IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 25, 2021, accepted April 12, 2021, date of publication April 15, 2021, date of current version April 23, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3073408

Internet of Things Applications, Security Challenges, Attacks, Intrusion Detection, and Future Visions: A Systematic Review

NIVEDITA MISHRA[1] **AND SHARNIL PANDYA**[2]
[1]Symbiosis Institute of Technology, Symbiosis International (Deemed) University, Pune 412115, India
[2]Symbiosis Centre for Applied Artificial Intelligence, Symbiosis International (Deemed) University, Pune 412115, India

Corresponding author: Sharnil Pandya (sharnil.pandya@sitpune.edu.in)

ABSTRACT Internet of Things (IoT) technology is prospering and entering every part of our lives, be it education, home, vehicles, or healthcare. With the increase in the number of connected devices, several challenges are also coming up with IoT technology: heterogeneity, scalability, quality of service, security requirements, and many more. Security management takes a back seat in IoT because of cost, size, and power. It poses a significant risk as lack of security makes users skeptical towards using IoT devices. This, in turn, makes IoT vulnerable to security attacks, ultimately causing enormous financial and reputational losses. It makes up for an urgent need to assess present security risks and discuss the upcoming challenges to be ready to face the same. The undertaken study is a multi-fold survey of different security issues present in IoT layers: perception layer, network layer, support layer, application layer, with further focus on Distributed Denial of Service (DDoS) attacks. DDoS attacks are significant threats for the cyber world because of their potential to bring down the victims. Different types of DDoS attacks, DDoS attacks in IoT devices, impacts of DDoS attacks, and solutions for mitigation are discussed in detail. The presented review work compares Intrusion Detection and Prevention models for mitigating DDoS attacks and focuses on Intrusion Detection models. Furthermore, the classification of Intrusion Detection Systems, different anomaly detection techniques, different Intrusion Detection System models based on datasets, various machine learning and deep learning techniques for data pre-processing and malware detection has been discussed. In the end, a broader perspective has been envisioned while discussing research challenges, its proposed solutions, and future visions.

INDEX TERMS Anomaly detection, DDoS attacks, deep learning, machine learning, Internet of Things, intrusion detection system.

I. INTRODUCTION

Internet of Things (IoT) is an emerging field of collection and transfer of data without human intervention. It is referred to as a system of connected objects embedded with sensors, software, control systems. Technological advances such as machine learning have resulted in the evolution of IoT technology [1]. IoT applications are increasingly making their presence felt in almost every area. Some of the widespread applications of IoT are shown in Fig. 1. In the present era, every sector is moving towards connected things to meet the world's pace. Education is not limited to the traditional way; now classrooms are connected, and with the use of technology, differently-abled students with hearing issues can learn using connected gloves and tablets. Similarly, IoT technology can turn out to be a significant boon for other disabled students. In fast-running lives, homes and cities are getting smart to fulfill humankind's basic needs such as security, waste management, air quality improvement, and entertainment [2].

The healthcare sector has transformed with the introduction of IoT, be it wearables or telemedicine and remote monitoring of patients [3]. IoT in agriculture has changed the way of traditional farming with better water management and soil monitoring [4]. IoT has been a game-changer in smart vehicles by introducing connected vehicles [5]. Also, in electric grids introduction of IoT has given energy management new heights [6]. IoT has evolved into this big industry after

The associate editor coordinating the review of this manuscript and approving it for publication was Jiafeng Xie.

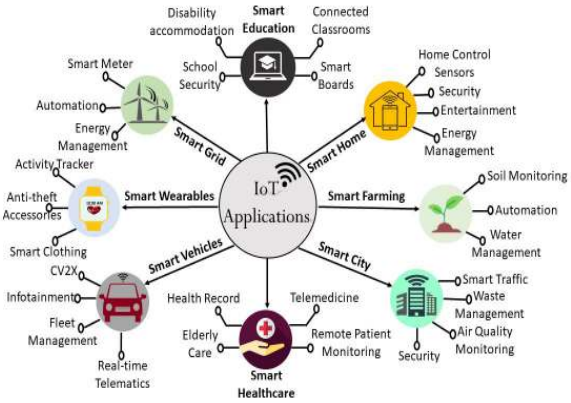

FIGURE 1. A representation of various IoT applications.

many advancements; the chronological advancement of IoT is depicted in Fig. 2. From the internet-enabled refrigerator to IoT-based smart city [7], the industry has come a long way, and it has become an integral part of everyday lives. IoT has marked its presence in the defense sector by introducing the Internet of Battlefields in 2017, and in the subsequent year, it was introduced in the healthcare sector. In the year 2005, the first Wi-Fi enabled rabbit was manufactured in Japan [8]. 2011 was a milestone year for IoT as it was added in the Hype cycle of emerging technologies; later on, it has been frequent in the Hype cycle and even recently in 2017, IoT remains at the peak of hype in Gartner cycle [9]. The focus is now shifting towards IoT security after the first large-scale IoT-based attack in 2016 [10].

Fig. 3 represents a variety of research challenges associated with growing IoT devices. The security protocols are not unified. Each device has a particular solution depending on the vendor. Also, as the IoT network is comprised of heterogeneous devices, protocols are not standardized. Traditional security methods do not work with IoT devices because they have a small processor, and adding security becomes impractical. Power resources and the location is also a concern for implementing security techniques. Mobility is a significant concern for connected cars, i.e., in smart automobiles [11]. In other applications where mobility is low, achieving trust is comparatively more comfortable than in fast-moving connected vehicles. IoT devices face significant Resource Constraint challenges as there are multiple facet constraints for IoT: cost, power, size. Heterogeneity of IoT devices is another major issue as most IoT applications work in a distributed environment with sensors, actuators, and other devices [12]. Table 1 represents various IoT-related open Issues and research challenges. These challenges make industry players skeptical about using IoT technology. Governments have not come up with a standard framework and regulation for IoT, giving freehand to service providers for implementing IoT. This creates a challenge in coming up with a common platform for any IoT solution. Expectations are on the rise from IoT in terms of performance with technological advancements, specifically artificial intelligence and data engineering.

TABLE 1. IoT research challenges.

Key Reference	Year	Research Challenges	Discussion
Pal *et al.* [13]	2018	Standardization and Regulatory Framework	Regulations and standardization are required for data ownership as data handling can even involve legal obligations in some cases like medical data.
Srivastava *et al.* [14]	2020	Security Requirements	Securing IoT devices becomes all the more difficult due to the range involved with different IoT devices and also varying issues of different IoT layers.
Pal *et al.* [13]	2018	Interoperability	Due to the diversity associated with heterogeneity involved in IoT, standard interfaces are significant for maintaining interoperability.
Ding *et al.* [15]	2020	Connectivity	Connectivity solutions are both licensed and unlicensed. Thus arises the need for a standard connectivity solution to address the decision-making issue of using specific connectivity solutions for IoT.

Similarly, interoperability is required as connected devices are increasing and common platforms are less for IoT. The connectivity issue is of concern as several highly critical IoT devices are involved in data transfer. Maintaining reliable data transfer for heterogeneous IoT devices poses a serious technical challenge.

The security requirement is essential in the case of IoT to maintain trust among consumers. Security management takes a back seat in IoT because of cost, size, and power. This consequently makes IoT vulnerable to security attacks, ultimately causing enormous financial and reputational losses. Research Contributions: In the literature, several surveys are available; key contributions of some of the highly cited research works are represented in Table 2. A detailed comparison of surveys and the proposed work is also depicted in Table 2. Comparison between papers is made based on IoT security, DDoS attack discussion, Intrusion Detection System, analysis of IDS datasets, and IDS techniques based on Machine learning and Deep learning techniques.

The undertaken work is a walkthrough from the initial discussion of IoT evolution and application to security issues in different layers and finally to various Intrusion Detection techniques. The undertaken study's scope is limited to the challenges posed by the security requirements of IoT technology. Some major contributions of the study are listed below:

1. Discussion of the evolution of IoT, applications, and challenges associated with IoT is presented.

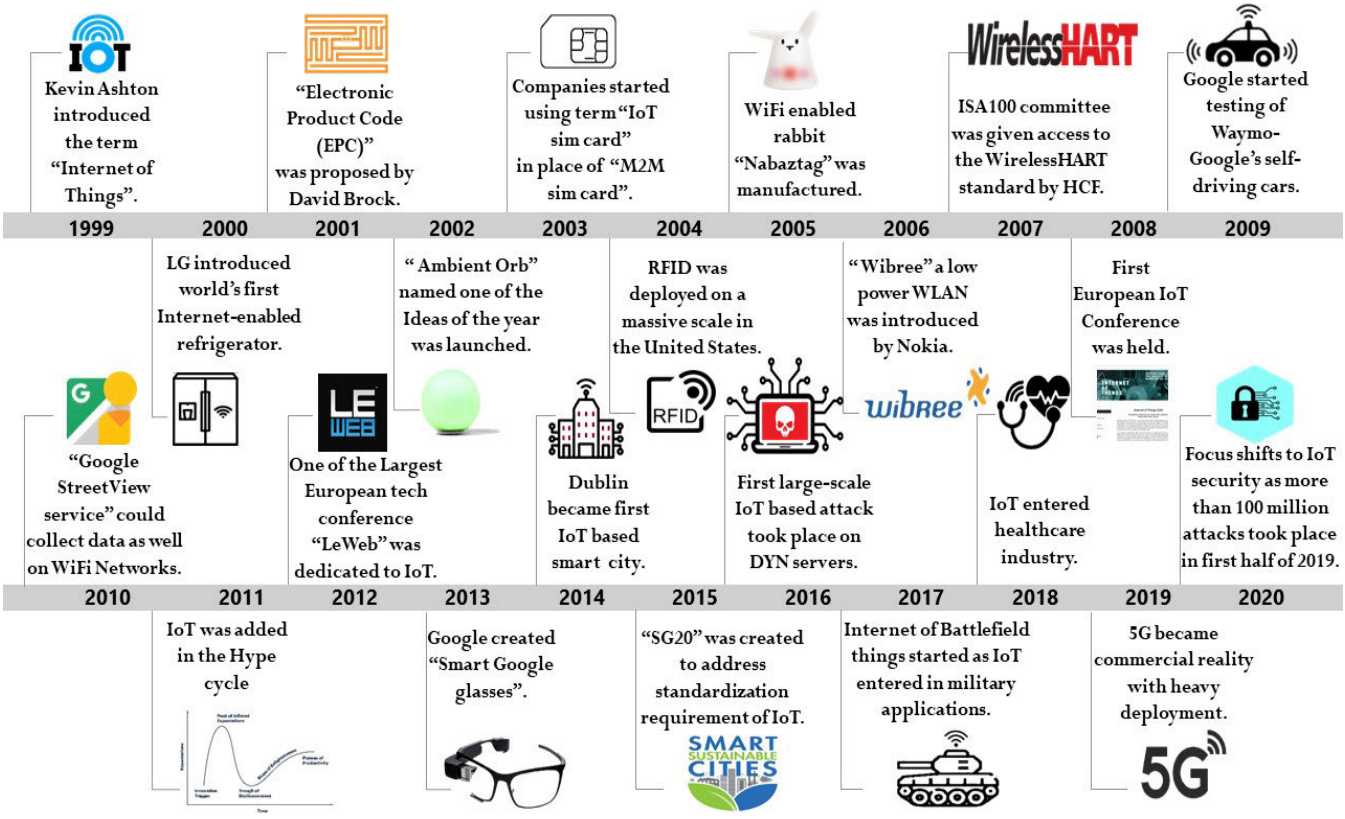

FIGURE 2. A chronological representation of the evolution of IoT technologies from 1999-2020.

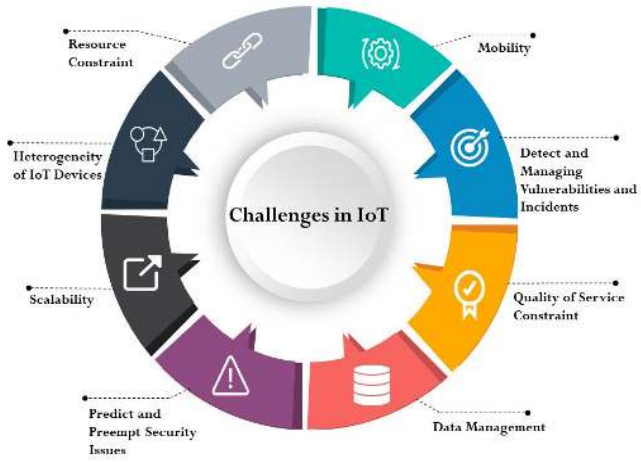

FIGURE 3. A graphical representation of various research challenges in IoT.

2. Security issues at different IoT layers are specifically discussed.

3. Risks associated with DDoS attacks and solutions for the same are analyzed in detail.

4. Several anomaly detection techniques are compared and analyzed.

5. Detailed review of recent Intrusion Detection System techniques is presented.

6. Finally, research gaps as understood from the survey and possible solutions for the same are proposed.

Methods and Materials: The roadmap of the conducted rigorous and detailed review is represented in Fig. 4. A systematic literature review is done in the undertaken study focusing on IoT security analysis, critical aspects of IoT security, and Intrusion Detection Systems in IoT. Papers from reputed journals and conferences, namely IEEE, Springer, Elsevier, Willey, and more, are considered while writing this survey to maintain quality of work. Scopus and web of science search engines are used for reliable results. Recent papers are focused upon keeping in mind that the IoT field is developing fast. For searching relevant articles, keywords used are DDoS attacks, IoT Security, Intrusion Detection System, Botnets, Machine Learning, Deep Learning.

Publisher and year-wise details of surveyed articles are depicted in Fig. 5 (a), and indexing-wise information for a similar duration is shown in Fig. 5 (b). Articles from reputed publishers, for example, IEEE, MDPI, Springer, and Elsevier, are explicitly focused, as evident from Fig. 5 (a). Science citation index (SCI) is also used for classifying surveyed articles, as depicted in Fig. 5 (b). Most of the investigated articles are Quartile-1 indexed and are from recent years, viz. 2020 and 2019.

Table 3 shows classified references based on different areas investigated in the proposed study. In the undertaken

TABLE 2. A detailed comparison of state-of-the-art surveys in the IoT security domain.

Authors	Year	Contribution	1	2	3	4	5	6
Yang *et al.* [16]	2017	The survey inspects four IoT security aspects: limitation of IoT devices and solutions, classification of IoT attacks, IoT authentication, and security attacks in different IoT layers.	✓	✓	×	×	×	×
Yu *et al.* [17]	2017	Edge computing and its use in IoT is thoroughly analyzed. Advantages and disadvantages associated with edge computing-based IoT are discussed.	✓	×	✓	×	×	×
Kouicem *et al.* [18]	2018	A top-down survey of IoT security solutions is conducted with focus on security solutions addressing resource constraints and scalability issues.	✓	✓	✓	×	×	×
Frustaci *et al.* [19]	2018	Different security issues and the availability of solutions for these issues are discussed in detail. Security issues raised due to communication protocols are also discussed.	✓	✓	✓	×	×	×
Noor *et al.* [20]	2019	New technologies related to IoT security, along with tools and simulators, are discussed in depth.	✓	×	×	×	×	×
Hassija *et al.* [21]	2019	A detailed review of security-related issues in IoT and discussion on emerging technologies for building a high trust level is presented.	✓	✓	✓	×	✓	×
Meneghello *et al.* [22]	2019	Security issues of different communication protocols and solutions are analyzed, particularly the weakness of commercial IoT solutions are discussed.	✓	×	✓	×	×	×
Srivastava *et al.* [14]	2020	Discussion on detection and defense against DDoS, Sybil, collusion attacks is presented along with different Intrusion Detection strategies.	✓	✓	✓	×	×	×
Anand *et al.* [23]	2020	Vulnerabilities associated with IoT in the backdrop of sustainable computing are analyzed, and a multifold study is presented, accompanied by a case study on smart agriculture.	✓	✓	✓	×	✓	×
The Proposed Survey	2021	Evolution of IoT, applications, and challenges associated with IoT, security issues in IoT are presented. Different types of DDoS attacks, their impacts, solutions, and anomaly detection are discussed in detail.	✓	✓	✓	✓	✓	✓

1. IoT Security issues, 2. DDoS attacks discussion, 3. Intrusion Detection System, 4. Database discussion for IDS, 5. Machine Learning Techniques for IDS, 6. Deep Learning Techniques for IDS

research, relevant work has been surveyed and acknowledged. A list of used terminologies is represented in Table 4.

