Machine Learning and Deep Learning Techniques for IoT-based Intrusion Detection Systems: A Literature Review

Laiby Thomas¹ & Subramanya Bhat²
¹Research Scholar, College of Computer Science and Information Science, Srinivas University, Mangaluru, India, and Assistant Professor, Dept. of Computer Science, NIMIT, Pongam, Kerala, India
OrcidID: 0000-0002-2608-3866; E-Mail: laibymary@gmail.com
²College of Computer Science and Information Science, Srinivas University, Mangaluru, India. OrcidID: 0000-0003-2925-1834; E-Mail: itsbhat@gmail.com

Area/Section: Computer Science.
Type of the Paper: Literature Review.
Type of Review: Peer Reviewed as per [C]O[PE] guidance.
Indexed in: OpenAIRE.
DOI:
Google Scholar Citation: IJMTS

How to Cite this Paper:
Laiby Thomas & Subramanya Bhat, (2021). Machine Learning and Deep Learning Techniques for IoT-based Intrusion Detection Systems: A Literature Review. International Journal of Management, Technology, and Social Sciences (IJMTS), 6(2), 296-314. DOI: 

International Journal of Management, Technology, and Social Sciences (IJMTS)
A Refereed International Journal of Srinivas University, India.

CrossRef DOI: https://doi.org/10.47992/IJMTS.2581.6012.0172

© With Author.

This work is licensed under a Creative Commons Attribution-Non-Commercial 4.0 International License subject to proper citation to the publication source of the work.
Disclaimer: The scholarly papers as reviewed and published by the Srinivas Publications (S.P.), India are the views and opinions of their respective authors and are not the views or opinions of the SP. The SP disclaims of any harm or loss caused due to the published content to any party.
Machine Learning and Deep Learning Techniques for IoT-based Intrusion Detection Systems: A Literature Review

Laiby Thomas¹ & Subramanya Bhat²

¹Research Scholar, College of Computer Science and Information Science, Srinivas University, Mangaluru, India, and Assistant Professor, Dept. of Computer Science, NIMIT, Pongam, Kerala, India
OrcidID: 0000-0002-2608-3866; E-Mail: laibymary@gmail.com
²College of Computer Science and Information Science, Srinivas University, Mangaluru, India. OrcidID: 0000-0003-2925-1834; E-Mail: itsbhat@gmail.com

ABSTRACT

Purpose: The authors attempt to examine the work done in the area of Intrusion Detection System in IoT utilizing Machine Learning/Deep Learning technique and various accessible datasets for IoT security in this review of literature.

Methodology: The papers in this study were published between 2014 and 2021 and dealt with the use of IDS in IoT security. Various databases such as IEEE, Wiley, Science Direct, MDPI, and others were searched for this purpose, and shortlisted articles used Machine Learning and Deep Learning techniques to handle various IoT vulnerabilities.

Findings/Result: In the past few years, the IDS has grown in popularity as a result of their robustness. The main idea behind intrusion detection systems is to detect intruders in a given region. An intruder is a host that tries to connect to other nodes without permission in the world of the Internet of Things. In the field of IDS, there is a research gap. Different ML/DL techniques are used for IDS in IoT. But it does not properly deal with complexity issues. Also, these techniques are limited to some attacks, and it does not provide high accuracy.

Originality: A review had been executed from various research works available from online databases and based on the survey derived a structure for the future study.

Paper Type: Literature Review.

Keywords: Internet of Things (IoT), Attacks, Machine Learning (ML), Deep Learning (DL), Intrusion Detection System (IDS), ABCD Analysis

1. INTRODUCTION:

The Internet of Things (IoT) has recently attracted a lot of attention due to its unique applications and aid in a range of fields, such as industrial operations, medical care, automation, smart surroundings, and so on [1]. The internet of things (IoT) is a network of physical objects with sensors, software, and connections that allow them to communicate with other networked devices over the internet. The capacity to monitor and operate such devices from afar has sparked a slew of new applications in sectors as diverse as connected industrial and manufacturing systems, smart homes, wearable gadgets, health surveillance systems, and energy management systems, to name a few [2-5]. The IoT is an unavoidable part of our daily lives. There are several items available to make our lives easier and more convenient, such as virtual companions [6]. In the Telecommunications era, the trending use of IoT is accelerating in a dramatic fashion to link with a variety of gadgets. Recently, there has been a surge in interest in examining the potential of Machine Learning technology, to assess and improve the security associated with digital devices [7, 8]. This scenario has led to the dissemination of large volume of data aimed at preventing such dangers. Applying AI to the IoT is one of the approaches for reducing the number of intensive job requests for each procedure. Despite the fact that the IoT offers a diverse set of services and applications, it is vulnerable to cyber-attacks. The major concerns in IoT systems are the security of the physical devices and safeguarding the data from external threats and attacks. Cyber assaults are the deliberate exploitation or illegal access to another person's or organization's information or infrastructure. The diversity of IoT devices and
protocols, devices openly interfacing with internet, and the limited computing power on devices makes it difficult to protect IoT devices from assaults [9-11].

Because the Internet of Things is a diverse ecosystem with a lack of interoperability, traditional security solutions are ineffective [12]. However, other aspects of IoT security concerned with maintaining integrity, confidentiality, and secure user access of data are improved. These security methods, despite being designed with the user and the IoT ecosystem in mind, contain a number of shortcomings. As a result, a separate module is required to provide IoT network security. An example of a security measure that is already being used in wireless networks is intrusion detection system (IDS) [13, 14]. Adding IDS features to wireless networks will aid IoT in protecting the network from assaults and other issues.

The proposed IT network mitigation strategy is incompatible with the Internet of Things. Machine learning or deep learning-based solutions for detecting threats by analysing IoT network traffic patterns are now in use, although they are few and few between. These techniques are appropriate because these are used for data classification, grouping, and anomaly detection [15]. Machine Learning is a method for extracting knowledge patterns from well-defined inputs without using any mathematical procedures to generate logical outputs.

