Research Roadmap for IoT Forensics

Manas Kumar Yogi[1], K.Mahesh [2]

[1,2]Dept. of CSE, Pragati Engineering College (Autonomous), Surampalem, A.P. (India)

ABSTRACT

With the advent of IOT ,most of modern technological devices consume and dissipate data at a level never imagined .It has been admitted by IOT researchers that usage of data trails can be used to provide evidence in court of law for safeguarding the common man's interests. our paper is a sincere effort to throw light on current principles applied for IOT forensics as well as difficulties faced by researchers in this field. This paper is a readymade guide for understanding the crucial research challenges lying ahead in the field of IoT forensics. In this paper we present the issues pertaining to IOT forensics which force us to think about leveraging the current techniques used in digital forensics, cloud forensics, network forensics.

Keywords : IoT, Digital Forensics, Cloud Forensics , anonymisation techniques, Digital Investigations

I. INTRODUCTION

What is IoT

Imagine a world in which every device in the home, workplace and car are connected. A world where the lights automatically turn on when the car approaches the driveway, the coffee starts brewing when the morning alarm goes off and the front door automatically unlocks when approached by a member of the household, but stays locked when a stranger arrives on the front step. That is the type of world the Internet of Things can create.

The Internet of things (IoT) is the inter-networking of physical devices, vehicles (also referred to as “connected devices” and “smart devices”), buildings, and other items—embedded with electronics, software, sensors, actuators, and network connectivity that enable these objects to collect and exchange data.

Internet of Things (IoT) is an ecosystem of connected physical objects that are accessible through the internet. The ‘thing’ in IoT could be a person with a heart monitor or an automobile with built-in-sensors, i.e. objects that have been assigned an IP address and have the ability to collect and transfer data over a network without manual assistance or intervention. The embedded technology in the objects helps them to interact with internal states or the external environment, which in turn affects the decisions taken.

IoT is to offers advanced connectivity of devices, systems, and services that goes beyond machine-to-machine (M2M) communications and covers a variety of protocols, domains, and applications. The interconnection of these embedded devices (including smart objects), is expected to usher in automation in nearly all fields, while also enabling advanced applications like a smart grid, and expanding to areas such as smart cities.

Currently, the focus in the IoT domain centers on its benefits and applications as well as security and privacy issues that apply. There is little by way of a dedicated incident response methodology for Digital Forensics (DF) responders within the IoT domain. This gap is what this paper aims to fill: to propose a high-level incident response strategy for approaching IoT-based crime scenarios

Importance of IoT

1. In total, there will be 34 billion devices connected to the internet by 2020, up from 10 billion in 2015. IoT devices will account for 24 billion, while traditional computing devices (e.g. smartphones, tablets, smartwatches, etc.) will comprise 10 billion.

2. Nearly $6 trillion will be spent on IoT solutions over the next five years.

3. Businesses will be the top adopter of IoT solutions. They see three ways the IoT can improve their bottom line by 1) lowering operating costs; 2) increasing productivity; and 3) expanding to new markets or developing new product offerings.

4. Governments are focused on increasing productivity, decreasing costs, and improving their citizens' quality of life. We believe they will be the second-largest adopters of IoT ecosystems.

5. Consumers will lag behind businesses and governments in IoT adoption. Still, they will purchase a massive number of devices and invest a significant amount of money in IoT ecosystems.

6. Although the specific predictions and the numbers differ, what is remarkable is that the numbers predicted for 2020 have been consistently, extremely large over the years," Petschow says.

