University of Salford

School of Computing, Science & Engineering, Salford

MSc by research dissertation report

Optimum parameter machine learning classification and prediction of Internet of Things (IoT) malwares using static malware analysis techniques

By

Syed Usman Shaukat

2018

Supervised By

Dr Mo Saraee

DECLARATION ON CONDUCT OF ASSESSED WORK

Programme	MSc by Research Cyber Security 2016/18
Module	Dissertation
Supervisor	Dr. Mo Saraee
Assignment title	Optimum parameter machine learning classification and prediction of Internet of Things (IoT) malwares using static malware analysis techniques

Briefly, unfair means in assessed work is likely to fall into one or more of the following categories:

- **Plagiarism.**

Plagiarism involves taking the work of another person or source and using it as if it were one's work. Work includes, but is not restricted to, written work, ideas, musical compositions, computer programs, laboratory or survey results, diagrams, graphs, drawings and designs.

- **Collusion.**

Collusion involves working with others on tasks that should be carried out on an individual basis. Collusion should not be confused with collaborative work which is sometimes used as a means of learning. I will be clearly stated when collaborative work is permitted in an assessment. Unless advised otherwise, any work which is submitted for assessment must be produced by individual students.

- **Falsifying experimental or other investigative results.**

This could involve a range of things that make it appear that information has been collected by scientific investigation, the compilation of questionnaire results, etc. whereas in reality it has been made up or altered to provide a more favourable result.

- **Contracting another to write a piece of assessed work.**

This involves any means whereby a person does work on behalf of another. It includes assessments done for someone else in full or in part by a fellow student, a friend or family member. It includes sitting an examination for someone else. It also covers obtaining material from internet 'cheat sites' or other sources of work. Penalties for this type of unfair means are likely to apply both to a student who does work on behalf of another and one who has work done for him/her.

I declare that

- This work is my own
- If this is a group project, each student has contributed to the work in accordance with the set criteria
- The work of others used in its completion has been duly acknowledged
- Experimental or other investigative results have not been falsified
- I have read and understood the Academic Misconduct Procedure*

***http://www.salford.ac.uk/__data/assets/pdf_file/0005/653648/AcademicMisconductProcedure.pdf**

It is the student's responsibility to be aware of this policy and procedure.

Signature 

Name **Syed Usman Shaukat**

ID Number: [redacted] **Date: 28/06/2018**

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

Abstract

Application of machine learning in the field of malware analysis is not a new concept, there have been lots of researches done on the classification of malware in android and windows environments. However, when it comes to malware analysis in the internet of things (IoT), it still requires work to be done. IoT was not designed to keeping security/privacy under consideration. Therefore, this area is full of research challenges. This study seeks to evaluate important machine learning classifiers like Support Vector Machines, Neural Network, Random Forest, Decision Trees, Naive Bayes, Bayesian Network, etc. and proposes a framework to utilize static feature extraction and selection processes highlight issues like over-fitting and generalization of classifiers to get an optimized algorithm with better performance. For background study, we used systematic literature review to find out research gaps in IoT, presented malware as a big challenge for IoT and the reasons for applying malware analysis targeting IoT devices and finally perform classification on malware dataset. The classification process used was applied on three different datasets containing file header, program header and section headers as features. Preliminary results show the accuracy of over 90% on file header, program header, and section headers. The scope of this document just discusses these results as initial results and still require some issues to be addressed which may effect on the performance measures.

Acknowledgments

I would like to thank everyone who has stood by me through this journey. My gratitude goes to Dr. Mo Saraee mainly. Thanks for your patient, guidance, serenity, and encouragement.

I am thankful to the staff at Salford Energy Hub for their assistance and guidance about energy devices.

To my family (My wife Bushra Javed) for all the constant support and encouragement.

Finally, to Allah God Almighty for giving life and everything I am today. You are just as great as your name.

Chapter 1: Introduction

In the last few years, the internet has started playing a vital role in everyone's life when the concept of "Internet of Things" (IoT) emerged. IoT has fundamentally changed the paradigm of traditional information technology (IT) to a new era of innovation involving sensor chips. In a concise time, the current conceptualization of IoT led to deploying over 9 billion interconnected devices in the market; this figure is expected to rise to 24 billion in 2020, it seems like nearly everything will become connected in future for effective communication (Gubbi et al., 2013). With the increased connectivity of the internet with all the luxuries, IoT devices bring the concept of security to the consumer (e.g., controlling home automation). This development of technology makes their hectic life enjoyable and under controlled. For instances, the idea of the customized theme in smart-lights is a best example of an IoT device, another concept of smart fridges with content management not only gives the consumer a useful real-time experience, but it also provides a flexible solution for this busy generation. Ease of life made IoT as an inevitable reality of life what we could only think in the early 80's. Now researchers claim that IoT devices are the first choice of the consumers. Therefore, an enormous number of IoT devices are getting developed, according to a survey the total number of smart IoT devices are expected to reach the figure of approximately 50 billion in 2020 and an estimate of over 75 billion in 2025. With the rise of IoT associated insecurities have also evolved rapidly in the form of malware, denial of service attacks (DoS) and botnets. Figure 1 shows the evolution of famous malicious attempts (Nordrum, Aug 2016, IHS, 2016).

In recent years, cybercriminals have developed malware targeting these IoT devices. Imagine! Your smart fridge sends fake content update requests to your local Tesco, you receive a bill that you were not supposed to pay, or your smart home meter has been infected by malware which is sending wrong meter reading, it would cause insecurity to the users and lack of trust to the product or service provider. Different kind of malware targeting IoT device, Figure 2 summarises some common malware attacks.

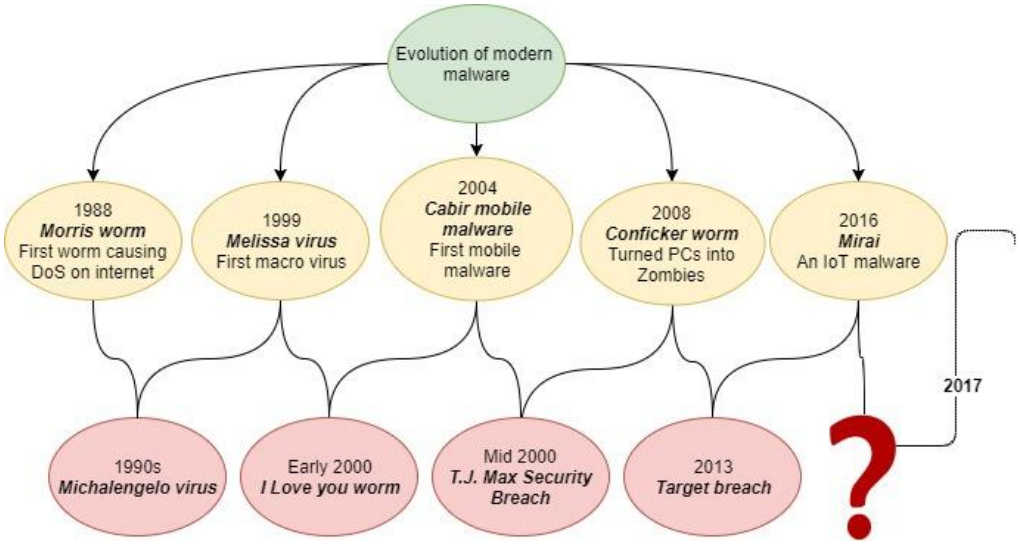


Figure 1: History of malware

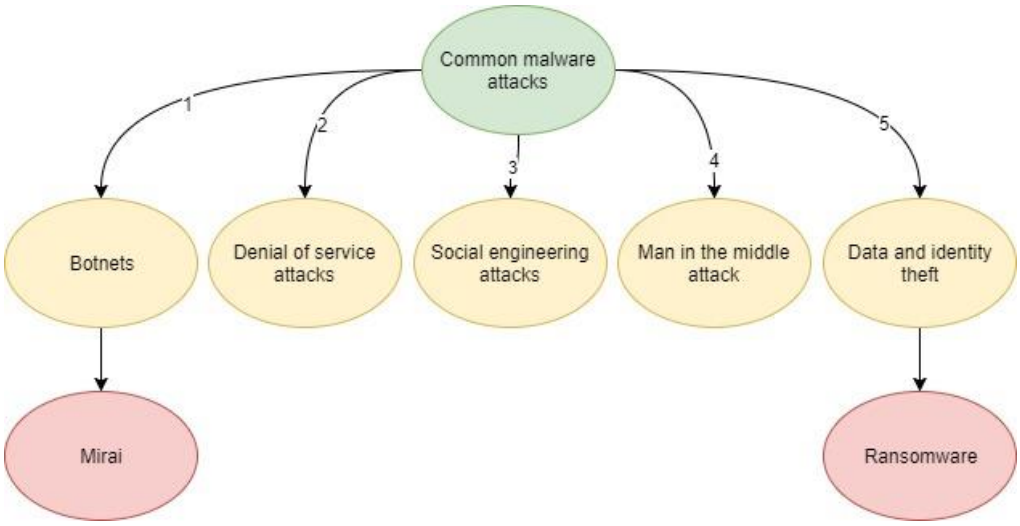


Figure 2: Common malware attacks

Mirai is a recent example of IoT malware which converts its target into a Botnet. Mirai got the focus of researcher's attention for targeting more than 148k IoT devices, making CCTV cameras and DVR recorders running on ports 23 and 2323 particularly and turning them into Botnets to cause Distributed Denial of Service Attack (Angrishi, 2017). To complicate the situation, malware authors used techniques to camouflage the primary activity of a malware causing more damage to the infected devices. Therefore, there is a significant need to employ malware detection techniques which incorporate Data Mining and Machine Learning Algorithms to enhance the detection capability.

Over the last one decade, plenty of research has been conducted in the field of malware analysis; all proposed techniques had some limitations leaving billions of devices still vulnerable to new

malware. Therefore, we need research which predominantly focuses on enhancing automated malware analysis using potentially malicious IoT executables that detect their malicious activities with the help of classification and prediction for better performance.

1.1 - Problem statement

Malware have been threatening the privacy of internet users for a very long time, with every day passing cybercriminals are using advanced programming skills to create more destructive malware. A malware of any class contributes to the revenue loss to the society. Recently, in a study, it was claimed that in 2016 the famous attack on Dyn (Oracle's infrastructure service provider company specializing in DNS and Email related services) by Mirai Botnet caused a revenue loss of $110 million (Kochetkova, 2016). The damage did not stop here, author of the malware released its source code on GitHub leaving the doors open for more powerful attacks.

IoT malwares have special features that enable them to run on huge range of architectures and target multiple platforms. Of course, this special feature makes the IoT malware a smarter form of malware, however, IoT malwares also take the advantage of weaknesses present in IoT devices such as firmware loopholes, no encryption mechanism etc. and exploit these weakness in a clever way. These malwares have an ability to scan heterogeneous devices using open ports, hostile these devices by applying various commands /or saved default usernames / or passwords in their database. Furthermore, another feature that contribute to make it different from other malwares is the instruction set to avoid IPs of Government Agencies.

Billions of IoT devices are already present in the market on vast number of architectures. Main weak point of these devices is their resource constrained nature. Due to rapid growth in their importance and the weaknesses present, these devices soon became a prime target of malware attacks. It has also been noted from published work that there have been numerous ongoing studies done in Windows and Mobile devices related malware mainly involving Android. Unfortunately, this area did not get much attention and till this date there is no published literature available to analyse malware with respect to IoT. This research work attempts to systematically review the literature to study the vulnerabilities reported in IoT and how these are related to malware attacks. This work also aims to analyse IoT malware and work focused on various distinct directions of malware analysis. First such direction was the collection of IoT malware/goodware. Second was extraction of meta-information, headers, strings, and symbols etc. Another direction was classification of malware with the help of machine learning. The

study gives us insight useful information about the patterns of IoT malware/goodware, classification and prediction.

Furthermore, there is no optimum framework for the classification of IoT application. Additionally, no work has been done to observe generalization of classifiers, hyper tuning the parameters (to figure out best parameter for a chosen algorithm) and analyze whether a particular classifier is overfitting or not? These factors motivate us to carry out work in this domain.

1.2 - Research aims and objectives

Based on the literature reviewed it is clear that further studies into IoT malware are needed. The proposed project will investigate the published literature through the bibliometric approach and summarize the gaps reported in the previous section of the problem statement, IoT environments suffer malware detection, classification, prediction and pattern recognition. Meanwhile, we can also observe the lack of application of feature scoring algorithms which can be used for feature selection. Here, I will not discuss the individual status of static or dynamic features of IoT malware; both require comprehensive research. In the light of this, the dissertation has three overriding objectives:

1. **Malware as a severe threat to IoT:** To review the literature to identify what kind of threats and vulnerabilities being faced by IoT devices. And how it becomes security, privacy, and trust related concern.
2. **Systematically review the literature:** To capitalize on public available malware detection tools and libraries of machine learning algorithms for data mining to support annotation of critical malware in IoT for which no information is available.
3. **IoT malware analysis:** To conduct IoT malware analysis using selected classifiers (shown in figure 3) and by observing optimum parameters. Sub-objectives are following:
   a) To study the results using a minimum number of features determined with the help of feature reduction/scoring techniques.
   b) To evaluation different feature scoring techniques and choosing best feature selection/reduction method for our datasets.
   c) To analyze the performance of the algorithms using false positive rates, overfitting issues and most importantly analyze the generalization of the algorithms.

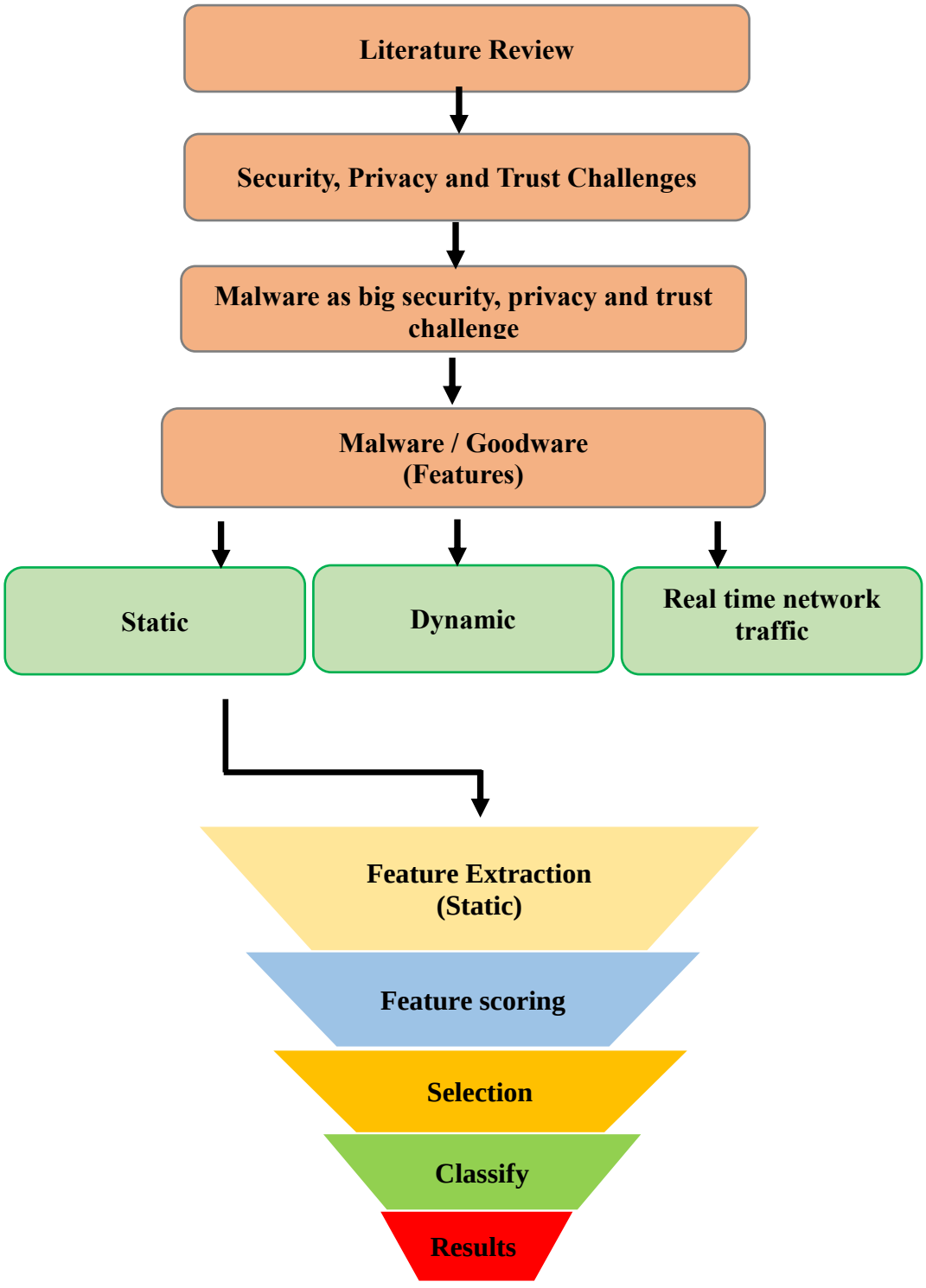


Figure 3: Showing how we reached at the stage of static malware analysis

1.3 - Research contribution

This section discusses the research contributions; figure 4 shows the representation of our research contributions. Supervised learning-based machine learning algorithms help to predict/determine relevant IoT malware features (only static features). These relevant features also help to detect different patterns of the applications being analyzed and create a logical linking between good or bad, predicting the ability of the malware attack on the basis of feature ranking and provide a coherent framework to classify the differences between malware and goodware.

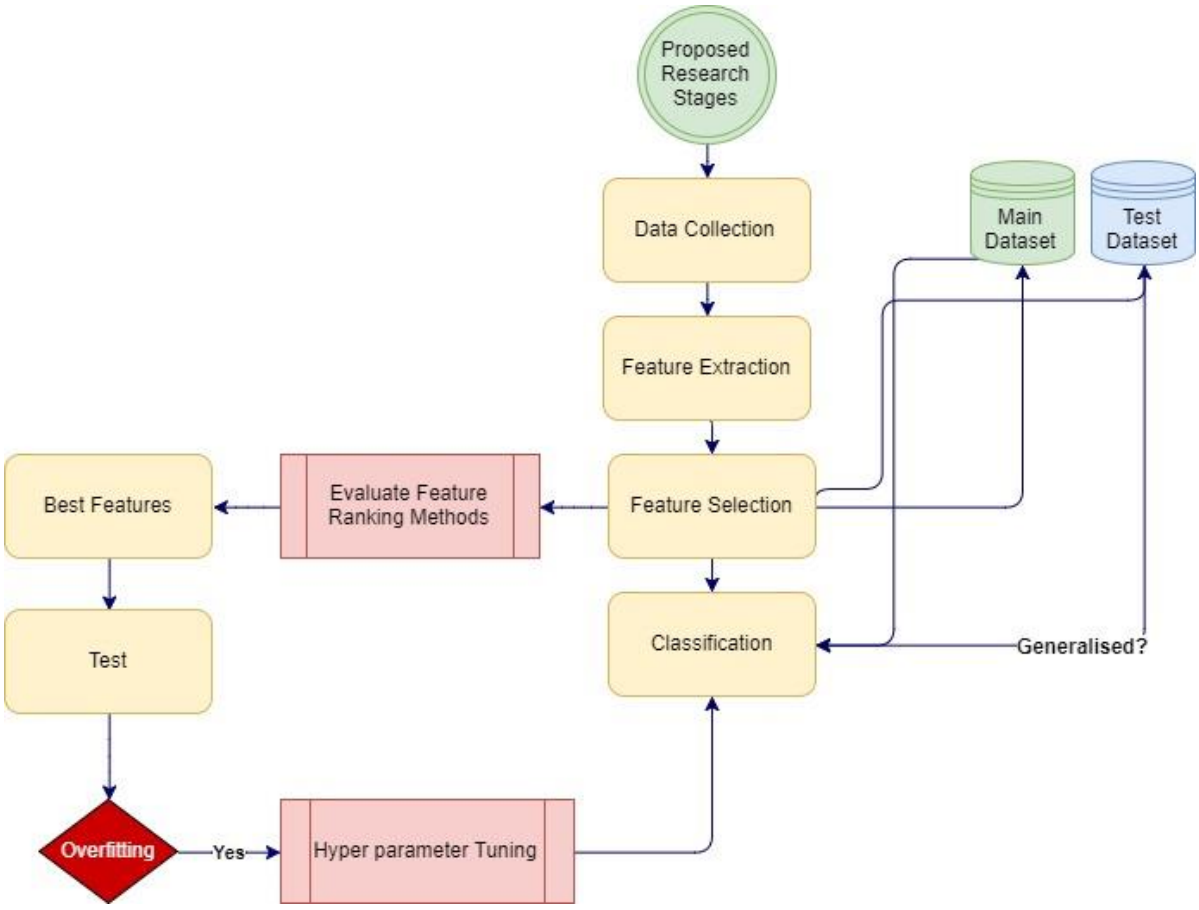


Figure 4: Proposed research stages for static malware analysis with machine learning

This section presents the research contribution and personal commitment. In addition to this, these contributions also help to perform the comparative analysis between IoT malware and goodware (malware belonging to UNIX System V malware family and clean ELF executables collected from Raspberry Pi based Qemu emulator) and also to differentiate between them by a minimum set of features. Therefore, we propose providing an optimum framework for malware classification and evaluation using selected classifiers. In other words, the recommended tool to help better prediction by involving a comparative analysis of various feature selection methods as well. I would explain these research contributions in more detail in later sections.

The rest of the document can be structured as follows: Chapter 2 contains information about malware analysis primarily static malware analysis and background literature review of static malware analysis. Chapter 3 (Data Collection and The Proposed Approach) describes research methodology to achieve research aims and objectives, while CHAPTER FOUR: Research Planning.

CHAPTER 2: LITERATURE REVIEW

Chapter 2. Literature review

The overall aim of the literature review is to perform background study of existing literature related to the internet of things, its related challenges, identify the research gaps to understand the theoretical and conceptual framework of the research in question. The presentation of literature being reviewed in chronological order helps to observe the quality of work, its growth, current gaps in the knowledge and future progress. Therefore, a literature review can be treated as an essential part of the research used to remodel the elements of the work done and create the basis of future work (Webster and Watson, 2002).

2.1 - Internet of Things

In simple words internet of things (IoT) consists of two important elements “internet” and “things.” The name IoT was first used in 1999 in which radio frequency (RF) application extension took place forming the basis of IoT infrastructure (Khodadadi et al., 2017). During early days of the introduction of IoT, RF chips used to be the force behind IoT but later on replaced by wireless sensor chips. There are various definitions of IoT in the literature as shown in table 1.

Table 1: Showing some definitions of IoT

Definition	Ref
"A network of objects which revolves around three pillars of the internet, sensors, and knowledge in such a way that there is always an intersection in the applicability."	(Atzori et al., 2010)
"A network of things, interconnecting with the help of sensors, actuators, data analytics and cloud computing to exchange the information and bring innovative solutions via unified frameworks."	(Gubbi et al., 2013)
"An environment that utilizes the information and enabling technologies to make infrastructure, its components and makes various services interactive, more aware and efficient."	(Bélissent, 2010)
A network of actively interacting, communicating, exchanging information and anonymously behaving things in information and social environments.	(Al-Fuqaha et al., 2015)
A network in which things with unique identities and virtual personalities operating with the help of intelligent interfaces and network protocols to enable them to interconnect and communicate with each other.	(Vermesan et al., 2011, Singh et al., 2014)
IoT can also be defined as a self-configuring global network which is based on protocols that are standard and interoperable. Furthermore, the connected objects have identities/physical attributes and consist of intelligent interfaces incorporated to form an information network.	(Van Kranenburg, 2008, Ray, 2016)

Based on above-mentioned definitions, IoT can be considered as a complex cyber-physical ecosystem with following six characteristics: (i) Dynamic infrastructure; (ii) Self-configuring; (iii) Well integrated; (iv) Interoperable; (v) Identities and physical attributes; (vi) Intelligent Interfaces.

These six characteristics make a "smart device" that is considered as a significant part of the scientific revolution. The tremendous growth in smart IoT devices and their innovative features is self-evident of their importance. IoT technology became more promising when researchers claimed that in the future every ordinary device will be transformed into a smart device and IoT will turn into Internet of Everything (IoE).

Along with the rise of IoT, top technology giants have also started focusing towards it for example Google acquired Nest to step into smart home automation business, Apple introduced HomeKit and Cisco launched various products in the areas of network connectivity, data analytics, embedded systems, security and digital transformation (DevNet) and made a bold statement that IoT market worth at least 14 trillion dollars (Dijkman et al., 2015). This is the beginning as numerous European companies are making their products in the areas of health care, energy and in the vehicle industry, etc. Furthermore, the good thing about IoT is that researchers have started performing quality research in it and trying to address associated challenges of security, privacy and in particular malware threats. As of now, research scientists continue to produce high-quality technical approaches including Robotics, Cloud Computing, Big Data and development of IoT machine learning frameworks.

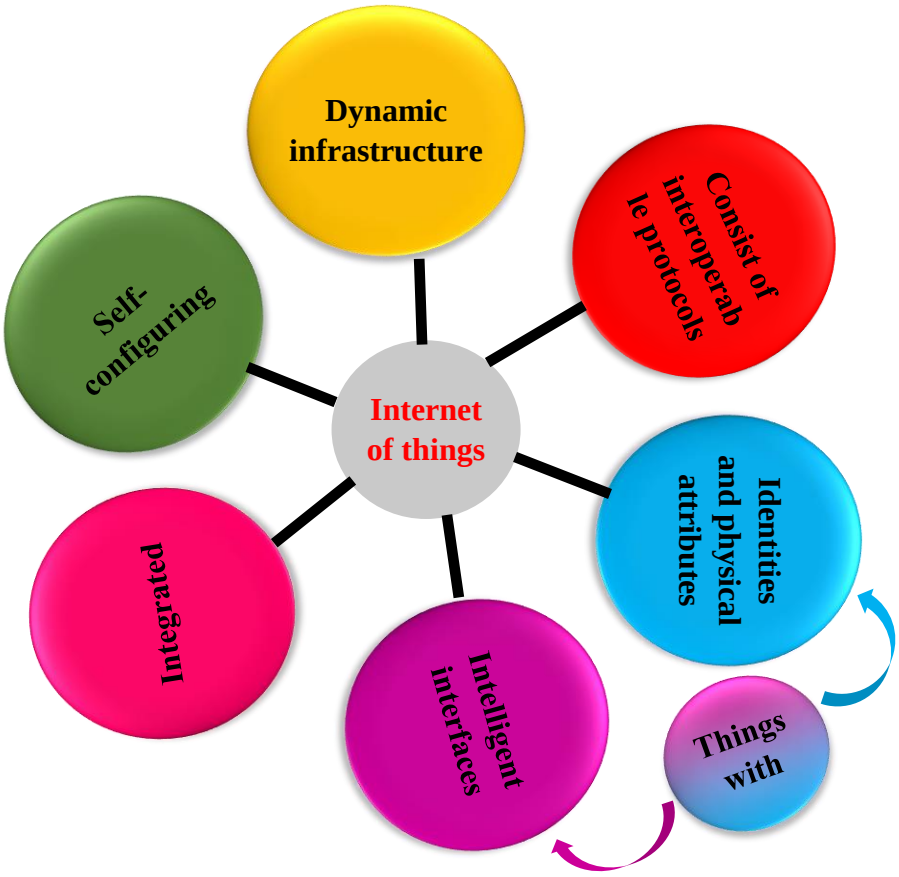


Figure 5: Illustrating the characteristics of IoT

2.2 – Architecture of IoT

Due to billions of heterogeneous IoT devices, all proposed IoT architectures in the literature have considerable variations. There is no single consensus on IoT architecture, and everyone is agree with this statement (Sethi and Sarangi, 2017). Consequently, researchers often described the architecture based on their study purpose.

Generic architecture

The most basic architecture of IoT devices consists of three layers (Figure 1):

(i) **Physical or Perceptive layer or remote sensing layer:** this layer is responsible for sensing and collecting information about the environment. It has the ability to sense the physical parameters or identify the other objects present in the environment.

(ii) **Network or communication layer:** this layer is responsible for the communication by connecting with other smart things, network devices, and servers. This layer provides the means of data transmission by processing the sensor data.

(iii) **Application layer or software layer:** this layer is responsible for communicating with the end-user to deliver various application-specific services and defines various IoT applications that can be deployed.

This underlying architecture was introduced in the early stage of research in the field of IoT (Jammes and Smit, 2005, Yan and Huang, 2009), but this kind of architecture is not sufficient to provide in-depth aspects of IoT for the advanced research (Sethi and Sarangi, 2017). Therefore, detailed, layered architecture has been proposed in the literature where some researchers have also included some additional layers (Figure 1): (i) physical layer; (ii) transport layer: (iii) processing layer; (iv) application layer; and (v) business layer. The roles of physical layer, network layer, and application layer are same as the three-layer architecture, and other layers have been described below (Rayes and Salam, 2017, Sethi and Sarangi, 2017, Bozdogan and Kara, 2015).

(iv) **Transport layer:** transport layer is responsible for transferring the sensor data from the physical layer to the processing layer through networks (such as wireless, 3G, LAN, Bluetooth, RFID, and NFC) to perform network operations and identify the connected devices.

(v) **Processing layers:** processing layer is the middle layer that is responsible for performing multiple functions such as storing data, analyzing data and processing data that comes from the transport layer. Many enabling technologies such as cloud computing, big data processing modules and database management, etc. are incorporated to provide a diverse set of services.

(vi) **Business layers:** business layer is responsible for managing whole IoT system including user details, applications, etc.

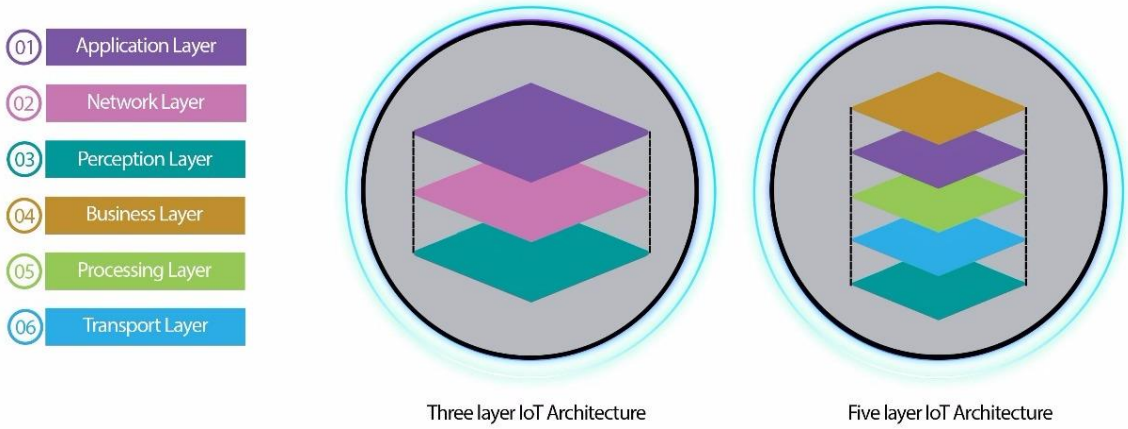


Figure 6: IoT architectural layers

Advanced architecture

In a recent article, another architecture has been proposed which is a somewhat advanced form of IoT. This architecture consists of seven layers, e.g. (i) application layer; (ii) application support and management layer; (iii) service layer; (iv) communication layer; (v) network layer; (vi) hardware layer; (vii) environment layer, etc.