The most popular industry use was cost-cutting measures, even for small automated processes that were controlled by sensors that were all connected to a single device. In the Internet of Things, where environmental inputs are received from several sensors, a single stream is incorporated with a broader application [16, 17]. In this regard, not only having current data, but also having huge volume of historical data to analyse and learn from is essential to make accurate predictions. Machine learning can be applied to several sectors, especially in the Internet of Things, where it offers numerous cost-cutting benefits [18]. Machine learning also plays an important part in industry applications. The greater predictive measure on a typical example of large machinery with sensors, which enables Machine Learning algorithms to detect flaws with high accuracy, as well as lower maintenance costs and carry over with on time [19].

2. RESEARCH OBJECTIVE AND METHODOLOGY :

The process of detecting intrusions in IoT networks has grown considerably more difficult since the attack behavior and medium of propagation used by malwares such as email and social networking platforms have become much more complicated. Many existing intrusion detection systems (IDS) are unable to detect any backdoor ransomware that has been placed on a system. This weakness in many IDS has prompted interest in creating a security solution to implement, in case the existing IDS fails to detect an intrusion. Machine Learning algorithms can serve as diverse and effective tools in this use case. A literature review has been undertaken to understand IDSs in detail and how Machine Learning can be integrated with IDS. The role of Machine Learning in IoT security is also investigated.

The following are the objectives of this study:

1. The importance of IoT Security
2. Different threats to IoT security
3. Importance of IDS
4. The role of deep learning and machine learning in IDS design
5. Various datasets involved in IoT security.

3. OVERVIEW OF IDS IN IoT :

Because of the disparate nature of IoT architecture it is very difficult to build an effective security mechanism. Many of the methods developed are giving importance to authentication, confidentiality access control, etc., but still the IoT ecosystem is vulnerable to attacks. In such situations Intrusion Detection System plays an important role which is already used in traditional networks. In this section, we present an architecture of IoT ecosystem, different attacks in IoT, role of IDS in IoT architecture and various datasets which can be used in IoT security.

3.1 IoT architecture:

During these years, IoT has become the priority of many organizations. IoT can connect together a number of devices and hence it could acquire many functionalities. According to DataProt [20], 10 billion or more IoT devices will be active by 2021 and by 2025, every minute, an estimated 152,200 devices will connect to the internet. IoT and its heterogeneity demands a layered architecture. Many
research works propose architecture designs for the IoT environment and the key points are discussed in this session.

The most common model is a layered architecture consisting of different layers as depicted in Figure 2. 3-layered, 4-layered and 5-layered architecture is proposed by Javed et al. in [21]

ITU-T proposed a reference model for IoT, shown in fig 3 which consists of four layers and in each layer could be able to manage the security faults.

[22], in their research gives the functions of each layer in IoT architecture. Perception layer, network layer, and application layer are the three layers that make up a typical IoT architecture. The perception layer could be referred to as the 'device layer' because it is made up of actual devices. The perception layer transfers information to the upper layers through interfaces. Data transfer is handled by the network layer, often known as the 'transmission layer. The network layer will determine the necessary routes after getting the processed data from the perception layer. In network layer, various communication technologies and networking devices need to be integrated, and
is an important layer in the architecture of IoT.
Application layer, dubbed ‘the business layer’, manages the various applications in IoT systems. This layer will receive the data from network layer and will provide services accordingly.

3.2 IoT Attacks:
Internet technology has expanded its application field in many different domains in our lives over the last few decades, including banking operations, on-line auctions, electronic commerce applications, social networking, and on-line applications, among others. Existing security holes in computer systems, on the other hand, have enabled hackers to access many electronic networks using denial-of-service (DoS) or distributed denial-of-service (DDoS) attacks. Anomaly detection systems are primarily built on depicting usual user, host, and network connection behaviour. It takes a lot of effort to spot unusual requests in the network. The number and complexity of different types of network assaults is continually increasing. Despite the fact that much effort has been done on attack detection, a network cannot be guaranteed to be attack-free.

Confidentiality, integrity, availability, and authorization are all major concerns with IoT security. The combination of real-world objects with IoT has the potential risk of introducing plenty of cyber security risks into everyday operations. Denial of service (DoS) and man-in-the-middle (MITM) attacks can be used to target critical infrastructure in the IoT. They can compromise any device, including the main server, which can cause the entire system to shut down if it is hacked [23]. IDS, which is already an important tool in network security, acts as a critical component of the IoT security architecture for traditional networks as well as information systems. Not only does it detect known attacks, but it also detects unknown ones. There are two different categories of intruders:

1. External Intruders – These are persons who are not part of the network and thus do not have network access. They get access to computers by transmitting malware or exploiting vulnerabilities.
2. Internal Intruders – These individuals have network access rights and privileges, but they are abusing them. Changing the substance of crucial data or stealing confidential data are examples of these types of attacks. All of these risks can be carried out physically by breaking into a computer system or gaining remote access to the network without authorization.

IoT is prone to four types of threats:

1. Denial of Service (DoS) – By injecting worthless or undesired traffic, this threat denies or stops users from accessing network resources.
2. Malware – To interrupt devices on the IoT network, attackers employ executable code. They could collect sensitive data or obtain illegal access to equipment. The attacker can take advantage of holes in the devices’ firmware and use their software to create disruption with the IoT architecture.
3. Data breaches – This is the attack in which network data that is sensitive, protected, or confidential is recovered. Spoofing ARP packets allows attackers to listen in on network conversations between peers.
4. Weakening Perimeters – Currently, the design of IoT network equipment does not conform to the concept of ubiquitous security. As network security features are not usually integrated to gadgets, the network is resistant to threats.

In [24] they have given a detailed description about IoT attacks and threats in different layers and it is depicted in Fig.4.
3.3 IDS in IoT:
An Intrusion Detection System (IDS) is anticipated to recognize both known as well as new assaults and alert sensor hubs to safeguard IoT from a variety of security issues. When an interruption occurs, IDS detects suspicious or unusual activities and issues a warning. In this section, an overview of Intrusion Detection System and its role in IoT is given. Any piece of software or hardware designed for the detection of attacks or malicious events directed towards the network or the system as a whole can be considered an intrusion detection system [25]. An IDS collects data from different sources and after examining, if any security faults are noticed it will give an alert about the intrusion to the admin people. Signature-based Intrusion Detection System (SIDS) and Anomaly-based Intrusion Detection System (AIDS) are the two forms of IDS (AIDS). The authors give a full review of IDS for IoT systems in [26], and figure 5 depicts the classification of IDS for IoT as presented in the publication.

