International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-1, October, 2019

Formalization in Digital Forensic Triage for Identification of Malicious IoT Devices

MohammedIbrahim, MuhammedBasheerJasser, Mohd TaufikAbdullah, Azizol Abdullah

Abstract: Considering the large number of devices connected to the Internet of Things(IoT), identifying malicious devices for the purpose of "search & seizure" remainsa critical issue for digital investigators. Consequently, the need for techniques that automatically identify malicious devices can speed up the process of digital investigation. However, few conceptual approaches were proposed to identify malicious devices during IoT forensic investigation. To overcome that, a formal approach is proposed to automatically triage and fingerprint IoT Malicious devices with their respective states. It is expected that with the proposed formal approach, investigators can simply identify malicious devices, their states as well as determine the scope of investigation.

Keywords: Formalization, IoT Forensic, Sensor, Triage

I.INTRODUCTION

Sequel to the inter connection of embedded digital devices including smart objects, a new revolution of internet service emerged, known as Internet of Things(IoT). IoT is capable of integrating all gadgets use in our daily basis, thereby increases the size of attack vector horizon and enables cyber criminals to launch various types of attack(Perumal et al., 2015). Internet of Things possesses various challenges to digital forensic investigation due to the size of the objects of forensic interest. The size of data for digital investigation keeps perpetually increasing. This is due to the effect of the increasing technology evolution when scale and bounds of the Internet quickly change to everyday use(Jusas et al., 2017). The ever increasing connectivity of devices in IoT resulted to novel types of cyber-physical evidentiary data(Meffert et al., 2017). Nonetheless, how to retrieve forensically useful data and how to examine it from various IoT devices without a common interface, internal storage or standard protocols is a challenge (Meffert et al., 2017). On the other hand, finding the existence of an IoT devices during digital investigations process remains a critical issue in IoT forensic, thereby limits the recognition of a particular user's data. This brings the question of how to perform what law enforcement term "search & seizure" when it is not obvious where the investigated data is being saved, or where its originatedfrom(Hegarty et al., 2014).

Revised Manuscript Received on October 30, 2019.

* Correspondence Author

Mohammed Ibrahim, Faculty of Computer Science, University Putra Malaysia,43400 UPM Serdang, Selangor, Malaysia

Muhammed Basheer Jasser, Faculty of Computer Science, University Putra Malaysia,43400 UPM Serdang, Selangor, Malaysia

Mohd Taufik Abdullah, Faculty of Computer Science, University Putra Malaysia,43400 UPM Serdang, Selangor, Malaysia

Azizol Abdullah, Faculty of Computer Science, University Putra Malaysia,43400 UPM Serdang, Selangor, Malaysia

Also, because of the large quantity of data aggregated from the cyber-crime scene, conducting digital investigation process is becoming cumbersome and time consuming. In this regard, the dynamical nature of life, and the circumstancesmake it necessary that cyber-crime information has to be acquired as quickly as possible without being restrictedwith the prolong legal formalism and make it necessary that certain information needs to be obtained. However, information acquired in a such manner cannot be straightaway utilized in the court; nevertheless, timely acquiring such information can fast up the upcoming process of digital investigation and in some circumstance someone's life can be protected as well with this information(Jusas et al., 2017).

In this regard, with digital triage, valuable intelligence information about a crime is obtainable without processing digital evidences and without a complete examination process. This agile intelligent approach can be utilized at the point of investigation to guide the search and seizure,and in the laboratory to ascertain if a media is valuable to be analyzed. Recently, the need for visualization for IoT states and exploring practical methods of automatically fingerprinting and triage IoT devices on a network under investigation is emphasized(Meffert et al., 2017). However, almost no research works of digital triage have connected to the huge area of social networks and the growing area of the Internet of Things is left virtually without consideration(Jusas et al., 2017). The objective of this work is to propose a formal approach to digital forensic triage that can automatically fingerprint and triage IoT devices under investigation.

