Journal of Network and Computer Applications 161 (2020) 102630

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

Machine learning based solutions for security of Internet of Things (IoT): A survey

Syeda Manjia Tahsien, Hadis Karimipour*, Petros Spachos

School of Engineering, University of Guelph, Guelph, ON N1G 2W1, Canada

ARTICLE INFO

Keywords:
Architecture
Attack surfaces
Challenges
Internet of Things
IoT attacks
Machine learning
Security solution

ABSTRACT

Over the last decade, IoT platforms have been developed into a global giant that grabs every aspect of our daily lives by advancing human life with its unaccountable smart services. Because of easy accessibility and fast-growing demand for smart devices and network, IoT is now facing more security challenges than ever before. There are existing security measures that can be applied to protect IoT. However, traditional techniques are not as efficient with the advancement booms as well as different attack types and their severeness. Thus, a strong-dynamically enhanced and up to date security system is required for next-generation IoT system. A huge technological advancement has been noticed in Machine Learning (ML) which has opened many possible research windows to address ongoing and future challenges in IoT. In order to detect attacks and identify abnormal behaviors of smart devices and networks, ML is being utilized as a powerful technology to fulfill this purpose. In this survey paper, the architecture of IoT is discussed, following a comprehensive literature review on ML approaches the importance of security of IoT in terms of different types of possible attacks. Moreover, ML-based potential solutions for IoT security has been presented and future challenges are discussed.

1. Introduction

The Internet of Things (IoT) interlink electrical devices with a server and exchanges information without any human intervention (Li et al., 2011; Abane et al., 2019; Sheng et al., 2013; HaddadPajouhet al., 2019). Users can remotely access their devices from anywhere, which makes them vulnerable to different attacks. The security of IoT system is, therefore, a matter of great concern with the increasing number of smart devices nowadays as the devices carry private and valuable information of the clients (Sengupta et al., 2019; Guan et al., 2019). For example, smart home devices and wearable devices hold information about the client's location, contact details, health data, etc. which need to be secured and confidential. Since most of the IoT devices are limited to resources (i.e., battery, bandwidth, memory, and computation), highly configurable and complex algorithm-based security techniques are not applicable (Zhou et al., 2017).

In order to secure IoT systems, Machine learning (ML) based methods are a promising alternative. ML is one of the advanced artificial intelligence techniques which does not require explicit programming and can outperform in the dynamic networks. ML methods can be used to train the machine to identify various attacks and provide corresponding defensive policy. In this context, the attacks can be detected at an early stage. Moreover, ML techniques seem to be promising in detecting new attacks using learning skills and handle them intelligently. Therefore, ML algorithms can provide potential security protocols for the IoT devices which make them more reliable and accessible than before.

Although there are several review articles available in the literature since 2017, none of them focused on particularly applying ML techniques for the security of IoT and covered different types of IoT attacks and all the possible solutions. Starting with, Cui et al. (2018) presented a review on different security attacks in IoT and demonstrated various machine learning based solutions, challenges, and research gap using 78 articles till 2017. In 2018, Xiao et al. (2018a) reviewed different IoT attack models such as spoofing attacks, denial of service attacks, jamming, and eavesdropping and mentioned their possible security solutions based on IoT authentication, access control, malware detections and secure offloading using machine learning techniques. A total of 30 papers have been cited where four different possible ML-based security solutions of IoT devices have been presented. Another research group (Alaa et al., 2017) in 2017 and Chaabouni et al. (2019) in 2019 also published a survey paper on machine learning based security of IoT where the authors specifically focused on intrusion detection and

[*] Corresponding author.
E-mail address: hkarimi@uoguelph.ca (H. Karimipour).

https://doi.org/10.1016/j.jnca.2020.102630
Received 12 November 2019; Received in revised form 20 February 2020; Accepted 20 March 2020
Available online 7 April 2020

various ML related works for the IoT system, respectively. Recently, Zeadally and Tsikerdekis (2020) published a ML based security of IoT review article where the authors have emphasized on the characteristics of IoT devices, presented specific ML techniques (a generalized description on supervised, unsupervised, and Reinforcement learning techniques) and their limitation for securing IoT devices. A total number of 61 papers have been cited in this literature review which indicates a research gap in terms of citation. Therefore, in this literature review a comprehensive research has been conducted to focus more on searching related ML based IoT security papers till 2020 in order to make the current work up to date for the readers.

Contribution of this Review Paper

Based on the information found so far from literature, the contribution of this paper is as follows:

- This literature review concentrates on the ML-based security solutions for IoT systems until most recent articles published in this field till 2020.
- At first, an IoT system with its taxonomy of various layers has been presented. Moreover, security in IoT and different potential attacks have been described with their possible layer-wise effects.
- This survey will also present different machine learning techniques and their applications to address various IoT attacks.
- Besides, a state-of-the-art review has been presented on possible security solutions of IoT devices. It mainly focuses on using different ML algorithms in three architectural layers of the IoT system based on published papers till 2019.
- The authors' present possible challenges/limitations in ML based security of IoT system and their perspective research direction.
- At the end, a statistical overview of published articles on ML based IoT security has been presented which will help researchers to give an idea about published research on particular area of interest and its potentiality.

The rest of the paper is arranged as follows: section 2 presents an overview of security of IoT consisting of IoT layers and importance of security in IoT; section 3 demonstrates attacks in IoT, their affects, and different attack surfaces; section 4 describes ML in IoT security including different types of learning algorithms and solutions for IoT security; research challenges in ML-based security of IoT has been presented in section 5; section 6 shows an analysis of published articles on ML-based security of IoT till date; finally, a conclusion of the survey including future recommendations are presented in section 7.

2. Security of Internet of Things

Security of IoT devices has become a burning question in the twenty-first century. In one side, IoT brings everything close and connects the whole world, on the other hand, it opens various windows to be victimized by different types of attacks.

Although the term IoT is short in its context wise, it contains the entire world with its smart technologies and services that can be imagined. The word IoT was first used by Kevin Ashton in his research presentation in 1999 (Ashton, 2011). From then, IoT is being used to establish a link between human and virtual world using various smart devices with their services through different communication protocols.

What was a dream 25 years ago is now a reality with the help of IoT. In one word, today's advanced world is wrapped by smart technology and IoT is the heart of it. Now, people cannot think a single moment by themselves without using IoT devices and their services. A survey shows that nearly 50 billion things are going to be connected with internet by 2020 and it will increase exponentially as time passes by (da Cruz et al., 2018). An estimated percentage of IoT device users by 2020 is presented in Fig. 1, (Statista, 2019). It is also estimated that IoT is going to capture around 3.9–11.1 trillion USD economical market by 2025 (Manyika et al., 2015). The number of connected IoT devices and global market of IoT system so far and future prediction as well until 2025 (Juniper Research, 2015; Statista, Technology & Telecommunication, 2019), is illustrated in Fig. 2. Therefore, research on IoT and its development and security has received huge attention over the last decades in the field of electrical and computer since. This following two sections will discuss IoT layers and security challenges.

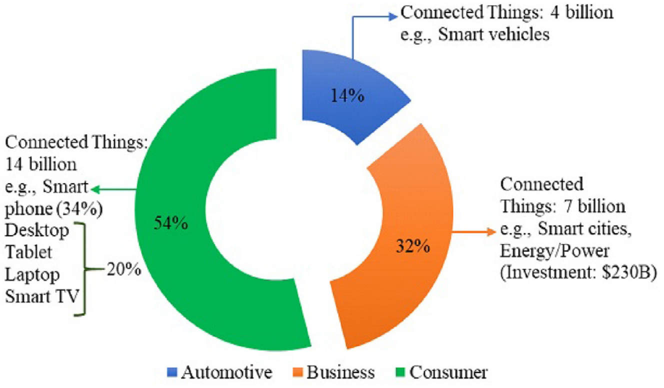


Fig. 1. Estimated IoT device users by 2020.

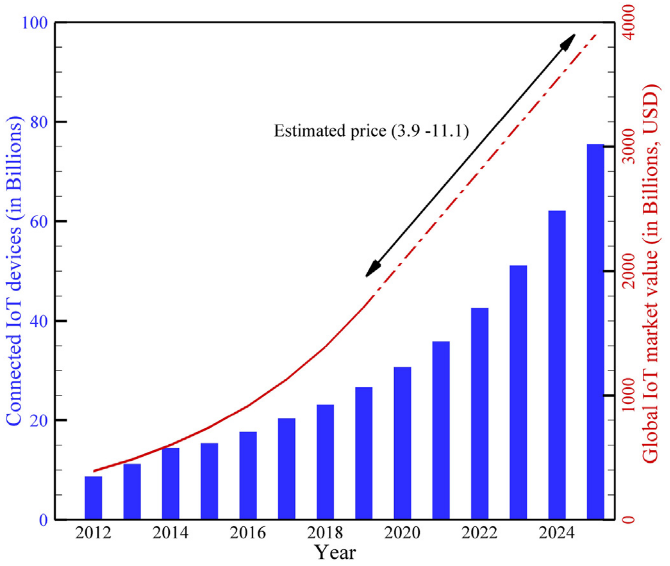


Fig. 2. Graphical presentation of total connected IoT devices and global IoT market so far and future prediction.

2.1. IoT layers

The architecture of IoT, which is a gateway of various hardware applications, is developed in order to establish a link and to expand IoT services at every doorstep. Different communication protocols, including Bluetooth, WiFi, RFID, narrow and wideband frequency, ZigBee, LPWAN, IEEE 802.15.4, are adopted in different layers of IoT architecture to transmit and receive various information/data (Saadeh et al., 2018; Gazis, 2017).

Moreover, large scale high-tech companies have their own IoT platforms to serve their valuable customers, such as Google Cloud, Samsung Artik Cloud, Microsoft Azure suite, Amazon AWS IoT, etc. (Xu et al., 2018). A standard architecture of IoT consists of mainly three layers i.e., perception/physical layer, network layer, and web/application layer (Elazhary, 2019) as shown in Fig. 3.

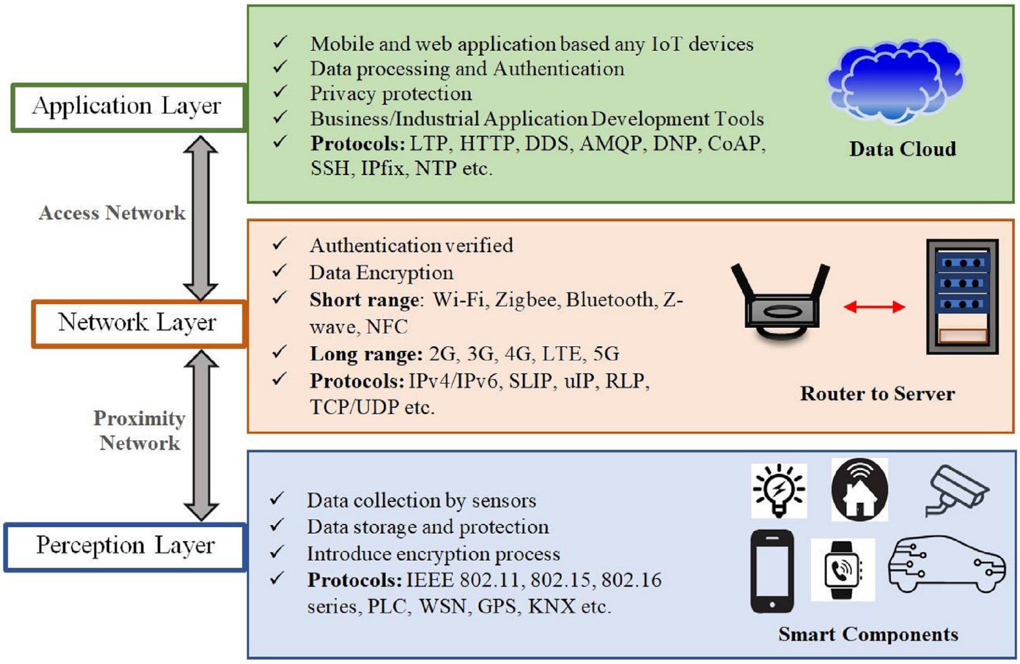

Application Layer
✓ Mobile and web application based any IoT devices
✓ Data processing and Authentication
✓ Privacy protection
✓ Business/Industrial Application Development Tools
✓ Protocols: LTP, HTTP, DDS, AMQP, DNP, CoAP, SSH, IPfix, NTP etc.
Data Cloud
Access Network
Network Layer
✓ Authentication verified
✓ Data Encryption
✓ Short range: Wi-Fi, Zigbee, Bluetooth, Z-wave, NFC
✓ Long range: 2G, 3G, 4G, LTE, 5G
✓ Protocols: IPv4/IPv6, SLIP, uIP, RLP, TCP/UDP etc.
Router to Server
Proximity Network
Perception Layer
✓ Data collection by sensors
✓ Data storage and protection
✓ Introduce encryption process
✓ Protocols: IEEE 802.11, 802.15, 802.16 series, PLC, WSN, GPS, KNX etc.
Smart Components

Fig. 3. IoT layers architecture.

2.1.1. *Application layer*

The application layer is the third layer in IoT systems which provides service to the users through mobile and web-based softwares. Based on recent trends and usages of smart things, IoT has numerous applications in this technologically advanced world. Living space/homes/building, transportation, health, education, agriculture, business/trades, energy distribution system, etc. have become smart by the grace of IoT system and it uncounted service (Amendola et al., 2014; Camara et al., 2015).