**Fig. 5:** Classification of IDSs for IoT

**Signature-based intrusion detection systems (SIDS)**
For already existing intrusions, signature-based detection, also known as knowledge-based detection, can deliver superior results. Pattern matching is used by SIDS to detect a known attack [27]. A database stores the known intrusion signatures and if a positive match is found, an alarm gets triggered. Fig.6 shows the working of an SIDS. For previously known attacks, SIDS provides excellent results. SIDS can identify intrusions with the least incidence of false alarms (FA) and it has a simple design. SIDS is inefficient if a zero-day attack comes since no signatures are available. As the traditional intrusion detection model is incapable of addressing this issue, Anomaly-based Intrusion Detection System (AIDS) can be used in its place.

**Fig. 6:** Signature-based Intrusion Detection System (SIDS)

**Anomaly-based Intrusion Detection System (AIDS)**
To detect new attacks, AIDS is the preferred type. It could also be used to create intrusion signatures. Instead of looking at the signature database it will look into the behaviour of the devices and interfaces. If any deviation or anomaly occurs AIDS will trigger an alert. The process of detecting intrusion has grown considerably more difficult since the attack behaviour and medium of propagation used by
malwares such as email and social networking platforms have become much more complicated. Generally, the implementation of AIDs occurs as two separate phases. The first phase involves training, during which the typical behaviour is learned. In the second phase, which is the testing phase, previously unseen intrusions are learned using a new dataset. AIDs can be divided into statistics-based, knowledge-based, and machine-learning methods [28].

An IDS identifies a breach before, during, or after it has occurred. The IDS is known to have a significant frequency of false alarms based on anomalies. Efforts are being made to cut down on the large number of false positives. Intrusion detection is a data analysis process that may be examined as a problem of proper data classification. From the perspective of anomaly-based IDS, this means that if we extract characteristics that can clearly demarcate normal data samples from the abnormal data, the false positives can be greatly reduced. Similarly, most intrusion detection approaches based on data mining or Machine Learning employ well-known methodologies and technologies. However, the more data there is, the longer it takes to analyze it, delaying the detection of assaults. An IDS will be more useful if it can raise an alarm early enough to limit the harm that a persistent assault can cause. As a result, IDS must be as quick as possible when operating online. This is thought to be possible if we can reduce the amount of data to be evaluated while maintaining its quality. Researchers have been inspired to use distributed IDSs in combination with other ML techniques as a result of recent advances in intelligent systems. To assist protect the system framework, use machine learning inside an IoT gateway to handle the challenges of safeguarding IoT devices.

Fig. 7: Classification of AID methods

Fig.8 shows the Machine Learning based IDS working approach. The two most prevalent ML approaches are supervised and unsupervised, which are based on the type of learning.

3.4 Different ML/DL Techniques for IDS:
This section outlines the various ML/DL algorithms that are useful for intrusion detection in IoT. In general, an IDS is made up of three modules: (1) gathering, (2) analysis, and (3) reporting [29].

The Data Gathering module collects data from the IoT system and even at this stage, the behavior can be examined and at the malicious behavior, if any, can be detected if possible.

The analysis Module will be processing the data which may contains the evidence of the attack and could detect the attacks. In this phase, ML/DL models can be leveraged for analysis of anomalies. These techniques are able to identify and predict a new attack after learning from the previous cases.

The Reporting module -When an anomalous activity is investigated, the Reporting module is a method that will report an assault. The structure of an IDS based on ML/DL techniques is shown in Figure 9.
3.4.1 ML techniques used in IoT IDS

Many studies have been directed towards the adoption of machine learning techniques in the IoT environment. There are several Machine Learning approaches suitable for classification tasks, including decision trees, support vector machines, rule-based systems, and neural networks, but these systems should be able to handle novel records even if they can train the data. Figure 10 shows an overview of machine learning approaches used in IDSs for IoT systems.

3.4.2 DL techniques used in IoT IDS

Due to its large data handling capacity, DL algorithms can also effectively be used in IoT security. In most cases DL algorithms master ML techniques due to its large data set handling capacity and its capability to model complex features from sample data [30]. DL algorithms can be easily linked to an IoT network, so that without any human intervention it can perform the assigned tasks. As shown in Figure 11, DL can be used in IDS by combining different techniques.

3.5 Datasets used for IoT Security:

Datasets are critical in determining the accuracy and efficacy of an IDS project. This section will highlight some of the most important publicly available datasets.
1) KDD99: - Knowledge Discovery and Datamining (KDD) cup-99 dataset, was launched in 2008. KDD99 is a publicly available dataset, popular in the domain of network intrusion detection (NIDS). This dataset contains 41 features that can be readily processed by ML algorithms [31]. The dataset is available in three versions: one is the full dataset and 10% subsets of both the training set and the test set (https://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html) [31].

2) NSL-KDD: - This dataset was launched in 2009 and is an enhancement over KDD99. This dataset has fixed the issue of duplicate records found in the original KDD99 dataset. The more balanced resampling of KDD99 data helped to get less biased results when applying the classification algorithms [32].

3) Bot-IoT: - The Bot-IoT dataset was developed to address both normal and botnet traffic. It replicates the real-world IoT ecosystem, including attack scenarios including data collecting, data theft, and denial of service. [33]. The statistical measures to be used in evaluating the features of the Bot-IoT dataset are also provided by the authors [34].

4) N-BaloT Dataset: - This is the most recent one among IoT IDS datasets. This dataset contains traffic from two networks: one from a video camera and the other from an IoT network. Authors give detailed description of attacks and related works in their work [35].

4. LITERATURE REVIEW:

In the RPL-based IoT, SD Bhosale et al. [36] proposed a real time implementation of an IDS for wormhole attack in the RPL based IoT. Wormhole attack is one of the most serious attacks that occurs at the RPL network&#39;s 6LoWPAN adaption layer. In this sort of attack, a group of attacker nodes creates a tunnel between two nodes that appears to be directly connected, causing network traffic to be misdirected. Using Cooja Simulator, the proposed IDS is implemented in Contiki OS. The attack and attacker node were identified using the received signal strength indicator (RSSI).

