ZKG INTERNATIONAL

ISSN: 2366-1313

Securing the Internet of Things: A Study on Machine Learning-Based Solutions for IoT Security and Privacy Challenges

Aziz UllahKarimy[1] ,Dr. P Chandrasekhar Reddy[2]
[1, 2]Electronics and Communication Engineering, Jawaharlal Nehru Technological University Hyderabad-JNTUH, India.

Abstract

The Internet of Things (IoT) is a rapidly growing technology that connects and integrates billions of smart devices, generating vast volumes of data and impacting various aspects of daily life and industrial systems. However, the inherent characteristics of IoT devices, including limited battery life, universal connectivity, resource-constrained design, and mobility, make them highly vulnerable to cybersecurity attacks, which are increasing at an alarming rate. As a result, IoT security and privacy have gained significant research attention, with a particular focus on developing anomaly detection systems. In recent years, machine learning (ML) has made remarkable progress, evolving from a lab novelty to a powerful tool in critical applications. ML has been proposed as a promising solution for addressing IoT security and privacy challenges. In this article, we conducted a study of the existing security and privacy challenges in the IoT environment. Subsequently, we present the latest ML-based models and solutions to address these challenges, summarizing them in a table that highlights the key parameters of each proposed model. Additionally, we thoroughly studied available datasets related to IoT technology. Through this article, readers will gain a detailed understanding of IoT architecture, security attacks, and countermeasures using ML techniques, utilizing available datasets. We also discuss future research directions for ML-based IoT security and privacy. Our aim is to provide valuable insights into the current state of research in this field and contribute to the advancement of IoT security and privacy.

Keywords: Internet of Things (IoT), Machine Learning (ML), Cybersecurity, security, privacy, attacks

1. Introduction

The Internet of Things (IoT) is one of the emerging technologies that aims to simplify human lives. However, there are significant security and privacy concerns associated with this technology that can be exploited. IoT devices often operate in unattended environments, making them easily susceptible to manipulation, and communicate through wireless technologies, which makes them vulnerable to eavesdropping(1). The IoT is a network of interconnected devices that enable seamless data exchange between physical devices, such as wearable technology, autonomous cars, industrial robots, medical and healthcare equipment, smart TVs, and smart city infrastructures that can be monitored and controlled remotely(2). It is predicted that IoT devices, which have access to sensitive data like personal information and bank details, will become more prevalent than mobile devices(3). In fact, IoT systems today often span across cloud and fog/edge layers, composed of multiple interconnected devices, resulting in a significant attack surface area.

The unique features of IoT networks raise serious security and privacy concerns. Traditional computer security technologies are ineffective when applied to IoT networks due to the low computing capacities of IoT devices, limited power resources, communication

ISSN: 2366-1313

technologies, software vulnerabilities, and lengthy security software update cycles(3). While security-enhancing techniques such as encryption, certification, and authentication mechanisms like DTLS (Datagram Transport Layer Security) can significantly improve security, they may not completely protect against unauthorized access to IoT network devices(4). Therefore, ensuring the security of IoT networks is one of the primary objectives in the development of information security technology. All layers of the IoT application, including the hardware level where data is gathered, the network level where data is transmitted to the data processing center, and the cloud level/databases where data is stored, are potential targets for IoT attacks(5). Pooja Chaudhary et al. proposed an Intrusion Detection System (IDS) that uses a machine learning technique called Self-Organizing Map (SOM) to protect IoT devices from cross-site Scripting (XXS) attacks at the application layer. The performance of this approach was validated with real-time datasets and found to be effective(6). SemihCakir et al. conducted a study on the vulnerabilities of the Routing Protocol for Low-Power and Lossy Networks (RPL) at the network layer of IoT applications. They proposed a deep learning approach based on Gated Recurrent Unit (GRU) network model, considering energy consumption and power resources, to detect and prevent Hello Flooding (HF) attacks in this routing protocol(7).

The use of machine learning techniques is one of the popular modern approaches to identifying anomalies and identifying attacks in computer networks. In the current application of Machine Learning in IoT intrusion detection systems, there are certain problems with using machine learning (ML) approaches to detect IoT intrusion threats. To illustrate, the first forms of intrusion attacks that have been investigated are rather simple, and more sophisticated attacks have not been taken into account. The second procedure of processing a large volume of data is quite difficult; in order to identify useful features for training ML models, a vast number of features must be extracted, which uses a lot of resources. Therefore, a lightweight method is required to automatically extract a small number of features for the ML model to detect various numbers of attacks(8). In our earlier research(4)(9), feature selection issues were investigated and robust features were chosen by putting forth several kinds of techniques for traffic identification and attacks traffic detection. Similarly, in(5),(10) and(11), for accurate network traffic classification using ML algorithms, many feature selection techniques are given to address the issue of selecting features. But based on the previous study, we concluded that taking more feature sets is ineffective for correct identification using ML approaches and this might reduce the accuracy of ML classifiers and increase computing complexity. However, no efficient ML model has yet been established for the identification of IoT network cyberattack traffic. In order to propose a novel method that addresses this problem, it is essential to analyze the effective feature selection problem for anomaly and malicious traffic in the IoT network(12), in this paper we systematically studied the most recent approaches proposed by the researchers.

The main objectives of this research study are:

- Study existing proposed IoT architectures and identify layer-wise security and privacy challenges.
- Investigate attacks targeting each layer of IoT architectures.
- Study existing countermeasures using machine learning techniques, analyze challenges, and compare findings from relevant papers.
- Study available datasets for training machine learning-based models for securing IoT applications and compare significant points.
- Propose further research ideas to fill existing gaps.

ISSN: 2366-1313

From thousands of published papers in this research area, we have systematically studied 25 recently published papers from international journals (IEEE, Springer, MDPI, and ELSEVIER) in the years 2020, 2021, and 2022, which propose various machine learning-based techniques for securing IoT applications. Our study highlights the most recent machine-learning approaches for IoT security and their existing limitations. This paper is organized as follows: Section I introduces the topic, Section II discusses existing IoT architectures layer-wise with tabulated security mechanisms, Section III covers IoT threats classified according to the layered architecture, Section IV provides the most recent ML/DL-based security mechanisms and solutions for the IoT platform, Section V discusses available datasets for training ML/DL algorithms, Section VI presents future research directions briefly, and Section VII concludes the paper.

2. Architecture of IoT

The choice of an IoT architecture is crucial and depends on the unique requirements of the application at hand, including factors such as scalability, performance, security, and cost-effectiveness. A well-designed IoT architecture can greatly benefit organizations by enhancing customer experience, improving operational efficiency, and fostering innovation.

However, there is currently no universally accepted IoT architecture that fits all use cases. Therefore, researchers have proposed various IoT architectures with different layers to cater to different application requirements. In this section, we will delve into the most commonly used existing IoT architectures, thoroughly studying their features and functionalities. Towards the end of this section, we will summarize our findings in a tabular format Table1, categorizing the architectures layer-wise and detailing the existing security mechanisms employed in each layer. This will provide a comprehensive overview of the functionality and security measures of each architecture, aiding in the understanding of their strengths and limitations.

2.1. Three layers Architecture

The fundamental and widely used architecture for IoT is the three-layer architecture, comprising the perception layer, networking layer, and application layer (13),(14).

2.1.1. Perception Layer

The Perception Layer, analogous to the facial skin and senses of IoT, is responsible for object identification and data collection. It encompasses components such as 2-D bar code labels and readers, RFID tags and reader-writers, cameras, GPS, sensors, terminals, and sensor networks. Its main function is to identify objects and gather relevant data.

2.1.2. Networking Layer

An overview of the data flow throughout the system is given by the network layers. Data Acquiring Systems (DAS) and Internet/Network gateways are present in this layer. Data aggregation and conversion tasks are carried out by a DAS (collecting and aggregating data from sensors, then converting analog data to digital data, etc.). Data gathered by the sensor devices must be transmitted and processed, the network layer performs that function. It enables connections and communication between these devices and other servers, smart gadgets, and network devices. Additionally, it manages each device's data transmission.

2.1.3. Application Layer

The application layer, which provides the user with application-specific services, is where user interaction occurs. A dashboard that displays the status of the devices in a system or a smart home application where users may turn on a coffee maker by touching a button in an

ISSN: 2366-1313

app are two examples. The Internet of Things can be used in a variety of applications, including smart homes, smart cities, and smart health.

2.2.Four Layers Architectures

This approach reorganizes the layers of the IoT architecture, utilizing an application layer, data processing layer, network layer, and perception or sensor layer, in a slightly different sequence compared to the traditional three-layer concept (15). The application layer defines all IoT deployment apps and serves as the network interface for end IoT devices in a four-layer IoT architecture. This layer approves the supply of services to various apps based on the data collected by sensors. The data processing layer is responsible for verifying the legitimacy and security of data transmitted from users, after receiving information from the perception layer. The network layer, also known as the transmission layer, acts as a bridge connecting network devices and networks, and facilitates the transport and transfer of data obtained from physical objects through sensors. The perception layer, also referred to as the sensor layer, is responsible for identifying IoT devices and collecting data from them. It must be able to differentiate between different types of sensors on a network and account for their varying operating principles.

2.3.Five Layers Architecture

Miao Wu et al. in (16) proposed a five-layer architecture to explain the features and implications of the Internet of Things (IoT), which includes the Business layer, Application layer, Processing layer, Transport layer, and Perception layer (14). In this architecture, the transport layer functions as the network layer, and the processing layer is introduced to handle the vast amount of data available in the IoT network. Techniques such as database management, cloud computing, intelligent data processing, and ubiquitous computing are commonly used in this layer. Typically, data is pre-processed, evaluated, and stored here before being transmitted to the data center, where it is accessed by software applications that handle the data and prepare subsequent actions. The business layer is another layer suggested in this architecture, serving as an IoT manager that oversees the applications, the fundamental business model, and other business operations. The success of any IoT device depends not only on the technologies it employs, but also on how it is delivered to its users. The business layer is responsible for creating graphs and flowcharts, analyzing data, and determining how to enhance the device, among other tasks.

2.4.Six Layers Architecture

In the six-layer architecture, data from the Perception Layer is transmitted to the Observer Layer or Monitor Layer. If no signs of threats are detected, the received data is thoroughly examined before being forwarded to the middleware layer. This layer also carries out data sender authentication. The Security Layer ensures that all components of the IoT system are secure. It receives data from the processing layer and utilizes keys for encryption. Only encrypted data is transmitted further over the network in a manner that can only be accessed by authorized users(13),(17) and(18).

ISSN: 2366-1313

Table 1. Comparison of IoT layered Architectures

	Layer name	Protocols	Functions	Security mechanism
Three layers Architecture	**Application**	CoAP, MQTT, AMQP, XMPP, DSS, Service Discovery: mDMS, DNS-SD, SSDP Security: TLS, DTLS	providing all kinds of applications for each industry	Authentication/key- agreement, access control/privacy protection, intrusion detection and prevention systems
	Network	Addressing:IPv4/IPv6 Routing: RPL, CORPL, CARP, etc.	transmitting data	Encryption, Identity authentication, intrusion detection and prevention, access control, firewall
	Perception	IEEE 802.15.4, IEEE 802.15.1, IEEE 802.11, IEEE 802.3, IEEE 1901, LPWAN, RFID, NFC, Z-Wave etc	perceive the physical properties of objects by various sensors	Sensor data protection, encryption, secure boot, physical security, authentication/authorization, firmware/software updates
Four Layer Architecture	**Application**	CoAP, MQTT, AMQP, XMPP, DSS, Service Discovery: mDMS, DNS-SD, SSDP Security: TLS, DTLS	providing all kinds of applications for each industry	Authentication/key- agreement, access control/privacy protection, intrusion detection and prevention systems
	Data processing /support layer	database, intelligent processing, cloud computing, ubiquitous computing, etc	store, analyses and process the information's of objects	Data encryption, Data Anonymization, threats detection, access control
	Network	Addressing:IPv4/IPv6 Routing: RPL, CORPL, CARP, LoWPANs,Zigbee, LoraWLAN, Sigfox, Z-wave etc.	transmitting data	Encryption, Identity authentication, intrusion detection and prevention, access control, firewall
	Perception	IEEE 802.15.4, IEEE 802.15.1, IEEE 802.11, IEEE 802.3, IEEE 1901, LPWAN, RFID, NFC, Z-Wave etc	perceive the physical properties of objects by various sensors	Sensor data protection, encryption, secure boot, physical security, authentication/authorization, firmware/software updates
Five layer	**Business**	--	Managing various	Authentication and authorization,