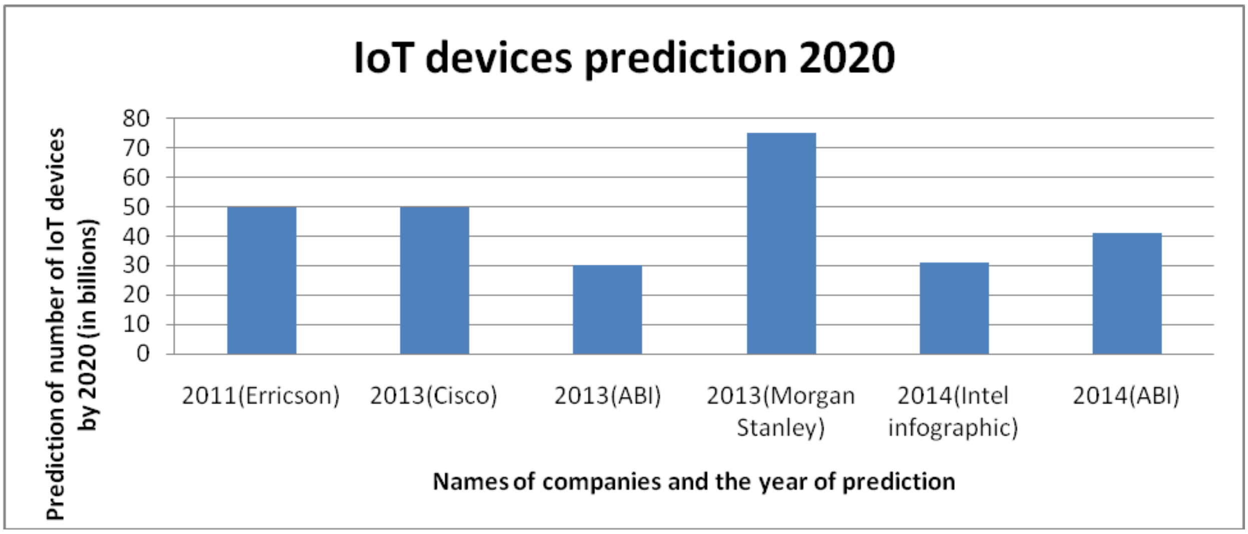

Fig.1.Evolution Of IoT Devices

II. IOT FORENSICS

IoT forensics involves numerous application domains apart from traditional forensics. these include wired, Wi-Fi, wireless and mobile network. Also, IoT comprises of the RFID sensor network. Various IoTware like appliances, tags and medical devices can be regarded as sources of evidence during investigation as well.

The most important challenge in investigating an IoT crime is due to the dynamic nature of IoT solutions. IoT is a combination of many major technology areas, which includes cloud computing, mobile devices, computers and tablets, sensors and RFID technologies. Consequently, forensics for IoT will engulf all of these domains.

IoT forensics can be classified based on the storage of data produced by the IoT devices.

2.1 . IoT Forensics with Local storage of data (Digital Forensics)

Digital forensics, also known as digital forensic science, is a branch of forensic science encompassing the recovery and investigation of material found in digital devices, often in relation to computer crime.

It involves investigation of electronic devices to find data that can be used to solve a digital crime. With the proliferation of Internet of Things (IoT), a new field in digital forensics has emerged. It is called IoT Forensics.

IoT forensics is a branch of digital forensics which deals with IoT related crimes and includes investigation of the connected devices, the sensors as well as the cloud of data.

2.2. Cloud Forensics

Cloud forensics is a cross discipline of cloud computing and digital forensics. Cloud computing is a shared collection of configurable net- worked resources (e.g., networks, servers, storage, applications and ser- vices) that can be reconfigured quickly with minimal effort . Digital forensics is the application of computer science principles to recover electronic evidence for presentation in a court of law.

Cloud forensics is a subset of network forensics. Network forensics deals with forensic investigations of networks. Cloud computing is based on broad network access. Therefore, cloud forensics follows the main phases of network forensics with techniques tailored to cloud computing environments.

IoT devices generate enormous amount of data. Cloud when combined with IoT is economical and efficient way to store the data.