Advanced architecture

In a recent article, another architecture has been proposed which is a somewhat advanced form of IoT. This architecture consists of seven layers, e.g. (i) application layer; (ii) application support and management layer; (iii) service layer; (iv) communication layer; (v) network layer; (vi) hardware layer; (vii) environment layer, etc.

2.3 – IoT environment

The key objectives of IoT devices are (i) To exchange secure and reliable information between connected devices; (ii) To identify the relevant objects; and (iii) To take part as an essential element in ubiquitous/mobile communication. These three objectives collectively form an IoT

environment (IE). The IoT devices with multiple architectural layers, associated protocols and enabling technologies to increase the heterogeneity in the IoT environment. It creates a model where numerous smart things or objects connected with each other by using compatible wired/wireless networks. The multiple interactions bring interoperability to the environment and help to reach common objectives for various IoT user-specific applications (Figure 7).

IoT environment typically consists of six important features (Al-Fuqaha et al., 2015):

(a) Identification: Identification is an essential feature of IoT. There are various techniques to identify a smart thing/object within the IE such as unique object code (present at the hardware level) (Koshizuka and Sakamura, 2010) and unique addresses, e.g. IPv4or IPv6 present at network layers (Al-Fuqaha et al., 2015).

(b) Sensing: In the IoT, sensing feature means communicating with IoT device, gathering data and report it back to the data handling or storing mechanism. Example of sensors/communicator is smart sensors, actuators, and radio frequency tags.

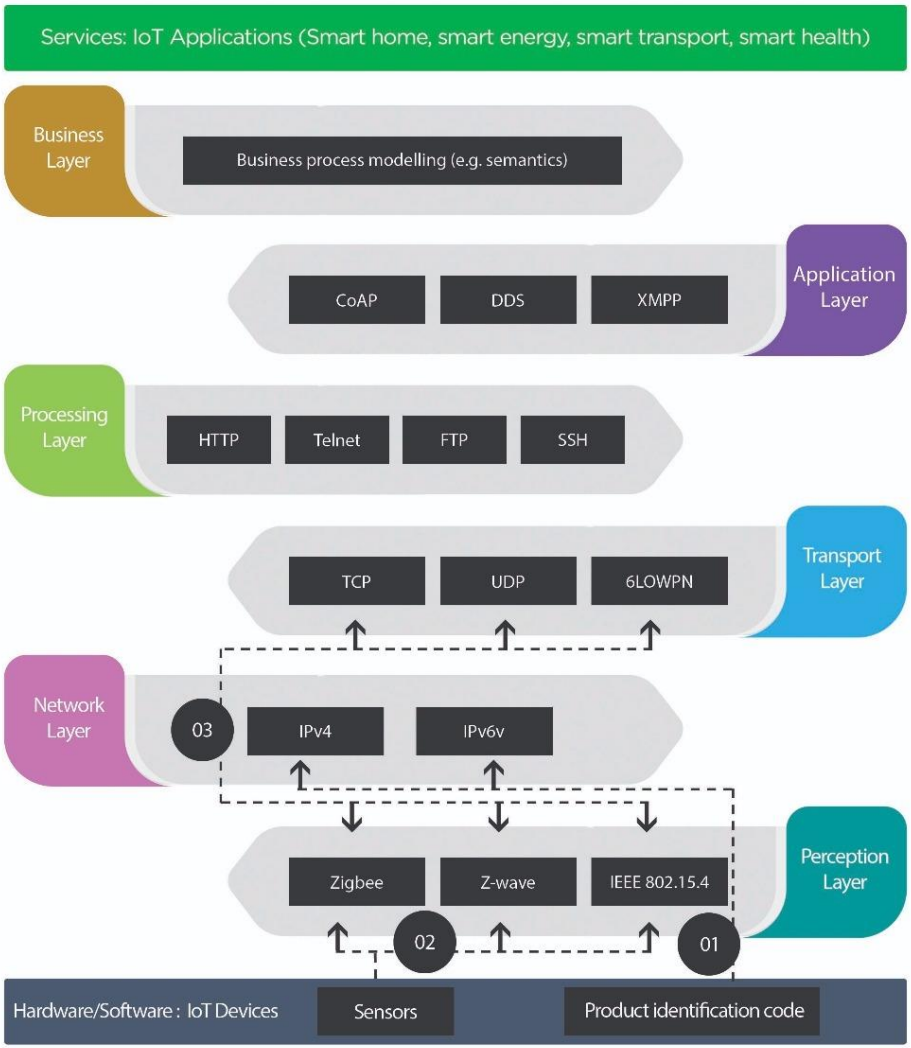
(c) Communication services: These features help IoT devices to connect with each other and provide various services while working in a low power mode and establishing a communication channel to send or receive the information to the sensor. Typical examples of communication services include Zigbee, Z-wave, WIFI etc.

(d) Hardware/software: IoT environment uses various hardware platforms and operating systems software to manage individual IoT devices, the combination of both these referes to as a computation platform or a controller. There are various examples of IoT hardware platforms like Intel Edison, Intel Galileo and Raspberry Pi etc, while Contiki OS, TinyOS etc are the examples of operating systems or softwares.

(e) Services: IoT devices, protocols and enabling technologies work on interoperability framework for providing various services explained in following section.

(f) Semantic: This feature creates an ability to extract information from an IoT device and apply knowledge representation techniques to bring sense into a raw data delivered by the sensor. Semantic services bring intelligence into IoT by using data analytics. Example of IoT semantic is Semantic Web Ontology (SWO)

Figure 7: IoT Environment containing architectural layers and protocols for bringing interoperability among IoT devices and applications where various IE features facilitate this process.



2.4 – IoT ecosystem and the world of connected services

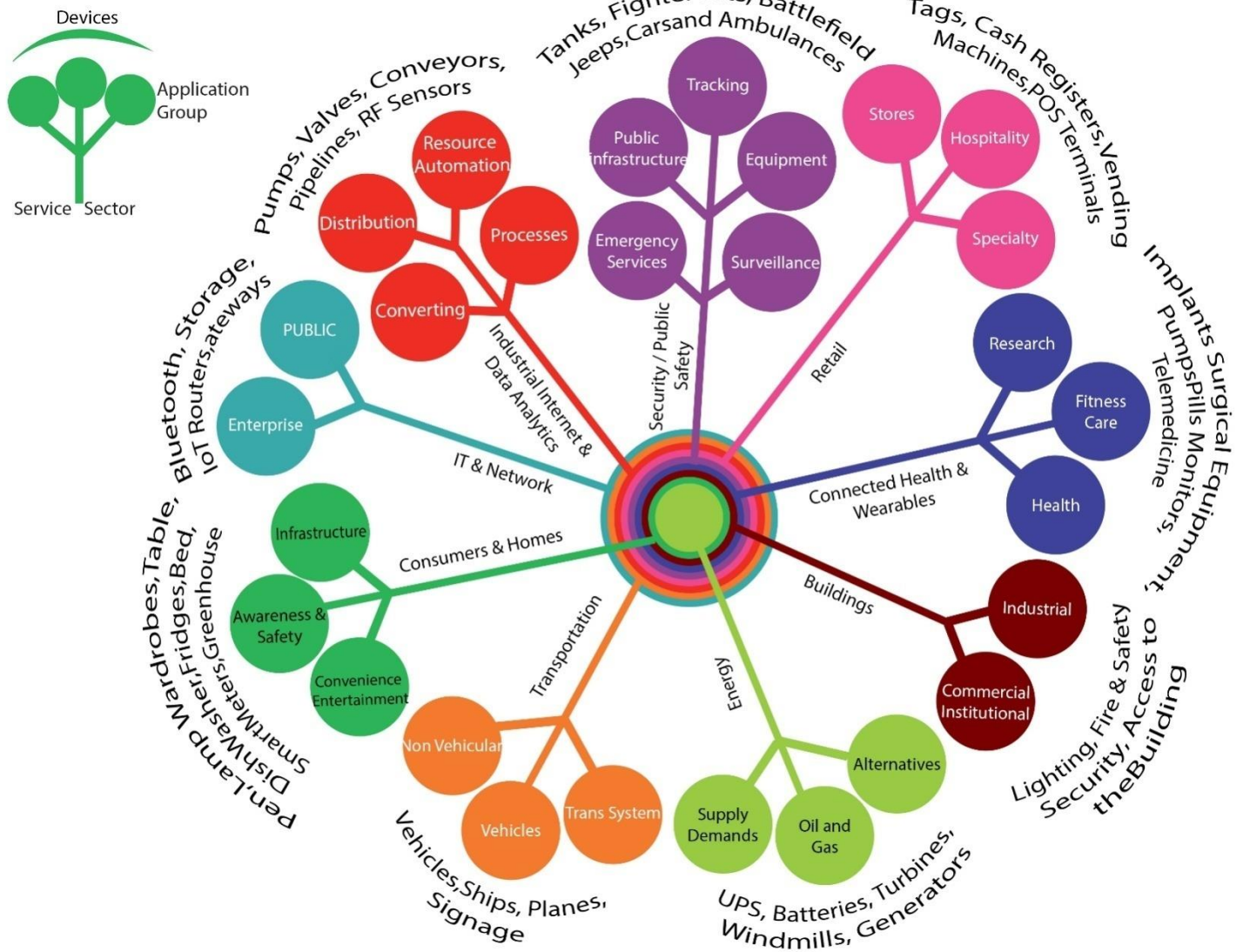
The world of IoT is redefining the relationship between the machines and humans; it allows the automation of everything around us. IoT devices and its enabling technologies are working on interoperability framework where they exchange information with each other and coordinate decisions and making a smart "cyber-physical ecosystem" (ENISA, 2017). This is an innovative ecosystem with billions of heterogeneous physical devices manufactured around the world to improve quality of life, efficiency, productivity, profitability, effectiveness, and decision-making. The application of the IoT devices is diverse, distributed into various services and benefitting ordinary people, industrialists, governments, health professionals, energy providers in simple words everyone. Pictorial view in Figure 8 illustrates the proliferation of devices and most exciting applications within different services.

2.4.1. Smart Energy

Smart energy is the conversion of a traditional energy distribution system that consists of some distribution lines, substations, transformers and Supervisory Control and Data Acquisition (SCADA) field devices and converts into modern networks that are not only smart, intelligent but also having the default capabilities of information exchange, maintainability and easy interoperability of individual components (Sajid et al., 2016). The examples of smart energy include UPS, batteries, generators, fuel cells, ambient energy harvesting, telemetry, power stations, smart grid and power controls, etc. (Fraga-Lamas et al., 2016). Among these smart grid is a popular application of IoT which consists of devices such as meters (gas, electric and water) and other energy appliances. Traditional energy transmission/distribution systems used to be unidirectional, smart grid makes them intelligent to sense the transmission requirements to avoid the congestion, effective communication between the utility services and the customer. Moreover, there are specific advantages smart grid gives to the users, for example, the efficient transmission, congestion control, smarter restoration of energy, less hassle to maintain, controlled peak hour energy supply and better integration of various controls (Min et al., 2014).

Smart grid consists of four major components including power plants (source), transmission (energy transfer in bulk from power source generators and the sub-station), distribution (a connection between power source and the customer), customer area to generate store power at home or anywhere and service provider to deliver the energy products to end-users (Min and Varadharajan, 2015).

Figure 8: The Internet of Things (IoT) - The World of Connected Services



2.4.2. Smart Home

Smart home or smart living is often referred to as Home Area Network (HAN) is the incorporation of communication techniques to form a network that connects the essential home appliances and services like lighting, heating, air conditioning and security in such a way that they can control and monitor these appliances anytime anywhere. These devices often use twisted pair cables and RF/IR sensor chips. Most importantly many of the device use mains power and use a gateway or controller to manage (connection and authentication) for all the devices associated with the home network (Jiang et al., 2004).

There have been numerous smart home devices in the market these days; prominent "smart" home devices include pen, wardrobes, table, lamp, picture frame, fridges/freezers, bed, pillow, digital cameras, power system, dishwasher, e-readers, smart utility meters, greenhouses and home surveillance system (Alvarez et al., 2017).

2.4.3. Smart Buildings

Incorporating energy management system (EMS) and security/safety systems within the building is the key aspect of smart building. It is useful for competitive management of resources, improve building visibility and manageability. It provides tools to reduce the operational cost and provide cost-effective benefits to the consumers, employees, and tenants. It also brings intelligent IoT ecosystem that includes sensor and gateway vendors, system integrators and application developers (Shenoy, 2016).

2.4.4. Smart Health

IoT being a game-changer in every industry playing its role in the healthcare sector as well by transforming healthcare into smart healthcare or connected health in which all medical appliances are always connected to give more useful and important information about patients. The healthcare market is going to hit 117 billion USD in 2020. With the help of connected health we may get the benefit of efficient risk analysis and healthcare asset management, drug management, monitoring for patients and hospitals e.g., smart pills a major contribution towards smart healthcare in which patient's clinical trials, activity monitoring, and self-reporting is performed and finally early medical intervention for critically ill patients (Patel et al., 2017, Weinberg et al., 2015).

2.4.5. Smart Wearables

Smart wearables include medical appliances, fitness equipment, and smartwatches help to monitor not only the health of the patients but also lets the athletes to maintain their routine fitness chart. According to Cisco till 2020, there will be at least 600 million smart-wearables used around the globe (Sun et al., 2017a). Smart wearables play an essential role in IoT world, according to the researchers, the reliance of people-centric aspect of IoT is heavily on these wearables to handle remote objects (Liu and Sun, 2016). There have been various products in the market, e.g., Samsung Gear and Fitbit, etc. are the widely used products. These intelligent devices have low-constrained architecture particularly the sensors.

2.4.6. Smart Security and defense

Various IoT devices are getting deployed throughout the cities that are undeniably transforming the public safety aspects. Wherever unplanned, emergency events and catastrophic disasters occur in the cities, these devices enable the interoperability and transform the critical information to the organization who deals with scenarios in which defense and public safety could influence and respond to emergency events (Fraga-Lamas et al., 2016).

2.4.7. Smart Retail

Another implementation of IoT concepts where analytical abilities, predicted outcomes and efficient results help retail industry. Smart retail includes supply chain, in-store applications, and customer specific applications as well. It helps to identify when the certainty customer needs help. Furthermore, smart retail also helps (i) identifying when there a maintenance required on a retail machine, (ii) transportation of merchandise by intelligently optimizing the route, tracking and temperature control, (iii) making retail warehouses automated by monitoring sales, stock levels and smart pallets which automatically report missing stock, (iv) proactive customer focus to identify when a customer needs an incentive as a highly valued customer retention scheme or autonomous doorstep product delivery and (v) to help to monitor automatic foot count for retail stores, analyse that information with other stores and re-modelling store to maintain customers (SAS).

2.4.8. Smart Industries

The concept of smart industries or industrial internet (II) was introduced by a company called GE, according to this concept complex machinery is used along with RF sensors and software forming a specialized IoT environment (Greenough and Camhi, 2015). II is a complex environment which widely uses machine learning, big data, and both homogeneous and

heterogeneous device communication as enabling technologies giving high-performance analytics to the users (Kevin, 2009, Jeff Kelly, 2013). With the evolution in IoT, data from the government is getting readily available to relevant people and with the help of industrial internet and big data analytics people can handle raw information and transform into more structured and valuable form of knowledge (Lohr, 2012). The examples of industrial devices ready for transformation into II include Pumps, Valves, Conveyors, Pipelines, Motors, and RF Sensors, etc.

2.4.9. Smart Transportations

In recent years, some cities have become more crowded, and millions of people commute through public transports on a daily basis. Integration of IoT into transportation system helps to optimize public transportation routes, assess congestions and plan safer roads for the journey by avoiding traffic congestions and reduce infrastructure costs (Gubbi et al., 2013)

Connected cars are another important IoT applications, vehicles connected with their gateways with the help of WLANs providing drivers benefits, e.g. automatic breakdown support, location services, driver assistance, entertainment, eHealth and fitness and advanced road traffic assistance in case of an incident (Kirk, 2015). Recently the CEO of Apple Tim Cook confirmed about the work on self-driving cars and the ability of Apple mobile devices to control the vehicles (Harris, 2015) which may give IoT another dimension to enable car to car communications, interaction with smart traffic lights, and most important connection with external access points (Bonomi et al., 2012). By 2020, it has been reported that approximately 75% of cars worldwide will be IoT enabled (Javed et al., 2018). For the public transportation, the concept of smart taxi system has also been introduced. Furthermore, another use of smart transportation system can be in train services by providing smart ticket for data collection & analytics, the management of public safety/security and the inclusion of smart tablets with the drivers (Zanella et al., 2014).

2.4.10. Smart IT & Network

IoT is constantly offering new devices and tools that helps to interact and connect IT and network professionals to perform network administration, monitoring network traffic, status of nodes connected and software updates etc. to assess the status of working and delivering better solutions. Furthermore, IoT also helps network support teams to perform better network fault finding and tolerance, handle cloud services, effective network usage and manage organizational assets (Fraga-Lamas et al., 2016).

2.4.11. Smart Cities

A smart city is a new vision of the technology that incorporates all infrastructures of a city into a controllable network. The infrastructure of a smart city may include almost everything for example homes, schools, streets (lights), hospitals, transportation system, gas/electric/water supplies and much more (Zanella et al., 2014). All of these technologies use different sort of gateways depending on the scope and use, but one thing in the smart city in comparison with other application is the use of cloud services for information exchange. With the advancement of IoT smart city has been emerged as an essential concept to bring comfort to the lives of residents of the town. Not only this, this idea has become a hub of other emerging technologies as well like data science (big data, malware threat analysis, and information governance) but also a complex challenge of heterogeneous and scalable computing challenges (Zhang et al., 2017, Schaffers et al., 2011, Hernández-Muñoz et al., 2011).

2.4.12. Smart Supply Chain

IoT is bringing revolutionary advancements into supply chains and not only shaping up the industry but also solving the majority of the problems consumers face. With the rapid expansion in the business everyday organizations require smart solutions that make companies connect their systems efficiently, communicate with different businesses, share information and reinforce their sales/supply chains departments. Not only this, it can help customers in placement, delivery, and tracking of orders. In addition to this, IoT can help consumers of supply chain products by introducing smart labeling system to give total control of the products. Therefore, with the help of IoT sensor chips, every aspect of supply chains can be controlled to provide efficient service (Javed et al., 2018, Kärkkäinen, 2003, McFarlane and Sheffi, 2003).

2.4.13. Smart Agriculture

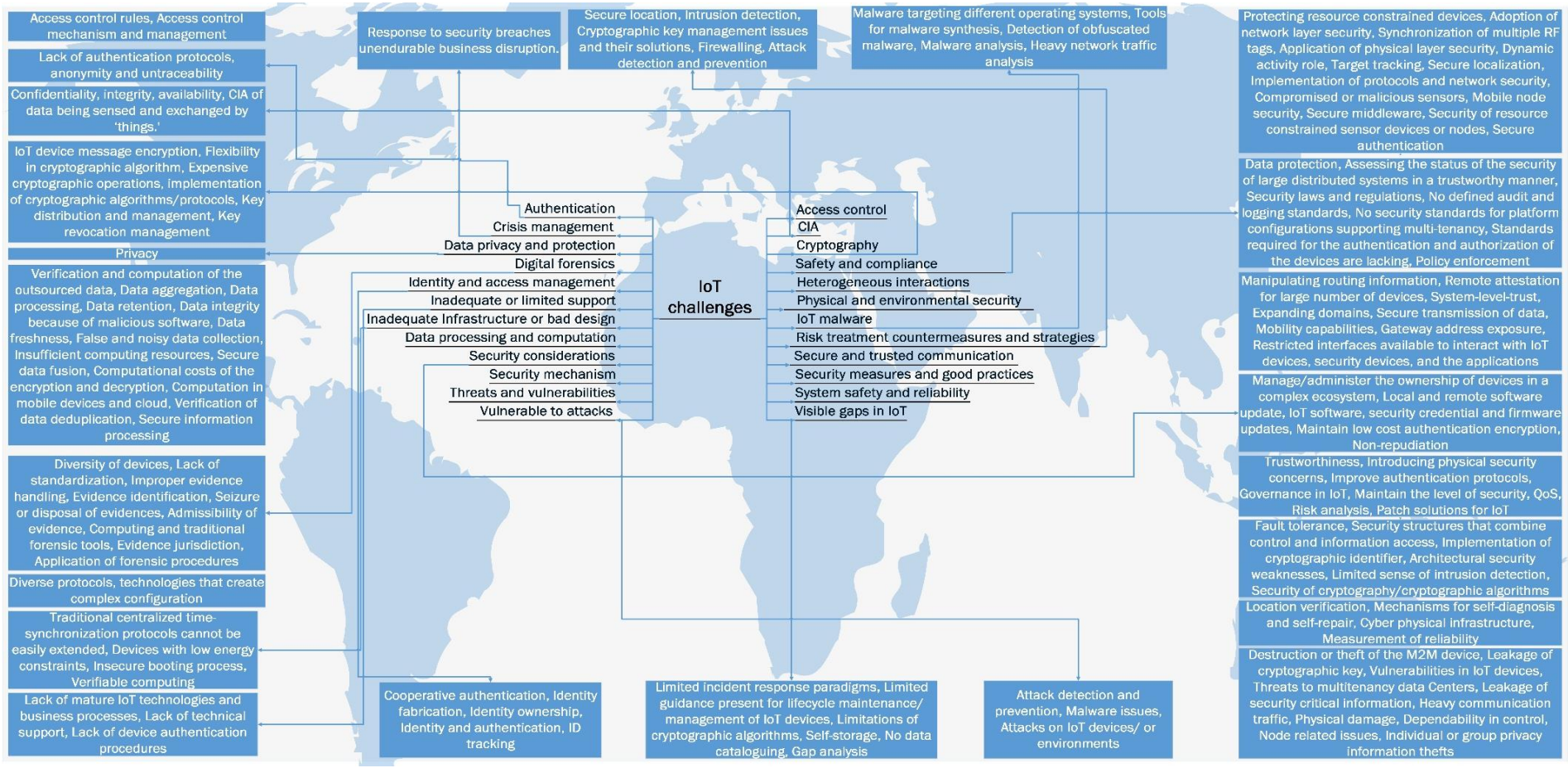
In the last few decades, climate has been changed drastically that brings various challenges for local and global food security. Agricultural commodities are highly sensitive to climatic conditions such as temperature, rain, and humidity. These changes are likely to affect agriculture production (crops, fruits and vegetables, etc.) and livestock (fisheries, poultry farming, etc.). In this situation, it is prerequisite to monitor the climate change and utilize all available resources for sustainable agriculture production by monitoring and better resource utilization effectively which is only possible with the help of smart IoT devices (Javed et al., 2018, Na and Isaac, 2016, Nukala et al., 2016). Smart agriculture system is an automated concept that has been recently introduced and getting worldwide attention. Thanks to IoT that

is making a significant contribution towards a better agriculture system people were looking for (Patil and Kale, 2016).

2.5 – Security and privacy challenges related to the IoT

IoT is connecting more and more devices every day; this emerging technology promises to provide access to the devices anywhere to carry out everyday tasks using different IoT applications discussed in the previous section. This advancement offers undoubtedly several benefits to the humanity. Although IoT is playing a transformational role in the lives of people, on the other hand, it also brings a large number of challenges. Security, privacy, and trust are probably the most challenging issues in IoT, and various authors have extensively discussed them in published literature. In this section, we tried to collate a list of most important challenges reported in the literature; these challenges were divided into 24 key groups defined. This grouped taxonomy is depicted in figure 9, table 2 summarising the challenges and followed by an overview of each group defined.

Figure 9: Summary of IoT security, privacy and trust challenges



2.5.1. Access control:

Access control let only authorized users to access a resource. It enables to control the software update, data sharing, maintenance and protection of sensitive data. Access control usually consists of three important building blocks: (i) access control mechanism; (ii) access control rules and (iii) access control management (Ouaddah et al., 2017). Due to the low power requirements and low constrained environment, access control is one of the major challenges in IoT (Alrawais et al., 2017a, Lin and Bergmann, 2016, Roman et al., 2013, Sicari et al., 2015, Pirbhulal et al., 2017, Yu et al., 2017, Tiburski et al., 2015, Ouaddah et al., 2017). Rules/policies/privileges for the access control are not well defined for most of the IoT devices. These rules are usually implemented on a high-level architecture which is ineffective due to resource constraints in most of the IoT devices. Without clear access control rules, access to the IoT devices get compromised that leads to more sophisticated attacks (Yaqoob et al., 2017, Ouaddah et al., 2017).

Table 2: Showing security, privacy and trust challenges:

Group	Challenge	Finding and comments	References
Access control	Access control rules	Effective implementation of access control is only possible with the help of predefined rules, policies or privileges that are not well defined for IoT devices. Without clear access control rules, access to the IoT devices may get compromised.	(Yaqoob et al., 2017, Zarpelão et al., 2017, Ouaddah et al., 2017)
	Access control mechanism and management	Application of adequate access control is a critical element of information security that requires comprehensive work on its mechanisms and management system. Due to the low power and low constrained environment, access control is one of the major challenges in IoT.	(Du and Chen, 2008, Zarpelão et al., 2017, Sicari et al., 2015, Ouaddah et al., 2017, Miorandi et al., 2012, Alrawais et al., 2017a, Pirbhulal et al., 2017)
Authentication	Lack of authentication protocols	A large number of heterogeneous IoT devices being manufactured, their diverse protocols, poor architecture, and complex configuration makes implementation of authentication protocols a challenge. This deficiency causes more complex security and privacy issues.	(Alrawais et al., 2017a, Amadeo et al., 2016, Zhang et al., 2013, Al-Fuqaha et al., 2015, Tiburski et al., 2015, Venckauskas et al., 2016b, Pirbhulal et al., 2017, Sicari et al., 2015)

	Anonymity and untraceability	Heterogenous IoT devices are developed and distributed in the dynamic environment. The big challenge lies in the design and development of secure and privacy-preserving services that are not well-defined. Personally identifiable information has maximum disclosure due to *anything*, *anytime* and *anywhere* nature of the IoT that raises issues of anonymity and untraceability. Here anonymity refers to something that nobody knows about yourself or your real identity, while the untraceability refers that no one can predict your actions.	(Gope and Hwang, 2015, Challa et al., 2017)
CIA	Confidentiality	Confidentiality or privacy is a severe issue in IoT because of many reasons, e.g. the massive amount of data/traffic being generated, and the ineffectiveness of security controls, etc. The design and build of these devices have a limited sense of hiding the sensitive information from unauthorized people to view it.	(Mendez et al., 2017, Tiburski et al., 2015, Liu and Sun, 2016, Lin and Bergmann, 2016, Pirbhulal et al., 2017, Maple, 2017, Sicari et al., 2015)
	Integrity	The issues of IoT like fault tolerance, malware attacks and untrusted communication effect integrity of devices resulting in physical damage or unavailability of the resources.	(Mendez et al., 2017, Tiburski et al., 2015, Liu and Sun, 2016, Lin and Bergmann, 2016, Pirbhulal et al., 2017,

			Maple, 2017, Sicari et al., 2015, Juma et al., 2008, Kwon et al., 2016)
	Availability	Availability makes sure that resources are available for the use when needed. The availability of IoT resources becomes an issue when confidentiality and integrity of objects are compromised (e.g., malicious attacks etc.). It directly or indirectly affects CIA where availability of resources becomes a big challenge.	(Mendez et al., 2017, Pirbhulal et al., 2017, Maple, 2017, Sicari et al., 2015)
	CIA of data being sensed and exchanged by 'things.'	CIA of data being sensed and exchanged means (i) a node is free from malware; (ii) no unconcerned party has access to data generated or stored; (iii) maintain the reliability and privacy of communication where sometimes both relevant and irrelevant information is kept as well. maintaining (i), (ii) and (iii) simultaneously without compromising CIA is a serious challenge.	(Mayer, 2009, Liu and Sun, 2016, Juma et al., 2008)
Crisis management	Response to security breaches unendurable business disruption.	Maintaining uninterrupted and safe operation, even when the system is compromised is the highest priority target for the IoT industry.	(Chiang and Zhang, 2016)
Cryptography	IoT device message encryption	A massive number of IoT devices increase network vulnerability. A process of encoding a message/information in a way that only authorized parties can access is a big challenge.	(Al-Fuqaha et al., 2015)

	Flexibility in a cryptographic algorithm	Most of the IoT devices were designed without considering security and privacy. Moreover, IoT lacks flexibility in hardware operations and only support limited operations. The flexible cryptographic operations can give more support to hardware and offer better protection. These algorithms have not been implemented so far and require many efforts.	(Amadeo et al., 2016, Ambrosin et al., 2016)
	Expensive cryptographic operations	Cryptographic operations often require much more resources and computation power to implement which is a challenge in IoT.	(Li et al., 2014a, Zhou et al., 2017)
	Implementation of cryptographic algorithms/protocols	Cryptography techniques of storing/transmitting the data are concerned due to the architectural limitation of IoT. IoT devices are based on either 8 or 16-bit architectures, and implementation of cryptographic algorithms for getting the right security is a challenge.	(Ning et al., 2015, Venckauskas et al., 2016a, Roman et al., 2013)
	Key distribution and management	The distribution and management of cryptographic keys is a critical issue when integrating cryptographic algorithms. If these keys are compromised then entire communication process may be disturbed. So there is a need to store keys at a safe/centralized location and distribute them when needed.	(Xiao et al., 2017, Alrawais et al., 2017a, Bu et al., 2017, Ciccozzi et al., 2017, Venckauskas et al., 2016a, Chandramouli et al., 2014)