ISSN: 2366-1313

			applications & user's privacy, decision making, performance monitoring	backup data and recover
	Application	CoAP, MQTT, AMQP, XMPP, DSS, Service Discovery: mDMS, DNS-SD, SSDP Security: TLS, DTLS	providing all kinds of applications for each industry	Authentication/key- agreement, access control/privacy protection, intrusion detection and prevention systems
	Processing	database, intelligent processing, cloud computing, ubiquitous computing, etc	store, analyses and process the information's of objects	Data encryption, Data Anonymization, threats detection, access control
	Transport	Addressing:IPv4/IPv6, 6LoWPAN Routing: RPL, CORPL, CARP, LoWPANs,Zigbee, LoraWLAN, Sigfox, Z-wave etc.	Networking and transmission of data, information	Encryption, Identity authentication, intrusion detection and prevention, access control, firewall
	Perception	IEEE 802.15.4, IEEE 802.15.1, IEEE 802.11, IEEE 802.3, IEEE 1901, LPWAN, RFID, NFC, Z-Wave etc	perceive the physical properties of objects by various sensors	Sensor data protection, encryption, secure boot, physical security, authentication/authorization, firmware/software updates
Six layer Architecture	**Application**	CoAP, MQTT, AMQP, XMPP, DSS, Service Discovery: mDMS, DNS-SD, SSDP Security: TLS, DTLS	Data management, device management, user interface, integration with other systems	Authentication/key- agreement, device identity management, access control/privacy protection, intrusion detection and prevention systems
	Network	Addressing:IPv4/IPv6, 6LoWPAN Routing: RPL, CORPL, CARP, LoWPANs,Zigbee, LoraWLAN, Sigfox, Z-wave etc.	Routing and Addressing, quality of services, device connectivity	Encryption, Identity authentication, intrusion detection and prevention, access control, firewall
	Security	SSL/TLS , AES,	Security related	Encryption, hash encryption,

ISSN: 2366-1313

	DTLS	functions	decryption
Middleware	cloud computing, big data, RDBS	Store, analyze and process data	Data processing, exact useful data, unrelated data removal
Observer	--	Monitor received data	Authentication
Perception	IEEE 802.15.4, IEEE 802.15.1, IEEE 802.11, IEEE 802.3, IEEE 1901, LPWAN, RFID, NFC, Z-Wave etc		Sensor data protection, encryption, secure boot, physical security, authentication/authorization, firmware/software updates

3. Security and Privacy

In traditional cyberspace, various types of attacks, such as those targeting connected computer networks, services, computer systems, embedded processors, and information storage or sharing, have existed for a long time. However, the sheer size of connected systems in the IoT and the relative simplicity of attacks present new vulnerabilities. This means that millions of connected devices with limited resources could potentially become targets of cyberattacks. In this section, we provide an overview of the most common attacks in an IoT environment. From a technical perspective, the multiple layers of the IoT ecosystem architecture, as outlined in this study, encounter various attacks, as illustrated in Fig 1.

3.1.Perception Layer

Perception layer is responsible for gathering data from various devices, like sensors, actuators, RFID tags etc. these devices are connected to internet and that makes them vulnerable(19),(20)if they are not properly protected. The sensors and data collected are the main target of attackers, who wants to utilize or replace them with their own. The common security issues of perception layers are(21):

3.1.1. Eavesdropping

These attacks are age old security issue, it takes merits of unsecure transmission, the attackers passively listen to the communication to gain access to private information, such as device identification number, application sensitive data, etc.

3.1.2. Node Capturing

It is the most significant attack, in this attacks the attackers physically catch the nodes and extracts the confidential data from the memory of the nodes.

3.1.3. Fake Node And Malicious

ISSN: 2366-1313

At this scenario, attackers adds a node of their own to the system and input fake data to the network with a purpose of stopping real information transmission. The added node tries to consume the precious energy of real nodes and potentially controls in order to damage the network.

3.1.4. Replay Attack

In this attack the intruder eavesdrop on communication and takes authentic information of sender, intruder send same authentication information to the victims showing proof of his authenticity and identity. Here the message is encrypted so the receiver consider it as reliable data and takes action as intruder desired.

3.1.5. Timing Attack

It is a security exploit that permits the attackers to discover vulnerabilities in the system, it is commonly used in devices that have weak computing capabilities. Here the attackers extract secrets maintained in the security of the system by studying how long it will take to respond to different inputs.

3.1.6. Jamming

In this a jammer can be used to completely or partially disrupt the signal of a node. IoT nodes are built on shared support that allows intruders to easily do radio interface or scrambling and that cause denial of services in the transmission. These attacks can done by intruder with bypassing any physical protocol layer or by emitting of radio signal to scramble specific channel. There are various kind of jammers such as constant jammers, decoy jammers, random jammers and reactive jammers. Proactive jammers, the purpose of proactive scramblers are to make the functional nodes nonreactive, broadcast signals without transferring any data communication in network by putting all the nodes in single channel until its energy is completely exhausted. Reactive Jammers, these kind of jammers block the signals, when they observe a network activity on data channel, a reactive jammer aims to compromise the reception of a message.

3.1.7. Tampering

Physical components of IoT system is the target, in this attack, the intruder gains direct access to a node's microcontroller or other hardware component. IoT Nodes are susceptible to quenching attacks since they are frequently utilized in a field and left unattended.

3.1.8. Collision

Here the intruder send his/her own signals, when he/she hears that a real node is transmitting a message, to interfere. Packets sent by two nodes at same time and frequency will collide, that can cause a huge disruption in the network.

3.1.9. Exhaustion, Repeatedly

Authentic requests are sent to implement a power exhaustion attack on resource constraint IoT devices, which keeps the device on and avoid device entering into power saving mode.

ISSN: 2366-1313

3.2.Network Layer

The movement of data within the application is outlined by the network layers. Data gathered by the sensor devices must be transmitted and processed. The network layer performs that function. It enables connections and communication between these gadgets and other servers, smart gadgets, and network gadgets.It manages each device's data transmission, on the other hand less secure wireless protocols, such as ZigBee, 802.15.4e, SigFox, LoRa, and 802.11x are used by IoT devices to connect with gateway or Internet.All these make it highly sensitive and vulnerable to attacks from the side of attackers.

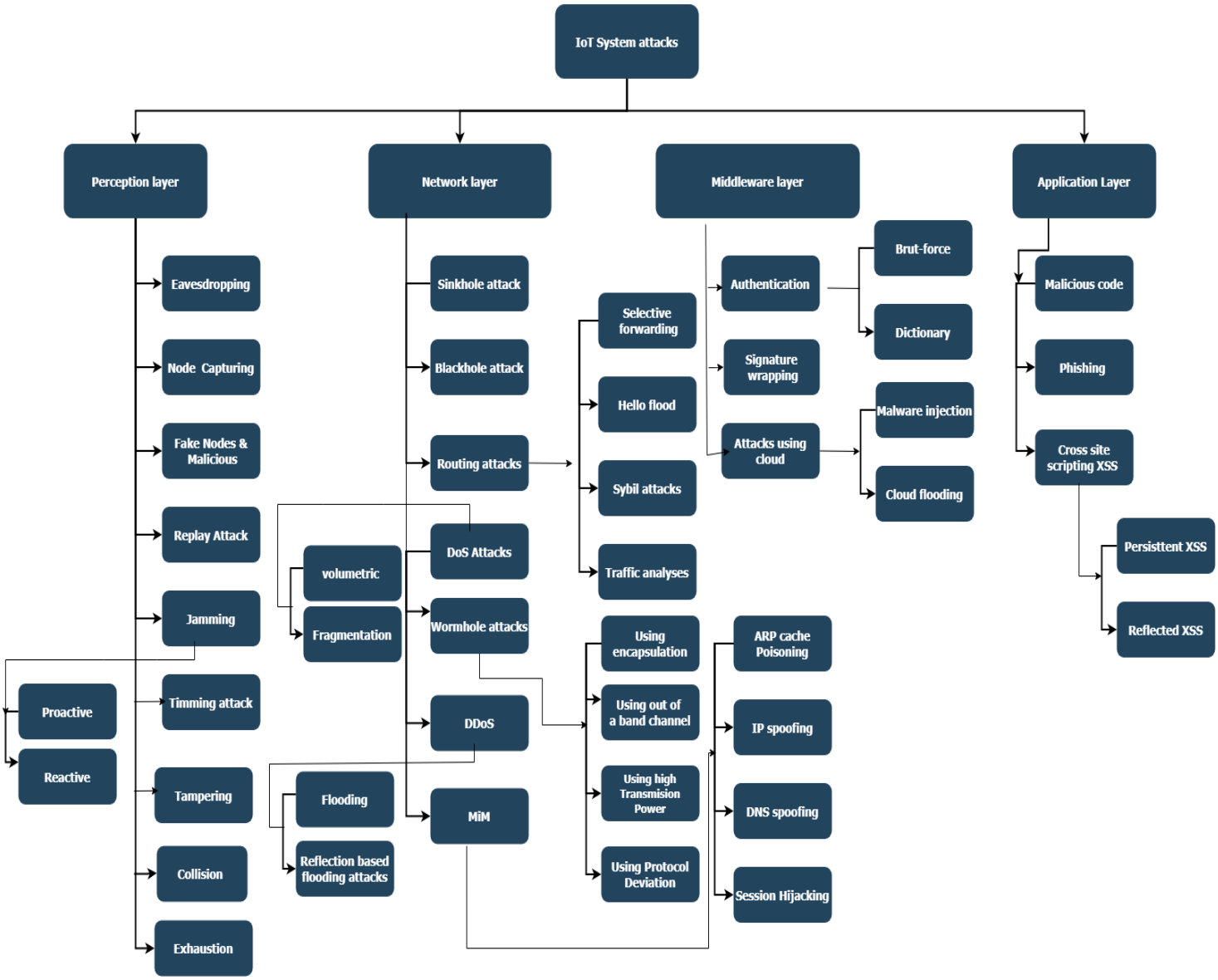

Fig 1. Internet of Things System's Attacks

. It has serious security problems with the authenticity and integrity of the data being transmitted across the network(22). The following are typical security issues that affect network layers:

3.2.1. Sinkhole Attack

ISSN: 2366-1313

An active type of attack is a sinkhole attack which is one of the most destructive routing attacks in IoT environment. With the help of fake routing information, the intrusive node draws the attention of its neighbour's, after which it performs selective forwarding or modifies the data travelling through it(23).

3.2.2. Blackhole Attack

A malevolent intruder drops all the packets that it is meant to forward during a blackhole attack. When coupled with a sinkhole attack, this attack can be exceedingly destructive and result in the loss of a significant portion of the traffic. It might be categorized as a denial-of-service attack(24). Several nodes can be cut off from the network if the attacker occupies a key position in the graph. Another variation of this attack is known as a "grayhole" attack, or "selective forwarding attack," in which the attacker merely discards a particular segment of the network traffic.

3.2.3. Routing Attacks

Packets are forwarded selectively in this attack. These attacks enable the launch of DoS (Denial of Service) attacks. Intruders want to block routing routes and filter any protocol. The RPL intruder could drop the rest of the packet and forward all RPL control messages(25).

- Selective Forwarding Attack, an attack that uses selective packet forwarding by malicious nodes can be used to launch a Denial-of-Service (DoS) attack. Disrupting routing paths is the general objective of this attack. It can be used to filter any protocol, although(26).
- Hello flood attack, one of the most common attacks at the network layer is the hello flood. Using the hello flood attack, the attacker can make conventional nodes to use a lot of power to transmit huge hello packets, which will cause them to lose power.
- Sybil attacks, here an attacker pretends to be numerous persons at once, it is known as a Sybil attack. It is one of the most significant problems when joining a P2P network. By creating numerous false identities, it controls the entire network and manipulates it. From a single perspective, each of these identities appears to be a typical user, but in reality, an unknown attacker—a single entity—controls each of these fake identities simultaneously. The Sybil attacker attacks the whole network.
- Traffic analysis attacks, Using the characteristics and patterns of the traffic on a link, traffic analysis seeks to obtain routing information. Even if the packets are encrypted, this attack is still possible to carry out. Similar to sniffing attacks, the goal is to identify parent/child links in order to acquire information about the RPL network, such as a partial image of the topology.

3.2.4. Denial Of Services Attacks

Denial of Service attacks are meant to shut down a device or a network in order to make inaccessible to the users and can be accomplished by flooding the target with traffic or delivering it data that can cause a crash. IoT devices are especially vulnerable to permanent denial of service (PDoS) exploits, which fully disable a system or device(27). In order to achieve this, batteries, power systems, or—more commonly—malware attacks, can be overloaded. In a software attack, the attacker may make use of vulnerabilities to replace a device's essential software, typically its operating system, with a corrupted or broken version of the software, turning the device useless. There are other various kind of DoS attacks to establish the desired goal of attacker, here we point out the most common types of DoS attacks. Lowe Rate DoS attack(28),(5),it makes use of the timeout feature of TCP/IP protocol. A TCP flow is made to enter a retransmission timeout state by precise periodic

ISSN: 2366-1313

traffic bursts that are sent during LR DoS attacks. Their network behaviour is similar to that of legitimate traffic, allowing the perpetrator of an LR DoS attack to remain undetected within the network and making it challenging for network managers to identify these attacks. IoT devices typically communicate data at a relatively low data rate, making LR DoS attacks of particular concern because they can go unnoticed in situations with minimal network traffic.