The rest of the paper is organized as follows. Section II discusses security issues in the IoT domain and a detailed discussion of DDoS attacks focusing on DDoS attacks in IoT devices. Section III presents Intrusion Detection Systems, a comparative analysis of Intrusion Detection and Prevention systems, Anomaly detection techniques, and IDS performance metrics. A review of several steps involved in Intrusion Detection, namely data collection, data preprocessing, anomaly detection, is done in detail in Section IV.

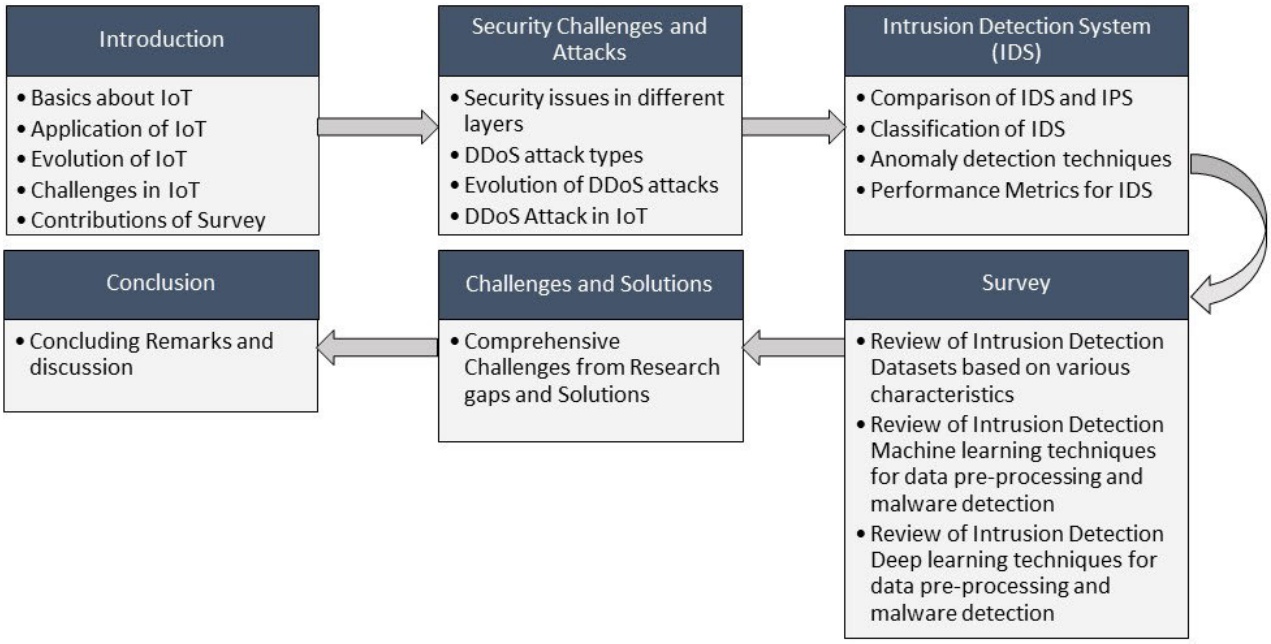

FIGURE 4. A roadmap of the conducted review work.

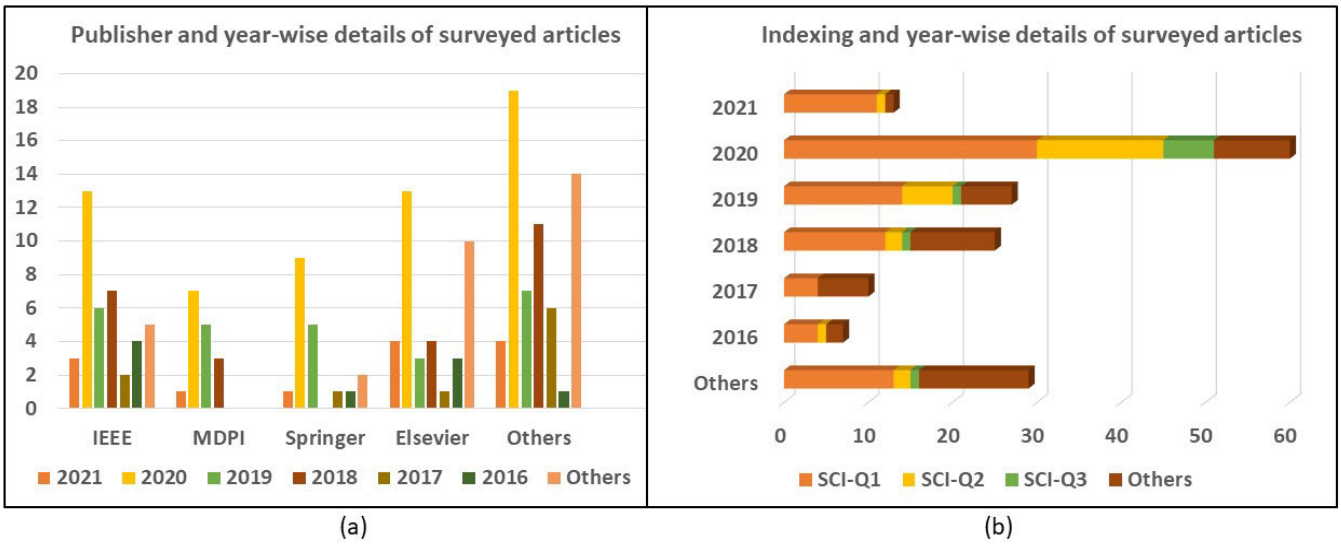

FIGURE 5. A graphical representation of year-wise details of surveyed articles (a) Publisher-wise analysis, (b) Indexing wise analysis for a similar duration.

Future direction for research is given by providing comprehensive research gaps from the survey and solutions proposed in Section V. Finally, concluding remarks are given in Section VI.

II. SECURITY ISSUES IN IoT DOMAIN

Internet of Things (IoT) devices are increasingly growing in numbers, and lack of security in these devices has resulted in transforming IoT devices into a hotbed for malicious activities [24]. Fig. 6 represents various cybersecurity attacks that can impact IoT layers such as the Perception layer, Support layer, Network layer, and Application layer. A widely accepted four-layered design is considered in the conducted review work [25].

PERCEPTION LAYER: The perception layer comprises sensors and actuators [26]. Sensors sense the environment around them while actuators act as controllers to take action based on sensed data. Sensors are also known as nodes, and these are vulnerable to node capturing attacks where an attacker may either capture the node or replace it with a malicious node. The over the air update of these nodes' firmware or software gives the attacker a chance to inject malicious or false code in the node causing Malicious code injection or False data injection attacks [27].

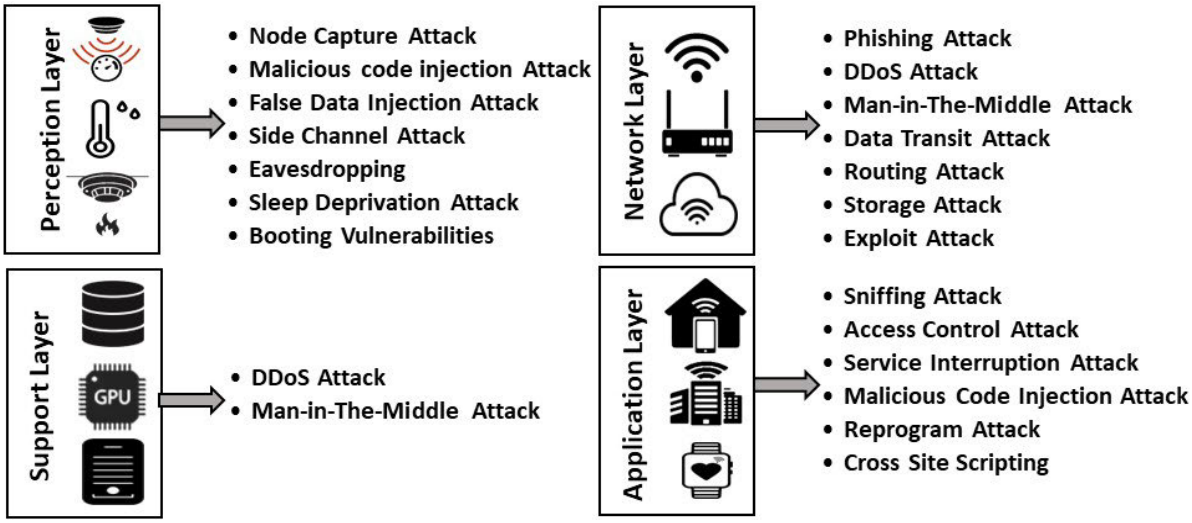

FIGURE 6. An illustrative representation of various security attacks in different IoT layers.

TABLE 3. Classification of referenced articles based on surveyed topics.

Surveyed Topic	Number of Papers
IoT Applications, challenges	12
IoT Security issues	15
DDoS attacks discussion	33
Intrusion Detection System	11
Database discussion for IDS	31
Machine Learning Techniques for IDS	32
Deep Learning Techniques for IDS	41
Generalized Papers	10

TABLE 4. A list of terminologies and abbreviations.

Terminology	Description
IoT	Internet of Things
IIoT	Industrial Internet of Things
QoS	Quality of Service
DDoS	Distributed Denial of Services
IDS	Intrusion Detection System
IPS	Intrusion Prevention System
CoAP	Constrained Application Protocol
ARMS	Apple Remote Management Services
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
DNS	Domain Name System
SNMP	Simple Network Management Protocol
WS-Discovery	Web Services Dynamic Discovery
NetBIOS	Network Basic Input Output System
SSDP	Simple Service Discovery Protocol
RIPv1	Routing Information Protocol version 1
LDAP	Lightweight Directory Access Protocol
NTP	Network Time Protocol
MQTT	Message Queuing Telemetry Transport
UPnP	Universal Plug and Play
Gbps	Gigabits per second
Tbps	Terabits per second

Side channel attack based on laser, power consumption, and timing can occur in this layer [28]. The nodes present in an open environment are vulnerable to eavesdropping attacks at the time of data transmission or similar events [29]. IoT devices are power constraint, and the attackers exploit this issue by draining the power source and causing Sleep deprivation. Typically, IoT devices' security process is enabled after booting, giving the attacker opportunity to launch an attack at boot time.

NETWORK LAYER: The network layer sends information from the sensing layer for further processing to the computational unit. This layer is highly vulnerable to attacks comprising several IoT devices [30]. Phishing attack targets several IoT devices in an attempt to at least take control of a few of them [31]. In a DDoS attack, an attacker tries to overwhelm the target by sending spoofed requests. IoT devices act as botnets in DDoS attacks and can create a massive flood of requests to deny the target further access to resources [32]. Worm-hole, Sinkhole attacks are examples of Routing attacks in which the attacker tries to route the traffic to a different path by gaining access to nodes [33]. At the time of data transfer, IoT devices are vulnerable to Data Transit attacks as data is critical and data breach is easier at the data transfer stage.

TABLE 5. A representation of various type of DDoS attacks.

Type of Attack	Goal of Attacker	Size Measured in	Examples	Existing Solution for preventing the Attack
Volumetric Attack	To consume all bandwidth between target and internet.	Bits per Second (bps)	NTP Amplification, UDP Flood, TCP Flood, DNS Amplification	Traffic is diverted to on-demand scalable scrubbing centers, where it can be absorbed.
Protocol Based Attack	To consume resources of server, firewall, load balancers.	Packets per Second (pps)	TCP SYN Flood, Ping of Death, Smurf Attack	Identification technique is used commonly to differentiate between legitimate and illegitimate traffic to block the attack before reaching the target server.
Application Layer Attack	To exhaust target resources.	Requests per Second (rps)	HTTP Flood, DNS Flood	These are generally slow attacks and mitigated by identifying bot behavior using captchas and similar techniques.

SUPPORT LAYER: The Support layer acts between Network and Application layer. This layer helps in resource allocation, computing, and data storage. Security of database is essential at this layer, and it is prone to DDoS, Man-in-the-middle, SQL injection kind of attacks. A broker like MQTT protocol is used for communication between client and service provider. In a Man-in-the-middle attack, the attacker takes control of the broker, thus controlling all the communication [34]. The target of attack in the Support layer is usually to access data; therefore, database and cloud security are crucial in this layer.

APPLICATION LAYER: The application layer contains smart applications viz. smart city, smart home, healthcare, and more. This layer directly deals with end-users; hence privacy and data theft are major concerns at this layer [35]. Similar to other layers, this layer is also affected by the Malicious code injection attack. A service interruption attack is similar to a denial of service attack as it causes service disruption. Some users are given the unique privilege to allow legitimate user access at the time of an attack, but the whole system can come under attack if this access is compromised. This makes access control attack one of the major concerns at the application layer [36]. Sniffing attack takes place with the help of sniffing tools where attacker sniffs network traffic data, and confidential data is compromised in this attack [37].

A. DISTRIBUTED DENIAL OF SERVICES (DDoS) ATTACKS

The DDoS attack, as the name suggests, is launched to overwhelm the target and disrupt services. The DDoS attack requires a large number of devices for launching an attack, and for this, IoT devices are well suited. As in most cases, users will not understand that the device is compromised; for example, baby monitors and smart toys have a user interface with limited access. They may generally work even after being part of a Botnet army. With the increasing volume of IoT devices, there is an urgent need to detect attacks timely to remove compromised devices. IoT devices were used as Botnets by Mirai in a significant DDoS attack, and also several such attacks have taken place [38]. Table 5 represents a comparative analysis illustrating three categories of DDoS attacks.

The severity of DDoS attacks can be understood from Fig. 7, which comprises major DDoS attacks from 2013 to 2020. Leading service providers like Amazon Web Services have been the victim of DDoS attacks. KrebsOnSecurity, Cloudflare, AWS, and more DDoS attack victims are themselves security providers against such attacks [39], [40]. Therefore, as shown in Fig. 7, attacks on these major establishments hamper companies financially and impede their reputation. In a Denial of Service (DoS) attack, the attacker tries to disrupt the services of the target by utilizing its resources with the help of fake requests.

Distributed Denial of Services (DDoS) is an amplified DoS attack. In a DDoS attack, requests are initiated from many sources, and hence it is named as distributed DoS. Due to this, it becomes challenging to mitigate DDoS attacks. There are many types of DDoS attacks: TCP SYN Flood attack, Teardrop attack, Smurf attack, Ping of Death attack, Botnets. DDoS attacks can also be classified as Reflection and Amplification attacks. In a reflection attack, the size of the request and response is the same [41], whereas, in an amplification attack, the size of the response is many times bigger than that of the request [42]. In Table 6, the chronological evolution of DDoS attack vectors is depicted.

B. DDoS ATTACK IN IoT DEVICES

DDoS and DoS attacks differ in the attack surface used to launch the attack. In a DDoS attack, multiple systems are used to launch the attack. These systems might be desktops, servers, IoT devices, and other connected devices [56]. Gaining access to these devices is the first step to launch an attack. The attackers exploit vulnerabilities of devices for taking control. In IoT, there are several security concerns that attackers are actively abusing [57].

Common vulnerabilities exploited by attackers for launching an attack using IoT:

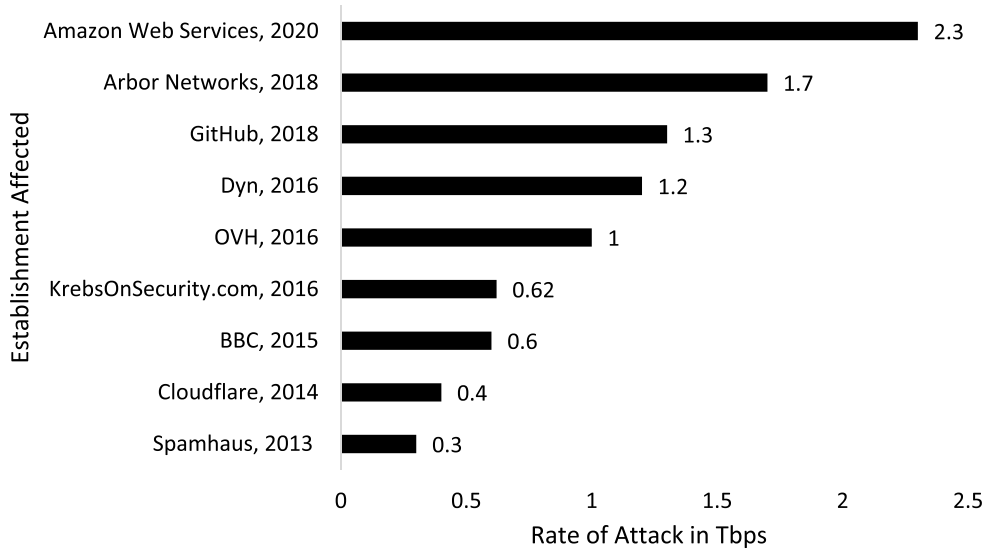

FIGURE 7. A bar chart representation of DDoS Attacks on major establishments for 2013-2020.