II.RELATED WORK

In an effort to identify and reduce IoT infrastructural complexity under investigation, (Oriwoh et al., 2013) conceptualized IoT forensic into 1-2-3 zones. These zones can guide forensic examiners on where to commence work in preparation of digital investigation. Zone1 of the model focuses on identifying device under the internal network and application, Zone2 which is the middle zone concerns about the gateway and border networks and finally zone 3 is aimed at focuses on forensic at cloud level. In addition to 1-2-3 zones, a next best thing triage model was introduced in order to identify object of forensic interest (OOFI). However, techniques that determine what this next best thing will be in any situations and any scenarios remain a challenge for future research(Oriwoh et al., 2013).

Similarly, (Oriwoh & Sant, 2013) proposed a forensic concept titled Forensic Edge Management System(FEMS). The concept aimed at developing an automatic security and forensic services for a Home IoT system. FEMS concept was built from three layers; the perception, network and application layers. The perception layer collects data from the sensors and transmits them to the application layer via the network layer. FEMS records data and save them for a specific period of time. The recording is determined by an event that passes through a predefined set of threshold. However, their work has some degree of application into our work but was not specifically aimed to identify IoT devices and their states for the purpose of "search and seizure".

In another perspective, (Zawoad &Hasan, 2015) defined IoT Forensic "as a special branch of digital forensics, where the identification, collection, organization, and presentation processes deal with the IoT infrastructure to establish the facts about a criminal incident". Consequently, they categorized IoT forensic into three digital forensic schemes: device level, network leveland cloud level forensics. In addition to that, a Forensic-Aware IoT (FAIoT) was proposed to ease the process of evidence collection and analysis. With FAIoT, a centralized evidence collection point is provided allowing for ease of access and organization with respect to the collected evidence.

(Meffert et al., 2017)argued that by knowing the state changes of an IoT device, it provides a practical solution to the IoT forensic challenges. In this regard, a Forensic State Acquisition from Internet of Things(FSAIoT) was proposed. In FSAIoT, state data is gathered from IoT devices through a centralized controller titled Forensic State Acquisition Controller(FSAC). The FSAC has three modalities of operations as follows: controller to IoT device, Controller to cloud and Controller to control. These modalities of operations enable the state acquisition of multitude IoT devices. However, despite its insightful and practical approach to IoT State acquisition, it has highlighted numerous limitations among which is forensic soundness. Some devices states are found to be modified during implementation using openHAB controller.

III.IoT FORENSIC CHALLENGES

Advancement and rapid adoption of IoT devices will attract more criminal mindset that can adversely exploit the weakness of IoT devices for their criminal purposes. However, researchers such as (Hegarty et al., 2014) and (Zawoad & Hasan, 2015) have outlined various challenge with regards to IoT forensic.

According to (Zawoad & Hasan, 2015) "The current tools and processes of digital forensics cannot stand the highly distributed and heterogeneous infrastructure of the IoT. Forensics investigators will face difficulties while identifying essential pieces of evidence from the IoT environment, as well capturing and examining those evidence" (Zawoad & Hasan, 2015). Also, the volatile nature of evidence in the IoT is much more difficult than in the usual forensics ; data may be saved locally by a thing; and the lifetime of the data before it is replaced or compressed using a lossy technique is finite(Hegarty et al., 2014).

Similarly, the nature of devices undergoing examinations in IoT complicates IoT investigations matters further. This is consequence to the possibility of replacing/compressing of data at a crime scene if there will be no connection between the IoT devices with a cloud service provider to save their data, and they aggregate large data greater than what they can hold. Consequently, this creates a challenge for a first responder, who must take a decision of whether to conserve the evidence on the devices or to enable data transfer from the scene and then visage the dispute of an inter-jurisdiction evidence aggregation process(Hegarty et al., 2014).