- Volumetric Attacks, any attack in which the bandwidth resources of the target network are intentionally used by the attacker is referred to as a volumetric attack. Once network bandwidth has been used, it is no longer available to authorized users and devices connected to the network. When an attacker performs a volumetric attack, they bombard network devices with ICMP echo requests until there is no more bandwidth left.
- Fragmentation Attacks, any attack that makes a network reassemble altered network packets is referred to as a fragmentation attack. In a fragmentation attack, the attacker manipulates packets sent to a network so that when the network tries to put them back together, it is unable to do so. The reason for this is that the packets contain more packet header information than is legal. As a result, packet headers that can't be reassembled in bulk are produced.

3.2.5. Distributed Denial Of Services

This is one of the most used and powerful DoS attack, here multiple systems are used to perform an attack on single target system, this multiple systems make this attack very difficult to detection for the victim, and it is because of the unknown origin of the attacks(29). There are various DDoS attacks, here we point out the most significant attacks:

- Flooding Attacks, this kind of attack involves flooding the network with irrelevant data, making the target system unavailable. More specifically the system drain is done by a big volume of requests from the intruder, for instance intruders flood various UDP (User Datagram Protocol) in various victim ports, therefore system will attend these requests over and over which make the victim system exhaustion of resources. Another sort of DDoS attack uses the TCP connection sequence to disable the victim's network, known as a SYN Flood attack. The victim's network responds to the attacker's SYN requests with a SYN-ACK response. The attacker should then reply with an ACK response, but instead, the sender doesn't (or uses a spoofed source IP address to send SYN requests instead). Network resources are used up by unanswered requests until no devices can establish connections.
- Reflection-based flooding Attacks, in this kind of attack, the attacker snoops on the real connection and repeatedly sends false requests to reflectors. These reflectors simultaneously respond to the target system, making it unreachable(29). An Example of DDoS attack is turning Up the Freeze was a Distributed Denial-of-Service (DDoS) attack against two apartment buildings in eastern Finland's environmental control systems. The DDoS attack fully shut down all climate control systems in the two apartments, leaving the occupants in the cold. The systems were rebooted to address the problem. The systems, however, became trapped in an endless cycle(30).

3.2.6. Man-In-The-Middle (MIM)

An attacker has complete control of a communication link between two legitimate entities during a MIM attack. Additionally, the attacker has the ability to alter, delete, and add messages to the communication channel in addition to reading them(31).

- Address Resolution Protocol (ARP) Cache Poisoning, spoofing is the way toward linking an attacker's mac address with the IP address of a legitimate user on a local

ISSN: 2366-1313

area network using fake ARP messages. Subsequently, information sent by the client to the host IP deliver is instead transmitted to the attacker.

- IP spoofing, IP spoofing includes a hacker impersonating a program by changing the packet headers in an IP address. Clients are consequently directed to the attacker's website when they attempt to access a URL linked to the application.
- DNS Spoofing, a DNS combines IP addresses with symbolic names. By storing misleading mapping data between symbolic names and IP addresses, a DNS spoofing, also known as DNS cache poisoning, affects the DNS resolver. An attacker may poison the DNS server by taking over an authoritative DNS server or by faking an answer to a recursive DNS query.
- Session Hijacking, when an attacker captures the user's session identifier and uses it maliciously to move the session to his or her own machine, the attack is known as a session hijacking attack.

3.2.7. Wormhole Attacks

Wormhole attacks are internal attacks that observe network activity without altering it, making it very challenging to detect the attack as the invaders are already part of the network. This attack can be performed in different modes:

- Wormhole attack using Encapsulation, two attacker nodes are utilised to launch the attack in this manner. To build a tunnel between them, the first attacker node is embedded close to the source and the second attacker node is embedded close to the destination. Attacker nodes are given instructions to encapsulate legitimate packets in malicious packets and tunnel them to a different end. Attacker nodes act as the shortest route when a request packet is sent by the source node and tunnel it to the other end without considering any other routes, achieving fastest route discovery.
- Wormhole attack using Out of Band Channel, this technique connects two different attacker nodes via a wired communication link with high-quality bandwidth or unusual frequency. Since there is no need for an intermediary node, it provides the quickest response time. Because the attacker nodes serve as the shortest path's endpoint, it can respond quickly when determining the route. It results in poor network performance as well as insecure communication.
- Wormhole attack using High transmission Power, the source node must be inside the attacker node's range in order for this mode to work with just one attacker node. It is the hardest to detect and the simplest to grow. Neither the routing table nor the header information are altered. Malicious nodes increase the transmission power and antenna height to communicate across long distances.
- Wormhole attack using protocol Deviation, in this attack, the communication routing protocol is disrupted to allow the attacker node to draw network traffic to itself. Invader monitors the request packet to initiate the attack and orders the attacker node to pass it to the target without deviating from other nodes so that it can be included to the route to the target. This technique is also used as a springboard for DoS assaults.

3.3. Middleware Layer

IoT middleware systems are the links that join IoT devices with higher-level services and applications because they pervasively integrate compute, networking, data management, and physical processes. As a result, it acts as an interface between IoT system components, enabling communication across diverse devices and applications that otherwise could not(32)(33). The middleware in the IoT architecture is made up of elements like the cloud. An assault on middleware directly targets the middleware components of the IoT system.

ISSN: 2366-1313

3.3.1. Attacks Using The Cloud

In a cloud-based attack, the attackers target a cloud platform directly for a variety of objectives, including data theft, flooding attacks, and so on. Popular cloud-based attacks consist of:

- Malware Injection in the Cloud, a cloud malware injection attack occurs when a hacker accesses a victim's data in the cloud and uploads a malicious copy of the victim's service instance, allowing the victim's service to be processed inside the malicious instance.
- Cloud Flooding Attack, with the help of an innocent host on the network, the attacker can flood the target with a large number of packets via a cloud flooding attack. These massive packets may contain a mix of TCP, UDP, and ICMP. This kind of attack may also damage the service's capacity to provide for authorised customers. Additionally, because the cloud cannot distinguish between legal and malicious traffic, its use may increase.

3.3.2. Authentication Attacks

Attacks that rely on authentication to verify a user, service, or application are known as authentication-based attacks.

- Brute-force, in a brute-force attack, the attacker tries a variety of login credentials in the hopes that they would be properly matched. Until the right password is determined, the attacker types a number of possible passwords.
- Dictionary attack, when an attacker creates a collection of possible passwords, it is known as a dictionary attack or a password-guessing attack. By listening in on the channel, the attacker carries out this and records the transcript. After then, attempts are made to generate passwords that match the ones that were previously recorded. If a match is found, the attacker has been able to obtain the password.

3.3.3. Signature Wrapping Attack

An attacker can make any web service request by using a signature wrapping technique to pass as a legitimate user. This is accomplished by including a malicious element in the message structure, which guarantees a reliable signature for the legitimate elements and processes the malicious element using the application logic(34).

3.4.Application Layer

All applications that make use of or have adopted IoT technology are categorized under the application layer. Smart homes, smart cities, smart health, tracking creatures, and other uses are possible with IoT. Around the world, application developers prioritize the efficiency and dependability of the product's service delivery over security. As a result, applications can be hijacked easily, preventing genuine users of authorized services. The following are some major vulnerabilities to the application layer:

3.4.1. Malicious Code

It is a piece of software code designed to have negative effects and harm the system. It is a threat kind that anti-virus software may not be able to stop or manage. It can either start up automatically or operate more like a software that demands the user's attention in order to take action. Malware that easily compromises nodes mostly targets IoT device vulnerabilities. The captured equipment is subsequently utilized as helpful nodes in the form of bots to attack other endpoints and network applications.

ISSN: 2366-1313

3.4.2. Phishing Attack

Phishing is a type of attack that seeks to obtain users' usernames and passwords by making them look to be a reliable entity. Cybercriminals may later utilise the sensitive information to harm the person or system(35).

3.4.3. Cross Site Scripting (XSS)

Web attacks like Cross-Site Scripting (XSS) are well-known. It happens when malicious web code, typically in script form, is delivered or executed from the victim's device's browser via certain web applications. As a result, it gives attackers the chance to steal sensitive information or possibly gain control of specific machines. With this execution, you might filter personal information or steal user cookies to hijack the identity in a fake session(36). Web interface of IoT devices mostly vulnerable to two main kind of XSS attacks(6).

- Persistent XSS, Persistent XSS involves an attacker inserting a malicious attack string into a website that stays there permanently in the database. As a result, anytime a user accesses an infected web page, the server will generate a response with an embedded attack string, which the web browser will eventually display.
- Reflected XSS, in reflected XSS, the attacker creates a URL with a malicious string and sends it to the user, causing the script to execute when the victim clicks on it and sends a request to the application server.

4. Machine Learning In Enhancing Security And Privacy Of IoT

In this segment of our research, we delve into the diverse Machine Learning (ML) algorithms that have been employed in the realm of Internet of Things (IoT) security and privacy. Figure 2provides a visual representation of the different categories of ML/DL techniques that have been utilized for IoT security and privacy purposes. Furthermore, we have conducted a thorough review of multiple international journal papers with the goal of exploring the potential of machine learning and deep learning techniques in addressing the security and privacy concerns associated with Internet of Things (IoT) applications. The findings of our review have been summarized and presented in Table 2 for easy comparison and reference.

In IoT networks, supervised learning techniques are used to solve challenges related to spectrum sensing, channel estimation, adaptive filtering, security, and location by using labeled data. Regression (Nearest neighbors and logistic regression) and classification (SVM, Naive Bayes, Random Forest, and Decision Tree) are two different sorts of approaches that fall under this category. In order to learn data representations with various levels of abstraction, deep learning methods offer a computational architecture that integrates many processing levels (layers).

4.1.Naïve Bayes

Naïve Bayes(37) is a classification algorithm, used for binary and multiclass environments, it has shown effective results on anomaly and intrusion detection problems. This algorithm performs well with discrete data. Because of its simplicity and computational easiness, the naive Bayes classification is one of the most widely used models in IDS. The chance that network traffic is abnormal or normal has been determined using a separate set of detected traffic parameters, such as status flags, protocols, and latency.

ISSN: 2366-1313

Features selection plays an important role in good training of machine learning model, Muhammad Shafiq et al.(12)implement various techniques in the features selection step of their proposed model for intrusion detection in an IoT environment to improve the detection accuracy. The authors applied the correlation-based metric BoT-IoT(38) dataset at the first stage to find in-depth the correlation between independent and class features, Area Under the Curve-based metric to find the most robust features carry accurate information for the BoT-IoT dataset attacks. They designed an algorithm based on Correlation and AUC metrics to filter the features and find out correlations between features and classes. Integrated TOPSIS and Shannon entropy were applied to validate the selected features. Five important features of the Bot-IoT dataset were selected and used to train four basic machine learning classifiers (SVM, DT, RF, and NB), the performance of classifiers is evaluated in terms of accuracy, precision, sensitivity, and specificity, and found that C4.5 decision tree and random forest algorithms outperform the other two achieve 96% average result.

In conclusion, Naïve Bayes is a widely used classification algorithm for intrusion detection in binary and multiclass environments. It has shown effective results on anomaly and intrusion detection problems, particularly with discrete data.