TABLE 6. A representation of the evolution of DDoS attack vectors.

Key Reference	Attack Vector with Year	Amplification Factor	Working Methodology
Hesselman *et al.* [43]	DNS (2013)	28-54	The functionality of open DNS resolvers is used to launch the attack.
Kawamura *et al.* [44]	NTP (2014)	556.9	The monlist command enabled NTP servers are abused to launch the attack.
Gondim *et al.* [45]	SSDP (2014)	30.8	UPnP protocol is exploited in SSDP reflection attacks.
Vasques *et al.* [46]	SNMP (2014)	6.3	Directly exposed servers with SNMP service are exploited under this attack.
Wisam *et al.* [47]	RIPv1 (2015)	131.2	Routers running RIPv1 with multiple routes are exploited in this attack.
Burch *et al.* [48]	CHARGEN (2015)	358.8	Internet-enabled devices running CHARGEN can be exploited to launch amplified attacks.
Noor *et al.* [49]	NetBIOS (2015)	3.8	Servers with open NetBIOS service are exploited for DDoS attacks.
Sieklik *et al.* [50]	TFTP (2016)	60	The protocol was intended for file transfers; its simple design omitted authentication capabilities and was exploited for the attack.
Choi *et al.* [51]	CLDAP (2017)	56-70	The gaming industry was the most affected by this attack. Servers with open UDP port 389 were targeted.
Agathe *et al.* [52]	Memcached (2018)	10000-51000	Attackers target Memcached servers with open TCP and UDP ports on 11211 to launch a DDoS attack.
Kondoro *et al.* [53]	CoAP (2018)	34	UDP garbage flood is created using IoT devices as amplifiers for launching this attack.
Malaimalavathani *et al.* [54]	WS Discovery (2019)	10-500	Being a UDP-based protocol, attackers use this to launch UDP flood attacks.
Bjarnason *et al.* [55]	ARMS (2019)	45	Operational management of Apple Remote Desktop (ARD) protocol running on UDP port 3283 was used to launch a DDoS amplification attack.

1. INSECURE CONNECTION: Connected devices are used for launching attacks, and it becomes a cakewalk for the attacker to infect a device when there is no firewall at work.

2. WEAK PASSWORD: Brute force attack comes into action when passwords are weak. In IoT, especially companies tend to keep the default password same for all the devices, and if not changed by the device user, it becomes effortless for the attacker to infect the device. Due to this practice, attackers try to find out more devices from the same manufacturer after successfully infecting some of a particular manufacturer's devices.

3. FIRMWARE UPDATES: Most IoT devices remain insecure because of outdated firmware. In some cases, the firmware update is not secure, which can leave the device unprotected.

4. SOFTWARE VULNERABILITIES: Software-related vulnerabilities remain on the watch list of attackers before the manufacturer releases any patch. These vulnerabilities are

TABLE 7. A detailed analysis indicating why IoT is preferred over other devices for DDoS attacks.

Parameter	Other Devices	IoT Devices
Maintenance	Servers require maintenance from its handler.	Almost no maintenance is required for IoT devices.
Security	Servers, laptops, and other similar devices are usually challenging to infect because of user awareness for security.	IoT devices are not so user-friendly, and people tend to neglect the security of these devices because of ignorance, making them more vulnerable to attacks.
Updates	Servers and similar devices are updated regularly and follow security protocols.	Firmware updates are rarely provided for IoT devices, and also, mostly these updates do not follow secure protocols resulting in insecure IoT devices.
Access	Power and internet services to these devices are limited; subsequently, access gained by the attacker also gets affected.	Often IoT devices work on very low power and remain connected to the internet, for example, CCTV, refrigerators, etc. This provides uninterrupted access to the attacker.

usually exploited by skilled attackers as these are difficult to identify and exploit.

5. VULNERABILITIES IN DATA HANDLING: Data transfer is a crucial step for connected devices, and at this phase, any loophole in the connection to the cloud or server can very well be exploited by the attacker.

A detailed analysis indicating the cause of preference of IoT devices over other devices is presented in Table 7. Conventional devices are usually secure as compared to IoT devices because of traditional security practices. This specifies the reason behind the drastic increase in IoT attack surface.

For launching a DDoS attack, the attacker runs malware scripts and tries to find out vulnerable devices. This process is similar in the case of IoT and other devices. It is a bit easier for the attacker to infect IoT devices because of a lack of security [58]. Once the attacker achieves access to these devices, unpretentious devices get mutated to bots, and the collection of these devices is called a botnet [32]. This botnet army can further be used for launching an attack by the master, or the master can decide to sell the army of bots for the use of other malicious actors. The DDoS attack, as shown in Fig. 8, takes place when multiple requests are sent using a botnet army, ultimately overwhelming the victim.

TABLE 8. Comparative analysis of IDS And IPS systems.

Intrusion Detection System (IDS)	Intrusion Prevention System (IPS)
IDS recognizes the threat and monitors the system.	IPS is a regulatory system as it takes monitors and defends the system.
Human intervention is required for action.	IPS takes action based on the rule set, and no human intervention is required.
IDS does not impact system performance.	IPS may slow down the system.
False alarm rate does not impact performance to the same extent as that for IPS.	False alarm rate is of high concern.
Legitimate users are not blocked as the system does not take action.	Legitimate traffic might be blocked due to false alarms.

All the servers are designed to handle a particular number of requests at a time, and attackers try to reach this threshold and exceed that by an enormous amount to overwhelm and ultimately deny the server to perform further [59]. IoT devices are exploited by attackers because of software, hardware, physical vulnerabilities. Along with these vulnerabilities, IoT devices use protocols that are vulnerable to DDoS attacks, viz. CoAP, RPL, 6LoWPAN [60].

III. INTRUSION DETECTION SYSTEM

As we move into the era where almost everything being used by humans will be connected to the internet, the security of these devices becomes paramount. Two major solutions are found in the literature for preventing DDoS attacks, namely Intrusion Detection System (IDS) and Intrusion Prevention System (IPS). IDS is a precautionary measure where the system itself takes no action in case of intrusion; instead, an alarm is raised. IPS is the punitive measure where an action is taken by the system in case of intrusion [61]. In IPS, an issue arises in the case of false positives as legitimate users can also get blocked. Table 8 represents a detailed comparative analysis of IDS and IPS systems.

Further in this study, IDS is focused upon as false alarm rate is a significant concern in malware classification. Also, punitive action taken on legitimate users can ruin the whole reason for creating a detection system. Fig. 9 represents a graphical representation of the category of Intrusion Detection Systems (IDS). Based on the target location, it can be classified as Host-based, Network-based, or Hybrid IDS. Host-based IDS is specific to a system, detection of an inside intruder is strong, and it can very well assess the extent of the compromise, but it is expensive as one IDS is required per host [62]. In Network-based IDS, the outside intrusion is very well detected, and it can protect all hosts, but there is too much traffic to analyze [63]. Hybrid IDS is flexible and provides more security as it combines features of both Host-based and

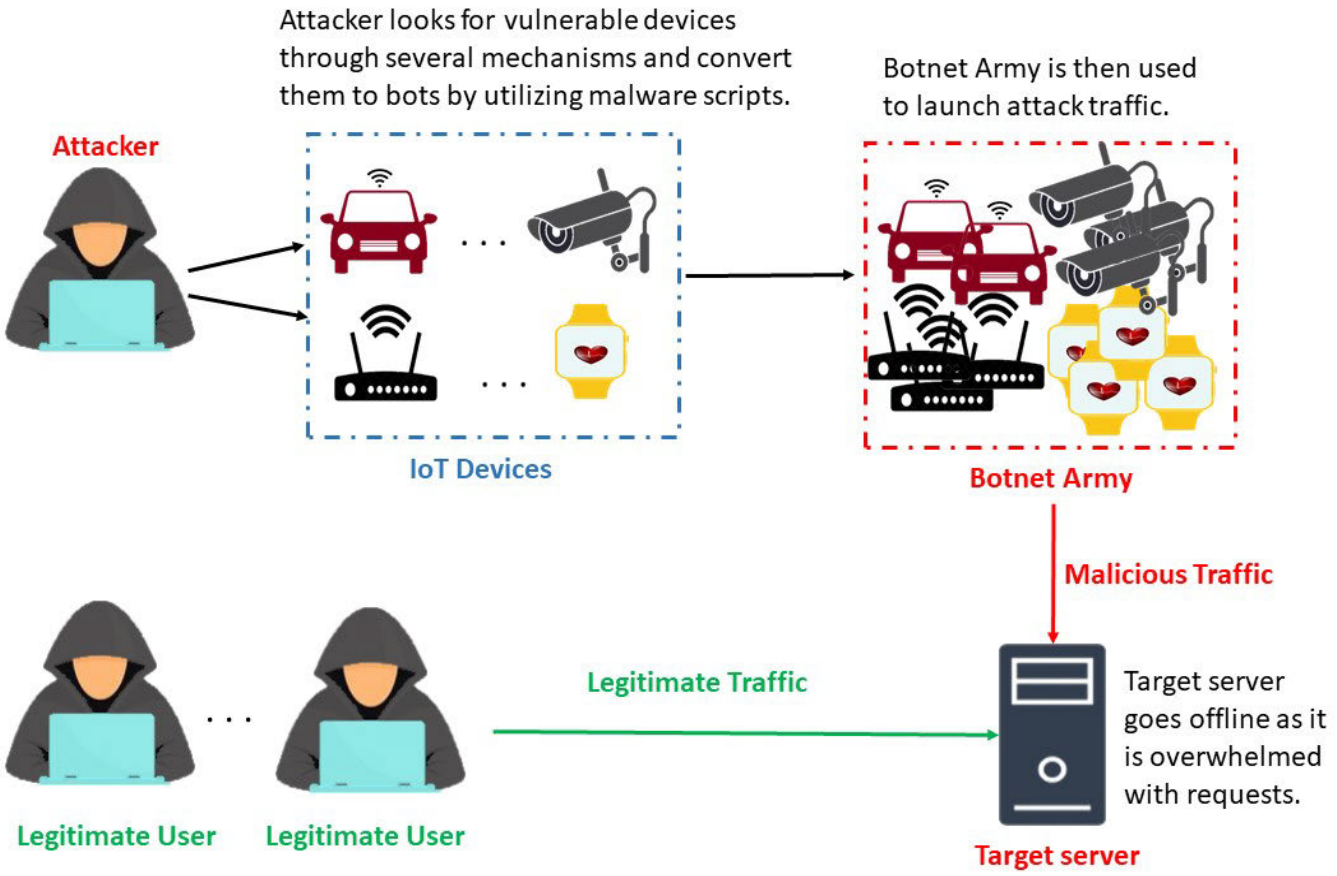

FIGURE 8. A representation of attacker gaining access to IoT devices and launching DDoS attacks.

Network-based IDS [64]. In Active IDS, definite action is taken for certain alerts, whereas only reports are generated, or alarms are raised in Passive IDS. Centralized IDS use individual monitors for monitoring each host, as it does not scale according to requirement, thus providing less flexibility. Moreover, centralized IDS is prone to a single point of failure. Distributed IDS, on the other hand, works as a Peer-to-Peer (P2P) architecture, and in this case, each monitoring unit doubles up as an analysis unit as well.

A. ANOMALY DETECTION TECHNIQUES

Two major approaches being used for Malware Detection are Signature-based Detection and Anomaly-based Detection. The signature-based detection technique [65] is not reasonably successful in the case of Botnets as Botnets keep on mutating, and as a result, Bot signature also keeps on changing. These techniques are not useful in real-world situations when the target is to detect new variants of Botnets. Anomaly-based detection techniques [66], on the other hand, are quite popular, as these techniques presume that Botnet traffic will be behaviorally different from normal traffic.

There are some other approaches, like Community Base Anomaly Detection [67], where Bots are identified using Communication Graph; for this to work accurately, a full graph should be available. Particular protocols/structures are also used for detection in a research study, but this approach is not practical if the same structure is not operated by other Botnets [68]. Bad Neighborhood is also one of the methods used in Spam and Phishing Detection; it is defined as a cluster of IP addresses that perform malicious activities over a certain period. Moura *et al.* [69] used this approach for IPv4 attacks and generated a blacklist of IPs.

This approach is not entirely practical as DDoS attacks are widespread, and it is challenging to assign clusters for blacklisting. Another method is whitelisting IPs, as Yoon [76], where a VIP list is created assuming that VIPs will log in from a particular IP address, i.e., IP address not very dynamic for personal laptops. This way, critical people can still use the websites which are under DDoS attack. This approach is also not very useful as it may only benefit a small set of people. Several Anomaly detection techniques are present in the literature. Table 9 represents a detailed comparative analysis of different Anomaly detection techniques. As seen from the Table, the machine learning-based approach is advantageous and is accepted globally for a wide range of malware classification. Therefore, the undertaken study focuses on the machine learning-based approach.

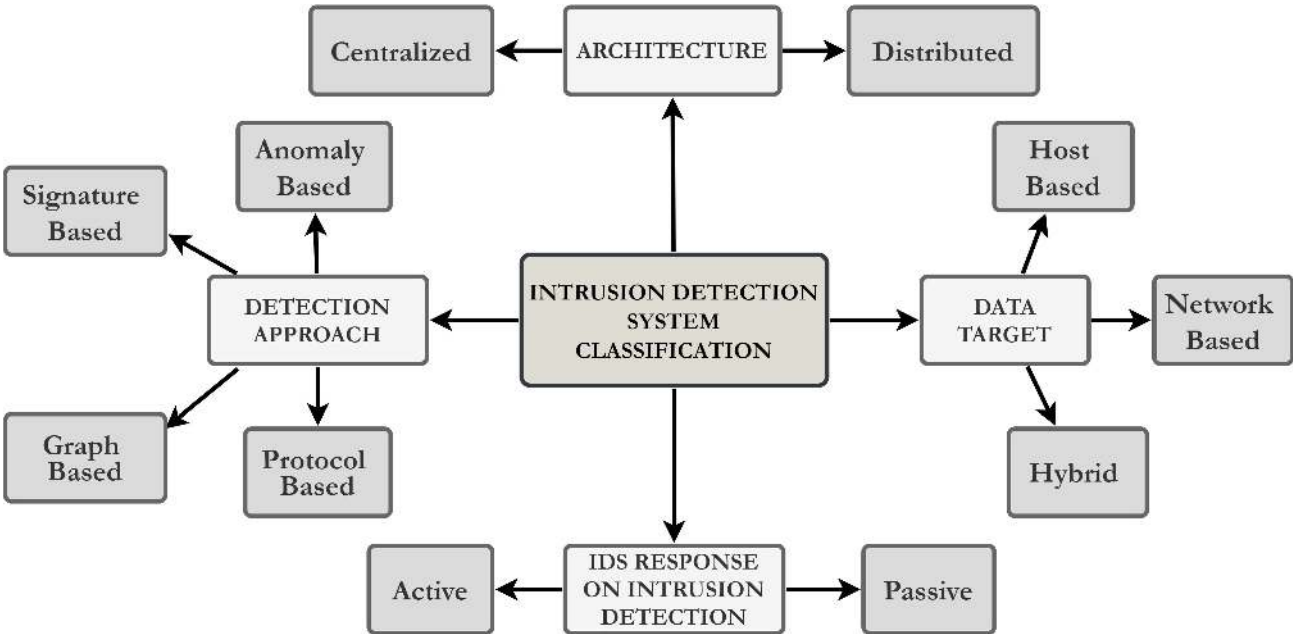

FIGURE 9. A graphical representation of the classification of various IDS techniques.

		Actual: Positive	Actual: Negative
Predicted	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

FIGURE 10. Confusion matrix.

