Section: Research Paper

Analysis Of Methodologies, Deployment Strategy, Validation Framework, Vulnerabilities, Available Datasets And Issues For Intrusion Detection Systems In Internet Of Things (IoTs)

Authors - Shubham Kumar 1, Dr. Nishant Kumar Pathak 2*

1 - Research Scholar, Department of Computer Science & Engineering, Shobhit Institute of Engineering & Technology ('A' NAAC Accredited Deemed-to-be University), NH-58 Modipuram, Meerut, 250110, Uttar Pradesh, India.

2 - Assisstant Professor, Department of Computer Science & Engineering, Shobhit Institute of Engineering & Technology ('A' NAAC Accredited Deemed-to-be University), NH-58 Modipuram, Meerut, 250110, Uttar Pradesh, India.

Corresponding author(s). E-mail(s): 1- shubhamkumar3593@gmail.com;
*Contributing authors: 2 - nishant.pathak@shobhituniversity.ac.in ;

ABSTRACT

The Internet of Things (IoT) is evolving swiftly in order to impact major industrial systems and people's daily lives more significantly. Most of the IDS technologies, that may be generally categorised based on detection approach, validation strategy, and deployment strategy have been presented in the literature to address assaults on the IoT ecosystem. This survey article provides a thorough analysis of current IoT IDS as well as an overview of the methods, deployment strategies, validation strategies, and datasets that are frequently used to create IDS. Moreover, we examine how current IoT IDS identify and protect exchanges over the IoT. In order to make IoT more secure, it also gives a taxonomy of IoT threats and highlights upcoming research challenges to defend against them. By bringing together, comparing, and combining disparate research efforts, these goals aid IoT security experts. Hence, in order to shed light on IoT IDS methodologies, their benefits and drawbacks, IoT attacks that take use of IoT communication networks, and related sophisticated IDS and detection capabilities to identify IoT assaults, we offer an original IoT IDS taxonomy**.**

Keywords: Assault, Attacks on IoT, Intrusion detection system, IoT, Machine learning, Artificial Neural Networks, Intrusion Detection Datasets, Challenges of IoT, IoT security.

Introduction

A network of networked devices called the Internet of Things (IoT) enables smooth data flow between physical items. These devices might be tracked and controlled remotely and might include medical and healthcare equipment, autonomous cars, industrial robots, smart Televisions, wearable technology, and smart city infrastructures. IoT devices will have access to the most private information and are predicted to outnumber mobile devices in terms of

prevalence. As a result of this attack probability will rise as attack surface area increases. IoT intrusion detection systems must be created in order to protect communications made possible by such IoT technologies since security will be a crucial supporting component of most IoT applications [45].

IoT IDS has improved in recent years thanks to advancements in artificial intelligence (AI), including machine learning and deep learning approaches (Intrusion Detection System). An updated, detailed taxonomy and a critical analysis of this most recent study are currently required. Although being a crucial component for the efficacy of "on-line" IDSs, the time spent developing and testing IoT IDS is not taken into account in the assessment of about IDSs procedures [59][1][3].

This paper offers analysis of methodologies, deployment strategy, validation framework, vulnerabilities, available datasets, and issues for intrusion detection systems in internet of things (IoTs).

In order for a researcher to rapidly become familiar with the essential components of IoT IDS, it offers an organised and comprehensive overview of the available IoT IDSs. A critical analysis of the machine learning and deep learning techniques used to create IoT IDS is also provided in this research. A number of approaches employed in each method are discussed, along with the detection techniques, validation tactics, and deployment strategies. Following a discussion of the difficulty of various detection systems, intrusion deployment strategies, and their evaluation approaches, a list of Depending on the IoT IDS, tips for recommended practises are provided. The difficulties facing the existing IoT IDSs are also explored. The research community in the field of IoT intrusion detection systems is primarily concerned with IoT methodologies, IoT deployment strategy, and IDS dataset challenges. In this work, these issues are discussed in comparison related to earlier survey releases [59][27][34][117][48].The datasets, difficulties, and strategies of IoT IDSs have not been fully explored in other research like [114][115][4]. This article offers an organised, modern, in-depth analysis of IDS in terms of methodologies and IoT attacks.