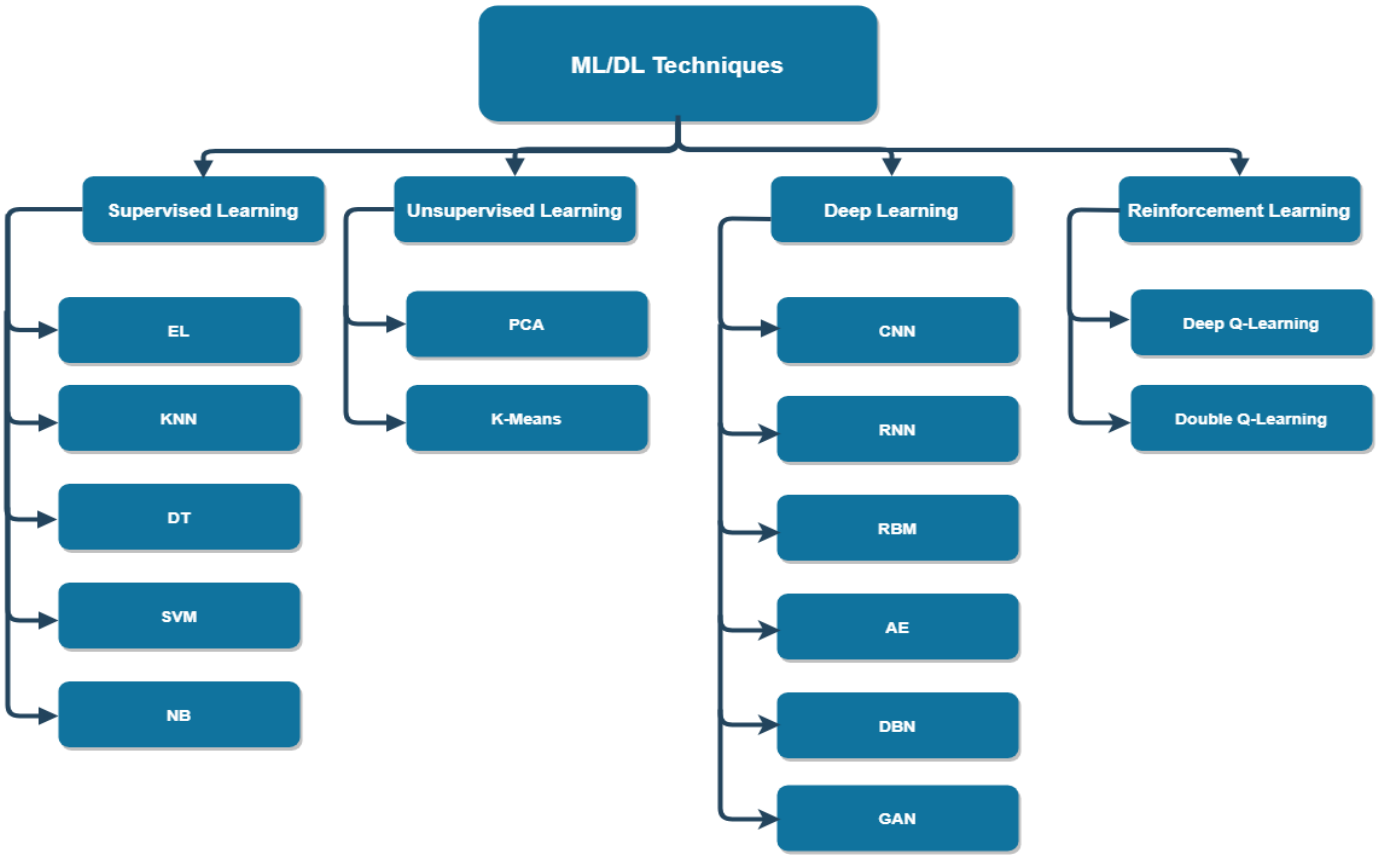

Fig 2. ML / DL Techniques for IoT Security and Privacy

4.2.K-Nearest Neighboring KNN

K-Nearest Neighboring KNN(39) is a supervised learning algorithm used for associating new data points with existing data points by searching through available datasets. This

ISSN: 2366-1313

algorithm performs well with fewer data, if the dataset is large and the network is huger the performance of this algorithm will be graded. KNN is very fast in learning and does not require any training, it stores the training data and learns from only when it makes the prediction. Smart homes are one of the IoT application areas, and securing IoT devices are a big challenge, Taotao Li et al. (8) Proposed a two-layer security approach for securing home-based IoT devices based on machine learning techniques. They created a real-time application that combined IoT devices and recorded all the data transmission using the TCPDUMP tool in PCAP format for a period of three days. The recorded data includes five types of attacks generated using a kali system, T-Shark tool was used to extract features from PCAP format and convert CSV files. The authors applied Kolmogorov Smirnow (KS) and Pearson Correlation (PC) approaches to select the best features. Five machine learning algorithms were tested for this experiment (MNB, SVM, DT, RF, and ANN), at the first layer classifiers identify attacks and benign and at the second layer the types of attacks are identified. Decision Tree algorithm obtains 98.71% accuracy in identifying malicious traffic and 99.00% accuracy in identifying attack types outperforms the other algorithms.

The results show high accuracy in identifying malicious traffic and attack types, which demonstrates the potential of these algorithms for enhancing security in smart homes.

4.3. Random Forest And Decision Tree (RF,DT)

Random forest and Decision Tree(40) are supervised learning algorithms, they define a model by implementing certain rules inferring from data features. DT is used for classification as well as regression problems. RF works well with large volumes of data, but it can cost huge storage prices. This algorithm can handle unbalanced data as well. Numerous DTs are used to derive several sophisticated machine-learning algorithms, including random forest (RF) and XGBoost. Anika Tasnim et al. in(41) focused on the performance of various machine learning techniques for identifying intrusions in an IoT environment, the latest datasets (TON-IOT) created in the Cyber Range and IoT Lab of UNSW university, including heterogeneous data sources collected from Telemetry datasets of IoT and IIoT sensors, Operating systems datasets of Windows 7 and 10 as well as Ubuntu 14 and 18 TLS and Network traffic datasets. In this study DT, RF, Adaboost, XGBoost, ANN, and Multi-layer Perceptron algorithms were trained and tested to classify attacks and normal network traffics. Accuracy, Precision, Recall, and F1-Score, Confusion Matrix were used in this study to evaluate the performance of the model. Overall accuracy achieved is 97% and above for both binary and multi-class classification of balanced and imbalanced data. Muhammad FasihAshfaq et al. in(42) carried out a detailed study on Logistic Regression and Decision Tree algorithms to classify DDoS attacks and normal traffic in an IoT environment. In this experiment, two different datasets were used for low-rate DDoS and high-rate DDoS real-time data collected from Wireshark which has two features, and KDD-Cup datasets containing 7 features. Accuracy and confusion matrices were considered for evaluating the performance of the model. The accuracy they achieved in this study in both the algorithms and datasets is very high compared to other methods being used as of now. RoumaissaBekkouche et al. in(43) proposed and deployed a very interesting approach to detecting and preventing malicious activity in an IoT environment using Decision Tree (DT) classifier. After preprocessing, discarding some nominal attributes, specific attributes, missing values, and correlated features, they down-sampled Avast IoT-23(44) dataset in two samples of 80% to train the model and 20% to test the model and achieved 99.9% accuracy for classifying each label. They also deployed the model with real-time traffic in a virtual environment, created using four machines. They retrieved incoming/outgoing traffic

ISSN: 2366-1313

generated from each system, preprocessed, and predicted the existence of attacks, when attacks were predicted from a specific machine its IP address was stored and blocked connection from the same IP address directly without going for further processing. Rajiv Yadav et al. in(11)proposed a lightweight method for Intrusion Detection System (IDS) based on Fast Correlation-Based Feature Selection (FCBFS) technique at the Feature Selection phase to reduce the complexity of IDS and make it more compact which in total makes nodes in IoT environment to reduce the power consumption and increase the lifetime of nodes. In this study, they evaluated and compared the accuracy, precision, recall, and F1-score for various ML techniques, FCBFS with XG-Boost with an accuracy score of 99.84 % gave a better performance compare to Decision Tree, Random Forest, NB, and ET.

These studies demonstrate high accuracy levels, ranging from 97% to 99.9%, in classifying attacks and normal traffic, and highlight the potential of these algorithms for enhancing security in IoT systems.

4.4.Support Vector Machine (SVM)

SVM (45) is a supervised learning technique with low computational complexity, used for classification and regression problems. It classifies data into n-dimensional space and draws an n-1 hyperplane to divide entire data into groups. It can do binary and multiclass classification and works with structured and semi-structured data. Its demerit is that cannot handle a large volume of data. SVM has been introduced for the anomaly detection of DoS attacks and malware detection in IoT networks, and it outperforms other machine learning algorithms in terms of accuracy, authors in(10)the performance of SVM on the latest IoT dataset(46) for multiclass attack classification. Mohammad DawoodMomand et al. in(9)proposed a support vector machine-based protocol for attack detection in RPL protocol for an IoT environment. The authors created a virtual IoT environment using the Contiki/Cooja simulator, to create a dataset containing three types of attacks by capturing traffic between simulated devices in the form of PCAP files. The attacks they considered in this experiment were Version number, rank, and DoS, generated data processed applying principal component analysis technique to find low dimension features to enhance routing efficiency and reduce energy consumption which also improves the detection accuracy of the SVM classifier. YakubKayodeSaheed et al. in(10)analyzed SVM, XGBoost, Cat Boost, KNN, QDA, and NB machine learning classifiers to classify attacks in the UNSW-NB15 dataset, in this study, they have used the min-max concept for the normalization of the dataset at the first stage to limit information leakage in test data and further applied Principal Component Analysis to reduce the dimension of the dataset, the attacks of the dataset were categorized into nine different types of attacks as follow: Analysis, Backdoor, DoS, Exploit, Generic, Reconnaissance, Fuzzers, Shellcode, Worm. XGBoost and Cat Boost classifiers give an accuracy of 99.99% with a training time of 0.7094 seconds, and 18.090 seconds respectively. The accuracy of the NB classifier is about 97.14% with a training time of 0.0102 seconds.

In conclusion use of techniques such as principal component analysis (PCA) and normalization of datasets can improve the accuracy and efficiency of attack detection, as shown in these studies. The high accuracy levels achieved with SVM and other classifiers highlight their potential for enhancing the security of IoT systems against various types of attacks.

4.5.Principal Component Analysis (PCA)

ISSN: 2366-1313

Principal Component Analysis (PCA) is not a method for detecting anomalies, it is frequently used to select or reduce features from huge datasets. To find irregularities and anomalies in an IoT network, the chosen feature sets can later be used in conjunction with a few machine learning classifiers. A big number of features can be reduced to a smaller set of features using the PCA technique without losing any important data(47). AaishaMakkar et al. in(48) did an interesting study on spam detection in IoT devices using machine learning techniques, for this study they examined five machine learning models and proposed an algorithm to evaluate the spamicity of each model. Based on the spamicity computation reliability of IoT devices was analyzed in terms of accuracy, precision, and recall metrics. To improve the performance of the proposed model they carried out a deep data preprocessing applying Principal Component Analysis (PCA) to reduce the variance among features in the dataset, followed by a features selection process using the entropy-based filter to find out the correlation between features.

Generalized Linear Model (GLM) has the best performance obtaining 91.8% accuracy and Bagged Model has the worst performance in this comparison obtaining 79.81% Accuracy. Time and energy consumption are two important parameters in the IoT environment ALEXANDER BRANITSKIY et al.(4) Proposed an interesting approach for intrusion detection in the IoT environment based on machine learning and parallel processing to overcome the time and energy constrained of IoT devices while securing these devices. The authors created a three layer-architecture, in the first layer min-max technique is used followed by the PCA technique and in the second layer basic machine learning classifiers are integrated and the third layer is for boosting the basic classifiers implementing one of the PV, WV, or SV composition techniques. The detection of the IoT-Botnet Attacks(49),(50) dataset was used for this experimental study, authors divided this dataset into several independent blocks inside each block the correlation coefficient is calculated and all blocks are processed in parallel mode to reduce the time-consuming. Parallel processes were done by creating 11 CSV files placed in each device in SPARK distributed data processing, each of them corresponding to one of the 11 attack classes in the dataset. They evaluated the performance of various basic classifier in both local and parallel processing mode, for local mode both binary and multiclass classification was carried out, but for parallel mode, only binary classification was evaluated, in local mode SVM has the best accuracy of 99.31% but time is 2698.25seconds which much higher compare to other classifies. For parallel mode Decision Tree classifier got the best performance achieving 99.99% in 156.62 seconds.

Overall, PCA is a valuable technique in these studies as it aids in reducing variance among features, identifying correlations, and enhancing the performance of machine learning models, particularly in IoT environments where time and energy constraints are crucial considerations.

4.6.Artificial Neural Network (ANN)

Artificial Neural Networks (ANN) recently, a machine learning technique called artificial neural networks have become very popular. When McCulloch and Pitts released a seminal research in 1942 speculating on how neurons in the human nervous system may function, this is when neural networks first came into existence(51). Mohammed Thakir Mahmood et al.(52)examined the performance of the ANN algorithm and some basic machine learning algorithms Decision Tree (DT), K-Nearest Neighboring (KNN), Support Vector Machine (SVM) and Random forest (RF) using self-generated dataset and KDD-Cupp99 dataset. First, they have done some preprocessing techniques on their dataset and split it into training, validating, and testing sets. For the testing of the model they combined attacks with a test set

ISSN: 2366-1313

of data and analyzed performance, in this experiment ANN outperforms the remaining machine learning classifier obtaining 97.77% accuracy in an execution time of 2.11 seconds.