2.1.2. *Network layer*

The network layer is more important in IoT systems because it acts as a transmission/redirecting medium for information and data using various connection protocols, including GSM, LTA, WiFi, 3-5G, IPv6, IEEE 802.15.4, etc, which connect devices with smart services (Singh et al., 2019). In the network layer, there are local clouds and servers that store and process the information which works as a middle-ware between the network and the next layer (Razzaque et al., 2016; Neely et al., 2006).

Big data is another important factor in the network layer because it attracts the attention of today's ever-growing economical market. The physical objects from the physical layer are producing a huge amount of information/data continuously which are being transmitted, processed, and stored by IoT systems. Since information/data are important for smart services in the network layer, ML and Deep Learning (DL) are extensively used nowadays to analysis the stored information/data to utilize better analysis techniques and extract good uses from it for smart devices (Ahmedet al., 2017).

2.1.3. *Perception layer*

The first layer of IoT architecture is the perception layer which consists of the physical (PHY) and medium access control (MAC) layers. The PHY layer mainly deals with hardware i.e., sensors and devices that are used to transmit and receive information using different communication protocols e.g., RFID, Zigbee, Bluetooth, etc (Asghari et al., 2018; Sethi and Sarangi, 2017).

The MAC layer establishes a link between physical devices and networks to allow to for proper communication. MAC uses different protocols to link with network layers, such as LAN (IEEE 802.11ah), PAN (IEEE 802.15.4e, Z-Wave), cellular network (LTE-M, EC-GSM). Most of the devices in IoT layers are plug and play types from where a huge portion of big data are produced (Tsai et al., 2014; Saggi and Jain, 2018; Gil et al., 2016; Alam et al., 2017; Sezer et al., 2018).

2.2. *Importance of security in IoT*

IoT devices are used for various purposes through an open network which makes the devices, therefore, more accessible to the users. In one hand, IoT makes human life technologically advance, easy going, and conformable; on the other hand, IoT puts the users' privacy more in danger due to different threats/attacks (Makhdoom et al., 2018; Farris et al., 2019). Since anyone can access certain IoT devices from anywhere without the user permission, the security of IoT devices has become a burning question. A wide range of security systems must be implemented to protect the IoT devices. However, the physical structure of IoT devices limits its computational functionality which restricts the implementation of complex security protocol (Abomhara, 2015). When an intruder accesses a system and exposes private information without the corresponding user's permission, this is considered as a threat/attack (Benkhelifa et al., 2018).

3. Attacks in IoT

Over the last few years, the IoT system has been facing different attacks which make the manufacturers and users conscious regarding developing and using IoT devices more carefully. This section describes different kind of attacks, their effects, and attack surfaces in IoT.

3.1. *Types of attack*

IoT attacks can be classified mainly as cyber and physical attacks where cyber attacks consist of passive and active attacks (see Fig. 4). Cyber attacks refer to a threat that targets different IoT devices in a wireless network by hacking the system in order to manipulate (i.e., steal, delete, alter, destroy) the user's information. On the other hand, physical attacks refer to the attacks that physically damage IoT devices.

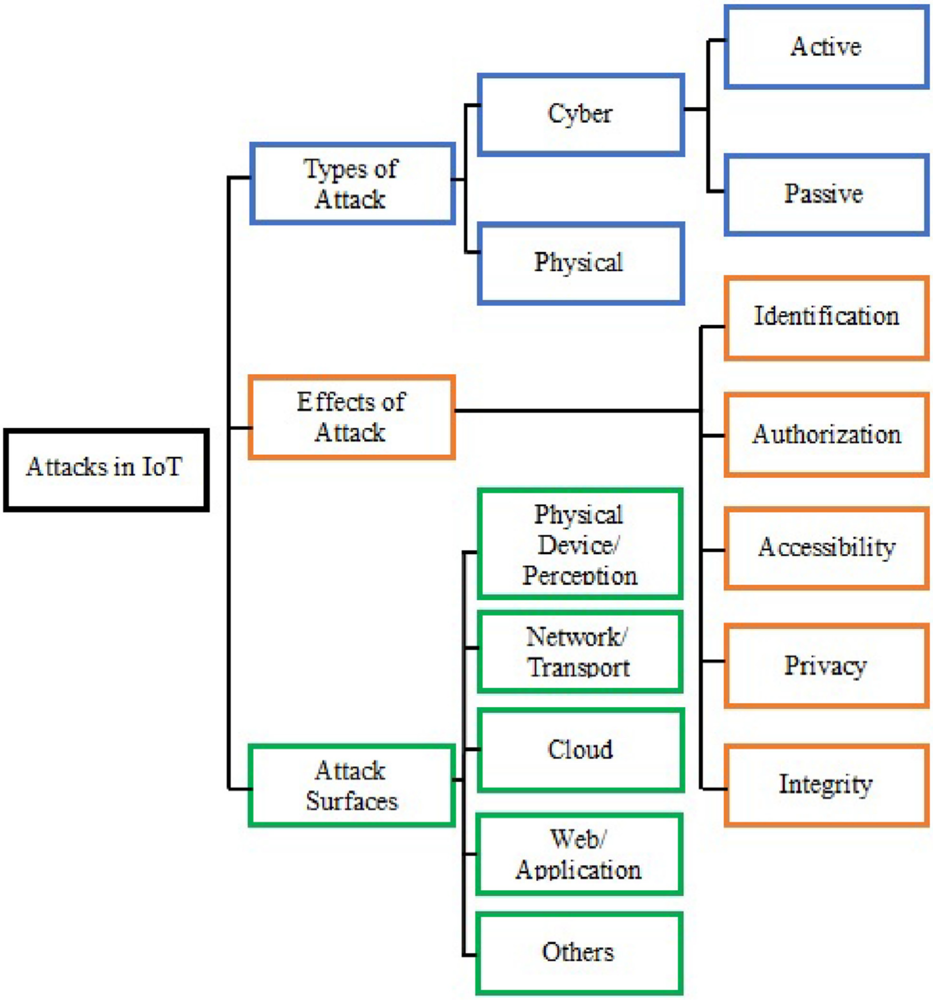


Fig. 4. A diagram of a detailed list of IoT security attacks that includes different types of attack, attack surfaces, and attack effects.

Here, the attackers do not need any network to attack the system. Therefore, this kind of attacks are subjected to physical IoT devices e.g., mobile, camera, sensors, routers, etc., by which the attackers interrupt the service (Roman et al., 2013).

The following subsections mainly focus on the different types of cyber attacks according to their severeness in IoT devices with Active and passive being the two main categories of a cyber attacks.

3.1.1. Active attacks

An active attack happens when an intruder accesses the network and its corresponding information to manipulate the configuration of the system and interrupt certain services. There are different ways to attack IoT device security, including disruption, interventions, and modifications under active attacks. Active attacks such as DoS, man-in-the-middle, sybil attack, spoofing, hole attack, jamming, selective forwarding, malicious inputs, and data tampering, etc. are illustrated in Fig. 5.

3.1.1.1. Denial of service attacks. Denial of Service (DoS) attacks are mainly responsible for disrupting the services of system by creating several redundant requests (see Fig. 5). Therefore, the user can not access and communicate with the IoT device which makes it difficult to take the right decision. In addition, DoS attacks keep IoT devices always turned on, which can ultimately affect the battery lifetime. A special type of an attack named Distributed DoS (DDoS) attack occurs when consists several attacks happen using different IPs to create numerous requests and keep the server busy. This makes it hard to differentiate between the normal traffic and attack traffic (Andrea et al., 2015). In recent years, a unique IoT botnet virus named Mirai was responsible for introducing destructive DDoS attacks that have damaged thousands of IoT devices through interferences (Bertino and Islam, 2017; Karimipour et al., 2019a; Yao et al., 2015; Ahmed et al., 2019; Mohammadi et al., 2018a).

3.1.1.2. Spoofing and Sybil attacks. Spoofing and Sybil attacks mainly target the identification (RFID and MAC address) of the users in order to access the system illegally in the IoT system (see Fig. 5). It is noticed that TCP/IP suite does not have strong security protocol which makes the IoT devices more vulnerable, especially to spoofing attacks. Moreover, these two attacks initiate further severe attacks, including DoS and man in the middle attacks (Xiao et al., 2016a).

3.1.1.3. Jamming attacks. ongoing communication in a wireless network by sending unwanted signals to the IoT devices which causes problems for the users by keeping the network always busy (Han et al., 2017) (see Fig. 5). In addition, this attack degrades the performance of the IoT devices by consuming more energy, bandwidth, memory, etc.

3.1.1.4. Man in the middle attacks. Man in the middle attackers pretend to be a part of the communication systems where the attackers are directly connected to another user device (see Fig. 5). Therefore, it can easily interrupt communications by introducing fake and misleading

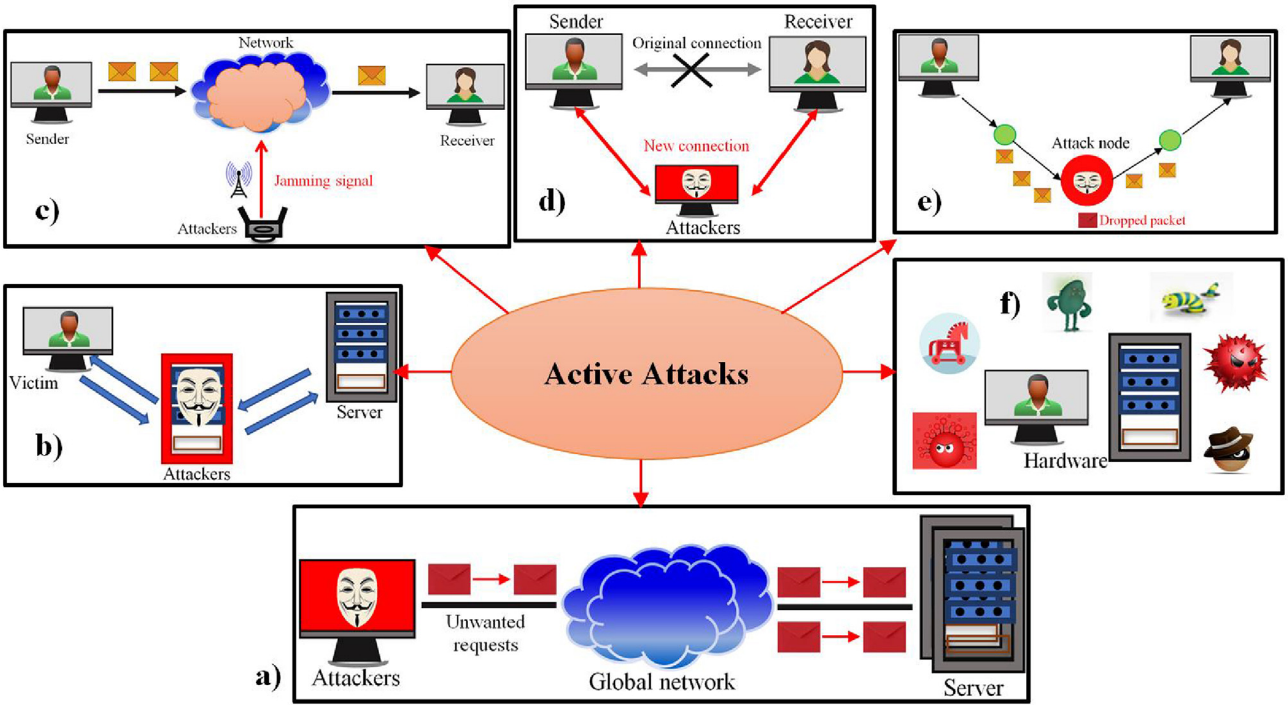


Fig. 5. Schematic diagram of different types of cyber attacks: a) Denial of Service attack, b) Spoofing and Sybil attacks, c) Jamming attack, d) Man in the middle attack, e) Selective Forwarding attack, f) Malicious input attack.

data in order to manipulate original information (Andrea et al., 2015).

3.1.1.5. Selective forwarding attacks. Selective forwarding attack acts as a node in the communication system which allows dropping some packets of information during transmission to create a hole in the network (see Fig. 5). This type of attack is hard to identify and avoid.

3.1.1.6. Malicious input attacks. Malicious input attacks include malware software attacks, such as trojans, rootkit, worms, adware, and viruses, which are responsible for the damage of IoT devices such as financial loss, power dissipation, degradation of the wireless network performance (Zhou et al., 2017; Xiao et al., 2017; Karimipour et al., 2019b) (see Fig. 5).

3.1.1.7. Data tampering. In data tampering, the attackers manipulate the user's information intentionally to disrupt their privacy using unwanted activities. The IoT devices that carry important user's information such as location, fitness, billing price of smart equipment are in great danger to encounter these data tampering attacks (Bekara, 2014).

3.1.2. Passive attack

Passive attacks try to gather the user's information without their consent and exploit this information in order to decrypt their private secured data (AlTawy and Youssef, 2016). Eavesdropping and traffic analysis are the main two ways to perform a passive attack through an IoT network. Eavesdropping mainly deploys the user's IoT device as a sensor to collect and misuse their confidential information and location (Wamba et al., 2013; Malasri and Wang, 2009; Spachos et al., 2018).