RK Deka et al. [37] presented a parallel cumulative ranker algorithm to rank the attributes of a dataset for cost-effective classification of network traffic. Also mentioned was the value of active learning in training an SVM binary classifier for detection of DDoS attack traffic in an unsupervised manner by an expert module. To achieve high accuracy in network traffic classification, the suggested method picks small batches of training samples from a dataset. With fewer training samples, our method on huge data improves classification accuracy.

S Rathore et al. [38] presented a distributed attack detection system for IoT, based on semi-supervised learning. For attack detection, ELM-based Semi-supervised Fuzzy C-Means (ESFCM) algorithm is suggested. The ESFCM technique is a combination of semi-supervised fuzzy c-means algorithm and an Extreme Learning Machine (ELM), which gives great generalization performance at a faster detection rate, to cope with the labelled data issue. Fog computing is similar to cloud computing and can detect network edge attacks and distributed attacks.

For IoT security, Y Liu et al. provided a design and study of probing routes to defend against sink-hole attacks (PRDSA). The PRDSA technique uses a routing approach that combines far-sink reverse routing, equal-hop routing, and minimum hop routing to avoid sinkholes and identify the safest route to the actual sink, thus enabling the detection of sink-holes with greater accuracy. The drawbacks of the previous schemes that prevented them from locating the sinkhole have been overcome in PRDSA. The PRDSA technique requires the nodes and the sink node to return the signature of the information so that the sinkhole's location can be determined. In addition, the PRDAS approach is based mostly on network energy consumption characteristics.

B Xu et al. [40] suggested a security design mechanism for detecting buffer overflow attacks in Internet of Things (IoT) devices. As the proposed alternatives do not alter the compiler or existing instruction set, they do not impose any restrictions on software developers. The automatic extraction tools extract the monitoring model and secure tag of each memory segment throughout the build process. During execution, the dynamic execution trail is evaluated by the intended hardware to see if it conforms to the authorized behaviour; if it doesn't, the appropriate response mechanisms are triggered.

An adaptive hybrid IDS based on a timed automata controller technique was developed by S Venkatraman and B Surendiran [41]. The suggested Hybrid IDS possessed additional knowledge of common multimedia file types, which it applied to a full evaluation of packets containing multimedia
data. The authors also designed a crowd-sourced online library for signature-based harmful pattern set generation and a self-tuning timed automaton to identify the intruder in IoT networks.

To prevent intrusions, R Patil et al. [42] created a virtual environment monitoring system. Security is a major issue in today's developments in virtualization-based computer systems. Attacks could originate at the system-level or network-level and the virtual environments operating in the system must be protected to improve the overall safety of internet computing services. It looked for weaknesses in the system and network models of newly connected virtual machines.

RAR Ashfaq et al. [43] experimented with a unique fuzziness-based semi-supervised learning strategy to increase the classifier's performance for IDSs by introducing unlabeled samples to a supervised learning algorithm. The fuzzy membership vector was created with a simple feed-forward neural network having only one hidden layer, and the fuzzy quantity was utilized to categorize unlabeled data. After adding a category to the original training dataset, the classifier was retrained.

AA Diro et al. [44] developed a new cybersecurity strategy based on deep learning (DL) to detect threats in the social internet of things. Developments in CPU architecture and neural network models have facilitated the easy and quick adoption of DL. Deep learning can effectively extract features from the input data, making it a resilient strategy to detect even small changes and zero-day attacks. The compression and learning ability of DL architectures helps to identify hidden patterns in the training data, allowing assaults to be separated from innocuous traffic. The performance of the deep learning model is compared to that of a traditional ML approach, and the detection of distributed attacks is compared to that of a centralized system.

An IDS against web application malware was proposed by Alazab et al. [45]. The study's purpose was to minimize web application vulnerabilities; the most important security techniques were security detection and prevention. The majority of present research, on the other hand, concentrates on ways to avoid an attack at the web application layer, with less attention paid to how to respond in the event of an attack. The proposed Intelligent Intrusion Detection and Prevention System (IIDPS) was a mix of a Signature-based Intrusion Detection System (SIDS) and an Anomaly-based Intrusion Detection System (AIDS).

HA Bany Salameh et al. [46] investigated the channel assignment problem in jamming attacks, both of the reactive and proactive type. The authors presented a new probabilistic-based channel assignment technique that aimed to reduce the invalidity ratio of CR packet transmissions while keeping latency limitations in mind. Legacy wireless networks have the same security flaws as Cognitive Radio (CR) networks. Jamming was once thought to be a widespread attack. Networks can be jammed in proactive mode or reactive mode. The suggested approach used statistical data from licensed primary users' activities, fading circumstances, and jamming attacks on idle channels to give the most secure channels with the lowest invalidity ratios to connecting CR IoT devices.

N Tariq et al. [47] proposed an energy-efficient, software-defined network-based Mobile Code-driven Trust Mechanism (MCTM) to address the problem by evaluating SN trust based on their forwarding behaviours. To evaluate the trust, MCTM iterates through the SNs in the specified manner and collects relevant information about them. The widespread adoption of the Internet of Things (IoT) has also led to greater adoption of wireless Sensor Nodes (SNs), which have vulnerabilities just like any other hardware or software systems. The results of the studies indicated that the method proposed by the authors outperforms a state-of-the-art technique for energy-efficient SN management based on Software-Defined Network (SDN).

The performance of different machine learning models has been examined to accurately anticipate attacks and abnormalities on IoT systems by M Hasan et al. [48]. The authors evaluated Logistic Regression (LR), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), and Artificial Neural Network (ANN) algorithms. In the sphere of IoT, attack and anomaly detection in the IoT infrastructure is a growing concern. As the use of IoT infrastructure increases, the threats and attacks against IoT infrastructure also aggravate. Attacks and anomalies that might cause an IoT system failure include Denial of Service, Data Type Probing, Malicious Control, Malicious Operation, Scan,
Spying, and Wrong Setup.

