SIT - 792

MINOR THESIS

Survey on Internet of Things (IoT) Malware

Under the Supervision of Dr. Lei Pan & Dr. Sutharshan Rajasegarar

BABANPREET KAUR
Student ID - 215462895

Table of Contents

Abstract

Internet-of-Things (IoT) is a developing concept in which number of smart devices are connected to each other over public and private networks to exchange data and information. This may prove to be an advantage of IoT but comes with a cost of security of smart devices. It is found that usually the IoT devices and smart devices get infected with malicious codes called malwares that are very difficult to mitigate as their variants are continuously rising. In IoT, there are number of attack vectors that influence the security and privacy of the smart devices. In most of the cases, the smart devices contain heaps of personal information and data that gets compromised due to malware attacks that can be used by the attackers. In this document, I have provided a review of IoT and malwares along with the prominent malwares existing in the realm of IoT. Also, there is an analysis of different details about certain characteristics of malwares in IoT. Also, the prominent malware families are found along with the malwares in Android and iOS platforms. The anatomy of IoT malwares is also presented with an example of Mirai IoT botnet and its operation. Following this, few detection methods of IoT malwares are presented. At last, a problem analysis of the findings is done along with the results.

Chapter – 1 - Introduction

Internet of things is a broad concept which deals with connecting various devices with the help of internet and establishing a communication between devices. Internet of Things has brought up an evolution in the realm of IT by introducing the world with huge range of applications The IoT devices act as small computational devices with specialized tasks. There are wide range of applications with IoT Devices such as Home automation (smart home), smart energy solutions, healthcare, autonomous connected vehicles, complicated industrial control systems. The domain of IoT is depicted in figure 1. According to a careful research, which is depicted in the figure 2 below, we have around 9 billon smart devices excluding the tablets, smart phones and personal computers. This is anticipated to grow tremendously up to 28.1 billon by the year 2020 which may grow in trillions by the year 2025 (K Angrishi, (2017)). The Internet of Things has emerged into new era of application based platforms where most of the smart devices can be connected to an application in a smart phone or device in order to remotely control the connected smart devices. There are plenty of app-based IoT platforms such as Samsung Smart things, Google, Brillo, Google Weave, Apple HomeKit and so on (YJ Jia, QA Chen, S Wang, A, Rahmati, E Fernandes, ZM Mao, A Prakash (2017)).

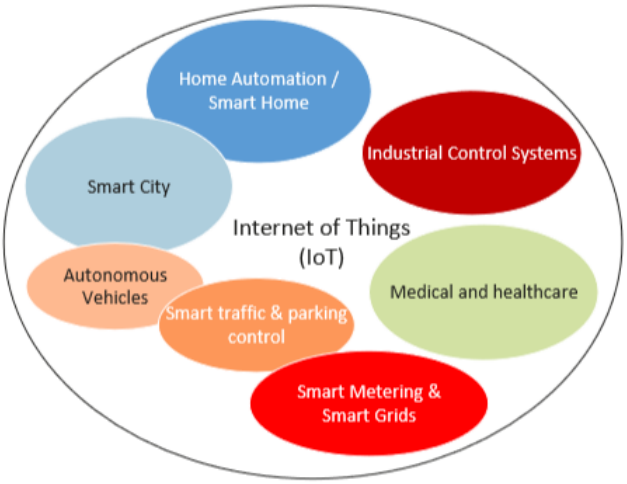

Figure 1 – IoT Domains (K Angrishi, (2017))

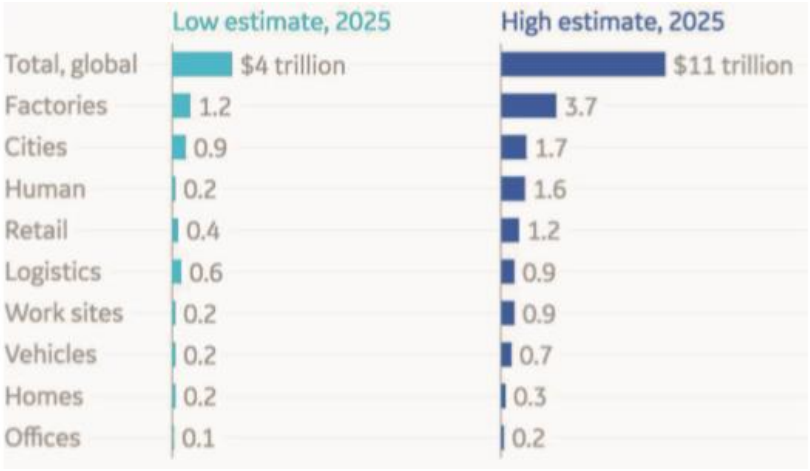

Figure 2 – Rise of IoT devices in Different Sectors (K Angrishi, (2017))

Interesting point about these smart devices is that most of them are designed with low security and some without and security features. The consequences from security and privacy issues in the domain of IoT can be life threatening and severe in some cases. There can be security issues from different types of vectors such as viruses, malwares, intrusions from hackers, DDoS attacks and so on. The most prominent and difficult to eradicate are the attacks due to malwares. Although the PCs and Internet are already immune to many different types of malwares but the nine billion IoT devices that have low security struggle to remain secured from IoT malwares that are increasing rapidly (K Angrishi, (2017)). Malwares are usually set of malicious programs. IoT malwares tend to affect the security and privacy of IoT devices by installing the malicious software that is helpful in getting the important information like sensor data, personal information. Once the smart devices are infected by the malwares they tend of form botnets. Also, IoT malwares are Linux based malware and have no impact on the host device until there is command from the herders of botnet (Suarez-Tangil, G, Tapiador, Juan E, Peris-Lopez, Pedro, Ribagorda (2014)).

Background of Topic –

The Internet of Things (IoT) is said to be a prevalent network of devices and objects that communicate with each other via the Internet. Mostly, these interconnected devices are smart devices such as sensors, mobile phones, smart phones, etc. These smart devices can easily transfer the information without any restrictions. Nowadays there is a huge range of smart

devices available in the market such as smart watches, cars, glasses smart TVs, etc., those are deployed by using networking and computing capabilities. These devices have become very popular in the IoT domains such as healthcare which are providing innovative solutions (Suarez-Tangil, G, Tapiador, Juan E, Peris-Lopez, Pedro, Ribagorda (2014)).

There is a risk of security and privacy issues in the IoT devices. The IoT devices are more likely to get malware attacks due to which the malicious software gets installed in the devices and compromises the important information present on the device. This is because of the reason that the smart devices are fed by low power due to which it becomes difficult to implement security architecture. Also, low memory contributes to the same as for IoT devices full encryption isn't possible. All these factors contribute in the additional requirement of bandwidth for the transmission of encrypted data (A Shifa, MN Asghar, M Flcury, (2016)). Moreover, these smart devices tend to incorporate with the third-party applications very easily which installs the malwares automatically without notifying the user.

The IoT malware differ from the traditional PC malware. The evolution of PC malwares was seen during 2000s when there was increase in computing resources along with the expansion of Internet resources. PC malware was set of malicious codes that had affected the Palm platforms. Most prominent malwares were Symb/Vapor, Symb/Liberty and Symb/Skuller. These malwares used to compromise the important information and used to corrupt the system files that lead to system failure. Whereas the IoT malwares are the Linux based malwares have their roots related to malwares found in mobile phones that became immensely famous after mid-2000s. The main attack vectors were network services such as Wi-Fi, Bluetooth, SMS and MMS. Symb/Cabir was prominent malware. Moreover, recent IoT smart devices have same network and sensor capabilities that becomes vulnerable to malwares leading to the remote control of smart devices by the attackers. Such an example is IoT botnet formed from infected IoT devices due to DDoS attacks that allow the hackers to gain the control of the IoT smart devices (Suarez-Tangil, G, Tapiador, Juan E, Peris-Lopez, Pedro, Ribagorda (2014)).

The major security issues of IoT based on layered architecture and IoT malware exists in two layers of IoT architecture i.e. perception and application layers. Malware and botnet based

attacks are more prominent in perceptual layer and are mostly software-related attacks in which the attacker injects harmful codes like Trojans, Worms and Viruses into the smart devices in order to get the important system information. Also, during the updations of particular applications the malicious codes are installed in the devices (A Shifa, MN Asghar, M Flcury, (2016)).