ANN's ability to learn complex patterns from data and make predictions based on those patterns makes it a valuable tool for intrusion detection in IoT devices.

4.7.Extreme Learning Machine (ELM)

Extreme Learning Machine (ELM) a single hidden layer feed-forward network (SLFN) called the Extreme Learning Machine (ELM) was developed. Data feature representation and overall performance depend heavily on the neural architecture. The ELM was first designed to learn SLFNs, but it has since been modified to train generalized SLFNs in which the hidden layer need not be neuron-like(53).

Sawssen Bacha et al. in (54) Suggested anomaly-based intrusion detection using a kernel extreme learning machine for an IoT environment. For the improvement of the proposed IDS, they used the kernel principal component analysis technique to minimize the dimension of the dataset and to overcome the linearity of the PCA technique for non-linear features that exist in IoT datasets. They used two latest developed datasets for evaluating their proposed model, UNSW-NB15(46) and N-BaIoT(49) performance were evaluated in term of accuracy, specificity, sensitivity, f-score, and prediction time. 98.64% accuracy was achieved in the UNSW-NB15 dataset and 99.4% accuracy with the N-BaIoT dataset. The important point in this study is the prediction time, ELM algorithm process is much higher compared to other deep learning algorithms, and the prediction time for this model was 0.0010 and 0.00999 seconds in both datasets.

Overall, the study highlights the important qualities of ELM, such as its fast prediction time, generalization capability, dimensionality reduction, high accuracy, and adaptability, making it a valuable tool for intrusion detection in IoT environments.

4.8. Deep Learning (DL)

Deep Learning (DL)(55) hierarchical representations in deep architectures are learned using supervised or unsupervised learning methods based on several layers of artificial neural networks (ANNs). Multiple processing levels are present in DL architectures. Based on the information from its input layer, each layer can develop non-linear responses. The mechanisms of the human brain and neurons for signal processing are replicated in DL technology. They are plenty of studies that have used the DL algorithm to classify attacks in IoT applications, mentioning(56),(1) examined deep learning algorithms for intrusion detection and classification in IoT environments. RegondaNagaraju et al. in(57) combined deep learning with hybrid optimization techniques of Grey Wolf Optimization (GWO) and Whale optimization algorithm to enhance malware detection in IoT IDS, four separate datasets are used in this technique, and raw network traffic is stored in database-1, database-2 is filled with prior information, fresh characteristics of malware and viruses are stored in database-3 and lastly, IoT gadgets with pirated software is stored in database-4.

The approach was tested using Google Code Jam, first, the information is refined to extract the relevant tokens then the weight of the token’s determined. In this study researcher studied the impact of various malware image proportions in this approach, 255*255 and 228*228 were the image proportion. Leorpard Smartphone High-dimensional data was used to test the method, here 14,733 malware and 2486 benign items were used, and 228*228 proportion images outperform the other in categorization and accuracy obtaining an accuracy of 98% in 35 seconds.

ISSN: 2366-1313

SarikaChoudhary et al. in(1)evaluated the performance of DNN for intrusion detection in an IoT environment, three datasets (KDD99, NSL-KDD, UNSW-NB15) were used to train the model. Eight attributes were extracted from the mentioned datasets, used as input and attribute as output. The datasets were divided into 70,15,15 for training, validation, and testing respectively. This was implemented on 6BR (IPv6 Border Router) so that it would monitor the traffic, and based on their training, it would detect attacked behavior. UNSW-NB15 (46) dataset used with DNN produced the best performance of 99.2% at epoch 19 and NSL-KDD(58) used with DNN produced an accuracy of 91.5% attack detection at epoch 43. The limitation of this study is the datasets they used are outdated and they did not specify the features extraction method used in this study. Monika Vishwakarma and NishthaKesswani(56)studied the performance of a deep neural network algorithm on attack detection in a real-time IoT environment, developing a fire alarm and intelligent room lighting system. In this interesting experiment firs they developed a deep neural network-based intrusion detection system and trained the system with the latest dataset of NF-UQ-NIDS which is a combination of four different types of NetFlow-based benchmark datasets containing 20 different attacks. Early dropping and dropout were also implemented in this model to the complexity of the model. Binary and multiclass classifications were done in this model, obtaining 99.4% accuracy and 97.48% respectively. In a real-time environment, they test the proposed system for detecting DoS and MITM attacks, in this scenario by pyshark python model was used to capture live packets.

S Thavamani et al.(59)evaluated the performance of a few deep learning algorithms (GRU, CNN, RNN, LSTM) taking the KDD-CUPP public dataset to forecast intrusions in a specific protocol (MQTT) of IoT application layer. Since MQTT is one of the lightest protocols used in the application layer of IoT applications, attackers focus on vulnerabilities of this protocol to target. In this experimental study, the author found that the LSTM algorithm outperforms other algorithms for intrusion detection. ZhihanLv et al. in(60)designed and proposed a deep learning-based intrusion detection system for the IoT environment, for this study they examined various auto-encoder in the preprocessing step of training the model to improve the accuracy of the intrusion detection system. They constructed a Stacked DenoisingAutoencoder Support Vector Machine (SDAE-SVM), SDAE extracts features of the data and stores them in the features database, followed by a dimension reduction phase and SVM classifies attack behaviors in the model. They evaluated the performance of the proposed model for four layers of IoT architecture and obtained a promising accuracy above 97%. A comparison was carried out between the proposed model SDAE-SVM and shallow learning IDS, multi k-nearest neighboring (ML-KNN), semi-supervised fuzzy clustering algorithm (SFCA), fuzzy c-mean clustering (FCM), K-mean clustering, found that accuracy of deep learning based IDS much higher compare to the rest of machine learning based IDS.

Enhancing security in a multi-cloud IoT environment is a recent research study carried out by D. Selvapandianand R. Santhosh1 in(61), they proposed a deep learning-based intrusion detection model to detect attacks and classify them in an IoT environment. NSL-KDD(58) dataset was preprocessed for this research study one-hot encoding was applied for features extractions and mapped 41 features to a 122-dimensional features LeNet-based model selected for this research study to detect the intrusions in the IoT environment, the proposed model outperform some of the existing neural network based IDS obtaining overall of 97.5% accuracy. They also did a comparative analysis with SVM and RNN-based intrusion detection systems and performance was evaluated in terms of accuracy, precision, detection rate, and false positive rate, the proposed model gave the best performance among the three.

ISSN: 2366-1313

In summary, DNNs are of great importance in IoT-based IDS due to their ability to learn hierarchical representations, process data non-linearly, replicate human brain mechanisms, achieve high accuracy and performance, integrate with optimization techniques, automate feature extraction and preprocessing, and outperform other machine learning-based IDS. DNNs have the potential to significantly enhance the security of IoT environments by accurately detecting various types of attacks and mitigating their impact.

4.9. Transfer Learning (TL)

Transfer Learning (TL), promising machine learning techniques for deep learning include transfer learning, which focuses on transferring information across domains. The goal of transfer learning is to use information from a related field (referred to as the source domain) to enhance learning outcomes or reduce the number of labeled examples needed in the target field. It is important to note that the application of transferred knowledge to new tasks is not always beneficial(62). Eva Rodríguez et al. in (63)studied the performance of Transfer Learning (TL) to detect zero_day intrusions in an IoT environment, two phases were introduced in this study, in phase one BoT-IoT [38] dataset was used as the source domain dataset to train the TL with 75% and 25% ratio as training and validation sets respectively, in this phase model is referred as base ID-model and applied knowledge learned in this source domain to target domain in the second phase. In the second phase, USNW-NB15 (46)datasets were used to further train the model. The trained model is validated using two sub-categories of UNSW-NB15 datasets containing zero-day attacks and combined zero-day and known attacks, the result they achieved was extremely well, with 99.04% of accuracy in zero-day attacks and 97.89% in zero-day and known attacks.

In summary, the importance of TL in this study is significant as it demonstrates the effectiveness of TL in improving the performance of intrusion detection systems in the IoT environment. TL allows the model to leverage existing knowledge and adapt to different datasets, which can lead to improved accuracy, data efficiency, generalization capability, and enhanced security. TL has the potential to be a valuable approach in IoT security research, and further exploration of TL techniques can contribute to the development of more robust and effective security solutions for IoT systems.

4.10. Convolution Neural Network (CNN)

Convolutional Neural Networks (CNNs)(64) is a particular kind of discriminative DL model that has been largely applied to handle massive training data sets using hierarchical-based feature extraction and representation. Instead of using conventional fully linked networks, the network makes advantage of local connections and share weights to fully exploit the 2-D input data structure. The network can function more quickly and easily for training as a result of the process, which greatly reduces the number of parameters. Convolutional layers, pooling layers (subsampling layers), and activation units are the 2 different layers that make up a CNN architecture. Since CNN takes minimal training time, it is now perfectly suited for highly effective and quick feature extraction from the raw data set. However, CNN has been found to have a limitation in that it requires considerable processing power. Consequently, implementing CNN on IoT networks with limited resources could be quite difficult (65).

Harun SurejIlango et al. in(5)researched detecting LR DoS attacks in the SDN environment at the network layer using combined FFCNN and CNN, CICDoS2017 dataset being used for this research in two phases of pre-processing and LR DoS detection phases.

ISSN: 2366-1313

After preprocessing of the dataset, a wrapper-based feature selection using SVM was used to select a subset of important features. Due to time-constrained in the IoT environment for data processing feature reduction plays a more important role and this has been done in this study. FFCNN is used to further classify the attacks and benign with only seven features of the dataset in the network, the performance of FFCNN is compared to the machine learning algorithms J48, Random Forest, Random Tree, REP Tree, SVM, and Multi-Layer Perceptron to further identify attacks and benign behavior in the network. Accuracy, precision, recall, F1 score, detection time per flow, and ROC curves are used to evaluate the models' performance. According to the empirical analysis, FFCNN performs better than other machine learning algorithms across the range. This study is more specific in detecting one kind of DoS attack and the performance of the model to detect other attacks was not tested, so the hybrid of the method with other existing methods may increase the processing time again for IoT devices.

Davide Di Monda et al. in (66)experimented performance of ML and DL to classify attacks in IoT applications, in this experiment two different architectures were proposed, single-modal architectures in two forms of 1D-CNN classifier & hybrid 2D-CNN+LSTM classifier and multimodal architecture of 1D-convolution, bidirectional Gated Recurrent Unit (GRU). A comparison was carried out between the proposed architecture and the ML-based classifiers of Gaussian Naïve Bayes, Decision Tree, Random Forest, and Gagging Classifier. The latest IoT-23 (44) dataset was used for this study to train the model with a down-sample of 75% and tested the model with a down-sample of 25%, performance was evaluated in terms of accuracy, F-Measure and G-mean, DT with the result of 95.62% accuracy outperforms the ML-based classifiers and multimodal with an accuracy of 99.93% outperform single-modal and hybrid modal.

IMTIAZ ULLAH et al.(67)conduct a detailed DL model for intrusion detection in an IoT environment. Based on CNN the authors designed three different models of CNN1D, CNN2D, and CNN3D to examine this study. Combined datasets were created from BoT-IoT(38), IoT Network Intrusion(68) and IoT-23 to increase the number of attacks 15 and one normal and divided this into three subsets (training, validating and testing). The Recursive Features Elimination technique was used to extract the relevant features from the created dataset followed by a random forest algorithm for estimation of the overall importance of features 64 features. The performance of this research was evaluated in terms of accuracy, precision, recall, and F1 metrics. Transfer learning principle was used to build multiclass and binary classification, multiclass classification was carried out for each dataset separately and the created dataset to classify normal and existing attacks in datasets using the three designed models obtained a detection rate of 99.6% for normal as highest and 88.23% for MITM attacks as lowest. The Binary classification was also carried out for all three models and the minimum detection rate was 99.71%. At the end of the research paper authors compared the experimental study based on CNN with other research studies and declared the proposed study was more effective in anomaly detection.

CNN plays a critical role in these studies by providing effective feature extraction and classification capabilities, improving the accuracy and performance of intrusion detection and anomaly detection in SDN and IoT environments, and outperforming other machine learning algorithms in various evaluation metrics.

4.11. Recurrent Neural Network (RNN)

A discriminative DL algorithm is the RNN. If the application data must be processed sequentially, such as with voice text or sensor data, and if there is a dependent relationship between the current state and prior states, RNN would be a suitable method. However, there

ISSN: 2366-1313

is no interdependency between input and output in the conventional neural network(69)(64)Long short-term memory (LSTM), gated recurrent unit (GRU), and bi-RNN are only a few of the more advanced RNN variations that have been proposed. These help solve the vanishing gradient problem and long-term dependency problem.