3.2. Affects of attacks

The affects of IoT attacks are threatening for the network in order to protect the user's privacy, authentication, and authorization. A detailed list of different types of attacks including their affects on IoT devices are presented in Fig. 6. The following features need to be considered while developing any security protocol to encounter the attacks for the IoT system.

3.2.1. Identification

Identification refers to the authorization of the user in the IoT network. Clients need to be registered first to communicate with the cloud server. However, trade-offs and robustness of IoT systems create challenges for identification (Bose et al., 2015). Sybil and spoofing attacks are responsible for damaging the security of the network and the attackers can easily get access to the server without proper identification. Therefore, an effective identification scheme for the IoT system is necessary which can provide strong security while having system restrictions (Yao et al., 2015).

3.2.2. Authorization

Authorization deals with the accessibility of the user to an IoT system. It gives permission to only the authorized clients to enter, monitor and use information data of the IoT network. It also executes the commands of those users who have authorization in the system. It is really challenging to maintain all user's logs and give access based on the information, since users are not only confined to humans but also sensors, machines, and services (Ahmed et al., 2019). Moreover, the formation of a strong protective environment is a difficult task while processing the client's large data sets (Moosaviet al., 2015).

3.2.3. Accessibility

Accessibility ensures that the services of the IoT system are always rendered to their authorized users. It is one of the important requirements to create an effective IoT network while DoS and jamming attacks disrupt this service by creating unnecessary requests and keep the network busy. Hence, a strong security protocol is needed to maintain the services of IoT devices to be available to their clients without any interruption (Restuccia et al., 2018).

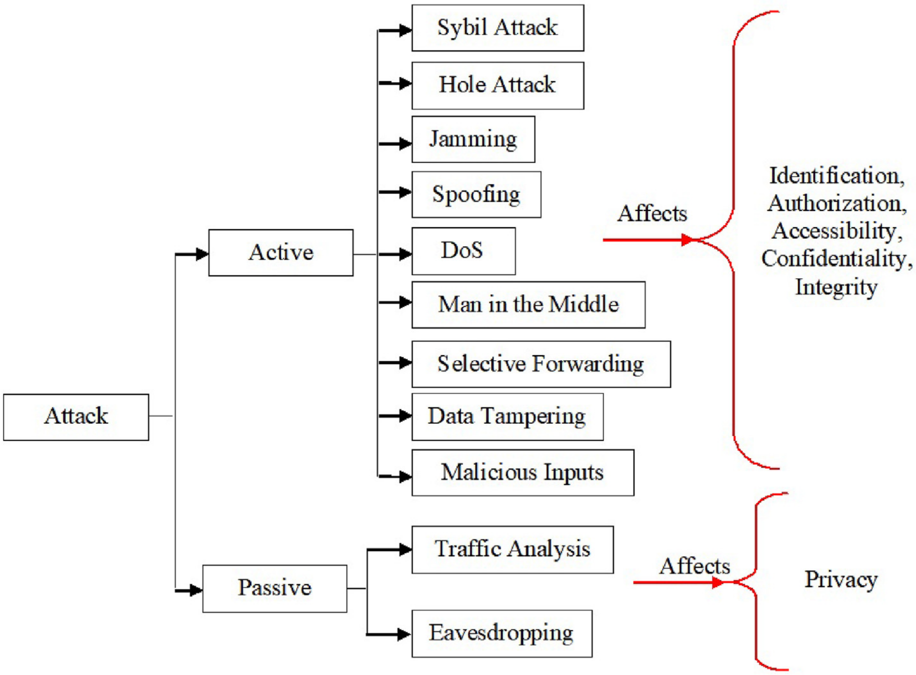


Fig. 6. A list of different kinds of active and passive attacks including their affects.

3.2.4. Privacy

Privacy is the only factor that both active and passive attacks are facing in IoT system. Nowadays everything, including sensitive and personal information, medical reports, national defense data, etc., are stored and transferred securely through the internet using different IoT devices which are supposed not to be disclosed by any unauthorized users (Roman et al., 2013). However, it is hard to keep most data confidential from unauthorized third parties since attackers can identify the physical location by tracking the IoT device and decrypt the information (Camara et al., 2015).

3.2.5. Integrity

Integrity property ensures that only authorized users can modify the information of the IoT devices while using a wireless network for communication. This requirement is fundamental for the security of IoT system to protect it from various malicious input attacks such as structured query language (SQL) injection attacks (Karimipour and Dinavahi, 2017a). If this feature is compromised somehow by irregular inspection during data storage in IoT devices, it will affect the functionality of those devices in the long run. In some cases, it can not only reveal the sensitive information but also sacrifice human lives (Camara et al., 2015).

3.3. Surface attacks

The architecture of IoT includes mainly three layers which have been demonstrated in section 2; however, four potential surfaces of IoT have been presented in order to describe attack surfaces more precisely possible attacks besides those three layers in this section (see Fig. 7). Here, the IoT surface attacks are categorized as a physical device/perception surface, network/transport surface, could surface, application/web surface. Moreover, considering the development of smart technologies in IoT system (e.g., smart grid, smart vehicles, smart house, etc.), new surface attacks such as attacks by interdependent, interconnected, and social IoT system are also discussed in this section.

3.3.1. Physical device/perception surface attacks

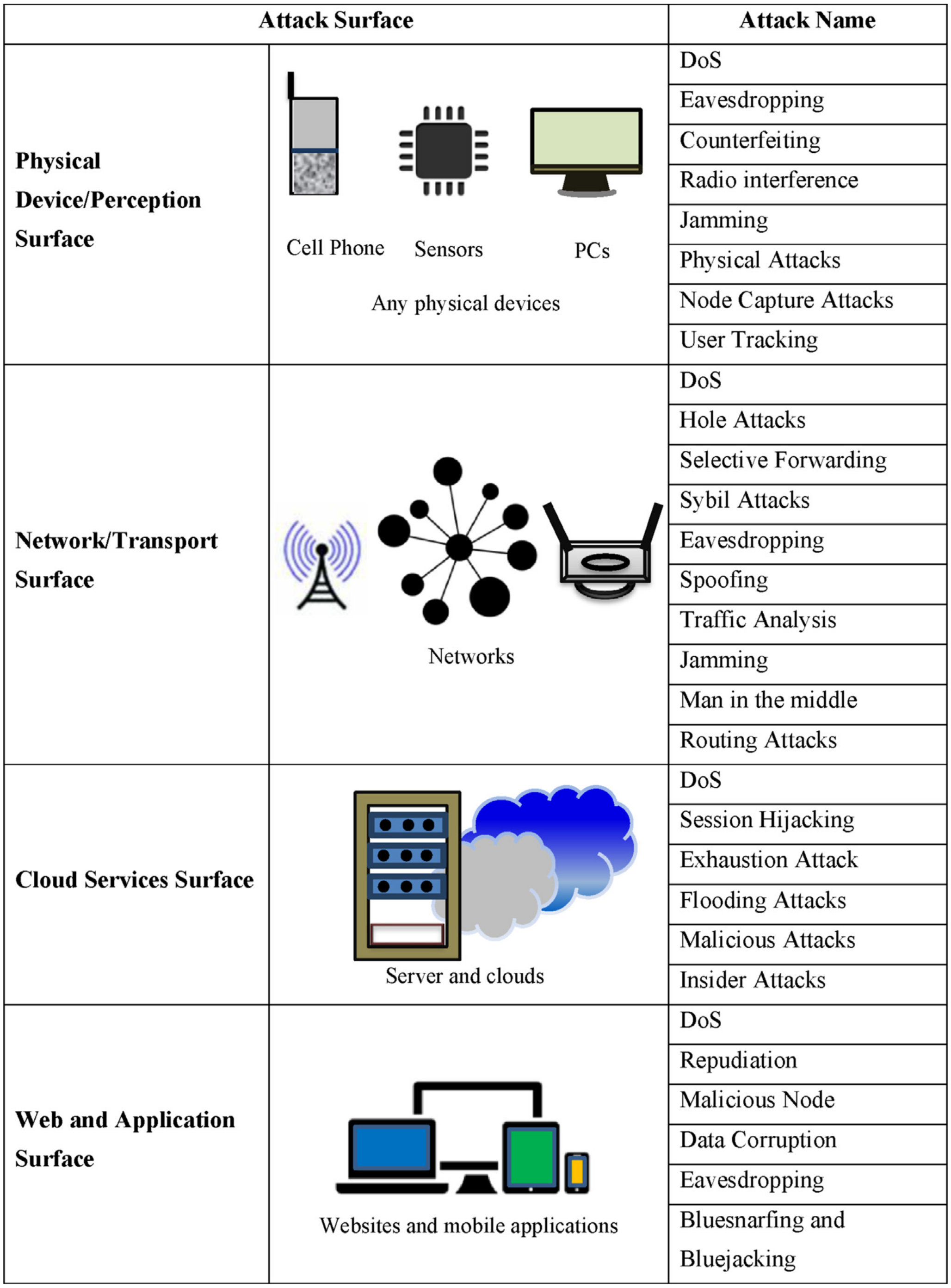
Physical devices are known as a direct surface attack of the IoT system, since they carry confidential and important information of users. Moreover, attackers can easily access the physical layer of IoT devices. RFID tags, sensors, actuators, micro-controllers, RFID readers are some units of physical devices which are used for identification, communication, collecting and exchanging information, (Atzori et al., 2010). These parts are vulnerable to DoS, eavesdropping, jamming, radio interference (Jing et al., 2014). However, physical attacks are the most alarming for physical device surface.

3.3.2. Network/transport surface attacks

Physical devices are connected through network services, including wired and wireless networks in IoT systems. Sensor networks (SNs) play an important role to develop an IoT network. Therefore, wired and wireless sensor network needs to be integrated to construct a large scale IoT surface. This large scale IoT surface is a potential target for different types of attacks as the user's information transfer openly through the sensor networks without any strong security protocol (Asghari et al., 2018; Jing et al., 2014). In order to launch an attack in a network service surface, attackers will always try to find an open ports or weak routing protocol to access the user network by using their IP address, gateway, and MAC address to manipulate the sensitive information (Mohammadi et al., 2018a; Modi et al., 2013). Network surface attacks are prone to DoS, jamming, man in the middle, spoofing, Sybil, selective forwarding, traffic analysis, hole attacks, internet attacks, routing attacks, and so on (Liu et al., 2016).

3.3.3. Cloud surface attacks

Besides self-storage capacity, the IoT devices now rely upon the could system which connects most of the smart devices and has unlimited storage capacity (Daz et al., 2016). This cloud computing technology enables its stored resources to share remotely for other users (Armbrustet al., 2010). Cloud computing, therefore, has become the base platform for IoT devices to transport a user's information and store it. Moreover, this could service makes IoT systems dynamic and updates



Attack Surface		Attack Name
Physical Device/Perception Surface	Cell Phone Sensors PCs Any physical devices	DoS
		Eavesdropping
		Counterfeiting
		Radio interference
		Jamming
		Physical Attacks
		Node Capture Attacks
		User Tracking
Network/Transport Surface	Networks	DoS
		Hole Attacks
		Selective Forwarding
		Sybil Attacks
		Eavesdropping
		Spoofing
		Traffic Analysis
		Jamming
		Man in the middle
		Routing Attacks
Cloud Services Surface	Server and clouds	DoS
		Session Hijacking
		Exhaustion Attack
		Flooding Attacks
		Malicious Attacks
		Insider Attacks
Web and Application Surface	Websites and mobile applications	DoS
		Repudiation
		Malicious Node
		Data Corruption
		Eavesdropping
		Bluesnarfing and Bluejacking

Fig. 7. Different attack surfaces of IoT including possible attacks (Mamdouh et al., 2018; Jing et al., 2014).

it in a real-time manner (Fremdt et al., 2013; Ukil et al., 2014). Therefore, users who are utilizing similar clouds can have their data hacked, stolen, and manipulate through surface attacks. Also, DoS, flooding attacks, insider attacks and malicious attacks can be exposed to cloud surfaces (Modi et al., 2013).

3.3.4. *Web and application surface attacks*

Over the last decades, the smart technology is growing very fast which results in increasing demand of IoT devices in order to remote access and control smart devices, such as smart cars, home assistance, watches, glasses, lights and fitness devices. Web and mobile applications make it possible to remotely access and control IoT devices. IoT devices are connected with the network through servers and clouds using a web mobile software based applications. Since there is a technological boom and a merge between the real and virtual world, it is difficult to distinguish between them in the near future. In addition, real-time technology makes IoT devices more alive using smart technologies (S. O. Technologies, 2019). Smart devices, such as android operating system based gadgets has attracted the market's attention due to their relatively simple and open architecture and application programming interface (Farukiet al., 2015). Therefore, third parties can easily upload their applications on the cloud which creates a way for malware developers to launch different malicious attacks to access IoT devices with/without a user's permission (Huang et al., 2014). Therefore, smart devices that utilize web and mobile applications are vulnerable to DoS, data corruption, eavesdropping, bluejacking, bluesnarfing etc (Bekara, 2014; Alaba et al., 2017).