Symantec in the year 2013, confirmed the discovery of the first IoT malware known as Linux.Darlloz, which enlightens the malware security issue in IoT security. If any victim user connects any of the infected devices, it may destroy various other devices as well. This is the first step in intruding IoT devices with an interest of pulling out massive amount of information which is stored in many devices. In addition to these problems, malwares can crouch inside the IoT devices, which is not designed with proper security mechanisms. Such issues hamper the quality of the IoT devices and violates the privacy of the end users. Moreover, the attacks from IoT Malwares are serious as the conventional security mechanisms fail in case of IoT as they have low computing power and low security. (ZK Zhang, M Wang, CW Hsu, CK Chen, S Shich, (2014)).

Motivation/ rationale of Topic -

The rationale for this research is to identify and present the information about malware in the field of Internet of Things. Although the domain of Internet of Things is very vast, so only fields related to smart phones, smart devices and smart home applications (that are used to connect all the smart devices present in homes to one application) will be examined in order to obtain some useful data about the prominent malwares existing in the realm of IoT. The main focus of this survey is to identify the malware families in specified domains of IoT. Although, there has been significant rise in the number of malwares in recent years. These malwares are of different types such as PC malware, Phone malwares (Android malware or iOS malwares) and so on. So, it is very important to identify the types and families of malwares present in the field of IoT in order to take significant steps to mitigate those malwares and stop them from exploiting the smart devices. Furthermore, a survey is presented in relation with IoT malwares with analysis from research papers. This research will be helpful in differentiating the IoT

malwares from PC malwares. Also, the analysis carried out will present the most suitable methods to identify the malwares that will be helpful in future for the deploying the appropriate security mechanisms for IoT smart devices which will make them less vulnerable to malware attacks.

Problem Statement

This area focuses on the identification of major research questions and challenges that are related to this topic. The main objective is to identify those challenges and find the suitable techniques to solve those research problems and question from an evidence from research. Let us briefly look at each question –

- What are existing malware families in IoT?

  There is a need to identify all the prominent IoT malware families that carry out major attacks in IoT domain.
- How are the attacks in IoT are carried out by malwares?

  It is very vital to know that how the malware-related attacks are being carried out. There should be some information regarding the attack vectors, approach followed and how attack gets triggered.
- What is the main difference of IoT Malware from PC malware – In order to mitigate the malwares in IoT, it is very important to understand their nature and evolution which makes it a highly important aspect to differentiate between IoT malware and PC Malware.

Apart from these research questions, there are some minor questions and problems that are listed below –

- What is the existing security structure for smart devices that is followed nowadays?
- What damage is caused by malwares attacks to IoT devices?
- What are the existing detection techniques, if any, for IoT malware attacks?

Chapter – 2 - Literature Review

This section of literature survey represents details regarding the malware existing in realm of IoT. Its main objective is to analyze and discuss the main findings and information from the existing literature. For fulfilling this purpose, some research papers and journal articles are selected for the literature review that represent the main details about malwares in IoT. In order to achieve this, this section of literature review is further divided into four sections –

I. IoT Malware in Smart Devices and their Families
II. IoT Malwares in Smartphones
III. Detection Techniques for IoT Malwares

Lastly the discussion and analysis of the findings from presented literature review is carried out.

I. IoT Malware in Smart Devices and their Families –

The field of Internet of Things has a very heterogeneous nature that poses out to be a challenge from security point of view. It has been seen that IoT devices have multiple security related issues that are not only confined to communication and sensor networks but also related to few privacy issues such as configuration issues in network, authentication issues and so on. Moreover, the IoT devices are poorly secured that makes them potentially vulnerable. Additionally, the IoT devices like sensors are usually designed to be used at a large scale due to which there are huge number of links that connect the smart devices or sensors. This solely becomes the reason for the security threats in the field of IoT. This issue is also amplified by the ability of some devices to connect automatically to other IoT devices and fielding the devices into insecure environments (Rodriguez-Mota, A, Escamilla-Ambrosio, PJ, Happa, J & Nurse, JRC (2016)). Also, according to Angrishi, the IoT devices have following characteristics that makes it easy for the attackers to launch malware attacks –

- They have very low memory (RAM).
- They have very low flash memory that is prominently used for storing the operating system and firmware.
- IoT devices support ARM and MIPS architectures.

- Does not have any User Interface (UI).
- These devices are usually networked via Internet.

These are some features that pose out to be the threat for IoT devices (Angrishi (2017)).

One of the main threats for IoT devices is malwares that are the malicious software. Malwares are a set of instructions that have the authority to affect the devices and system without any user's permissions. According to the researchers the malwares mainly relied on the sequences of the bytes and API calls. A study by Symantec states that almost every day around millions of malwares along with their variants were used for attacks. Malwares refer to range of intrusive software that consists of computer worms, viruses and malicious software and programs. It also takes the form of executable script and codes that contain malicious information in it. That malicious information gets installed in the device and performs tasks such as stealing of data and passwords. According to the author, there are two types of malware analysis, namely static analysis and dynamic analysis. The static analysis is the one that works on malware variants' signatures and the dynamic analysis works on the method to check the code on API calls to ensure if its malicious or not (Makandar, A, Patrot, A (2015)).

The evolution of IoT malwares has been through PC malwares that modified over the time with the advancement of communication and wireless technologies such as 3G, Bluetooth and so on. It is notable that malwares found in IoT smart devices can be characterized according to following features –

- Attack goals
- Distribution strategies
- Privilege acquisition strategies

The attack goals and behavior characteristics of malwares usually focus on fraud, theft SPAM, sabotage and service misuse activities. Smartphone botnets are perfect example for these particular characteristics as in 2009 when SymbOS/Yxes botnet targeted the Symbian OS platforms to steal important information from the Symbian based smartphones. Similarly, another example for this could be Grayware applications that collect the sensitive information

from the smart devices and use it for dubious activities where user's approval is not required. This has become one of the main privacy threats due to malwares. Furthermore, the characteristic of malwares to distribute by themselves by two means of self-propagation and social engineering. Under the self-propagation technique, the malwares tend to automatically install itself on the smart devices using different strategies. Whereas under the social engineering technique of distribution the malwares exploit the security unawareness of the users to put them into illusion to install the applications to allow the hackers to perform the malicious activities. For instance, Andr/Opfake-C malware has the ability to spread through Facebook and install by itself. This allows the hackers to make all the premium calls. Also, there are few other distribution vectors that are listed below –

- Market to Device
- Application to Device
- Network to Device
- Web-browser to Device

All these vectors have the ability to infect the smart devices. In Market to Device distribution strategy the attackers tend to add some malicious applications to the application market and users get affected once they install the applications. In Application to Device propagation strategy the attacker relies on the vulnerability application and exploit to spread that automatically. The Network to Device propagation strategy is based upon the exploitation of vulnerabilities in the device. In the Web-browser to Device distribution strategy the malwares propagate through the web-borne attacks. Lastly, the characteristic of privilege acquisition consists of various techniques utilized by the malwares. In some cases, malwares tend to gain privileges by exploiting the vulnerabilities such as API based vulnerabilities, buffer overflows, system vulnerabilities and so on. Also, in some cases the privileges are granted by the users unknowingly for example social engineering techniques (Suarez-Tangil, G, Tapiador, Juan E, Peris-Lopez, Pedro, Ribagorda (2014)). The author also lists the malware classes. These are Worm, Spyware, Trojan, Backdoor, Rootkit and Botnet. The worm is a type of malware that tends to slip into the systems and device without the user's permission. It also operates without

any acknowledgement to the user. These can have deadly effects on the websites, devices, PCs, etc. The spyware is a type of malware that collects the information with a motive of advertising that is from a third-party. It is usually successful to obtain the credit card numbers, login credentials of the social networking sites and so on. Also, spyware with an access to cameras tend to record the video and send it to people automatically. Trojan is harmful piece of malicious software that that tricks the user and create the backdoors to provide the hackers the remote access of the devices. Backdoor is designed to avail the opportunity to get the remote access and administration of the device. If installed in the device, it gains the control of machines and devices without any knowledge to the users. Rootkits tend to hide itself and loads certain programs into the system with a goal to modify the kernel of the operating system. Lastly, the botnet tends to gain the control of compromised devices remotely by using attacks such as DDoS attacks (Peng, S, Yu, S, Yang, A (2014)). The figure below shows the attacks in the smart devices by type from 2004 to 2014 –