Authors (59)used some variants of RNN algorithms to design an IDS for attack detection in an IoT environment. Xiaoyong Yuan et al. in(70) proposed a RNN based approach to detect DDoS attacks, four RNN models (LSTM, CNNLSTM, GRU, 3LSTM) were used to evaluate DDoS attacks detections on the two versions of ISCX2012 dataset Data14 and Data15, this approach utilizes a sequence of continues network traffic. In the Second model, they used CNN to correlate network fields and maximize the efficiency of the approach, the accuracy achieved in this model was 95.89%. The 3LSTM model outperformed the remaining three models by achieving an accuracy of 98.41% and an error rate of 1.590. In this approach features extraction and features selection were not explained clearly.

Dr.Janardhana et al.(71)studied performance deep learning and machine learning algorithms trained using two known datasets NSL-KDD(58) for binary classification and UNSW-NB15(46) for multiclass classification in an IoT environment. For this study, the authors did preprocessing steps for the chosen datasets to improve the detection accuracy and trained CNN, RNN, NB, DT, and SVM algorithms, the performance of each algorithm was evaluated in terms of accuracy, precision, recall, and F1-score. RNN with hyperparameter optimization produced the highest accuracy of 96.6% in detecting security and privacy attacks.

Overall, the studies mentioned highlight the importance of RNN in designing IDS for attack detection in IoT environments, capturing sequential patterns in time-series data, feature extraction and selection, performance comparison with other algorithms, and hyperparameter optimization for achieving high accuracy in IoTattack detection.

4.12. Reinforcement Learning (RL)

Reinforcement Learning (RL) is a type of machine learning where an AI agent attempts to complete a task by choosing the optimal next step that can result in a greater ultimate reward overall(72). In a real-world scenario, the agent goes through a lot of trial-and-error stages and seeks to optimize the reward it receives from the environment. An agent interacts with an environment, which can be a simulator, a game, the actual world, etc. RL is a Markov decision process (MDP) where the results of actions taken from states are entirely dependent on the current state, regardless of previous states and actions.

Integration of deep learning and reinforcement learning is a current research topic in the field of cybersecurity Sunder Ali Khowaja et al. (73)Proposed an integrated framework of Q-learning and LSTM for Industrial Internet of Things (IIoT) malware detection. The authors used a combined public dataset for this experiment and applied phase space embedding (PSE) and space autoencoder (SAE) transformation techniques to transform the static and dynamic extracted features for training the active learning model to predict malware in the IIoT environment, here the model learn the policy using Q-learning objective. LSTM network is trained with action-value function on both PSE and SAE and further adaptive weighting and meta-learner (NB) decision-level fusion approach is applied to combine the result from both the stream and improve prediction accuracy. The authors created a hypothetical framework of the suggested model to detect malware in the IIoT environment. Data generated by sensors and IoT devices are passed through an IIoT gateway, a sniffer is placed to sense packets and store them in a database, and stored packets undergo through preprocessing step, first Androguard tool is used to extract static features and MobileSecurityFramework (MobileSF)

ISSN: 2366-1313

is used to extract dynamic features. The proposed model is implemented at this stage to predict malware, experiment showed that the efficiency of the proposed model is sufficiently high compared to supervised learning by using only 50% of the training data.

Over Overall, the study suggests that RL can be a promising algorithm for IoT security, particularly for IIoT malware detection, by integrating with deep learning techniques and leveraging its ability to actively acquire knowledge, combine information from different sources, and achieve efficient utilization of data.

5. Datasets

To handle the security issues in the Internet of Things environment with innovative solutions including machine learning-based spam, fraud, and virus detection, various data types are required. The data includes system logs, network traffics, application logs, binary or raw alerts, event traces, and threat information. A well-structured dataset is required to train the machine learning model using supervised and unsupervised algorithms. . In order to design and validate efficient and accurate protection systems to detect IoT attacks, the availability of public datasets is a critical point in the research world. In this section we studied and tabulated Table 3. The most popular available IoT datasets and also outlined the available attacks in datasets and their demerits.

6. Future Research Ideas

Applying ML/DL in IoT security and privacy has the potential to significantly improve the detection and mitigation of threats and attacks targeting IoT devices. These are some possible paths for future research in the use of ML/DL in IoT security and privacy

- There is growing interest in employing IoT devices themselves to process and analyze data without storing it at a central place due to privacy issues around generated data and relevant legislation surrounding managing this data. A decentralized ML/DL strategy called federated learning enables numerous IoT devices to work together to train a model withoutsharing their data with a central server. Future studies could concentrate on the potential of

Table 2. Comparison of Machine Learning and Deep Learning Approaches for IoT Security and Privacy

ISSN: 2366-1313

Reference	Classifier Type	Feature extraction/selection	Classifier algorithm	Attacks	Dataset	No of features	Accuracy
(74) 2022	multiclass classification	--	CNN	IoT device classification	Self recorded dataset	--	92%
(11) 2022	multiclass classification	(FCBFS)	XG-Boost	DoS, R2L,U2R, Probe	NSL-KDD	10	99.84 %
(5) 2022	multiclass classification	wrapper-based feature selection using SVM	FFCNN	LR DoS	CIC DoS 2017	7	97%
(10) 2022	multiclass classification	min–max concept, Principal Component Analysis (PCA)	XGBoost, CatBoost, KNN, SVM, QDA, and NB	Analysis, Backdoor, DoS, Exploit, Generic, Reconnaissance, Fuzzers, Shellcode, ,Worm	UNSW-NB15	10 out of 49	99.99%
(57) 2022	Multiclass	GWO, Whale Optimization	DCNN,	General IoT attacks	Google code jam (GCJ)	--	99%
(70) 2017	Muliclass	--	LSTM, CNNLSTM, GRU, 3LSTM	DDoS	ISCX2012	20	98.41%
(56) 2022	Binary & multiclass	Early stopping, drop out, adaptive gradient	DNN	20 types of attacks	NF-UQ-NIDS	9	99.4% 97.48%
(1) 2020	Multiclass	---	DNN	9 attacks	Kdd99, nslkdd, unsw-nb15	9	91.5%,96.3% and 99.2%
(63) 2022	Multiclass	one hot encoding, standard normalization	Transfer Learning (TL)	(generic, exploits, DoS,reconnaissance, fuzzers , analysis, backdoor, shellcode, and worms)	BoT-IoT, UNSW-NB15	46 and 49	99.04% & 97.89%
(42) 2022	Binary & multiclass	Random forest	Logistic Regression & Decision Tree	DDoS	Wireshark collected data & KDD-Cup	2 for Wireshark dataset, 7 for KDD -cup	99.99%
(41)	Binary &	SMOTE (Synthetic	DT, RF,	backdoors, DoS,	TON-IoT		

ISSN: 2366-1313

2022	multiclass	Minority Oversampling Technique)	Adaboost, XGBoost, ANN and Multi-layer Perceptron	DDoS, injection, scanning, Man-in-the-Middle (MitM), ransomware, password assaults, and Cross-site Scripting (XSS)			
(59) 2022	Multiclass	--	GRU, CNN, RNN and LSTM	DDOS attack-HOIC, DDOS attack-LOIC-HTTP, DDOS attack Hulk, Bot FTP-bruteforce, SSH-Bruteforce, Infilteration, DOS-SlowHTTPTest, DoS-attacks GoldenEye	KDDCUPP MQTT	9	78%
(66) 2022	Multiclass	Adam optimizer , early-stopping and softmax activation	ML-based classifier : Guassian Naïve Bayes, Decision Tree, Random Forest, Bagging Classifier DL-based: 1D-CNN, 2D-CNN+LSTM, 1D-CNN+GRU	20 malicious traces and 3 benign	Avast IoT-23	23	95.62% & 99.93%
(43) 2022	Multiclass	--	DT	20 malicious traces and 3 benign	Avast IoT-23	23	99.9%
(54) 2022	Binary and multiclass	Kernel principal component analysis KPCA	Kernel Extreme Learning Machine KELM	Analysis, Backdoor, DoS, Exploit, Generic, Reconnaissance, Fuzzers, Shellcode,	UNSW-NB15, N-BaIoT	--	98.64% & 99.4%

ISSN: 2366-1313

				,Worm			
(48) 2021	Multiclass	---	Bagged Model, Bayesian Generalized Linear Model (BGLM), Boosted Linear Model, eXtreme Gradient Boost (xgBoost), Generalized Linear Model (GLM) with stepwise feature selection	--	REFIT Smart Home dataset	15	
(61) 2021	Binary and multiclass	One-hot encoding	CNN	4	NSL-KDD	41	97.5%
(67) 2021	Binary and multiclass	Recursive Features Elimination	CNN	15	BoT-IoT, IoT Network Intrusion, MQTT-IoT-IDS2020 and IoT-23	64	99.6% & 99.71%
(71) 2021	Binary and multiclass	Hyper parameter optimization	CNN,RNN, SVM, NB, DT	4 & 9	NSL-KDD, UNSW-NB15	-	96.6%
(9) 2021	Multiclass	PCA	SVM	3	Self generated dataset	3	90-92%
(4) 2021	Binary and multiclass	PCA, PV,WV, SV	SVM,DT ,MP	10	N-BaIoT	115	99.99%
(12) 2021	Multiclass	Correlation, AUC, Shannon, TOPSIS	SVM, DT, NB, RF	5	BoT-IoT	46	96%
(60) 2020 IEEE	Multiclass	One-hot encoding	Stacked Denoising Autoen	4	NSL-KDD	41	98%

ISSN: 2366-1313

			coder-SVM				
(52) 2020 IEEE	Binary	k-mean clustering	ANN, KNN, SVM DT	4	Self generated dataset and KDD-Cup99	4 AND 41	97.77%
(8) 2020 IEEE	Binary & multiclass	KolmosgorovSmirnow, pearson correlation	ANN, KNN, SVM DT , MNB	9 and 29	Self generated	88	98.71% & 99%
(73) 2020 Springer	Binary	Phase space embedding, spare autoencoder	Q-Learning & LSTM	-	Combined public dataset	--	

Table 3. Existing datasets ML/ DL Approaches for IoT Security and Privacy

Serial #	Reference /Dataset	Attacks	Description	Demerits
1	(75)AIoT-Sol	MITM, PPPD RCE, SSDP flood, SYN flood, SSL regeneration, directory brute, directory traversal, command injection, open redirect, SQLi, XXE, XSS, SSRF, CSRF, MQTT brute force	This dataset created incorporating physical IoT devices and networking devices mapped with OWAP Top 10 IoT security risk.	Does not contains all available attacks
2	(76)MQTTset	Bruteforce, DoS,Flood, malformed slowite	This dataset is created in real time scenario from 10 different sensors using MQTT and CoAP protocols	Focused mainly on MQTT attacks
3	(77)MQTT-IoT-IDS2020	Aggressive scan, UDP scan, MQTT brute force, SSH brute force	provides a dataset with Message Queuing Telemetry Transport (MQTT) protocolrelated benign or attack instances	Focused mainly on MQTT attacks
4	(78)IoTID20	Syn Flooding, Brute Force, HTTP Flooding, UDP Flooding, ARP Spoofing, Host Port, OS Scan	Dataset is generated in a simulated smart home testbed consisting of SKT NGU and EZVIZ WiFi cameras. It has three label features which are binary, category, and subcategory.	Does not contains all available attacks
5	(79) CIDDS-001/ CIDDS-002	DoS, Ping & SYN scan, SSH bruteforce, {ACK, FIN, UDP, SYN, Ping} Scans	Both datasets are created from an emulated business environment for a period of four weeks	they are provided as a unidirectional flow-based traffic, contain less number of attacks
6	(80)CSE-	botnet, XSS, DoS, DDoS,	The generation of CIC-	Not focused on IoT

ISSN: 2366-1313

	CICIDS2017/2018	heartbleed, infiltration, SSH bruteforce, SQLi , botnet, bruteforce, port scan, DDoS, DoS, web attack, infiltration attack	IDS2018 was done in an emulated environment for 5 days. CIC introduced a network traffic flow generator called CICFlowMeter that extracts 80 statistical features	context
7	(46)UNSW-NB15	backdoors, DoS, exploits, fuzzers, port scans, recon, shellcode, spam, worms	This dataset combines actual modern regular network traffic with synthetic attack activity	Not focused on IoT context
8	(49)N-BaIoT	BASHLITE Attacks and Mirai attacks	It suggests real traffic data, gathered from 9 commercial IoT devices authentically infected by Mirai and BASHLITE.	Focused on Wi-Fi communication
9	(38)BoT-IoT	DDoS, DoS, OS and Service Scan, Keylogging and Data exfiltration attacks, with the DDoS and DoS attacks further organized,	The BoT-IoT dataset was created by designing a realistic network environment in the Cyber Range Lab of UNSW Canberra.	Authentication and disconnection phase not found
10	(81)TON_IoT	DoS,DDoS, Scanning, ransomware, backdoor, injection, xss, password, MITM	The TON_IoT datasets are new generations of Industry 4.0/Internet of Things (IoT) and Industrial IoT (IIoT) datasets for evaluating the fidelity and efficiency of different cybersecurity applications based on Artificial Intelligence (AI)	raw PCAP traffic data related to MQTT was not released
11	(82)MedBIoT	Bashlite,Mirai, tori	This dataset is created in medium IoT network consist of 80 devices	Authentication phase not found, no MQTT attacks
12	(44)IoT-23	20 malwares	These IoT network traffic was captured in the Stratosphere Laboratory	Focused on DNS traffic for IoT context
13	(58)KDD-Cup/ NSL-KDD	DoS, Probing, Privilege Escalation (Remote to Local and User to Root)	It contains primary attributes about TCP connections, high-level attributes such as the number of failed logins and contains more than 20 different types of attacks. NSL-KDD comes with an enhancement by removing the duplicates	Not focused on IoT context

federated learning for IoT security and privacy, as it can facilitate efficient threat detection while protecting user privacy.