	Key revocation management	Cryptographic key management becomes an issue in IoT due to the lack of key revocation techniques that may allow cybercriminals to utilize the keys obtained in the process of a system breach.	(Ambrosin et al., 2016, Touati and Challal, 2015, Chandramouli et al., 2014, Sawand et al., 2015)
Data privacy and protection	Privacy (General, attacks, preservation, and privacy-preserving data mining)	Data privacy is a major concern for the people which always requires preservation. The unrestricted access to the data poses significant security and privacy risks to consumers. The critical data privacy-challenges are: (i) most IoT devices fail to encrypt data that are being transferred; (ii) user sensitive information can be compromised due to unencrypted data.	(Barki et al., 2016b, Mayer, 2009, Sicari et al., 2015, Wang et al., 2014, Liu and Sun, 2016, Ning et al., 2015, Pirbhulal et al., 2017)
Data processing and computation	Verification and computation of the outsourced data	The process of outsourced data (a data produced or governed by another company) to the cloud to perform computational operations, and then request results may lead to security problems, e.g., password crack and DoS/DDoS attacks. Non-verified outsourced data in IoT lead to duplication of data or opens up doors for further complexities.	(Yu et al., 2017, Liu et al., 2015)
	Data aggregation	One of the biggest challenges in IoT is the gathering of unprecedented data generated from a multitude of devices every second. Due to multiple related issues with IoT like heterogeneity, complexity, an ever-increasing number of devices, data	(Pandey et al., 2010, Sawand et al., 2015, Luong et al., 2016)

		aggregation is an issue where both active and passive attacks, eavesdropping, lack of confidentiality/integrity/trust, etc. play an essential role to degrade the value of the precious data collected. Moreover, computation is more difficult when noisy or duplicated data is also present in the collection.	
	Data processing	A huge amount of unprecedented IoT data processing (i.e., acquiring and managing) is a challenge for the data analytics particularly when data is coming from multiple sensors, devices of complex configurations and various vendors, from outsourced companies, etc. The processing of massive data, elimination of ambiguities, noise and deduplication for processing without violating/harming data confidentiality and integrity is a very tough job. Processing or manipulation of data in this complex environment becomes an issue when data comes from edge devices in which computation and handling is a very tough challenge.	(Mineraud et al., 2016, Alrawais et al., 2017a, Gaona-Garcia et al., 2017, Luong et al., 2016)
	Data retention	IoT devices generate a massive amount of data every day and preservation of that data for continued storage, for compliance or business reasons is a nightmare for organizations. Due to undefined/agreed governance laws defined by the government to	(Kumarage et al., 2016, Rose et al., 2015)

		help legal matters related to IoT, it is a challenge for the businesses to maintain a set standard for data retention.	
	Data integrity because of malicious software	Data integrity is a fundamental aspect of IoT security and reliability; a malware may cause issues related to data integrity in which an attacker may gain administrative permissions to make changes in the environment.	(Kwon et al., 2016, Mendez et al., 2017)
	Data freshness	Dealing with massive amount of data is a big challenge not only to store recent /relevant data without any adversaries replayed old messages but also manage the uncertainties in the data as well.	(Pirbhulal et al., 2017, Jing et al., 2014, Islam et al., 2015, Chen et al., 2009)
	False and noisy data collection	Malicious attacks may cause issues (such as hardware failures or unreliable communication etc.). As a result IoT sensors may give noisy/false data. The resolution of this problem is an open issue.	(Sawand et al., 2015, Mavromoustakis et al., 2016, Chen et al., 2015)
	Insufficient computing resources	IoT devices have limited computing resources especially when considering IoT enabled medical devices where authentication schemes employ complex algorithms that require more computational resources. The techniques that focus on the need to do any computation of data with limited resources are challenging to apply.	(Yasin et al., 2017)

	Secure data fusion	The process of combining or synthesizing multiple data sources to produce more reliable information that is consistent and accurate as compared to the information provided by any individual data source is known as data fusion. The archival of data that is being generated in every second from in IoT and managing the dimension of uncertainty associated with data fusion is a big challenge.	(Venckauskas et al., 2016a, Chen et al., 2009)
	Computational costs of the encryption and decryption	In IoT environment, many copies of encrypted data get generated which requires a computational cost. This issue is also linked with the resource-constrained environment.	(Xiao et al., 2017, Usman et al., 2017, Yu et al., 2017)
	Computation in mobile devices and cloud	IoT devices are unable to deal with big size databases; these devices share their data with cloud environment get the advantage of computing. But due to the limitations of power, storage, and computation capabilities results obtained from cloud may get compromised. Therefore, there is a need for IoT devices to have onboard computing capabilities.	(Yu et al., 2017)
	Verification of data deduplication	In the cloud, there exist a lot of highly redundant data, which wastes the storage and bandwidth of the cloud servers. The correctness and verification of this redundant data is a significant challenge.	(Yu et al., 2017, Yan et al., 2016a, Yan et al., 2016b)

	Secure information processing	Processing of information in such a way that is only detectable by the analyst and provides a secure mechanism to avoid information mishandling when data is coming from numerous sources is a serious issue that may lead to privacy breaches.	(Zhang et al., 2017)
Digital forensics	Diversity of devices	IoT devices have a high diversity/heterogeneity (i.e., different operating systems, vendors, and methods of communication) making it difficult for traditional forensic tools to work. There is a continuous need to update these tools to support varied architecture of IoT to conduct the examination effectively.	(Zulkipli et al., 2017)
	Lack of standardization	Various authors have reviewed IoT challenges and proposed their solutions, yet they do not provide implementation guidelines that fit almost every scenario and their possible frameworks for future development in the area.	(Harbawi and Varol, 2017, Zulkipli et al., 2017)
	Improper evidence handling	IoT evidence has some important characteristics (i.e., they are volatile, fragile and with short lifespan) that make forensic examination difficult. These evidence can be easily tampered or even overwritten. Another issue is that, to collect/preserve evidence, IoT devices needed to be switched off to avoid change in metadata (i.e., accessed time), but it is not possible to shut down these devices. Therefore, investigators need to equip themselves with techniques to cope with these situations.	(Zulkipli et al., 2017)

	Evidence identification	In IoT where data is coming from a variety of different sources (i.e., from different vendors, data centers, clouds or even from different countries), an important forensic challenge is not only to identify the potential sources of evidence but make them accessible as well which at the moment is impossible. This challenge becomes more complicated when we consider it in a relationship with interoperability, heterogeneity and scalability issues of IoT.	(Harbawi and Varol, 2017, Brown et al., 2005, Quick and Choo, 2014, Taylor et al., 2010, Zulkipli et al., 2017, Liu, 2015)
	Seizure or disposal of evidence	IoT environments are full of both reliable and unreliable information in which seizure or disposal of crime scene evidence may also take place either by forensic experts or by cybercriminals to hinder the investigation process. Orientation and location of digital evidence at such places where the collection, disposal or seizure of the evidence may not possible makes forensic examination a challenging task.	(Yakubu et al., 2016, Conlan et al., 2016, Liu, 2015)
	Admissibility of evidence	Researchers claim that the ever-increasing number of IoT devices and volume of data generated by them required more time to conduct the forensic investigation. On the other hand, the vulnerability of IoT creates doubts on the admissibility of evidence that is an open challenge.	(Vlachopoulos et al., 2013, Quick and Choo, 2014, Sheldon, 2005)

	Computing and traditional forensic tools	Current forensic tools/techniques available for computation are not capable of dealing with IoT environment.	(Zawoad and Hasan, 2015)
	Evidence jurisdiction	An important challenge in evidence collection is to handle jurisdiction issues where evidence is beyond the scope/reach of an investigator. In IoT environment, data roaming/ traveling is usually possible particularly when cloud computing is involved which makes it impossible for an investigator to identify, collect, seize, or dispose of evidence (i.e., a country having ownership of specific evidence may refuse to handover it to the investigator from another country).	(Oriwoh and Sant, 2013, Zulkipli et al., 2017, Liu, 2015)
	Application of forensic procedures	There are six steps involved in digital forensics examination with clear guideline to apply but when it comes to IoT, there are a lot of factors that make the forensic investigation tough challenge to use (i.e., massive amount of data generated/exchanged between devices, volume of the heterogeneous devices in the network and various other factors). In this scenario, evidence finding (identification, preservation, and collection) requires extra research efforts to apply whole forensic framework under extraordinary constrained environment of IoT.	(Zulkipli et al., 2017)

Heterogeneous interactions	Diverse protocols, technologies that create a complex configuration	Usage of various protocols and mixture of heterogeneous technology cause configuration issues requiring considerable attention. Furthermore, cybersystems have various kinds of interactions between entities, these interactions are not limited to cyber and physical characteristics but also include social attributes, which are particularly crucial for across-space interactions.	(Kim, 2017, Gubbi et al., 2013, Ning et al., 2013)
Identity and access management (IAM)	Cooperative authentication	In mentioned literature cooperative authentication has been reported as a network security challenge in a smart community environment to filter false data traffic in the community network.	(Ning et al., 2015, Li et al., 2011)
	Identity Fabrication (IF)	An attacker may fabricate and create a fake identity, RFID identity can be duplicated or spoofed, and the existence of multiple identities is an issue. There is a need to differentiate between fake/fabricated or duplicated identities vs. original identities which is a very challenging task.	(Ning et al., 2015, Roman et al., 2013, Babar et al., 2010, Meghanathan, 2010)
	Identity ownership	Things or objects in the IoT often have a relationship to real persons and in many cases to other objects. These objects can be the owners, manufacturers, users, administrators, or many other functions. Ownership of objects and their identities becomes a critical challenge when they move from one network to another,	(Ning et al., 2015, Lam and Chi, 2016)

		in this case, the same object gets another owner. If this issue is not handled properly, the device can be compromised.	
	Identity and authentication	Mutual authentication is a critical issue in IoT when it comes to managing a large number of objects having a variety of data sources. For a trustworthy communication between devices, there should be a centralized authentication system which deals with object identities and provides the right level of authentication.	(Roman et al., 2013, Mahalle et al., 2010)
	ID tracking	In a heterogeneous environment where numerous smart devices are communicating, the process of tracing an object with the help of their identifiers is a crucial process. If handled, ID tracking may help in better asset handling, verification and audit process. Currently, this challenge is in debates.	(Wang et al., 2014)
Inadequate infrastructure or bad design	Traditional centralized time-synchronization protocols cannot be easily extended	Application of traditional time synchronization protocols, e.g., network time protocol (NTP) centrally in the low constrained environment is difficult because of a diverse range of devices, and their extension to adopt security features is even harder.	(Dong and Liu, 2015)
	Devices with low energy constraints	IoT devices come with low-resources and limited battery power; this energy constraint becomes a challenge when applying security and privacy controls.	(Venckauskas et al., 2016b)
	Insecure booting process	Boot process requires sensitive, trusted or verified and secure protection to avoid the device getting compromised or corrupted	(Yaqoob et al., 2017)

		while booting. Management of integrity metrics such as software and firmware in the boot process is not well designed at the developmental stage of IoT devices. It becomes a big challenge of the developmental process that leads to insecure booting process and may result in problems with device integrity metrics.	
	Verifiable computing	A challenge of fast-paced environment to develop security systems as a part of the architecture that verifies the computation performed. A significant problem in IoT is the lack of trust while performing data processing and computation when dealing with heterogeneous devices with a complex configuration. Under these circumstances verifiability while performing computation becomes a serious issue which requires immediate attention.	(Alrawais et al., 2017a)
Inadequate or limited support	Lack of mature IoT technologies and business process	Despite rapid growth in IoT based technologies, still, there is a lacking of maturity in the technologies and the business processes.	(Kim, 2017)
	Lack of technical support	IoT organizations consider the shortage of staff experienced in cybersecurity, hardware/software and data science, etc. to run IoT related projects efficiently.	(Lin and Bergmann, 2016)
	Lack of device authentication procedures	IoT lacks device authentication procedures so that anonymous devices can be added by the attackers and scalability/heterogeneity makes it more challenging.	(Yaqoob et al., 2017)

IoT malware	Malware targeting different operating systems	IoT devices support a variety of operating systems (i.e., ARM Mbed OS, Contiki and Windows 10 for IoT, etc.) and it has been noticed through published literature that majority of malware target Windows and Android operating systems. It is important to consider which operating system is being targeted by most of the attackers. To study malware targeting each IoT device, it is important to consider: (i) device architecture; (ii) firmware and (iii) operating system. Unfortunately, there is not enough literature to cover this aspect.	(Karanja et al., 2017)
	Tools for malware synthesis	Synthesis of IoT malware is still in early days, real-time data for malware analysis is usually not available, and therefore, the usage of emulators/simulators comes into the picture. There are various tools in the market to synthesize the IoT malware (i.e., emulators, honeypots, testbeds, etc.) but all of these tools have some limitations when it comes to the resource-constrained nature of IoT. Moreover, no research has been performed yet focusing on tools for malware synthesis.	(Karanja et al., 2017)
	Detection of obfuscated malware	Detection of malware is challenging in IoT due to: (i) Author's ability to write complex obfuscation techniques; (ii) Use of polymorphic/metamorphic malware; (iii)higher latency of IoT devices being online 24/7;(iv) weak security mechanism to	(Karanja et al., 2017)

		discourage malware; (v) and no support of the anti-malware system.	
	Malware analysis	Researchers claim that malware are becoming increasingly complex and adaptive, malware authors are continuously changing their strategies for infection and distribution. Along with the complexities of IoT, it is becoming more and more challenging to analyse malware targeting IoT environment efficiently. Malware analysis is very important to understand different perspective of malware that helps in correct identification and classification.	(Dulaunoy et al., 2017, Suarez-Tangil et al., 2014)
	Heavy network traffic analysis	One of the major problem in IoT is the massive volume of traffic generated by billions of devices communicating together. Analysis of network traffic plays an important role in cybersecurity, it helps in anomaly detection and building up a better defense. In case of malicious attacks detection of malicious traffic becomes a challenging task.	(Conti et al., 2018)
Physical and environmental security	Protecting resource constrained devices	IoT environment consists of limited resource devices where implementing protective measures is a key challenge.	(Chiang and Zhang, 2016)
	Adoption of network layer security	Resource-constrained IoT environment is making the adoption of network layer security approaches (such as IPSec and IKE in 6LoWPAN environments) a challenge in IoT.	(Granjal et al., 2015)

	Synchronization of multiple RF tags	To handle (identify, verify) multiple RF tags centrally to secure IoT environment against attacks has been reported as a challenge.	(Liu et al., 2016)
	Application of physical layer security	IoT has multiple practical constraints due to which the application of physical layer security becomes a challenge as reported.	(Mukherjee, 2015)
	Dynamic activity role	Cyber entities might be simultaneously idle in some scenarios and active in others, this activity has been reported as one of the obstacles for the network and application security in IoT.	(Ning et al., 2013)
	Target tracking	Target tracking deals with finding/tracking the objects in IoT environment, and the capability to track their movements has been reported as a security obstacle.	(Ning et al., 2015)
	Secure localization	Wireless network sensors are deployed in IoT devices. Secure localization in wireless sensor networks is an unattended area that can give passage for the malicious attacks.	(Pirbhulal et al., 2017, Chen et al., 2017, Sen, 2010)
	Implementation of protocols and network security	Heterogenous, resource-constrained devices influence significantly on protocols and network security of IoT devices during device interaction making the implementation of cryptographic protocols for network security a tough challenge.	(Roman et al., 2013)
	Compromised or malicious sensors	A situation is reported in which a legitimate sensing device gets compromised; the adversary usually makes the clones or replicas to cause more damage. This kind of compromise is a serious issue which requires addressing.	(Sawand et al., 2015, Qiu and Ma, 2016)

	Mobile node security	IoT nodes are usually mobile and frequently move from one cluster to another. During this process, there is a possibility of potential exploitation, so there is a need to efficiently handle node mobility using effective cryptographic mechanisms to provide rapid identification, authentication, and privacy protection. Because of unavailability of these services mobile node security is an issue.	(Sicari et al., 2015)
	Secure middleware	Many IoT systems have been derived using middleware frameworks that increase the need for the application of security matrix to protect middleware from getting compromised. This issue has been addressed by researchers, but still, this issue is getting reported.	(Sicari et al., 2015)
	Security of resource-constrained sensor devices or nodes	The security of resource-constrained IoT sensor nodes is a serious security challenge due to resource-constrained sensor nodes. In this situation application of traditional security measures is not practical because these security measures put enormous computation/communication overhead on the devices.	(Dong and Liu, 2015, Mineraud et al., 2016, Sen, 2010)
	Secure authentication	Application of security during the process of authentication, while IoT devices are communicating with each other, has been reported as an important challenge in various literature.	(Mineraud et al., 2016, Liu et al., 2016, Mendez et al., 2017, Borgia, 2014)

Risk treatment countermeasures and strategies	Secure location	A process to implement, detect and maintain the security of node locations around the wireless network environment has been reported as a challenge.	(Venckauskas et al., 2016a)
	Intrusion Detection	Intrusion detection techniques detect misbehavior or malicious IoT devices and notify others in the network to take appropriate actions. The nature of IoT environments with limited resource makes it challenging to detect the insider and outsider attacks.	(Ning et al., 2015, Alrawais et al., 2017a)
	Cryptographic key management issues and their solutions	In a cryptosystem, key management ensures to provide data confidentiality in IoT, distributed/diverse nature of IoT raises this issue as a security challenge.	(Tiburski et al., 2015, Sen, 2010)
	Firewalling	Application of right kind of firewalls on IoT is an issue because most of the traditional firewalls do not perform efficiently on the network traffic generated by IoT environment.	(Yaqoob et al., 2017)
	Attack detection and prevention	Detection and prevention of malicious attacks (predominantly DoS/DDoS) attacks is one of the most serious challenge ever reported.	(Alrawais et al., 2017a, Venckauskas et al., 2016a)
Safety and compliance	Data Protection	Data generated by IoT devices is huge, and it is not preserved at any level of communication/computation. Due to resource constraints, lack of encryption/decryption of data of IoT devices is an open challenge where breaches of data protection may occur particularly in the cloud and distributed IoT environment.	(Alrawais et al., 2017a, Ziegeldorf et al., 2014)

	Assessing the status of the security of large distributed systems in a trustworthy manner	IoT supports large-scale network environments including cloud-based, large-scale distributed environments and complex interconnected networks, i.e. smart cities that includes hospitals, transports, logistics and many more. An ability of a system to tell in a trustworthy way whether it will operate securely or not is a challenge in IoT.	(Chiang and Zhang, 2016)
	Security laws and regulations	The adoption and harmonization of security laws and regulations in the presence of contradicting stakeholders, viewpoints, and complex devices is a real challenge that requires continuous attention.	(Ciccozzi et al., 2017, Suo et al., 2012, Hu, 2011)
	No defined audit and logging standards	Audit standards make sure that security controls have been effectively placed in the environment, while logging standards make sure that data being logged is complying with specific standards and access standards, this is something not present at the moment. The implementation of audit/logging standards without any conflict is a complex challenge	(Kim, 2017, CSA, 2015)
	No security standards for platform configurations supporting multi-tenancy	It is a complex challenge to integrate security standards to harmonize multi-tenancy platforms. It has been reported that IoT systems were designed without consideration of security standards. Furthermore, when it comes to the multi-tenancy data	(Kim, 2017, CSA, 2015)

		centers, the IoT devices associated with them and their configurations make this issue more complicated to resolve.	
	Standards required for the authentication and authorization of the devices are lacking	Implementation of standards for the authorization and authentication still lacks in IoT devices/environment.	(Kim, 2017, CSA, 2015)
	Policy enforcement	Policy enforcement refers to the mechanisms used to force the application of a set of defined actions in a system. When an entity or a user is authenticated to access an IoT resource, it becomes crucial to enforce the security policy to restrict the user to perform only the actions they are allowed to do, which is missing in IoT at the moment.	(Sicari et al., 2015)
Secure and trusted communication	Manipulating routing information	Routing protocols and a mechanism to secure IoT are not well defined. Therefore, route manipulation is possible in most of the IoT devices. It has been reported in the literature that route modification towards a malicious node or to an illegitimate destination can lead to information manipulation where the security and privacy can be compromised.	(Du and Chen, 2008)
	Remote attestation (RA) for a large number of devices	Attestation services allow a user or application to authenticate an IoT node. In IoT devices, remote attestation is possible at the individual device level to prove trustworthiness. It has been	(Chiang and Zhang, 2016)

		reported, RA for a large number of devices require high cost and management complexity. Attestation is also challenging in case of malware attacks.	
	System-level-trust (SLT)	(Ciccozzi et al., 2017) States SLT as an important challenge in IoT, they did not discuss this in detail either the reference they quoted explained anything about it. However, in general, SLT is an important feature that refers to the trustworthiness of system. When considering SLT concerning IoT, it means it should give some level of security, reliability, privacy, and trust and keeps the IoT device and the environment secure from compromises.	(Ciccozzi et al., 2017)
	Expanding domains (ED)	ED is one of the main obstacles to securing the cyber entities in IoT; ED has been described as the mapping of objects in IoT with networking and communication of cyber entities.	(Ning et al., 2013)
	Secure transmission of data	How to securely transmit the collected data from the sensor nodes to the destination that remain insecure due to low power and small size IoT nodes has been reported as a major challenge.	(Pirbhulal et al., 2017)
	Mobility capabilities	In IoT mobility capabilities have been reported as one of the important challenges when users are on the move, it may cause one of the following: (i) service interruption; (ii) service continuity; (iii) mobility management of the components	(Venckauskas et al., 2016b, Fraga-Lamas et al., 2016)

	Gateway address exposure	IoT devices send data to the local gateway on a daily basis especially in the healthcare sector. IoT connected devices have only been bound to HTTP for the interactions with the gateway. HTTP is insufficiently insecure for many of the interactions in the IoT while sending the data; attackers may guess gateway address and manipulate the data for harmful purposes and cause issues related to authorization, authentication, and accounting.	(Fantacci et al., 2014)
	Restricted interfaces available to interact with IoT devices, security devices, and the applications	IoT environments face challenges when an organization needs to integrate an IoT device into existing infrastructure. Unfortunately, there are no interfaces available to do this work.	(Kim, 2017, CSA, 2015)
Security considerations	Manage/administer the ownership of devices in a complex ecosystem	IoT is a single complex ecosystem where various infrastructures are condensed to form a dynamic and interactive environment. Where the main challenges are: (i) tracking of multiple devices; (ii) identify malicious identities; (iii) bad manufacturer; (iv) identify a malicious attacker. In this situation,	(Furfaro et al., 2017)
	Local and remote software update	IoT devices are vulnerable, to keep them secure there is a need to design software: (i) to update remote software that handles security updates; (ii) track firmware updates.	(Alrawais et al., 2017a)

	IoT software, security credential and firmware updates	Fixed firmware is an open issue; there are very few appliances that keep updating the firmware/software regularly. Many vulnerabilities may occur because of outdated software, so the best approach is to secure profiles/roles/access, etc. And related software up to date.	(Lin and Bergmann, 2016, Chiang and Zhang, 2016)
	Maintain low-cost authentication encryption	To implement and maintain cost effectiveness of encryption mechanism in the authentication process has been reported as a challenge.	(Mazumder et al., 2017)
	Non-Repudiation (NR)	NR is referred to as the ability to ensure that a person cannot deny something. Nonrepudiation becomes an issue in case of malware attacks or when the security controls are not properly implemented. Although people have tried to address this challenge but still the weak areas in IoT environment may trigger this issue at any time.	(Pirbhulal et al., 2017)
Security measures and good practices	Trustworthiness	Security in IoT is the most important consideration where maintaining trust is a key challenge whether it is required in cryptographic systems, device firmware or other at any stage of IoT environment.	(Ambrosin et al., 2016, Alrawais et al., 2017a, Chiang and Zhang, 2016, Roman et al., 2013, Pirbhulal et al., 2017, Sicari et al., 2015, Pan et al., 2011)

	Introducing physical security concerns	Deployed IoT devices often remain exposed to various threats and vulnerabilities. Mostly used software-based solutions to protect sensitive information. The key challenge is attackers with efficient solutions can reverse the software solutions. In this situation, it is important to introducing the physical security instead of a software solution to protect the IoT asset. With the evolution of IoT, the domain of physical security is changing, in particular, with the emergence of malware threats, vulnerabilities, and other destructive activities. Thus, IoT is showing considerable potential implications in this regard and requiring comprehensive research.	(Kim, 2017, CSA, 2015)
	Improve authentication protocols	IoT devices confront various security challenges, and authentication protocols are required to improve performance, security, and effectiveness in IoT environment.	(Liu et al., 2016)
	Governance in IoT	Governance guidelines are unclear for IoT, but no one has described it in detail. However, it has been reported that there are at least two important aspects of governance that need to be addressed: (i) Information Governance in IoT: IoT creates enormous amount of information that requires management by the implementation of governance practices.	(Roman et al., 2013, Roman et al., 2011, Ning et al., 2015, Hoepman, 2011)

		(ii) Trust governance framework: Governance frameworks consist of record management, risk analysis, asset management, etc. these are badly missing elements in IoT. There is a serious need to implement governance frameworks in IoT that brings trustworthiness in this business.	
	Maintain the level of security	To attain/maintain the right level of protection and continuous improvement to bring resilience to it has been reported as a challenge, although authors of these literature have proposed some solution but full security is still an issue.	(Mazumder et al., 2017)
	QoS	QoS in IoT is an area which is yet to be explored, the heterogeneity, limited resource constraints, mixed network traffic, and complex network topologies, etc. make implementation of QoS a tough task to do. Few other authors have also reported QoS needs at different levels within the IoT.	(Venckauskas et al., 2016b, Atzori et al., 2010)
	Risk analysis	In the IoT environments comprising of numerous smart devices, it's important to effectively evaluate the security, measure the amount of risk involved in order to present a holistic view of the whole system.	(Wen et al., 2017, Riahi et al., 2013)
	Patch solutions for IoT	One of the apparent research gaps in IoT is the inability of the devices to upgrade the software (security related) and patch them	(Yaqoob et al., 2017, Min et al., 2014,

		in a non-disruptive way. It has been reported that IoT devices must be able to accept updates and patches of security software.	Kuusijärvi et al., 2016, Sicari et al., 2015)
Security mechanism	Fault tolerance	Fault tolerance is the process that enables a system to continue operating smoothly/adequately in the event of the failure. This feature is lacking and reported as a tight constraint in IoT.	(Ning et al., 2015, Roman et al., 2011, Pan et al., 2011)
	Security structures that combine control and information access	It has been reported that IoT needs the implementation of security structures that combine control and information access but no further detail has been specified.	(Ciccozzi et al., 2017)
	Implementation of cryptographic identifier	Asymmetric keys (based on large numbers) are extensively used as cryptographic identifiers, have a significant overhead as compare to symmetric keys, therefore, require substantial computational resources that is a current challenge in IoT due to architectural limitations.	(Mayer, 2009)
	Architectural security weakness	IoT device manufacturers have been showing little or no focus on the implementation of security mechanisms at architectural level that gives rise to architectural security concerns and associated threats and vulnerabilities.	(Anantharaman et al., 2017)
	Limited sense of intrusion detection	The architectural design of IoT doesn't allow it to be easily extended to adopt security mechanisms. Therefore, the ability to deal with the intrusion detection to deter malicious attacks is insufficient and has been reported as a challenge.	(Pajouh et al., 2016)

	Security of cryptography/cryptographic algorithms	Application of security measures to protect cryptographic algorithms and the keys remains a significant challenge.	(Xiao et al., 2017, Continella et al., 2017)
System safety and reliability	Location Verification	A challenge to design secure system for the verification of the locations of IoT devices in harsh environments such as transport systems. Some authors attribute the presence of location verification to better security improvement, but still, this area requires considerable work.	(Alrawais et al., 2017a, Chen et al., 2009)
	Mechanisms for self-diagnosis and self-repair	It has been reported that the dependability of IoT can be increased when it has the mechanism of self-diagnosis and self-repair in order to provide better fault tolerance and smooth operations. It is easier to provide this facility at application/device level, but when it comes to the whole system, it is considered as one of the hardest problems.	(Garlan et al., 2003)
	Cyber-physical infrastructure	With the boost of urbanization, smart city concepts is on the rise. In which multiple IoT applications are getting deployed. There is a major concern of handling physical damage or undesirable risk of injury to the infrastructure and their components in case of cyber attacks. Moreover, there is need to make whole infrastructure (such as electricity supply, water distribution,	(AlDairi, 2017)

		streets, buildings, etc.) safe and reliable to provide a better way of life.	
	Measurement of reliability	Reliability is one of the important elements to attain the quality of service a user expects from the manufacturer of a product. It can be interpreted as a measurement of unreliability at which failure may occur. Authors have reported this challenge concerning QoS.	(Venckauskas et al., 2016b, Venckauskas et al., 2016a)
Threats and vulnerabilities	Destruction or theft of the M2M device	Being deployed in reachable locations, M2M devices or their cards can be easily stolen.	(Barki et al., 2016a)
	Leakage of cryptographic key	Cryptographic key leakage occurs in public-key systems when the system gets attacked by inside or outside attackers.	(Chiang and Zhang, 2016)
	Vulnerabilities in IoT devices	With the rapid advancements in IoT technologies, every day new vulnerabilities are getting discovered, some of the vulnerabilities are inter-related to existing ones, but the complexity in IoT environment allows room for newer ones and makes this challenge more difficult to handle.	(Kim, 2017, CSA, 2015)
	Threats to multitenancy data centers	Multi-tenancy is a system building block in which a single instance of an application serves multiple customers. The information may be leaked from the main data centers that are dealing with different client services that leads to security/privacy breaches.	(Kumarage et al., 2016)

	Leakage of security-critical information	Problems faced during the transmission of data is the leakage of critical information, this may be due to malware as well.	(Kwon et al., 2016)
	Heavy communication traffic	IoT devices generate heavy communication traffic that is difficult to handle.	(Li et al., 2014a)
	Physical damage	In malware attacks, an attacker can destroy physical device or important hardware modules of targeted devices. Keeping various architectural weaknesses of IoT, it's a challenging task to keep IoT devices safe and secure.	(Roman et al., 2013, Challa et al., 2017)
	Dependability in control	With the advancement of IoT, the applications like smart cities have been thriving where heterogeneous network infrastructure exists with multiple sensoring devices, information processing, and control systems interact with each other. It has been reported that dependability in control is considered as the topmost priority challenge in this kind of IoT application where it can be a prime target for attackers and terrorists etc. an active attacker can try to gain partial or full control over an IoT entity or a system.	(Roman et al., 2013, Challa et al., 2017, Zhang et al., 2017)
	Node related issues	A malicious IoT node could pretend to be legitimate to exchange and collect the data generated by other IoT devices for malicious purposes. With ever-increasing volume of nodes can cause (i) node capture issues; (ii) detection of captured, rogue or unreliable nodes is a serious challenge.	(Roman et al., 2013, Challa et al., 2017, Alrawais et al., 2017a, Ning et al., 2015)

	Individual or group privacy information thefts	Privacy related thefts are a nightmare for any user. It was always a big challenge and will still be a serious issue whether a privacy theft occurs at an individual or group level resulting in disclosure of sensitive information.	(Ning et al., 2013)
Visible gaps in IoT	Limited incident response paradigms	Authors have reported that there are limited best practices available for incident response and existing incident response mechanisms will not be adequate for emerging IoT infrastructures.	(Chiang and Zhang, 2016, CSA, 2015)
	Limited guidance present for lifecycle maintenance/management of IoT devices	IoT usually has limited capability operating systems in which guidance on secure configuration of these devices is either limited or not present.	(CSA, 2015)
	Limitations of cryptographic algorithms	The applicability of cryptographic algorithms in IoT is limited and requires further analysis to ensure that algorithms can be successfully implemented given the constrained memory and processor speed expected in the IoT.	(Trappe et al., 2015)
	Self-storage	Self-storage is an open issue in IoT that is linked with limited resource-constrained environment. The problem occurs when alternative storage is needed, in which handling of data ownership and access control management is difficult.	(Mineraud et al., 2016)
	No data cataloging	The need of addressing modern data processing is readily becoming essential that is only possible if data catalogs are	(Mineraud et al., 2016)

		available, unfortunately, this feature is unavailable in IoT and the author mentions resource limitations as the reason for this unavailability.	
	Gap analysis	ENISA performed gap analysis and identified following research areas requiring attention: • Fragmentation in existing security approaches and regulations. • Lack of awareness and knowledge. • Insecure design and/or development. • Lack of interoperability across different IoT devices, platforms, and Frameworks. • Lack of economic incentives. • Lack of proper product lifecycle management.	(ENISA, 2017)
Vulnerable to attacks	Attack detection and prevention	Nowadays manufacturers have been rapidly introducing new products in the market, throwing more and more devices in which security and privacy are already questionable, the need to detect and prevent cyber security attacks particularly malicious attacks is increasing every day. All sorts of vulnerabilities targeting physical interfaces, hardware profiles, wired/wireless communication protocols, user interfaces, ports, etc. are posing the most significant challenge for IoT in which there is a need to develop approaches to detect and prevent the maximum number of threats.	(Chen et al., 2009, Alrawais et al., 2017b, Venckauskas et al., 2016a, Venckauskas et al., 2016b, Kliarsky, 2017)

	Malware issues	The reported cases of malware targetting IoT rising significantly. Authors of various literature have reported numerous reasons for malware attacks (i.e., architectural limitations or design implications, deficiencies in security mechanisms, Internal/external threats, software vulnerabilities, software modification and many more). The varied possibilities of malware related issue in IoT making this one of the most significant challenges.	(Yu et al., 2017, Chiang and Zhang, 2016, Ning et al., 2013, Kumarage et al., 2016, Alrawais et al., 2017a, Liu and Sun, 2016, Sawand et al., 2015, Barki et al., 2016b)
	Attacks on IoT devices/ or environments	Various attacks discussed in literature have been summarised in table 3	Table 3

2.5.2. Authentication:

Every day a large number of IoT devices are getting manufactured based on diverse protocols where the heterogeneity and ad-hoc nature of agents create complex configurations. IoT technology aims to provide accessibility (*anytime*, *anything* and *anywhere*), better services and seamless communication to connected smart devices, where authentication is prerequisite. The big challenge lies in the design and development of IoT infrastructure where security and privacy were not considered. Rigorous authentication is based on efficient, user-friendly and scalable procedures/protocols requires resources such as storage and computation which are lacking in IoT due to poor architecture. In this scenario, identifiable device/user information have maximum disclosure giving criminals access to resources and raising further complex issues of anonymity and untraceability. This deficiency of authentication procedures also cause more complex security and privacy issues (Alrawais et al., 2017a, Amadeo et al., 2016, Zhang et al., 2013, Al-Fuqaha et al., 2015, Tiburski et al., 2015, Venckauskas et al., 2016b, Pirbhulal et al., 2017, Sicari et al., 2015).