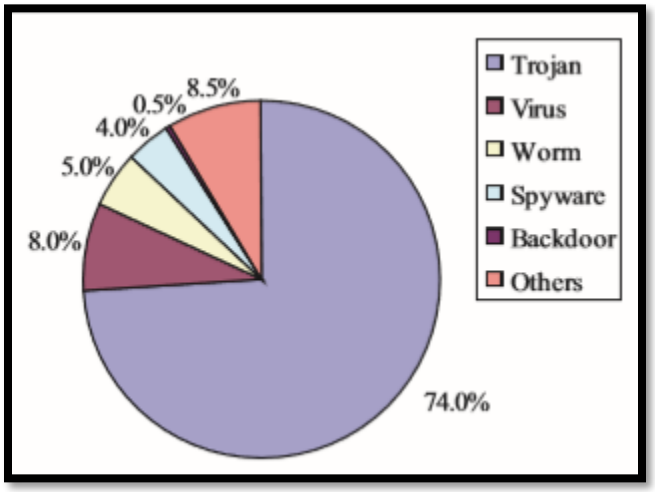

Figure 3 – Smart Device Threats (by Type) from 2004 to 2014 (Peng, S, Yu, S, Yang, A (2014))

According to Angrishi, malware in IoT are incorporated to IoT devices and launch DDoS attacks to convert them into IoT Botnets. The DDoS attacks are launched in the IoT devices by using two techniques of reflection and amplification. In the reflection technique, the DDoS attack is launched when the attacker tends to send the distinguished packets forged IP address that resembles the IP address of the target/victim as the main source address of packets to all the

destinations. Whereas in amplification, the attackers send very less number of packets to elicit the big number of packets that are directed towards the target/victim. These both techniques can be combined to launch a big attack on IoT devices. Mostly malformed TCP, ICMP or UDP traffic is used to launch DDoS attack. The IoT devices have become very vulnerable to be used as the botnets in order to launch the DDoS attacks at large scale. Nowadays 96% of DDoS attacks have been launched by the IoT devices and rest 4% from the other routers and compromised Linux servers. The IoT botnets effect both IoT devices and the also the persons using Internet (Angrishi (2017)).

Also, author specifies the main characteristics of the IoT malwares that incorporate the DDoS attacks. These are –

- These are Linux based Malwares.
- They become active whenever there is Control and Command from the attacker to carry out the DDoS attacks, otherwise they don't affect the performance of the host site.
- The IoT malwares don't get propagated by the reflection technique.
- The IoT malwares are usually written in language C.

The following list discusses the examples of IoT malwares that were mainly utilized for launching DDoS attacks in form of IoT Botnets –

- Linux/Hydra - One of the oldest known IoT malwares is Linux/Hydra, released in 2008 and it was the earliest malware known to attack the IoT devices by DDoS attack functionalities.
- Psyb0t – It's the IoT malware that targets the DSL modems and routers and it was released in 2009. Also, these attacks use brute force attacks that access the telnet and SSH of IoT devices.
- LightAidra/Aidra – This IoT malware is designed to extensively search out the telnet ports that are opened and usually attacks the IoT devices with MIPS, ARM architecture.
- Linux.Darlloz – It is a type of IoT worm that has the ability to spread by exploiting PHP vulnerability and affects the platforms same as LightAidra. This IoT malware results in

inaccessibility of a device using Telnet ports. In a study, it was found that almost 31000 IoT devices were infected with this.

- Spike/Dofloo – It is a type of IoT malware that was found in 2014 and used to target Windows PC (both 32-bit and 64-bit), Linux PCs and IoT devices based upon ARM and MIPS architecture. It can launch the attacks with payloads like UDP floods, DNS queries and so on.
- BASHLITE/Torlus/Lizkebab – This is one the most popular IoT malwares that affects the IoT devices that have Linux operating system. It is noted that this malware has been responsible to enslave nearly 1 million IoT devices. It is also known as the predecessor of Mirai malware botnet.
- Mirai – It has been recognized as one of the most dominant IoT malware botnet. It has been noted that it affected around 4000 IoT devices in an hour. It is also responsible for around DDoS attack of 1.1Tbps and affected 148000 IoT devices that included DVR's, CCTV cameras, home routers (Angrishi (2017)).

The following figure shows the operation of Mirai botnet that causes the DDoS attack by targeting number of servers and weakly configured IoT devices.

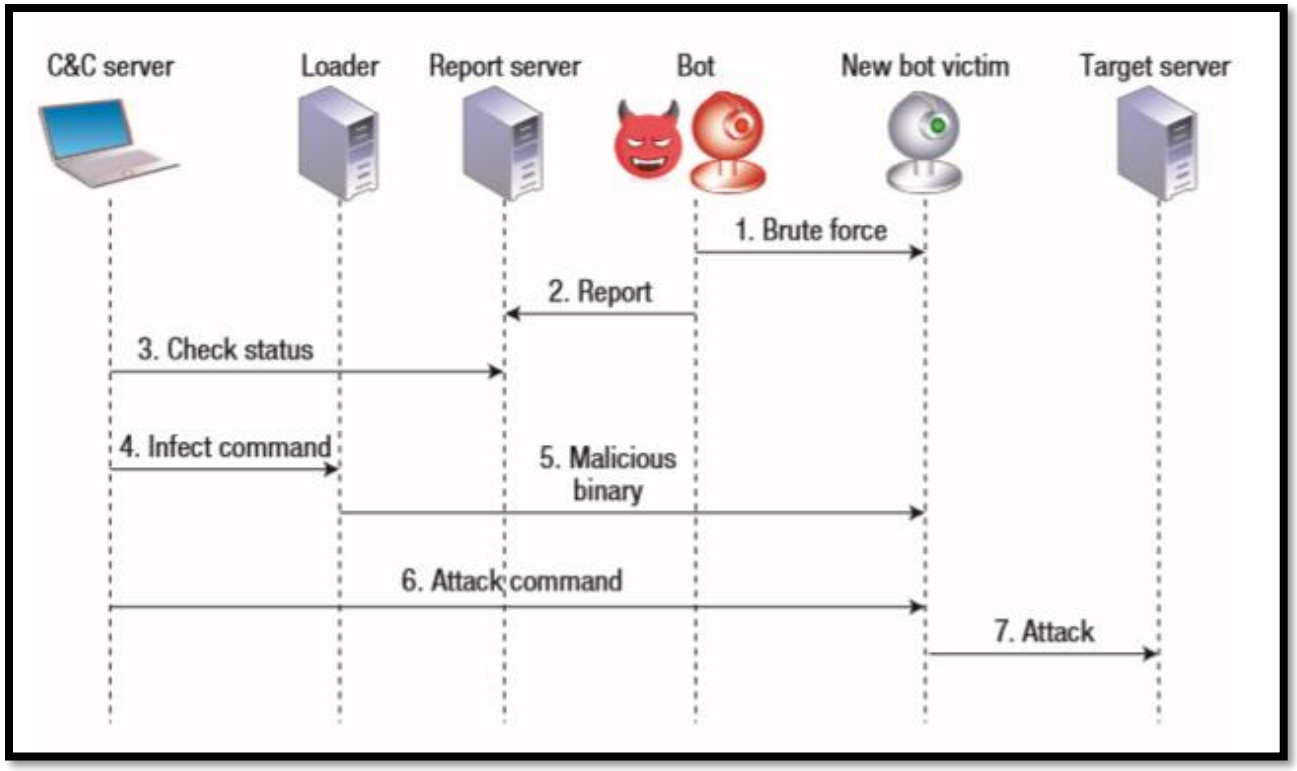

Figure 4 – Mirai IoT Malware Botnet operation (Kolias, C, Kambourakis, G, Stavrou, A & Voas, J, (2017))

The main components of this botnet are bot, Command and control (C&C) server, scanner, reporting server, loaders and distribution server for malwares. The Mirai botnet first starts the brute-force attacks to find out the credentials for weakly configured IoT devices. Then it gains a shell (CLI or GUI) to report various characteristics of device to the reporting server. Using C&C server the botnet scans for new targets and issues the infect command to exploit the vulnerabilities. Then the loader tends to install the binary of the malware in the IoT device. Then the malware executes and it shuts down all other ports such as SSH and telnet. Then an attack is launched after the botmaster issues the instruction due to which the bots tend to attack the server by using any of the attack variations such as GRE (Generic Routing Encapsulation) or HTTP flooding attacks (Kolias, C, Kambourakis, G, Stavrou, A & Voas, J, (2017)).

The above paragraphs have quoted plenty of information regarding the IoT malwares, their characteristics, their attack vectors and most importantly there is a list of most prominent IoT malwares along with the operation of Mirai malware botnet. Now a further research is carried out in order to discuss the most prominent IoT malware families that affect the IoT devices.