To protect against DoS attacks, C Lyu et al. [49] suggested a selective authentication-based geographic opportunistic routing (SelGOR), which meets the requirements of authenticity and dependability in WSNs. SelGOR uses an SSI-based trust model to increase data delivery efficiency by studying statistic state information (SSI) of wireless networks. SelGOR, unlike earlier opportunistic routing protocols, uses an entropy-based selective authentication mechanism to preserve data integrity. It was able to isolate DoS attackers, while lowering computational costs. A distributed cooperative verification system was also created to speed up the isolation of attackers. SelGOR additionally avoids redundant data transfer and unnecessary signature verification caused by opportunistic routing owing to this method.

TND Pham et al. [50] suggested a method for detecting flooding attacks in delay-tolerant networks and tolerating burst traffic. The proposed method used encounter records (FDER) to detect flooding attacks while allowing valid burst traffic at the same time. Nodes share their encounter records (ER), which keep track of the messages they've sent during past encounters. This ER history can be utilized to calculate a node's new message transmission rate and number of forwarded replicas per message over time.

Under replay attacks, B Chen et al. [51] established a safe fusion estimation for bandwidth constrained devices. MIFE (Multisensor Information Fusion Estimation) is an appealing alternative for studying secure estimate problems because it has the potential to improve estimation accuracy while also improving reliability and robustness against attacks. The secure distributed Kalman fusion estimation problem was also investigated from the defender's perspective. To characterize replay attacks and bandwidth restrictions, a novel mathematical model with compensation technique was presented, and then a recursive distributed Kalman fusion estimator (DKFE) in the linear minimum variance sense was designed.

4.1 Summary of Related Work:

Table1: Summary of findings from 2015-2021 presented by various authors and Comparison of existing techniques.

| Reference | Author | Type of attack addressed | Techniques Used | Advantages | Disadvantages |
|-----------|--------|--------------------------|-----------------|------------|--------------|
| [36]      | Deshmukh-Bhosale, S. and Sonavane, S.S. (2019) | wormhole attack | Hybrid approach | RSSI is high | For high number of network, the detection rate is low |
| [37]      | Deka, R.K., Bhattacharyya, D.K. and Kalita, J.K. (2019) | DDoS attack | developed a parallel cumulative ranker technique for ranking dataset attributes for cost-effective network traffic classification. | with fewer training data, improves classification accuracy | It does not work on power system attack dataset |
| [38]      | Rathore, S. and Park, J. H. (2018) | distributed attack | A fog-computing based attack detection system and an | The fog computing paradigm is capable of supporting | It lacks the ability to self-learn and compress data. |
| Reference | Authors | Attack Type | Approach Description | Benefits | Drawbacks |
|-----------|---------|-------------|-----------------------|----------|-----------|
| [39]      | Liu, Y et al (2018) | Sinkhole attack | The PRDSA technique is suggested to prevent Sinkhole attacks and provide IoT security. | The proposed method outperforms previous methods. | The proposed scheme’s parameters aren't well optimised. |
| [40]      | Xu, B., Wang et al (2018) | Buffer overflow attacks | A security hardware design with architectural enhancements to detect buffer overflow attacks. | This approach is appropriate for IoT devices with minimal resources and high security requirements. | The detection speed can be improved |
| [41]      | Venkatraman, S. and Surendiran, B. (2020) | DoS attack | Based on a timed automata controller technique, adaptive hybrid IDS is used. | It has a profound impact on the long-term viability of cyberspace and our smart society. | Detection accuracy can be improved |
| [42]      | Patil, R. and Modi, C. (2019) | Zero-day attacks | To prevent intrusions, created a virtual environment monitoring system. | It meets the virtual environment's security requirements on both the system and network levels. | The detection rate can be improved |
| [43]      | Ashfaq et al (2016) | Classification of attacks | Using unlabeled data and a supervised learning algorithm, a unique fuzziness-based semi-supervised learning strategy is developed. | Normal and anomaly problems were detected and reported. | It does not detect multiple type of attacks |
| [44]      | Diro et al (2018) | Distributed attack detection | To enable the identification of assaults in social IoT, a new deep | Distributed attack detection was found to be better than | The performance can be improved |
| Reference | Authors | Type of Attack | Description | Methodology | Performance |
|-----------|---------|----------------|-------------|-------------|-------------|
| [45]      | Ammar Alazab, Michael Hobbs, et al (2014) | Web application vulnerability | IIDPS is a syndrome that combines SIDS and AIDS. | IIDPS can detect and prevent a wide range of inappropriate behaviour. | The performance of the proposed system can be improved |
| [46]      | Salameh, H.A.B., Almajali et al (2018) | Jamming Attack | A new channel assignment technique based on probabilistics. | Under time restrictions, reducing the invalidity ratio of CR packet transfers. | The quality of CR links across multiple channels is not taken into account. |
| [47]      | Tariq, N., Asim, M, et al (2019) | Internal attacks targeting to IoT | Energy-efficient, software-defined-network-based Mobile Code-driven Trust Mechanism (MCTM) | Detecting and isolating harmful internal SNs in IoT applications that use SNs | Routing attacks such as Sybil, sink hole, and wormhole are not explored for resource-constrained SNs. |
| [48]      | Lyu, C., Zhang, X., Liu, Z (2019) | DDoS attacks | The accuracy of many machine learning models in predicting attacks and abnormalities on IoT devices has been compared. | High system performance | Big data and other unknown issues are not taken into account. |
| [49]      | Lyu, C., Zhang, X., Liu, Z., & Chi, C. H. (2019) | DoS attack | Geographic opportunistic routing based on selective authentication (SelGOR) | As a result of opportunistic routing, avoid duplicate data transfer and redundant signature verification. | Behavior model of DoS attackers is not considered |
| [50]      | Pham et.al (2019) | flooding attack | FDER can identify flooding attacks while yet allowing valid burst traffic. | To accommodate a bursty traffic environment, provide performance fairness. | Does not reduce the impact caused by the adversaries |
| [51]      | Brun et.al (2018) | Threats to IoT enabled | Examines the different | High performance | It does not detect broad |
devices in home environment | possible attack vectors against IoT gateways. | range of attacks
---|---|---