3.3.5. Other attacks

Other new surface attacks are initiated by IoT systems because of a smart technology that is attacked by interdependent, interconnected and social IoT systems (Zhou et al., 2018). Attacks that are caused by interdependent IoT systems refer to where the attacker does not need to identify a user's device to attack. For example, a smart building has different kinds of sensors which controls the temperature, air-condition, lighting system. These sensors also depend on other sensors which are connected to the clouds for updating and real-time operation. Since most of IoT devices are interconnected through a global network that creates a wide range of surface attacks for IoT devices, it increases the potential of different types of attacks. Any contaminated treats can easily spread out to other IoT devices because of the interconnected systems. Social surface attacks are new to IoT system due to the increasing number of social sites which involve the user to share their private information with another user. Thus, these social sites may exploit the user's information for any illegal actions (Nitti et al., 2015; Atzori et al., 2014).

4. Machine learning (ML) in IoT security

ML is one of the artificial intelligence techniques which trains machines using different algorithms and helps devices learn from their experience instead of programming them explicitly (Jordan and Mitchell, 2015). ML does not need human assistance, complicated mathematical equations, and can function in the dynamic networks. In the past few years, ML techniques have been advanced remarkably for IoT security purposes (Alsheikh et al., 2014; Butun et al., 2014). Therefore, ML methods can be used to detect various IoT attacks at an early stage by analyzing the behaviour of the devices. In addition, appropriate solutions can be provided using different ML algorithms for resource-limited IoT devices. This section is divided into following two subsections i.e., ML Techniques and ML-based solutions for IoT security.

4.1. ML techniques

ML techniques including supervised techniques, unsupervised techniques, and reinforcement learning can be applied to detect smart attacks in IoT devices and to establish a strong defensive policy. Fig. 8 illustrates different machine learning algorithms used for the security of the IoT systems.

4.1.1. Supervised learning

Supervised learning is the most common learning method in machine learning where the output is classified based on the input using a trained data set which is a learning algorithm. Supervised learning is classified as classification and regression learning.

Classification Learning: Classification learning is a supervised ML algorithm where the output is a fixed discrete value/category e.g., [True, False] or [Yes, No], etc. The following subsections will demonstrate different types of classification learning, including Support Vector Machine, Bayesian Theorem, K-Nearest Neighbor, Random Forest, and Association Rule.

4.1.1.1. Support Vector Machine (SVM). SVM algorithm is used to analyze data that use regression and classification analysis. SVM creates a plane named hyperplane between two classes. The goal of the hyperplane is to maximize the distance from each class which distinguishes each class with a minimum error at maximum margin (Buczak and Guven, 2015). If the hyperplane becomes nonlinear after analysis, then SVM uses kernel function to make it linear by adding new features. Sometimes it is hard to use the optimal kernel function in SVM. However, SVM possesses a high accuracy level which makes it suitable for security applications in IoT like intrusion detection (Liu and Pi, 2017; Modiri et al., 2018), malware detection (Ham et al., 2014), smart grid attacks (Karimipour and Dinavahi, 2017b) etc.

4.1.1.2. Bayesian Theorem. The Bayesian theorem is based on the probability of statistics theorem for learning distribution which is known as Bayesian probability. This kind of supervised learning method gets new results based on present information using Bayesian probability. This is known as Nave Bayes (NB). Therefore, NB has been a widely used learning algorithm that needs the prior information in order to implement the Bayesian probability and predict probable outcomes. This is one of the challenges that can successfully be deployed in IoT. NB is usually used in IoT to detect intrusion detection in the network layer (Panda and Patra, 2007; Mukherjee and Sharma, 2012) and anomaly detection (Agrawal and Agrawal, 2015; Swarnkar and Hubballi, 2016). NB has some advantages, such as simple to understand, requiring less data for classifications, easy to implement, applicable for multi-stage calcification. NB depends on features, interactions between features, and prior information which might resist getting accurate outcome (Box and Tiao, 2011).

4.1.1.3. K-nearest neighbor (KNN). KNN refers to a statistical nonparametric method in supervised learning which usually uses Euclidian distance (Chen et al., 2015). Euclidian distance in KNN determines the average value of unknown node which is k nearest neighbors (Deng et al., 2016). For instance, if any node is lost, then it can be anticipated from the nearest neighbor's average value. This value is not accurate but helps to identify the possible missing node. KNN method is used in intrusion detection, malware detections, and anomaly detection in IoT. KNN algorithm is simple, cheap, and easy to apply (Adetunmbi et al., 2008; Tsai et al., 2009). In contrast, it is a time-consuming process to identify the missing nodes which are challenging in terms of accuracy.

4.1.1.4. Random Forest (RF). RF is a special ML method which uses a couple of Decision trees (DTs) in order to create an algorithm to get an accurate and strong estimation model for outcomes. These several trees are randomly developed and trained for a specific action that becomes the ultimate outcome from the model. Although RF uses DTs, the learning algorithm is different because RF considers the average of the output and requires less number of inputs (Breiman, 2001; Cutleret al., 2007). RF is typically used in DDoD attack detection (Doshi et al., 2018), anomaly detection (Chang et al., 2017), and unauthorized IoT devices identification (Meidanet al., 2017a) in network surface attacks. A previous literature shows that RF gives better result in DDoS attack detection over SVM, ANN, and KNN (Doshi et al., 2018). Despite RF not being useful in real time applications, it needs a higher amount of training data sets to construct DTs that identify sudden unauthorized intrusions.

4.1.1.5. Association Rule (AR). AR method is another kind of supervised ML technique which is used to determine the unknown variable depending on the mutual relationship between them in a given data set (Agrawal et al., 1993). AR method was successfully used in intrusion detection in Tajbakhsh et al. (2009) where fuzzy AR was used to detect the intrusion in the network. AR is also simple and easy to adopt; however, it is not commonly used in IoT as it has high time complexity and gives results on assumptions that may not provide an accurate outcome for a large and complex model (Kotsiantis and Kanellopoulos, 2006).

Regression Learning: Regression learning refers to where the output of the learning is a real number or a continuous value depending on the input variables. Different RLs like Decision Tree, Neural Network, Ensemble Learning are presented in the follows subsections.

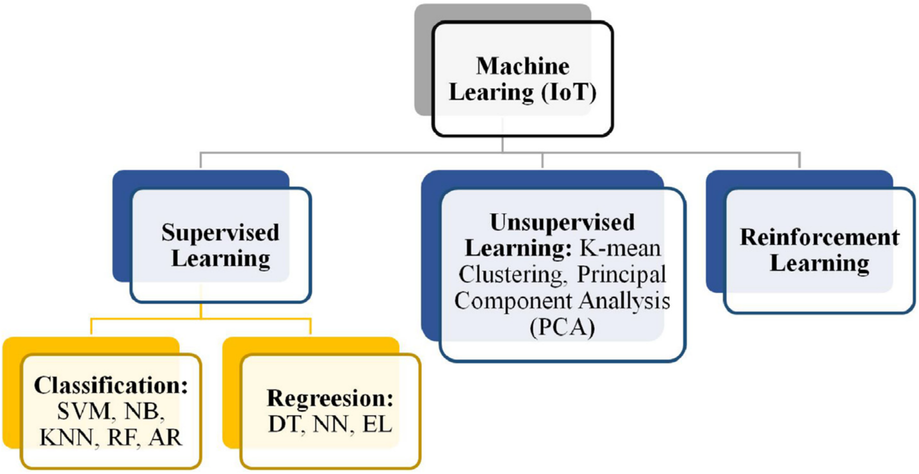

Machine Learing (IoT)
Supervised Learning
Unsupervised Learning: K-mean Clustering, Principal Component Anallysis (PCA)
Reinforcement Learning
Classification: SVM, NB, KNN, RF, AR
Regreesion: DT, NN, EL

Fig. 8. Machine learning and its classification.

4.1.1.6. Decision Tree (DT). DT is a natural supervised learning method which is like a tree that has branches and leaves. DT has different branches as edges and leaves as nodes. DTs are used to sort out the given samples based on the featured values. DT in ML is mainly categorized as classification and regression (Kotsiantis, 2013). DT has advantages over other ML techniques like simple construction, easy to implement, handling large data samples, and being transparent (Kotsiantis et al., 2007; Quinlan, 1986). In contrast, this technique has some disadvantages such as requiring a big space to store the data due to its large construction. This makes the learning algorithm more complex if several DTs are considered to eliminate the problem (Kotsiantis et al., 2007; Quinlan, 1986). DTs are widely used as classifier in security application like DDoS and intrusion detection (Kim et al., 2014).

4.1.1.7. Neural Network (NN). NN technique is constructed based on the human's brain structure which uses neuron. NN has widely used ML techniques that can deal with complex and nonlinear problems (Gondhi and Gupta, 2017; Hush and Horne, 1993). Hierarchical and interconnected are the two main network categories in NN algorithm based on different functional layers of the neuron (typically: input, hidden and output layers. NN techniques reduce the network response time and subsequently increases the performance of the IoT system. However, NN are computationally complex in nature and hard to implement in a distributed IoT system.

4.1.1.8. Ensemble Learning (EL). EL is a rising learning algorithm in ML where EL uses different classification techniques to get an acceptable outcome by increasing its performance. EL usually combines homogeneous or heterogeneous multi-classifier to get an accurate outcome. Since EL uses several learning algorithms, it is well fitted to solve most problems. However, EL has a high time complexity compared to any other single classifier method. El is commonly used for anomaly detection, malware detection, and intrusion detection (Aburomman and Reaz, 2016; Mohammadi et al., 2018b).

4.1.2. Unsupervised learning

In Unsupervised learning, there is no output data for given input variables. Most of the data are unlabeled where the system tries to find out the similarities among this data set. Based on that, it classifies them into different groups as clusters. Many unsupervised learning techniques have been used for security of IoT devices to detect DoS attacks (using multivariate correlation analysis) and privacy protection (applying infinite Gaussian mixture model (IGMM)) (Tan et al., 2013; Xiao et al., 2013). The following sub-section will focus on the types of unsupervised learning that includes Principal Component Analysis (PCA) and K-means Clustering technique.

4.1.2.1. Principal Component Analysis (PCA). PCA which is also known as a feature reduction technique converts a large data set into smaller ones but holds the same amount of information as in the large set. Therefore, PCA decreases the complexity of a system. This method can be used for selecting a feature to detect real-time intrusion attacks in an IoT system (Wold et al., 1987). The combination of PCA and some other ML methods can be applied to provide a strong security protocol. A model proposed by (Zhao et al., 2017) uses PCA and classifier algorithms, such as KNN and softmax regression to provide an efficient system.

4.1.2.2. K-mean clustering. This unsupervised learning technique creates small groups in order to categorize the given data samples as a cluster. This is a well-known algorithm that uses clustering methods. There are some simple rules to implement this method such as i) Firstly, differentiate the given data set into various clusters where each cluster has a centroid (k-centroid) where the main target is to determine k-centroid for each cluster; ii) Then, select a node from each cluster and relate this with the nearest centroid and keep doing this until every node is contacted. Then, recalculation is performed based on the average value of node from every cluster; iii) Finally, the method redo its prior steps until it coincides to get the K-mean value (Hartigan and Wong, 1979; Jain, 2010). K-mean learning techniques are useful especially for smart city to find suitable areas for living. K-mean algorithms are also useful in IoT system when labeled data is not required due to its simplicity. However, this unsupervised learning algorithm is less effective compared to supervised learning. K-mean clustering method is usually used in anomaly detection (Bhuyan et al., 2014; Muniyandi et al., 2012) and Sybil attack detection (Xie et al., 2017).

4.1.3. Reinforcement learning (RL)

RL allows the machine to learn from interactions with its environment (like humans do) by performing actions to maximize the total feedback (Mnihet al., 2015). The feedback might be a reward that depends on the output of the given task. In reinforcement learning, there are no predefined actions for any particular task while the machine uses trial and error methods. Through trial and error, the agent can identify and implement the best method from its experience to gain the highest reward.

Many IoT devices (e.g., sensors, electric glass, air conditioner) use reinforcement learning to make changes according to the environment. Moreover, RL techniques have been used for security of IoT devices, including Q-learning, deep Q-network (DQN), post-decision state (PDS), and Dyna-Q to detect various IoT attacks and provide suitable security protocols for the devices. In Xiao et al. (2016a), Xiao et al. (2017), Xiao et al. (2016b) and Li et al. (2016), Q-learning has been used for authentication, jamming attacks, and malicious inputs whereas Dyna-Q in malware detection and authentication. In addition, DQN and PDS can provide security for jamming attacks and malware detection, respectively (Han et al., 2017).

4.2. ML based solution for IoT security

ML-based security solutions field for IoT devices has become an emerging research area and is attracting the attention of today's researchers to add more to this field over the last few years. In this section, different ML methods have been presented as a potential solutions for securing IoT systems. These solutions have been investigated based on three main architectural layers of an IoT system, including physical/perception layer, network layer, and web/application layer-wise.

4.2.1. Physical/perception layer

Traditional authentication methods used for securing the physical surface is not quite sufficient due to the exact threshold value to detect the unwanted signals which give fake alarm (Xiao et al., 2016a). Therefore, ML-based learning methods can be an alternative for authentication in the physical layer. Xiao et al. (2016a) reported that Q-learning based learning methods reduces the authentication error by about 64.3% and shows better performance than usual physical layer authentication methods using 12 transmitters. In another study, supervised ML techniques such as Distributed Frank Wolf and Incremental Aggregated Gradient were applied to determine the logistics regression model's parameters in order to reduce the communication overhead and increase the efficiency of spoofing detection (Xiao et al., 2018b). Besides, unsupervised learning like IGMM is also used to secure the physical surface and ensure the authentication of IoT devices (Xiao et al., 2018b).