According to the research carried out by Mohaisen, the classification and analysis of malwares is very important and in order to fulfill this the authors have proposed a system known as CHATTER that examines all the events in a device. As per the authors, it is very vital to identify the malware families as it helps in the mitigation and damage assessment due to attack. Moreover, there are two techniques that are followed for classifying malwares, namely, Signature-based and Behavior-based. Signature-based techniques is based upon the examination of patterns retrieved from reverse engineering and analyzing malware manually, which requires lots of efforts. Behavior-based technique utilize run-time execution to extract the important characteristics. CHATTER is also based upon behavior-based technique. The flow diagram for CHATTER is shown below –

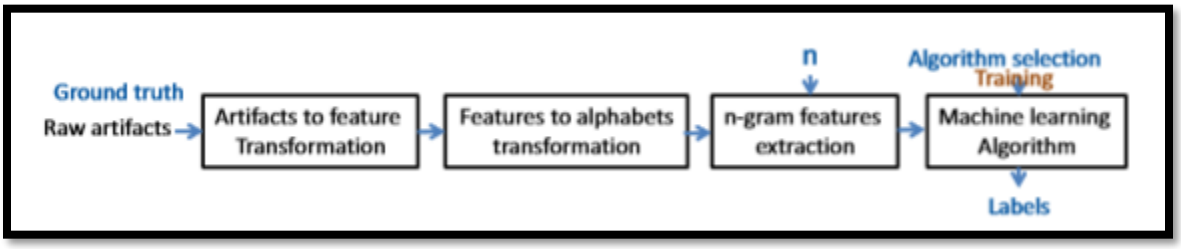

Figure 5 – Flow Diagram for CHATTER

The evaluation of malware samples is for malwares belonging to Darkness, Zeus and SRAT malware families that are included in network based behavior. Events that are utilized in order to examine malware families based upon n-gram approach are connection events, request and response events and the following results are obtained when used for different algorithms.

	n-grams	1				4				8			
	Algorithms	P	R	A	F1	P	R	A	F1	P	R	A	F1
Zeus	*k*-NN	80.79	79.68	81.48	79.97	79.07	83.90	82.25	81.35	78.29	78.17	79.64	78.09
	SVM	67.41	82.67	72.69	73.92	75.96	80.47	78.67	77.84	80.41	82.87	82.45	81.50
	Decision Trees	80.14	80.90	81.74	80.42	81.13	81.82	82.67	81.35	80.82	82.82	83.02	81.77
Dark.	*k*-NN	76.22	73.13	76.08	74.56	80.40	71.52	77.70	75.57	71.38	69.58	71.65	70.20
	SVM	76.82	32.38	62.24	45.05	78.18	71.32	76.45	74.35	76.62	76.36	77.22	76.27
	Decision Trees	80.45	72.56	78.20	76.07	81.75	72.89	79.04	76.93	80.50	68.37	76.39	73.59
SRAT	*k*-NN	81.38	76.78	82.78	78.45	83.87	81.83	85.51	81.95	83.99	74.28	82.93	78.16
	SVM	76.88	65.43	75.88	69.55	83.70	82.94	86.23	83.03	85.68	80.86	86.33	82.71
	Decision Trees	85.16	81.11	86.44	82.60	88.28	81.65	88.01	84.45	86.13	78.92	85.54	81.85

Table 1 – Results obtained with CHATTER based upon n-gram approach, determining Precision, Recall and Accuracy based upon Algorithms, k-NN, SVM and Decision Tree

Results obtained from above table state that the Darkness malware family is most difficult to identify where as SRAT can be identified easily. Also, it is seen that the accuracy of identifying the malware families is maximum when using Decision Tree algorithm in CHATTER (Mohaisen, A, West, AG, Mankin, A (2014)).

Chatter has an ability to characterize the samples of malwares by executing them, thus using this intelligence for training purposes. This capability is not required for online operation. Let us go through the design goals and requirements:

• Cost-effectiveness: Feature extraction should be computationally inexpensive in the online operations. CHATTER tends to ignore solutions requiring deep analysis of a large number of artefacts.

• Less-invasiveness: A system can be deployed externally to observe malware is ideal. This is only a design objective for the final system and not a requirement on how that system is trained or arrived.

• Generalizable and multi-purpose: The main aim of CHATTER is to characterize malware samples. The system should be flexible for re-purposing outside the malware realm. In this paper, they have described two such applications that benefit from these generalized techniques.

• Flexible to behavioral changes: Malware families can evolve over time to prevent behavioral techniques. In the current design of the system, the main aim of the author is to resist these evolution techniques by providing flexible and longitudinal aware techniques.

• Accuracy: The main goal of the system is to provide greatest coverage and minimize the false positives. In this paper, they have demonstrated operationally acceptable accuracy (Mohaisen, A, West, AG, Mankin, A (2014)).

As per the research by Liu, utilizing Phylogeny models for the purpose of malware forensics focuses on determining the evolution-based relation among different malwares. This method is helpful in identifying the malware evolution and also assists in identifying the nature of new malware. Phylogeny models are usually represented with the help of tree-like structures that represent the relation between distinctive instances. In this research, the authors utilize the concept of API sequences for systems that consist DLL libraries containing particular information. Also, utilize MJ (median-joining) network (Liu, J, Wang, Y, Wang Y (2016)). Whereas the research done by Adeel, focuses on introduction of malware detection framework in mobile phones. The analysis is carried on Cabir and Commwarrior malware families. It is well-known that mobile phone malware has been emerging since the first malware outbreak in 2004. As per the reports, there have been hundreds of worms, Trojans, viruses and spywares found along with their variants. These malwares cause certain loss in privacy by transfer of important information that leads in malfunctioning of mobile device. These malwares propagate through mobile phone networks such as 2G, 3G, WLAN and so on. Also, the author states that Cabir and Commwarrior families of malware are one of the most dangerous families of mobile devices due to their characteristics of mutation, proliferation and high variants. For Cabir family around thirty-three variants have been discovered since 2004 that utilize Bluetooth a medium to spread whereas Commwarrior family was found in 2005 and propagate through MMS. Analysis on these families is done with the help of simulation technique known as MPeersim through which mobiles phones are examined using different capabilities as shown in the table below. Also, some results have been shown in second table (Adeel, M, Tokarchuk, LN (2011)).

Normal File Types (Bluetooth, MMS, SMS)	Virus File Types (All families from Section 1)	Initial Virus Population (Variable)
Communication Types (Bluetooth, MMS, SMS)	Node Battery Power (Variable)	Node State (Active, Idle, Dead)
Node Associations (Bluetooth Neighbors, Phonebook Contacts)	Battery Usage Modes (Bluetooth, MMS, SMS)	Node Type (Agent, Normal Node)
Node Leave-Join Rate (High, Medium, Low)	Global Probability of Infection (0<GPI<1)	Initial Normal Files Population (Variable)
Node Leave-Join Time (Anytime)	Immunization (Bluetooth, MMS)	Mobile Node Type (Type 1, 2, & 3)
Node Mobility (Variable)	Bluetooth Neighbor Density (Variable)	AI Support (Neural Nets MLP & DTree C4.5)
Payload Type (Normal, Malware)		

Table 2 – MPeerism Analysis based upon File, communication, Node, Payload types (Adeel, M, Tokarchuk, LN (2011))

In terms of propagation behaviours, mobile viruses and worms do have some common features. By carefully investigating these common features we can potentially group these malwares into different categories (also known as malware families). Behaviour based anomaly detection systems are known to perform better than signature based techniques in the situations where malwares perform constant mutations. 'Cabir' the most widely recognized mobile malware families, is very vulnerable and rapidly evolving. Hence it has become very important to develop certain detection and mitigation techniques for mobile devices and networks which can not only detect these anomalies but also can safeguard the malware evolution (Adeel, M, Tokarchuk, LN (2011)).

The development of the 'P2P' malware detection framework can be uncharacterized into four distinct phases: classification of mobile P2P malware, development of simulation environment, propagation analysis and lastly detection of mobile P2P malware. As per the research conducted in the paper, there are altogether 13 families that comprises of 25% of malware existence. They have adopted a malware detection technique, which has reduced malware

propagation and without compromising the detection accuracy. They have adopted a mobile P2P simulation environment which is capable of taking various mobile P2P characteristics and also provides a strong platform for analysis of mobile P2P malwares. This strong platform will benefit various research scholars to study the propagation behaviour of future malwares. They have evaluated various kinds of viruses with strong results depriving the behaviour of 'Commwarrior' family malwares and dangerous malwares such as 'Cabir' which are responsible for battery depletion, propagation speed and attack strategy (Adeel, M, Tokarchuk, LN (2011)).