2.5.3. Confidentiality, Integrity, and Availability (CIA):

The design and build of IoT devices have a limited sense of hiding the information from unauthorized people to view it. Data sensed and exchanged by things, and the privacy of humans/things must be ensured to prevent unauthorized identification and tracking (Liu and Sun, 2016, Tiburski et al., 2015, Lin and Bergmann, 2016, Pirbhulal et al., 2017, Sicari et al., 2015). Confidentiality (privacy), integrity and availability are the true aspects of security. CIA is a model, guide and a complementary requirement for information security, if ignored or not followed correctly may raise very complex security and privacy issues. Unfortunately, CIA model was not given much attention in IoT. This model should be enforced at every level of IoT infrastructure, i.e. access control, data sensed or exchanged, and authentication, etc.

2.5.4. Crisis management:

Due to immense popularity, IoT environments/devices have been a target of malicious attacks (in particular botnets and ransomware) resulting in a denial of service. Moreover, it’s a part of an effective governance plan to devise strategies to respond to the crisis situations that may cause intolerable business disruption to resume to the state of smooth operations (Chiang and Zhang, 2016). Effective brute-force solutions cannot be applied in crisis situations. Therefore, there is need to address the associated challenges comprehensively.

2.5.5. Cryptography:

IoT devices usually come with 8 or 16-bit chips, the constraint of limited energy (small batteries), storage and onboard memory that prevents the implementation of efficient, flexible and inexpensive cryptographic algorithms (Ning et al., 2015, Roman et al., 2013, Venckauskas et al., 2016b)**.** There is a need to design efficient cryptographic algorithms which can be applied throughout the device to offer an end-to-end secure communication channel. In this regard, lightweight security protocols are the best choice for 8-bit or 16-bit devices. The implementation of these security protocols is widely influenced by the heterogeneity and scalability issues of IoT. Along with this, a competent algorithm requires storage and energy in the devices which is a very challenging task in IoT (Riahi Sfar et al., 2017). Moreover, the revocation of old/useless keys (public/private) is another issue in IoT, and various authors have proposed their approaches to address this issue, but this challenge is still a nightmare because of several other inter-related problems.

2.5.6. Data privacy and protection:

A tremendous amount of data is generated by IoT devices every day. The unrestricted access to that information poses a significant security and privacy risks to the consumers due to the fact that devices not only connected with cloud but also forming local intranets to exchange data between them (Barki et al., 2016b, Mayer, 2009, Sicari et al., 2015, Wang et al., 2014, Liu and Sun, 2016, Ning et al., 2015, Pirbhulal et al., 2017). The key data privacy-challenges are:

- Most IoT devices fail to encrypt data that are being transferred.
- User sensitive information can be compromised due to unencrypted data.

2.5.7. Data processing and computation

One of the most important aspects of IoT is "data," which poses very complex challenges due to the presence of 3Vs, i.e. "Volume," "Velocity" and "Variety." The process of analyzing or computing the huge amount of data (volume) being generated at a rapid pace (velocity) from various sources (variety) that sometimes includes outsourced data is a serious problem. Issues related to data processing and computation include aggregation, retention, integrity, freshness, handling of false or noisy data, secure fusion, computational cost, verification of outsourced data, verification of de-duplication data in mobile and cloud-based IoT. Various researchers are using different techniques for the processing and computation, but still it's an open research challenge in IoT (Alrawais et al., 2017a, Kumarage et al., 2016, Kwon et al., 2016, Pirbhulal

et al., 2017, Sawand et al., 2015, Tiburski et al., 2015, Venckauskas et al., 2016b, Xiao et al., 2017, Yasin et al., 2017, Yu et al., 2017, Zhang et al., 2017).

2.5.8. Digital forensics:

Digital forensics is an important discipline that helps to identify illicit activities from the digital world for a variety of matters, e.g. criminal cases and cyber malware attacks (inside/outside), etc. There are six important pillars of the forensic investigation, but when it comes to the application of these six elements in IoT, forensic examiners face numerous issues. Many of these issues have been reported in the literature, but the solution to these issues still requires work. For example, a traditional forensic tool like ENCASE works well with conventional technologies like laptops, desktop computers or mobile devices but when it comes to IoT devices (home appliances, i.e. smart fridge), we have no answer. IoT is a fast-moving technology concerning the ever increasing number of devices and massive volume of data. It means IoT digital forensics needs to constantly develop IoT-specific forensic tools by addressing the challenges reported in the literature also summarised in the table (Attwood et al., 2011, Brown et al., 2005, Caviglione et al., 2017, Conlan et al., 2016, D'Orazio et al., 2017, Harbawi and Varol, 2017, Liu, 2015, Oriwoh and Sant, 2013, Quick and Choo, 2014, Sheldon, 2005, Taylor et al., 2010, Vlachopoulos et al., 2013, Yakubu et al., 2016, Zawoad and Hasan, 2015, Zulkipli et al., 2017).

2.5.9. Heavy network traffic:

Despite all the benefits of IoT, the weaknesses of one-time, low cost, resource-constrained and unsecured sensors generating a massive amount of network traffic gives birth to another challenge of how to analyze network traffic particularly if that traffic is malicious to understand the pattern of clean and malicious behaviors. There is not one possible answer to this situation because the volume of the data requires more hardware/software resources and highly analytical techniques and continuous improvement in them to cope with the ever-increasing amount of information and cope with newer threats (Gan et al., 2011).

2.5.10. Heterogeneous interactions:

IoT technology inherent the complexity due to a considerable number of heterogeneous devices associated with interoperability system to exchange the information with each other. Constrained entities share internet with non-constrained devices, either directly or through the gateways. In this situation, security and privacy of IoT devices are compromised due to their

lack of support of heterogeneity and incompetent architectural model (Gubbi et al., 2013, Kim, 2017).

2.5.11. Identity and access management (IAM)

In our daily lives, we remain occupied in billions of IoT devices, and there is a continuous growth in numbers which is a challenging task regarding identity and access management. IAM refers to the process of representing/recognizing entities as digital identities in virtual networks. It enables (right) people to access only concerned objects only when needed. The functions of IAM are increasing rapidly where it is important to identify not only the authorized people, tracking of their object, and handling of their privileges towards a variety of different data sources. From the perspective of cyber security identity management of IoT devices is the most critical and vital area towards securing the environment (i.e., identify people, devices, monitors, sensors and secure data access, etc.). Without IAM an attacker may fabricate and create a fake identity, RFID identity can be duplicated or spoofed, and finally, the existence of multiple identities is an issue as well (Mayer, 2009, Roman et al., 2013). Furthermore, an object may have either multiple or fake owners as well which may cause some serious issues unless we have a right defense in place (Babar et al., 2010, Lam and Chi, 2016, Li et al., 2011, Mahalle et al., 2010, Meghanathan, 2010). There is a need to identify between fake/fabricated or duplicated identities vs. original identities which is a very challenging task.

2.5.12. Inadequate infrastructure or bad design:

IoT is a hot favorite topic for the researchers because of its bright future perspective, being a consumer's first choice, and extraordinary long list of security issues that cause massive service outage worldwide. IoT devices were designed to improve lives of the people, but because of lack of incentives for the vendors, security features were entirely ignored from hardware/software infrastructure making IoT device or environmental design a "Bad design." It has been noticed that in the event of any DoS attack, the attackers utilize weaknesses of the target system in which IoT is full. At first place, if the prototype of IoT device was not made considering security as a part of its build, it may be considered unfit to be placed in the list of safe devices. Moreover, if the same prototype with weak or no security features is adopted, then it may be regarded as "inadequate infrastructure" or "bad design." With this critical infrastructure, IoT is facing many challenges described in the table (Alrawais et al., 2017a, Dong and Liu, 2015, Venckauskas et al., 2016a, Yaqoob et al., 2017).

2.5.13. Inadequate or limited support:

Provision of support is one of a most important factor related to customer satisfaction. With the rapid increase in a number of devices with heterogeneous/complex configuration, support process has become more challenging. IoT industry will require a change in support process; this includes staff competencies to reflect strong IT/cyber security skills (i.e., knowledge about IoT devices and environment), management guidelines, cyber security skills, maturity in technical/business processes, and authentication processes within an organization. (Kim, 2017, Lin and Bergmann, 2016, Yaqoob et al., 2017)

2.5.14. Physical and environmental security

The fundamental issue in IoT is not only to protect its intrinsic elements (i.e., object or entities), but also to protect the information assets from malicious threats. Physical and environmental security is not a new concept, but its implementation in IoT is a challenging task that focuses on detection/prevention of unauthorized entities from gaining access to a resource and steal valuable assets. However, diverse nature of IoT devices, immature build and frequent changes in infrastructural design creates obstacles in this regard (Sen, 2010, Mendez et al., 2017, Borgia, 2014, Chen et al., 2017, Luong et al., 2016, Trappe et al., 2015, Liu et al., 2016).

2.5.15. Risk treatment, countermeasures, and strategies (RTCS):

Various authors have reported challenges i.e. the development of intrusion detection mechanism, firewalls, attack detection/prevention system and devising strategies for the secure management of cryptographic keys in order to save them from unauthorised access (Alrawais et al., 2017a, Ning et al., 2015, Tiburski et al., 2015, Trappe et al., 2015, Venckauskas et al., 2016b, Yaqoob et al., 2017). We grouped these challenges under RTCS. Risk management is considered as a most critical part of the organizational governance; if ignored, it's not possible for the organizations to survive for long. Therefore, it would not be wrong to consider it as a crucial challenge for an IoT environment to have risk treatment, countermeasures and their strategies as a part of their ongoing process. It is not a one-time process but rather an ongoing process, important phases in risk management include: (i) "Plan" (identification and evaluation of risks); (ii) "Do" (risk treatment, make strategies to overcome future risks); (iii) "Check" (analyse changes) and (iv) "Act" (plan for future events). All of these phases are recursive and help the analysts to discover all associated risks and develop strategies to act accordingly to effectively treat them. When we talked about IoT environment RTCS should come in top priority list (which at the moment it is not).

2.5.16. Safety and compliance (SAC):

Safety and compliance are considered as important driving forces of IoT since IoT devices have been a favorite target of cybercriminals who may steal or destroy important confidential information. This is why the need for safety and compliance becomes crucial and likely to grow in importance in IoT. By nature, IoT brings various security risks directly to the organizations, every organization must follow specific compliance standards to avoid substantial payouts in terms of fines and loss of reputation.

Unfortunately, this is the area in IoT where organizations did not pay much attention which may cause potential implications. IoT industry needs practices for recognizing the safety requirements while designing the products, delivering them to the broader markets and customers. IoT devices without safety, compliance, and conformity may rise safety-related compliance issues (Alrawais et al., 2017a, Chiang and Zhang, 2016, Ciccozzi et al., 2017, CSA, 2015, Gubbi et al., 2013, Hu, 2011, Kim, 2017, Sicari et al., 2015, Suo et al., 2012, Trappe et al., 2015, Ziegeldorf et al., 2014).

2.5.17. Secure and trusted communication (SATC)

SATC helps to improve availability/accessibility of IoT resources. Network communication is an integral part of IoT, where security and privacy are the indispensable but neglected tools. IoT relies on various communication channels such as sensing nodes, routing systems, etc. When a user or a sensor node share or exchange their data through a communication channel, in this process authentication and authorization play an important role, if communication channel(s) is insecure attacks like a man in the middle are possible. It's a challenging task to discover, verify, identify, and authenticate the devices/data in an IoT network to preserve privacy and whole CIA-triad. There is a need to establish a framework or set of rules to bring trust and security in entire communication process instead of at a particular level.

2.5.18. Security considerations

It is a known fact that in this scientific age people are increasingly relying on IoT devices facilitating them in every walk of life. Literally, we have billions of interconnected devices, and this figure is likely to be increased in the near future. As stated previously these devices are a favorite target of malware, intrusions of various kinds and interfaces to steal/expose personal information and jeopardize the safety of the people. Therefore, it's a major concern of security to address: (i) how to manage/administer the ownership of these devices in a complex ecosystem; (ii) how to update local/remote security credentials and other related

software or firmware; (iii) how to maintain low-cost features for implementing encryption for authentication roles;(iv) and finally how to provide non-repudiation while considering the limitations of them like heterogeneity, complexity, scalability and volume etc.

2.5.19. Security measures and good practices (SMAGP)

It is a known problem with IoT that most of the vendors have been developing their products without sufficient inclusion of security-related features in them. The lack of important features results in serious issues for not only the organizations but also for the public intending to use them. To provide consumers security & privacy concerning their data, and trustworthiness in the communication, government policy agencies should enforce vendors to implement proper security measures and good practices as a part of their hardware and software design. At this stage of IoT progress, due to various complex factors, it is not possible to govern a set of rules that are applicable universally on all sorts of devices, but it is possible to define a list of top most important security measures and a list of good practices to benefit the whole environment. In this regard, it is important to identify the gaps and then merge the technical solutions for security measures and good practices.

2.5.20. Security mechanisms:

Security of IoT has become a ubiquitous issue in which all traditional concepts have become upside down. One of the fundamental elements of securing IoT device is to implement strong security mechanisms. However, the architectural model of IoT devices doesn't allow to be easily extended to adopt security mechanisms. Therefore, the ability to deal with the intrusion detection is limited. Various challenges in this regard have been reported in the literature, we have listed few of the challenges in the table. The presented challenges indicate that there is a need for robust IoT security mechanisms to acquire a secure IoT infrastructure that penetrates well in all IoT applications.

2.5.21. System safety and reliability (SSaR)

It is known the fact that there is no dependability in IoT, dependability comes from the pursuance of two components, i.e. safety and reliability which are prerequisite of a better quality of service (Zin et al., 2016). These two components form the basis of failure-free communication environment.

With the urbanization, a smart city is the revolutionary concept where different IoT technologies are being deployed (such as IoT objects, cloud computing, real-world user interfaces, semantic web, etc.) and forming a smart system. These systems remain incomplete

without considering the safety and reliability. The “safety” and “reliability” should be assessed and measured at various levels (i.e., sensor level and communication level etc.) to provide a secure system (Li et al., 2014b, Li et al., 2015, Zheng et al., 2014, Kharchenko et al., 2016).

2.5.22. Threats and vulnerabilities

There is no doubt about it that IoT technologies contain serious vulnerabilities that are undeniable. Along with this, the rapid expansion of devices inherits same or sometimes even more complex vulnerabilities exponentially expanding the possibility of various threats. It would be highly dangerous for an IoT infrastructure if we ignore or misjudge the importance of vulnerability management and threat detection solution. Therefore, we need to put it into our priority list to identify the existing threats and vulnerabilities before they compromise security/privacy/trustworthiness of a system.

2.5.23. Visible gaps in IoT:

A solid understanding of a domain is only possible when we perform technology gap analysis. Purpose of this group is to provide a summary of few of the most prominent gaps in IoT that are co-related with security/privacy/trust challenges. Another reason to include this group is to highlight important aspects of modern technology that are clearly lacking in IoT, and their solution may bring resilience to IoT by making the technology mature.

2.5.24. Vulnerable to attacks:

Protecting the IoT devices or infrastructures from threats is a complex and challenging task. Security experts believe that global connectivity (access *anywhere*), accessibility (access *anyhow*, *anytime*) are the fundamental tenets for numerous types of attacks. However, there is no uniformity in the attacks and hard to expect where and when the attack may target. Attackers may focus on various communication channels, sensors, hardware profiles, information exchanged, etc. causing either fabrication, denial of service, jamming, identity theft, etc. In addition to this, the co-inherent complexity of IoT networks, highly scalable nature, heterogeneity of the entities located at various locations attract more attacks to give an attacker enough room to break in the system.

2.5.25. Attack detection and prevention:

Nowadays manufacturers have been rapidly introducing new products in the market, throwing more and more devices in which security and privacy are already questionable, the need to detect and prevent cyber security attacks particularly malicious attacks is increasing every day. All sorts of vulnerabilities targeting physical interfaces, hardware profiles, wired/wireless

communication protocols, user interfaces, ports, etc. are posing a significant challenge in which there is a need to develop approaches to detect and prevent a maximum number of threats (Alrawais et al., 2017a, Venckauskas et al., 2016a, Kliarsky, 2017).

2.6– Malware attacks as a big security/privacy risk and related work

The benefits of IoT are undoubtedly most attractive. Therefore, this technology has been adopted by various big organizations. On the other hand, security and privacy challenges of IoT is creating a global impression that "Internet of Things is the new Windows XP-malware's favorite target!" (Kuusijärvi et al., 2016). This statement is quite convincing when you see the published literature about IoT attacks. The table 3 shows 45 different IoT attacks reported in the literature; these attacks are directly or indirectly related to malware.

Table 3: The summary of IoT attacks reported in literature related directly or indirectly with malware

Attack	Description	Reference	Target	References
Availability attacks	Availability attacks mainly utilize the limitations of bandwidth and transmission power resulting in communication failure in IoT setup.	(Sun et al., 2017b)	Network Layer Transport Layer	Our contribution
Camouflage attacks	A camouflage node hides itself under a false identity and utilizes this appearance from a legitimately authenticated node, and spreads fake and harmful messages, or executes blackhole attacks, or other fatal attacks	(Sun et al., 2017b)	Network layer	(El Mouaatamid et al., 2016)
Chosen ciphertext attack	A CCA is an attack model for cryptanalysis in which the cryptanalyst gathers information, at least in part, by selecting a ciphertext and obtaining its decryption under an unknown key.	(Li et al., 2014a)	Transport Layer	Our contribution
Clone attack	In these kinds of attacks, the attacker captures and compromises legitimate node usually makes the clones.	(Du and Chen, 2008, Gope and Hwang, 2015)	Application Layer Network Layer Transport Layer	(El Mouaatamid et al., 2016)

Collision attacks	An attack on a cryptographic hash which tries to find two inputs producing the same hash value.	(Du and Chen, 2008)	Data Link layer, Network Layer	(El Mouaatamid et al., 2016, Borgohain et al., 2015, Ghildiyal et al., 2014, Karlof and Wagner, 2003)
Cryptographic overhead and WSN attacks	Security in IOT and WSN requires further research, although security elements exit in many protocols but the still specific type of security analysis is needed which causes crypto overhead and related cyber attacks.	(Fantacci et al., 2014)	Network Layer Transport Layer	Our contribution
Data Attacks (Modification and Injection)	Data can be compromised during its transmission as well as at rest on a device or an application's server.	(Barki et al., 2016b)	Application Layer	(El Mouaatamid et al., 2016)
Delay attack	In these kinds of attacks, an attacker intentionally puts a delay in sending or receive messages for some time to fail the time synchronization process.	(Du and Chen, 2008)	Physical Layer /Data Link Layer	MINE
DoS and DDoS	A denial of service attack occurs when one or multiple systems gets flooded with bandwidth or services	(Barki et al., 2016b, Challa et al., 2017,	Physical Layer Data link Layer	(El Mouaatamid et al., 2016, Ghildiyal et

	intentionally or unintentionally by an attacker.	Ning et al., 2013, Luong et al., 2016, Qiu and Ma, 2016, Roman et al., 2013, Du and Chen, 2008, Liu and Sun, 2016, Sun et al., 2017b, Giuliano et al., 2017)	Network Layer Transport Layer	al., 2014, Jing et al., 2014)
Eavesdropping	Eavesdropping is an unethical process of listening to the private conversation between people without their consent	(Challa et al., 2017, Ning et al., 2013, Roman et al., 2013, Wang et al., 2014, Sun et al., 2017b, Barki et al., 2016b)	Network Layer Transport Layer	(El Mouaatamid et al., 2016, Jing et al., 2014)
Exhaustion	The exhaustion attacks are computer security threats capable of crashing, hanging, or other sorts of	(Du and Chen, 2008)	Data Link layer	(El Mouaatamid et al., 2016, Borgohain et

	interferences with the victim.			al., 2015, Ghildiyal et al., 2014)
Fake sensing attacks	Crowdsensing networks are vulnerable to faked sensing attacks by users causes sensing costs and privacy leakage	(Luong et al., 2016)		
Firmware attack	Flaws in a firmware of IoT devices leading an attacker to modify the firmware and replace it with his malicious one to achieve his goals.	(Liu and Sun, 2016)	Data link Physical layer	
Flooding (incl. ICMP & Hello flooding)	To bring down the entire network or the services by flooding it with enormous amounts of traffic.	(Du and Chen, 2008)	Network Layer Transport Layer	(El Mouaatamid et al., 2016, Borgohain et al., 2015, Ghildiyal et al., 2014, Karlof and Wagner, 2003)
GPS deception	In this kind of attack, an adversary can provide a node with fake information about its location.	(Sun et al., 2017b)		
Hardware-Based Attacks	Various hardware-based attacks have emerged for example stealth backdoor circuits or trojans to steal precious patient information.	(Yasin et al., 2017)	Physical layer	

Illusion attacks	In this kind of attacks, some voluntary sensors that generate false or meaningless information in the network will be placed. These malicious sensors are always properly authenticated and identified in some way or other.	(Sun et al., 2017b)		
Impersonation attacks	An attack in which an attacker successfully assumes the identity of one of the legitimate parties in a system or in a communications protocol.	(Wang et al., 2014, Barki et al., 2016b, Gope and Hwang, 2015, Challa et al., 2017)	Network Layer Transport Layer	(El Mouaatamid et al., 2016)
Insider attacks	Malicious attacks executed (intentionally or unintentionally) on a network or computer system by a person with authorized system access. This attack has also been named as "privileged insider attack."	(Kumarage et al., 2016, Challa et al., 2017)	Physical layer, Application layer Link Layer	(Karlof and Wagner, 2003)
Internal attacks, vulnerabilities caused, software vulnerabilities and software	A negative use of programming to harm people or network environments.	(Yu et al., 2017, Chiang and Zhang, 2016, Ning et al., 2013,		

modification, etc.		Kumarage et al., 2016, Alrawais et al., 2017a, Liu and Sun, 2016, Sawand et al., 2015, Barki et al., 2016b)		
Jamming	Jamming attacks prevent nodes from using the channel to communicate by occupying the channel that they are communicating.	(Du and Chen, 2008, Ning et al., 2013)	Physical Layer Data link layer	(El Mouaatamid et al., 2016, Borgohain et al., 2015, Ghildiyal et al., 2014, Karlof and Wagner, 2003)
Logical Attacks	Targeting the proper functioning of a system without making any changes to the device's software	(Barki et al., 2016b)		
Man-in-the-Middle attacks	In MIMA, the attacker secretly relays and possibly alters the communication between two parties who believe they are directly communicating with each other.	(Liu and Sun, 2016, Challa et al., 2017, Giuliano et al., 2017, Qiu and Ma, 2016)	Network Layer Transport Layer	(Jing et al., 2014)

Masquerading	Pretending to be something or someone it's not.	(Giuliano et al., 2017, Sun et al., 2017b, Du and Chen, 2008)	Application Layer	
Message manipulation attack (MMA)	A MMA is an attack to manipulate a communication message sent by the devices.	(Du and Chen, 2008)	Application Layer	(Borgohain et al., 2015)
Mole attacks	Motion sensors in a smartwatch could leak personal information when a user types on a laptop keyboard, which is referred to as a mole attack	(Liu and Sun, 2016)		
Mule attacks	Adversaries may manipulate the local environment to fool sensors to record activities to achieve credits	(Liu and Sun, 2016)		
Network traffic analysis	Analysis of network traffic behavior/patterns by a passive attacker to steal information.	(Giuliano et al., 2017, Ning et al., 2013)	Network Layer	
Offline password guessing attack	An offline attack attempts to emulate the password and requires a known output of that process.	(Challa et al., 2017)		
Physical attacks targeting physical layer	Any malicious attack focusing on the physical layer	(Barki et al., 2016b)	Physical layer	(Borgohain et al., 2015)

Physical-level malicious attacks	Attacks targeting physical layer for example data slurping in which an attacker can steal data using even an iPod.	(Ciccozzi et al., 2017)	Physical layer	(Borgohain et al., 2015)
Relay attacks	An adversary may conduct a relay attack to make an entity believe that it is in the vicinity of the sender or receiver	(Barki et al., 2016b)	Physical Layer Data link Layer	(El Mouaatamid et al., 2016)
The release of Message Content attack	A passive attack in which a mail message, phone call or important messages would be intercepted or listened to is called the RoMCA.	(Giuliano et al., 2017)		
Replay attacks	Replay attacks have a unique feature, i.e., it can be conducted by illegitimate nodes. A lot of message replays increase the cost of precious bandwidth, resulting in the dropping of priority messages from the queue.	(Ning et al., 2013, Du and Chen, 2008, Wang et al., 2016c, Sun et al., 2017b, Giuliano et al., 2017, Gope and Hwang, 2015, Challa et al., 2017)	MultiLayer Attacks Transport layer	(El Mouaatamid et al., 2016, Borgohain et al., 2015, Jing et al., 2014, Karlof and Wagner, 2003)

Route attacks including Denial of services	Malicious nodes in the network modify the routing information or change the number of hops in forwarding routing request packets causing a denial of service	(Sun et al., 2017b)	Network Layer Transport Layer	(El Mouaatamid et al., 2016)
Secrecy attack	The secrecy attacks steal data by eavesdropping or interception.	(Sun et al., 2017b)	Network Layer Transport Layer	
Selective forwarding attack	Where compromised node drops packets selectively	(Du and Chen, 2008)	Application Layer, Network Layer, Routing layer	(El Mouaatamid et al., 2016, Pongle and Chavan, 2015, Borgohain et al., 2015, Ghildiyal et al., 2014, Karlof and Wagner, 2003)
Side channel attack	These attacks could be based on either power consumption, timing information, fault or electromagnetic leaks and enable the retrieval of the "used" secret keys.	(Barki et al., 2016b)	MultiLayer Attacks	(El Mouaatamid et al., 2016)

Sinkhole (black hole) attack	where compromised node tries to attract network traffic by advertising its fake routing update and attract more attacks	(Du and Chen, 2008)	Network Layer Transport Layer Application Layer	(El Mouaatamid et al., 2016, Pongle and Chavan, 2015, Borgohain et al., 2015, Ghildiyal et al., 2014, Jing et al., 2014)
Skimming attack	The wireless interception of information extracted from RFID chip-based debit, credit and ID cards and other documents, such as passports.	(Ning et al., 2013)	Network Layer Transport Layer	
Spoofing attack	Emulate/Imitate/Reproduce something while exaggerating its characteristic features for comic effect.	(Barki et al., 2016b, Ning et al., 2013)	Network Layer Transport Layer	(El Mouaatamid et al., 2016, Borgohain et al., 2015, Ghildiyal et al., 2014, Karlof and Wagner, 2003)
Stolen smart card attack	An attack to steal information about smart card	(Challa et al., 2017)		
Surface attacks	Attacks on a physical surface layer of an IoT	(Minoli et al., 2017)		

	device to steal private information.			
Sybil attack	A Sybil attack is an attack that the adversary forges one or multiple identities.	(Du and Chen, 2008, Sun et al., 2017b, Qiu and Ma, 2016, Dong and Liu, 2015)	Physical Layer Data link Layer Network Layer	(El Mouaatamid et al., 2016, Pongle and Chavan, 2015, Borgohain et al., 2015, Ghildiyal et al., 2014, Karlof and Wagner, 2003)
Tampering	Interfere with a system or process to cause damage or make unauthorized alterations.	(Ning et al., 2013, Du and Chen, 2008)	Physical layer	(El Mouaatamid et al., 2016, Borgohain et al., 2015, Ghildiyal et al., 2014, Karlof and Wagner, 2003)
Target tracking		(Ning et al., 2015)		
Traceability attack	Traceability attacks pose a threat to the privacy of users carrying the compromised IoT device.	(Bu et al., 2017)		
Unavailability of	Privacy-preserving methods are open issues in IoT	(Mayer, 2009)	Physical Layer	

communication caused by botnets and Distributed Denial of services attacks	causing communication between devices disturbed. This often results in Denial of Services.		Data link Layer Network Layer Transport Layer	
User account injection attack	Attack similar to SQL injection but with the user account to compromise a device.	(Liu and Sun, 2016)	Application layer, Presentation, Session layer	(El Mouaatamid et al., 2016)
Wormhole attack	The fundamental idea of wormhole attack is that two or more malicious nodes hide the true distances among them entice other normal nodes to route across these dangerous nodes to absorb data flow and cause network conjunction or cooperate with other attackers.	(Sun et al., 2017b, Du and Chen, 2008)	Physical Layer Data link layer Network Layer	(El Mouaatamid et al., 2016, Pongle and Chavan, 2015, Borgohain et al., 2015, Ghildiyal et al., 2014, Karlof and Wagner, 2003)

The table mentioned above is self-explanatory that series of malicious attacks are targeting IoT. Now a day's most of the security professionals have to defend their organizations from cyber criminals trying continuously to steal the wealth of public or private information. This information can be passwords, financial information, health records or anything valuable. Attackers reap the benefits of their malicious intent by trying and infecting more and more resources, for example, email servers, databases, surveillance camera, search engines, corporate servers and use them as sources to cause more damage. No matter, what's the intent, cybercrime cannot be started or completed without a malware. Nowadays, malware development is an industry where the author's contribution is only to write the code and get their share, a team of criminal marketers does rest of the job.

There is no such thing like 100% security or safety, as soon as billions of interconnected IoT devices join the network, they get a gift of associated challenges. These challenges can be of any type described in above tables (X, X). Solution to these challenges requires not only human efforts but also involve massive investment of resources to mitigate. It is understood that malware plays a vital role in the breach of security, privacy, and trust. Therefore, it is necessary to study the various aspects of IoT malware to prevent future attacks.

2.7 – Overview of malware

The purpose of this section is to introduce malware analysis, detection approaches and various studies aiming to perform the proposed research, i.e. malware analysis, detection and classification.