B. PERFORMANCE METRICS FOR IDS

Some of the predominantly used performance metrics for Intrusion Detection Systems are discussed below.

1) CONFUSION MATRIX

Confusion Matrix (CM) is not directly a performance measure in itself. Still, it is one of the most instinctive metrics for defining a classification model's correctness. Almost all performance metrics are computed using CM parameters.

In Confusion Matrix, there are two ways to reduce errors: reducing False Negatives and reducing False Positives. There is no set rule for the same, and it depends on the requirement. For instance, for email spam classification, False Positives should be minimized, and for cancer patient classification, false negatives should be minimized.

2) ACCURACY

Accuracy is defined as the number of correct predictions over total predictions. This metric is ideal for use in the case of a balanced dataset. When there is a majority class in a dataset, the results provided by this metric may not reflect the model's actual performance.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FN + FP} \cdots \quad (1)$$

3) PRECISION

Precision is a measure to calculate the Machine Learning Model's accuracy in finding the number of actual positives out of total predicted positives. This metric is useful when False Positive is of high cost for Model quality, for example, email Spam Detection Model.

$$\text{Precision} = \frac{TP}{TP + FP} \cdots \quad (2)$$

4) RECALL/SENSITIVITY

Recall is a measure to calculate the Machine Learning Model's accuracy in finding the number of positives out of total actual positives. This metrics is useful when False Negative is of high cost for Model quality, for example, Fraud Detection Model.

$$\text{Recall} = \frac{TP}{TP + FN} \cdots \quad (3)$$

5) F-1 SCORE

It is calculated as a Harmonic Mean of precision and recall metrics to better evaluate model performance. This is a metric of importance for an imbalanced dataset as in this; equal importance is given to both Precision and Recall.

$$\text{F-1 Score} = \frac{2 \times Precision \times Recall}{Precision + Recall} \cdots \quad (4)$$

6) SPECIFICITY

Specificity is the opposite of Sensitivity (Recall), and it is a measure of False Positive Rate.

$$\text{Specificity} = \frac{TN}{TN + FP} \cdots \quad (5)$$

TABLE 9. Comparative analysis of anomaly detection techniques.

Key Reference	Technique	Examples	Approach	Advantage (+) / Disadvantage (-)
Hai *et al.* [70]	Machine Learning Model	Neural Networks, Genetic Algorithms, Clustering, Classification, Outlier Detection	• Machine Learning models have two stages: Training and Testing. • Broadly it can be divided into two categories: Supervised Learning and Unsupervised Learning.	+ Can identify patterns quickly. + Wide application range. + Suitable for online datasets. + Can improve continuously. - Long training time is required. - Larger dataset is needed for better results.
Kalkan *et al.* [71]	Statistical Model	Univariate, Multivariate, Time Series, Markov Process	• The statistical model approach depends on mathematical calculations. • Model is created for normal behavior from the historical data of the user.	+ Model is simple. - Model depends heavily on statistical or mathematical modeling, thus affects accuracy.
Kim *et al.* [72]	Payload Based Model	Grained Model, N-gram analysis,	• The Payload-based approach model learns characteristics of the normal packet payload, and deviation is considered anomalous behavior.	+ Model works very well for known attacks. - Longer handling time because of computational overhead. - Lesser accuracy is achieved for new attacks.
Kumar *et al.* [73]	Rule-Based Model	Association Rule, Fuzzy Rule, Behavior Rule	• In a Rule-based model, rules are created from data traffic patterns. • If the rule is broken, it is considered anomalous behavior.	+ Model is simple. - Longer monitoring of traffic is required for rule creation. - High false-positive rate.
Choudhary *et al.* [74]	Protocol Based Model	Application protocol, Communication Protocol	• The protocol-based model works with monitoring protocols at different layers. • Computing techniques are used to identify anomalies associated with a particular protocol.	+ Detection accuracy is high for a particular type of attack. - Suitable for specific attack type so fails for other attacks.
Choraś *et al.* [75]	Signal Processing Model	Wavelet, Entropy Analysis	• Signal processing methods are used for analyzing traffic. • The normal pattern is drafted from the traffic pattern, and deviation from the same is considered an anomaly.	+ Detection accuracy is high with lower false-positive rate. - Method for identifying patterns is complex.

7) AUC-ROC CURVE

Receiver Operating Characteristic (ROC) curve is a measure to determine the stability between precision and recall by a varying threshold. The Area Under Curve (AUC) represents the quality of the classification model.

$$\text{True Positive Rate (TPR)} = \frac{TP}{TP + FN} \ldots \quad (6)$$

$$\text{False Positive Rate (FPR)} = \frac{FP}{FP + TN} \ldots \quad (7)$$

It is a curve between TPR, i.e., Recall (Sensitivity), and FPR, i.e., (1 – Specificity). In general, AUC near to one represents a better classification model. Several other metrics are also used in the performance evaluation of IDS, namely the KAPPA metric, Root Mean Square Error (RMSE), and many more, depending on the requirement.

IV. REVIEW OF STEPS INVOLVED IN IDS

A. REVIEW OF VARIOUS DATASETS IN IDS

Data can be collected in two ways in IDS: first by using existing datasets else by creating own dataset. There are several datasets being extensively used for Anomaly Detection. KDD-99 [77] is one of the oldest and extensive dataset, and despite it being highly imbalanced, it is used even now due to the lack of its alternatives. NSL-KDD [78] was created to remove the issues associated with KDD-99. It is one of the benchmark datasets used for Anomaly detections regardless of whether it is obsolete, as the attacks incorporated in this dataset are mostly outdated. The skewedness of KDD-99 and NSL-KDD is almost removed in UNSW-NB15 [79], [80], consisting of 49 features and 10 target classes, whereas KDD consists of 41 features and 5 target classes. For Botnets, CTU-13 [81] having 13 scenarios; each of different Botnet samples is being used nowadays. A separate malware is executed for each capture, which performs different actions using different protocols. ISOT is also one of the popular datasets [82], particularly for IoT Botnet attack databases. A detailed review of some of the popular existing network-based datasets is represented in Table 10. There are some tools and techniques available for creating own dataset in the literature. Wireshark is used to capture and present network traffic. Detection rates achieved are high as DoS tools produce quite predictable traffic. Spleen is a software tool for creating a dataset similar to DARPA [95]. In this application, some additional features can also be added for better functionality. CICFlowmeter [96] is a java based tool used for extracting network features from raw network captures. It captures a set of 80 features and prepares a pcap or CSV file to be used for further analysis. Sharafaldin *et al*. [96] created two networks: 1.) Victim-Network consisting of 3 servers,1 firewall, 2 switches, and 10 PCs, 2.) Attack-Network includes one router, one switch, and four PCs. This kind of network is generally used for data generation.

Dataset generated in this work contains attacks based on McAfee report 2016. This dataset's shortcoming is the number of systems used is less, and it needs to be updated for more recent attacks, and more Botnet attack data should be used. Commercial IoT devices were infected in the lab using Mirai and Bashlite Botnets by Meidan *et al.* [97] to launch an attack. The network is sniffed using port mirroring on the central switch, and data is recorded using Wireshark. Djanie *et al.* [98] used eight DoS attack tools for launching an attack on a Virtual Ubuntu 19 Machine. For identifying a suitable dataset for research work, the number of records plays a crucial role. Fig. 11 depicts the bar chart representation of the number of attack records in Percentage. As shown in Fig. 11, some of the datasets like KDD-99 consist of many attack records, i.e., 80.14% in the training set and 80.52% in the test set. Whereas some of the datasets like CDX 2009 consists of comparatively smaller number of attack records, i.e., 0.76%.

B. REVIEW OF VARIOUS MACHINE LEARNING TECHNIQUES IN IDS

Data pre-processing comprises several steps: adding missing values, normalizing data, removing unwanted features/outliers. Feature analysis and extraction is the backbone of any Machine Learning Model. For feature extraction, different optimization techniques are used by researchers: Principal Component Analysis (PCA), Genetic Algorithms (GA), and Boosting Algorithms. Machine learning techniques are mainly used for feature engineering as they are lightweight and less complicated [99]. Botnet analysis has two major subdivisions, specifically Flow-based traffic analysis and Graph-based traffic analysis [100]. These analyses differ mainly on the feature selection part as statistical features are selected for Flow-based analysis; otherwise, Graph-based features are chosen. Usually, the same set of features are extracted for different time windows to create a real-time scenario [97]. These are statistical features and can be highly useful for capturing malicious activities as features indicate anomaly if the behavior is unseen, which happens in the case of a spoofed IP address. A botnet can be Detected using Graph-based features, as done by Chowdhury *et al.* [101]. In this detection, efficiency was improved by removing inactive nodes, and detection methodology was given, where, by using only six nodes, Bots can be detected effectively. Features are selected by Ghasemi *et al.* [65] using a genetic algorithm. For each class, a two labeled dataset is created, which is then given to classifiers, and after applying the voting mechanism best classifier is predicted, and a new dataset is created where number of classes is equal to the number of columns. This dataset is further used for the training model; this way model will be trained for different behaviors of all classes. Some of the critical data pre-processing techniques based on machine learning are discussed in Table 11.

In Intrusion Detection System, classification and clustering algorithms are used extensively. On network traffic, data classification algorithms such as Support Vector Machine (SVM) [111], [112], Decision Tree [113], [114], KNN [100] have given important and powerful results. Incremental Learning is one of the recent techniques being used with machine learning techniques for real-time results. It is often used in image classification [116], target recognition [117], and used less in Intrusion Detection or information security. SVM is one of the most prevalent techniques being used in Incremental Learning. An incremental SVM Learning algorithm based on Incremental Clustering for Intrusion Detection is proposed by Du *et al.* [118]. Yi *et al.* [119] used a kernel function modified for incorporating Mean and Mean Square

TABLE 10. A detailed review of various network-based datasets.

Datasets	Year	Dataset Publically Available	Nature of Traffic		Nature of Data	Labeled Dataset	Balanced Dataset	Network Type	Traffic Type
			Normal	Attack					
DARPA [83], [84]	1998-99	Yes	✓	✓	Packet, Logs	✓	×	Small Network	Emulated
KDD-99 [77]	1998-99	Yes	✓	✓	Other	✓	×	Small Network	Emulated
UNIBS [85]	2009	On Request	✓	×	Data Flows	×	×	University Network	Real
CDX [86]	2009	Yes	✓	✓	Packet	×	×	Small Network	Real
ISOT [82]	2010	Yes	✓	✓	Packet	✓	×	Small Network	Emulated
ISCX [87]	2012	Yes	✓	✓	Packet, Bidirectional Flows	✓	×	Small Network	Emulated
CTU-13 [81]	2013	Yes	✓	✓	Uni and Bidirectional Flows	✓	×	University Network	Real
Botnet [88]	2014	Yes	✓	✓	Packet	✓	×	Various Network	Synthetic
UNSW-NB15 [79]	2015	Yes	✓	✓	Packet, Other	✓	×	Small Network	Emulated
AWID [89]	2015	On Request	✓	✓	Other	✓	×	Small Network	Emulated
NDSec-1 [90]	2016	On Request	×	✓	Packet, Logs	✓	×	Small Network	Emulated
CICDOS [91]	2017	Yes	✓	✓	Packet, Bidirectional Flows	✓	×	Small Network	Emulated
Unified Host & n/w [92]	2017	Yes	✓	-	Logs, Bidirectional Flows	×	×	Enterprise Network	Real
CICIDS [93]	2018	Yes	✓	✓	Packet, Bidirectional Flows	✓	×	Small Network	Emulated
DDoS [94]	2019	Yes	✓	✓	Packet, Bidirectional Flows	✓	×	Small Network	Emulated

- Not Specified

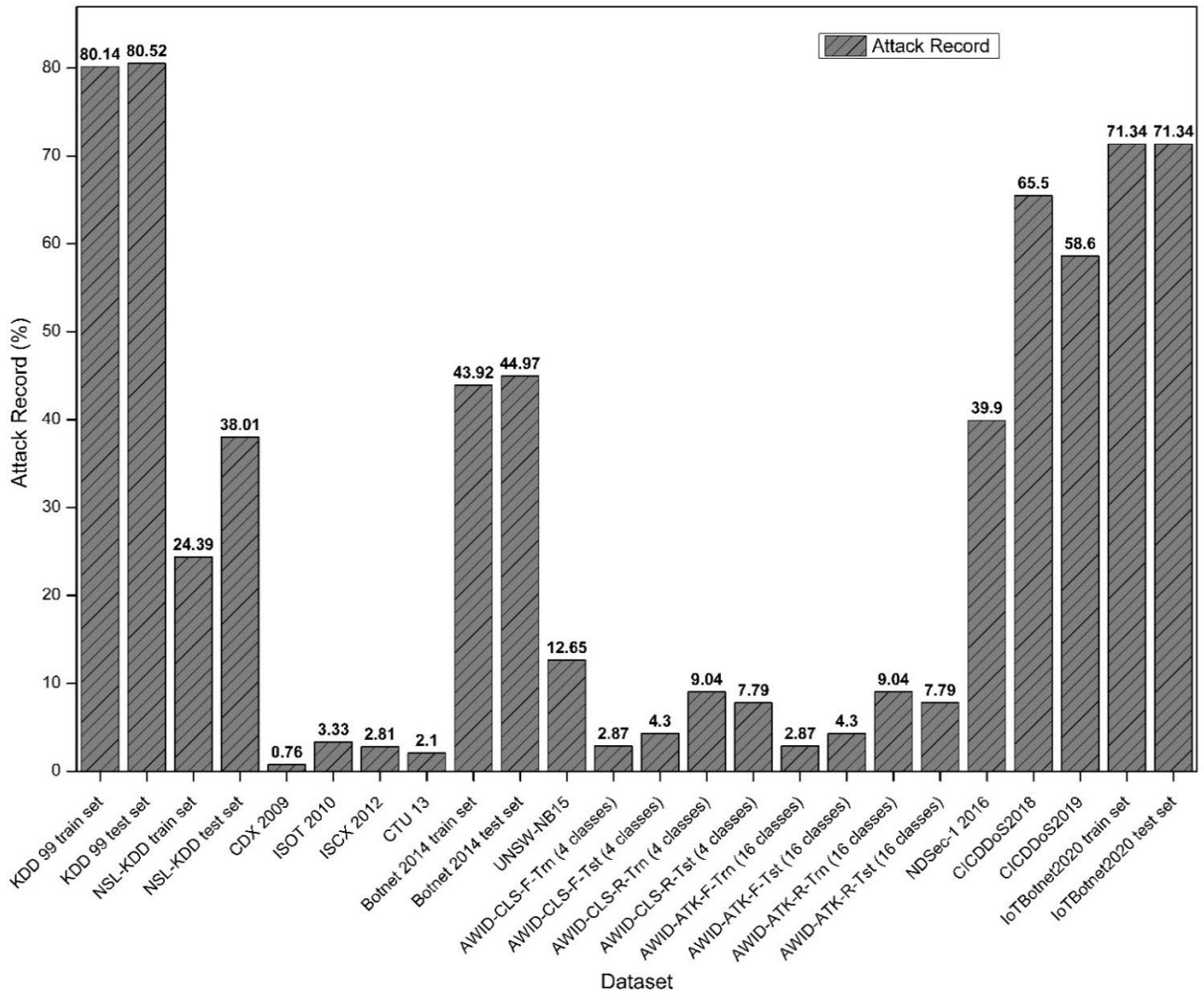

FIGURE 11. A bar chart representation of the percentage of attack record in different datasets.

Value of attributes to develop an upgraded Incremental SVM algorithm, which solves the oscillation problem of follow-up learning process of Incremental SVM. Incremental Learning is used by Zhuang *et al.* [120] for proposing a Malware Detection algorithm to preserve support vectors obtained from old data of SVM. Intrusion Detection is done using Incremental Multiclass SVM by Li *et al.* [121]. Although using Incremental Learning is required for real-time detection, here small dataset is used, so the sample size used for classification is significantly less, and SVM performance is not suitable for such scenarios. Online and Real-time version of unsupervised network anomaly detector is introduced using Incremental Grid Clustering for Intrusion Detection by Dromard *et al.* [122].

Results in the study demonstrate the scalability of the algorithm and can be used online. It uses a time sliding window, which is discrete for traffic collection. Speed is a bit slow, which can be improved using a better streaming tool. A synthetically generated log file is used; the applied algorithm learns by itself, creating clusters in sliding time as presented by Landauer *et al.* [123].