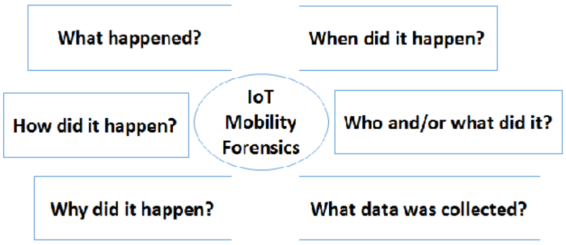

Fig.2. Guiding Principles In IoT Forensics

III. IOT FORENSIC CHALLENGES AND APPROACHES

The IoT will undoubtedly provide a richer source of evidence from the physical world than conventional computer systems. The way in which IoT is realising Zelkha vision of ambient intelligence (Zelkha, Epstein, Birrell, & Dodswoth, 1998) means that environments are beginning to react to the user's requirements, without the need for conscious interaction by the user. As a result, IoT environments are likely to contain contextual evidence of which the perpetrators are simply oblivious. This paradigm shift means that digital investigations will increasingly encounter evidence from events taking place in the physical world

The four main phases of digital forensics investigation from figure 1 face a number of challenges from the IoT. We discuss the implications of the IoT for each phase under the headings below and identify areas in which solutions should be targeted.

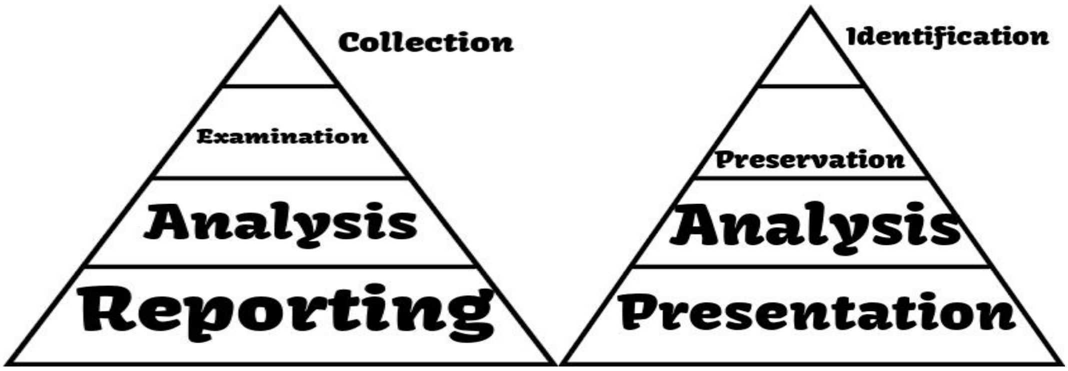

Fig.3. Process Models Deployed To Develop IoT Forensics Tools

3.1. Identification

Detecting the presence of IoT systems poses challenges to digital forensic investigations, as does the identification of a particular user's data. This raises the question of how to carry out what law enforcement term "search & seizure" when it is not apparent where the data being investigated is being stored, or where the data came from.

A potential solution to identification of data may be the integration of IoT device data into Building Information Modelling ("National BIM Standards," 2013);

Building Information Modelling (BIM) is a digital representation of physical and functional characteristics of a facility. A BIM is a shared knowledge resource for information about a facility forming a reliable basis for decisions during its lifecycle; defined as existing from earliest conception to demolition.

By combining the information about the IoT capabilities of a building or structure, it may be possible to answer the questions of; where has the information come from? Where is the information stored? It is also crucial to identify in what format the data is stored or encoded. This would narrow the scope of the investigation, and enable the selection of features or data that identifies an individual user from a much smaller data set. A composite picture of the data gathered about an individual user could be constructed from the data stored or forwarded by the buildings they have inhabited

3.2. Preservation

There are established procedures in place to capture volatile evidence before it becomes unavailable, for example first responders can create memory dumps prior to a machine being shut down. Evidence volatility in the IoT is much more complex; data may be stored locally by a thing, in which case the lifespan of the data before it is overwritten or compressed using a lossy technique is finite. The data from a thing may be transferred and consumed by another thing or a local ad-hoc network of things, alternatively it may transferred to the cloud for aggregation and processing.