2.7.1 What is malware?

Malware is a menace to the society, in other words, an adverse use of application development aiming to harm the general public and organization. There are various definitions of malware, but in simple words, it is a negative programming force that is being used for destructive purposes by cybercriminals. Software developed by certain individuals and spread by one or many with intentions to cause loss of money, reputation and grievance to the people or organizations.

2.7.2 Characteristics of malware and its variants

Malware has various characteristics could be referred as deception/destruction capabilities, but following four features make them more harmful: (i) Stealthy behavior; (ii) poly/metamorphic nature; (iii) armor capabilities; (iv) obfuscation of code (Chen et al., 2012).

Stealthy behavior: This characteristic makes it capable of hiding the system activities while during infection and later stages. During stealth, process attacker tries to control by occupying registry and user/system files, etc. Furthermore, once it holds the system entirely, it hides itself from system processes to avoid discovery by anti-malware software.

Poly/metamorphic nature of malware: A perfect implementation of object-oriented programming to implement malware code in such a way that a malware automatically creates multiple variants without changing core functionality but targeting different victims. In metamorphic malware not only code sequence may be altered dynamically but also changes functionalities as well.

Armor capabilities: A modern-day malware is capable of detecting that someone is trying to debug or reverse engineer (a standard function used for this called "IsDebuggerPresent" it and automatically goes into an isolation state, this capability comes from the fact that they can find whether they are being run in a virtual environment like VMWare based virtual machine. Therefore, they hide their actual functions, imports, exports and sometimes system calls as well. That makes a malware analyst's job even harder.

Obfuscation of code: Refers to confuse, mislead, compress, encrypt or decode various coding elements of malware to hide their actual functionality and avoid presence in the system.

2.7.3 Malware variants

Malware has various forms classed into a Botnet, Worm, Ransomware, Rootkit, and Trojan (shown in table), each of these families is dangerous enough to cause too much financial loss particularly if we talk about Botnets and Ransomware, they are alarmingly dangerous. In chapter 1 we discussed the severity of IoT Botnet Mirai which caused above 1TB distributed denial of service attack on Dyn and caused over $110 million to the company. Furthermore, ransomware can be considered as modern-day kidnaps, kidnapping of highly valuable system information and ask for money to release them.

Different types of malware	
Malware type	**Description**
Botnet	Bots have the ability to compromise one or more machines, use them as attack source to target more victim machines.
Worm	Worms replicate themselves in a compromised machine by disabling security features.
Ransomware	Modern-day form of kidnapping and compromising data by encrypting it and getting the money to release.
Rootkit	Compromise machines without getting detected.
Trojan	Disguised as a good application to gain access.

Table 4: Different types of a malware

2.7.4 Malware analysis techniques

Malware analysis techniques are used to read the patterns of malware by either reading their code statically or getting information by executing them (whether it is an isolated/artificial environment on live). In general malware analysis techniques can be divided into two categories, i.e. (i) static malware analysis; (ii) dynamic malware analysis techniques, but we will be dividing dynamic malware analysis into the third one of network traffic analysis technique as shown in figure 9 (Damodaran et al., 2017). In static analysis, researchers use two approaches to reading the code either by de-compilation /reverse engineering or just with the help of some text reading parsers developed using scripting languages like Python, C++ or Java. It's relatively faster and less time-consuming. While dynamic analysis deals with the study of malware behavior while executed (Ravula et al., 2011).

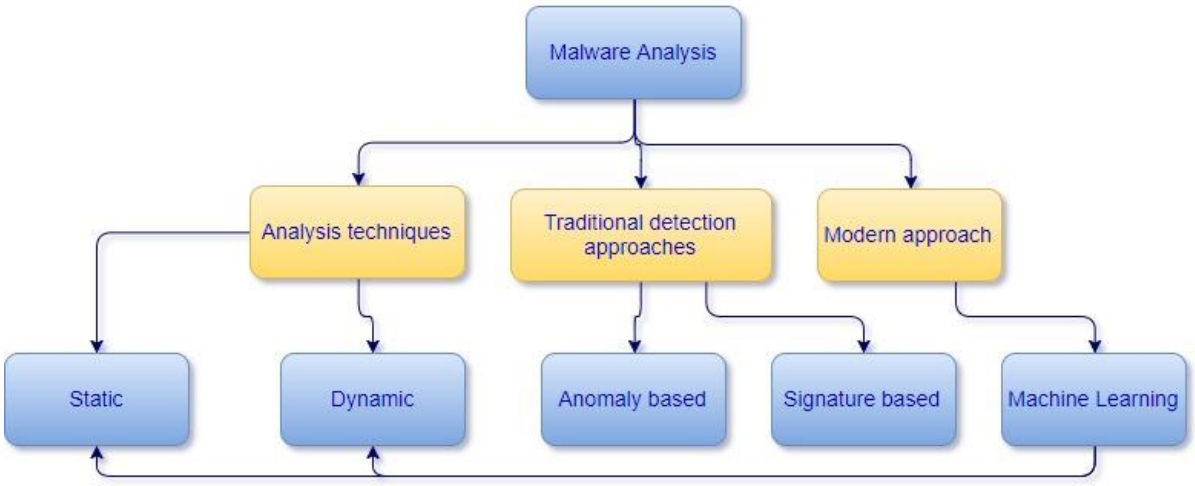


Figure 10: A structural explanation of malware analysis

Both static and dynamic methods are cost-effective and accurate like others both of these techniques have their pros and cons. For example, when researchers started analyzing malware, malware authors began obfuscating the malware to hide malicious intentions and making static analysis hard to perform (Borello and Mé, 2008). Nowadays, polymorphism and metamorphism are the familiar concepts those used in malware creation with the purpose of deception, code reuse, and faster penetration. Table 8 shows some features of static malware. Malware investigation is always performed in a controlled environment either with the help of some specialized sandboxes or some virtual machines to avoid infection to other machines (Wrench and Irwin, 2015). For the scope of this document, we will be focusing on only static malware analysis.

2.8- Machine Learning

In the field of computer sciences machine learning is a sub-branch of artificial intelligence that uses the algorithms to automate the analytics. Application of machine learning is everywhere and there is a possibility that you have been using it without knowing anything about it, e.g. to find out daily trends of stock in the stock market, another example is biomedical research where expression of some medical concept can be studied, etc. Mostly machine learning is applied in those research areas where future prediction is required by previously held information to improve the performance of existing software solutions.

The used algorithms in machine learning can be either supervised or unsupervised. When the dataset is labeled then supervised learning is used. In supervised learning dataset, each algorithm has a common principal, i.e. “predictive modeling” that model the data to find the trends or structure of the data to make predictions. Algorithms used in supervised learning are also known as classifiers or classification algorithms. Whereas, in unsupervised learning, the dataset is unlabeled and there is no target class defined that helps to discover the unknown classes. Algorithms used in unsupervised learning are known as regression algorithms.

For data analysis where a target class is present classification algorithms are mainly used. The following table shows the most commonly used classification algorithms.

Table 5: Machine learning algorithms and their purpose

Algorithm	Purpose
Bayesian Network	Probabilistic/statistical modeling in complex or uncertain cases, fast analytical response, support best decision analysis, possibility to use information from different sources for analysis (Uusitalo, 2007).
Decision tree (C4.5)	Builds decisions like a tree, feature selection, classification (with minimum effort), easy interpretation of results, handles both continuous and discrete attributes, handles missing values, Prunes tree once created (Ali et al., 2012).
Chi-square Automatic Interaction Detector	Classification/prediction, interaction between variables, easy implementation/interpretation, non-parametric, powerful/quicker and cost effective (Rygielski et al., 2002).
Classification and Regression Trees	Good for both classification and regression analysis avoid data exhaustion
Decision Stump	One level decision tree, discriminates between two of three classes, help ensembles
Gradient Boosting	Good for both classification and regression, used for ensemble learning, boost weaker models.
High-Performance Neural	High-performance analytical procedure, runs on single/distributed modes, utilizes all available system cores and threads.
K-Nearest Neighbor	Pattern recognition, used for both classification and regression, from multiple points of multiple classes separate a new class.
Least Absolute Shrinkage and Selection Operator	Variable selection, regularization, enhance prediction accuracy, interpretable results and to reduce overfitting (Tibshirani, 1996).

Least Angle Regression	Fast forward feature selection, stepwise regression model, computationally fast, simple, cross-validation, identify possible correlation (EFRON et al., 2004).
Linear Regression	Prediction/forecasting, error reduction, to explain data, its relationship between a dependent variable and multiple independent variables.
Logistic Regression	To describe data, the relationship between one binary target and one or more other variables including categorical variables.
Naive Bayes	A family of classifiers, provide high scalability, so simple to implement, statistically independent classification.
Neural Network Multilayer Perception	Three layer neural network, helps to distinguish from inseparable data, capable to stochastically solve machine learning problems.
Neural Network Multilayer Perception Back Propagation	Most general, to calculate gradients, simplest, for non-linear approximation, the minimum value of error function is looked.
Partial Least Squares	Predicted modeling, help overfitting, feature reduction (Tobias, 1995).
Random Forest	The ensemble of classification and regression, form multiple trees, flexible, best results, no need to tune parameters, select the best feature, overfitting, outlier detection and variable priority (Ali et al., 2012).
Stepwise Regression	Automatic independent feature selection, stepwise removal of predictors to build regression model (Lewis, 2007).
Support Vector Machine	Work well in high dimensional spaces, risk minimization, handle overfitting, the margin of separation, where no. of dimensions > samples, memory efficiency (Schwarm and Ostendorf, 2005).

Malware detection and role of machine learning

Tradional malwares are a history, it is a time in which highly-sophisticated malware attacks are happening specially when the importance of IoT devices is growing day by day, various research studies have been conducted on malware analysis, but still, continuous threats were coming up every day. So, researchers started involving machine learning to study the different aspects of malware. It is important to note that till this date there has been no evidence of research on IoT malware analysis with the help of machine learning. The behaviour of both IoT and non-IoT malware is still un-tested.

With the help of machine learning algorithms, it is possible to learn the pattern of new and old malware, prediction of further attacks and enhancing the malware detection systems.

Traditional malware detection systems and intrusion detection systems used both signature-based analysis and anomaly-based malware detection. Anomaly-based detection uses behavior patterns from network traffic. Both of these methods again have strengths and weaknesses covered in the dynamic analysis. Dynamic malware investigation is time-consuming and slow. Besides weaknesses of both static and dynamic, they are effectively used along with supervised and unsupervised learning models. Machine learning in malware detection helps in predicting abnormal patterns in an efficient way which saves computational overheads. It also describes the quality of the classifier on given circumstances and its ability to detect outliers (Srndic and Laskov, 2016). Most popular classification methods include Neural Network, Support Vector Machine, Decision Tree, and Random Forest, also shown in our classifier selection in the figure and table 6 shows a summary of the literature of malware analysis with machine learning.

Table 6: Overview of literature concerning different platforms, their features extracted, classifiers and accuracies about static malware analysis

Reference	Platform	No. of Malware	No. of Benign	Feature Extracted	Classifier	Accuracy
(Wang et al., 2016a)	Android	20045	20023	Meta data	1,2,3,4,5	94
(Lo et al., 2016b)	Windows	7630	1818	Static (File Info, version, properties, PE Info, metadata, behavioral info)	1,5,6	99.6
(Kühnel and Meyer, 2016)	Mobile	2441	4539	Character set	5, 7, 4,1,2	90+
(Kang et al., 2015)	Android	4554	51179	APIs, Permissions, serial number of certificate	2	90
(Ding et al., 2016)	Windows	2000	2000	Opcode (n-grams)	5, 7, 4	96.2
(Cepeda et al., 2016)	Windows	7630	1818	Static (File Info, version, properties, PE Info, metadata, behavioral info)	1,5,6	99.6
(Baldangombo et al., 2013)	Windows	236756	10592	Header, DLLs, API calls,	5, 8, 2	99.6
(Santos et al., 2013a)	Windows	13189	13000	Opcodes	7,8,5 1, 2,9	95.90
(Islam et al., 2013a)	Windows	2398	2008	Function length frequency vectors, Printable Strings	5, 4,1,10	87.81
(Kolter and Maloof, 2004)	Windows	1651	1971	Byte sequences	2,4,5,11,12,13, 14,15	-
(Milosevic et al.)	Android	200	200	Permissions and source code	1,2,5,16,17,18	95.6
(Nath and Mehtre, 2014b)	Windows	9458	123	Image Visualisation	7	98.08
(Yuan et al., 2016)	Android	1760	20000	Permissions, file contents, API	6	96.76
(Adebayo et al., 2014)	Android	1000	500	Byte code	19	97.20
(Fereidooni et al., 2016)	Android	18677	11187	Intents, used permissions, APIs, IMEI	1,2,3,4,5,6,7,8	97.30

Key: 1= Random Forest, 2=Naïve Bayes, 3=Logistic Regression, 4=Decision Tree, 5=Support Vector Machine, 6=Neural Network, 7=K-Nearest Neighbour, 8=Decision Tree (J48), 9= Bayesian Network, 10=IB1, 11=IBK, 12=Boosted naive Bayes, 13=boosted SVM, 14=Boosted decision trees, 15=TFIDF, 16=C4.5 Decision Tree, 17=JRip, 18=AdaBoost, 19=Association rule

Static malware analysis

Since smart devices have emerged in the industry particularly mobile phones, their security issues increased as well. People use these devices to store historical data, passwords, contacts, picture, videos, account information and much more valuable information (Arzt et al., 2014). To understand the threats related to mobile devices, their analysis (static) and application of machine learning. We selected papers from ISI Web of Science database from 2010-2017. Table 9 presents information about various static features related to mobile research. Our chosen articles were consisting of Android mobile devices which is one of the most extensive mobile operating system in the market (Atwal).

In addition to the growing importance of Android devices, people tend to analyze loopholes behind these devices. Malware authors create malware targeting weaknesses of these devices and steal sensitive information from these devices. Recent studies show that mobile devices mainly android platform are the prime targets of attackers (Symantec, 2017). An example of a simplest everyday attack, if someone's mobile gets compromised and gets charged for sending premium rate messages then this becomes a serious issue, malware detection is a serious concern for not only the general public but for organizations as well using Android devices to control IoT devices.

Further to our current discussion, Table 9 illustrates a list of some important static features and few relevant articles in which these features have been used. These selected articles also apply machine learning classifiers to them and explain various performance measure which we will explain later on. Here we will discuss two important papers regarding static feature extraction, classification, and efficient detection mechanism. One article is ANASTASIA, a framework which analyses various static features including API calls, IMEI addresses, malicious intentions and user permissions. ANASTASIA performs classification using important classifiers, e.g. SVM, KNN, DT, NB, Boosting techniques and Deep Learning as well. Along with this author also handle imbalanced dataset which now a day's researchers ignore to analyze. As a result of classification, we can get the accuracy of 97.3%, the reason to choose this paper is that it performs various performance measures include Accuracy, TPR, FPR, Precision, Recall, and F1-score, etc.

Another example of an excellent paper, Droid Api Miner (Aafer et al., 2013) performed static malware analysis using features like API calls, Opcodes extracted from Byte Code, etc. and uses DT, C4.5 DT, KNN, and linear SVM as classifiers resulting in an accuracy of 99%. These

continuously growing research trends are because of its enormous importance in IoT domain, and the majority of Android apps are free of charge.

Feature extraction and current literature	
Feature extracted	**Reference**
API	(Cho et al., 2017, Kang et al., 2015, Dhaya et al., 2014, Fereidooni et al., 2016, Geneiatakis et al., 2015, Yerima et al., 2015b)
Strings	(Cho et al., 2017, Sanz et al., 2013)
Bytes	(Santos et al., 2011, Adebayo et al., 2014)
URL	(Thomas et al., 2011)
Permissions	(Xu et al., 2013, Kim et al., 2015, Kang et al., 2015, Fereidooni et al., 2016, Geneiatakis et al., 2015, Su et al., 2016b, Su et al., 2016a, Feizollah et al., 2017, Lopez and Cadavid, 2016, Akhuseyinoglu and Akhuseyinoglu, 2016, Yerima et al., 2015b)
Java code	(Feizollah et al., 2017)
Network address	(Feizollah et al., 2017)
Hardware components	(Feizollah et al., 2017)
Intent filters	(Fereidooni et al., 2016, Su et al., 2016b, Su et al., 2016a, Feizollah et al., 2017, Yerima et al., 2015b)
User flow	(Brown et al., 2016)
Opcode	(Damodaran et al., 2017)

Table 7: Showing static features extracted in various literature

Feature selection methods in machine learning

The discussion on machine learning and classification algorithms will be incomplete without discussion of feature selection methods. Therefore, for our research as well, this is an important step. When performing malware analysis and feature list is too big, it is important to reduce the features to a minimum set without compromising the performance of classification algorithms being used and making analysis process smooth. It will not be wrong if we say that in majority of instances least the features are higher will be the detection rate, and overall accuracy of the

algorithms. Furthermore, it is also important for forensic experts to put more focus on important features filtered out during process of feature selection.

For background knowledge of feature selection methods, we used comprehensive literature review on feature selection methods being used in the literature and summarised our findings in the tables below. Table 8 shows various feature selection methods used in diverse literature. This table can used to compare the methods used in the literature and our contribution towards feature selection methods.

Feature selection methods used in literature	
Method	**Reference**
Information Gain	(Akhuseyinoglu and Akhuseyinoglu, 2016)
Gain ratio	(Yerima et al., 2015b, Yerima et al., 2015a)
Chi-square	(Feng et al., 2017, Lopez and Cadavid, 2016) (Akhuseyinoglu and Akhuseyinoglu, 2016)
Association rule with apriori algorithm	(Adebayo et al., 2014)
Ensemble of randomized decision tree	(Fereidooni et al., 2016)
Fischer score	(Cohen et al., 2016)
Top feature	(Cohen et al., 2016)
Deep belief network	(Yuan et al., 2016)

Table 8: Showing various feature selection method

Performance measures of Static malware analysis

In this section, we will discuss various performance measures widely used in the literature by the researchers to predict the performance of classification algorithms, for example, better the accuracy of an algorithm, better would be the power of differentiation between malware and benign application. We are listing standard performance measures and their formulas in table 10 below. In Table 10 and 11, some recent papers have been mentioned with their performance measures and formulas to calculate them. TP represents a malware correctly identified, FP represents a goodware correctly identified, TN and FN are the false alarms of both malware and goodware wrongly identified. To conclude the discussion on performance measures, we can say that these measures play an important role in evaluating the work.

Performance Measure	**Formula**
TPR or Sensitivity	TP/(TP+FN)
FPR or (1-TNR)	TN/(TN+FP)
TNR or Specificity	FP/(FP+TN)
FNR	FN/(FN+TP)
Accuracy	(TP+TN)/(TP+TN+FP+FN)
Precision	TP/(TP+FP)
Recall	TP/(TP+FN)
F1-score	(2 ⁎ TPR ⁎ Precision) / (TPR + Precision)

Table 9: Showing performance evaluation measures used in this study

Important classifiers employed in the literature					
Reference	**Classifier**	**TPR**	**FPR**	**AUC**	**ACC**
(Wang et al., 2016a, Wang et al., 2016d)	Naïve Bayes	0.77	0.07	0.92	85
	Bayesian Network	0.87	0.08	0.96	89
	Support Vector Machine	0.92	0.07	0.92	92
	Logistic Regression	0.90	0.07	0.97	91
	Decision Tree with J48	0.92	0.06	0.96	93
	Random Forest	0.94	0.05	0.98	94
(Santos et al., 2013a, Santos et al., 2013b)	Support Vector Machine	x	0.02	0.95	95.90
	K-Nearest Neighbour	0.95	0.05	0.95	94.83
	Decision Tree	0.93	0.08	0.93	92.61
	Random Forest	0.96	0.06	0.99	95.26
	Naïve Bayes	0.90	0.10	0.93	90.02
	Bayesian Network	0.91	0.04	0.98	93.40
(Feng et al., 2017)	V-SVM	0.91	18.21	0.9643	91.29
(Yerima et al., 2015a)	Decision Trees	0.95	0.04	0.964	95.4
	Random Trees	0.96	0.04	0.960	95.9
	Naïve Bayes	0.82	0.08	0.88	91.5
	Random Forest	0.97	0.02	0.992	97.4
(Cohen et al., 2016)	Random Forest	0.97	0.05	0.9927	97

Table 10: Recent papers and their performance measures

Preventive approaches to cope with malware attacks on IoT

Possible prevention is to know what kind of malware attacking, knowing the characteristics Due to the scope of this report, we will only be discussing one static malware analysis. Future works may involve addressing other challenges as well.

The conclusion of literature review

The literature review showed us that in static malware analysis sufficient work in being done in Windows and Android-based environments but there is no work done in the field of IoT malware analysis (Please note: with IoT malware analysis we mean malware analysis of information obtained from IoT). Further to this, there is a need to address efficient malware detection system by calculating misclassification cost, class imbalance, and optimization of used classifiers, we aim to cover these issues with the help of cost-sensitive learning. Various classifiers and feature selection methods have been used for malware classification and prediction using WEKA and Matlab. However, we opted to use SAS Enterprise Miner 14.2 (latest version) for the classification and feature selection. Because of being a licensed software, researchers usually do not often use this software but still an excellent software. This research will be giving us a chance to evaluate this software and pinpoint pros and cons of this software.

CHAPTER 3: RESEARCH METHODOLOGY

Chapter 3: Research methodology

3.1 - The proposed research stages

A pilot study has been performed and proposed research stages consist of six main activities of data collection, feature extraction, feature selection and evaluation, classification, test and generalization check. The figure below represents all of these stages.

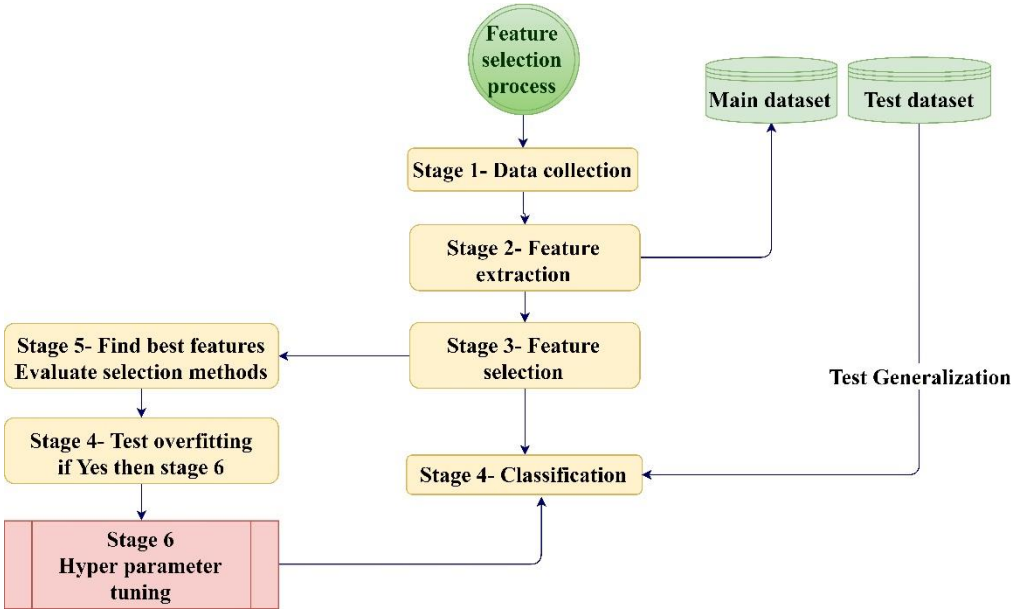


Figure 11: Research proposed framework

3.2. Data collection

This section presents the process of data collection and feature extraction. Collected a sample dataset of 1102 IoT ELF malware downloaded from VXHeaven, having a class of UNIX System V and platforms. A total of 1083 clean application (ARM) collected using Qemu emulator from the Raspbian operating system. Figure 12 shows both malware and goodware samples.

In total, there are 2085 files, these malware from various platform help us to test issues like class imbalance and generalization of the algorithms. Using a controlled environment, all clean applications were analyzed via VT to check the possibility of being infectious, ignored some so-called clean applications downloaded from Contigo website found infected while scanning through VT. Some of the malware were found to be packed with UPX packer while analyzing them VT (Cross checked using our selected DE complier-IDA Pro 6.9), unpacked them with the corresponding unpacker downloaded from UPX website.

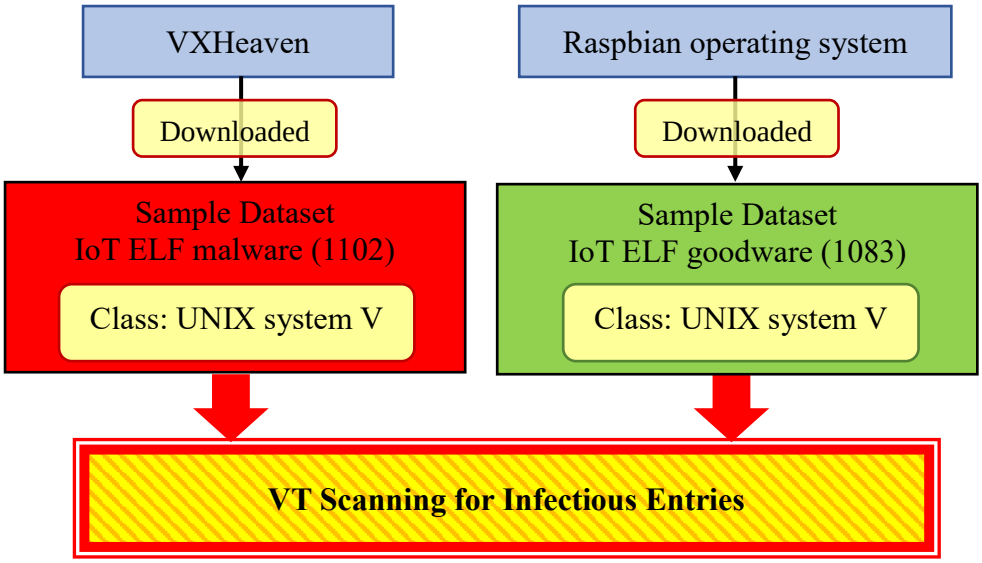


Figure 12: Data collection process

3.3. Feature extraction

After malwares get checked up for packing/unpacking issues, next step is feature extraction. "ReadELF" a python script (A static feature extraction script) was used to extract features from each malware and goodware, and the data was cross-checked with three static feature extraction tools in Linux called "ReadELF," "ObjDump" and "ELFParser." Table 12 shows datasets created after features extraction.

3.3.1. Datasets created

During data collection process, following datasets were generated, some of these datasets are complete and operational, and some are requiring further processing for example segment headers, functions, imports, and exports. Table 12 shows the list of datasets created in this research.

Datasets created			
Feature set	**Disassembled?**	**features**	**Stage**
File header	No	33	Done
Program header	No	10	Done
Section header	No	18	Done
Segment header	Yes	15	Future work
Symbol table	No	7	Future work
Strings	No	5	Future work
Function	Yes	10	Future work
Process names	Yes	5	Future work
Imports	Yes	5	Future work
Exports	Yes	5	Future work

Table 11: Datasets created

3.4 - Feature selection and evaluation

During our discussion about Machine learning in section 2.8, we discussed about feature selection and evaluation and presented a table in which various feature selection methods were introduced in almost all studies related to malware analysis. As stated previously, our purpose is to give readers an overview of some of the existing feature selection methods and use some additional methods which were never used in any of the research related to malware analysis which gives approximately zero percent possibility of use in IoT research.

The process of feature selection is used to perform dimension reduction of malware dataset and makes data easier to analyse. Data analysis with a vast amount of data requires more computational resources and a considerable amount of time. So, feature selection lets us remove noisy/useless features without losing efficiency and improves results. In our research we aim to select best feature selection methods to support our framework.

Our feature selection process includes the application of principal component analysis, using linear/logistic regression, decision trees, variable selection, variable clustering, statistics explorer methods backed by some additional feature selection methods applied using SAS Enterprise Miner 14.2 shown in table 14. Furthermore, our aim to use these feature selection methods as an input to our classifiers using same binary target variable and compare the list of selected features and find out best selector. In section 3.5- Figure 13 shows the proposed process of feature selection.