In this, false positives were detected in large amounts, which remains a big problem in Real-Time Anomaly Detection. Yang *et al.* [129] extracted features using damped Incremental statistics in place of a commonly used time sliding window, which reduced memory consumption. Incremental Possibilistic clustering (IPC) is used to detect outliers as new centers for clusters and use Mahalanobis distance for merging clusters. Various Genetic algorithms are used for anomaly detection. There is a recent shift of focus on swarm intelligence, which is used for optimization. For optimizing hyperparameters, different techniques are used by researchers. Shorman *et al.* [130] achieved impressive results using Grey Wolf Optimization for tuning the hyperparameters of one-class support vector Machine (OCSVM) and selecting the best Botnet features. Table 12 represents a detailed

TABLE 11. A detailed review of various machine learning-based data pre-processing techniques.

Authors	Dataset	Technique	Discussion
Lin *et al.* [102]	KDD-99	Feature Engineering - Cluster Center and Nearest Neighbor Classifier – K-NN	This algorithm performs better than K-NN and SVM in terms of performance metrics and computational efficiency for testing and training time.
Bijalwan *et al.* [103]	ISCX	Feature Engineering – Dataset is segregated into normal and attack classes Classifier – Ensemble Classifier	The use of Ensemble Classifier provides better results than a single classifier.
Alejandre *et al.* [104]	1. ISOT 2. ISCX	Feature Engineering - Genetic Algorithm Classifier – C 4.5 algorithm	A genetic algorithm was used as an optimizer, and because of this higher detection rate is achieved.
Garg *et al.* [105]	1. NSL-KDD 2. Kyoto	Feature Engineering - 1. Horizontal Feature Selection - Infinite Feature Selection 2. Vertical Feature selection – Abridging Algorithm Classifier – SVM	This analysis helps understand the effect of feature selection and can be used to reduce execution time.
Chellammal *et al.* [106]	1. KDD-99 2. NSL-KDD 3. Koyoto 2006	Feature Engineering – Correlation Detection - Ensemble Learning	Data is partitioned by segregating majority and minority classes and creating multiple datasets by sampled data.
Devan *et al.* [107]	NSL-KDD	Feature Engineering - XGBoost Classifier – Deep Neural Network	Features are selected analytically, and therefore, results are also good, but the drawback is optimal learning rate is chosen from experience, not analytically.
Rajadurai *et al.*[108]	NSL-KDD	Feature Engineering - PCA Classifier – Deep Learning model	PCA retains significant features, thus giving better results. This work is suited for detecting known attacks.
Li *et al.* [109]	Two data subsets from VirusShare, Four data subsets from VXHeavens	Feature Engineering - Random forest algorithm, Feature grouping Classifier - RMSE is calculated using Autoencoder and classified using Kmeans	Because of the three-layer neural network structure, it is efficient and lightweight. Autoencoder technique can efficiently solve the sample imbalance problem.
Khare *et al.* [110]	NSL-KDD	Feature Engineering - Min-Max Normalization, 1-N Encoding, Spider Monkey Optimization Classifier – Deep Neural Network	The use of nature-inspired algorithms for dimensionality reduction reduces the issues of quantity and quality of high dimensional data.

review of significant machine learning-based malware detection techniques.

C. REVIEW OF VARIOUS DEEP LEARNING TECHNIQUES IN IDS

Deep learning techniques are being extensively used for feature engineering because of their ability to learn high-dimensional features. Generative Adversarial Network (GAN) is one of the most common feature engineering techniques, specifically for their application in synthetic data creation and learning better about minority classes. Ferdowsi *et al.* [131] used GAN for feature engineering as well as detection. In this study, a distributed GAN is proposed to provide a fully distributed IDS for the IoT to detect anomalous behavior without reliance on any centralized controller. Learning more about the characteristics of the minority class seems to have improved the performance of classifying normal classes since the characteristics are significantly different from those of other classes, as depicted by Lee *et al.* [139] by deploying GAN for feature engineering. Features engineering is achieved using the Flow Wasserstein GAN model and Attention GRU Model by Han *et al.* [140]. An attention model is used to detect the payload-based attack. Yang *et al.* [141] used a Supervised adversarial Variational autoencoder for feature engineering. Regularization is achieved using Wasserstein GAN with gradient penalty. The advantage of SAVAER DNN is that it can effectively detect lower frequent attacks along with frequent attacks. Taylor Series is used along with Elephant Herd Optimization (TEHO) to train Deep Belief Network [142].

Deep networks' training time issue is resolved in this study by using Bhattacharya Distance for feature classification. Jiang *et al.* [143] addressed the issue of Data imbalance by

TABLE 12. A detailed review of various machine learning-based Malware detection techniques.

Authors	Dataset	Technique	Discussion
Meidan *et al.* [97]	Data is collected from nine IoT devices. Alexa Rank and GeoIP are used for enriching the dataset.	Feature Engineering – Twenty-three sets of features were extracted from five time windows. Classifiers -XGBoost, RandomForest, GBM	IoT and Non-IoT devices are classified using ML classifiers. This can be utilized for finding unauthorized IoT devices also.
Prokofiev *et al.* [124]	Data from 100 botnets is collected.	Feature Engineering - Logistic Regression Detection – Logistic Regression	Model is created to identify if a connection initiating device is running a bot.
Mazini *et al.* [125]	1. NSL-KDD 2.ISCXIDS2012	Feature Engineering - Artificial Bee Colony Evaluation – AdaBoost	The complexity of the model is less, and because of boosting algorithm, performance is also good.
Bezerra *et al.* [126]	Mirai, Bashlite, Hajime, Aidra, Tsunami, Dofloo are botnets used for attacks.	Feature Engineering – Scaling and Normalization Classifier - Elliptic Envelope, Isolation Forest, Local Outlier Factor, One-class Support Vector Machine	This work is mainly valuable when botnet details are end-to-end encrypted.
Khan *et al.* [127]	Five botnets were used for attacks; Wireshark was used to obtain CSV files.	Feature Engineering - Wrapper Method Evaluation – 10 fold cross-validation	Detection time is not mentioned, and the decision tree algorithm's depth is kept at eight, and the classification tree is set to 100 without explanation.
Djanie *et al.* [98]	The attack is launched using eight DoS attack tools. Wireshark is used to capture and display network traffic.	Feature Engineering – Manual Normalization Detection - SVM classifier	Snort IDS is used for testing, and a high detection rate is achieved.
Wang *et al.* [128]	Data is simulated using five new Botnets, namely Zues, Athena, Mirai, Ares, and Black Energy	Feature Engineering – Statistical and graph-based features are extracted. Evaluation - K means clustering, least-square technique, and Local Outlier Factor	In this study, a hybrid of both flow-based and graph-based detectors is used hence performs better than individual detectors.

using One-Side Selection (OSS) to reduce noise in majority classes and synthetically enhancing minority classes with the help of the Synthetic Minority Over-Sampling (SMOTE) technique. In this study, CNN is used for extracting spatial features, and BiLSTM for extracting temporal features. Simple feedforward neural networks are used as weak learners for feature engineering by Nabil *et al.* [144]. Binary classification is achieved in this study, and neural networks are used as specialized networks for a subset of features. Table 13 represents a detailed review of the critical Deep Learning based data pre-processing techniques.

Supervised Learning based on Deep Reinforcement Learning (DRL) is used by Martin *et al.* [154]. A new algorithm is used, which can also be applied in Online Learning. The response is very fast in DRL. Four DRL techniques, namely DQN, DDQN, Policy gradient, and Actor critic, are used in this work. Out of these, DDQN gives the best results. The results are also compared with existing Machine Learning algorithms. Sparse Autoencoder and Outlier detection method is used. This method can forecast IoT Botnets in advance, making it easier to detect and mitigate the attack, as shown by Kumar and Bhama [155]. Restricted Boltzmann Machine is used by Otoum *et al.* [156] for feature engineering and detection. The study indicates that RBC IDS and adaptively supervised and clustered hybrid IDS achieve the same detection and accuracy rates. However, the detection time of RBC IDS is approximately twice that of ASCH IDS. Wei *et al.* [157] used Deep Belief Network for feature engineering and Deep Belief Network and Particle Swarm Optimization for detection. In this study, an optimal network

TABLE 13. A detailed review of various deep learning-based data pre-processing techniques.

Authors	Dataset	Technique	Discussion
Erfani *et al.* [112]	An experiment is done on six real and two synthetic datasets.	Feature Engineering - Deep Belief network Classifier - One-Class SVM	This Model is faster than a deep autoencoder with comparable results. The linear kernel can be used as this model is scalable.
Sun *et al.* [132]	Mimicking attack is generated using LSGAN and GAN.	Feature Engineering - LSGAN and GAN	LSGAN and GAN are compared, results obtained could not clearly establish the need to use LSGAN in place of GAN.
Ma *et al.* [133]	1. ISCX-IDS-2012 2. CIC-IDS-2017	Feature Engineering – 1. 1D CNN: sequence features 2. Deep Neural Network: statistical and environmental features Classifier - Neural Network	A hybrid solution is given for feature selection, and a Shallow neural layer is used for anomaly detection.
Huang *et al.* [134]	1. NSL-KDD 2. UNSW-NB15 3. CIC-IDS-2017	Feature Engineering - Imbalanced data filter and convolutional layers are added to GAN.	In this study, by conducting several experiments, it is observed that synthesized samples are necessary for better performance, especially the ones of minority classes.
Saraeian *et al.* [135]	1. NSL-KDD 2. ISCXIDS 2012	Feature Engineering and detection-Convolutional Neural Network	The deep learning techniques have stronger learning ability which is intuitive from achieved higher accuracy.
Merino *et al.* [136]	KDD-99	Feature Engineering - Generative Adversarial Network Classifier - Neural Network Binary Classifier	In this study, the quality of data generated using GAN is ensured by using a neural network for evaluating model.
Manimurugan *et al.* [137]	CICIDS 2017	Feature Engineering – Duplicating technique Classifier - Deep Belief Network	A greedy layer-wise training algorithm was used to train DBN one layer at a time. Minority samples were combined together to avoid getting misclassified as benign.
Kim *et al.* [138]	1. KDD-99 2. CSE-CIC-IDS2018	Feature Engineering – Numerical Samples are converted to RGB and grayscale images. Classifier – Convolutional Neural Network	Study shows that RGB images in both binary and multiclass classifications have higher accuracy than grayscale images.

structure is acquired by comparing the DBN structure obtained from five optimization algorithms.

Xu *et al.* [158] normalized features manually and used RNN as a classifier. The detection rate achieved for DoS and Probing attacks was higher as compared to R2L and U2R attacks. This is due to the property of RNN of working well with time-series tasks, and DoS and Probing attacks have more obvious timing characteristics than R2L and U2R attacks. LSTM and Autoencoder classifiers are used as an ensemble by Zhong *et al.* [159]. This heterogeneous ensemble learning ensured that the model has better adaptability and accuracy when compared with other methods. The study by Muhuri *et al.* [160] shows that the performance of the LSTM-RNN model is better in binary classification as compared to multiclass classification. In binary classification, the proposed model outperforms SVM and RF models. In multiclass Classification, RF outperforms the LSTM-RNN model, as minority attack types do not show obvious timing characteristics. A study done by Hsu *et al.* [161] shows that after applying CNN, accuracy increases considerably when compared with only LSTM model. Some of the noteworthy Deep Learning based detection techniques are depicted in Table 14.

Parra *et al.* [162] proposed a model working on two security mechanisms simultaneously. A distributed convolutional neural network is used to secure IoT devices at the origin of

TABLE 14. A detailed review of various deep learning-based data Malware detection techniques.

Authors	Dataset	Technique	Discussion
Yousefi-Azar *et al.* [145]	Malware is collected from different sites like virustotal.	Feature Engineering - Hashing algorithm viz. tf -simhashing Classification - Novel extreme learning model	It extracts static features of any given binary file to distinguish malware from benign; hence, it helps mitigate the zero-day attack.
McDermott *et al.* [146]	A labeled dataset was created for four attack vectors (UDP, ACK, DNS, SYN Flood)of the Mirai botnet.	Feature Engineering - BLSTM-RNN with word Embedding Detection - Bidirectional LSTM-RNN	Dataset was generated in this study. Although results are promising but processing time is more, and for comparison, out of ten attack vectors, only four were considered.
Meidan *et al.* [97]	Mirai and Bashlite are used to infect devices. Data is collected using IoT devices; for sniffing, Wireshark is used.	Feature Engineering – Manual Normalization Classification - Deep Autoencoder	Back propagation is used, so as is the case for deep learning algorithms, the time taken for detection is more.
Pektaş *et al.* [147]	Six types of Botnets were used for the attack. Also, CTU 13 and ISOT were used as Benchmark datasets.	Feature Engineering - Graph structure is used to extract statistical-based network flow features. Detection - Convolutional and Recurrent neural network.	In this study, execution time is high even though higher configuration hardware is used.
Yeo *et al.* [148]	Nine different malware datasets from Stratosphere IPS are used.	Feature Engineering - Netmate Classification - CNN, MLP, SVM & RF	No analysis is given for the selection of parameters for models.
Ghasemi *et al.* [65]	KDD 99 and NSL KDD are used, and a new dataset is created based on five different labels using a Genetic Algorithm.	Feature Engineering - Genetic Algorithm Classification - Kernel Extreme Learning Machine	This model is trained for different behavior of all attacks individually; hence its performance is good.
Jahromi *et al.* [149]	VXHeaven, Kaggle, Windows ransomware, IoT malware, and a combined dataset of ransomware and IoT malware samples.	Feature Engineering – Not Required Classification - Novel extreme learning Machine model	Back propagation is avoided for training the network; thus, it is very fast compared to other approaches.
Qureshi *et al.* [150]	A subset of KDD dataset is used.	Feature Engineering - Pre-trained network on regression-related task is used for feature extraction. Detection - Novel Incremental SVM technique RS-ISVM	Oscillation problem of traditional SVM is reduced by retaining old samples, which are likely to become support vectors.
Kim *et al.* [151]	CTU-13	Feature Engineering - Manual sampling of features Detection - Recurrent Variational Autoencoder	At every time window, anomaly scores of every flow are calculated, which provides the degree of maliciousness of individual connections.
Assis *et al.* [152]	1. CICDDoS 2019 2. CICIDS 2018	Feature Engineering – Manual sampling of features Detection - Gated Recurrent Unit	GRU can learn long-term dependencies, making it a better option for mitigating sneakier attacks where attack profile is changing rapidly.
Rehman *et al.* [153]	CICDDoS 2019	Feature Engineering - Dataset is divided into different attack categories. Min-max normalization and SMOTE are used for data processing. Detection – Gated Recurrent Unit, Recurrent Neural Network, Naïve Bayes, Sequential Minimal Optimization	Best results are obtained using GRU and SMO. Results obtained differed according to the attack type: reflection or exploitation. On average GRU mechanism was proved to be better.

attack. At the back-end, the LSTM model is used. Using two models in tandem reduces the resource and communication requirement. Lee *et al.* [163] compared LSTM, CNN, MLP, Stacked Autoencoder (SAE) for detecting attacks based on packet length in SDN switch. Results show that MLP performs best. The reason might be the choice of number of layers as in this work for LSTM, only one layer is used, and three layers were used for all others. The results obtained vary accordingly. Bots are utilized to send traffic to the victim at comparatively lower speed in link-flooding attacks. These attacks are defended by Chen *et al*. [164] using an ensemble of CNN and LSTM models. The link-flooding DDoS attacks are difficult to mitigate; LSTM is utilized in this work to review the attack patterns periodically. For a similar low rate DDoS attack pattern in wireless systems, Liu and Yin [165] used a combination of LSTM and CGAN. This is because LSTM works well in identifying patterns in sequenced packages.

The study by Zavrak and Iskefiyeli [166] shows that Variational Autoencoder gave similar ROC curves for the attacks displaying similar behavioral characteristics. Adversarial Autoencoder (AAE) was used by Hara and Shiomoto [167] as a semi-supervised learning technique to address the issue of more extensive labeled data needed by other methods. Comparable results were obtained in this study with only 0.1% of labeled data. However, the concern is longer training time as AAE trains one neural network for the AE and two neural networks for Discriminators.