Type	Description
Malware Prevalence	Prevalence of each family malware at any instance during simulation
Battery Power	Node, network and cumulative battery power of 2^{nd} & 3^{rd} degree neighbors
Throughput	Bluetooth, MMS, SMS throughput for node, network and cumulative throughput of up to 2^{nd} & 3^{rd}
Node Status	Number of Active, Dormant an Dead Nodes in the network at any instance during simulation
Threat Identification	Adaptive classification of activities into normal, and malicious activities
Sub-Threat Conditions Identification	Adaptive identification of flags (i.e. sub-threat conditions) on agent nodes through classification of instance/activities
Malware Family Identification	Adaptive identification of malware families based on agent nodes through classification of instance/activities
Node Relationship Diagram	Elaborates the Bluetooth associations formed by network nodes to constitute a network
Malware Propagation Modeling	Propagation modeling of generic mobile malware (i.e. Bluetooth, MMS, Hybrid) and actual mobile malware (i.e. all the families in Figure 3)

Table 3 – MPeerism Results (Adeel, M, Tokarchuk, LN (2011))

II. IoT Malwares in Smartphones -

This section discusses about the IoT in the field of smartphones. The smartphones nowadays are based upon different operating systems but most prominent platforms are Android and iOS. So here, I have discussed the IoT malwares in both Android and iOS based smartphones. It is seen that there has been very fast growth of smartphones, especially Android based

smartphones, but there has been plenty of application in the android application market such as backdoor, adware, bot, if installed, then it could get the control of whole device. According to a recent research there has been more than 1000 malwares in Android. There are around 40 most prominent IoT malware families in Android. Some of them are ADRD, DogWars, DroidDelux, DroidDream, DroidKungFu1, DroidKungFu2, oldKungFuSapp, GoldDream, Plankton, zHash, Zsone, etc. There has also been number of methods used to depict the relations and similar behaviors among different malware families (Hsiao SW, Sun YS, Chen CM (2016)). Also, it has been seen that almost 97% of the malwares target the Android smartphones with a motive to steal important information and money. Malwares in smartphones tend to conduct the activities such as stealing emails, texts, calendars, contact lists, sending texts illegally, steals important information such as documents, bank account information, social networking sites' credentials and so on (Chakraborty T, Pierazzi F, Subrahmanian VS (2017)). The figure below shows that there has been extensive increase in the Android malwares as compared to other operating system smartphones.

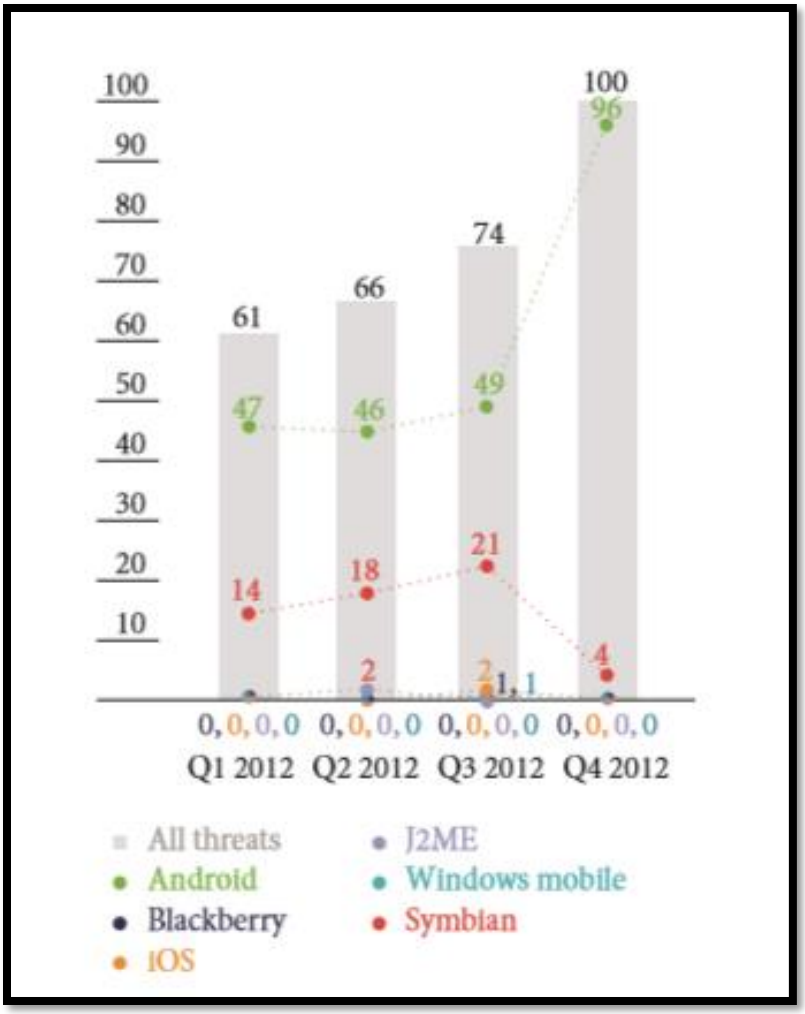

Figure 6 – Increase in Smartphone Malwares (Ham, H, Kim, H, Kim, M & Choi, M (2014))

Detection can be one of the possible countermeasure against these malwares. There has been lots of literature on the basis of mobile malware analysis which focuses mainly on detection. Some of these works consists of outdated datasets which do not characterize malware families through the feature explanation, because they use features which are low level and very difficult to interpret. Most of these existing works have drawbacks as they mainly focus on large families (malware) for which the training data is available and neglect the small families (malware) which comprises of novel malware variants on which the security analysts should mainly focus and prevent the malwares. In this research paper, the author has used the Android Malware Genome Project that is utilized for the academic research. Some of the works such as DREBIN, DroidAPIMiner, CrowDroid carry out detection of malware using static analysis while

TaintDroid and DroidScope utilize dynamic analysis. The figure below shows the detected analysis of malware families detected using cumulative distribution (Chakraborty T, Pierazzi F, Subrahmanian VS (2017)).

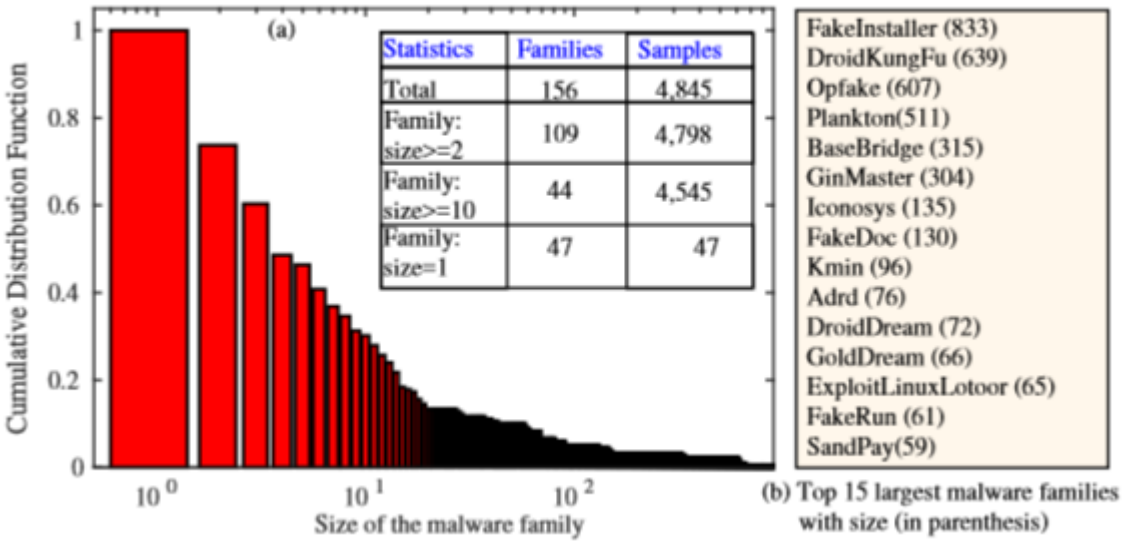

Figure 7 – (a) Detection Analysis from DREBIN dataset using Cumulative Distribution, (b) Top 15 Malware families (Chakraborty T, Pierazzi F, Subrahmanian VS (2017))

A research by Cimitile (2017), stated that constructing phylogenic trees for determining the evolution of Malware families in Android is the best approach. This technique focuses on building the evolutionary relation between the malware families. The groups of same malware family tend to show common behavior and properties that relate in other families of malwares as well. Construction of phylogenic trees is mainly based upon the dynamic analysis of the malwares as it has been seen that the authors of codes of malwares have the property to hide them from the revealing the relations with other families. The derivation of relations can be done through several techniques such as code reordering, garbage insertion and so on. This paper utilizes the CCS i.e. Calculus of Communicating Systems to specify certain calculus properties that lead to results of common behaviors of malware families. It assists in analyzing the new malware that can be identified as most of times the new malwares are generated by adding new codes to the existing codes of malwares. Also, it is useful in determining the naming

technique for the malwares. This technique also reveals the ancestor-descendant relationship between the malwares (Cimitile A, Mercaldo F, Martinelli, Nardone V, Santone A, Vaglini G (2017)).