5. DISCUSSION & FUTURE WORK:

This review tries to give an overview of IDS and its use in IoT Security. In various research works, researchers discussed different methods of designing IDS for IoT. In some of the studies their main focus is on layer wise security issues and they have made a detailed analysis on layer wise threats, security mechanisms required to control it. Majority of the research works use the publicly available datasets to demonstrate the evaluation of their work. When it comes to real-world datasets, the model's accuracy may suffer. It is also noted that it is difficult to develop an IDS which is addressable to all attacks. A systematic review of ML based IoT and its application in IoT security is the main focus of this study. By using ML-DL techniques, it may help in effective handling of IoT attacks. According to peer-reviewed articles, there is a great deal of study being done in this area of IDS, but scalability constraints in IDS deployment, network latency issues, and resource limits with IoT devices are still challenges for researchers. At the same time combined use of semi-supervised learning and reinforcement learning in the deployment of IDS in IoT security still need to be explored.

6. RESEARCH GAP:

After the literature survey, it is examined that the current methodology, models, and publications appear to be lagging in addressing the various challenges in the field of IoT security. There is a need to reduce the complexity in the techniques and also these techniques lack the accuracy. Researchers have been inspired to use distributed IDSs in combination with other ML techniques as a result of recent advances in intelligent systems. When it comes to dealing with the complexity of IDSs, traditional machine learning algorithms have a number of limitations. One method to ensure that IDSs deliver on their promise in the real world is to improve technology to address these flaws. During the review, it was determined that various aspects of IoT security do not effectively leverage machine learning.

7. RESEARCH AGENDA:

1. Which Machine Learning Algorithms are the most effective in combating assaults in the IoT environment?
2. What technology can improve data security in a large user network?
3. What system can prevent attacks and detect attacks over the network?
4. What new development framework will be able to connect IDS to the most effective machine learning algorithms?
5. What Machine Learning Technology can you recommend with a high detection rate, accuracy, and low false positive rate?
6. Which Machine Learning Algorithms are the best for creating an IDS for IoT Security?

8. ANALYSIS OF RESEARCH AGENDA:

This review paper explains the problems that are being addressed. Most of the IDS currently available are not suitable for IoT architecture because for detection and prevention they need more memory, power, and bandwidth. This means that if we extract characteristics that correctly distinguish normal data from aberrant data, the rate of false positives can be considerably reduced. Similarly, most data mining and machine learning-based intrusion detection systems employ well-known methodology and technologies. However, the more data there is, the longer it takes to analyze it, delaying the detection of assaults. As a result, machine learning algorithms are employed as a diverse and effective tool.

9. RESEARCH PROPOSAL:

Based on the study, identified gaps, and discussions, a model will be constructed to process the data and detect attacks early using ML/DL approaches.

10. ABCD ANALYSIS OF RESEARCH PROPOSAL:

ABCD analysis is primarily concerned with exploring advantages, Benefits, Constraints, and Demerits.
in a systematic manner [52-55]. We conducted an ABCD analysis of the IDS [56-59]

**Advantages:**
- Detection of attacks which could not be detected by firewalls or other security mechanisms.
- It can be a valid method to prevent network damage

**Benefits:**
- Intrusion detection systems can improve security responses.
- Pre-host Detection.
- Easier to deploy.

**Constraints:**
- Larger size of data
- Higher dimensionality is required
- Data preprocessing

**Disadvantages:**
- Reliability of the data
- Complexity in data handling
- Daily monitoring is required to analyze the findings of IDS.

11. **CONCLUSION:**

The Internet of Things (IoT) is a network of networked devices that allows for effortless data exchange among physical items, and its popularity has grown in recent years. Medical and healthcare equipment, autonomous vehicles, industrial robots, smart TVs, wearables, and smart city infrastructures are all examples of items that may be monitored and controlled remotely. At the same time, it has the potential to disrupt IoT network security and pose serious privacy and security risks to consumers. As a result, IoT networks require an efficient security system. Many security techniques are in use, but an Intrusion Detection System, similar to those used in traditional networks, is a useful tool for IoT. The most appropriate and reliable approach for identifying attacks in IoT networks is an IDS based on Machine Learning. In this review paper a survey is conducted on IDS in IoT architecture from papers published between 2014 and 2021. IoT architecture, different layers of IoT, different levels of protocols, threats and challenges, Machine Learning and Deep Learning based IDS, and various datasets suited for IoT security are all explored. In future a hybrid model for IDS in IoT environment is planned which need further analysis of various ML and DL techniques.

**REFERENCES:**

[1] Sarkar, S., Chatterjee, S., & Misra, S. (2015). Assessment of the Suitability of Fog Computing in the Context of Internet of Things. *IEEE Transactions on Cloud Computing, 6*(1), 46-59. Google Scholar

[2] Lu, Y., & Da Xu, L. (2018). Internet of Things (IoT) cybersecurity research: A review of current research topics. *IEEE Internet of Things Journal, 6*(2), 2103-2115. Google Scholar

[3] Moustafa, N., Turnbull, B., & Choo, K. K. R. (2018). An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things. *IEEE Internet of Things Journal, 6*(3), 4815-4830. Google Scholar

[4] Liu, C., Yang, J., Chen, R., Zhang, Y., & Zeng, J. (2011, July). Research on immunity-based intrusion detection technology for the Internet of Things. In 2011 *Seventh International Conference on Natural Computation, 1*(1), 212-216. IEEE. Google Scholar

[5] Al-Yaseen, W. L., Othman, Z. A., & Nazri, M. Z. A. (2017). Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system. *Expert Systems with Applications, 67*(1), 296-303. Google Scholar
Huda, S., Abawajy, J., Alazab, M., Abdollahian, M., Islam, R., & Yearwood, J. (2016). Hybrids of support vector machine wrapper and filter based framework for malware detection. *Future Generation Computer Systems, 55*(2), 376-390. [Google Scholar]

Zarpelão, B. B., Miani, R. S., Kawakani, C. T., & de Alvarenga, S. C. (2017). A survey of intrusion detection in Internet of Things. *Journal of Network and Computer Applications, 84*(4), 25-37. [Google Scholar]