V. SECURITY-BASED CHALLENGES AND PROPOSED SOLUTIONS

IoT sector faces several security challenges that need to be fixed at priority for the IoT domain's further progress. The heterogeneous nature of IoT devices is a major concern as due to this issue single solution is not possible. Fig. 12 represents a detailed analysis of various research challenges, research gaps, and possible solutions.

CHALLENGE 1: ROBUST MACHINE LEARNING MODEL

Broadly robustness is defined as the property where results obtained in the training set are similar to that of the test set. A robust machine learning model is required for real-world applications [168]. SVM and LASSO algorithms can be written as robust algorithms [169]. The models work well in theory, but all are not accepted at enterprise levels because of robustness.

PROPOSED SOLUTION: Incremental Learning can be used as a solution for achieving robustness in actual terms. As the data is changing continuously in a real-world scenario, a model might get trained on a quite different set from the actual validation set. In Incremental Learning, the model keeps learning continuously, making it more robust. Deep learning techniques like the generation of adversarial data for checking the system's robustness are being used in some works. Combining the benefits of both Incremental and Deep Learning can provide astonishing results.

CHALLENGE 2: GENERALIZABILITY OF MODEL

Robustness is a necessary and sufficient condition for generalizability. Generalizability is defined by assessing the performance of a model on unseen test situations [170]. Robustness and generalizability are usually not seen together to evaluate a model, whereas a robust, generalizable model should be the target to make a sustainable model.

PROPOSED SOLUTION: Incremental and transfer Learning, if used collectively, might solve this issue. Incremental Learning is often used in image classification [116], target recognition [117], and used less in Intrusion Detection or information security. Similar is the case for transfer learning; the fusion of these two techniques can pave way for incredible results in terms of robustness and generalizability in the field of anomaly detection.

CHALLENGE 3: REAL-TIME ANALYSIS

Real-time analysis is essential for any model to be adopted at the enterprise level. IDS models presented in the literature are typically offline. In malware classification, the challenge is in identifying patterns to distinguish between legitimate and malware traffic. In offline mode, machine learning models work on static datasets while the online stream of data is analyzed in online Learning. Till now, real-time analysis of data is not explored much for malware analysis.

PROPOSED SOLUTION: Incremental Learning can be the solution for real-time analysis as the model can get updated according to newly added features. The significant old results can be kept aside to extract information from the same for upcoming similar data. The related approach is used by Qureshi *et al*. [150] in case of support vectors, by retaining old samples which are likely to become support vector.

CHALLENGE 4: RESOURCE CONSTRAINTS OF IOT DEVICES

IoT devices are known to be constraint devices in terms of power, cost, and size. With constraints in place, maintaining security is a challenge. As for low-cost IoT devices, keeping all the security requirements is a major concern. Capacities of Deep Learning techniques could not be utilized because of these constraints.

PROPOSED SOLUTION: A solution for this could be to use Deep Learning techniques with powerful hyperparameter tuning techniques. Although, Deep learning does not necessarily require feature engineering, using it makes the model lightweight. Then the model can be used at the node itself, allowing faster action in case of attack. Thus, properly using Deep Learning techniques can help mitigate the effect of resource constraint and employ newly available techniques.

CHALLENGE 5: LONGER TRAINING TIME OF INTRUSION DETECTION MODEL

Most of the Intrusion Detection models suffer from longer training time, which affects the performance of the model to such an extent that sometimes compromise has to be made on overall system performance to reduce training time. This becomes a more significant challenge while using Deep Learning models because of the number of layers involved.

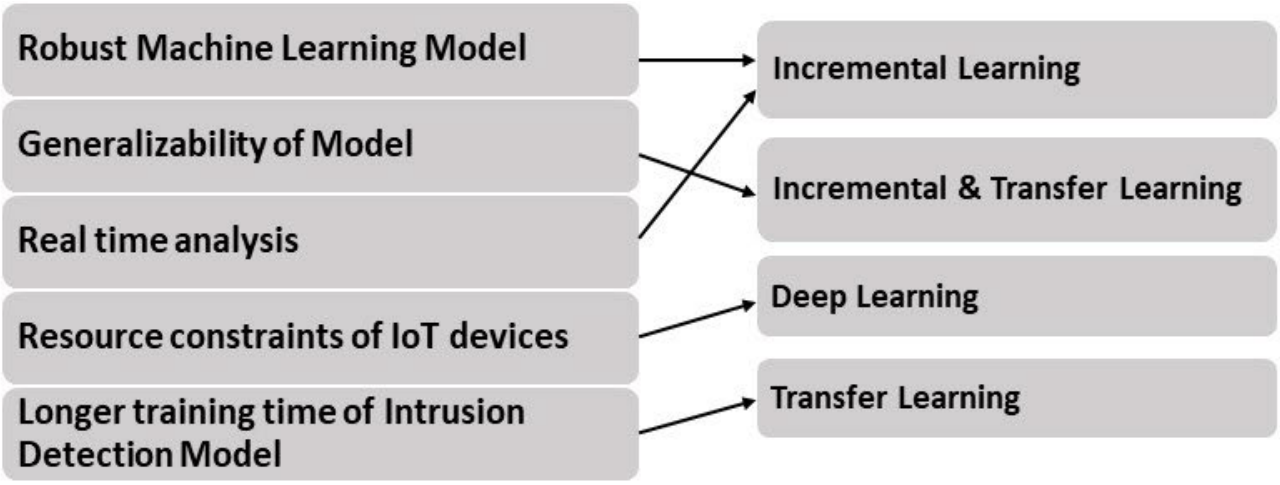

FIGURE 12. A logical mapping of comprehensive challenges, Research gaps and possible solutions.

PROPOSED SOLUTION: Transfer Learning is defined as the ability to use a pre-trained model for different yet similar work.

This concept is being actively used in various use cases of Machine Learning, but it is not explored much in malware analysis. When used with Transfer Learning, Deep Learning enhances system performance multifold as better results can be obtained in less training time.

VI. CONCLUSION

The evolution of IoT has been incredible, and it has paved the way for several endeavors in the field of technology. IoT security plays a crucial role in further technology progression as investors will move ahead in this domain only when state-of-the-art security measures are met. In general, cybersecurity works on the CIA model, i.e., confidentiality, integrity, and availability. The attackers tend to utilize vulnerabilities of communication protocols for launching attacks. Better mitigation techniques are required for mitigating attacks as these attacks jeopardize service providers' reputations. All three aspects, viz. confidentiality, integrity, and availability, are affected due to the attacks, which is the primary concern for service providers. IoT devices have scales for generating the data from small-scale applications with few bytes every second to Kbytes every second depending upon the application being addressed. This data is sometimes very crucial, viz. medical data, military data. Distributed denial of services (DDoS) attacks are significant threats for the cyber world because of their potential to bring down the victims. DDoS attacks require a large number of devices for launching attacks, and for this, IoT devices are well suited. As in most cases, users will not understand that the device is compromised; viz., baby monitors, smart toys have a user interface with limited access and may usually work even after being part of a Botnet army. With the increasing volume of IoT devices, there is an urgent need to detect Botnet attacks timely to remove compromised devices.

As we move into the era where almost everything being used by humans will be connected to the internet, the security of these devices becomes supreme. Two major solutions are found in the literature for preventing DDoS attacks, namely Intrusion Detection System (IDS) and Intrusion Prevention System (IPS). The Intrusion Detection System is analyzed in the undertaken review work, and various intrusion detection models have been evaluated. Furthermore, we have also discussed the classification of Intrusion Detection Systems, different anomaly detection techniques, various Intrusion Detection System models based on datasets, diverse machine learning, and deep learning techniques for data pre-processing and malware detection. In the end, a broader perspective has been envisioned while surveying different intrusion detection techniques and future visions.

In the future, we plan to implement these solutions and develop a robust and generalized intrusion detection model.

REFERENCES

[1] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for Internet of Things (IoT) security," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1646–1685, 3rd Quart., 2020.

[2] R. Hassan, F. Qamar, M. K. Hasan, A. Hafizah, M. Aman, and A. S. Ahmed, "Internet of Things and its applications: A comprehensive survey," *Symmetry*, vol. 12, no. 10, p. 1674, 2020.

[3] S. P. Dash, "The impact of IoT in healthcare: Global review," *J. Indian Inst. Sci.*, vol. 100, no. 4, pp. 773–785, 2020.

[4] K. Demestichas, N. Peppes, and T. Alexakis, "Survey on security threats in agricultural IoT and smart farming," *Sensors*, vol. 20, no. 22, p. 6458, Nov. 2020.

[5] M. A. Rahim, M. A. Rahman, M. M. Rahman, A. T. Asyhari, M. Z. A. Bhuiyan, and D. Ramasamy, "Evolution of IoT-enabled connectivity and applications in automotive industry: A review," *Veh. Commun.*, vol. 27, pp. 1–15, Jan. 2021.

[6] N. K. Suryadevara and G. R. Biswal, "Smart plugs: Paradigms and applications in the smart city-and-smart grid," *Energies*, vol. 12, no. 10, pp. 1–20, 2019.

[7] S. B. Atitallah, M. Driss, W. Boulila, and H. B. Ghézala, "Leveraging deep learning and IoT big data analytics to support the smart cities development: Review and future directions," *Comput. Sci. Rev.*, vol. 38, Nov. 2020, Art. no. 100303.

[8] S. C. Eimler, N. C. Krämer, and A. M. Von Der Pütten, "Empirical results on determinants of acceptance and emotion attribution in confrontation with a robot rabbit," *Appl. Artif. Intell.*, vol. 25, no. 6, pp. 503–529, 2011.
[9] M. Miraz, M. Ali, P. Excell, and R. Picking, "Internet of nano-things, things and everything: Future growth trends," *Future Internet*, vol. 10, no. 8, p. 68, Jul. 2018.
[10] Q.-D. Ngo, H.-T. Nguyen, L.-C. Nguyen, and D.-H. Nguyen, "A survey of IoT malware and detection methods based on static features," *Opt. Commun.*, 2020, Art. no. 126175.
[11] I. Zafeiriou, "IoT and mobility in smart cities," in *Proc. 3rd World Symp. Commun. Eng. (WSCE)*, 2020, pp. 91–95.
[12] T. S. Bharati, "Internet of Things (IoT): A critical review," *Int. J. Sci. Technol. Res.*, vol. 8, no. 10, pp. 227–232, 2019.
[13] A. Pal, H. K. Rath, S. Shailendra, and A. Bhattacharyya, "IoT standardization: The road ahead," in *Proc. IntechOpen*, 2018, pp. 53–74.
[14] A. Srivastava, S. Gupta, M. Quamara, P. Chaudhary, and V. J. Aski, "Future IoT-enabled threats and vulnerabilities: State of the art, challenges, and future prospects," *Int. J. Commun. Syst.*, vol. 33, no. 12, pp. 1–40, 2020.
[15] J. Ding, M. Nemati, C. Ranaweera, and J. Choi, "IoT connectivity technologies and applications: A survey," *IEEE Access*, vol. 8, pp. 67646–67673, 2020.
[16] Y. Yang, L. Wu, G. Yin, L. Li, and H. Zhao, "A survey on security and privacy issues in Internet-of-Things," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1250–1258, Oct. 2017.
[17] W. Yu, F. Liang, X. He, W. G. Hatcher, C. Lu, J. Lin, and X. Yang, "A survey on the edge computing for the Internet of Things," *IEEE Access*, vol. 6, pp. 6900–6919, 2018.
[18] D. E. Kouicem, A. Bouabdallah, and H. Lakhlef, "Internet of Things security: A top-down survey," *Comput. Netw.*, vol. 141, pp. 199–221, Aug. 2018.
[19] M. Frustaci, P. Pace, G. Aloi, and G. Fortino, "Evaluating critical security issues of the IoT world: Present and future challenges," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2483–2495, Aug. 2018.
[20] M. B. M. Noor and W. H. Hassan, "Current research on Internet of Things (IoT) security: A survey," *Comput. Netw.*, vol. 148, pp. 283–294, Jan. 2019.
[21] V. Hassija, V. Chamola, V. Saxena, D. Jain, P. Goyal, and B. Sikdar, "A survey on IoT security: Application areas, security threats, and solution architectures," *IEEE Access*, vol. 7, pp. 82721–82743, 2019.
[22] F. Meneghello, M. Calore, D. Zucchetto, M. Polese, and A. Zanella, "IoT: Internet of threats? A survey of practical security vulnerabilities in real IoT devices," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8182–8201, Oct. 2019.
[23] P. Anand, Y. Singh, A. Selwal, M. Alazab, S. Tanwar, and N. Kumar, "IoT vulnerability assessment for sustainable computing: Threats, current solutions, and open challenges," *IEEE Access*, vol. 8, pp. 168825–168853, 2020.
[24] K. Angrishi, "Turning Internet of Things (IoT) into Internet of vulnerabilities (IoV): IoT botnets," 2017, *arXiv:1702.03681*. [Online]. Available: https://arxiv.org/abs/1702.03681
[25] M. Burhan and R. A. Rehman, "IoT elements, layered architectures and security issues: A comprehensive survey," *Sensors*, vol. 18, no. 9, pp. 1–37, 2018.
[26] H. A. Khattak, M. A. Shah, S. Khan, I. Ali, and M. Imran, "Perception layer security in the Internet of Things," *Futur. Gener. Comput. Syst.*, vol. 100, pp. 144–164, Jan. 2019.
[27] O. S. J. Nisha and S. M. S. Bhanu, "A survey on code injection attacks in mobile cloud computing environment," in *Proc. 8th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence)*, Jan. 2018, pp. 1–6.
[28] M. Devi and A. Majumder, "Side-channel attack in Internet of Things: A survey," in *Applications of Internet of Things* (Lecture Notes in Networks and Systems), J. K. Mandal, S. Mukhopadhyay, and A. Roy, Eds. Singapore: Springer, 2021, pp. 213–222.
[29] K. O. A. Alimi, K. Ouahada, A. M. Abu-Mahfouz, and S. Rimer, "A survey on the security of low power wide area networks: Threats, challenges, and potential solutions," *Sensors*, vol. 20, no. 20, pp. 1–19, 2020.
[30] H. Mrabet, S. Belguith, A. Alhomoud, and A. Jemai, "A survey of IoT security based on a layered architecture of sensing and data analysis," *Sensors*, vol. 20, no. 13, pp. 1–20, 2020.
[31] K. Nirmal, B. Janet, and R. Kumar, "Analyzing and eliminating phishing threats in IoT, network and other Web applications using iterative intersection," *Peer-Peer Netw. Appl.*, pp. 1–13, Jun. 2020.
[32] R. Vishwakarma and A. K. Jain, "A survey of DDoS attacking techniques and defence mechanisms in the IoT network," *Telecommun. Syst.*, vol. 73, no. 1, pp. 3–25, Jan. 2020.
[33] A. Raoof, A. Matrawy, and C.-H. Lung, "Enhancing routing security in IoT: Performance evaluation of RPL's secure mode under attacks," *IEEE Internet Things J.*, vol. 7, no. 12, pp. 11536–11546, Dec. 2020.
[34] D. Dinculean and X. Cheng, "Vulnerabilities and limitations of MQTT protocol used between IoT devices," *Appl. Sci.*, vol. 9, no. 5, p. 848, Feb. 2019.
[35] A. Tewari and B. B. Gupta, "Security, privacy and trust of different layers in Internet-of-Things (IoTs) framework," *Future Gener. Comput. Syst.*, vol. 108, pp. 909–920, Jul. 2020.
[36] B. Ahlawat, A. Sangwan, and V. Sindhu, "IoT system model, challenges and threats," *Int. J. Sci. Technol. Res.*, vol. 9, no. 3, pp. 6771–6776, 2020.
[37] B. Prabadevi and N. Jeyanthi, "A review on various sniffing attacks and its mitigation techniques," *Indonesian J. Elect. Eng. Comput. Sci.*, vol. 12, no. 3, pp. 1117–1125, 2018.
[38] G. Perrone, M. Vecchio, R. Pecori, and R. Giaffreda, "The day after Mirai: A survey on MQTT security solutions after the largest cyber-attack carried out through an army of IoT devices," in *Proc. Big Data Secur. (IoTBDS)*, Apr. 2017, pp. 246–253.
[39] Cloudflare. (2020). *Famous DDoS Attacks | Cloudflare*. Accessed: Mar. 20, 2021. [Online]. Available: https://www.cloudflare.com/learning/ddos/famous-ddos-attacks/
[40] P. Nicholson. (2020). Five most famous DDoS attacks and then some. A10 Blog. Accessed: Mar. 20, 2021. [Online]. Available: https://www.a10networks.com/blog/5-most-famous-ddos-attacks/
[41] S. Hussain, R. Atallah, and A. Kamsin, "DDoS reflection attack based on IoT: A case study," in *Proc. Comput. Sci. Line Conf.* Cham, Switzerland: Springer, 2019, pp. 44–52.
[42] A. Colella and C. M. Colombini, "Amplification DDoS attacks: Emerging threats and defense strategies," in *Proc. Int. Conf. Availability, Reliability, Secur.*, in Lecture Notes in Computer Science: Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, vol. 8708, 2014, pp. 298–310.
[43] C. Hesselman, M. Kaeo, L. Chapin, K. Claffy, M. Seiden, D. McPherson, D. Piscitello, A. McConachie, T. April, J. Latour, and R. Rasmussen, "The DNS in IoT: Opportunities, risks, and challenges," *IEEE Internet Comput.*, vol. 24, no. 4, pp. 23–32, Jul. 2020.
[44] T. Kawamura, M. Fukushi, Y. Hirano, Y. Fujita, and Y. Hamamoto, "An NTP-based detection module for DDoS attacks on IoT," in *Proc. IEEE Int. Conf. Consum. Electron.-Taiwan (ICCE-TW)*, Jun. 2017, pp. 15–16.
[45] J. J. C. Gondim, R. de Oliveira Albuquerque, and A. L. S. Orozco, "Mirror saturation in amplified reflection distributed denial of service: A case of study using SNMP, SSDP, NTP and DNS protocols," *Future Gener. Comput. Syst.*, vol. 108, pp. 68–81, Jul. 2020.
[46] A. T. Vasques and J. J. C. Gondim, "Amplified reflection DDoS attacks over IoT mirrors: A saturation analysis," in *Proc. Workshop Commun. Netw. Power Syst. (WCNPS)*, Oct. 2019, pp. 1–6.
[47] W. M. Lafta, A. A. Alkadhmawee, and M. A. Altaha, "Best strategy to control data on Internet-of-robotic-things in heterogeneous networks," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 2, pp. 1830–1838, 2021.
[48] H. Burch and B. Cheswick, "Tracing anonymous packets to their approximate source," in *Proc. 14th Conf. Syst. Admin. LISA*, 2000, pp. 319–327.
[49] N. Z. M. Safar, N. Abdullah, H. Kamaludin, S. A. Ishak, and M. R. M. Isa, "Characterising and detection of botnet in P2P network for UDP protocol," *Indonesian J. Elect. Eng. Computer Sci.*, vol. 18, no. 3, pp. 1584–1595, 2020.
[50] B. Sieklik, R. Macfarlane, and W. J. Buchanan, "Evaluation of TFTP DDoS amplification attack," *Comput. Secur.*, vol. 57, pp. 67–92, Mar. 2016.
[51] S.-J. Choi and J. Kwak, "A study on reduction of DDoS amplification attacks in the UDP-based CLDAP protocol," in *Proc. 4th Int. Conf. Comput. Appl. Inf. Process. Technol. (CAIPT)*, Aug. 2017, pp. 1–4.
[52] A. Blaise, M. Bouet, V. Conan, and S. Secci, "Detection of zero-day attacks: An unsupervised port-based approach," *Comput. Netw.*, vol. 180, Oct. 2020, Art. no. 107391.
[53] A. Kondoro, I. B. Dhaou, H. Tenhunen, and N. Mvungi, "Real time performance analysis of secure IoT protocols for microgrid communication," *Future Gener. Comput. Syst.*, vol. 116, pp. 1–12, Mar. 2021.
[54] M. Malaimalavathani and R. Gowri, "A survey on semantic Web service discovery," in *Proc. Int. Conf. Inf. Commun. Embedded Syst. (ICICES)*, Feb. 2013, pp. 222–225.