The research by Hsiao, states about the grouping of malwares in Android platform based upon the behavior of malware families. This is carried out with the help of phylogenic tree technique that examines different malware families. Also, some significant characteristics have been identified along with designing analysis tool for Android apps. The following phylogenic tree was obtained for two malware families that reveals that both the families DroidKungFu1 and DroidKungFu2 are variants for each other (Hsiao SW, Sun YS, Chen CM (2016)).

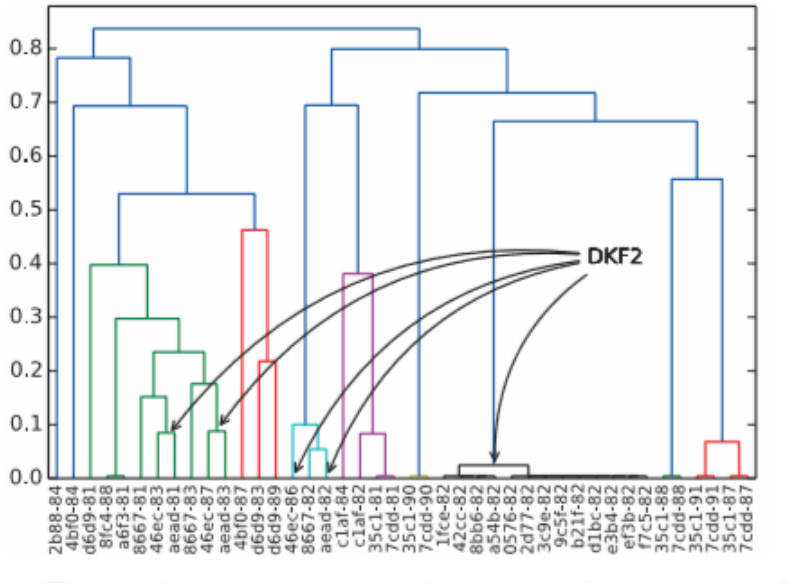

Figure 8 – Phylogenic Tree for DKF1 and DKF2 malware families (Hsiao SW, Sun YS, Chen CM (2016))

The author designed phylogenetic tree for two families of malwares, namely, DroidKungFu2 and DroidKungFu1. It was found that for DroidKungFu2 the programmers simply wrote new codes for the variants just with few changes in the codes. Similarly, for the malware families of DroidDream and Droid DreamLight it was found that there was a total of eleven major behavior groups. These families belonged to different families but the threads belonged to the same

families. Although, it was found from this study that the classification method using phylogenetic tree was good way to identify the evolution of the malware families from the thread profiles. But this method was not successful in depicting the mechanisms and classifying rules of the malware families (Hsiao SW, Sun YS, Chen CM (2016)).

From the above research, it is clear that the following are the malware families existing in the realm of smartphones. Let us now know these malware families:

- Ransomware: It makes inaccessible the files or the devices of the victim. There is only one way to unlock the device which is to pay a ransom to the attacker.
- Adrd: It has less server side commands and it is close to Geinimi.
- DroidDream: It can access unique identification information by gaining root access to the device.
- DroidKungFu 1, DroidKungFu 2, DroidKungFu 3, DroidKungFu 4: It allows hackers to access the smartphones at the time they wish to use it by installing a backdoor.
- Fakeinstaller: Antireversing techniques, server-side polymorphism, frequent recompilation and obfuscation.
- Geinimi: It can receive commands from remote server which will allow the owner of that server to control the phone.
- Kmin: It does not have an ability to kill antimalware process and is similar to the BaseBridge.
- Opfake: Through premium text messages, it demands payment of an application content.
- Plankton: The Dex code which is downloaded has and ability to launch the connection to the command server and listens for the commands when they get executed (Bernardi ML, Cimitile M, Martinelli F, Mercaldo F (2017)).

IoT malware in iOS Platform

Whereas for iOS platform there is less number of malicious code or malware as compared to the Android platform. For iOS operating system, malicious software only penetrates the system when the phone is jailbroken. Also, the distribution mechanism of iOS is designed that the third-party applications can only be a part of Apple's App Store if they get approved and signed by Apple Inc. This process includes the step of auditing by Apple Inc. that further includes checking the application for its content and function to ensure that the applications does not consist of any malicious functions before hitting the App Store. Also, the developers deploy the encrypted firmware package to make it hard for the third-party applications to get the sandbox access and analyze the kernel of iOS devices. This way iOS platform seems out to be more secure than any other platform. It is seen from the studies that the iKee is one of the earliest known worm in iOS platform and was found in most of the jailbroken devices. This worm scans the OpenSSH ports and automatically logs on into the account. Once an iOS device is infected with iKee then it gets added into a botnet to spread the malicious codes to other devices. Although there are some more applications from non-jailbroken devices that spread out the malware, these are SpyPhone, Mactans, XcodeGhost, and Jeky11. SpyPhone installs a malicious code that tends to steal the phone numbers from the iOS device. Mactans also tend to steal information from the iOS devices by connecting them to a particular kind of charger that enables privacy theft and monitors user clicks. Jeky11 triggers the backdoor in its code to execute number of malicious activities on the iOS devices. (Yixiang Z, Kang Z (2017)).

The research conducted by Peng states the propagation methods utilized by the malwares to spread in the IoT devices. The malwares tend to follow the dynamic process for the propagation, hence the authors have classified the propagation models in order to detect certain things related to the propagation mechanisms followed by the malwares, decide the extent of effects that malware attack can have on the devices, effect of malware propagation on the networks, classify the infection traits on devices and at last design the methods to mitigate malware attacks on the IoT devices and smartphones. For this purpose, the authors have discussed the different types of propagation models such as deterministic models, spatial temporal models, mathematical-based propagation models, SIP model, SIS model, SEIR model,

etc. Also, the problems in the existing models are discussed. These are particularly related to the diversity of models that one particular propagation method tends to research about the propagation method of the particular type of malware, for instance, SEIRD model analyses the Bluetooth worm. Another drawback is that the propagation models are based upon the partial data and information on malware propagation. At last there is a difficulty in carrying out the comparison of different propagation methods (Peng, S, Yu, S, Yang, A (2014)).

III. Detection of Malwares in IoT devices –

As IoT malware is a potential threat and in order to mitigate that there need to detect the malwares in order to design countermeasures for the same. According to Ham et. al there are number of malware detection techniques available at present that can classified into three broad categories –

- Signature-based Detection
- Behavior-based Detection
- Dynamic Analysis Detection

Signature-based detection techniques usually refer to the traditional method of detection in which both static and dynamic based detection techniques are used. As discussed, the static and dynamic detection techniques work simultaneously and focus on the examination of source code and traffic/data flow respectively. Behavior-based detection is related to examining the attacking patterns and behaviors in a system. The dynamic analysis detection technique is usually related to the taint analysis for the source code of the malwares (Ham, HS, Kim, H, Kim M, Choi, M (2014)).

Bernardi et. al has done a formidable research which describes an approach for dynamic malware detection which is based on combination of both the Process Mining and Fuzzy Logic techniques. Fuzzy logic mainly helps to classify the identified malware applications and verify their relations with the existing malware variants. Process mining is a step used to characterize the behavior of an application. The combination of these two allows to obtain a powerful application that is used to verify their relations with existing malware variants. It can establish if

it detects a known malware family and identify the differences between the detected malware behavior and the other variants of some malware family.