Research in Wang et al. (2017a), Shi et al. (2017) and Namvar et al. (2016) showed that RL techniques can effectively address jamming attacks for the security of IoT. A method was proposed for the aggressive jamming attack in Namvar et al. (2016) where a centralized system scheme was considered. An intelligent power distribution strategy and IoT access point were used to work against the jamming attackers. In another study (Han et al., 2017), RL and deep CNN were combined to avoid jamming signals for cognitive radios that increase RL performance. Cognitive radio (CR) devices have dynamic changing capabilities according to working environments (Bkassiny et al., 2013).

Recently, ML-based a new centralized scheme was proposed in Kiran et al. (2018) for the security of IoT devices. Basically, it permits certain users with authorization to communicate with the system and safely store authorized users's information. In the proposed peer-to-peer security protocol scheme, clients need to be registered first to the cloud server before starting communication in the IoT system. Besides, Alam et al. (2018) proposed a model to avoid attacks and secure IoT devices using Neural Network (NN) and ElGamal algorithm. Here private and public keys were used to control its cryptosystem. Manipulated data have been segmented into groups and then compared with the training data. In addition, a novel defense strategy for detecting and filtering poisonous data collected to train an arbitrary supervised learning model has been presented in Baracaldo et al. (2018).

4.2.2. Network layer

While attack becomes a normal phenomenon, securing network layers becomes a challenge that connects real life to the virtual world. Accordingly, different supervised ML algorithms like SVM, NN, and K-NN are being used to detect the intrusion attack (Buczak and Guven, 2015; Branch et al., 2013; Diro and Chilamkurti, 2017). In one study, NN was used to detect DoS attacks in IoT networks by adopting the multilayer perception based control system (Kulkarni and Venayagamoorthy, 2009). Saied et al. (2016) proposed a model for DDoS attack detection using an ANN algorithm. In the proposed scheme, only real information packets have permission to transmit through the network instead of fake ones. ANN performed better in detecting DDoS attack only if it was trained with updated data sets. Yu et al. (2008) research has experimentally showed that SVM based ML method in IoT system was capable of getting a high number of attack detection rate (99.4%).

Miettinen et al. (2017) presented an IoT SENTINEL model in which the classifier categorizes the IoT devices using RF algorithm to secure it from any unprotected device connection and avoid damage. Meidan et al. (Meidanet al., 2017b) used ML classifier algorithms for the identification of IoT devices. Considering various attributes, ML techniques classify the devices according to the connection with the IoT network into two categories (i.e., IoT devices and non-IoT devices). Then, the classifier controls the access of non- IoT devices and prevents possible attacks. A previous study (Lee et al., 2017) investigated the abnormal behaviour of IoT devices and the impact of detection accuracy on ML algorithms (i.e., SVM and k-means) with the partial change of training data sets. A decrement was noticed in accuracy rate for ML techniques and therefore, identification in the variation of accuracy and training data set can be a potential research topic.

An intrusion detection scheme was proposed by (Nobakht et al., 2016) at the network layer using ML algorithms for security of IoT devices. Recall, accuracy and precision matrices were used here to evaluate the classifier's performance due to the unbalanced data set. On the other hand, the area under the receiver operating characteristic curve (AUC) can be used as performance matrices for better results (Bradley, 1997; He and Garcia, 2009). Along the same direction, ANN techniques were used in Suarez et al. (2016) to train the machines to detect anomalies in IoT systems. Though the authors found good results from experiments, there is still a scope of further investigations to observe performance with larger data sets in which more data are tampered with attacks. Using unsupervised ML methods, Deng et al. (2018) integrate c-means clustering with PCA and propose an IDS with better detection rate for IoT. In another study, unsupervised ML algorithm (i.e., Optimum-path forest) was also used to develop an intrusion detection framework for the IoT network (Rocha et al., 2009; Dean and Ghemawat, 2008).

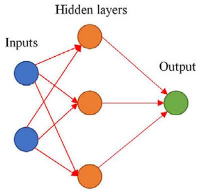
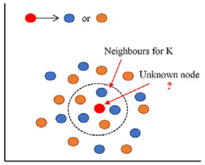
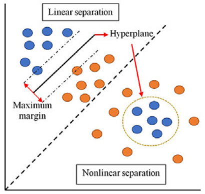
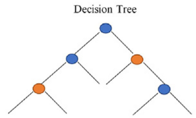
In 2018, Doshi with his colleagues in Doshi et al. (2018) presented a way to detect DDoS attacks in local IoT devices using low-cost machine learning algorithms and flow-based and protocol-agnostic traffic data. In this proposed model, some limited behaviors of IoT network such as calculation the endpoints and time taken to travel from one packet to another (time intervals between packets) have been considered. They compared a variety of classifiers for attack detection, including KNN, KDTree algorithm, SVM with the linear kernel (LSVM), DT using Gini impurity scores, RF using Gini impurity scores, NN. It was reported that the proposed techniques can identify DDoS attacks in local IoT devices using home gateway routers and other network middle boxes. The accuracy of the test set for five algorithms is higher than 0.99.

4.2.3. Web/application layer

K-NN, RF, Q-learning, Dyna-Q- based ML methods have been widely used to secure IoT devices from web/application based attacks, especially for malware detection (Xiao et al., 2017). Andrea et al. (2015) used supervised ML techniques (both K-NN and RF) to detect malware attacks and reported that RF methods with data set of MalGenome give better detection rate than K-NN. In another research, Q-learning shows better performance in terms of detecting latency and accuracy than Dyna-Q-based detection learning method (Xiao et al., 2017).

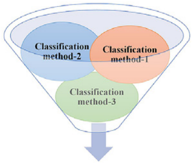
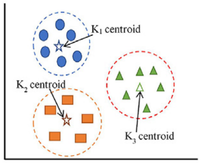
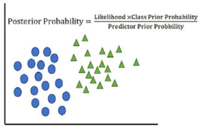
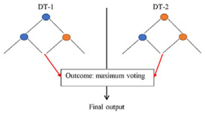
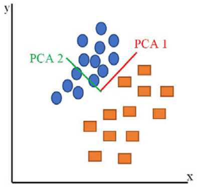
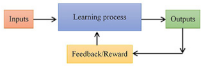
Table 1 presents a list of ML techniques used in different applications to detect attacks as a different layers wise solution of security of IoT.

Table 1
ML based solutions for securing IoT system.

ML Method	Illustration	Applications/Attack Detections	Protected Layer	Acc.(%)	Ref.
NN		Security of IoT Networks		99	Altaf et al. (2019)
DoS			Kulkarni and Venayagamoorthy (2009)		
Intrusion/Malware Detection			Buczak and Guven (2015) and Sedjelmaci et al. (2016)		
Privacy of an IoT element			Jeong et al. (2017)		
Security of Mobile Networks			Do et al. (2016)		
KNN		Intrusion/Malware Detection			Branch et al. (2013) and Narudin et al. (2016)
Detection of Intrusions and Anomalies, False Data Injection Attacks, Impersonation Attacks	Application Network		Karimipour and Dinavahi (2017b) and Aminanto and Kim (2017)		
Authentication of an IoT Element		80	Baldini et al. (2017)		
SVM	 Linear separation Hyperplane Maximum margin Nonlinear separation	Intrusion Detection		97.23	Sedjelmaci et al. (2016) and Bamakan et al. (2016)
		99–99.7	Kabir et al. (2018) and Wang et al. (2017b)		
		90–92	Zarpelã et al. (2017) and Zissis (2017)		
Security of Mobile Networks			Do et al. (2016)		
False Data Injection Attacks, Authentication, Data Tampering, Abnormal Behaviour	Application Network, Perceptino		Karimipour and Dinavahi (2017b) and Nobakht et al. (2016)		
DT		Detection of Intrusion and Suspicious Traffic Sources			Goeschel (2016)

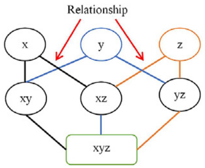
(*continued on next page*)

Table 1 *(continued)*

ML Method	Illustration	Applications/Attack Detections	Protected Layer	Acc.(%)	Ref.
Intrusion Detection		50–78	Stroeh et al. (2013)		
EL	 Classification method-2 Classification method-1 Classification method-3	Intrusion/Malware Detection, False Data Injection Attacks, Authentication, Data Tempering	Application Network, Perceptino		Karimipour and Dinavahi (2017b), Lee et al. (2017) and Nobakht et al. (2016)
K-means		Sybil Detection in Industrial WSNs and Private Data Anonymization in an IoT System, Data Tampering, Abnormal Behaviour	Network		Lee et al. (2017)
Intrusion detection			Rathore and Jha (2013)		
Network attack detection		80.19	IoT Analytics		
NB	 Posterior Probability = Likelihood ×Class Prior Probability / Predictor Prior Probbility	Intrusion Detection		50–78	Stroeh et al. (2013) and Usama et al. (2017)
Anomaly Detection			Mehmood and Rais (2016)		
Security of an IoT Element			Jincy and Sundararajan (2015)		
Traffic Engineering		80–90	Hogan and Esposito (2017)		
RF		Intrusion/Malware Detection		99.67	Narudin et al. (2016) and Farnaaz and Jabbar (2016)
Anomalies, DDoS, and Unauthorized IoT Devices	Network		Miettinen et al. (2017) and Zarpelã et al. (2017)		
PCA		Real-Time Detection System, Intrusion Detection	Network		Deng et al. (2018)
RL		DoS			Li et al. (2016)

(*continued on next page*)

Table 1 *(continued)*

ML Method	Illustration	Applications/Attack Detections	Protected Layer	Acc.(%)	Ref.
Spoofing			Xiao et al. (2016a)		
Eavesdropping			Xiao et al. (2016b)		
Jamming			Han et al. (2017)		
Malware Detection			Xiao et al. (2017)		
AR		Intrusion Detection			Tajbakhsh et al. (2009)

5. Research challenges

Currently, the field of IoT and its significance has been reaching at every doorste. Also, the security of IoT has been gaining attention from various networks and application researchers. The application of IoT, its usage, and impact on networks define different challenges and limitations that open new research directions in the future. In order to establish a secured and reliable IoT system, these probable challenges must be addressed. A list of possible challenges and future research fields have been presented based on research that has been conducted so far as well as future predictions in IoT network. In this section, possible research challenges have been presented as follows:

1) *Data Security*: Any learning algorithm needs a clear and reliable data sample based on what that method can be trained to secure the system. Learning techniques usually observe various attributes of the available data sets and use them to prepare training data sets. In that case, the availability of data, data quality, and data authentication play a vital role to train the data set of the learning methods. Unlike other learning techniques, machine learning also needs large, high quality, and available training data sets to develop an accurate ML technique. If a training data set contains low-quality data which carries noise can interrupt the deploy of a comprehensive and precise learning method. Therefore, authentication of the training data sets is an important challenge in ML techniques for effective security of the IoT network (Nweke et al., 2018).

In order to properly implement ML algorithms in IoT system, sufficient data sets are required which are often very difficult to gather based on if the system can identify threats and take necessary actions. In this context, data augmentation is a considerable approach to generate enough data set based on the existing real data. However, the challenge exists where the produced new data samples must properly be distributed in a different class in order to attain maximum accuracy from ML algorithms (Nweke et al., 2018).

Besides, an exact identification of any attack is another big issue in the security of IoT in order to properly distinguish good from bad state of IoT network. The challenge is if any intruder knows the attack type and has the ability to manipulate the training data set that is used for ML techniques, then it becomes easy for the attackers to modify their attack types and its effects on the network. Therefore, identifying different kinds of attacks and the probability of their occurrence in the network is a critical future research field in IoT.

2) *Infrastructure Problem*: When Vender (software-programmer) launches the software, they do not know the weakness of their product which paves a way for the attackers to investigate the infrastructure and hack the system through the software. This type of attack is alarming, and known as zero-day attack, which is very complicated to predetermine with traditional security techniques. Therefore, a strong software infrastructure needs to be developed for the proper security of IoT system. Security must be embedded in every stage in the IoT system starting from hardware to software which will ensure a vulnerable free environment in the overall system.

3) *Computational Restriction and Exploitation of Algorithms*: To compile any advanced machine learning algorithm is always challenging because it consumes a large memory and additional energy during processing extensive IoT systems. IoT devices deal with large data sets and with limited resources. Also, if ML methods are incorporated with the IoT system, then they will create more computational complexity for the system. Therefore, there is a need to minimize this complexity using machine learning techniques.

ML methods have been considered for cryptanalysis by attackers which is a potential threat for the IoT system. Though it is usually hard to break the system's cryptography, advanced ML algorithms, such as SVM and RF are implemented to break strong cryptographic system (Lerman et al., 2015).

4) *Privacy Leakage*: The most common issue in IoT nowadays is privacy. People use smart devices to exchange their data and information for various purposes. Slowly, the information of the clients is being collected and shared which is unknown to the clients. The users are unaware of what, how and where are their private information has been shared. All IoT devices have basic security protocols such as authentication, encryption and security updates. Therefore, IoT devices require message encryption before sending over the cloud to keep them secret. However, privacy protection must be a security concern in the IoT device design criteria. o illustrate, Google home assistance (Google home speaker and Chromecast have leaked a user's location. Thus, as IoT devices carry confidential and sensitive information/data of the users, there is a possibility for it to be misused if its leaked.