Another challenge lies in the recovering of the existence of an IoT system to digital forensic investigation and in finding a specific user's data. This brings the problem of how to perform what law enforcement term "search & seizure" since it is unclear where the investigated data is being saved, or originated from (Hegarty et al., 2014). However, we argue that with a systematic approach,which can fingerprint IoT devices and their respective states as well as their locations, can guide law enforcement agents where to carryout "search & seizure". Consequently, investigators will be provided with a mechanism that enablesthem to serve "digital warrant" that can entails the scope of evidence to be seize and examine.

In an effort to resolve the aforementioned obstacle, a formal approach is proposed to automatically fingerprint and triage IoT device under investigation as well as identifying their states for the purpose of "search and seizure". Consequently, it can speed up the process of digital investigation of IoT based environment.

IV.THE FORMAL APPROACH

Rewrite Logic

In this work, we adopt rewrite logic of (Martí-Oliet & Meseguer, 1996) to formally build a model of IoT system and digital triage capability that can automatically fingerprint out the respective IoT devices, states and locations under investigation.

Rewrite logic enables us to define the functions of a system and its states changes rules that transform the device from one state to the next. Based on rewrite logic, we can ascertain the behavior of the device and determine which states of IoT devices are of forensic interest. Upon building the digital triage, we can query the transitional path to determine the correctness of the system.

According to formal theory, a rewrite theory is defined as a four tuple$\mathfrak{R}=(\sum, E, \emptyset, R)$, where$\sum$ is the set of functions (operation). E is the set of all equations in $\mathfrak{R}$ which can be used to define all functions (operations) in $\mathfrak{R}$. R is the set of labeled rewriting rules and $\emptyset$ is a function assigning to each operator. Rewriting rules prescribe state changes in the device and define explicitly the way the system evolves. Also rewrite rules can be used concurrently to determine the system behavior. Consequently, this will enable us to capture the system behavior along with the triage to determine which IoT device is of forensic interest.

In this regard, we will apply the formal model of the IoT

Information Model in rewrite logic using Maude(Martí-Oliet & Meseguer, 1996). Maude is a model checker that supports rewriting logic and provides the user with platform for executing rewriting logic rules and formally verifying them(Martí-Oliet & Meseguer, 1996).

Maude is capable of performing model checking of the system with respect to invariants.Maude executes state search using breadth first search, beginning from the initial state *s* and subsequently applying a rule to move to the next state. Consequently, all possible model transitionsare checked, which is referred to as explicitly model checking. Rewrite logic has been used in security analysis of IoT device(Tabrizi& Pattabiraman, 2016).

The Formalization Technique

This section presents a formal approach for fingerprinting IoT devices under investigation. A set of basic actions of triage,which can capture various IoT device(s) state changes, is proposed.Model checking is applied to search automatically all the possible scenarios by implementing those actions on the model of IoT devices. The search finds the device state scenarios and guarantees to find all the evidentiary data with respect to the state space of our model (i.e IoT Device and the triage).To build our proposed model, the following steps are followed:

Step 1: We start by formalizing the IoT system and its operations. This can be achieved through the IoT Information Model (Carre et al., 2013). The model defines the operations with regards to IoT devices. IoT Information Model is built on three main aspects: virtual entity, service and Resource descriptions. In this paper, we express the abstraction of IoT Information Model formally in rewriting logic.

Step 2: We define the set of capability actions of digital forensic triage in rewriting logic. Modeling both the IoT devices and digital triage capableness in rewriting logic enables us to automatically search possible scenarios in which the digital triage's actions on IoT devices can detect the IoT device of forensic interest. An example of IoT devices of forensic interest is a device state resulted from an alarm service description.

Step 3: We comprise the model of the IoT devices with that of digital triage concurrently. Subjecting them to model checking, our model searches through all the execution paths of the model, which will result in the identification of a particular device under investigation.