The transfer and aggregation of data/evidence presents a challenge when securing the chain of evidence. In order to overcome this challenge and leverage the resilient nature of data in IoT in digital investigations, techniques are required to track and filter the transit of data across an IoT environment. Such techniques will facilitate the identification and extraction of data assumed to have been modified or deleted due to the constraints of IoT devices.

Preservation of the scene is a contentious issue in digital forensics. IoT investigations will complicate matters further due to the nature of the devices undergoing analysis. It is possible that data at a crime scene will be overwritten/compressed if the devices cannot interact with a cloud service provider to store their data, and they collect more data than they can store. This presents a problem for first responders, who must decide whether to preserve the evidence on the devices by allowing data transfer from the scene and then face the challenges of an inter-jurisdiction evidence collection process. Alternatively, they may sever the connection between the devices and the cloud and attempt local extraction of evidence from devices that may be of a proprietary nature. However, the physical placement, power availability or connectivity of each device may render this approach impractical. This also raises the question of whether – and how – a first responder or investigator should prevent devices recording information once a scene has been secured. Principle two of the ACPO guide indicates that a person may access the original data during in an investigation if they are capable of explaining the relevance and implications of doing so. Further research is required to determine what the implications are under a variety of circumstances. Ideally, a mechanism should be in place to enable an investigator to serve a "digital warrant" that prevents evidence being compromised.

We consider the interplay between the legal and technical challenges associated with gathering evidence from an IoT environment. A warrant is served during conventional investigations as the first part of the evidence preservation process. The warrant details the scope of evidence to be seized and examined. In the case of the IoT service providers, they often store data on behalf of their users. This means that individuals may not have direct access to their own data, or it may be presented to them in a different format than that in which it is stored. This complicates the preservation process, as the warrant may have to be served to individuals and their service providers.

3.3. Analysis

Analysis of data from an IoT environment will have to consider the provenance of evidence in order to demonstrate the evidence is reliable and authentic. Data provenance in the IoT differs from conventional digital forensic investigations in which the temporal dimension is often the main consideration e.g. file modified, accessed, and created time, email time lines .The interaction between IoT and cloud computing facilitates the aggregation and processing of data from the IoT. The vast quantities of data generated by IoT and stored in large-scale distributed cloud environments is likely to be the subject of a cloud investigation. From a technical perspective the image, analyze present paradigm of current digital forensics practice does not map well onto the IoT domain. This is aside from the ethical issues of imaging these devices in multitenancy cloud environment, There are a number of technical barriers; IoT data is either stored on proprietary devices that are difficult to interface with or in cloud computing platforms where the scale, distribution and remote nature of the data preclude imaging as a viable extraction process. Distributed analysis techniques are required to analyze the data stored in cloud computing platforms. Some work has already been carried out in this area to tackle the challenges posed by cloud computing investigations.

3.4. Presentation

Presenting the findings of IoT investigations poses a new challenge; data will often have undergone aggregation and processing using analytic functions that can alter the structure and meaning of data. At the device level, lossy compression techniques may reduce the granularity of the data order in to preserve limited resources such as memory, battery life, network bandwidth, etc. The granularity and semantics of evidence from the IoT will create challenges to digital forensic investigations. For example, one system may store temperature ranging from 0-5 as cold, 6-10 as average, 11-16 as warm and 16+ as hot. Another system may use different figures and describe the same temperature readings using different terminology resulting in a semantic gap. Ontological descriptors and standardization of metadata has limited adoption, with a view to moving IoT devices towards a semantic sensor web. However, from a forensics perspective, the issue is that devices may adopt differing descriptor formats or may retain a proprietary format. Presentation poses a challenge regardless of the underlying format of the data, as the conflicting grammar describing data from IoT systems has the potential to be misleading.