5) *Real-Time Update Issue*: As IoT devices are increasing rapidly, updating IoT devices' software, firmware update needs to be observed properly. But it is challenging to keep track and apply updates to millions of IoT devices while all devices are not supportive of air update. In that case, applying manual updates is required, such as is real time and data consuming which is cumbersome for users sometimes. Therefore, the term life long learning concept has been introduced to help machines continuously search for updates and makes their firewall strong for updated threats.

Due to the dynamic nature of IoT systems, every day new applications and electronic devices are connected to the network which results in unknown new attacks. Therefore, this is a challenge of IoT security to adopt an intelligent and real-time updated machine learning algorithm to detect unknown attacks (Suthaharan, 2014; Chen et al., 2018).

6. Analysis on published articles on ML-based IoT security

Literature shows that ML has been incorporated with IoT since 2002 (Samie et al., 2018). Therefore, probable research statistics on ML in IoT, ML in the security of IoT, and review on ML in the security of IoT has been presented in Fig. 9 based on the search engine like Elsevier, IEEE, Springer, Wiley, Hindawi, MDPI, Arxiv, and Taylor & Francis by sorting out to cross-check the title, abstract, and keywords from journal and conference papers. Authors tried their best to incorporate all possible related articles and in this regard, authors manually checked the titles and keywords especially to short out the articles. Fig. 9 illustrates that the rate of publication in all cases increases exponentially. Moreover, the publication in ML-based security of IoT starts in 2016 and the growth of publication is very fast which indicates that there is a huge potential of doing research in this field.

Fig. 10 presents statistical results on different ML algorithms based publication in IoT security up to March 2019 which is still increasing

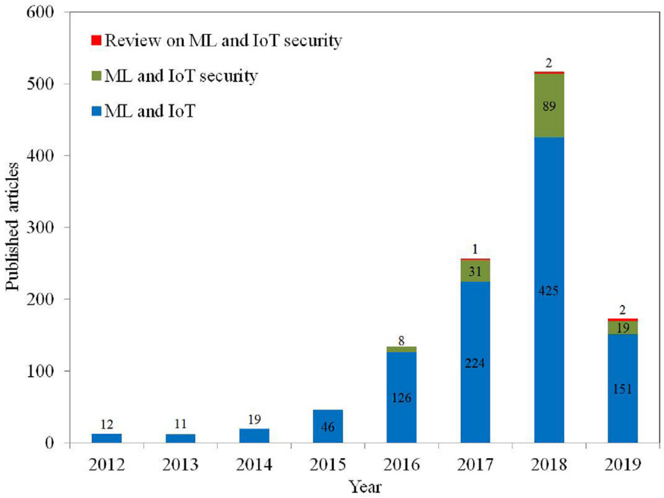


Fig. 9. A Statistic on paper published on ML and IoT, ML and security of IoT, and survey on ML and security of IoT till March 2019.

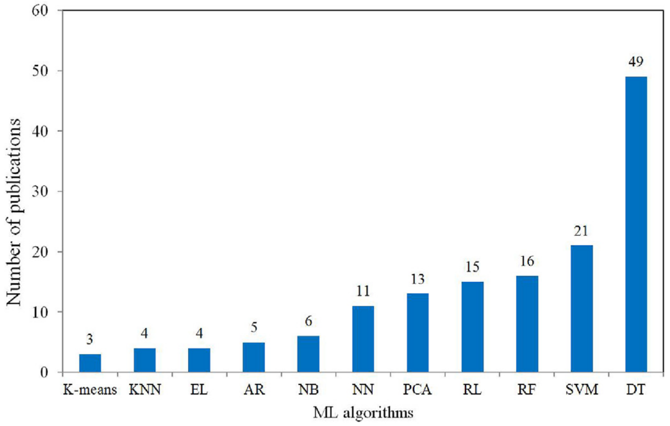


Fig. 10. A Statistic on paper published on various ML algorithms used in security of IoT until March 2019.

with time. It is found that DT was mostly used (32%) in the security of IoT compared to other learning methods. In addition, these statistics help direct the work of future researchers in potential fields.

7. Conclusion

Internet of Things (IoT) have the ability to change the future and bring global things into our hand. As a result, anyone can access, connect, and store their information in the network from anywhere using the blessing of smart services of IoT. Although, the empowerment of IoT connects our lives with the virtual world through smart devices to make life easy, comfortable, and smooth, security becomes a great concern in IoT system to care for its services. Therefore, to enhance the security with time and growing popularity, challenges and security of IoT has become a promising research in this field which must be addressed with novel solutions and exciting strategic plans for uncertain attacks in upcoming years. In this paper, a state of the art comprehensive literature review has been presented on ML-based security of IoT that includes IoT and its architecture, a thorough study on different types of security attacks, attack surfaces with effects, various categories of ML-based algorithms, and ML-based security solutions. In addition, research challenges have been demonstrated. Comparing with other review papers, this literature survey includes all papers on IoT and ML-based security of IoT up to 2019. During 2018, there was a huge acceleration in research on security of IoT. This literature review has focused on ML embedded algorithms on security of IoT from where anyone can get a general idea about different potential IoT attacks and their surface wise effects. Also, ML algorithms have been discussed with possible challenges that can aid future researchers to fix their ultimate goals and fulfill their aim in this field.

CRediT authorship contribution statement

Syeda Manjia Tahsien: Conceptualization, Methodology, Writing - original draft. **Hadis Karimipour:** Visualization, Writing - review & editing, Investigation, Supervision. **Petros Spachos:** Writing - review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abane, A., Daoui, M., Bouzefrane, S., Muhlethaler, P., 2019. A Lightweight forwarding strategy for named data networking in low-end IoT. J. Netw. Comput. Appl. 148, 1–12.

Abomhara, M., 2015. Cyber security and the internet of things: vulnerabilities, threats, intruders and attacks. J. Cyber Secur. Mobil. 4 (1), 65–88.

Aburomman, A.A., Reaz, M.B.I., 2016. A novel SVMkNNPSO ensemble method for intrusion detection system. Appl. Soft Comput. 38, 360–372.

Adetunmbi, A.O., Falaki, S.O., Adewale, O.S., Alese, B.K., 2008. Network intrusion detection based on rough set and knearest neighbour. Int. J. Comput. Intell. ICT Res. 2 (1), 60–66.

Agrawal, S., Agrawal, J., 2015. Survey on anomaly detection using data mining techniques. Procedia Comput. Sci. 60, 708–713.

Agrawal, R., Imieliski, T., Swami, A., 1993. Mining association rules between sets of items in large databases. Acm Sigmod Record 22 (2), 207–216 ACM.

Ahmed, A.I.A., Ab Hamid, S.H., Gani, A., Khan, S., Khan, M.K., 2019. Trust and reputation for Internet of Things: fundamentals, taxonomy, and open research challenges. J. Netw. Comput. Appl. 145, 1–13.

Ahmed, E., et al., 2017. The role of big data analytics in Internet of Things. Comput. Network. 129, 459–471.

Alaa, M., Zaidan, A.A., Zaidan, B.B., Talal, M., Kiah, M.L.M., 2017. A review of smart home applications based on Internet of Things. J. Netw. Comput. Appl. 97, 48–65.

Alaba, F.A., Othman, M., Hashem, I.A.T., Alotaibi, F., 2017. Internet of Things security: a survey. J. Netw. Comput. Appl. 88, 10–28.

Alam, F., Mehmood, R., Katib, I., Albogami, N.N., Albeshri, A., 2017. Data fusion and IoT for smart ubiquitous environments: a survey. IEEE Access 5, 9533–9554.

Alam, M.S., Husain, D., Naqvi, S.K., Kumar, P., 2018. IOT security through Machine Learning and homographic encryption technique. In: International Conference on New Trends in Engineering & Technology. ICNTET, Chennai.

Alsheikh, M.A., Lin, S., Niyato, D., Tan, H.-P., 2014. Machine learning in wireless sensor networks: algorithms, strategies, and applications. IEEE Commun. Surv. Tutor. 16 (4), 1996–2018.

Altaf, A., Abbas, H., Iqbal, F., Derhab, A., 2019. Trust models of Internet of smart things: a survey, open issues, and future directions. J. Netw. Comput. Appl. 137, 93–111.

AlTawy, R., Youssef, A.M., 2016. Security tradeoffs in cyber physical systems: a case study survey on implantable medical devices. IEEE Access 4, 959–979.

Amendola, S., Lodato, R., Manzari, S., Occhiuzzi, C., Marrocco, G., 2014. RFID technology for IoT-based personal healthcare in smart spaces. IEEE Internet Things J. 1 (2), 144–152.

Aminanto, M.E., Kim, K., 2017. Improving detection of WiFi impersonation by fully unsupervised deep learning. In: Information Security Applications: 18th International Workshop, WISA 2017.

Andrea, I., Chrysostomou, C., Hadjichristofi, G., Feb. 2015. Internet of things: security vulnerabilities and challenges. In: Proc. IEEE Symp. Computers and Communication, Larnaca, Cyprus, pp. 180–187.

Armbrust, M., et al., 2010. A view of cloud computing. Commun. ACM 53 (4), 50–58.

Asghari, P., Rahmani, A.M., Seyyed Javadi, H.H., 2018. Service composition approaches in IoT: a systematic review. J. Netw. Comput. Appl. 120, 61–77.

Ashton, K., 2011. That internet of things thing. RFID J. 22 (7), 1.

Atzori, L., Iera, A., Morabito, G., 2010. The internet of things: a survey. Comput. Network. 54 (15), 2787–2805.

Atzori, L., Iera, A., Morabito, G., 2014. From smart objects to social objects: the next evolutionary step of the internet of things. IEEE Commun. Mag. 52 (1), 97–105.

Baldini, G., Giuliani, R., Steri, G., Neisse, R., 2017. Physical layer authentication of internet of things wireless devices through permutation and dispersion entropy. In: 2017 Global Internet of Things Summit. GIoTS, Geneva, pp. 1–6.

Bamakan, S.M.H., Wang, H., Yingjie, T., Shi, Y., 2016. An effective intrusion detection framework based on mclp/svm optimized by timevarying chaos particle swarm optimization. Neurocomputing 199, 90–102.

Baracaldo, N., Chen, B., Ludwig, H., Safavi, A., Zhang, R., 2018. Detecting poisoning attacks on machine learning in IoT environments. In: 2018 IEEE International Congress on Internet of Things. (ICIOT), San Francisco, CA, pp. 57–64, https://doi.org/10.1109/ICIOT.2018.00015.

Bekara, C., 2014. Security issues and challenges for the IoTbased smart grid. Procedia Comput. Sci. 34, 532–537.

Benkhelifa, E., Welsh, T., Hamouda, W., Fourthquarter 2018. A critical review of practices and challenges in intrusion detection systems for IoT: toward universal and resilient systems. IEEE Commun. Surv. Tutor. 20 (4), 3496–3509, https://doi.org/10.1109/COMST.2018.2844742.

Bertino, E., Islam, N., 2017. Botnets and internet of things security. Computer 50 (2), 76–79.

Bhuyan, M.H., Bhattacharyya, D.K., Kalita, J.K., 2014. Network anomaly detection: methods, systems and tools. IEEE Commun. Surv. Tutor. 16 (1), 303–336.

Bkassiny, M., Li, Y., Jayaweera, S.K., 2013. A survey on machinelearning techniques in cognitive radios. IEEE Commun. Surv. Tutor. 15 (3), 1136–1159.

Bose, T., Bandyopadhyay, S., Ukil, A., Bhattacharyya, A., Pal, A., 2015. Why not keep your personal data secure yet private in IoT?: our lightweight approach. In: Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2015 IEEE Tenth International Conference on. IEEE, pp. 1–6.

Box, G.E., Tiao, G.C., 2011. Bayesian Inference in Statistical Analysis. John Wiley & Sons.

Bradley, A.P., 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recogn. 30 (7), 1145–1159.

Branch, J.W., Giannella, C., Szymanski, B., Wolff, R., Kargupta, H., Jan. 2013. Innetwork outlier detection in wireless sensor networks. Knowl. Inf. Syst. 34 (1), 23–54.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Buczak, A.L., Guven, E., 2015. A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Commun. Surv. Tutor. 18 (2), 1153–1176.

Butun, I., Morgera, S.D., Sankar, R., 2014. A survey of intrusion detection systems in wireless sensor networks. IEEE Commun. Surv. Tutor. 16 (1), 266–282.

Camara, C., Peris-Lopez, P., Tapiador, J.E., 2015. Security and privacy issues in implantable medical devices: a comprehensive survey. J. Biomed. Inf. 55, 272–289.

Chaabouni, N., Mosbah, M., Zemmari, A., Sauvignac, C., Faruki, P., 2019. Network intrusion detection for IoT security based on learning techniques. In: IEEE Communications Surveys & Tutorials.

Chang, Y., Li, W., Yang, Z., 2017. Network intrusion detection based on random forest and support vector machine. In: Computational Science and Engineering (CSE) and Embedded and Ubiquitous Computing (EUC), 2017 IEEE International Conference on, vol. 1. IEEE, pp. 635–638.

Chen, F., Deng, P., Wan, J., Zhang, D., Vasilakos, A., Rong, X., 2015. Data mining for the internet of things: literature review and challenges. Int. J. Distributed Sens. Netw. 11, 431047.