The Formal Model

To develop the formal model, IoT Information Model (Carrez et al., 2013) is to be usedin theformalization of IoT devicesvia rewriting logic. The Information Model defines the operations of each attribute of an entity with respect to its associated services and resources. Using the Information Model, we explore the functionality of execution paths of the components of the IoT devices and formalize them into rewrite logic. The following are the major operations of IoT devices based on the IoT information model.

To simplify our model, only the operationsthat arerelevant to our approach are considered. Devices are represented in IoT Information model using *Virtual Entity*. This represents physical entities in the digital world. Every *virtual entity* needs to have unique attributes. For every attribute specified in the *virtual entity*, associated services are defined to provide information about the attribute (sensing). Sensors provide information, knowledge, or data about the physical entity they monitor. Information from sensors can be recorded for later retrieval from on-device resources(device(s) that stores data on its memory). Furthermore on-device resources specify the geographical location of the resource. To develop the formal model based on the Figure 1 above, the following operations are introduced:

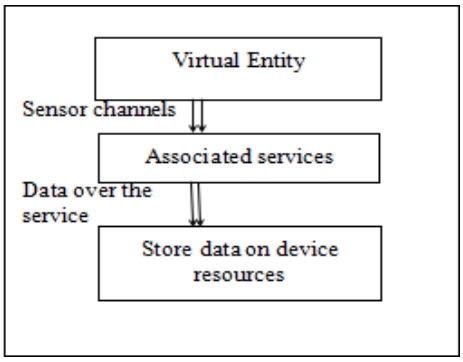

Fig. 1 Communication paths in IoT Information Model.

Passing Virtual entity attribute to the Sensor.

A Virtual Entity has many attributes and behaviors. Theseattributes and behaviors can be read and monitored via sensor channels. Data about an entity from the sensor can then be stored to the On-device resources via associated services as shown in Figure.1 above. Sensors monitor and produce information about the identity of the device and its states changes. Open request associated services that act as an interface between sensor and on device resources collect those sensor data and store them on On-device resources for later retrieval. On-device resources not only store sensor information, but they also provide information about the location of device and state-tracking information.

To formalize the sensor operations, the sensor data of a given device is defined as the collection of device attributes from a sensors. For any given device, a sensor *s* produce an attribute of a device *a*, the collection of sensors and the device attributes is called *sensorComp*,given as a tuple <*s*, *a*>, where *s* is a particular sensor from a collection of sensors in an IoT environment. Then the collection of a sequence <*s*, *a*> from a sequence of sensors is formally referred to as *sensorSeq*. Then in rewrite logic a sensor *s*, an attribute *a*, *sensorComp* and *sensorSeq* are all referred to as data type and can be defined using a keyword *sort*. However, due to dynamical nature of IoT environment, devices can be added at any point in time, then an operation *createSensor* has to be added that will recursively check additional device to be added to the sequence of sensor (i.e*sensorSeq*). If no additional device is being added, only sensor data from existing devices can be read and save.

Sending and storing Sensor Data To the On-Device Resource.

Following the generation ofsensor data from the device,

the next step is to send and store them on On-device resources. This can be achieved through an interface known as associated services such as open request from On-device resources defined by *getAsensorData* operation, which requests a sensor data from the sensor. If there is any data from the sensor, the associated service wait to receive the data otherwise timeout for non-response from the sensor. Since we are interested in identifying object of forensic interest, we define an operation *hasanalarm*. This operation will check among the sequence of *sensorComp* (i.e*sensorSeq*), which *sensorComp* has raised an alarm to be distinguished from other *sensorComp*.Therefore, for any data collected from a sensor, a test of alarm using the associated service is performed to determine which sensor data is of forensic interest.

Given a received sensor data, the sensor data will then be stored in On-device resources. To store the data in On-device resources, we convert the *sensorComp*, which is a tuple <*s*, *a*>of a sensor and the device attribute to a new tuple known as *saveSensorComp*. A collection of *saveSensorComp* will be converted to *saveSensorSeq*. On the other hand, On-device resources have two other management parts, fault detection and state tracking. For forensic interest, once the operation *hasanalarm* raises against *sensorComp*, then the *saveSensorComp*, and *saveSensorSeq* will be savedon fault detection component. Otherwise, *saveSensorComp*, and *saveSensorSeq*can be saved on state tracking component.