[55] S. Bjarnason and R. Dobbins. (2019). A call to ARMS: Apple remote management service UDP reflection/amplification DDoS attacks. ASERT Blog. Accessed: Mar. 22, 2021. [Online]. Available: https://www.netscout.com/blog/asert/call-arms-apple-remote-management-service-udp

[56] Y. Al-Hadhrami and F. K. Hussain, "DDoS attacks in IoT networks: A comprehensive systematic literature review," *World Wide Web*, pp. 1–31, Jan. 2021.

[57] K. Doshi, Y. Yilmaz, and S. Uludag, "Timely detection and mitigation of stealthy DDoS attacks via IoT networks," *IEEE Trans. Dependable Secure Comput.*, early access, Jan. 8, 2021, doi: 10.1109/TDSC.2021.3049942.

[58] S. E. Belanda, C. F. M. Foozy, A. Mustapha, P. S. S. Palaniapan, and Z. Abdullah, "Detecting botnet attack in Internet of Things (IoTs) environment by using machine learning technique: A review," *J. Crit. Rev.*, vol. 7, no. 8, pp. 1324–1329, 2020.

[59] H. Wu, H. Han, X. Wang, and S. Sun, "Research on artificial intelligence enhancing Internet of Things security: A survey," *IEEE Access*, vol. 8, pp. 153826–153848, 2020.

[60] A. Munshi, N. A. Alqarni, and N. A. Almalki, "DDOS attack on IoT devices," in *Proc. 3rd Int. Conf. Comput. Appl. Inf. Secur. (ICCAIS)*, Mar. 2020, pp. 5–9.

[61] C. V. Martínez and B. Vogel-Heuser, "Towards industrial intrusion prevention systems: A concept and implementation for reactive protection," *Appl. Sci.*, vol. 8, no. 12, pp. 1–29, 2018.

[62] T. R. Glass-Vanderlan, M. D. Iannacone, M. S. Vincent, Q. Chen, and R. A. Bridges, "A survey of intrusion detection systems leveraging host data," *ACM Comput. Surv.*, vol. 52, no. 6, p. 128, 2018.

[63] R. Panigrahi, S. Borah, A. K. Bhoi, and P. K. Mallick, "Intrusion detection systems (IDS)—An overview with a generalized framework," *Adv. Intell. Syst. Comput.*, vol. 1040, pp. 107–117, Jan. 2020.

[64] A. N. Cahyo, A. K. Sari, and M. Riasetiawan, "Comparison of hybrid intrusion detection system," in *Proc. 12th Int. Conf. Inf. Technol. Electr. Eng. (ICITEE)*, Oct. 2020, pp. 92–97.

[65] J. Ghasemi, J. Esmaily, and R. Moradinezhad, "Intrusion detection system using an optimized kernel extreme learning machine and efficient features," *Sādhanā-Acad. Eng. Sci.*, vol. 45, no. 1, Dec. 2020, Art. no. 2.

[66] W. N. H. Ibrahim, S. Anuar, A. Selamat, O. Krejcar, R. G. Crespo, E. Herrera-Viedma, and H. Fujita, "Multilayer framework for botnet detection using machine learning algorithms," *IEEE Access*, vol. 9, pp. 48753–48768, 2021.

[67] W. Wang, B. Fang, Z. Zhang, and C. Li, "A novel approach to detect IRC-based botnets," in *Proc. Int. Conf. Netw. Secur., Wireless Commun. Trusted Comput. (NSWCTC)*, Apr. 2009, pp. 408–411.

[68] D. Zhao, I. Traore, B. Sayed, W. Lu, S. Saad, A. Ghorbani, and D. Garant, "Botnet detection based on traffic behavior analysis and flow intervals," *Comput. Secur.*, vol. 39, pp. 2–16, Nov. 2013.

[69] G. Moura, R. Sadre, and A. Pras, "Bad neighborhoods on the Internet," *IEEE Commun. Mag.*, vol. 52, no. 7, pp. 132–139, Jul. 2014.

[70] T. H. Hai, L. H. Hoang, and E.-N. Huh, "Network anomaly detection based on late fusion of several machine learning algorithms," *Int. J. Comput. Netw. Commun.*, vol. 12, no. 6, pp. 117–131, Nov. 2020.

[71] K. Kalkan, L. Altay, G. Gür, and F. Alagöz, "JESS: Joint entropy-based DDoS defense scheme in SDN," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 10, pp. 2358–2372, Oct. 2018.

[72] A. Kim, M. Park, and D. H. Lee, "AI-IDS: Application of deep learning to real-time Web intrusion detection," *IEEE Access*, vol. 8, pp. 70245–70261, 2020.

[73] V. Kumar, D. Sinha, A. K. Das, S. C. Pandey, and R. T. Goswami, "An integrated rule based intrusion detection system: Analysis on UNSW-NB15 data set and the real time online dataset," *Cluster Comput.*, vol. 23, no. 2, pp. 1397–1418, Jun. 2020.

[74] S. Choudhary and N. Kesswani, "A survey: Intrusion detection techniques for Internet of Things," *Int. J. Inf. Secur. Privacy*, vol. 13, no. 1, pp. 86–105, 2019.

[75] M. Choraś, Ł. Saganowski, R. Renk, and W. Hołubowicz, "Statistical and signal-based network traffic recognition for anomaly detection," *Expert Syst.*, vol. 29, no. 3, pp. 232–245, Jul. 2012.

[76] M. Yoon, "Using whitelisting to mitigate DDoS attacks on critical Internet sites," *IEEE Commun. Mag.*, vol. 48, no. 7, pp. 110–115, Jul. 2010.

[77] (1999). *KDD Cup 1999 Data*. Accessed: Mar. 21, 2021. [Online]. Available: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[78] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *Proc. IEEE Symp. Comput. Intell. Secur. Def. Appl. (CISDA)*, 2009, pp. 1–6.

[79] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *Proc. Mil. Commun. Inf. Syst. Conf. (MilCIS)*, Nov. 2015, pp. 1–6.

[80] N. Moustafa and J. Slay, "The evaluation of network anomaly detection systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set," *Inf. Secur. J., Global Perspective*, vol. 25, nos. 1–3, pp. 18–31, Apr. 2016.

[81] S. García, M. Grill, J. Stiborek, and A. Zunino, "An empirical comparison of botnet detection methods," *Comput. Secur.*, vol. 45, pp. 100–123, Sep. 2014.

[82] S. Saad, I. Traore, A. Ghorbani, B. Sayed, D. Zhao, W. Lu, J. Felix, and P. Hakimian, "Detecting P2P botnets through network behavior analysis and machine learning," in *Proc. 9th Annu. Int. Conf. Privacy, Secur. Trust*, Jul. 2011, pp. 174–180.

[83] R. P. Lippmann, D. J. Fried, I. Graf, J. W. Haines, K. R. Kendall, D. McClung, D. Weber, S. E. Webster, D. Wyschogrod, R. K. Cunningham, and M. A. Zissman, "Evaluating intrusion detection systems: The 1998 DARPA off-line intrusion detection evaluation," in *Proc. DARPA Inf. Survivability Conf. Expo. (DISCEX)*, 2000, pp. 12–26.

[84] R. Lippmann, J. W. Haines, D. J. Fried, J. Korba, and K. Das, "1999 DARPA off-line intrusion detection evaluation," *Comput. Netw.*, vol. 34, no. 4, pp. 579–595, 2000.

[85] F. G. L. Salgarelli, M. Dusi, and P. Torino, "GT: Picking up the truth from the ground for Internet traffic," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 39, no. 5, pp. 13–18, 2009.

[86] B. Sangster, T. J. O'Connor, T. Cook, R. Fanelli, E. Dean, C. Morrell, and G. J. Conti, "Toward instrumenting network warfare competitions to generate labeled datasets," in *Proc. 2nd Work. Cyber Secur. Express Test (CSET)*, 2009, pp. 1–6.

[87] A. Shiravi, H. Shiravi, M. Tavallaee, and A. A. Ghorbani, "Toward developing a systematic approach to generate benchmark datasets for intrusion detection," *Comput. Secur.*, vol. 31, no. 3, pp. 357–374, May 2012.

[88] E. B. Beigi, H. H. Jazi, N. Stakhanova, and A. A. Ghorbani, "Towards effective feature selection in machine learning-based botnet detection approaches," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Oct. 2014, pp. 247–255.

[89] C. Kolias, G. Kambourakis, A. Stavrou, and S. Gritzalis, "Intrusion detection in 802. 11 networks: Empirical evaluation of threats and a public dataset," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 184–208, 1st Quart., 2016.

[90] F. Beer, T. Hofer, D. Karimi, and U. Bühler, "A new attack composition for network security," *Lect. Notes Informat. Ser. Gesellschaft fur Inform.*, vol. 271, pp. 11–20, 2017.

[91] H. H. Jazi, H. Gonzalez, N. Stakhanova, and A. A. Ghorbani, "Detecting HTTP-based application layer DoS attacks on Web servers in the presence of sampling," *Comput. Netw.*, vol. 121, pp. 25–36, Jul. 2017.

[92] M. J. M. Turcotte, A. D. Kent, and C. Hash, "Unified host and network data set," 2017, *arXiv:1708.07518*. [Online]. Available: http://arxiv.org/abs/1708.07518

[93] AWS. (2018). *A Realistic Cyber Defense Dataset (CSE-CIC-IDS2018)*. Accessed: Mar. 21, 2021. [Online]. Available: https://registry.opendata.aws/cse-cic-ids2018

[94] I. Sharafaldin, A. H. Lashkari, S. Hakak, and A. A. Ghorbani, "Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy," in *Proc. Int. Carnahan Conf. Secur. Technol. (ICCST)*, Oct. 2019, pp. 1–8.

[95] E. Guillén, J. Rodriguez, R. Páez, and A. Rodriguez, "Detection of non-content based attacks using GA with extended KDD features," *Lect. Notes Eng. Comput. Sci.*, vol. 1, pp. 30–35, Oct. 2012.

[96] I. Sharafaldin, A. Habibi Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *Proc. 4th Int. Conf. Inf. Syst. Secur. Privacy*, 2018, pp. 108–116.

[97] Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, A. Shabtai, D. Breitenbacher, and Y. Elovici, "N-BaIoT—Network-based detection of IoT botnet attacks using deep autoencoders," *IEEE Pervas. Comput.*, vol. 17, no. 3, pp. 12–22, Jul. 2018.

[98] Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, A. Shabtai, D. Breitenbacher, and Y. Elovici, "N-BaIoT—Network-based detection of IoT botnet attacks using deep autoencoders," *IEEE Pervas. Comput.*, vol. 17, no. 3, pp. 12–22, Jul. 2018.