ISSN: 2366-1313

- Security systems must continually react to new threats as they emerge because IoT device threats and attacks are constantly changing. Future studies can concentrate on creating self-adaptive and self-learning ML/DL systems that can recognize and react to zero day attacks instantly.

- This study found that the majority of approaches with empirical investigation for example in (8),(9),(52),(74) were built and tested with just a few varied types of sensor nodes, proving that one of the basic tenets of IoT applications is the interoperability of heterogeneous devices. However, routing protocol performance and message processing speeds may be impacted by various hardware configurations. As a result, when developing IoT security solutions, researchers must take hardware heterogeneity into account.
- We discovered that the majority of studies employed well-known evaluation measures and characteristics to confirm the efficiency of their research. However, those criteria are frequently employed while creating systems for identifying attacks in conventional networks. There is also a necessity to examine the proposed mechanism in terms of computational cost, deployment strategy, etc. due to the constrained environment of IoT networks. A lightweight ML and DL solution that can operate in a limited environment and be adapted for deployment in small devices is also required.
- Number and complexity of threats and attacks targeting IoT applications are growing. Comprehensive datasets that include the most possible attack patterns, considering diversity of IoT devices to reflect the rapidly evolving landscape of IoT technology, incorporating adversarial attacks and contextual data to improve the accuracy, should be generated for training ML and DL algorithms to detect and mitigate these threats. Such datasets can also be utilized to compare the efficacy of recently proposed algorithms to that of current attack detection techniques. Although creating collaborative databases on IoT threats that can be regularly updated with new attacks is essential, the vast variety of IoT devices makes it technically difficult. Furthermore, there is a privacy concern because datasets, particularly for industrial and medical IoT devices, may contain sensitive or important information that is not intended to be shared publicly.

7. Conclusion

Due to the potential risks and vulnerabilities associated with IoT devices and systems. Researchers have recently been very interested in IoT security and privacy concerns. The dynamic nature of IoT networks causes a variety of challenges for traditional security and privacy solutions. Machine learning (ML) and deep learning (DL) play an increasingly important role in developing new security mechanisms for IoT systems by enabling faster and more accurate detection of security threats, improving the efficiency and effectiveness of security models. In this paper, we have studied the various ML/DL based approaches proposed by researchers in the IoT from security and privacy perspective. We have studied the security and privacy challenges in IoT architecture, existing attacks targeting each layer of IoT systems.

Then we have studied available datasets for training the machine learning models indicating significant points. We have also outlined some challenges in this domain as future research directions. In our upcoming work, we'll take into account the findings as we create and develop machine learning-based intrusion detection algorithms for protecting IoT devices.

ISSN: 2366-1313

References

1. *Analysis of KDD-Cup'99, NSL-KDD and UNSW-NB15 Datasets using Deep Learning in IoT* . Sarika Choudhary, Nishtha Kesswani. Rajasthan, India : ELSEVIER, 2020. 10.1016/j.procs.2020.03.367.
2. *A critical review of intrusion detection systems in the internet of things: techniques, deployment strategy, validation strategy, attacks, public datasets and challenges.* Alazab, Ansam Khraisat and Ammar. s.l. : Springer , 2021. https://doi.org/10.1186/s42400-021-00077-7.
3. *A decade of research on patterns and architectures for IoT security.* Tanusan Rajmohan, Phu H. Nguyen and Nicolas Ferry. s.l. : Springer , 2022. https://doi.org/10.1186/s42400-021-00104-7.
4. *Applying machine learning and parallel data processing for attack detection in IoT.* ALEXANDER BRANITSKIY, IGOR KOTENKO, IGOR SAENKO. Saint-Petersburg Russia : IEEE, 2021, Vol. 9. doi: 10.1109/TETC.2020.3006351.
5. *A FeedForward–Convolutional Neural Network to Detect Low-Rate DoS in IoT* . Harun Surej Ilango, Maode Ma, Rong Su. Qatar : ELSEVIER , 2022. https://doi.org/10.1016/j.engappai.2022.105059.
6. *Adaptive cross-site scripting attack detection framework for smart devices security using intelligent filters and attack ontology.* Pooja Chaudhary1 • B. B. Gupta2, 3,4,5 • A. K. Singh1. Germany : Springer, 2022.
7. *RPL Attack Detection and Prevention in the Internet of Things Networks Using a GRU Based Deep Learning.* Cakir, Semih, Toklu, Sinan and Yalcin, Nesibe. s.l. : IEEE, 06 October 2020, Vol. 8. DOI: 10.1109/ACCESS.2020.3029191.
8. *Machine Learning-based Intrusion Detection for IoT Devices in Smart Home.* Taotao Li, Zhen Hong, and Li Yu, Member. Singapore : IEEE, 2020. DOI: 10.1109/ICCA51439.2020.9264406.
9. *Machine Learning-based Multiple Attack Detection in RPL over IoT.* Momand, Mohammad Dawood and Mohsin, Mohabbat Khan. Indis : IEEE, 2021. 10.1109/ICCCI50826.2021.9402388.
10. *A machine learning-based intrusion detection for detecting internet of things network attacks.* Yakub Kayode Saheed, Aremu Idris Abiodun, Sanjay Misra, Monica Kristiansen Holone, Ricardo Colomo-Palacios. s.l. : IEEE, 2022. https://doi.org/10.1016/j.aej.2022.02.063.
11. *Augmentation in performance and security of WSNs for IoT applications using feature selection and classification techniques.* Rajiv Yadav, Indu Sreedevi, Daya Gupta. Delhi India : ELSEVIER, 2022. https://doi.org/10.1016/j.aej.2022.10.033.
12. *CorrAUC: A Malicious Bot-IoT Traffic Detection Method in IoT Network Using Machine-Learning Techniques.* Muhammad Shafiq, Zhihong Tian, Ali Kashif Bashir, Xiaojiang Du, Mohsen Guizani. 5, s.l. : IEEE, 2021, Vol. 8. doi: 10.1109/JIOT.2020.3002255.
13. *Internet of Things: Architectures, Protocols, and Applications.* Sarangi, Pallavi Sethi and Smruti R. Delhi India : Hindawi, Journal of Electrical and Computer Engineering, 2017, Vols. Volume 2017, Article ID 9324035, 25 pages. https://doi.org/10.1155/2017/9324035.
14. *Research on the architecture of Internet of things* . Miao Wu, Ting-lie Lu, Fei-Yang Ling, ling Sun, Hui-Ying Du. Chengdu : IEEE, 2010. 10.1109/ICACTE.2010.5579493.
15. *The key layers of IoT architecture.* Alae-Eddine Bouaouad, Adil Cherradi,Saliha Assoul, Nissrine Souissi. Marrakesh, Morocco : IEEE, 2021. 10.1109/CloudTech49835.2020.9365919.

ISSN: 2366-1313

16. *LoRaWAN Technology Mapping to Layered IoT Architecture* . F. Flammini, D. Dobrilović, A. Gaglione and D. Tokody. Zrenjanin, Serbia : AIIT - International Conference on Applied Internet and Information, 2020.
17. *The key layers of IoT architecture.* Alae-Eddine Bouaouad, Adil Cherradi,Saliha Assoul,Nissrine Souissi. Morocco : IEEE, 2020. 10.1109/CloudTech49835.2020.9365919.
18. *Improved Layered Architecture for Internet of Things.* Darwish, Dina Gamal. August 2015, e, ElManial, Cairo, Egypt : International Journal of Computing Academic Research (IJCAR), , 2015, Vol. Volume 4. 2305-9184.
19. *IoT Elements, Layered Architectures and Security Issues : A Comprehensive Survey.* Muhammad Burhan, Rana Asif Rehman,Bilal Khan and Byung-Seo Kim. Sejong City 30016, korea : MDPI sensors , 2018.
20. *Perception layer security in the internet of things.* K.Aarika, M.Bouhlal, , R.AitAbdelouahid , S.Elfilali , E.Benlahmar,. Belgium : Elsevier B.V, 2020. 175 (2020) 591–596.
21. *Study on Security Architecture in the Internet of Things.* Li, Lan. Harbin China : IEEE, 2012.
22. *An In-Depth Analysis of IoT Security Requirements, Challenges, and Their Countermeasures via Software-Defined Security.* Waseem Iqbal, Haider Abbas, Mahmoud Daneshmand ,Bilal Rauf, and Yawar Abbas Bangash. s.l. : IEEE, 2020, Vols. JOURNAL, VOL. 7, NO. 10,.
23. *Study of the Impact of Sinkhole Attack in IoT Using Shewhart Control Charts.* Fakhr-eddine HACHEMI, Mohammed MANA, Boucif Amar Bensaber. Tlemcen, Algeria : IEEE, 2020.
24. *Detection and Prevention of Black Hole Attacks in IoT & WSN.* Shoukat Ali, Dr. Muazzam A Khan, Jawad Ahmad,Asad W. Malik, and Anis ur Rehman. Islamabad : IEEE, 2018.
25. Rushabh Vaghela1, Prof. Deepak Upadhyay2. *A Survey on Routing Attacks in Internet of Things (IOT).* Gujrat India : International Research Journal of Engineering and Technology (IRJET), 2020.
26. *Deep learning and big data technologies for IoT security.* Mohamed Ahzam Amanullah, Riyaz Ahamed Ariyaluran Habeeb, Fariza Hanum Nasaruddin ant others. Malaysia : ELSEVIER, 2020.
27. *DoS Attacks in IoT Systems and Proposed Solutions.* Nada Abughazaleh, Ruba bin Jabal, Mai Btish. Jeddah, KSA : International Journal of Computer Applications, 2020, Vols. Volume 176 – No. 33,.
28. *Low-Rate TCP-Targeted Denial of Service Attacks.* Knightly, Aleksandar Kuzmanovic and Edward W. Houston,USA : Presented at the Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications. Karlsruhe, Germany., 2003.
29. *DDoS attack on IoT Devices* . Asmaa Munshi, Nouf Ayadh Alqarni, Nadia Abdullah Almalki. Jeddah, Kingdom of Saudi Arabia : IEEE, 2020.
30. *Deep Learning Approach for Intelligent Intrusion Detection System.* R. VINAYAKUMAR, MAMOUN ALAZAB, (Senior Member, IEEE), K. P. SOMAN and others. Online : IEEE, 2019. 10.1109/ACCESS.2019.2895334.
31. *Man-in-the-middle-attack: Understanding in simple words.* Avijit Mallik, Abid Ahsan,Mhia Md. Zaglul Shahadat and Jia-Chi Tsou. s.l. : International Journal of Data and Network Science , 2019. 10.5267/j.ijdns.2019.1.001 .

