initial step for studying attack patterns on IoT things and deriving a method for forensics analysis. Experiment study revealed that devices are highly vulnerable and hackers use many methods for getting access to IoT network without much effort.

Next, a HoneyNetCloud was made using four hardware honeypot and 100 software honeypots on public cloud (AWS, Azure, Cloudways, Google cloud, Linode, Hostwinds, Vultr and Siteground) for attracting a wide span of hacks realistically globally. Figure 1 represents HoneyNetCloud architecture. It has three primary modules: Access Controller, Shell Interceptor cum Fidelity Maintainer and Inference Terminal. Access controller ensures control on the repeated attacks by intruders, Shell Interceptor intercepts the packets of data that flows in, Fidelity maintainers role is strategy implementation to prevent hackers in recognizing honeypots and finally the Inference Terminal intercepts the plain text to the hpfeeds. Attackers were attracted to IoT networks for hacking IoT devices. A vast range of information was logged for a full year, providing multiple insights for developing a process methodology for the IoT network forensics. Setup had run for a tenure of one year (02/10/2018 to 01/10/2019). Figure 2 portrays the attack locations spotted by HIM.
Captured data include attack time, host, source, protocols, packet type, source port, destination port, IP addresses, country code, source location, geo location and commands. Tools like Kibana, RStudio, Bulk extractor, Nmap, Snort and Wireshark were utilised to organise the data captured and analysed to categorise the attack types (Type I to VIII) and patterns/behaviours of attacks. DST (Dempster-Shafer evidence theory) algorithm was used for predicting the attack types [11]. The HoneyNetCloud architecture could stop the intrusions in to the n/w since the honeypots mimicked the real devices and attackers attempt went in vain by trying to intrude the fake IoT devices. This analysis brought in the conclusion of having a mechanism of bringing HoneyNetCloud setup in forensics methodology, as the organizations where IoT network is implemented can easily reduce the likelihoods of vulnerabilities in the real-time scenario.

On top of malware based attacks, HoneyNetCloud has also captured eight types of file-less attacks too. The classifications of types are based on behaviors and intentions of IoT attacks. An innovative process model which is preventive model for the IoT forensics termed HoneyNetCloud Investigation - HIM Model is proposed. It uses honeypots for attracting hackers and recording their behaviours and intentions in the first phase of HIM and the remaining of it provides hierarchical steps to the forensic analysts for carrying out the investigation. Incorporating the findings from the previous objectives, a methodology was developed baselining to the generic forensic process methodologies and its stages in a top down manner. Proposed model has been named as HIM – HoneyNetCloud Investigation Methodology for IoT Forensics. In the hierarchical model, the output of each process is a feed to its next process as the model itself is intrinsically a process oriented model. The structure consists of Class, Process, Stage, Sub-Stage and Parallel process classes as laid in the Figure 3.

4. DESIGN & DEVELOPMENT

Kohn [4], by making use of outcomes of the work by Bogen and Dampier [17], proposed the Unified Modelling Language (UML) for presenting the forensic procedure. Essentially, Ruan and Huebner [18] state that a conventional model is required to indicate the procedure utilized in the cyber forensic analysis, which can give very characterized and clear semantics to the advanced forensics procedure. Presentation of the UML system to the cyber forensics process gives formalism and an organized way to deal with advanced forensic procedures [19]. Considering these research works, the UML was considered as a proper method to display the whole procedure in a conventional way in the new model developed. The kind of UML diagram chosen for demonstrating the whole cyber forensic process with in the DFIM is the UML Activity Diagram under the Behavioural UML class. Building up the utilization of UML in a computerized crime scene investigation is another contribution of this research to the field of cyber forensics. The reasons for selecting the UML Activity Diagram over other
sorts of Behavioural UML Diagrams or Structural UML Diagrams are as follows: (i) The UML Activity Diagram can manage a wide range of a process’ flow controls and can model the parts of the whole digital forensic procedure and their relationships in an effective and clear manner, (ii) the Use case Diagrams do not benefit while describing process models would require customization for each scenario, thereby being less generic [19], (iii) the Activity Diagram gives an all-around characterized and unambiguous semantics to the cyber forensic procedure which is easily understandable to concerned authorities, e.g. a court of law [20, 21], (iv) the simplicity of a UML Activity Diagram enables CFIs to visualize the different parts of the forensic process more clearly, (v) being a type of stream graph, it is preferred in courts [19-21]. Since the UML displays and represents different parts of advanced criminological procedure, this research will use the UML Activity Diagrams in forensics to introduce both higher and lower level constituents in the DFIM.