Malware Detection process:

The adopted malware detection process which is been discussed in the above paper has following steps:

- Computing SEF of an application: For this step they used the term 'APK' for android applications. The process to compute SEF using APK consists following steps.
  - Syscall Traces Extraction
  - Syscalls Execution Fingerprints (SEF) computation
- Computation of SEF for families of malwares: In this stage, the mine consists the set of declare models from syscall traces generated from different APK infected with the corresponding malware family when it is simulated with a given system event. The SEF is the obtained result of the family of malware characterizing the behavior of the malware family (Bernardi ML, Cimitile M, Martinelli F, Mercaldo F (2017)).

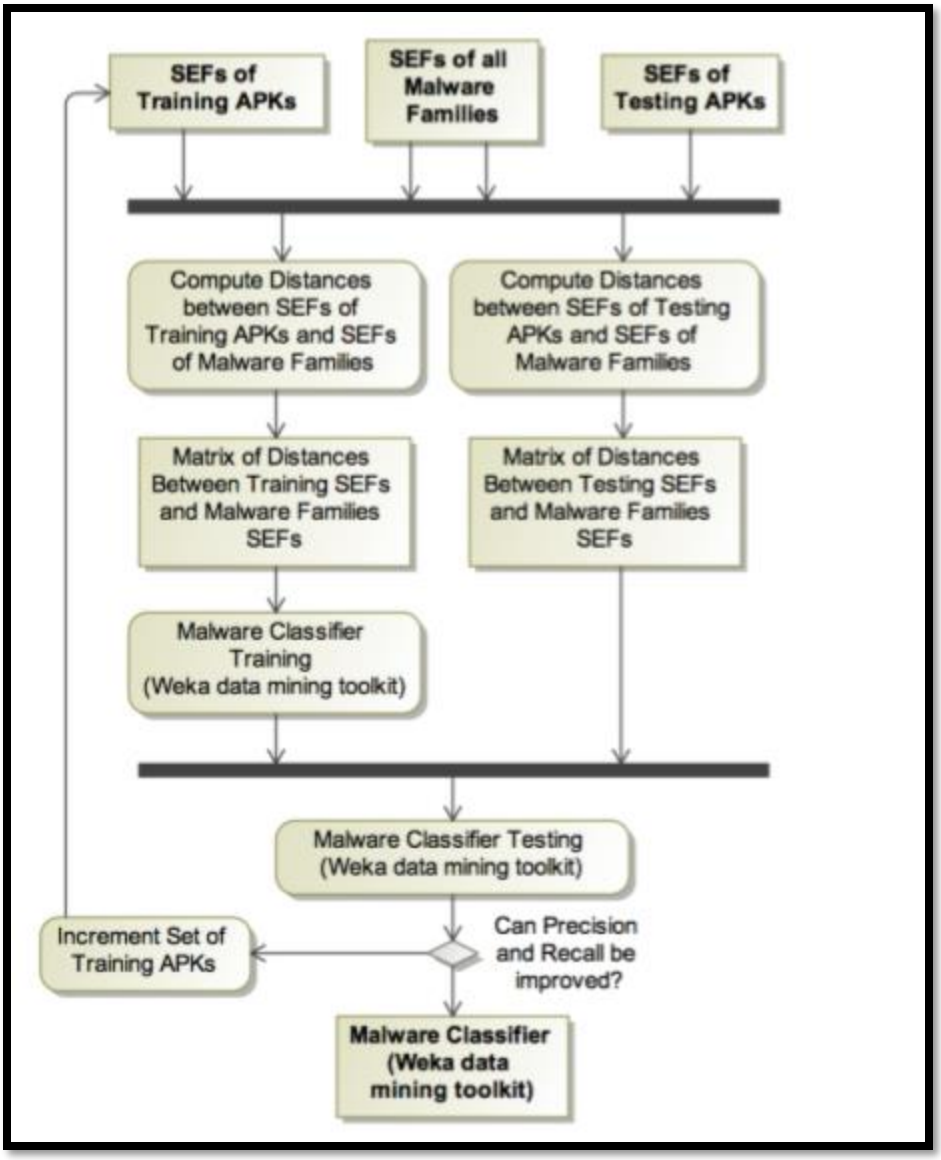

Figure 9 – Malware Classifier (Bernardi ML, Cimitile M, Martinelli F, Mercaldo F (2017))

Some researchers have proposed a Virtualised network as a type of testbed to detect and analyze the malwares. The main advantages of this setup are that it can be reused and it is very stable to detect the malwares. It has been setup using VMware virtualization technology that helps the user to create the virtual networks on the virtual machines that can be configured. Also, it provides a software support to implement the IDS and malware sensors to detect the attacks. Further the authors use a simulation tools to model an environment that resembles a

real network to study the malwares. Example of such tools are SSFNet, NS-3, etc (Ahmad, MA, Woodhead, Sm Gan, D (2016)).

For detection malwares Liang et. al proposed a framework that can be used for malware detection of smartphones. The architecture of framework is shown below –

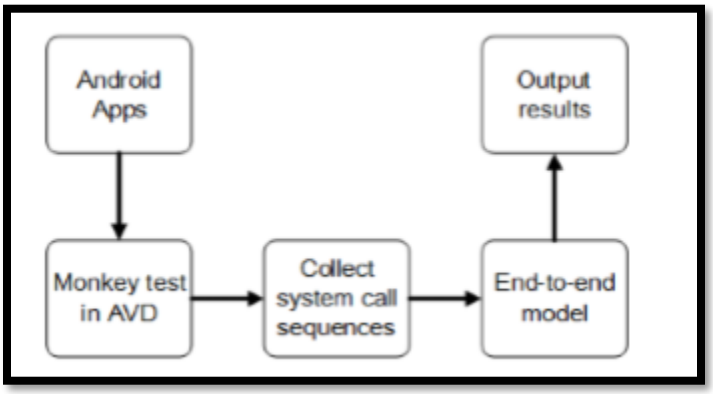

Figure 10 - Android Malware Detection Model (Liang, H, Song, Y, Xiao, D (2017))

It has three steps – System Call Embedding, System-call-level Vertical Multiple Convolution and Multilayer perceptron. In this the first layer maps the system calls then vertical multiple convolution to mine the large matrices. At last the perceptron is utilized to carry the classification (Liang, H, Song, Y, Xiao, D (2017)).

Chapter – 3 – Problem Analysis and Discussion of Results

The survey carried out for Malware in IoT depicts the main differences between the PC malware and the IoT malwares by presenting the characteristics of the IoT malwares and important attack vectors. The critical analysis also consists of the main types of malwares that exist in the IoT devices and what attack vectors they utilize to attack the smart devices. Angrishi et. al and Makandar et. al described the main characteristics of the IoT malwares and found that malwares can be classified into different types according to the attack vectors, distribution strategies, pace of spreading, propagation mode. It is shown in the table below –

Type	**Worm**	**Trojan**	**Backdoor**	**Spyware**	**Rootkit**	**Botnet**
Propagation Mode	Self-replicating	Deceptive	Deceptive	Deceptive	Deceptive	Deceptive
Attack Vector	M2D, A2D	N2D, W2D	A2D	N2D	N2D, A2D	N2D
Spreading Pace	Very Fast	Slow	Normal	Slow	Fast	Very Fast

Table 4 – Analysis of Different types of Malwares on basic properties

It is clear from the above table that except worms all other types of malwares are propagate in a deceptive way. Mostly malwares spread from the Network to device attack vector, some with low spreading pace and some with very fast spreading pace. Also, the literature represented above depicts that Trojan type of malwares have been very prominent in the smart devices since a decade.

Further, the most prominent IoT malwares are discussed such as Linux/Hydra, Psyb0t, BASHLITE and Mirai malware. It is found that these malwares mostly effect the IoT devices with ARM and MIPS architectures with the help of Command and Control (C&C) servers found after forming botnets. The IoT malwares also have an impact on Linux based smart devices. Continuing to this

the operation of Mirai IoT malware botnet is described to get an insight on how the IoT malwares tend to form botnets by compromising the smart devices and converting them into bots. The botnets also tend to block the Telnet and SSH ports for the smart devices. Interestingly, the user does not get to know about the fact that his smart device has been compromised unless and until the herder tends to carry malicious activities with the botnets.

Further, three methods were explained by Mohaisen et. al, Liu et. al and Adeel et. al to study the behavior of IoT malware families such as Darkness, Zeus, SRAT, Cabir, Commwarrior. It led to the inference that Darkness malware family was the most difficult to identify.