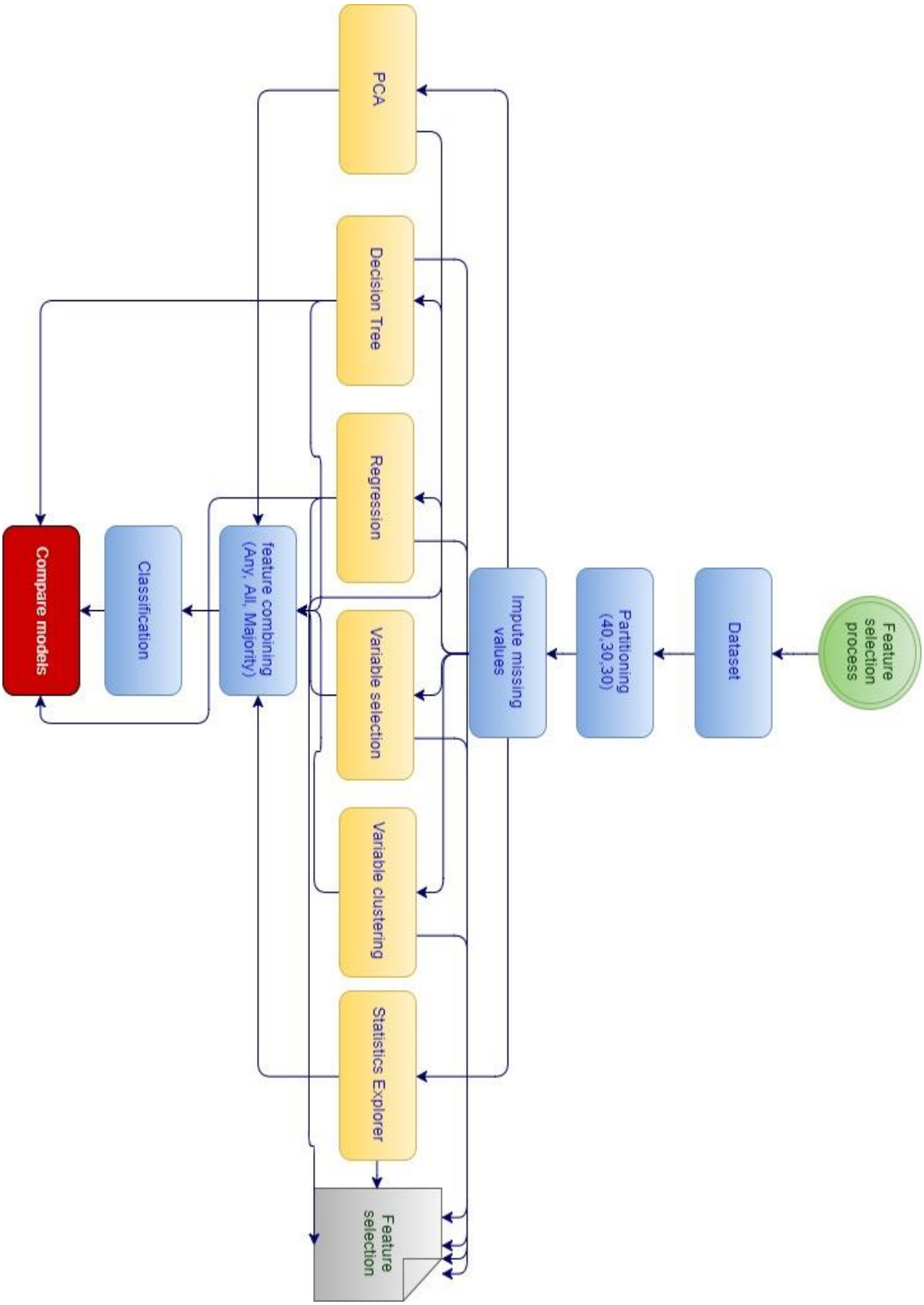


Figure 13: Feature selection process used in this dissertation

Our proposed feature selection methods	
Feature selection method	**Process**
R-square	Statistics explorer
Chi-square	Statistics explorer
R-square and Chi-square-both	Statistics explorer
Fast Selection	Variable selection
Least angle regression (LAR)	Variable selection
LASSO	Variable selection
Variable correlation	PCA
Full feature selection	Regression
Stepwise feature selection	Regression
Backward feature selection	Regression
Forward feature selection	Regression
Fast backward feature selection	Regression
Decision tree	Decision trees

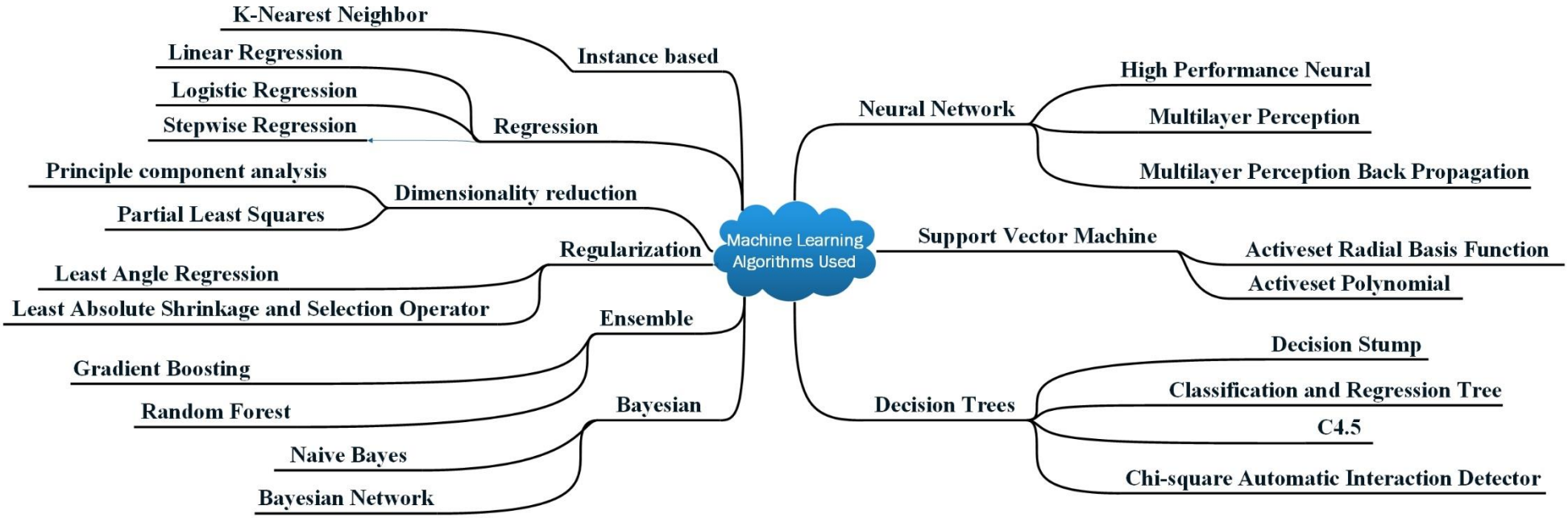
Table 12: Feature selection methods

3.5. Classification

Data classification is supervised learning technique in which there is a predefined target variable (In our case we have binary target also known as a positive class, represented by 0 for goodware and 1 for malware), in classification the original dataset was subdivided into training, test and validate having size 40%, 30%, and 30%. It helped us to determine the performance measures by calculating Accuracy of Classification Rate, the Area Under the Curve, Precision, Recall, and F1-score. For classification, the approach used contains process of importing a file, impute variables, pass variables to feature selection algorithms, apply classification algorithms and compare results. Figure 14 shows a list of our selected classifiers. During the classification process our aim was to apply classification/feature selection methods that have been used in the recent literature and available in SAS enterprise miner.

In machine learning we use multiple algorithms to perform spot checking on the dataset. The purpose of spot checking is to not only perform which algorithms performs well in the dataset which you do not know beforehand. To achieve better spot checking results, researchers use various kinds of algorithms with various implementations like trees, instances, linear or non-linear etc (Brownlee, 2016). Because datasets were created with the information extracted from malware full categorical values that is why we had to do spot checking for each dataset.

Figure 14: Machine learning classifiers used in our approach



3.6. Further experiments on the algorithms

During the process of classification, there are two important checks to be considered, to see whether (1) classifier is generalized? (2) Not overfitting? If overfitting, consider hyper tuning the parameters and find an optimized parameter. To check generalization, another dataset is being created which will contain malware samples to be checked for generalization. For example, our malware dataset consists of malware from advanced micro devices, **INTEL 80386, X86-64, SPARC, MIPS and MC63000**, same malware in a fraction present in primary dataset as well to train the classifiers.

CHAPTER 4: PRELIMINARY RESULTS AND DISCUSSION

Chapter 4: Results and discussion

This chapter presents the results of the current analysis. In section 3.4 we discussed the feature selection methods and classifiers used in our investigation. This chapter aims to present a quick look at feature selection results of the main header, program header, and section header features.

Feature worth

Figure 15 shows feature worth for the main header about the binary target variable category (category "1" represents malware and "0" represents benign app). As shown in the figure, the variables from sh4 onwards (includes many flag variables) can be ignored and may have next to none impact on the classification process.

Figure 16, shows variable worth for program header and indicates that only three flag variables containing malware or goodware access permissions have the least importance. These variables (Read, Write and Execute) can be dropped.

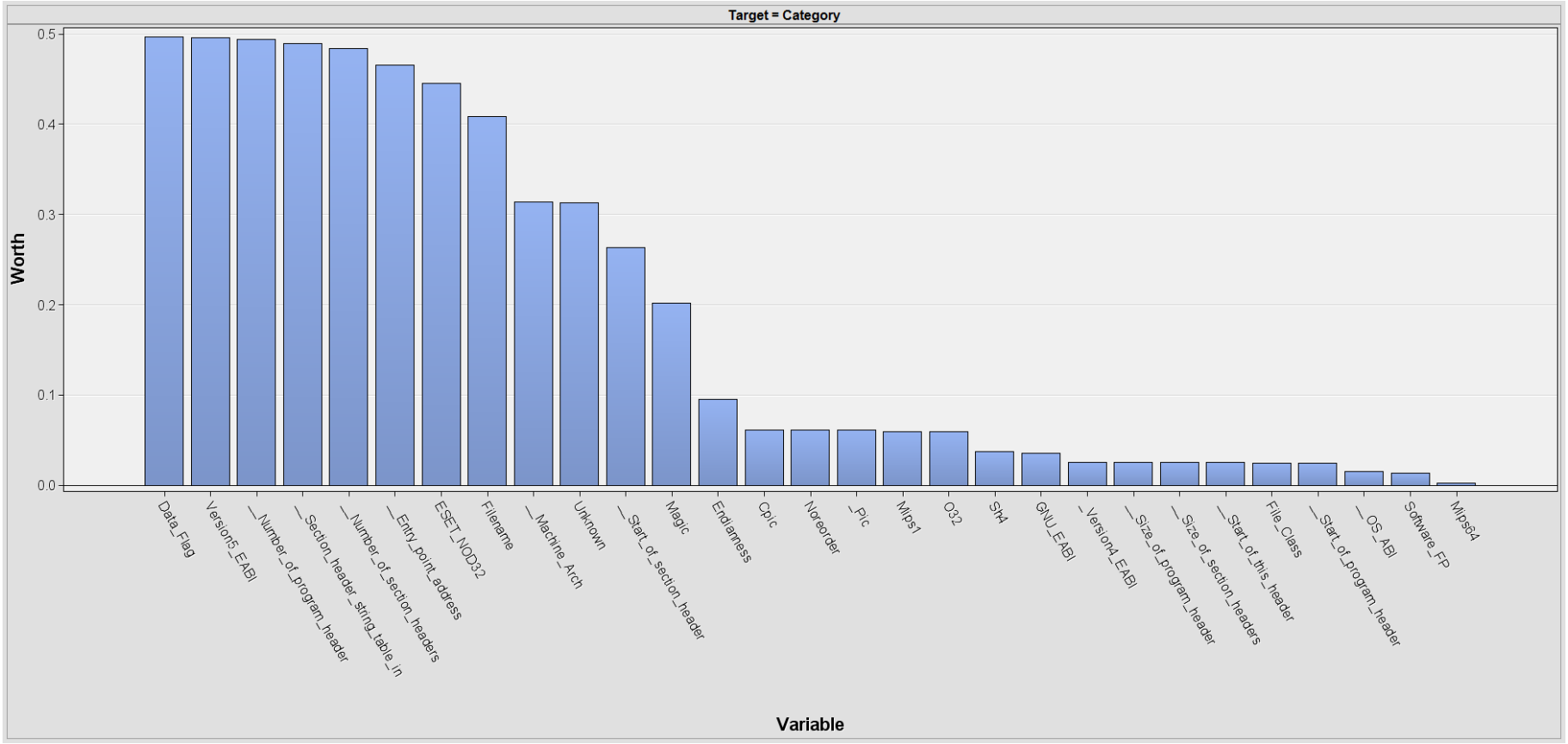


Figure 15: Main header variable worth

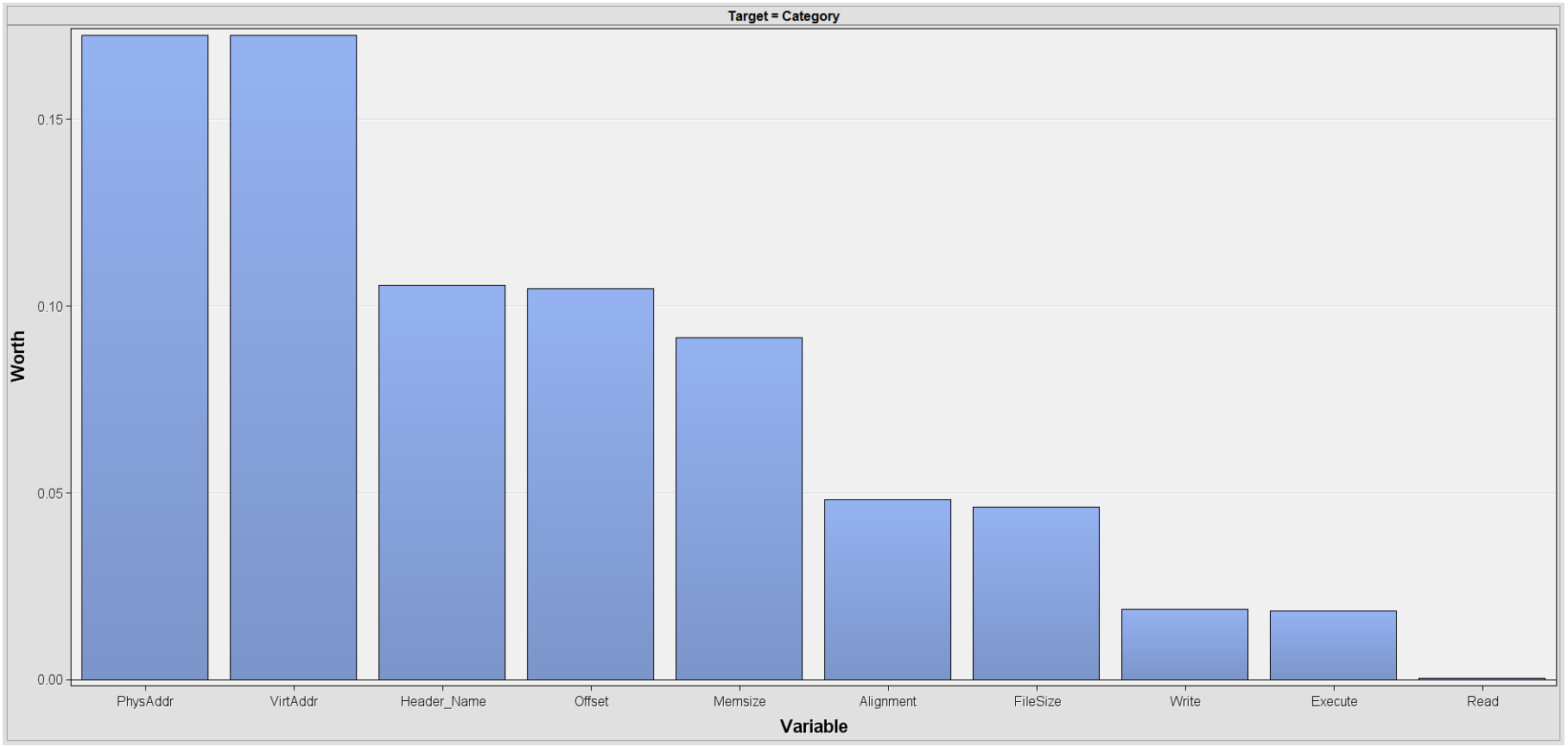


Figure 16: Program header variable worth

Similar kind of information can be seen in figure 15 for section headers. All **binary flag variables** have least worth. The variable of sh_size (size of current section header) and all flag variables can be ignored. In future investigations as shown in previous sections, we will try comparing various feature selection methods, group the results obtained from those methods using parameters in table 15 and feed them to the classifiers.

Feature combining rules	
Any	Any variable rejected by any of the feature selection methods will be ignored.
All	Variables rejected by all feature selection methods will be ignored.
Majority	Variable rejected by the majority of the feature selection methods will be ignored.

Table 13: Feature combining rules

In each dataset, Roc chart, output, and fit statistics were calculated for training, validation and test datasets with partitioning size of 40%, 30%, and 30% respectively.

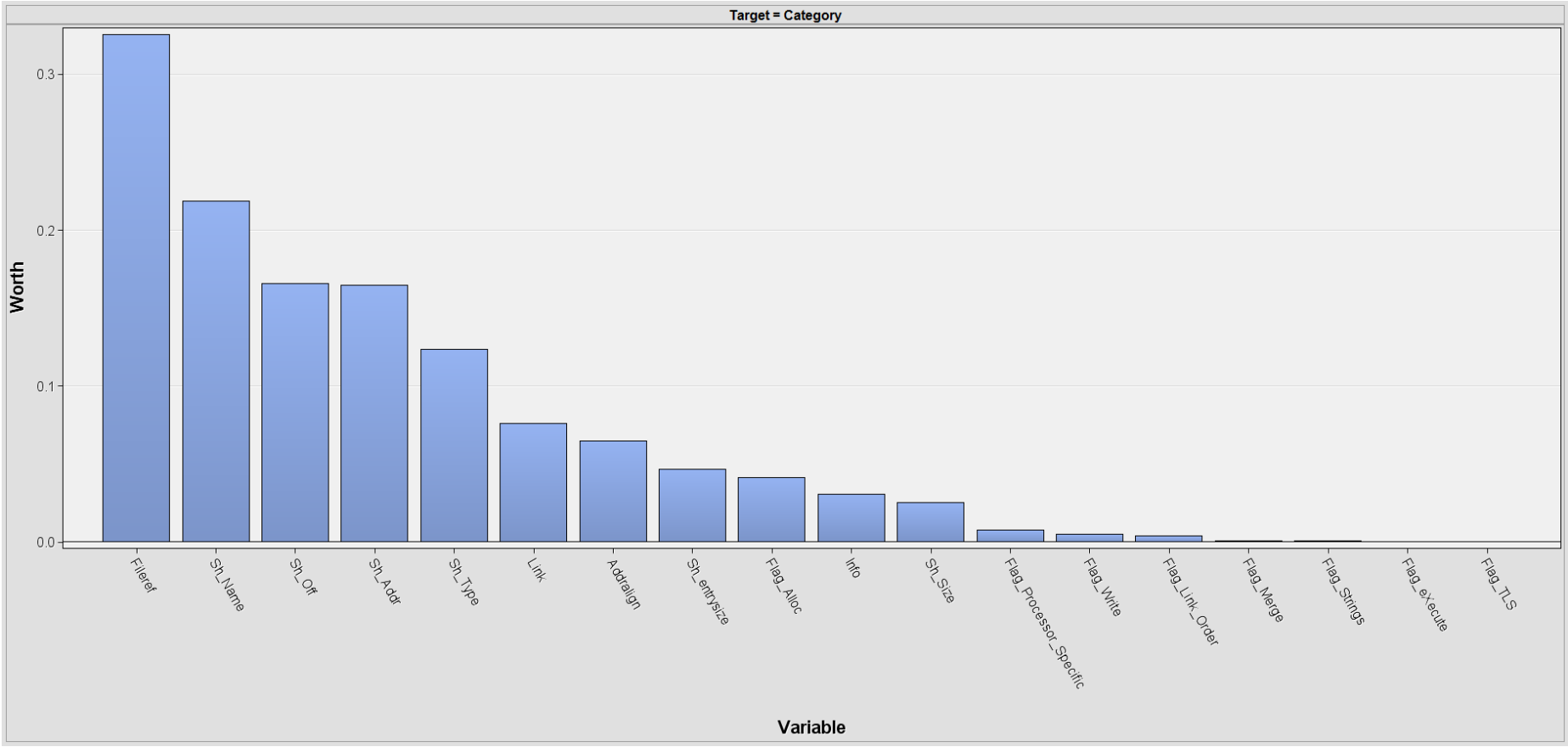


Figure 17: Section header variable worth

Malware classification and discussion of results.

For IoT malware classification, we used 20 classifiers on 3 datasets i.e. program header, main header and section header. Our results shown that each classifier gives different attributes of performance accuracies for the malware classification and prediction. Some of them performed very well and some of them did not. This situation prompts us to categories these classifiers in to 4 groups on the basis of their percentage of accuracy i.e. Group A (>=94%),

Group B (>=90%), Group C (>=80%) and Group D (<80%). In chapter-2 section 2.8, we summarized 15 published studies in Table 6, where classification was performed on Windows and Android based malware and goodware. Although there is no study published on IoT malware classification but it will be useful to compare our results with previously published studies.

There are five key findings from our research and key issues it raises for future malware analysis. First key finding by comparison of 20 different classifiers for IoT malware analysis was noted that Random forest outclassed all results with maximum performance in all datasets. This finding when comparing with previous studies, it has been noted that (Lo et al., 2016a, Baldangombo et al., 2013) and (Cepeda et al., 2016) also identified the highest performance (99.6%) with Random forest, although that classification was for Windows based malware. This finding justifies that Random Forest gives the best performance on IoT malware classification. However, an interesting thing was noted that (Lo et al., 2016b) and (Cepeda et al., 2016) used imbalanced dataset and did not specify the sampling method used for the classification, therefore, their accuracies may be overestimated. The same finding was observed in all the previous studies (Baldangombo et al., 2013, Nath and Mehtre, 2014a, Yuan et al., 2016) i.e. imbalanced dataset and their accuracies were over 98%. This indicates an unaddressed issue in these published articles. However, our datasets were balanced datasets with higher accuracies.

Second finding from our results is that hyper parameter tuning also effect the accuracy of classifiers. It was achieved by changing parameter and with the help of different kernels. For example, classification was performed by changing the attributes of the classifiers. In our study neural network was used with three different kernels: (i) High performance neural, (ii) Multilayer perception, and (iii) Multilayer perception back propagation. High performance neural shown slightly better performance for main header dataset of 99.85%, in section headers and program header, this classifier showed decrease in performance to 93.37% and 90.79%.

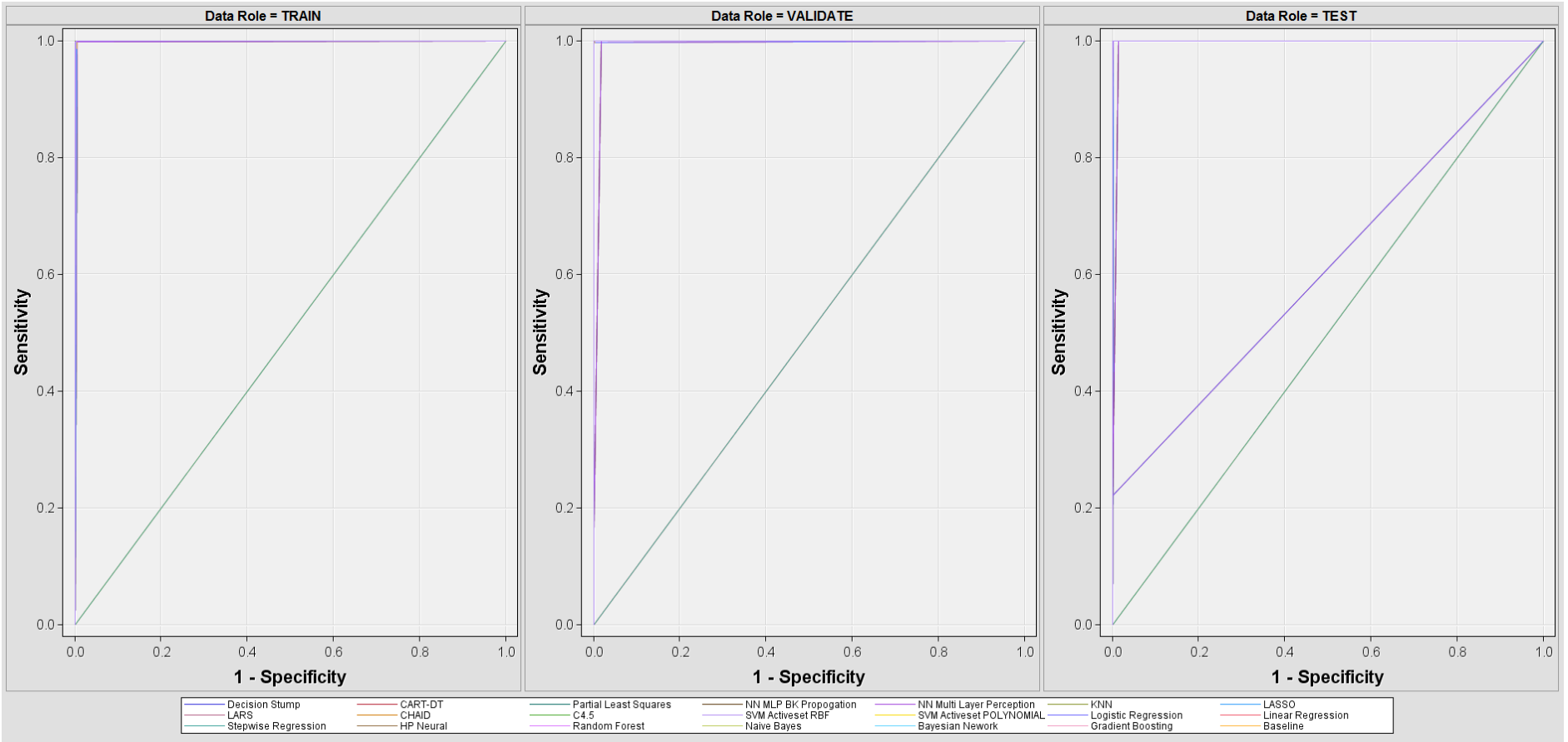


Figure 18: Main header ROC Chart

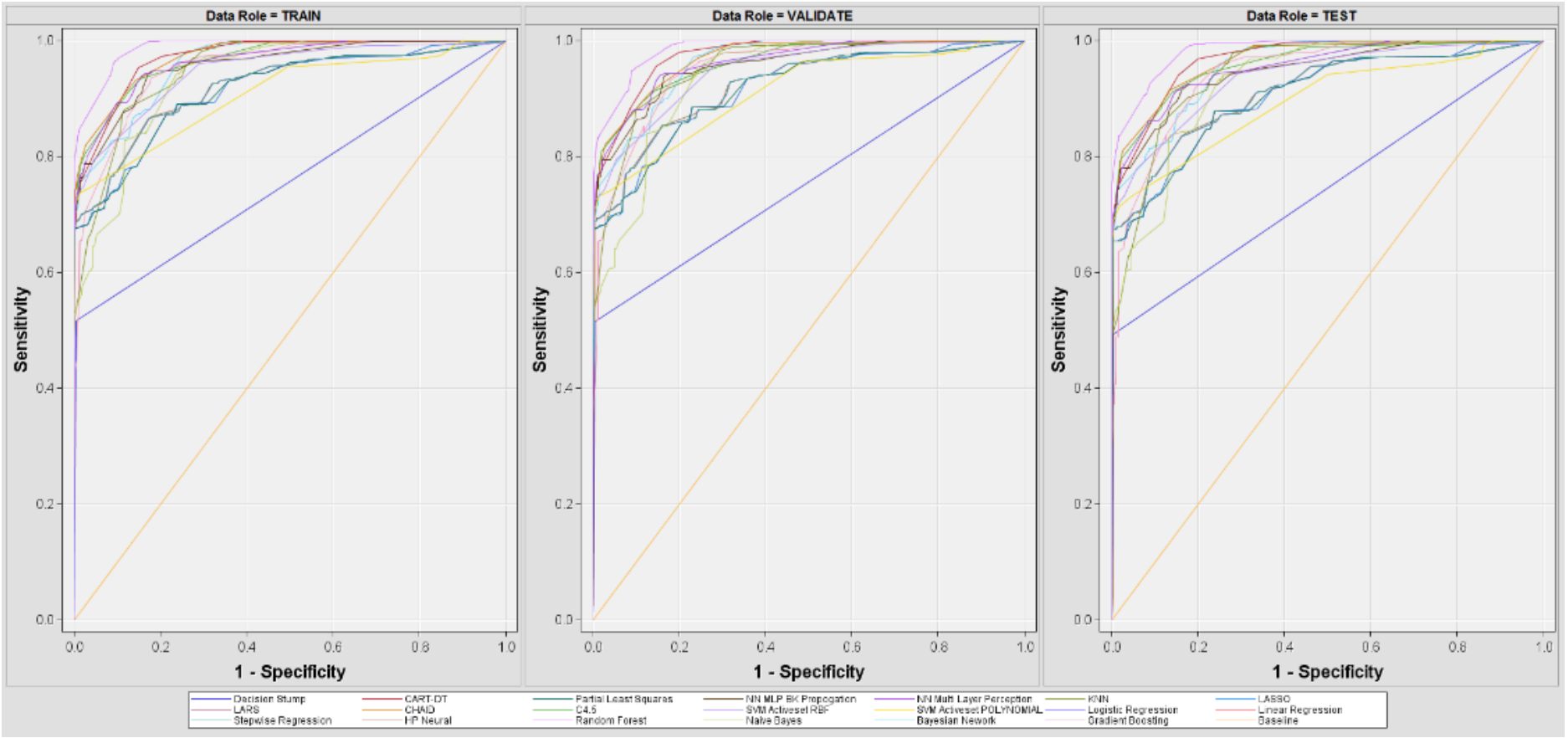


Figure 19: Program header ROC Chart

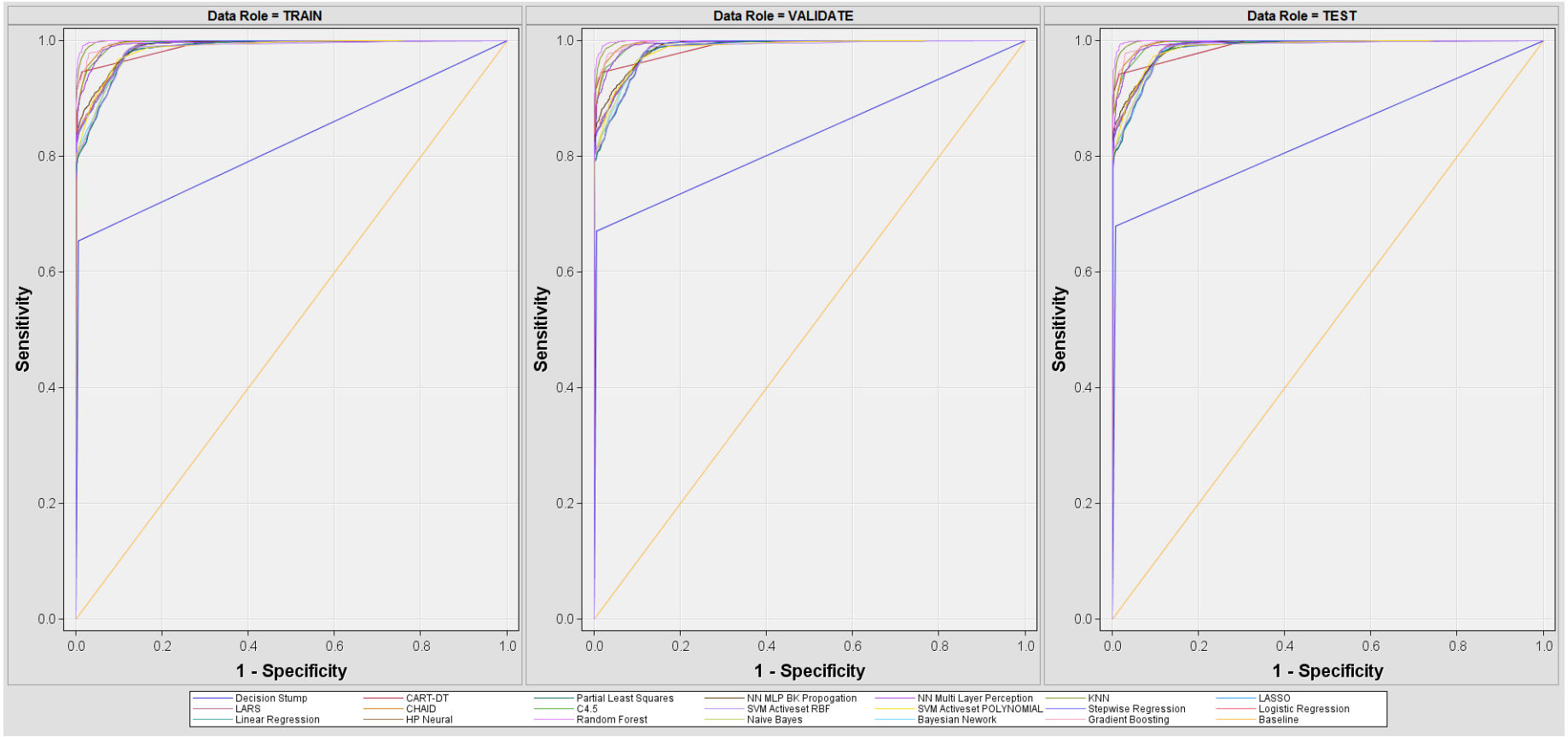


Figure 20: Section header ROC Chart

Multilayer perception and Back propagation shown better performance in program header. No published article specified this attribute to address performance related issues to help malware analysis. In main header all three of these variants of neural network were placed in Group A. In comparison to this Neural Network was used with default parameters in (Fereidooni et al., 2016, Lo et al., 2016a) but parameters or kernels were not specified. Three variants of tree algorithms were used and as predicted they performed well in all datasets.