Triaging and fingerprinting IoT devices under Investigation.

Given a sensor data saved on anOn-device resource, to access the sensor data, a triage system is to be modeled, which requests a stored data from an On-device resources defined by two operations *getAsaveSensorComp* and *getAsaveSensorSeq*. These operations generate all the sensor data and categorize them according to fault and state tracking data.

V.EXPECTED RESULTS AND RESEARCH SIGNIFICANCE

It is expected that our formal model will identify IoT devices under investigation and enable investigators to performthe process of digital forensic investigationfaster by isolating specific devices for forensic examinations. Similarly, it will assist in defining where search and seizure can be performed among large number of devices in an IoT environment. Consequently, it will guide law enforcement agents in issuance of digital warrant.

VI.CONCLUSION

Considering the number of devices connected to Internet of Things(IoT), Identifying malicious devices for the purpose of “search & seizure” remainsa critical issue for digital investigators due to the volume of devices concern. However, we argue that with a systematic approach that can fingerprint IoT devices and their respective states, it can guide law enforcement agents where to carryout “search & seizure”. We propose a formal approach to digital forensic triage that can automatically fingerprint and triage IoT devices under investigation. The formalization is built from IoT Information Model to avoid machine dependency.

Furthermore, as this is the starting point of the work, future work is required to introducea formal specification and to develop a tool. Also, IoT devices that store resources on a network can be formalized for digital forensic triage.

REFERENCES

1. Carrez, F., Bauer, M., Boussard, M., & Bui, N. (2013). Final architectural reference model for the IoT v3. 0. EC FP7 IoT-A Deliverable, 1.
2. Hegarty, R., Lamb, D. J., & Attwood, A. (2014). Digital Evidence Challenges in the Internet of Things. Paper presented at the INC.
3. Jusas, V., Birvinskas, D., & Gahramanov, E. (2017). Methods and tools of digital triage in forensic context: Survey and future directions. Symmetry, 9(4), 49.
4. Martí-Oliet, N., & Meseguer, J. (1996). Rewriting logic as a logical and semantic framework. Electronic Notes in Theoretical Computer Science, 4, 190-225.
5. Meffert, C., Clark, D., Baggili, I., & Breitinger, F. (2017). Forensic State Acquisition from Internet of Things (FSAIoT): A general framework and practical approach for IoT forensics through IoT device state acquisition. Paper presented at the Proceedings of the 12th International Conference on Availability, Reliability and Security.
6. Oriwoh, E., Jazani, D., Epiphaniou, G., & Sant, P. (2013). Internet of things forensics: Challenges and approaches. Paper presented at the Collaborative Computing: Networking, Applications and Worksharing (Collaboratecom), 2013 9th International Conference Conference on.
7. Oriwoh, E., & Sant, P. (2013). The forensics edge management system: A concept and design. Paper presented at the Ubiquitous Intelligence and Computing, 2013 IEEE 10th International Conference on and 10th International Conference on Autonomic and Trusted Computing (UIC/ATC).
8. Perumal, S., Norwawi, N. M., & Raman, V. (2015). Internet of Things (IoT) digital forensic investigation model: Top-down forensic approach methodology. Paper presented at the Digital Information Processing and Communications (ICDIPC), 2015 Fifth International Conference on.
9. Tabrizi, F. M., & Pattabiraman, K. (2016). Formal security analysis of smart embedded systems. Paper presented at the Proceedings of the 32nd Annual Conference on Computer Security Applications.
10. Zawoad, S., & Hasan, R. (2015). Faiot: Towards building a forensics aware eco system for the internet of things. Paper presented at the 2015 IEEE International Conference on Services Computing (SCC).