This was followed by the research on IoT malwares in smartphones and was found that the following were most prominent IoT malware families in the smartphones –

IoT Malware Family	Description
Adrd	It is close to Gemini but with less server-side commands.
DriodDream	It can gain root access of the smart device to steal information
DriodKungFu 1, DriodKungFu 2, DriodKungFu 3	These families install backdoor and allow the attackers to access smartphones
FakeInstaller	Responsible for server-side polymorphism and recompilation
Opfake	Demands payment for application through premium text messages
Plankton	Download Dex code to launch connection to C&C server.
Gemini	Receives commands from remote servers

Table 5 – Prominent IoT malware families existing in Smartphones

It was clear from this study that Android based smartphones incurred maximum number of malware attacks as compared to iOS based smartphones. Furthermore, another inference is drawn from the research carried out by Hsiao et. al, Cimitile et. al and Chakraborty et. al that detecting the evolution of Malware families can be effectively done by constructing the phylogenic tree as this method tends to relate the threads to the adjacent threads that

generates the connection between the malware families. Also, it makes it very clear that the malware authors could have simply put very few variants while deriving the new family of malware. This method is also very helpful in designing the countermeasures that should be taken in order to mitigate the malware attacks.

Lastly, some malware detection methods are discussed that have few drawbacks at this stage because the concept of IoT malwares is still in its infancy and it is entirely different from the malwares existing in PC's and conventional systems. The detection of malwares could be done by implementing the simulation-based tools that identify the behavior of the systems and devices. It can also be done by classifying the malwares according to their behaviors and using the mining tools to identify the important information. Also, few authors utilized propagation methods of malwares to identify them, but those methods have so many drawbacks such as partial information is available after the results as most of the steps for these methods is mathematical and sometimes all the propagation methods cannot even be compared as there is no decided standard to compare those.

In respect to the problem statements of the survey discussed in the introduction, this survey completely provides the supportive evidence to each and every problem statement. The most prominent IoT malware families have been identified and discussed as per their characteristics. Further, the propagation methods and the attack vectors are discussed that provide a suitable evidence to the problem question of ways the IoT malware attacks carried out. It is also supported by discussing the anatomy of botnet and how the IoT malware botnet attacks are carried out in IoT devices. Further, the background of topic has been successfully able to depict the difference between PC Malware and IoT malwares. This is really useful as PC malware has different mechanism in comparison to the IoT malware. Additionally, the minor problem questions of security structure of smart devices, damages caused due to malware attacks and existing detection techniques are completely supported by the evidence from the survey of literature review and is analyzed above.

Conclusion -

In this research, I have presented a survey from quality research papers to extract some basic details about the malwares in IoT. The very first IoT malware was released in 2008 and it was Linux/Hydra that carried out malicious activities of stealing important information from the infected devices. Following this, the main characteristics of IoT malwares and their differences from PC malware are presented to make it clear that both the malwares have different mechanism of propagation and different attack vectors. It is also found that the malware attacks in IoT usually take place in the application and perception layers of the IoT architectures. The main domain of IoT that was discussed in this document was the domain of smart devices such as cameras, lights, etc., and the domain of smart phones. Also, the IoT devices have very less RAM and flash memory due to which a good security architecture cannot be implemented in the IoT devices. It is found from the research that IoT malwares usually attack smart devices with MIPS and ARM based architectures. Further, the IoT malwares have been increasing at a very fast rate since a decade and have compromised millions of smart devices. It is found that there were some prominent and popular IoT malware families that compromised the smart devices and smart phones. These families followed a particular attack vectors to launch the attacks. Furthermore, the anatomy of botnet was discussed and the mechanism used by IoT malware botnet to propagates the attack was discussed. This gave an example of how the malware attacks propagate in the smart devices. Additionally, it is evident from the survey that the concept of IoT malwares is still in its initial stage and we still don't very effective detection and mitigation techniques for the same. But it is found that the best method to find out the evolution of malwares is constructing the phylogenetic trees for different malware families. This gives the connection between two or more malware families that can be helpful in mitigating the malware attacks. This survey paper will be very helpful in getting the idea of IoT malwares right from the basics and their effects in the realm of IoT.

List of References –

ZK Zhang, M Wang, CW Hsu, CK Chen, S Shich, 2014, 'IoT Security: Ongoing Challenges and Research Opportunities', 2014 IEEE 7th International Conference on Service-Oriented Computing and Application, National Chiao Tung University, Taiwan, pp. 230-234, <http://ieeexplore.ieee.org/document/6978614/?reload=true>

YJ Jia, QA Chen, S Wang, A, Rahmati, E Fernandes, ZM Mao, A Prakash, 2017, 'ContexIoT: Towards Providing Contextual Integrity to Appified IoT Platforms', Internet society, University of Michigan, pp. 1-15, <https://www.internetsociety.org/sites/default/files/ndss2017-08-2-jia_slides.pdf>

K Angrishi, 2017, 'Turning Internet of Things(IoT) into Internet of Vulnerabilities (IoV) : IoT Botnets', Turning IoT into IoV: IoT Botnets, <https://www.researchgate.net/publication/313671691_Turning_Internet_of_ThingsIoT_into_Internet_of_Vulnerabilities_IoV_IoT_Botnets>

A Shifa, MN Asghar, M Flcury, 2016, 'Multimedia Security Perspectives in IoT', The Sixth International Conference on Innovative Computing Technology (INTECH 2016), Dublin, pp. 550-555, <http://ieeexplore.ieee.org/document/7845081/>

Mota AR, Ambrosio E, Happa J, Nurse JRC 2016, 'Towards IoT Cybersecurity Modelling: From Malware Analysis Data to IoT system Representation', IEEE, Mexico City, <http://ieeexplore.ieee.org/document/7811597/?reload=true>

Makandar A, Patrot A 2015, 'Malware Analysis and Classification using Artificial Neural Network', IEEE, <http://ieeexplore.ieee.org/document/7492653/?reload=true>

Liu J, Wang Y, Wang Y 2016, 'Inferring Phylogenetic Networks of Malware Families from API Sequences', International Conference oMhn Cyber-Enabled Distributed Computing and Knowledge Discovery, Changsha, pp 14-17, < eeexplore.ieee.org/document/7864197/>

Adeel M, Tokarchuk LN 2011, 'Analysis of Mobile P2P Malware Detection Framework through cabir and Commwarrior Families', IEEE, pp 1335-1343, <http://ieeexplore.ieee.org/document/6113305/?reload=true>

Mohaisen A, Andrew GW, Mankin A 2014, 'Chatter: Classifying Malware Families Using System Event Ordering', IEEE, pp 283-291, <http://ieeexplore.ieee.org/document/6997496/?reload=true>

Chakraborty T, Pierazzi F, Subrahmanian VS 2017, 'EC2: Ensemble Clustering and Classification for Predicting Android Malware Families', IEEE, pp 1-14, <http://ieeexplore.ieee.org/document/8013726/?reload=true>

Hsiao SW, Sun YS, Chen CM 2016, 'Behavior Grouping of Android malware Family', IEEE, <http://ieeexplore.ieee.org/document/7511424/?reload=true>

Yixiang Z, Kang Z 2017, 'Review of iOS Malware Analysis', IEEE, pp 511-514, <http://ieeexplore.ieee.org/document/8005524/?reload=true>

Cimitile A, Mercaldo F, Martinelli, Nardone V, Santone A, Vaglini G 2017, 'Model Checking for Mobile Android Malware Evolution', IEEE, pp 24-28, <http://ieeexplore.ieee.org/document/7967989/?reload=true>

Kambourakis G, Kolias C, Stavrou A, Voas J 2017, 'DDoS in IoT: Mirai and Other Botnets', Cybertrust, pp- 80-84

Tangil GS, Tapiador JE, Lopez PP, Ribagorda A 2014, 'Evolution, Detection and Analysis of Malware for Smart Devices', IEEE Communications surveys and tutorials, pp- 961-970

Bernardi ML, Cimitile M, Martinelli F, Mercaldo F 2017, ' A Fuzzy-based process mining approach for dynamic malware detection', IEEE ANALYSIS MALWARE DETECTION, pp- 1-8

Peng, S, Yu S, Yang A, 2014, 'Smartphone Malware and Its Propagation Modelling: A Survey', IEEE Communications Surveys & Tutorials, vol. 6, no. 2, pp. 927-928, 937.

Ahmad, MA, Woodhead S, Gan, D, 2016, 'The V-Network Testbed for Malware Analysis', IEEE, Proceedings from International Conference on Advance Communication Control and Computing technologies, pp. 629.