Review articles that have already been written focus on intrusion detection techniques, dataset concerns, certain categories of computer assaults, and IDS evasion [20][34][40][29][5].There have been no in-depth studies of IoT IDS, dataset problems, implementation tactics, IoT intrusion approaches, or different attack kinds. Due to the growth of IoT IDS, several new solutions have been provided in the interim, necessitating an upgrade. This study complements taxonomies offered in by enhancing the taxonomy of the IoT IDS industry [59][27][34][67][5].

Thinking about the conversation of earlier discussions, the following are the main topics of this article:

- Classifying distinct IoT IDS types based on deployment, validation, and intrusion tactics.
- Outlining current efforts to enhance IoT security IDS.
- An analysis of IoT assaults.

INTRUSION DETECTION IN THE INTERNET OF THINGS

In this part a survey of the IoT IDS research that has already been done is presented. IDS placement technique, detection method, and validation approach were taken into account while classifying each study. The categorization of IDS for IoT networks is shown in Figure 1.

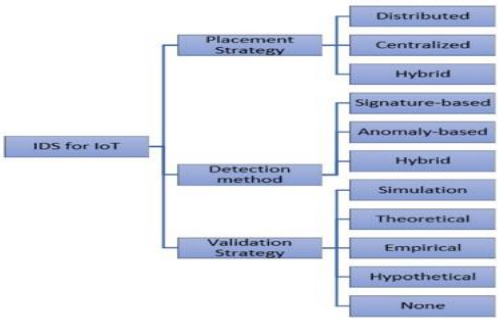

Figure 1: Classification Of IDS In IoT

The IDS methods, deployment approach, validation strategy, IoT threats, and datasets covered by this work and other research studies are displayed in Figure 1. The diversity of IoT IDS surveys suggests that a review of an IDS for IoT research is necessary. Due to the varied nature of the IoT ecosystem, it is specifically noted that none of these studies cover all IoT detection techniques. This analysis therefore examines IDS for IoT across a variety of technologies.

IoT Intrusion Detection Systems Methods

An illegal action or behaviour that has an impact on the IoT ecosystem is referred to as a "IoT intrusion." In other words, an assault is deemed invasive if it jeopardises the information's integrity, confidentiality, or accessibility in any way. An incursion, for instance, occurs when an assault prevents genuine computer users from using such services. An IDS is a software or hardware device that monitors computer systems for malicious activity in order to keep the system secure [67][5]. IDS's primary goal is to detect malicious network traffic and unauthorised computer usage, which is impossible with a conventional firewall. As a result, the computer systems become very guarded against malevolent acts that jeopardise computer systems' confidentiality, integrity, or availability. Signature-based Intrusion Detection System (SIDS) and Anomaly-based Intrusion Detection System are the two primary subcategories of IDS systems (AIDS).

Signature-based intrusion detection systems (SIDS)

Signature intrusion detection systems (SIDS), also known as Knowledge-based Detection or Misuse Detection, provide shape corresponding methods to realize a recognized assault [57]. SIDS, corresponding a prior incursion is located using several techniques. In other words, an alarm signal is triggered when an intrusion signature matches a previously recorded intrusion signature in the signature database. The host's logs are inspected in Detail to look for groups of instructions or behaviours that have been recognised as malware in the past. SIDS is also known as Knowledge-Based Detection or Misuse Detection in the literature [77].

Figure 2 illustrates how SIDS techniques function conceptually.

Figure 2: Working Of SIDS

Anomaly-Based Intrusion Detection System (AIDS)

Since it has the ability to circumvent SIDS's restrictions, AIDS has drawn the attention of many academics. A typical computer system behaviour model for AIDS is created utilising machine learning, statistical analysis, or knowledge-based methods. An anomaly, also known as an incursion, is any significant divergence from the model's expectations. These kinds of tactic takes use of the fact that malevolent behaviour differs from customary user behaviour. An incursion is described as anomalous user behaviour that deviates from expected user behaviour. A model of the normal traffic profile is learned the system during the training phase normal conduct. A fresh data set is employed during testing to build the system's ability to generalise to previously unidentified incursions. AIDS subcategories can be determined by the training methodology, such as statistical- based, based on knowledge, and machine learning-based [31].