[99] S. Chesney, K. Roy, and S. Khorsandroo, "Machine learning algorithms for preventing IoT cybersecurity attacks," in *Proc. SAI Intell. Syst. Conf.*, in Advances in Intelligent Systems and Computing, vol. 1252, 2021, pp. 679–686.
[100] M. A. Amanullah, R. A. A. Habeeb, F. H. Nasaruddin, A. Gani, E. Ahmed, A. S. M. Nainar, and N. M. Akim, "Deep learning and big data technologies for IoT security," *Comput. Commun.*, vol. 151, pp. 495–517, Feb. 2020.
[101] S. Chowdhury, M. Khanzadeh, R. Akula, F. Zhang, S. Zhang, H. Medal, M. Marufuzzaman, and L. Bian, "Botnet detection using graph-based feature clustering," *J. Big Data*, vol. 4, no. 1, pp. 1–23, Dec. 2017.
[102] W.-C. Lin, S.-W. Ke, and C.-F. Tsai, "CANN: An intrusion detection system based on combining cluster centers and nearest neighbors," *Knowl.-Based Syst.*, vol. 78, pp. 13–21, Apr. 2015.
[103] A. Bijalwan, N. Chand, E. S. Pilli, and C. Rama Krishna, "Botnet analysis using ensemble classifier," *Perspect. Sci.*, vol. 8, pp. 502–504, Sep. 2016.
[104] F. V. Alejandre, N. C. Cortés, and E. A. Anaya, "Feature selection to detect botnets using machine learning algorithms," in *Proc. Int. Conf. Electron., Commun. Comput. (CONIELECOMP)*, 2017, pp. 1–7.
[105] S. Garg, R. Singh, M. S. Obaidat, V. K. Bhalla, and B. Sharma, "Statistical vertical reduction-based data abridging technique for big network traffic dataset," *Int. J. Commun. Syst.*, vol. 33, no. 4, pp. 1–13, 2020.
[106] P. Chellammal and P. D. S. K. Malarchelvi, "Real-time anomaly detection using parallelized intrusion detection architecture for streaming data," *Concurrency Comput., Pract. Exper.*, vol. 32, no. 4, pp. 1–9, Feb. 2020.
[107] P. Devan and N. Khare, "An efficient XGBoost–DNN-based classification model for network intrusion detection system," *Neural Comput. Appl.*, vol. 32, no. 16, pp. 12499–12514, Aug. 2020.
[108] H. Rajadurai and U. D. Gandhi, "An empirical model in intrusion detection systems using principal component analysis and deep learning models," *Comput. Intell.*, pp. 1–14, Jun. 2020.
[109] X. Li, W. Chen, Q. Zhang, and L. Wu, "Building auto-encoder intrusion detection system based on random forest feature selection," *Comput. Secur.*, vol. 95, Aug. 2020, Art. no. 101851.
[110] N. Khare, P. Devan, C. Chowdhary, S. Bhattacharya, G. Singh, S. Singh, and B. Yoon, "SMO-DNN: Spider monkey optimization and deep neural network hybrid classifier model for intrusion detection," *Electronics*, vol. 9, no. 4, p. 692, Apr. 2020.
[111] C. A. Catania, F. Bromberg, and C. G. Garino, "An autonomous labeling approach to support vector machines algorithms for network traffic anomaly detection," *Expert Syst. Appl.*, vol. 39, no. 2, pp. 1822–1829, Feb. 2012.
[112] S. M. Erfani, S. Rajasegarar, S. Karunasekera, and C. Leckie, "High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning," *Pattern Recognit.*, vol. 58, pp. 121–134, Oct. 2016.
[113] S. S. Sivatha Sindhu, S. Geetha, and A. Kannan, "Decision tree based light weight intrusion detection using a wrapper approach," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 129–141, Jan. 2012.
[114] G. Kim, S. Lee, and S. Kim, "A novel hybrid intrusion detection method integrating anomaly detection with misuse detection," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1690–1700, Mar. 2014.
[115] W. Meng, W. Li, and L.-F. Kwok, "EFM: Enhancing the performance of signature-based network intrusion detection systems using enhanced filter mechanism," *Comput. Secur.*, vol. 43, pp. 189–204, Jun. 2014.
[116] C. Tang, W. Li, P. Wang, and L. Wang, "Online human action recognition based on incremental learning of weighted covariance descriptors," *Inf. Sci.*, vol. 467, pp. 219–237, Oct. 2018.
[117] M. Ristin, M. Guillaumin, J. Gall, and L. Van Gool, "Incremental learning of random forests for large-scale image classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 3, pp. 490–503, Mar. 2016.
[118] H. Du, S. Teng, M. Yang, and Q. Zhu, "Intrusion detection system based on improved SVM incremental learning," in *Proc. Int. Conf. Artif. Intell. Comput. Intell. (AICI)*, 2009, pp. 23–28.
[119] Y. Yi, J. Wu, and W. Xu, "Incremental SVM based on reserved set for network intrusion detection," *Expert Syst. Appl.*, vol. 38, no. 6, pp. 7698–7707, Jun. 2011.
[120] W. Zhuang, L. Xiao, J. Cui, and W. Zhuang, "Support vector machine based on incremental learning for malware detection," in *Proc. Int. Conf. Comput. Sci. Intell. Commun. (CSIC)*, 2015, pp. 204–207.
[121] J. Li, D. Xue, W. Wu, and J. Wang, "Incremental learning for malware classification in small datasets," *Secur. Commun. Netw.*, Feb. 2020, Art. no. 6309243.
[122] J. Dromard, G. Roudiere, and P. Owezarski, "Online and scalable unsupervised network anomaly detection method," *IEEE Trans. Netw. Service Manage.*, vol. 14, no. 1, pp. 34–47, Mar. 2017.
[123] M. Landauer, M. Wurzenberger, F. Skopik, G. Settanni, and P. Filzmoser, "Dynamic log file analysis: An unsupervised cluster evolution approach for anomaly detection," *Comput. Secur.*, vol. 79, pp. 94–116, Nov. 2018.
[124] A. O. Prokofiev, Y. S. Smirnova, and V. A. Surov, "A method to detect Internet of Things botnets," in *Proc. IEEE Conf. Russian Young Researchers Electr. Electron. Eng. (EIConRus)*, Jan. 2018, pp. 105–108.
[125] M. Mazini, B. Shirazi, and I. Mahdavi, "Anomaly network-based intrusion detection system using a reliable hybrid artificial bee colony and AdaBoost algorithms," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 31, no. 4, pp. 541–553, Oct. 2019.
[126] V. H. Bezerra, V. G. T. da Costa, S. B. Junior, R. S. Miani, and B. B. Zarpelão, "IoTDS: A one-class classification approach to detect botnets in Internet of Things devices," *Sensors*, vol. 19, no. 14, pp. 1–27, 2019.
[127] R. U. Khan, X. Zhang, R. Kumar, A. Sharif, N. A. Golilarz, and M. Alazab, "An adaptive multi-layer botnet detection technique using machine learning classifiers," *Appl. Sci.*, vol. 9, no. 11, p. 2375, Jun. 2019.
[128] W. Wang, Y. Shang, Y. He, Y. Li, and J. Liu, "BotMark: Automated botnet detection with hybrid analysis of flow-based and graph-based traffic behaviors," *Inf. Sci.*, vol. 511, pp. 284–296, Feb. 2020.
[129] T. Y. Yang, S. Y. Liu, and J. Y. Liu, "Network traffic anomaly detection based on incremental possibilistic clustering algorithm," *J. Phys., Conf. Ser.*, vol. 1284, no. 1, 2019, Art. no. 012067.
[130] A. Al Shorman, H. Faris, and I. Aljarah, "Unsupervised intelligent system based on one class support vector machine and grey wolf optimization for IoT botnet detection," *J. Ambient Intell. Humanized Comput.*, vol. 11, no. 7, pp. 2809–2825, Jul. 2020.
[131] A. Ferdowsi and W. Saad, "Generative adversarial networks for distributed intrusion detection in the Internet of Things," 2019, pp. 1–6, *arXiv:1906.00567*. [Online]. Available: http://arxiv.org/abs/1906.00567
[132] D. Sun, K. Yang, Z. Shi, and C. Chen, "A new mimicking attack by LSGAN," in *Proc. IEEE 29th Int. Conf. Tools with Artif. Intell. (ICTAI)*, Nov. 2017, pp. 441–447.
[133] C. Ma, X. Du, and L. Cao, "Analysis of multi-types of flow features based on hybrid neural network for improving network anomaly detection," *IEEE Access*, vol. 7, pp. 148363–148380, 2019.
[134] S. Huang and K. Lei, "IGAN-IDS: An imbalanced generative adversarial network towards intrusion detection system in ad-hoc networks," *Ad Hoc Netw.*, vol. 105, Aug. 2020, Art. no. 102177.
[135] S. Saraeian and M. M. Golchi, "Application of deep learning technique in an intrusion detection system," *Int. J. Comput. Intell. Appl.*, vol. 19, no. 2, Jun. 2020, Art. no. 2050016.
[136] T. Merino, M. Stillwell, M. Steele, M. Coplan, J. Patton, A. Stoyanov, and L. Deng, "Expansion of cyber attack data from unbalanced datasets using generative adversarial networks," in *Proc. Int. Conf. Softw. Eng. Res., Manage. Appl.* Cham, Switzerland: Springer, 2020, pp. 131–145.
[137] S. Manimurugan, S. Al-Mutairi, M. M. Aborokbah, N. Chilamkurti, S. Ganesan, and R. Patan, "Effective attack detection in Internet of medical things smart environment using a deep belief neural network," *IEEE Access*, vol. 8, pp. 77396–77404, 2020.
[138] J. Kim, J. Kim, H. Kim, M. Shim, and E. Choi, "CNN-based network intrusion detection against denial-of-service attacks," *Electronics*, vol. 9, no. 6, pp. 1–21, 2020.
[139] J. H. Lee and K. H. Park, "GAN-based imbalanced data intrusion detection system," *Pers. Ubiquitous Comput.*, pp. 1–8, Nov. 2019.
[140] L. Han, Y. Sheng, and X. Zeng, "A packet-length-adjustable attention model based on bytes embedding using flow-WGAN for smart cybersecurity," *IEEE Access*, vol. 7, pp. 82913–82926, 2019.
[141] Y. Yang, K. Zheng, B. Wu, Y. Yang, and X. Wang, "Network intrusion detection based on supervised adversarial variational auto-encoder with regularization," *IEEE Access*, vol. 8, pp. 42169–42184, 2020.
[142] S. Velliangiri, P. Karthikeyan, and V. V. Kumar, "Detection of distributed denial of service attack in cloud computing using the optimization-based deep networks," *J. Experim. Theor. Artif. Intell.*, pp. 1–20, Apr. 2020.
[143] K. Jiang, W. Wang, A. Wang, and H. Wu, "Network intrusion detection combined hybrid sampling with deep hierarchical network," *IEEE Access*, vol. 8, pp. 32464–32476, 2020.

[144] N. Moukafih, G. Orhanou, and S. El Hajji, "Neural network-based voting system with high capacity and low computation for intrusion detection in SIEM/IDS systems," *Secur. Commun. Netw.*, vol. 2020, Jul. 2020, Art. no. 3512737.
[145] M. Yousefi-Azar, L. G. C. Hamey, V. Varadharajan, and S. Chen, "Malytics: A malware detection scheme," *IEEE Access*, vol. 6, pp. 49418–49431, 2018.
[146] C. D. McDermott, F. Majdani, and A. V. Petrovski, "Botnet detection in the Internet of Things using deep learning approaches," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2018, pp. 1–8.
[147] A. Pekta and T. Acarman, "Deep learning to detect botnet via network flow summaries," *Neural Comput. Appl.*, vol. 31, no. 11, pp. 8021–8033, Nov. 2019.
[148] M. Yeo, Y. Koo, Y. Yoon, T. Hwang, J. Ryu, J. Song, and C. Park, "Flow-based malware detection using convolutional neural network," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Jan. 2018, pp. 910–913.
[149] A. N. Jahromi, S. Hashemi, A. Dehghantanha, K.-K.-R. Choo, H. Karimipour, D. E. Newton, and R. M. Parizi, "An improved two-hidden-layer extreme learning machine for malware hunting," *Comput. Secur.*, vol. 89, Feb. 2020, Art. no. 101655.
[150] A. S. Qureshi, A. Khan, N. Shamim, and M. H. Durad, "Intrusion detection using deep sparse auto-encoder and self-taught learning," *Neural Comput. Appl.*, vol. 32, no. 8, pp. 3135–3147, Apr. 2020.
[151] J. Kim, A. Sim, J. Kim, and K. Wu, "Botnet detection using recurrent variational autoencoder," 2020, *arXiv:2004.00234*. [Online]. Available: http://arxiv.org/abs/2004.00234
[152] M. V. O. Assis, L. F. Carvalho, J. Lloret, and M. L. Proença, "A GRU deep learning system against attacks in software defined networks," *J. Netw. Comput. Appl.*, vol. 177, Mar. 2021, Art. no. 102942.
[153] S. U. Rehman, M. Khaliq, S. I. Imtiaz, A. Rasool, M. Shafiq, A. R. Javed, Z. Jalil, and A. K. Bashir, "DIDDOS: An approach for detection and identification of distributed denial of service (DDoS) cyberattacks using gated recurrent units (GRU)," *Future Gener. Comput. Syst.*, vol. 118, pp. 453–466, May 2021.
[154] M. Lopez-Martin, B. Carro, and A. Sanchez-Esguevillas, "Application of deep reinforcement learning to intrusion detection for supervised problems," *Expert Syst. Appl.*, vol. 141, Mar. 2020, Art. no. 112963.
[155] C. U. O. Kumar and P. R. K. S. Bhama, "Detecting and confronting flash attacks from IoT botnets," *J. Supercomput.*, vol. 75, no. 12, pp. 8312–8338, 2019.
[156] S. Otoum, B. Kantarci, and H. T. Mouftah, "On the feasibility of deep learning in sensor network intrusion detection," *IEEE Netw. Lett.*, vol. 1, no. 2, pp. 68–71, Jun. 2019.
[157] P. Wei, Y. Li, Z. Zhang, T. Hu, Z. Li, and D. Liu, "An optimization method for intrusion detection classification model based on deep belief network," *IEEE Access*, vol. 7, pp. 87593–87605, 2019.
[158] C. Xu, J. Shen, X. Du, and F. Zhang, "An intrusion detection system using a deep neural network with gated recurrent units," *IEEE Access*, vol. 6, pp. 48697–48707, 2018.
[159] Y. Zhong, W. Chen, Z. Wang, Y. Chen, K. Wang, Y. Li, X. Yin, X. Shi, J. Yang, and K. Li, "HELAD: A novel network anomaly detection model based on heterogeneous ensemble learning," *Comput. Netw.*, vol. 169, Mar. 2020, Art. no. 107049.
[160] P. S. Muhuri, P. Chatterjee, X. Yuan, K. Roy, and A. Esterline, "Using a long short-term memory recurrent neural network (LSTM-RNN) to classify network attacks," *Information*, vol. 11, no. 5, pp. 1–21, 2020.
[161] C.-M. Hsu, M. Z. Azhari, H.-Y. Hsieh, S. W. Prakosa, and J.-S. Leu, "Robust network intrusion detection scheme using long-short term memory based convolutional neural networks," *Mobile Netw. Appl.*, pp. 1–8, Jul. 2020.
[162] G. De La Torre Parra, P. Rad, K.-K.-R. Choo, and N. Beebe, "Detecting Internet of Things attacks using distributed deep learning," *J. Netw. Comput. Appl.*, vol. 163, Aug. 2020, Art. no. 102662.
[163] T.-H. Lee, L.-H. Chang, and C.-W. Syu, "Deep learning enabled intrusion detection and prevention system over SDN networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2020, pp. 2–7.
[164] Y. H. Chen, Y. C. Lai, P. T. Jan, and T. Y. Tsai, "A spatiotemporal-oriented deep ensemble learning model to defend link flooding attacks in IoT network," *Sensors*, vol. 21, no. 4, pp. 1–29, 2021.
[165] Z. Liu and X. Yin, "LSTM-CGAN: Towards generating low-rate DDoS adversarial samples for blockchain-based wireless network detection models," *IEEE Access*, vol. 9, pp. 22616–22625, 2021.
[166] S. Zavrak and M. Iskefiyeli, "Anomaly-based intrusion detection from network flow features using variational autoencoder," *IEEE Access*, vol. 8, pp. 108346–108358, 2020.
[167] K. Hara and K. Shiomoto, "Intrusion detection system using semi-supervised learning with adversarial auto-encoder," in *Proc. IEEE/IFIP Netw. Oper. Manage. Symp. (NOMS)*, Apr. 2020, pp. 1–8.
[168] G. R. G. Lanckriet, L. El Ghaoui, C. Bhattacharyya, and M. I. Jordan, "A robust minimax approach to classification," *J. Mach. Learn. Res.*, vol. 3, pp. 555–582, Dec. 2002.
[169] H. Xu, C. Caramanis, and S. Mannor, "Robust regression and Lasso," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 1801–1808.
[170] M. Paschali, S. Conjeti, F. Navarro, and N. Navab, "Generalizability vs. robustness: Adversarial examples for medical imaging," 2018, *arXiv:1804.00504*. [Online]. Available: http://arxiv.org/abs/1804.00504

NIVEDITA MISHRA received the master's degree in engineering from the Department of Electronics and Communication, Thapar University, Patiala. She is currently pursuing the Ph.D. degree with the Department of Electronics and Communication, Symbiosis Institute of Technology, Pune. Her research interests include the Internet of Things security, machine learning, deep learning, and optical communication.

SHARNIL PANDYA received the master's degree from Swinburne University, Australia, in 2009, and the Ph.D. degree in security (WSNs) from India, in 2015. He is currently with the Symbiosis Centre for Applied Artificial Intelligence and the Symbiosis Institute of Technology, India. His research interests include the Internet of Things, security, and privacy issues, healthcare informatics, and computer vision. He is a Reviewer for reputed journals, such as IEEE Transactions on Industrial Electronics, IEEE Internet of Things, IEEE Sensors Journal, and *Transactions on Emerging Telecommunications Technologies*.