Chen, Z., Ma, N., Liu, B., 2018. Lifelong Learning for Sentiment Classification. arXiv:1801.02808.

Cui, L., Yang, S., Chen, F., Ming, Z., Lu, N., Qin, J., 2018. A survey on application of machine learning for Internet of Things. Int. J. Mach. Learn. Cybern. 9 (8), 1399–1417.

Cutler, D.R., et al., 2007. Random forests for classification in ecology. Ecology 88 (11), 2783–2792.

da Cruz, M.A.A., Rodrigues, J.J.P.C., Sangaiah, A.K., Al-Muhtadi, J., Korotaev, V., 2018. Performance evaluation of IoT middleware. J. Netw. Comput. Appl. 109, 53–65.

Daz, M., Martn, C., Rubio, B., 2016. Stateoftheart, challenges, and open issues in the integration of Internet of things and cloud computing. J. Netw. Comput. Appl. 67, 99–117.

Dean, J., Ghemawat, S., 2008. MapReduce: simplified data processing on large clusters. Commun. ACM 51 (1), 107–113.

Deng, Z., Zhu, X., Cheng, D., Zong, M., Zhang, S., 2016. Efficient kNN classification algorithm for big data. Neurocomputing 195, 143–148.

Deng, L., Li, D., Yao, X., Cox, D., Wang, H., 2018. Mobile Network Intrusion Detection for IoT System Based on Transfer Learning Algorithm. Cluster Computing, pp. 1–16.

Diro, A.A., Chilamkurti, N., 2017. Distributed attack detection scheme using deep learning approach for Internet of Things. Future Generat. Comput. Syst..

Do, V.T., Engelstad, P., Feng, B., Do, T.V., 2016. Strengthening mobile network security using machine learning. In: Younas, M., Awan, I., Kryvinska, N., Strauss, C., Thanh, D.V. (Eds.), Mobile Web and Intelligent Information Systems. Springer International Publishing, Cham, pp. 173–183.

Doshi, R., Apthorpe, N., Feamster, N., 2018. Machine Learning DDoS Detection for Consumer Internet of Things Devices. arXiv:1804.04159.

Elazhary, H., 2019. Internet of Things (IoT), mobile cloud, cloudlet, mobile IoT, IoT cloud, fog, mobile edge, and edge emerging computing paradigms: disambiguation and research directions. J. Netw. Comput. Appl. 128, 105–140.

Farnaaz, N., Jabbar, M.A., 2016. Random forest modeling for network intrusion detection system. Procedia Comput. Sci. 89 (Suppl. C), 213–217.

Farris, I., Taleb, T., Khettab, Y., Song, J., Firstquarter 2019. A survey on emerging SDN and NFV security mechanisms for IoT systems. IEEE Commun. Surv. Tutor. 21 (1), 812–837, https://doi.org/10.1109/COMST.2018.2862350.

Faruki, P., et al., 2015. Android security: a survey of issues, malware penetration, and defenses. IEEE Commun. Surv. Tutor. 17 (2), 998–1022.

Fremdt, S., Beck, R., Weber, S., 2013. Does cloud computing matter? An analysis of the cloud model softwareasa service and its impact on operational agility. In: System Sciences (HICSS), 2013 46th Hawaii International Conference on. IEEE, pp. 1025–1034.

Gazis, V., Firstquarter 2017. A survey of standards for machine-to-machine and the internet of things. IEEE Commun. Surv. Tutor. 19 (1), 482–511, https://doi.org/10.1109/COMST.2016.2592948.

Gil, D., Ferrnandez, A., Mora-Mora, H., Peral, J., 2016. Internet of things: a review of surveys based on context aware intelligent services. Sensors 16 (7), 1069.

Goeschel, K., 2016. Reducing false positives in intrusion detection systems using datamining techniques utilizing support vector machines, decision trees, and naive Bayes for offline analysis. In: SoutheastCon, 2016. IEEE, pp. 1–6.

Gondhi, N.K., Gupta, A., 2017. Survey on Machine Learning Based Scheduling in Cloud Computing, pp. 57–61.

Guan, Zh., Zhang, Y., Wu, L., Wu, J., Li, J., Ma, Y., Hu, J., 2019. APPA: An anonymous and privacy preserving data aggregation scheme for fog-enhanced IoT. J. Netw. Comput. Appl. 125, 82–92.

HaddadPajouh, H., et al., 2019. A Survey on Internet of Things Security: Requirements, Challenges, and Solutions. Internet of Things.

Ham, H.-S., Kim, H.-H., Kim, M.-S., Choi, M.-J., 2014. Linear SVMbased android malware detection for reliable IoT services. J. Appl. Math. 2014.

Han, G., Xiao, L., Poor, H.V., Mar. 2017. Twodimensional antijamming communication based on deep reinforcement learning. In: Proc. IEEE Int. Conf. Acoustics Speech and Signal Processing, pp. 2087–2091 New Orleans, LA.

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: a kmeans clustering algorithm. J. Royal Stat. Soc. Ser. C (Appl. Stat.) 28 (1), 100–108.

He, H., Garcia, E.A., 2009. Learning from imbalanced data. IEEE Trans. Knowl. Data Eng. 21 (9), 1263–1284.

Hogan, M., Esposito, F., 2017. Stochastic delay forecasts for edge trafc engineering via bayesian networks. In: IEEE International Symposium on Network Computing and Applications. IEEE, Cambridge, MA, pp. 1–4.

Huang, J., Zhang, X., Tan, L., Wang, P., Liang, B., 2014. Asdroid: detecting stealthy behaviors in android applications by user interface and program behaviour contradiction. In: Proceedings of the 36th International Conference on Software Engineering. ACM, pp. 1036–1046.

Hush, D.R., Horne, B.G., Jan. 1993. Progress in supervised neural networks. IEEE Signal Process. Mag. 10 (1), 8–39, https://doi.org/10.1109/79.180705.

IoT Analytics, Why the internet of things is called internet of things: defnition, history, disambiguation. https://iotanalytics.com/internetofthingsdefnition/.

Jain, A.K., 2010. Data clustering: 50 years beyond Kmeans. Pattern Recogn. Lett. 31 (8), 651–666.

Jeong, H.J., Lee, H.J., Moon, S.M., 2017. Workinprogress: cloudbased machine learning for iot devices with better privacy. In: 2017 International Conference on Embedded Software. EMSOFT, Seoul, pp. 1–2.

Jincy, V.J., Sundararajan, S., 2015. Classification mechanism for iot devices towards creating a security framework. In: Buyya, R., Thampi, S.M. (Eds.), Intelligent Distributed Computing. Springer International Publishing, Cham, pp. 265–277.

Jing, Q., Vasilakos, A.V., Wan, J., Lu, J., Qiu, D., 2014. Security of the internet of things: perspectives and challenges. Wireless Network 20 (8), 2481–2501.

Jordan, M.I., Mitchell, T.M., 2015. Machine learning: trends, perspectives, and prospects. Science 349 (6245), 255–260.

Juniper Research, 2015. Internet of things connected devices to almost triple to over 38 billion units by 2020. http://www.juniperresearch.com/press/press-releases/iot-connecteddevices-to-triple-to-38-bn-by-2020.

Kabir, E., Hu, J., Wang, H., Zhuo, G., 2018. A novel statistical technique for intrusion detection systems. Future Generat. Comput. Syst. 79, 303–318.

Karimipour, H., Dinavahi, V., Dec. 2017a. Robust massively parallel dynamic state estimation of power systems against cyberattack. IEEE Access 6, 2984–2995.

Karimipour, H., Dinavahi, V., 2017b. On false data injection attack against dynamic state estimation on smart power grids. In: IEEE Int. Conf. on Smart Energy Grid Engineering, pp. 1–7.

Karimipour, H., Dehghantanha, A., Parizi, R.M., Choo, R., Leung, H., 2019a. A Deep and scalable unsupervised machine learning system for cyberattack detection in largescale smart grids. IEEE Access 7, 80778–80788.

Karimipour, H., Geris, S., Dehghantanha, A., Leung, H., 2019b. Intelligent anomaly detection for largescale smart grids. In: IEEE CCECE, pp. 1–4.

Kim, G., Lee, S., Kim, S., 2014. A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Syst. Appl. 41 (4), 1690–1700.

Kiran, B.N., Radheshyam, S.G., Sagar, N., Balthar, S.A., Shrinath, 2018. SECURITY FOR IoT SYSTEMS USING MACHINE LEARNING. Int. J. Adv. Res. Innovat. Ideas Educ. (IJRIIE) 4 (2), 2707–2710.

Kotsiantis, S.B., 2013. Decision trees: a recent overview. Artif. Intell. Rev. 39 (4), 261–283.

Kotsiantis, S., Kanellopoulos, D., 2006. Association rules mining: a recent overview. GESTS Int. Trans. Comput. Sci. Eng. 32 (1), 71–82.

Kotsiantis, S.B., Zaharakis, I., Pintelas, P., 2007. Supervised machine learning: a review of classification techniques. Emerg. Artif. Intell. Appl. Comput. Eng. 160, 3–24.

Kulkarni, R.V., Venayagamoorthy, G.K., 2009. Neural network based secure media access control protocol for wireless sensor networks. In: Proc. Int. Joint Conf. Neural Networks. June, Atlanta, GA, pp. 3437–3444.

Lee, S.Y., Wi, S.-r., Seo, E., Jung, J.-K., Chung, T.-M., 2017. ProFiOt: abnormal Behavior Profiling (ABP) of IoT devices based on a machine learning approach. In: Telecommunication Networks and Applications Conference (ITNAC), 2017 27th International. IEEE, pp. 1–6.

Lerman, L., Bontempi, G., Markowitch, O., 2015. A machine learning approach against a masked AES. J. Cryptogr. Eng. 5 (2), 123–139.

Li, X., Lu, R., Liang, X., Shen, X., Nov. 2011. Smart community: an Internet of things application. IEEE Commun. Mag. 49 (11), 68–75.

Li, Y., Quevedo, D.E., Dey, S., Shi, L., Apr. 2016. SINRbased DoS attack on remote state estimation: a gametheoretic approach. IEEE Trans. Contr. Netw. Syst. 4 (3), 632–642.

Liu, Y., Pi, D., 2017. A novel kernel SVM algorithm with game theory for network intrusion detection. KSII Trans. Internet Inf. Syst. 11 (8).

Liu, Y., Cheng, C., Gu, T., Jiang, T., Li, X., 2016. A lightweight authenticated communication scheme for smart grid. IEEE Sensor. J. 16 (3), 836–842.

Makhdoom, I., Abolhasan, M., Lipman, J., Liu, R.P., Ni, W., 2018. Anatomy of threats to the internet of things. IEEE Commun. Surv. Tutor., https://doi.org/10.1109/COMST.2018.2874978.

Malasri, K., Wang, L., 2009. Securing wireless implantable devices for healthcare: ideas and challenges. IEEE Commun. Mag. 47 (7).

Mamdouh, M., Elrukhsi, M.A.I., Khattab, A., 2018. Securing the internet of things and wireless sensor networks via machine learning: a survey. In: 2018 International Conference on Computer and Applications. ICCA 2018, pp. 215–218.

Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J., Aharon, D., 2015. Unlocking the Potential of the Internet of Things. http://tinyurl.com/hnlhz8v.

Mehmood, T., Rais, H.B.M., 2016. Machine learning algorithms in context of intrusion detection. In: 3rd International Conference on Computer and Information Sciences (ICCOINS). IEEE, https://doi.org/10.1109/iccoins.2016.7783243.

Meidan, Y., et al., 2017a. Detection of Unauthorized IoT Devices Using Machine Learning Techniques. arXiv:1709.04647.

Meidan, Y., et al., 2017b. ProfilIoT: a machine learning approach for IoT device identification based on network traffic analysis. In: Proceedings of the Symposium on Applied Computing. ACM, pp. 506–509.

Miettinen, M., Marchal, S., Hafeez, I., Asokan, N., Sadeghi, A.-R., Tarkoma, S., 2017. IoT Sentinel: automated device type identification for security enforcement in IoT. In: Distributed Computing Systems (ICDCS), 2017 IEEE 37th International Conference on. IEEE, pp. 2177–2184.

Mnih, V., et al., 2015. Humanlevel control through deep reinforcement learning. Nature 518 (7540), 529.

Modi, C., Patel, D., Borisaniya, B., Patel, H., Patel, A., Rajarajan, M., 2013. A survey of intrusion detection techniques in cloud. J. Netw. Comput. Appl. 36 (1), 42–57.

Modiri, E., Azmoodeh, A., Dehghantanha, A., Karimipour, H., Jan. 2018. Fuzzy pattern tree for edge attack detection and categorization in IoT. J. Syst. Architect. 1–15.

Mohammadi, S., Mirvaziri, H., Ahsaee, M.G., Karimipour, H., Feb. 2018a. Cyber intrusion detection by combined feature selection algorithm. J. Inf. Secur. Appl. 44, 80–88.

Mohammadi, S., Desai, V., Karimipour, H., 2018b. Multivariate mutual information feature selection for intrusion detection. In: IEEE Canada Electrical Power and Energy Conf. IEEE, pp. 1–6.

Moosavi, S.R., et al., 2015. SEA: a secure and efficient authentication and authorization architecture for IoTbased healthcare using smart gateways. Procedia Comput. Sci. 52, 452–459.