Mienye, I. D., Sun, Y., & Wang, Z. (2019). Prediction performance of improved decision tree-based algorithms: a review. *Procedia Manufacturing, 35*(7), 698-703. [Google Scholar]

Chowdhury, A., Karmakar, G., & Kamruzzaman, J. (2019). The co-evolution of cloud and IoT applications: Recent and future trends. In *Handbook of Research on the IoT, Cloud Computing, and Wireless Network Optimization*, 213-234. IGI Global. [Google Scholar]

Saha, H. N., Mandal, A., & Sinha, A. (2017, January). Recent trends in the Internet of Things. In *2017 IEEE 7th annual computing and communication workshop and conference (CCWC)*, 1-4. IEEE. [Google Scholar]

Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity, 2*(1), 1-22. [Google Scholar]

Javaid, A., Niyaz, Q., Sun, W., & Alam, M. (2016). A deep learning approach for network intrusion detection system. *Eai Endorsed Transactions on Security and Safety, 3*(9), 1-5. [Google Scholar]

Huda, S., Abawajy, J., Alazab, M., Abdollahian, M., Islam, R., & Yearwood, J. (2016). Hybrids of support vector machine wrapper and filter based framework for malware detection. *Future Generation Computer Systems, 55*(3), 376-390. [Google Scholar]

Zarpelão, B. B., Miani, R. S., Kawakani, C. T., & de Alvarenga, S. C. (2017). A survey of intrusion detection in Internet of Things. *Journal of Network and Computer Applications, 84*(1), 25-37. [Google Scholar]

Ioulianou, P., Vasilakis, V., Moscholios, I., & Logothetis, M. (2018). A signature-based intrusion detection system for the Internet of Things. *Information and Communication Technology Form, 2*(1), 1-5. [Google Scholar]

Alnaghes, M. S., & Gebali, F. (2015, May). A Survey on Some Currently Existing Intrusion Detection Systems for Mobile Ad Hoc Networks. In *The Second International Conference on Electrical and Electronics Engineering, Clean Energy and Green Computing (EEECEGC2015)*, 12(1), 12-18. [Google Scholar]

Gomez, C., Arcia-Moret, A., & Crowcroft, J. (2018). TCP in the Internet of Things: from ostracism to prominence. *IEEE Internet Computing, 22*(1), 29-41. [Google Scholar]

Midi, D., Rullo, A., Mudgerikar, A., & Bertino, E. (2017, June). Kalis—A system for knowledge-driven adaptable intrusion detection for the Internet of Things. In *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, 656-666. IEEE. [Google Scholar]

Shafi, Q., Basit, A., Qaisar, S., Koay, A., & Welch, I. (2018). Fog-assisted SDN controlled
framework for enduring anomaly detection in an IoT network. *IEEE Access*, 6(2), 73713-73723.

[20] *Meet the team that brings you cybersecurity tips*. DataProt. (2021, April 23). Retrieved November 21, 2021, from https://dataprot.net/about-us/.

[21] Asharf, J., Moustafa, N., Khurshid, H., Debie, E., Haider, W., & Wahab, A. (2020). A review of intrusion detection systems using machine and deep learning in internet of things: Challenges, solutions and future directions. *Electronics*, 9(7), 11-77.

[22] Santos, L., Rabadao, C., & Gonçalves, R. (2018, June). Intrusion detection systems in Internet of Things: A literature review. In *2018 13th Iberian Conference on Information Systems and Technologies (CISTI)*, 1-7. IEEE.

[23] Prabavathy, S., Sundarakantham, K., & Shalinie, S. M. (2018). Design of cognitive fog computing for intrusion detection in Internet of Things. *Journal of Communications and Networks*, 20(3), 291-298.

[24] Chaabouni, N., Mosbah, M., Zemmari, A., Sauvignac, C., & Faruki, P. (2019). Network intrusion detection for IoT security based on learning techniques. *IEEE Communications Surveys & Tutorials*, 21(3), 2671-2701.

[25] Othman, S. M., Alsohybe, N. T., Ba-Alwi, F. M., & Zahary, A. T. (2018). Survey on intrusion detection system types. *International Journal of Cyber-Security and Digital Forensics*, 7(4), 444-463.

[26] Khraisat, A., & Alazab, A. (2021). A critical review of intrusion detection systems in the internet of things: techniques, deployment strategy, validation strategy, attacks, public datasets and challenges. *Cybersecurity*, 4(1), 1-27.

[27] Khraisat, A., Gondal, I., & Vamplew, P. (2018, June). An anomaly intrusion detection system using C5 decision tree classifier. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 7(2), 149-155. Springer, Cham.

[28] Jyothsna, V., & Prasad, K. M. (2019). Anomaly-based intrusion detection system. In *Computer and Network Security*, 2(1), 35-51. IntechOpen.

[29] Al-Garadi, M. A., Mohamed, A., Al-Ali, A. K., Du, X., Ali, I., & Guizani, M. (2020). A survey of machine and deep learning methods for internet of things (IoT) security. *IEEE Communications Surveys & Tutorials*, 22(3), 1646-1685.

[30] Li, H., Ota, K., & Dong, M. (2018). Learning IoT in edge: Deep learning for the Internet of Things with edge computing. *IEEE network*, 32(1), 96-101.

[31] Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009, July). A detailed analysis of the KDD CUP 99 data set. In *2009 IEEE symposium on computational intelligence for security and defense applications*, 1-6. IEEE.