In Group A, main header and section header datasets had Chi-Square Automatic Interaction Detection, C4.5, CART and Decision Stump (A tree with MAX Depth=1) but in program header decision stump was the least performing classifier and we placed it in Group D. In comparison to this, in reviewed articles (Wang et al., 2016b, Kuhnel et al., 2016, Baldangombo et al., 2013, Santos et al., 2011, Islam et al., 2013b, Milosevic et al., 2017, Fereidooni et al., 2016) decision tree was used along with two different tree variants i.e. J48, C4.5. C4.5.

Furthermore, in terms of future research, it was observed in this study that in training, test and validations all three datasets had similar behavior and had a similar ROC chart a straight line (in main header due to accuracies 100%) but a curve in program and section header. Our model comparison process shown that all algorithms had a closer performance except decision stump. Our ROC chart was a plot between True Positive Rate (TPR) and False Positive Rate (FPR) which we derived from Table 9 showing performance measures. Figures 19, 20, and 21 show the charts showing the positive likelihood of predicting a malware. Our classifiers had a positive prediction power of 100% using Group A classifier in main header and in over 94% in case of program header and section header. We can predict another thing from our ROC curve that closer the chart from left hand corner higher would be the accuracy which is apparent from these three figures. When this curve passes through top left corner then we can also conclude that both percentages of specificity and sensitivity are equal to 100%.

Dataset	Classifier	Accuracy	AUC	Precision	Recall	F1-score
Program header	Random Forest	94.36	0.988	0.96	0.76	0.85
Main header	Least Angle Regression	100.00	1	1.00	1.00	1.00
	Least Absolute Shrinkage and Selection Operator	100.00	1	1.00	1.00	1.00
	C4.5	100.00	0.997	1.00	1.00	1.00
	Chi-square automatic interaction detection	100.00	0.997	0.98	1.00	0.99
	Decision Stump	100.00	0.997	1.00	1.00	1.00
	Random Forest	100.00	1	1.00	1.00	1.00
	Partial Least Squares	100.00	1	1.00	1.00	1.00
	SVM Activeset RBF	100.00	1	1.00	1.00	1.00
	SVM Activeset Polynomial	100.00	1	1.00	1.00	1.00
	Gradient Boosting	100.00	0.997	1.00	1.00	1.00
	Classification and Regression Trees	99.85	1	1.00	1.00	1.00
	HP Neural	99.85	1	1.00	1.00	1.00
	KNN	99.85	1	1.00	1.00	1.00
	Bayesian Network	99.85	1	0.98	1.00	0.99
	Naive Bayes	99.85	1	1.00	1.00	1.00
	Linear Regression	99.69	1	1.00	1.00	1.00
	Logistic Regression	99.69	1	1.00	1.00	1.00
	NN Multilayer Perception	99.08	1	1.00	1.00	1.00
	NN Multilayer Perception Back Propagation	99.08	1	1.00	1.00	1.00
Section header	Random Forest	98.34	0.999	0.98	0.98	0.98
	KNN	97.84	0.999	0.96	0.98	0.97
	Gradient Boosting	96.96	0.997	0.98	0.93	0.96
	Classification and Regression Trees	96.9	0.99	0.97	0.94	0.96
	Chi-square automatic interaction detection	96.89	0.995	0.95	0.94	0.94
	C4.5	96.47	0.995	0.95	0.95	0.95
	NN Multilayer Perception	95.87	0.994	0.81	0.99	0.89
	NN MLP BK Propagation	94.05	0.988	0.85	0.97	0.90

Table 14: Shows the performance measures of different classifiers belong to Group A

Another finding from our results was that in program headers three classifiers C4.5, CHAID, and CART had best performance of 93.08%, 92.84% and 92.49% respectively. Comparing this to our results reviewed articles used just simple Decision Tree and C4.5 and in terms of performance, we seen slightly better performance, but in the articles where performance was better (97.30%, 95.6%, 96.2%, and 94%) the feature set were entirely different in nature and it was not even closer. One result was 87.81% but majority of these literature used Random Forest

as well which has been observed to give best results. Therefore, we had a chance of getting better performance in Group B as well with our tree-based classifiers.

Dataset	Classifier	Accuracy	AUC	Precision	Recall	F1-score
Program header	C4.5	93.08	0.966	0.96	0.76	0.85
	Chi-square Automatic Interaction Detection	92.84	0.972	0.93	0.81	0.86
	Classification and Regression Trees	92.49	0.971	0.94	0.81	0.87
	NN MLP BK Propagation	92.41	0.958	0.99	0.66	0.79
	NN Multi-Layer Perception	92.17	0.963	0.99	0.68	0.81
	SVM active set RBF	91.03	0.949	0.96	0.83	0.89
	SVM active set Polynomial	90.97	0.934	0.99	0.68	0.81
	HP Neural	90.79	0.921	0.98	0.51	0.67
	Stepwise Regression	90.79	0.92	0.99	0.68	0.81
	Linear Regression	90.79	0.92	0.98	0.67	0.80
	Logistic Regression	90.79	0.92	0.99	0.68	0.81
	LARS	90.54	0.914	0.99	0.69	0.81
	LASSO	90.54	0.914	0.62	0.93	0.74
	Partial Least Squares	90.20	0.913	0.98	0.67	0.80
Section header	SVM Active set POLYNOMIAL	93.61	0.987	0.94	0.89	0.91
	Linear Regression	93.46	0.988	0.93	0.91	0.92
	Logistic Regression	93.46	0.988	0.95	0.96	0.96
	HP Neural	93.37	0.988	0.92	0.90	0.91
	Stepwise Regression	93.31	0.987	0.92	0.90	0.91
	Bayesian Network	92.45	0.986	0.92	0.90	0.91
	SVM Active set RBF	92.36	0.982	0.91	0.88	0.90
	LASSO	92.26	0.984	0.90	0.89	0.89
	Least Angle Regression	92.25	0.984	0.90	0.89	0.89
	Partial Least Squares	92.17	0.984	0.91	0.88	0.89
	Naive Bayes	90.73	0.985	0.92	0.90	0.91

Table 15: Shows the performance measures of different classifiers belong to Group B

Stepwise regression was the only one classifier in main header was the worst performing classifier with 50% accuracy, which may be due to nature of the data and opens doors for further studies. We could clearly see the curve was away from upper left corner and the Point of Compromise (POC) between sensitivity and specificity was smaller. Our worst-case scenario was too far away from POC and could be seen as a straight line in ROC curve near baseline.

Dataset	Classifier	Accuracy	AUC	Precision	Recall	F1-score
Program header	Gradient Boosting	88.84	0.951	0.93	0.66	0.77
	Bayesian Network	86.16	0.96	0.70	0.89	0.78
	Decision Stump	86.00	0.756	0.94	0.77	0.85
	Naive Bayes	83.71	0.935	0.66	0.85	0.75
	KNN	81.97	0.952	0.99	0.69	0.81
Section Header	Decision Stump	87.35	0.823	0.98	0.67	0.80

Table 16: Shows the performance measures of different classifiers belong to Group C

Dataset	Classifier	Accuracy	AUC	Precision	Recall	F1-score
Main header	Stepwise regression	50.38	0.5	0.50	1.00	0.67

Table 17: Shows the performance measures of different classifiers belong to Group D

Furthermore, our feature selection methods outclassed reviewed articles as in all three of our datasets the majority of our classifiers were in Group A or B in fact more in Group A. Keeping feature scoring in views, features with least weightage were dropped from analysis to observe the classifier performance and we noticed that there was no difference in performance of classifiers. Our research was first research to combine feature selection methods to experience difference in results.

Overall finding from published literature shown that majority of the articles did not explain individual accuracies including our list of performance measures and just specified overall accuracy which is insufficient to compare with our results.

Chapter 5: Future work and achievements

In the future, our research will be expanded to develop further a robust IoT security framework based on analysis of malware attacking IoT networks. In this regard, efforts would be made to propose a preventive approach to cope with future security, privacy, and trust related threats.

We identify a few open research directions listed below to help us extend the research in this dissertation.

1. To perform further experiments on datasets created.
2. Application of machine learning on following data sets.

Feature set	features
Segment header	15
Symbol table	7
Strings	5
Function	10
Process names	5
Imports	5
Exports	5

3. Text mining of malware string, symbols, functions, imports, and exports.
4. The creation of physical IoT network and analysis of network traffic with/without malicious activities.
5. The creation of network simulation of IoT network and comparison of network traffic data analysis with physical network data analysis and propose a preventive framework.
6. To perform more in-depth research on identified security, privacy and trust challenges and malware attacks.

Achievements

The purpose of this section is to explain work done till the time of assessment.

Table 11 illustrates datasets created to support our static malware analysis process. In addition to this table 14 shows sample strings to give an idea about the string dataset. Following are the achievements at this stage.

1. Datasets including file header, program header and section header, strings and symbol table are 100% complete.
2. Paper for malware classification using file header, program header, and section header has been written. Results have been compiled as well, and I am currently addressing some questions raised by the supervisors.
3. Another paper that includes only symbol table and possibly strings as well is under classification phase, for this dataset as a test experiment TF/IDF and PCA were applied size of the dataset includes one million symbols from 2185 malware and goodware. Removed symbols with 0% frequency to avoid level limitation constraint of the analysis tools. (Missing values have been checked as well, need to address some issues with this dataset to avoid imbalanced dataset and increase generalization).
4. Applied text mining on malware/goodware strings and planning to apply sentiment analysis to observe some useful insights from strings. During feature extraction process, I experienced that there are important strings present in malware like IP Addresses, Ports, Linux directory access commands, error messages, abusive words as passwords, default passwords, etc. and in goodware nature of the strings present are entirely different. Following table 3 represents some strings from malware and goodware datasets.
5. Other datasets of functions, process names, imports, exports and segment headers were extracted using IDA Pro 6.9 Decompiler, requiring some preprocessing and data cleansing. (70% work done in those)

Table showing sample strings extracted from malware and goodware	
Malware	**Goodware**
142.54.191.34:23	@Fld
root	$%s,
admin	$Fld%d,
user	$%s)
login	$Fld%d)
guest	= split(/[%c\\n]/, $_, -1);\n
support	= split($FS, $_, -1);\n
toor	= split(' , $_, -1);\n
changeme	$FS = *+?.[]()\|^$\\ ;\t\t# field separator from -F switch\n $FS = ;\t\t# set field separator\n
1234	Could not parse message from stdin\n;
12345	Show forms dialog options; misc; Miscellaneous options; Show miscellaneous options;
123456	Zenity;version;3.4.0;copyright;Copyright Â© 2003 Sun Microsystems;
default	Error showing notification: %s;tooltip;visible;
password	Display;notification;Set the notification text;
(null)	Set dialog timeout in seconds
buf: %s\n	Set step size

/bin/sh	$, = ;\t\t# set output field separator\n
/proc/cpuinfo	$\\ = \"\\n\";\t\t# set output record separator\n
cd /tmp || cd /var/run || cd /mnt || cd /root cd /; **wget http://142.54.191.34/bins.sh;** **chmod 777 bins.sh;** **sh bins.sh; tftp 142.54.191.34 -c get tftp1.sh;** **chmod 777 tftp1.sh; sh tftp1.sh;** **tftp -r tftp2.sh -g 142.54.191.34;** **chmod 777 tftp2.sh; sh tftp2.sh;** **ftpget -v -u anonymous -p anonymous** **-P 21 142.54.191.34 ftp1.sh ftp1.sh;** **sh ftp1.sh; rm -rf bins.sh tftp1.sh tftp2.sh ftp1.sh;** **rm -rf *; exit\r\n**	program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY;

Table 18: Showing sample malware/goodware strings

Conclusion

Reaching the final step of this research, many interesting facts were presented not only concerning Internet of Things but also with about Malware Analysis and Machine Learning. Initially, during comprehensive literature review, it was identified that IoT is a favorite target of malware due to the presence of various security, privacy and trust challenges associated with IoT ecosystem. We recognized that all of the difficulties discovered contain great possibilities for future research. Furthermore, by literature review, it was found that the reason behind the statement that “IoT has been a favorite target of malware attacks” was our extracted list of malware attacks reported in the literature. Reported evidence states malware analysis a vital research area.

Talking about research objectives of the thesis, we aimed to discuss IoT with the perspective of malware. Therefore, the first objective was to perform a comprehensive literature review on IoT to build foundations for conducting the malware analysis on IoT devices. This successfully lead us to collect IoT malware and clean samples and perform malware analysis. Our research also has shown that due to complex nature of IoT devices, traditional antivirus mechanisms are not feasible on IoT and it was needed to conduct malware analysis focusing solely on IoT. Due to time constraints, we performed static malware analysis only and results shown that our machine learning process returned promising results with above 90% accuracy. Feature selection process showed that by removing unnecessary features, we could improve efficiency and reduce system overhead. These fewer features also help forensic experts to focus on available features as top priority features and investigate further.

Also, best classification results were obtained by using Random Forest. Least performance in all three datasets was given by Decision Stump (Decision tree with MAXDEPTH=1), KNN and Stepwise Regression. To mention our final suggestion for future research, further analysis is required to select best features that can be extracted in a faster manner; it would help to build a lightweight embedded program for monitoring suspicious behavior along with experimentation on sampling methods on datasets.

References

AAFER, Y., DU, W. L. & YIN, H. 2013. DroidAPIMiner: Mining API-Level Features for Robust Malware Detection in Android. *In:* ZIA, T., ZOMAYA, A., VARADHARAJAN, V. & MAO, M. (eds.) *Security and Privacy in Communication Networks, Securecomm 2013.* New York: Springer.

ADEBAYO, O. S., ABDULAZIZ, N. & IEEE 2014. Android Malware Classification Using Static Code Analysis and Apriori Algorithm Improved with Particle Swarm Optimization. *2014 4th World Congress on Information and Communication Technologies (WICT)***,** 123-128.

AKHUSEYINOGLU, N. B. & AKHUSEYINOGLU, K. 2016. AntiWare: An Automated Android Malware Detection Tool based on Machine Learning Approach and Official Market Metadata. *2016 Ieee 7th Annual Ubiquitous Computing, Electronics Mobile Communication Conference (Uemcon)***,** 7.

AL-FUQAHA, A., GUIZANI, M., MOHAMMADI, M., ALEDHARI, M. & AYYASH, M. 2015. Internet of things: A survey on enabling technologies, protocols, and applications. *IEEE Communications Surveys & Tutorials,* 17**,** 2347-2376.

ALDAIRI, A. 2017. Cyber Security Attacks on Smart Cities and Associated Mobile Technologies. *Procedia Computer Science,* 109**,** 1086-1091.

ALI, J., KHAN, R., AHMAD, N. & MAQSOOD, I. 2012. Random forests and decision trees. *IJCSI International Journal of Computer Science Issues,* 9**,** 272-278.

ALRAWAIS, A., ALHOTHAILY, A., HU, C. & CHENG, X. 2017a. Fog Computing for the Internet of Things: Security and Privacy Issues. *IEEE Internet Computing,* 21**,** 34-42.

ALRAWAIS, A., ALHOTHAILY, A., HU, C., XING, X. & CHENG, X. 2017b. An Attribute-Based Encryption Scheme to Secure Fog Communications. *IEEE Access,* 5**,** 9131-9138.

ALVAREZ, R., ARAQUE, J. & SIERRA, J. E. 2017. A Novel Smart Home Energy Management System: Architecture and Optimization Model. *Indian Journal of Science and Technology,* 10.

AMADEO, M., CAMPOLO, C., QUEVEDO, J., CORUJO, D., MOLINARO, A., IERA, A., AGUIAR, R. L. & VASILAKOS, A. V. 2016. Information-centric networking for the internet of things: challenges and opportunities. *IEEE Network,* 30**,** 92-100.

AMBROSIN, M., ANZANPOUR, A., CONTI, M., DARGAHI, T., MOOSAVI, S. R., RAHMANI, A. M. & LILJEBERG, P. 2016. On the Feasibility of Attribute-Based Encryption on Internet of Things Devices. *IEEE Micro,* 36**,** 25-35.

ANANTHARAMAN, P., LOCASTO, M., CIOCARLIE, G. F. & LINDQVIST, U. 2017. Building Hardened Internet-of-Things Clients with Language-theoretic Security.

ANGRISHI, K. 2017. Turning Internet of Things (IoT) into Internet of Vulnerabilities (IoV): IoT Botnets. *arXiv preprint arXiv:1702.03681*.

ARZT, S., RASTHOFER, S., FRITZ, C., BODDEN, E., BARTEL, A., KLEIN, J., LE TRAON, Y., OCTEAU, D. & MCDANIEL, P. 2014. FlowDroid: Precise Context, Flow, Field, Object-sensitive and Lifecycle-aware Taint Analysis for Android Apps. *Acm Sigplan Notices,* 49**,** 259-269.

ATTWOOD, A., MERABTI, M., FERGUS, P. & ABUELMAATTI, O. SCCIR: Smart cities critical infrastructure response framework. Developments in E-systems Engineering (DeSE), 2011, 2011. IEEE, 460-464.

ATWAL, R. Gartner Says Tablet Sales Continue to Be Slow in 2015. *Egham, UK*.

ATZORI, L., IERA, A. & MORABITO, G. 2010. The internet of things: A survey. *Computer networks,* 54**,** 2787-2805.

BABAR, S., MAHALLE, P., STANGO, A., PRASAD, N. & PRASAD, R. 2010. Proposed Security Model and Threat Taxonomy for the Internet of Things (IoT). *In:* MEGHANATHAN, N., BOUMERDASSI, S., CHAKI, N. & NAGAMALAI, D. (eds.) *Recent Trends in Network Security and Applications: Third International Conference, CNSA 2010, Chennai, India, July 23-25, 2010. Proceedings.* Berlin, Heidelberg: Springer Berlin Heidelberg.

BALDANGOMBO, U., JAMBALJAV, N. & HORNG, S.-J. 2013. A static malware detection system using data mining methods. *arXiv preprint arXiv:1308.2831*.

BARKI, A., BOUABDALLAH, A., GHAROUT, S. & TRAORE, J. 2016a. M2M Security: Challenges and Solutions. *Ieee Communications Surveys and Tutorials,* 18**,** 1241-1254.

BARKI, A., BOUABDALLAH, A., GHAROUT, S. & TRAORÉ, J. 2016b. M2M Security: Challenges and Solutions. *IEEE Communications Surveys & Tutorials,* 18**,** 1241-1254.

BÉLISSENT, J. 2010. Getting clever about smart cities: new opportunities require new business models. *Cambridge, Massachusetts, USA*.

BONOMI, F., MILITO, R., ZHU, J. & ADDEPALLI, S. Fog computing and its role in the internet of things. Proceedings of the first edition of the MCC workshop on Mobile cloud computing, 2012. ACM, 13-16.

BORELLO, J.-M. & MÉ, L. 2008. Code obfuscation techniques for metamorphic viruses. *Journal in Computer Virology,* 4**,** 211-220.

BORGIA, E. 2014. The Internet of Things vision: Key features, applications and open issues. *Computer Communications,* 54**,** 1-31.

BORGOHAIN, T., KUMAR, U. & SANYAL, S. 2015. Survey of security and privacy issues of Internet of Things. *arXiv preprint arXiv:1501.02211*.

BOZDOGAN, Z. & KARA, R. 2015. Layered model architecture for internet of things. *Journal of Engineering Research and Applied Science,* 4**,** 260-264.

BROWN, J., ANWAR, M., DOZIER, G. & IEEE 2016. Detection of Mobile Malware: An Artificial Immunity Approach. *2016 Ieee Symposium on Security and Privacy Workshops (Spw 2016)***,** 74-80.

BROWN, R., PHAM, B. & DE VEL, O. 2005. Design of a Digital Forensics Image Mining System. *In:* KHOSLA, R., HOWLETT, R. J. & JAIN, L. C. (eds.) *Knowledge-Based Intelligent Information and Engineering Systems: 9th International Conference, KES 2005, Melbourne, Australia, September 14-16, 2005, Proceedings, Part III*. Berlin, Heidelberg: Springer Berlin Heidelberg.

BROWNLEE, J. 2016. Spot-check classification algorithms. *Machine Learning Mastery with Python; Machine Learning Mastery Pty Ltd.: Vermont Victoria, Australia***,** 100-120.

BU, K., WENG, M., ZHENG, Y., XIAO, B. & LIU, X. 2017. You Can Clone but You Can't Hide: A Survey of Clone Prevention and Detection for RFID. *IEEE Communications Surveys & Tutorials*.

CAVIGLIONE, L., WENDZEL, S. & MAZURCZYK, W. 2017. The Future of Digital Forensics: Challenges and the Road Ahead. *IEEE Security & Privacy,* 15**,** 12-17.

CEPEDA, C., TIEN, D. L. C. & ORDÓÑEZ, P. Feature Selection and Improving Classification Performance for Malware Detection. 2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom) (BDCloud-SocialCom-SustainCom), 8-10 Oct. 2016 2016. 560-566.

CHALLA, S., WAZID, M., DAS, A. K., KUMAR, N., REDDY, A. G., YOON, E. J. & YOO, K. Y. 2017. Secure Signature-Based Authenticated Key Establishment Scheme for Future IoT Applications. *IEEE Access,* 5**,** 3028-3043.

CHANDRAMOULI, R., IORGA, M. & CHOKHANI, S. 2014. Cryptographic key management issues and challenges in cloud services. *Secure Cloud Computing.* Springer.

CHEN, F., DENG, P., WAN, J., ZHANG, D., VASILAKOS, A. V. & RONG, X. 2015. Data Mining for the Internet of Things: Literature Review and Challenges. *International Journal of Distributed Sensor Networks,* 11**,** 431047.

CHEN, L., THOMBRE, S., JARVINEN, K., LOHAN, E. S., ALEN-SAVIKKO, A. K., LEPPAKOSKI, H., BHUIYAN, M. Z. H., BU-PASHA, S., FERRARA, G. N. & HONKALA, S. 2017. Robustness, Security and Privacy in Location-Based Services for Future IoT: A Survey. *IEEE Access*.

CHEN, X., MAKKI, K., YEN, K. & PISSINOU, N. 2009. Sensor network security: a survey. *IEEE Communications Surveys & Tutorials,* 11**,** 52-73.

CHEN, Z., ROUSSOPOULOS, M., LIANG, Z., ZHANG, Y., CHEN, Z. & DELIS, A. 2012. Malware characteristics and threats on the internet ecosystem. *Journal of Systems and Software,* 85**,** 1650-1672.

CHIANG, M. & ZHANG, T. 2016. Fog and IoT: An Overview of Research Opportunities. *IEEE Internet of Things Journal,* 3**,** 854-864.

CHO, T., KIM, H. & YI, J. H. 2017. Security Assessment of Code Obfuscation Based on Dynamic Monitoring in Android Things. *Ieee Access,* 5**,** 6361-6371.

CICCOZZI, F., CRNKOVIC, I., RUSCIO, D. D., MALAVOLTA, I., PELLICCIONE, P. & SPALAZZESE, R. 2017. Model-Driven Engineering for Mission-Critical IoT Systems. *IEEE Software,* 34**,** 46-53.

COHEN, A., NISSIM, N., ROKACH, L. & ELOVICI, Y. 2016. SFEM: Structural feature extraction methodology for the detection of malicious office documents using machine learning methods. *Expert Systems with Applications,* 63**,** 324-343.

CONLAN, K., BAGGILI, I. & BREITINGER, F. 2016. Anti-forensics: Furthering digital forensic science through a new extended, granular taxonomy. *Digital Investigation,* 18**,** S66-S75.

CONTI, M., DEHGHANTANHA, A., FRANKE, K. & WATSON, S. 2018. Internet of Things security and forensics: Challenges and opportunities. *Future Generation Computer Systems,* 78**,** 544-546.

CONTINELLA, A., CARMINATI, M., POLINO, M., LANZI, A., ZANERO, S. & MAGGI, F. 2017. Prometheus: Analyzing WebInject-based information stealers. *Journal of Computer Security,* 25**,** 117-137.

CSA 2015. Security Guidance for Early Adopters of the Internet of Things (IoT).

D'ORAZIO, C. J., CHOO, K.-K. R. & YANG, L. T. 2017. Data exfiltration from Internet of Things devices: iOS devices as case studies. *IEEE Internet of Things Journal,* 4**,** 524-535.

DAMODARAN, A., DI TROIA, F., VISAGGIO, C. A., AUSTIN, T. H. & STAMP, M. 2017. A comparison of static, dynamic, and hybrid analysis for malware detection. *Journal in Computer Virology and Hacking Techniques,* 13**,** 1-12.

DHAYA, R., POONGODI, M. & IEEE 2014. Detecting Software Vulnerabilities in Android Using Static Analysis. *2014 International Conference on Advanced Communication Control and Computing Technologies (Icaccct)***,** 915-918.

DIJKMAN, R. M., SPRENKELS, B., PEETERS, T. & JANSSEN, A. 2015. Business models for the Internet of Things. *International Journal of Information Management,* 35**,** 672-678.

DING, Y., CHEN, S. & XU, J. Application of Deep Belief Networks for opcode based malware detection. 2016 International Joint Conference on Neural Networks (IJCNN), 24-29 July 2016 2016. 3901-3908.

DONG, W. & LIU, X. 2015. Robust and Secure Time-Synchronization Against Sybil Attacks for Sensor Networks. *IEEE Transactions on Industrial Informatics,* 11**,** 1482-1491.
DU, X. & CHEN, H.-H. 2008. Security in wireless sensor networks. *IEEE Wireless Communications,* 15.
DULAUNOY, A., WAGENER, G., MOKADDEM, S. & WAGNER, C. 2017. An extended analysis of an IoT malware from a blackhole network.
EFRON, B., HASTIE, T., JOHNSTONE, I. & TIBSHIRANI, R. 2004. LEAST ANGLE REGRESSION. *The Annals of Statistics,* 32**,** 407–499.
EL MOUAATAMID, O., LAHMER, M. & BELKASMI, M. 2016. Internet of Things Security: Layered classification of attacks and possible Countermeasures. *Electronic Journal of Information Technology*.
ENISA 2017. Baseline Security Recommendations for IoT in the context of Critical Information Infrastructures.
FANTACCI, R., PECORELLA, T., VITI, R. & CARLINI, C. 2014. A network architecture solution for efficient IOT WSN backhauling: challenges and opportunities. *IEEE Wireless Communications,* 21**,** 113-119.
FEIZOLLAH, A., ANUAR, N. B., SALLEH, R., SUAREZ-TANGIL, G. & FURNELL, S. 2017. AndroDialysis: Analysis of Android Intent Effectiveness in Malware Detection. *Computers & Security,* 65**,** 121-134.
FENG, P. B., MA, J. F. & SUN, C. 2017. Selecting Critical Data Flows in Android Applications for Abnormal Behavior Detection. *Mobile Information Systems***,** 16.
FEREIDOONI, H., CONTI, M., YAO, D. F. & SPERDUTI, A. 2016. ANASTASIA: ANdroid mAlware detection using STAtic analySIs of Applications. *2016 8th Ifip International Conference on New Technologies, Mobility and Security (Ntms)***,** 5.
FRAGA-LAMAS, P., FERNÁNDEZ-CARAMÉS, T. M., SUÁREZ-ALBELA, M., CASTEDO, L. & GONZÁLEZ-LÓPEZ, M. 2016. A review on internet of things for defense and public safety. *Sensors,* 16**,** 1644.
FURFARO, A., ARGENTO, L., PARISE, A. & PICCOLO, A. 2017. Using virtual environments for the assessment of cybersecurity issues in IoT scenarios. *Simulation Modelling Practice and Theory,* 73**,** 43-54.
GAN, G., LU, Z. & JIANG, J. Internet of things security analysis. Internet Technology and Applications (iTAP), 2011 International Conference on, 2011. IEEE, 1-4.
GAONA-GARCIA, P., MONTENEGRO-MARIN, C., PRIETO, J. D. & NIETO, Y. V. 2017. Analysis of Security Mechanisms Based on Clusters IoT Environments. *International Journal of Interactive Multimedia and Artificial Intelligence,* 4**,** 55-60.
GARLAN, D., CHENG, S.-W. & SCHMERL, B. 2003. Increasing system dependability through architecture-based self-repair. *Architecting dependable systems.* Springer.
GENEIATAKIS, D., SATTA, R., FOVINO, I. N. & NEISSE, R. 2015. On the Efficacy of Static Features to Detect Malicious Applications in Android. *In:* FISCHERHUBNER, S., LAMBRINOUDAKIS, C. & LOPEZ, J. (eds.) *Trust, Privacy and Security in Digital Business.* Berlin: Springer-Verlag Berlin.
GHILDIYAL, S., MISHRA, A. K., GUPTA, A. & GARG, N. 2014. Analysis of Denial of Service (DOS) Attacks in wireless sensor networks. *IJRET: International Journal of Research in Engineering and Technology,* 3**,** 2319-1163.
GIULIANO, R., MAZZENGA, F., NERI, A. & VEGNI, A. M. 2017. Security Access Protocols in IoT Capillary Networks. *Ieee Internet of Things Journal,* 4**,** 645-657.
GOPE, P. & HWANG, T. 2015. Untraceable Sensor Movement in Distributed IoT Infrastructure. *IEEE Sensors Journal,* 15**,** 5340-5348.

GRANJAL, J., MONTEIRO, E. & SILVA, J. S. 2015. Security for the Internet of Things: A Survey of Existing Protocols and Open Research Issues. *IEEE Communications Surveys & Tutorials,* 17**,** 1294-1312.

GREENOUGH, J. & CAMHI, J. 2015. The Internet of Things 2015: Examining How the OT Will Affect the World. *BI Intelligence Report. Available online: https://zh. scribd. com/document/288595065/the-internet-ofthings-2015-examining-how-the-iot-will-affect-the-world-pdf (accessed on 23 November 2016).*

GUBBI, J., BUYYA, R., MARUSIC, S. & PALANISWAMI, M. 2013. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems,* 29**,** 1645-1660.

HARBAWI, M. & VAROL, A. An improved digital evidence acquisition model for the Internet of Things forensic I: A theoretical framework. Digital Forensic and Security (ISDFS), 2017 5th International Symposium on, 2017. IEEE, 1-6.

HARRIS, M. 2015. Documents confirm Apple is building self-driving car. *The Guardian,* 14.

HERNÁNDEZ-MUÑOZ, J. M., VERCHER, J. B., MUÑOZ, L., GALACHE, J. A., PRESSER, M., GÓMEZ, L. A. H. & PETTERSSON, J. Smart cities at the forefront of the future internet. The Future Internet Assembly, 2011. Springer, 447-462.

HOEPMAN, J.-H. In things we trust? Towards trustability in the internet of things. AmI Workshops, 2011. Springer, 287-295.

HU, Z. The research of several key question of internet of things. Intelligence Science and Information Engineering (ISIE), 2011 International Conference on, 2011. IEEE, 362-365.

IHS 2016. IoT platforms: enabling the Internet of Things.

ISLAM, R., TIAN, R., BATTEN, L. M. & VERSTEEG, S. 2013a. Classification of malware based on integrated static and dynamic features. *Journal of Network and Computer Applications,* 36**,** 646-656.

ISLAM, R., TIAN, R. H., BATTEN, L. M. & VERSTEEG, S. 2013b. Classification of malware based on integrated static and dynamic features. *Journal of Network and Computer Applications,* 36**,** 646-656.

ISLAM, S. M. R., KWAK, D., KABIR, M. H., HOSSAIN, M. & KWAK, K. S. 2015. The Internet of Things for Health Care: A Comprehensive Survey. *IEEE Access,* 3**,** 678-708.