ISSN: 2366-1313

32. *Survey of DoS/DDoS attacks on IoT.* Rozan Khader, Derar Eleyan. Palestine : Sustainable Engineering and Innovation ISSN 2712-0562, 2021, Vols. Vol. 3, No. 1, January 2021, pp.23-28.
33. *The Role of Lightweight Approaches Towards the Standardization of a Security Architecture for IoT Middleware Systems.* Ramao Tiago Tiburski, Leonardo Albernaz Amaral, Everton de Matos,Dario F. G. de Azevedo, and Fabiano Hessel. Brazil : IEEE Communication Magazine , 2016.
34. *Analysis of Signature Wrapping Attacks and Countermeasures.* Sebastian Gajek, Meiko Jensen, Lijun Liao, and Jorg Schwenk. Germany : IEEE, 2009.
35. *Phishing Environments, Techniques, and Countermeasures: A Survey.* AHMED ALEROUD, LINA ZHOU. jordan,BALTIMORE COUNTY : sciencedirect, 2017. https://www.sciencedirect.com/science/article/pii/S0167404817300810.
36. *Cross-site scripting (XSS) attacks and mitigation: A survey.* Germán E. Rodríguez, Jenny G. Torres, Pamela Flores, Diego E. Benavides. Ecuador : ELSEVIER , 2020.
37. *"Mobile network intrusion detection for IoT system based on transfer learning algorithm,".* L. Deng, D. Li, X. Yao, D. Cox, and H. Wang,. s.l. : Clust. Comput., vol. 22, pp. 9889–9904, Jan. 2018.
38. *Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset,.* N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull,. s.l. : Future Gener. Comput. Syst, 2019, Vol. 100. doi: 10.1016/j.future.2019.05.041.
39. *Computational Complexity of Machine Learning Algorithms.* online : Available: https://www.thekerneltrip.com/machine/learning/, 2018.
40. *"Random forest classification for detecting android malware,"* . M. S. Alam and S. T. Vuong. s.l. : in Proc. IEEE Int. Conf. Green Comput.Commun. IEEE Internet Things IEEE Cyber, Phys. Soc. Comput.,, Aug. 2013,. pp. 663–669..
41. *Experimental Analysis of Classification for Different Internet of Things (IoT) Network Attacks Using Machine Learning and Deep learning.* Tasnim, Anika, et al. Chiangrai, Thailand : IEEE, 2022. 10.1109/DASA54658.2022.9765108.
42. *Classification of IoT based DDoS Attack using Machine Learning Techniques.* Ashfaq, Muhammad Fasih, et al. Seoul, Korea, Republic : IEEE, 2022. 10.1109/IMCOM53663.2022.9721740.
43. *Ultra-Lightweight and Secure Intrusion Detection System for Massive-IoT Networks.* Roumaissa Bekkouche, Mawloud Omar, Rami Langar, Bechir Hamdaoui. Seoul, Korea, Republic : IEEE, 2022. 10.1109/ICC45855.2022.9838257.
44. *IoT-23: A labeled dataset with malicious and benign IoT network traffic.* S. Garcia, A. Parmisano, and M. J. Erquiaga. Zenodo : s.n., Jan. 2020. . doi: 10.5281/zenodo.4743746.
45. *"A cloud-assisted malware detection and suppression framework for wireless multimedia system in IoT based on dynamic differential game,"* . W. Zhou and B. Yu. no. 2,pp. 209–223,, China Commun., , : s.n., Feb. 2018., Vol. 15.
46. *a comprehensive data set for network intrusion detection systems (UNSW-NB15 network dataset)* . Moustafa N, Slay J. Canberra, ACT, Australia : In: 2015 military communications and information systems conference (MilCIS). IEEE, 2015.
47. *A Taxonomy of Machine-Learning-Based Intrusion Detection Systems for the Internet of Things: A Survey.* Murali, Abbas Jamalipour and Sarumathi. VOL. 9, s.l. : IEEE INTERNET OF THINGS JOURNAL, JUNE 15, 2022, Vols. NO. 12,. Digital Object Identifier 10.1109/JIOT.2021.3126811.

ISSN: 2366-1313

48. *An Efficient Spam Detection Technique for IoT Devices Using Machine Learning.* Aaisha Makkar, Sahil (GE) Garg, Neeraj Kumar, Shamim Hossain, Ahmed Ghoneim, Mubarak Alrashoud. 2, s.l. : IEEE, 2021, Vol. 17. 10.1109/TII.2020.2968927.
49. *N-BaIoT—network-based detection of IoT botnet attacks using deep autoencoders.* . Comput 17:12–22., s.l. : IEEE Pervasive, 2018. https://doi.org/10.1109/MPRV.2018.03367.
50. *"N-BaIoT: Network-based detection of IoT botnet attacks using deep autoencoders.* al., Y. Meidan et. Jul.–Sep. 2018, : IEEE Pervasive Comput. no. 3, pp. 12–22,, Vol. 17. doi: 10.1109/MPRV.2018.03367731.
51. *Review on Various Machine Learning Algorithms Implemented in IoT Security* . Lokesh Babu C, Vanitha M. 2022 Third International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT) : IEEE, 2022. DOI: 10.1109/ICICICT54557.2022.9917738.
52. *Using Machine Learning To Secure IOT Systems.* Mohammed Thakir Mahmood, Saadaldeen Rashid Ahmed Ahmed , Moahmmed Rashid Ahmed Ahmed. Istanbul, Turkey : IEEE, 2020. DOI: 10.1109/ISMSIT50672.2020.9254304.
53. *A comprehensive survey of neural architecture search: challenges comprehensive survey of neural architecture search: challenges* . Ren P, Xiao Y, Chang X, Huang P-Y, Li Z, Chen X, Wang X (2020). . arXiv preprint : s.n. https://arxiv.org/abs/2006.02903.
54. *Anomaly-based intrusion detection system in IoT using kernel extreme learning machine.* Sawssen Bacha, Ahamed Aljuhani, · Khawla Ben Abdellafou,Okba Taouali,Noureddine Liouane, Mamoun Alazab. Germany : Springer , 2022. https://doi.org/10.1007/s12652-022-03887-w.
55. *"Robust malware detection for Internet of (battlefield) things devices using deep eigenspace learning,"* . A. Azmoodeh, A. Dehghantanha, and K. R. Choo,. 1, s.l. : IEEE Trans. Sustain. Comput.,, Jan-March 2019, Vol. 4. pp. 88–95,.
56. *DIDS: A Deep Neural Network based real-time Intrusion detection system for IoT.* Monika Vishwakarma, Nishtha Kesswani. Ajmer, India : ELSEVIER, 2022. https://doi.org/10.1016/j.dajour.2022.100142.
57. *Attack prevention in IoT through hybrid optimization mechanism and deep learning framework* . Regonda Nagaraju, Jupeth Toriano Pentang,Shokhjakhon Abdufattokhov, Ricardo Fernando CosioBorda, N. Mageswari, G. Uganya. INDIA : ELSEVIER, 2022. https://doi.org/10.1016/j.measen.2022.100431.
58. *Nsl-kdd data set for network-based intrusion detection systems.* https://www.unb.ca/cic/datasets/nsl.html.
59. *LSTM based Deep Learning Technique to Forecast Internet of Things Attacks in MQTT Protocol.* S Thavamani, U. Sinthuja. Bengaluru India : IEEE, 2022. 10.1109/ICAECC54045.2022.9716585.
60. *Deep-Learning-Enabled Security Issues in the Internet of Things.* Lv, Zhihan, et al. 12, s.l. : IEEE, 2020, Vol. 8. 10.1109/JIOT.2020.3007130.
61. *Deep learning approach for intrusion detection in IoT-multi cloud environment.* D. Selvapandian, R. Santhosh. s.l. : springer , 2021. https://doi.org/10.1007/s10515-021-00298-7.
62. *A Comprehensive Survey on Transfer Learning.* Zhuang, Fuzhen, et al. 1, s.l. : IEEE, 07 July 2020, Vol. 109. DOI: 10.1109/JPROC.2020.3004555.
63. *Transfer-Learning-Based Intrusion Detection Framework in IoT Networks.* Eva Rodríguez, Pol Valls, Beatriz Otero, Juan José Costa , Javier Verdú , Manuel Alejandro Pajuelo. Barcelona, Spain : MDPI, 2022. https://doi.org/10.3390/s22155621.

ISSN: 2366-1313

64. *"A survey of machine and deep learning methods for Internet of Things (IoT) security,"* . M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani,. s.l. : IEEE Commun. Surveys Tuts.,, 2020, Vols. vol. 22, no. 3, pp. 1646–1685, 3rd Quart., .
65. *"Classifying IoT security risks using deep learning algorithms,"* . W. Abbass, Z. Bakraouy, A. Baina, and M. Bellafkih,. in Proc. 6th Int. Conf. Wireless Netw. Mobile Commun. (WINCOM), Marrakesh, Morocco, : IEEE, 2022.
66. *Machine and Deep Learning Approaches for IoT Attack Classification.* Alfredo Nascita, Francesco Cerasuolo, Davide Di Monda,Jonas Thern Aberia Garcia, Antonio Montieri, Antonio Pescape. Italy : IEEE, 2022. 10.1109/INFOCOMWKSHPS54753.2022.9797971.
67. *Design and Development of a Deep Learning-Based Model for Anomaly Detection in IoT Networks.* IMTIAZ ULLAH, QUSAY H. MAHMOUD. online : IEEE, 2021. 10.1109/ACCESS.2021.3094024.
68. *'IoT network intrusion dataset,''* . H. Kang, D. H. Ahn, G. M. Lee, J. D. Yoo, K. H. Park, and H. K. Kim. s.l. : IEEE Dataport, Tech. Rep, 2019. 10.21227/q70p-q449.
69. *"A review of intrusion detection systems using machine and deep learning in Internet of Things: Challenges, solutions and future directions," Electronics,* . J. Asharf, N. Moustafa, H. Khurshid, E. Debie, W. Haider, and A. Wahab,. s.l. : Electronics , 2020, Vols. vol. 9, no. 7, pp. 1–45, .
70. *DeepDefense: Identifying DDoS Attack via Deep Learning.* Xiaoyong Yuan, Chuanhuang Li and Li, Xiaolin. Hong Kong, China : IEEE, 2017. 10.1109/SMARTCOMP.2017.7946998.
71. *Detecting Security and Privacy Attacks in IoT Network using Deep Learning Algorithms.* Janardhana, D R, et al. Nitte, India : IEEE, 2021. 10.1109/DISCOVER52564.2021.9663586.
72. *Reinforcement Learning for IoT Security:A Comprehensive Survey.* Rawat, Aashma Uprety and Danda B. 11, s.l. : IEEE, June 2021, Vol. 8. Digital Object Identifier 10.1109/JIOT.2020.3040957.
73. *Q-learning and LSTM based deep active learning strategy for malware defense in industrial IoT applications.* Khuwaja, Sunder Ali Khowaja & Parus. online : Springer, 2020. https://doi.org/10.1007/s11042-020-10371-0.
74. *Investigation on identify the multiple issues in IoT devices using Convolutional Neural Network* . Swapna Thouti, Nookala Venu , Dhruva R. Rinku, Amit Arora, N. Rajeswaran. Hyderabad, India : ELSEVIER , 2022. https://doi.org/10.1016/j.measen.2022.100509.
75. *OWASP IoT Top 10 based Attack Dataset for Machine Learning* . Min, Nay Myat, et al. PyeongChang Kwangwoon_Do, Korea, Republic of : IEEE, 11 March 2022. DOI: 10.23919/ICACT53585.2022.9728969.
76. *MQTTset, a New Dataset for Machine Learning Techniques on MQTT.* Ivan Vaccari, Giovanni Chiola , Maurizio Aiello and Maurizio Mongelli, Enrico Cambiaso. s.l. : MDPI, 2020. doi:10.3390/s20226578.
77. *''Machine learning based IoT intrusion detection system: An MQTT case study (MQTT-IoT-IDS2020 dataset),'.* H. Hindy, E. Bayne, M. Bures, R. Atkinson, C. Tachtatzis, and X. Bellekens,. ' in Proc. 12th Int. Netw Conf. (INC) in Lecture Notes in Networks and Systems, vol. 180. Cham Switzerland: : Springer, , 2020, , Vols. 73–84, . doi: 10.1007/978-3-030-64758-2_6.
78. *A Scheme for Generating a Dataset for Anomalous Activity Detection in IoT Networks." In: Goutte C., Zhu X. (eds) Advances in Artificial Intelligence. Canadian AI Lecture Notes in Computer Science,.* Ullah and Q. H. Mahmoud. s.l. : Springer, Cham., 2020. , Vol. 12109. https://doi.org/10.1007/978-3-030-47358-7_52.
79. *Flow-based benchmark data sets for intrusion detection.* M. Ring, S. Wunderlich, D. Grüdl, D. Landes, and A. Hotho,. online : Journal of Information Warfare , 2017,

ISSN: 2366-1313

Proceedings of the 16th European Conference on Cyber Warfare and Security, pp. page 361--369. ISSN 1445-3312 Print/ISSN 1445-3347 .

80. *"IDS 2018 | Datasets | Canadian Institute for Cybersecurity | UNB." https://www.unb.ca/cic/datasets/ids-2018.html.*

81. *Federated TON_IoT Windows Datasets for Evaluating AI-based Security Applications.* Nour Moustafa, Marwa Keshk, Essam Debie, Helge Janicke. Australia : s.n., 2020. https://research.unsw.edu.au/projects/toniot-datasets.

82. *MedBIoT: Generation of an IoT Botnet Dataset in a Medium-sized IoT.* Alejandro Guerra-Manzanares, Jorge Medina-Galindo, Hayretdin Bahsi ,Sven Nõmm. s.l. : SCITEPRESS, 2020. DOI: 10.5220/0009187802070218.