[32] McHugh, J. (2000). Testing intrusion detection systems: a critique of the 1998 and 1999 darpa intrusion detection system evaluations as performed by lincoln laboratory. *ACM Transactions on Information and System Security (TISSEC)*, 3(4), 262-294.
[33] Koroniotis, N., Moustafa, N., Sitnikova, E., & Turnbull, B. (2019). Towards the development of realistic botnet dataset in the internet of things for network forensic analytics: Bot-iot dataset. *Future Generation Computer Systems, 100*(2), 779-796. [Google Scholar]

[34] Shafiq, M., Tian, Z., Bashir, A. K., Du, X., & Guizani, M. (2020). CorrAUC: a malicious bot-IoT traffic detection method in IoT network using machine-learning techniques. *IEEE Internet of Things Journal, 8*(5), 3242-3254. [Google Scholar]

[35] Mirsky, Y., Doitshman, T., Elovici, Y., & Shabtai, A. (2018). Kitsune: an ensemble of autoencoders for online network intrusion detection. *arXiv preprint arXiv:1802.09089*. [Google Scholar]

[36] Liu, Y., Ma, M., Liu, X., Xiong, N. N., Liu, A., & Zhu, Y. (2018). Design and analysis of probing route to defense sink-hole attacks for Internet of Things security. *IEEE Transactions on Network Science and Engineering, 7*(1), 356-372. [Google Scholar]

[37] Deka, R. K., Bhattacharyya, D. K., & Kalita, J. K. (2019). Active learning to detect DDoS attack using ranked features. *Computer Communications, 145*(4), 203-222. [Google Scholar]

[38] Rathore, S., & Park, J. H. (2018). Semi-supervised learning based distributed attack detection framework for IoT. *Applied Soft Computing, 72*(1), 79-89. [Google Scholar]

[39] Liu, Y., Ma, M., Liu, X., Xiong, N. N., Liu, A., & Zhu, Y. (2018). Design and analysis of probing route to defense sink-hole attacks for Internet of Things security. *IEEE Transactions on Network Science and Engineering, 7*(1), 356-372. [Google Scholar]

[40] Xu, B., Wang, W., Hao, Q., Zhang, Z., Du, P., Xia, T., ... & Wang, X. (2018). A security design for the detection of buffer overflow attacks in iot device. *IEEE Access, 6*(1), 72862-72869. [Google Scholar]

[41] Venkatraman, S., & Surendiran, B. (2020). Adaptive hybrid intrusion detection system for crowd sourced multimedia internet of things systems. *Multimedia Tools and Applications, 79*(5), 3993-4010. [Google Scholar]

[42] Patil, R., & Modi, C. (2019). Designing a Virtual Environment Monitoring System to Prevent Intrusions in Future Internet of Things. In *Recent Findings in Intelligent Computing Techniques*, 345-351. Springer, Singapore. [Google Scholar]

[43] Ashfaq, R. A. R., Wang, X. Z., Huang, J. Z., Abbas, H., & He, Y. L. (2017). Fuzziness based semi-supervised learning approach for intrusion detection system. *Information Sciences, 378*(1), 484-497. [Google Scholar]

[44] Diro, A. A., & Chilamkurti, N. (2018). Distributed attack detection scheme using deep learning approach for Internet of Things. *Future Generation Computer Systems, 82*(2), 761-768. [Google Scholar]

[45] Alazab, A., Hobbs, M., Abawajy, J., Khraisat, A., & Alazab, M. (2014). Using response action with Intelligent Intrusion detection and prevention System against web application malware. *Information Management and Computer Security, 22*(5), 431-449. [Google Scholar]

[46] Salameh, H. A. B., Almajali, S., Ayyash, M., & Elgala, H. (2018). Spectrum assignment in cognitive radio networks for internet-of-things delay-sensitive applications under jamming attacks. *IEEE Internet of Things Journal, 5*(3), 1904-1913. [Google Scholar]

[47] Tariq, N., Asim, M., Maamar, Z., Farooqi, M. Z., Faci, N., & Baker, T. (2019). A Mobile Code-driven Trust Mechanism for detecting internal attacks in sensor node-powered IoT. *Journal of Parallel and Distributed Computing, 134*(1), 198-206.
[48] Hasan, M., Islam, M. M., Zarif, M. I. I., & Hashem, M. M. A. (2019). Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. Internet of Things, 7(1), 100-159.

[49] Lyu, C., Zhang, X., Liu, Z., & Chi, C. H. (2019). Selective authentication based geographic opportunistic routing in wireless sensor networks for Internet of Things against DoS attacks. IEEE Access, 7(2), 31068-31082.

[50] Pham, T. N. D., Yeo, C. K., Yanai, N., & Fujiwara, T. (2017). Detecting flooding attack and accommodating burst traffic in delay-tolerant networks. IEEE Transactions on Vehicular Technology, 67(1), 795-808.

[51] Brun, O., Yin, Y., Gelenbe, E., Kadioglu, Y. M., Augusto-Gonzalez, J., & Ramos, M. (2018, February). Deep learning with dense random neural networks for detecting attacks against IoT-connected home environments. In International ISCIS Security Workshop, 79-89. Springer, Cham.

[52] Aithal, P. S., Shailashree, V., & Kumar, P. M. (2015). A new ABCD technique to analyze business models & concepts. International Journal of Management, IT and Engineering, 5(4), 409-423.

[53] Aithal, P. S. (2016). Study on ABCD analysis technique for business models, business strategies, operating concepts & business systems. International Journal in Management and Social Science, 4(1), 95-115.

[54] Shenoy, V., & Aithal, P. S. (2016). ABCD Analysis of On-line Campus Placement Model. IRA-International Journal of Management & Social Sciences, 5(2), 227-244.

[55] Shenoy, V., & Aithal, P. S. (2017). Quantitative ABCD Analysis of IEDRA Model of Placement Determination. International Journal of Case Studies in Business, IT and Education (IJCSBE), 1(2), 103-113.

[56] Surana, J., Saraf, I., Puri, N., Navin, B., & Surana, J. (2017). A Survey on Intrusion Detection System. International Journal of Engineering Development and Research (IJEDR), 5(2), 960-965.

[57] Intrusion Prevention System Benefits / UM Information and Technology Services. Retrieved November 21, 2021, from https://its.umich.edu/enterprise/wifi-networks/network-security/ips/benefits

[58] The Pros & Cons of Intrusion Detection Systems: Rapid7 blog. Rapid7. (2020, October 27). Retrieved November 21, 2021, from https://www.rapid7.com/blog/post/2017/01/11/the-pros-cons-of-intrusion-detection-systems/.

[59] Basics of intrusion detection system, Classifications and advantages. ElProCus. (2020, May 6). Retrieved November 21, 2021, from https://www.elprocus.com/basic-intrusion-detection-system/