Mukherjee, S., Sharma, N., 2012. Intrusion detection using naive Bayes classifier with feature reduction. Procedia Technol. 4, 119–128.

Muniyandi, A.P., Rajeswari, R., Rajaram, R., 2012. Network anomaly detection by cascading kMeans clustering and C4. 5 decision tree algorithm. Procedia Eng. 30, 174–182.

Namvar, N., Saad, W., Bahadori, N., Kelley, B., 2016. Jamming in the Internet of Things: a gametheoretic perspective. In: Global Communications Conference (GLOBECOM), 2016 IEEE. IEEE, pp. 1–6.

Narudin, F.A., Feizollah, A., Anuar, N.B., Gani, A., Jan. 2016. Evaluation of machine learning classifiers for mobile malware detection. Soft Comput. 20 (1), 343–357.

Neely, S., Dobson, S., Nixon, P., 2006. Adaptive middleware for autonomic systems. In: Annals des tele-communications, vol. 61. Springer, pp. 1099–1118. 9-10.

Nitti, M., Atzori, L., Cvijikj, I.P., 2015. Friendship selection in the social internet of things: challenges and possible strategies. IEEE Internet Things J. 2 (3), 240–247.

Nobakht, M., Sivaraman, V., Boreli, R., 2016. A hostbased intrusion detection and mitigation framework for smart home IoT using OpenFlow. In: Availability, Reliability and Security (ARES), 2016 11th International Conference on. IEEE, pp. 147–156.

Nweke, H.F., Teh, Y.W., Al-garadi, M.A., Alo, U.R., 2018. Deep Learning Algorithms for Human Activity Recognition Using Mobile and Wearable Sensor Networks: State of the Art and Research Challenges. Expert Systems with Applications.

Panda, M., Patra, M.R., 2007. Network intrusion detection using naive bayes. Int. J. Comput. Sci. Netw. Secur. 7 (12), 258–263.

Quinlan, J.R., 1986. Induction of decision trees. Mach. Learn. 1 (1), 81–106.

Rathore, H., Jha, S., 2013. Bioinspired machine learning based wireless sensor network security. In: 2013 World Congress on Nature and Biologically Inspired Computing. IEEE, Fargo, ND, pp. 140–146.

Razzaque, M.A., Milojevic-Jevric, M., Palade, A., Clarke, S., 2016. Middleware for internet of things: a survey. IEEE Internet Things J. 3 (1), 70–95.

Restuccia, F., Daro, S., Melodia, T., Dec. 2018. Securing the internet of things in the age of machine learning and softwaredefined networking. IEEE Internet Things J. 5 (6), 4829–4842.

Rocha, L.M., Cappabianco, F.A., Falc, A.X., 2009. Data clustering as an optimumpath forest problem with applications in image analysis. Int. J. Imag. Syst. Technol. 19 (2), 50–68.

Roman, R., Zhou, J., Lopez, J., 2013. On the features and challenges of security and privacy in distributed internet of things. Comput. Network. 57 (10), 2266–2279.

S. O. Technologies, Mobile apps leveraging the internet of things (IoT). [Online]: https://www.spaceotechnologies.com/mobileappsleveragingtheinternetofthings/, (Accessed 15 January 2019).

Saadeh, M., Sleit, A., Sabri, K.E., Almobaideen, W., 2018. Hierarchical architecture and protocol for mobile object authentication in the context of IoT smart cities. J. Netw. Comput. Appl. 121, 1–19.

Saggi, M.K., Jain, S., 2018. A Survey towards an Integration of Big Data Analytics to Big Insights for Value-Creation. Information Processing & Management.

Saied, A., Overill, R.E., Radzik, T., 2016. Detection of known and unknown DDoS attacks using Artificial Neural Networks. Neurocomputing 172, 385–393.

Samie, F., Bauer, L., Henkel, J., Chen, Z., Ma, N., Liu, B., 2018. From cloud down to things: an overview of machine learning in internet of things. IEEE Internet Things J. (IoTJ) 1–14.

Sedjelmaci, H., Senouci, S.M., Al-Bahri, M., 2016. A lightweight anomaly detection technique for lowresource iot devices: a gametheoretic methodology. In: IEEE International Conference on Communications (ICC), pp. 1–6.

Sengupta, J., Ruj, S., Bita, S.D., Nov. 2019. A Comprehensive survey on attacks, security issues and blockchain solutions for IoT and IIoT. J. Netw. Comput. Appl. 1–50.

Sethi, P., Sarangi, S.R., 2017. Internet of things: architectures, protocols, and applications. J. Electr. Comput. Eng..

Sezer, O.B., Dogdu, E., Ozbayoglu, A.M., 2018. Context-Aware computing, learning, and big data in internet of things: a survey. IEEE Internet Things J. 5 (1), 1–27.

Sheng, Z., Yang, S., Yu, Y., Vasilakos, A., Dec. 2013. A survey on the IETF protocol suite for the Internet of things: standards, challenges, and opportunities. IEEE Wirel. Cmmun. 20 (6), 91–98.

Shi, C., Liu, J., Liu, H., Chen, Y., 2017. Smart user authentication through actuation of daily activities leveraging WiFienabled IoT. In: Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing. ACM, p. 5.

Singh, A., Payal, A., Bharti, S., 2019. A walkthrough of the emerging IoT paradigm: visualizing inside functionalities, key features, and open issues. J. Netw. Comput. Appl. 143, 111–151.

Spachos, P., Papapanagiotou, I., Plataniotis, K.N., Sept. 2018. Microlocation for smart buildings in the era of the internet of things: a survey of technologies, techniques, and approaches. IEEE Signal Process. Mag. 35 (5), 140–152.

Statista, 2019. Internet of things to hit the mainstream by 2020. https://www.statista.com/chart/2936/internet-of-things-to-hit-the-mainstream-by-2020/.

Statista, Technology & Telecommunication, 2019. Consumer Electronics, Source. IHS, https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/.

Stroeh, K., Mauro Madeira, E.R., Goldenstein, S.K., 2013. An approach to the correlation of security events based on machine learning techniques. J. Internet Serv. Appl. 4 (1), 7.

Suarez, J., Quevedo, J., Vidal, I., Corujo, D., Garcia-Reinoso, J., Aguiar, R., 2016. A secure IoT management architecture based on informationcentric networking. J. Netw. Comput. Appl. 63, 190–204.

Suthaharan, S., 2014. Big data classification: problems and challenges in network intrusion prediction with machine learning. Perform. Eval. Rev. 41 (4), 70–73.

Swarnkar, M., Hubballi, N., 2016. OCPAD: one class Nave Bayes classifier for payload based anomaly detection. Expert Syst. Appl. 64, 330–339.

Tajbakhsh, A., Rahmati, M., Mirzaei, A., 2009. Intrusion detection using fuzzy association rules. Appl. Soft Comput. 9 (2), 462–469.

Tan, Z., Jamdagni, A., He, X., Nanda, P., Liu, R.P., May 2013. A system for Denialof Service attack detection based on multivariate correlation analysis. IEEE Trans. Parallel Distr. Syst. 25 (2), 447–456.

Tsai, C.-F., Hsu, Y.-F., Lin, C.-Y., Lin, W.-Y., 2009. Intrusion detection by machine learning: a review. Expert Syst. Appl. 36 (10), 11994–12000.

Tsai, C.-W., Lai, C.-F., Chiang, M.-C., Yang, L.T., 2014. Data mining for internet of things: a survey. IEEE Commun. Surv. Tutor. 16 (1), 77–97.

Ukil, A., Bandyopadhyay, S., Pal, A., 2014. Iotprivacy: to be private or not to be private. In: Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on. IEEE, pp. 123–124.

Usama, M., Qadir, J., Raza, A., Arif, H., Yau, K.A., Elkhatib, Y., Hussain, A., Al-Fuqaha, A.I., 2017. Unsupervised Machine Learning for Networking: Techniques, Applications and Research Challenges. CoRR.. arXiv:1709.06599.

Wamba, S.F., Anand, A., Carter, L., 2013. A literature review of RFIDenabled healthcare applications and issues. Int. J. Inf. Manag. 33 (5), 875–891.

Wang, N., Jiang, T., Lv, S., Xiao, L., 2017a. Physicallayer authentication based on extreme learning machine. IEEE Commun. Lett. 21 (7), 1557–1560.

Wang, H., Gu, J., Wang, S., 2017b. An effective intrusion detection framework based on SVM with feature augmentation. Knowl. Base Syst. 130139.

Wold, S., Esbensen, K., Geladi, P., 1987. Principal component analysis. Chemometr. Intell. Lab. Syst. 2 (13), 37–52.

Xiao, L., Yan, Q., Lou, W., Chen, G., Hou, Y.T., Oct. 2013. Proximitybased security techniques for mobile users in wireless networks. IEEE Trans. Inf. Forensics Secur. 8 (12), 2089–2100.

Xiao, L., Li, Y., Han, G., Liu, G., Zhuang, W., Dec. 2016a. PHYlayer spoofing detection with reinforcement learning in wireless networks. IEEE Trans. Veh. Technol. 65 (12), 10037–10047.

Xiao, L., Xie, C., Chen, T., Dai, H., May 2016b. A mobile offloading game against smart attacks. IEEE Access 4, 2281–2291.

Xiao, L., Li, Y., Huang, X., Du, X.J., Oct. 2017. Cloudbased malware detection game for mobile devices with offloading. IEEE Trans. Mobile Comput. 16 (10), 2742–2750.

Xiao, L., Wan, X., Lu, X., Zhang, Y., Wu, D., 2018a. IoT security techniques based on machine learning: how do IoT devices use AI to enhance security? IEEE Signal Process. Mag. 35 (5), 41–49.

Xiao, L., Wan, X., Han, Z., Mar. 2018b. PHYlayer authentication with multiple landmarks with reduced overhead. IEEE Trans. Wireless Commun. 17 (3), 1676–1687.

Xie, M., Huang, M., Bai, Y., Hu, Z., 2017. The anonymization rotection algorithm based on fuzzy clustering for the ego of data in the Internet of Things. J. Electr. Comput. Eng. 2017.

Xu, X., Fu, S., Qi, L., Zhang, X., Liu, Q., He, Q., Li, S., 2018. An IoT-Oriented data placement method with privacy preservation in cloud environment. J. Netw. Comput. Appl. 124, 148–157.

Yao, X., Chen, Z., Tian, Y., 2015. A lightweight attributebased encryption scheme for the Internet of Things. Future Generat. Comput. Syst. 49, 104–112.

Yu, J., Lee, H., Kim, M.S., Park, D., Oct. 2008. Traffic flooding attack detection with SNMP MIB using SVM. Comput. Commun. 31 (17), 4212–4219.

Zarpel, B.B., Miani, R.S., Kawakani, C.T., de Alvarenga, S.C., 2017. A survey of intrusion detection in Internet of Things. J. Netw. Comput. Appl. 84, 25–37.

Zeadally, S., Tsikerdekis, M., 2020. Securing internet of things (IoT) with machine learning. Int. J. Commun. Syst. 1–16.

Zhao, S., Li, W., Zia, T., Zomaya, A.Y., 2017. A dimension reduction model and classifier for anomalybased intrusion detection in internet of things. In: Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence & Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech), 2017 IEEE 15th Intl. IEEE, pp. 836–843.

Zhou, J., Cao, Z., Dong, X., Vasilakos, A.V., Jan. 2017. Security and privacy for cloud-based IoT: challenges. IEEE Commun. Mag. 55 (1), 26–33.

Zhou, W., Zhang, Y., Liu, P., 2018. The Effect of IoT New Features on Security and Privacy: New Threats, Existing Solutions, and Challenges yet to Be Solved. arXiv: 1802.03110.

Zissis, D., 2017. Intelligent security on the edge of the cloud. In: International Conference on Engineering, Technology and Innovation. IEEE, Funchal. 1066 1070.

Syeda Manjia Tahsien received the B.Sc. degree in Electrical, Electronic, and Communication Engineering (EECE) with a major in Communication Engineering from the Military Institute of Science and Technology (MIST), Bangladesh in 2016. She is currently pursuing MASc degree in Engineering with collaborative specialization in AI. Her research interest includes Artificial Intelligence (AI), Manufacturing System Analysis, Optimization, and Internet of Things (IoT).

Hadis Karimipour received the Ph.D. degree in Energy System from the Department of Electrical and Computer Engineering in the University of Alberta in Feb. 2016. Before joining the University of Guelph, she was a postdoctoral fellow in University of Calgary working on cyber security of the smart power grids. She is currently an Assistant Professor at the School of Engineering, Engineering Systems and Computing Group at the University of Guelph, Guelph, Ontario. Her research interests include application of machine learning on security analysis, cyber-physical modeling, cyber-security of the smart grids, and parallel and distributed computing. She is member of IEEE and IEEE Computer Society. She serves as the Chair of the IEEE Women in Engineering (WIE) and chapter chair of IEEE Information Theory in Kitchener-Waterloo section.

Petros Spachos is an Assistant Professor with the School of Engineering, University of Guelph, Guelph, ON, Canada. His research interests include experimental wireless networking with a current focus on wireless sensor networks, smart cities, and the Internet of Things. He received the Ph.D. degree from the University of Toronto, Toronto, ON, Canada, and he is a Senior Member of IEEE. Contact him at: petros@uoguelph.ca