For attaining the objectives of the research, following steps were accomplished. The kind of UML diagram chosen for demonstrating the whole cyber forensic process within investigation process oriented model is the UML Activities Diagram under the Behavioral UML class. Building up the utilization of UML in a computerized IoT crime forensic investigation is another contribution of the research to the fraternity of forensics. The processes are linked together by the process data stream along with the investigation conventions, e.g. case management & data flow. Hence, every processes and classes of the HIM are connected through the data flow. HIM model has been structured utilizing a top-down methodology so that to empower cyber-forensic experts to get a better idea of components, in particular, Classes, Processes, Stages, Sub-Stages, and Parallel processes [22]. There are five Classes in the HIM, each one of which contains various Processes. There are 24 Processes within the 5 classes. The HoneyNetCloud architecture has been applied in preparation class which will ensure capturing the attackers/hackers to the fake network, thus saving the real network. Figure 4 represents the overall architecture of HIM and portrays the model of forensic investigation with its components, various classes and processes.
5. PARALLEL CLASSES OF HIM

The HIM model has supporting processes that runs along with all the classes from top to bottom i.e. preparation class to the future ready class [23]. Eight parallel processes are designed and grouped together to make a class called parallel class which is an overriding class. These parallel classes run throughout the investigation timeframe assisting other layers of the model. Figure 5 illustrates the parallel classes of HIM.

![Figure 5. Parallel classes of HIM](image)

6. PROCESS FLOW

The essential elements needed for a digital forensic investigation model (DFIM) are summarized as follows. Figures 6 & 7 demonstrate the flowchart of processes.

1. Readiness: Organizations should be prepared in advance to deal with potential cybercrimes to get the evidences being available, should an investigation be required to detect and prosecute if violations have occurred.
2. Detection: The sooner the incidents are detected, the easier the root cause identified.
3. Preliminary Response: This includes the very first actions taken prior to the investigation of a detected crime.
4. Planning: Detailed action plan of investigation and its logistics need to be drawn.
5. Preparation: The stepwise action plans drafted in the planning stage.
6. Protecting the Crime Scene: This stage is a key point where the collection of suspected data and the protection of the evidences to be dealt with take place.
7. Identification: Identifying the issues that would have the digital evidence to help for the investigation.
8. Acquisition: This element consists of the actions to acquire the evidence, duplication of evidence for investigation, and verifying the collected data. It also includes the logistics and parking them in a safe storage so as for producing in court later.
9. Examination: This element consists of using various methods and tools by which the collected digital evidences are extracted and examined.

![Figure 6. Process flow of preparation, incident detection and primary response processes of HIM](image)
7. EVALUATION

The cyber forensics fraternity utilizes a model only when they can use it conveniently and easy to implement so that they can achieve the goal easily. So both “usability and capability” should be there in an effective model [23]. The evaluation process is basically defining the magnitude to which a model and associated data accurately represent reality from a viewpoint of the model’s intent [25].

7.1 Hypothetical case scenarios

This approach defines by applying two hypothetical case scenarios and then analysed using the model. Each scenario shows how HIM could be operative in the investigation. The scenarios are on real world situations intended to demonstrate the deployment of HIM.

Case 1 - United Condiments is a small-scale company located in Kerala which produces biscuits and confectionaries distributing in the local market. The company has an IoT controlled automated mixing machine, where the operator can control the required ingredients for each product according to the process requirement. An unknown attacker took control of the IoT machine and changed the recipe parameters increasing the measure of the salt to 3 times of original. The change in
parameters was unknown until the product reached customers and when they started complaining about it. The company had to refund the customers for the entire lot they made with the modified parameter and underwent a huge loss. The company has no incident response team and no forensic readiness process implemented. However, they have an in house IT department with a manager and two admin staff who take care of the IT systems. United Condiments suspects that an employee, Mr Sami, who left the company a few days before the incident occurred, had a hand in it. The company also suspects that Sami stole the confidential data of their unique products’ recipes and the price details before he left the job as he wanted to setup a similar company as a direct competitor. The company contacted their legal advisor regarding this case and the legal advisor in turn contacted a forensic service provider to scrutinize the issue. Investigation of this particular scenario does not appear to be very wide in scope since the involved components are only an IoT mixing device and one PC that runs the interface. Investigation begins at the planning stage as neither forensic preparation and readiness were done by the company, nor was there an in-house incident response team set up by the company. The forensic service provider appoints two investigators X and Y for dealing this case. Throughout the investigation, beginning with the planning process of inception class, they strictly adhere to the parallel classes of HIM.

**Case 2** - An email is received by Crescent Hospitals, a leading hospital group in Kerala, claiming that a serious vulnerability was found in their IoT automated operation theatre where remotely controlled surgeries by experts from Europe and United States are performed when a need arises. The email from the unknown source offers to divulge the details for a lump sum payment, else they threaten to hack the hospitals IoT systems and then damage the units.

The hospital has an IT department with technology experts handling various roles. Upon checking the logs of IoT devices monitoring server, the admin finds an unauthorised access into the server from outside. The hospital’s director, further, receives emails from the hacker threatening to reveal this vulnerability to the press and public, affecting reputation of the hospital group, if money asked is not paid. Further to this, the hospital group decided to report this to the national cybercrime investigation cell who initiated the investigation. The compromised server was located in Mumbai where the hospital has their head office. The source of email is also Mumbai, hence the investigation team appointed the Mumbai cyber cell to handle the case for further investigation.