JAMMES, F. & SMIT, H. 2005. Service-oriented paradigms in industrial automation. *IEEE Transactions on Industrial Informatics,* 1**,** 62-70.

JAVED, F., AFZAL, M. K., SHARIF, M. & KIM, B.-S. 2018. Internet of Things (IoTs) Operating Systems Support, Networking Technologies, Applications, and Challenges: A Comparative Review. *IEEE Communications Surveys & Tutorials*.

JEFF KELLY, D. F. 2013. *The Industrial Internet and Big Data Analytics: Opportunities and Challenges* [Online]. Available: http://wikibon.org/wiki/v/The_Industrial_Internet_and_Big_Data_Analytics:_Opportunities_and_Challenges [Accessed].

JIANG, L., LIU, D.-Y. & YANG, B. Smart home research. Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on, 2004. IEEE, 659-663.

JING, Q., VASILAKOS, A. V., WAN, J. F., LU, J. W. & QIU, D. C. 2014. Security of the Internet of Things: perspectives and challenges. *Wireless Networks,* 20**,** 2481-2501.

JUMA, H., KAMEL, I. & KAYA, L. On protecting the integrity of sensor data. 2008 15th IEEE International Conference on Electronics, Circuits and Systems, Aug. 31 2008-Sept. 3 2008 2008. 902-905.

KANG, H., JANG, J. W., MOHAISEN, A. & KIM, H. K. 2015. Detecting and Classifying Android Malware Using Static Analysis along with Creator Information. *International Journal of Distributed Sensor Networks***,** 9.

KARANJA, E., MWANGI , MASUPE, S. & MANDU, J. 2017. INTERNET OF THINGS MALWARE: A SURVEY. *International Journal of Computer Science & Engineering Survey (IJCSES),* 8

DOI:.

KÄRKKÄINEN, M. 2003. Increasing efficiency in the supply chain for short shelf life goods using RFID tagging. *International Journal of Retail & Distribution Management,* 31**,** 529-536.

KARLOF, C. & WAGNER, D. 2003. Secure routing in wireless sensor networks: Attacks and countermeasures. *Ad hoc networks,* 1**,** 293-315.

KEVIN, A. 2009. That 'Internet of Things' thing, in the real world things matter more than ideas. *RFiD Journal,* 22.

KHARCHENKO, V., KOLISNYK, M., PISKACHOVA, I. & BARDIS, N. Reliability and Security Issues for IoT-based Smart Business Center: Architecture and Markov Model. 2016 Third International Conference on Mathematics and Computers in Sciences and in Industry (MCSI), 27-29 Aug. 2016 2016. 313-318.

KHODADADI, F., DASTJERDI, A. V. & BUYYA, R. 2017. Internet of Things: An Overview. *arXiv preprint arXiv:1703.06409*.

KIM, J. Y. 2017. Efficiency of Paid Authentication Methods for Mobile Devices. *Wireless Personal Communications,* 93**,** 543-551.

KIM, Y., OH, T. & KIM, J. 2015. Analyzing User Awareness of Privacy Data Leak in Mobile Applications. *Mobile Information Systems***,** 12.

KIRK, R. 2015. Cars of the future: the Internet of Things in the automotive industry. *Network Security,* 2015**,** 16-18.

KLIARSKY, A. 2017. Detecting Attacks Against The 'Internet of Things'.

KOCHETKOVA, K. 2016. *How to not break the Internet* [Online]. Available: https://www.kaspersky.com/blog/attack-on-dyn-explained/13325/ [Accessed].

KOLTER, J. Z. & MALOOF, M. A. Learning to detect malicious executables in the wild. Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, 2004. ACM, 470-478.

KOSHIZUKA, N. & SAKAMURA, K. 2010. Ubiquitous ID: standards for ubiquitous computing and the Internet of Things. *IEEE Pervasive Computing,* 9**,** 98-101.

KÜHNEL, M. & MEYER, U. Classification of Short Messages Initiated by Mobile Malware. 2016 11th International Conference on Availability, Reliability and Security (ARES), Aug. 31 2016-Sept. 2 2016 2016. 618-627.

KUHNEL, M., MEYER, U. & IEEE 2016. Classification of Short Messages Initiated by Mobile Malware. *Proceedings of 2016 11th International Conference on Availability, Reliability and Security, (Ares 2016)***,** 618-627.

KUMARAGE, H., KHALIL, I., ALABDULATIF, A., TARI, Z. & YI, X. 2016. Secure Data Analytics for Cloud-Integrated Internet of Things Applications. *IEEE Cloud Computing,* 3**,** 46-56.

KUUSIJÄRVI, J., SAVOLA, R., SAVOLAINEN, P. & EVESTI, A. Mitigating IoT security threats with a trusted Network element. Internet Technology and Secured Transactions (ICITST), 2016 11th International Conference for, 2016. IEEE, 260-265.

KWON, D., HODKIEWICZ, M. R., FAN, J., SHIBUTANI, T. & PECHT, M. G. 2016. IoT-Based Prognostics and Systems Health Management for Industrial Applications. *IEEE Access,* 4**,** 3659-3670.

LAM, K.-Y. & CHI, C.-H. 2016. Identity in the Internet-of-Things (IoT): New challenges and opportunities. *Information and Communications Security*. Springer.
LEWIS, M. 2007. Stepwise versus Hierarchical Regression: Pros and Cons. *Online Submission*.
LI, D., AUNG, Z., WILLIAMS, J. & SANCHEZ, A. 2014a. P3: Privacy Preservation Protocol for Automatic Appliance Control Application in Smart Grid. *IEEE Internet of Things Journal,* 1**,** 414-429.
LI, L., LI, S. & ZHAO, S. 2014b. QoS-Aware Scheduling of Services-Oriented Internet of Things. *IEEE Transactions on Industrial Informatics,* 10**,** 1497-1505.
LI, S., XU, L. D. & ZHAO, S. 2015. The internet of things: a survey. *Information Systems Frontiers,* 17**,** 243-259.
LI, X., LU, R., LIANG, X., SHEN, X., CHEN, J. & LIN, X. 2011. Smart community: an internet of things application. *IEEE Communications Magazine,* 49.
LIN, H. C. & BERGMANN, N. 2016. IoT Privacy and Security Challenges for Smart Home Environments. *Information,* 7**,** 15.
LIU, C., YANG, C., ZHANG, X. & CHEN, J. 2015. External integrity verification for outsourced big data in cloud and IoT: A big picture. *Future Generation Computer Systems,* 49**,** 58-67.
LIU, J. IoT forensics issues, strategies and challenges. 12th IDF Conference, 2015 Tokyo, Japan.
LIU, J. & SUN, W. 2016. Smart Attacks against Intelligent Wearables in People-Centric Internet of Things. *IEEE Communications Magazine,* 54**,** 44-49.
LIU, W., LIU, H., WAN, Y. L., KONG, H. F. & NING, H. S. 2016. The yoking-proof-based authentication protocol for cloud-assisted wearable devices. *Personal and Ubiquitous Computing,* 20**,** 469-479.
LO, C. T. D., PABLO, O. & CARLOS, C. 2016a. Feature Selection and Improving Classification Performance for Malware Detection. *Proceedings of 2016 Ieee International Conferences on Big Data and Cloud Computing (Bdcloud 2016) Social Computing and Networking (Socialcom 2016) Sustainable Computing and Communications (Sustaincom 2016) (Bdcloud-Socialcom-Sustaincom 2016)***,** 560-566.
LO, C. T. D., PABLO, O. & CARLOS, C. M. Towards an effective and efficient malware detection system. 2016 IEEE International Conference on Big Data (Big Data), 5-8 Dec. 2016 2016b. 3648-3655.
LOHR, S. 2012. The age of big data. *New York Times,* 11.
LOPEZ, C. C. U. & CADAVID, A. N. 2016. Machine Learning Classifiers for Android Malware Analysis. *2016 Ieee Colombian Conference on Communications and Computing (Colcom)***,** 6.
LUONG, N. C., HOANG, D. T., WANG, P., NIYATO, D., KIM, D. I. & HAN, Z. 2016. Data Collection and Wireless Communication in Internet of Things (IoT) Using Economic Analysis and Pricing Models: A Survey. *Ieee Communications Surveys and Tutorials,* 18**,** 2546-2590.
MAHALLE, P., BABAR, S., PRASAD, N. R. & PRASAD, R. 2010. Identity management framework towards internet of things (IoT): Roadmap and key challenges. *Recent Trends in Network Security and Applications***,** 430-439.
MAPLE, C. 2017. Security and privacy in the internet of things. *Journal of Cyber Policy,* 2**,** 155-184.
MAVROMOUSTAKIS, C., MASTORAKIS, G. & BATALLA, J. M. 2016. *Internet of Things (IoT) in 5G mobile technologies*, Springer.
MAYER, C. P. 2009. Security and privacy challenges in the internet of things. *Electronic Communications of the EASST,* 17.

MAZUMDER, R., MIYAJI, A. & SU, C. H. 2017. A simple authentication encryption scheme. *Concurrency and Computation-Practice & Experience,* 29**,** 10.
MCFARLANE, D. & SHEFFI, Y. 2003. The impact of automatic identification on supply chain operations. *The international journal of logistics management,* 14**,** 1-17.
MEGHANATHAN, N. 2010. Design of a reliability-based source routing protocol for wireless mobile Ad Hoc networks. *Recent Trends in Network Security and Applications***,** 463-472.
MENDEZ, D. M., PAPAPANAGIOTOU, I. & YANG, B. 2017. Internet of things: Survey on security and privacy. *arXiv preprint arXiv:1707.01879*.
MILOSEVIC, N., DEHGHANTANHA, A. & CHOO, K.-K. R. Machine learning aided Android malware classification. *Computers & Electrical Engineering*.
MILOSEVIC, N., DEHGHANTANHA, A. & CHOO, K.-K. R. 2017. Machine learning aided Android malware classification. *Computers & Electrical Engineering*.
MIN, B. & VARADHARAJAN, V. 2015. Design and Analysis of a Sophisticated Malware Attack Against Smart Grid. *In:* DESMEDT, Y. (ed.) *Information Security.* Cham: Springer Int Publishing Ag.
MIN, B. H., VARADHARAJAN, V. & IEEE 2014. Design and Analysis of Security Attacks against Critical Smart Grid Infrastructures. *2014 19th International Conference on Engineering of Complex Computer Systems (Iceccs 2014)***,** 59-68.
MINERAUD, J., MAZHELIS, O., SU, X. & TARKOMA, S. 2016. A gap analysis of Internet-of-Things platforms. *Computer Communications,* 89**,** 5-16.
MINOLI, D., SOHRABY, K. & OCCHIOGROSSO, B. 2017. IoT Considerations, Requirements, and Architectures for Smart Buildings—Energy Optimization and Next-Generation Building Management Systems. *IEEE Internet of Things Journal,* 4**,** 269-283.
MIORANDI, D., SICARI, S., DE PELLEGRINI, F. & CHLAMTAC, I. 2012. Internet of things: Vision, applications and research challenges. *Ad Hoc Networks,* 10**,** 1497-1516.
MUKHERJEE, A. 2015. Physical-Layer Security in the Internet of Things: Sensing and Communication Confidentiality Under Resource Constraints. *Proceedings of the IEEE,* 103**,** 1747-1761.
NA, A. & ISAAC, W. Developing a human-centric agricultural model in the IoT environment. Internet of Things and Applications (IOTA), International Conference on, 2016. IEEE, 292-297.
NATH, H. V. & MEHTRE, B. M. Static Malware Analysis Using Machine Learning Methods. SNDS, 2014a. Springer, 440-450.
NATH, H. V. & MEHTRE, B. M. Static malware analysis using machine learning methods. International Conference on Security in Computer Networks and Distributed Systems, 2014b. Springer, 440-450.
NING, H., LIU, H. & YANG, L. T. 2013. Cyberentity Security in the Internet of Things. *Computer,* 46**,** 46-53.
NING, H., LIU, H. & YANG, L. T. 2015. Aggregated-Proof Based Hierarchical Authentication Scheme for the Internet of Things. *Ieee Transactions on Parallel and Distributed Systems,* 26**,** 657-667.
NORDRUM, A. Aug 2016. *Popular Internet of Things Forecast of 50 Billion Devices by 2020 Is Outdated* [Online]. Available: http://spectrum.ieee.org/tech-talk/telecom/internet/popular-internet-of-things-forecast-of-50-billion-devices-by-2020-is-outdated [Accessed].

NUKALA, R., PANDURU, K., SHIELDS, A., RIORDAN, D., DOODY, P. & WALSH, J. Internet of Things: A review from 'Farm to Fork'. Signals and Systems Conference (ISSC), 2016 27th Irish, 2016. IEEE, 1-6.

ORIWOH, E. & SANT, P. The forensics edge management system: A concept and design. Ubiquitous Intelligence and Computing, 2013 IEEE 10th International Conference on and 10th International Conference on Autonomic and Trusted Computing (UIC/ATC), 2013. IEEE, 544-550.

OUADDAH, A., MOUSANNIF, H., ABOU ELKALAM, A. & AIT OUAHMAN, A. 2017. Access control in the Internet of Things: Big challenges and new opportunities. *Computer Networks*, 112**,** 237-262.

PAJOUH, H. H., JAVIDAN, R., KHAYAMI, R., ALI, D. & CHOO, K. K. R. 2016. A Two-layer Dimension Reduction and Two-tier Classification Model for Anomaly-Based Intrusion Detection in IoT Backbone Networks. *IEEE Transactions on Emerging Topics in Computing*, PP**,** 1-1.

PAN, J., PAUL, S. & JAIN, R. 2011. A survey of the research on future internet architectures. *IEEE Communications Magazine*, 49.

PANDEY, V., KAUR, A. & CHAND, N. 2010. A review on data aggregation techniques in wireless sensor network. *Journal of Electronic and Electrical Engineering*, 1**,** 01-08.

PATEL, A., S. PATEL, R., SINGH, N. M. & KAZI, F. 2017. *Vitality of Robotics in Healthcare Industry: An Internet of Things (IoT) Perspective*.

PATIL, K. & KALE, N. A model for smart agriculture using IoT. Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), 2016 International Conference on, 2016. IEEE, 543-545.

PIRBHULAL, S., ZHANG, H. Y., ALAHI, M. E. E., GHAYVAT, H., MUKHOPADHYAY, S. C., ZHANG, Y. T. & WU, W. Q. 2017. A Novel Secure IoT-Based Smart Home Automation System Using a Wireless Sensor Network. *Sensors*, 17**,** 19.

PONGLE, P. & CHAVAN, G. A survey: Attacks on RPL and 6LoWPAN in IoT. 2015 International Conference on Pervasive Computing (ICPC), 8-10 Jan. 2015 2015. 1-6.

QIU, Y. & MA, M. 2016. A Mutual Authentication and Key Establishment Scheme for M2M Communication in 6LoWPAN Networks. *Ieee Transactions on Industrial Informatics*, 12**,** 2074-2085.

QUICK, D. & CHOO, K.-K. R. 2014. Impacts of increasing volume of digital forensic data: A survey and future research challenges. *Digital Investigation*, 11**,** 273-294.

RAVULA, R. R., LISZKA, K. J. & CHAN, C. C. 2011. *DYNAMIC ANALYSIS OF MALWARE USING DECISION TREES*, Setubal, Insticc-Inst Syst Technologies Information Control & Communication.

RAY, P. P. 2016. A survey on Internet of Things architectures. *Journal of King Saud University - Computer and Information Sciences*.

RAYES, A. & SALAM, S. 2017. Internet of Things (IoT) Overview. *Internet of Things From Hype to Reality.* Springer.

RIAHI, A., CHALLAL, Y., NATALIZIO, E., CHTOUROU, Z. & BOUABDALLAH, A. A systemic approach for IoT security. Distributed Computing in Sensor Systems (DCOSS), 2013 IEEE International Conference on, 2013. IEEE, 351-355.

RIAHI SFAR, A., NATALIZIO, E., CHALLAL, Y. & CHTOUROU, Z. 2017. A roadmap for security challenges in the Internet of Things. *Digital Communications and Networks*.

ROMAN, R., NAJERA, P. & LOPEZ, J. 2011. Securing the internet of things. *Computer*, 44**,** 51-58.

ROMAN, R., ZHOU, J. & LOPEZ, J. 2013. On the features and challenges of security and privacy in distributed internet of things. *Computer Networks*, 57**,** 2266-2279.

ROSE, K., ELDRIDGE, S. & CHAPIN, L. 2015. The internet of things: An overview. *The Internet Society (ISOC)***,** 1-50.
RYGIELSKI, C., WANG, J.-C. & YEN, D. C. 2002. Data mining techniques for customer relationship management. *Technology in society,* 24**,** 483-502.
SAJID, A., ABBAS, H. & SALEEM, K. 2016. Cloud-Assisted IoT-Based SCADA Systems Security: A Review of the State of the Art and Future Challenges. *IEEE Access,* 4**,** 1375-1384.
SANTOS, I., BREZO, F., SANZ, B., LAORDEN, C. & BRINGAS, P. G. 2011. Using opcode sequences in single-class learning to detect unknown malware. *Iet Information Security,* 5**,** 220-227.
SANTOS, I., DEVESA, J., BREZO, F., NIEVES, J. & BRINGAS, P. G. Opem: A static-dynamic approach for machine-learning-based malware detection. International Joint Conference CISIS'12-ICEUTE´ 12-SOCO´ 12 Special Sessions, 2013a. Springer, 271-280.
SANTOS, I., DEVESA, J., BREZO, F., NIEVES, J. & BRINGAS, P. G. 2013b. OPEM: A Static-Dynamic Approach for Machine-Learning-Based Malware Detection. *In:* HERRERO, A., SNASEL, V., ABRAHAM, A., ZELINKA, I., BARUQUE, B., QUINTIAN, H., CALVO, J. L., SEDANO, J. & CORCHADO, E. (eds.) *International Joint Conference Cisis'12 - Iceute'12 - Soco'12 Special Sessions.* Berlin: Springer-Verlag Berlin.
SANZ, B., SANTOS, I., NIEVES, J., LAORDEN, C., ALONSO-GONZALEZ, I. & BRINGAS, P. G. Mads: malicious android applications detection through string analysis. International Conference on Network and System Security, 2013. Springer, 178-191.
SAS. *5 IoT applications retailers are using today* [Online]. Available: https://www.sas.com/en_gb/insights/articles/big-data/five-iot-applications-retailers-are-using-today.html [Accessed].
SAWAND, A., DJAHEL, S., ZHANG, Z. & NAÏT-ABDESSELAM, F. 2015. Toward energy-efficient and trustworthy eHealth monitoring system. *China Communications,* 12**,** 46-65.
SCHAFFERS, H., KOMNINOS, N., PALLOT, M., TROUSSE, B., NILSSON, M. & OLIVEIRA, A. 2011. Smart cities and the future internet: Towards cooperation frameworks for open innovation. *The future internet***,** 431-446.
SCHWARM, S. E. & OSTENDORF, M. Reading level assessment using support vector machines and statistical language models. Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics, 2005. Association for Computational Linguistics, 523-530.
SEN, J. 2010. A survey on wireless sensor network security. *arXiv preprint arXiv:1011.1529*.
SETHI, P. & SARANGI, S. R. 2017. Internet of Things: Architectures, Protocols, and Applications. *Journal of Electrical and Computer Engineering,* 2017.
SHELDON, A. 2005. The future of forensic computing. *Digital Investigation,* 2**,** 31-35.
SHENOY, S. 2016. *Addressing IoT for Smart Buildings* [Online]. Intel. Available: http://blogs.intel.com/iot/2016/10/06/addressing-iot-for-smart-buildings/ [Accessed].
SICARI, S., RIZZARDI, A., GRIECO, L. A. & COEN-PORISINI, A. 2015. Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks,* 76**,** 146-164.
SINGH, D., TRIPATHI, G. & JARA, A. J. A survey of Internet-of-Things: Future vision, architecture, challenges and services. Internet of things (WF-IoT), 2014 IEEE world forum on, 2014. IEEE, 287-292.

SRNDIC, N. & LASKOV, P. 2016. Hidost: a static machine-learning-based detector of malicious files. *Eurasip Journal on Information Security*, 20.
SU, M. Y., FUNG, K. T., HUANG, Y. H., KANG, M. Z. & CHUNG, Y. H. 2016a. Detection of Android Malware: Combined with Static Analysis and. Dynamic Analysis. *2016 International Conference on High Performance Computing & Simulation (Hpcs 2016)*, 1013-1018.
SU, M. Y., FUNG, K. T. & IEEE 2016b. Detection of Android Malware by Static Analysis on Permissions and Sensitive Functions. *2016 Eighth International Conference on Ubiquitous and Future Networks.* New York: Ieee.
SUAREZ-TANGIL, G., TAPIADOR, J. E., PERIS-LOPEZ, P. & RIBAGORDA, A. 2014. Evolution, detection and analysis of malware for smart devices. *IEEE Communications Surveys & Tutorials,* 16**,** 961-987.
SUN, W., LIU, J. & ZHANG, H. 2017a. When Smart Wearables Meet Intelligent Vehicles: Challenges and Future Directions. *IEEE Wireless Communications,* 24**,** 58-65.
SUN, Y. C., WU, L., WU, S. Z., LI, S. P., ZHANG, T., ZHANG, L., XU, J. F., XIONG, Y. P. & CUI, X. G. 2017b. Attacks and countermeasures in the internet of vehicles. *Annals of Telecommunications,* 72**,** 283-295.
SUO, H., WAN, J., ZOU, C. & LIU, J. Security in the internet of things: a review. Computer Science and Electronics Engineering (ICCSEE), 2012 international conference on, 2012. IEEE, 648-651.
SYMANTEC 2017. Symantec Internet Security Threat Report.
TAYLOR, M., HAGGERTY, J., GRESTY, D. & HEGARTY, R. 2010. Digital evidence in cloud computing systems. *Computer Law & Security Review,* 26**,** 304-308.
THOMAS, K., GRIER, C., MA, J., PAXSON, V., SONG, D. & SOCIETY, I. C. 2011. Design and Evaluation of a Real-Time URL Spam Filtering Service. *2011 Ieee Symposium on Security and Privacy.* Los Alamitos: Ieee Computer Soc.
TIBSHIRANI, R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267-288.
TIBURSKI, R. T., AMARAL, L. A., MATOS, E. D. & HESSEL, F. 2015. The importance of a standard securit y archit ecture for SOA-based iot middleware. *IEEE Communications Magazine,* 53**,** 20-26.
TOBIAS, R. D. An introduction to partial least squares regression. Proceedings of the twentieth annual SAS users group international conference, 1995. Citeseer, 1250-1257.
TOUATI, L. & CHALLAL, Y. Efficient CP-ABE Attribute/Key Management for IoT Applications. 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing, 26-28 Oct. 2015 2015. 343-350.
TRAPPE, W., HOWARD, R. & MOORE, R. S. 2015. Low-Energy Security: Limits and Opportunities in the Internet of Things. *IEEE Security & Privacy,* 13**,** 14-21.
USMAN, M., AHMED, I., ASLAM, M. I., KHAN, S. & SHAH, U. A. 2017. SIT: A Lightweight Encryption Algorithm for Secure Internet of Things. *arXiv preprint arXiv:1704.08688*.
UUSITALO, L. 2007. Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Modelling,* 203**,** 312-318.
VAN KRANENBURG, R. 2008. *The Internet of Things: A critique of ambient technology and the all-seeing network of RFID*, Institute of Network Cultures.
VENCKAUSKAS, A., STUIKYS, V., DAMASEVICIUS, R. & JUSAS, N. 2016a. Modelling of Internet of Things units for estimating security-energy-performance

relationships for quality of service and environment awareness. *Security and Communication Networks,* 9**,** 3324-3339.
VENCKAUSKAS, A., STUIKYS, V., TOLDINAS, J. & JUSAS, N. 2016b. A Model-Driven Framework to Develop Personalized Health Monitoring. *Symmetry-Basel,* 8**,** 18.
VERMESAN, O., FRIESS, P., GUILLEMIN, P., GUSMEROLI, S., SUNDMAEKER, H., BASSI, A., JUBERT, I. S., MAZURA, M., HARRISON, M. & EISENHAUER, M. 2011. Internet of things strategic research roadmap. *Internet of Things-Global Technological and Societal Trends,* 1**,** 9-52.
VLACHOPOULOS, K., MAGKOS, E. & CHRISSIKOPOULOS, V. 2013. A model for hybrid evidence investigation. *Emerging Digital Forensics Applications for Crime Detection, Prevention, and Security,* 150.
WANG, J. Y., FLOERKEMEIER, C. & SARMA, S. E. 2014. Session-based security enhancement of RFID systems for emerging open-loop applications. *Personal and Ubiquitous Computing,* 18**,** 1881-1891.
WANG, K., SONG, T. & LIANG, A. Mmda: Metadata Based Malware Detection on Android. 2016 12th International Conference on Computational Intelligence and Security (CIS), 16-19 Dec. 2016 2016a. 598-602.
WANG, K., SONG, T., LIANG, A. L. & IEEE 2016b. Mmda: Metadata based Malware Detection on Android. *Proceedings of 2016 12th International Conference on Computational Intelligence and Security (Cis)***,** 598-602.
WANG, L. C., LI, J. & AHMAD, H. 2016c. Challenges of Fully Homomorphic Encryptions for the Internet of Things. *Ieice Transactions on Information and Systems,* E99D**,** 1982-1990.
WANG, P., CHAO, K. M., LO, C. C., LIN, W. H., LIN, H. C. & CHAO, W. J. 2016d. Using malware for software-defined networking-based smart home security management through a taint checking approach. *International Journal of Distributed Sensor Networks,* 12**,** 23.
WEBSTER, J. & WATSON, R. T. 2002. Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly***,** xiii-xxiii.
WEINBERG, B. D., MILNE, G. R., ANDONOVA, Y. G. & HAJJAT, F. M. 2015. Internet of Things: Convenience vs. privacy and secrecy. *Business Horizons,* 58**,** 615-624.
WEN, Z., YANG, R., GARRAGHAN, P., LIN, T., XU, J. & ROVATSOS, M. 2017. Fog Orchestration for Internet of Things Services. *IEEE Internet Computing,* 21**,** 16-24.
WRENCH, P. & IRWIN, B. 2015. A SANDBOX-BASED APPROACH TO THE DEOBFUSCATION AND DISSECTION OF PHP-BASED MALWARE. *Saiee Africa Research Journal,* 106**,** 46-63.
XIAO, M., ZHOU, J., LIU, X. J. & JIANG, M. D. 2017. A Hybrid Scheme for Fine-Grained Search and Access Authorization in Fog Computing Environment. *Sensors,* 17**,** 22.
XU, J. L., YU, Y. F., CHEN, Z., CAO, B., DONG, W. Y., GUO, Y. & CAO, J. W. 2013. MobSafe: Cloud Computing Based Forensic Analysis for Massive Mobile Applications Using Data Mining. *Tsinghua Science and Technology,* 18**,** 418-427.
YAKUBU, O., ADJEI, O. & NARENDRA, B. C. 2016. A Review of Prospects and Challenges of Internet of Things. *International Journal of Computer Applications,* 139.
YAN, B. & HUANG, G. Supply chain information transmission based on RFID and internet of things. Computing, Communication, Control, and Management, 2009. CCCM 2009. ISECS International Colloquium on, 2009. IEEE, 166-169.
YAN, Z., DING, W., YU, X., ZHU, H. & DENG, R. H. 2016a. Deduplication on encrypted big data in cloud. *IEEE transactions on big data,* 2**,** 138-150.

YAN, Z., WANG, M., LI, Y. & VASILAKOS, A. V. 2016b. Encrypted data management with deduplication in cloud computing. *IEEE Cloud Computing,* 3**,** 28-35.
YAQOOB, I., AHMED, E., HASHEM, I. A. T., AHMED, A. I. A., GANI, A., IMRAN, M. & GUIZANI, M. 2017. Internet of Things Architecture: Recent Advances, Taxonomy, Requirements, and Open Challenges. *IEEE Wireless Communications,* 24**,** 10-16.
YASIN, M., TEKESTE, T., SALEH, H., MOHAMMAD, B., SINANOGLU, O. & ISMAIL, M. 2017. Ultra-Low Power, Secure IoT Platform for Predicting Cardiovascular Diseases. *IEEE Transactions on Circuits and Systems I: Regular Papers,* 64**,** 2624-2637.
YERIMA, S. Y., SEZER, S. & MUTTIK, I. 2015a. High accuracy android malware detection using ensemble learning. *Iet Information Security,* 9**,** 313-320.
YERIMA, S. Y., SEZER, S., MUTTIK, I. & IEEE 2015b. Android Malware Detection: An Eigenspace Analysis Approach. *2015 Science and Information Conference (Sai)***,** 1236-1242.
YU, X. X., YAN, Z. & VASILAKOS, A. V. 2017. A Survey of Verifiable Computation. *Mobile Networks & Applications,* 22**,** 438-453.
YUAN, Z. L., LU, Y. Q. & XUE, Y. B. 2016. DroidDetector: Android Malware Characterization and Detection Using Deep Learning. *Tsinghua Science and Technology,* 21**,** 114-123.
ZANELLA, A., BUI, N., CASTELLANI, A., VANGELISTA, L. & ZORZI, M. 2014. Internet of things for smart cities. *IEEE Internet of Things journal,* 1**,** 22-32.
ZARPELÃO, B. B., MIANI, R. S., KAWAKANI, C. T. & DE ALVARENGA, S. C. 2017. A Survey of Intrusion Detection in I nternet of Things. *Journal of Network and Computer Applications*.
ZAWOAD, S. & HASAN, R. FAIoT: Towards Building a Forensics Aware Eco System for the Internet of Things. Services Computing (SCC), 2015 IEEE International Conference on, 2015. IEEE, 279-284.
ZHANG, K., NI, J., YANG, K., LIANG, X., REN, J. & SHEN, X. S. 2017. Security and Privacy in Smart City Applications: Challenges and Solutions. *IEEE Communications Magazine,* 55**,** 122-129.
ZHANG, Q., ALMULLA, M. & BOUKERCHE, A. 2013. An Improved Scheme for Key Management of RFID in Vehicular Adhoc Networks. *IEEE Latin America Transactions,* 11**,** 1286-1294.
ZHENG, X., MARTIN, P., BROHMAN, K. & XU, L. D. 2014. Cloud Service Negotiation in Internet of Things Environment: A Mixed Approach. *IEEE Transactions on Industrial Informatics,* 10**,** 1506-1515.
ZHOU, B., EGELE, M. & JOSHI, A. High-performance low-energy implementation of cryptographic algorithms on a programmable SoC for IoT devices. 2017 IEEE High Performance Extreme Computing Conference (HPEC), 12-14 Sept. 2017 2017. 1-6.
ZIEGELDORF, J. H., MORCHON, O. G. & WEHRLE, K. 2014. Privacy in the Internet of Things: threats and challenges. *Security and Communication Networks,* 7**,** 2728-2742.
ZIN, T. T., TIN, P. & HAMA, H. Reliability and availability measures for Internet of Things consumer world perspectives. 2016 IEEE 5th Global Conference on Consumer Electronics, 11-14 Oct. 2016 2016. 1-2.
ZULKIPLI, N. H. N., ALENEZI, A. & WILLS, G. B. 2017. IoT Forensic: Bridging the Challenges in Digital Forensic and the Internet of Things.