IV. KEY ISSUES AND EMERGING REQUIREMENTS

The emergence of the IoT will present new opportunities for data to be misused and lead to an expansion and development of new digital forensic techniques. We identify new approaches that may emerge out of the necessity to analyze the IoT.

4.1. Preservation Issues

Firstly we consider the preservation of evidence, forensic readiness is an area that is relatively well understood in conventional computing environments. We agree according to researchers who state that an alternative approach is required to enable IoT forensic environments to achieve the same. As we suggested in Section 3 a digital warrant would assist in the gathering of evidence. This approach could be extended to "digital preservation orders" that prevent evidence from be contaminated or overwritten by reducing the resolution at which data is captured by devices, or freezing the data stored by service providers. The warrant would be digitally signed by the serving authority to enable the providers or devices to check the authenticity of the warrant. The providers or devices would them submit the requested information to the authority over a standard set of interfaces.

4.2. Aggregation Issues

The financial motivation behind many IoT systems is the value that comes from the data aggregated in the providers' systems. Such data sets are valuable marketing commodities. Future investigations may benefit from such data to provide or substantiate evidence about an individual or sequence of events. However, to consider briefly an opposing standpoint; aggregated data may breach data privacy legislation, with the holders of data inferring information about individuals that breaches legislation. New techniques are required to reason over data and determine what can be inferred from large data sets, likewise techniques are required to investigate cases where "aggregation offences" are alleged to have taken place. Similarly, investigatory techniques are required to analyze cases where anonymisation techniques were inadequate, or rendered so by joint analysis of many data sets. Legal frameworks must be updated alongside the development of these techniques to ensure that the data gathered by the IoT is not misused. While the aggregation of data provides the possibility of inferring useful information about an individual, it also introduces some challenges such as the semantic gap discussed in Section 3. One approach to tackling this challenge is the development of digital forensic tools that can bridge the semantic gap. These tools would enable calculation and comparison of the granularity of data from different sources. This approach would be particularly useful when conflicting evidence emerges from different IoT devices or service providers. It may be possible to resolve semantic conflicts and even use characteristics of the measurements taken by different systems to provide evidence that is more accurate.

V. CONCLUSION

This paper has put together the intrinsic elements involved in meeting the research challenges in developing the field of IOT forensics. We have presented the required approaches for handling the current issues faced in this domain followed by the need to analyze the behavior of data flow in a IOT architecture. We conclude with the fact that given the current scenario and tools available to investigate the cases of offence occurring in IOT environment, considerable work has to be done so as to bridge the semantic gap. Our paper methodologically develops a roadmap for IOT forensics researchers to work in a direction which will be fruitful for the IOT user community.

REFERENCES

1. Burd, S. D., Jones, D. E., & Seazzu, a F. (2011). Bridging Differences in Digital Forensics for Law Enforcement and National Security. In 2011 44th Hawaii International Conference on System Sciences (pp. 1–6).
2. Grobler, C. P., Louwrens, C. P., & von Solms, S. H. (2010). A Multi-component View of Digital Forensics. In 2010 International Conference on Availability, Reliability and Security (pp. 647–652). IEEE. doi:10.1109/ARES.2010.61
3. Haines, L. (2007). Cops may check crash drivers' mobile records • The Register. The Register. Retrieved from http://www.theregister.co.uk/2007/02/27/mobile_phone_proposal/
4. Hegarty, R., Merabti, M., Shi, Q., & Askwith, R. (2012). Scalable Distributed Signature Detection. In Proceedings of the 7th International Workshop on Digital Forensics & Incident Analysis (pp. 27 – 37). Heraklion, Greece.
5. Kent, A. K., Chevalier, S., Grance, T., Dang, H., & Kent, K. (2006). Guide to integrating forensic techniques into incident response. NIST Special Publication, (August).
6. McKemmish, R. (1999). What is forensic computing. Trends and Issues in Crime and Criminal Justice, (118).