### 7.2 Forensic lab evaluation

This approach is being done at the forensic lab. The objective of this approach is that to demonstrate the model HIM using various investigation techniques and tools (EnCase, FTK, zenMap, Nessus, DiscImage etc.) that answers questions who, when, where and how.

### 7.3 Expert review evaluation

Other than above evaluation methods, the model has also been sent to the experts for independent evaluation which was also analysed and feedback incorporated to fine tune the model. Two attributes “Usability and Capability” are considered for the evaluation of forensic model [26]. Key performance indicators of Usability are Clarity, Specification, Readiness, Planning and Reporting. The metrics of the attribute Capability are Detection time, Efficiency and Volume of attacks to Honeypots [27]. Attributes and its corresponding metrics are listed in the Table 1.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarity</td>
<td>Specification</td>
</tr>
<tr>
<td>Readiness/Preparation</td>
<td>Planning</td>
</tr>
<tr>
<td></td>
<td>Reporting</td>
</tr>
<tr>
<td></td>
<td>Detection Time</td>
</tr>
<tr>
<td></td>
<td>Efficiency</td>
</tr>
<tr>
<td>Volume of attacks to honeypots</td>
<td></td>
</tr>
</tbody>
</table>

A simulator forensic tool based on the model has been executed to generate the values of the parameters for the Capability. With the simulator, decent volume of IoT attacks were tried and capability attribute was recorded to generate the graphical representation of the outputs (Figures 8 & 9). In the expert review evaluation method, carefully generated questionnaire was provided to the forensic experts. The feedback received from the team has been incorporated to generate the graph of KPIs under Usability. The metrics values as recorded are listed in the Table 2 (a & b).

<table>
<thead>
<tr>
<th>Cases</th>
<th>Detection Time (Man Hours)</th>
<th>Month</th>
<th>IoT Attacks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HIM</td>
<td>IDPM</td>
<td>Jan-20</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>22.5</td>
<td>Feb-20</td>
</tr>
<tr>
<td>8</td>
<td>21.9</td>
<td>29.2</td>
<td>Mar-20</td>
</tr>
<tr>
<td>12</td>
<td>26.1</td>
<td>34.4</td>
<td>Apr-20</td>
</tr>
<tr>
<td>16</td>
<td>27.4</td>
<td>36.9</td>
<td>May-20</td>
</tr>
<tr>
<td>20</td>
<td>31.1</td>
<td>40.2</td>
<td>Jun-20</td>
</tr>
<tr>
<td>24</td>
<td>32</td>
<td>42.6</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>33.4</td>
<td>44.8</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>34.1</td>
<td>49.8</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 8.** Capability: - Detection of root cause

Many activities in the cyber space attacks lifecycle may go undetected. Cyber resiliency overtly considers intrusions and compromises of attack resources [28]. Such incidence may fail to get noticed and detected. Hence, performance metrics are
very much necessary but they are insufficient to make out systems cyber resiliency. Metrics are indeed needed for capability [29]. Survey method was advised and used to quantify the qualitative parameters under Usability (Figure 10). In the expert review evaluation method, carefully generated questionnaire was provided to the forensic experts. The feedback received from the team has been incorporated to generate the graph of KPIs under Usability.

![Figure 9. Capability: - IoT attacks in HoneyPot](image)

**Figure 9.** Capability: - IoT attacks in HoneyPot

![Figure 10. Usability parameters of proposed model vs existing generic model](image)

**Figure 10.** Usability parameters of proposed model vs existing generic model

Proposed model HIM has been assessed and the results shows a usable model for IoT forensics compared to the generic model of forensic.

8. CONCLUSIONS

Cyber-attacks are generally dangerous, however IoT attacks will become significantly dangerous since it gives more exterior to attack. In empirical literature survey done to study the existing forensic models that are currently being used by fraternity of forensics, it was being identified that, a reliable, usable and capable model for the IoT forensics was lacking. With the purpose to learn the patterns of various attacks, an infrastructure environment was implemented and it was entitled HoneyNetCloud consisting of hundreds of software honeypots on various public clouds and four hardware based honeypots. DST based algorithm was designed for detection of attacks and finally a very efficient IoT forensic model named HIM – Honey Net Cloud Investigation Model with a newly setup infrastructure HoneyNetCloud was designed. For developing the HIM model, wide span of IoT attacks was extracted for a span of one full year until October 2019. The massive data extracted were analysed elaborately. Massive range of numerous attack types were captured with diverse profiles, behaviors, influences and characteristics. The proposed model HIM has been evaluated using hypothetical case scenario method, forensic lab method and expert review process and found to satisfy the research objectives.

9. FUTURE WORK

Though, the model HIM was evaluated using various methods and found to satisfy the research objectives, there could be some scope of enhancement in the real time use, which could be considered in later stage. As and when the 5G comes to reality, the face of IoT devices would change and so the connectivity to internet. This can be a potential area to study in future as DDoS attacks would be on the rise. Since the research work is done in the Indian jurisdiction, Indian forensic investigation concepts are considered. This can also be extended to the other countries forensic investigations too.

REFERENCES