As AIDS does not require a signature database, its ability to detect zero-day attacks is its main advantage to detect unusual user behaviour [21]. When the observed behaviour deviates from expected behaviour, AIDS sends out a warning signal. AIDS also offers a variety of advantages. They can first identify harmful internal activity. An alarm is raised if an intruder begins making transactions in a stolen account that aren't obvious from regular user activity. Second, because the system is built from personalised profiles, it is difficult for a cybercriminal to identify typical user activity without raising an alert.

Techniques For Implementing AIDS

An overview of contemporary AIDS techniques that aim to increase detection precision and decrease false alarms is provided in this section.

The four major categories of AIDS approaches are deep learning, reinforcement learning, unsupervised learning, and supervised education [35][44][32][75][29]. All input is collected and examined during supervised learning.

In order Identifying typical user behaviour from input to output, you utilise an algorithm and an input variable and an output variable. The goal is to sufficiently approximate the mapping function to be able to forecast the output variables for a given record when a fresh set of input records is gathered. Unsupervised learning, on the other hand, aims to recognise the required actions from the system data already there, such example in circumstances where you just have input data and no corresponding output variables, such protocol specifications and network traffic.

By trial and error and the use of feedback from its own actions and experiences, an agent may learn utilising reinforcement learning techniques in an interactive environment. The goal of reinforcement learning is to develop an optimal action model that will maximise the agent's overall cumulative reward. Artificial neural networks, specifically convolutional neural networks (CNNs), are the foundation of deep learning models. Fig. 3 displays these four classes together with illustrations of their subclasses.

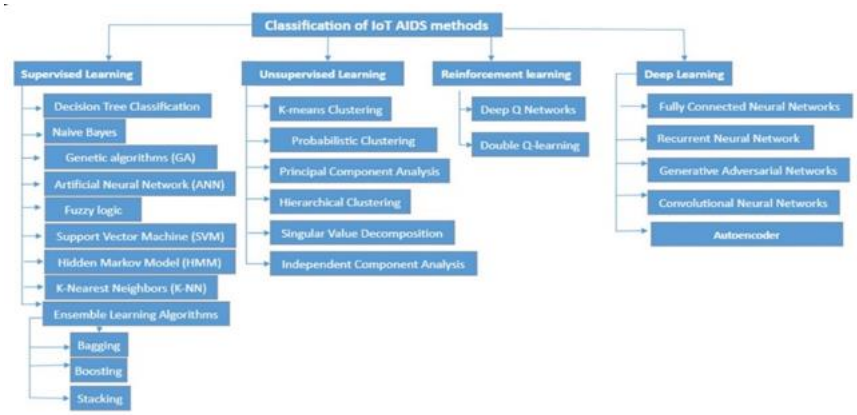

Figure 3: Classification Of AIDS

Supervised Learning In Intrusion Detection System

Several supervised learning methods for IDS are presented in this subsection. Each approach is explained in detail, and citations to significant research articles are provided.

Using labelled training data, supervised learning-based IDS algorithms find intrusions. Training and testing are the two processes that make up a supervised learning technique. Relevant features and classes are found during the training step, and the algorithm subsequently gains knowledge from these data samples. Each record in supervised learning IDS consists of a network or host data source and a corresponding output value (i.e., label), such as incursion or normal. The next use of feature selection is to remove features that are not essential. A classifier is then trained using a supervised learning approach using the training data for certain characteristics. to understand the natural link between the input information and the labelled output value. The literature has examined a wide range of supervised learning techniques, each with advantages and disadvantages. Unknown data are classified into incursion or normal using the learned model classes during testing. As a model, the resulting classifier then predicts the class to which the input data may belong given a collection of feature values. A general method for using classification algorithms is shown in Figure 5. The majority of currently planned IDSs get supervised training.

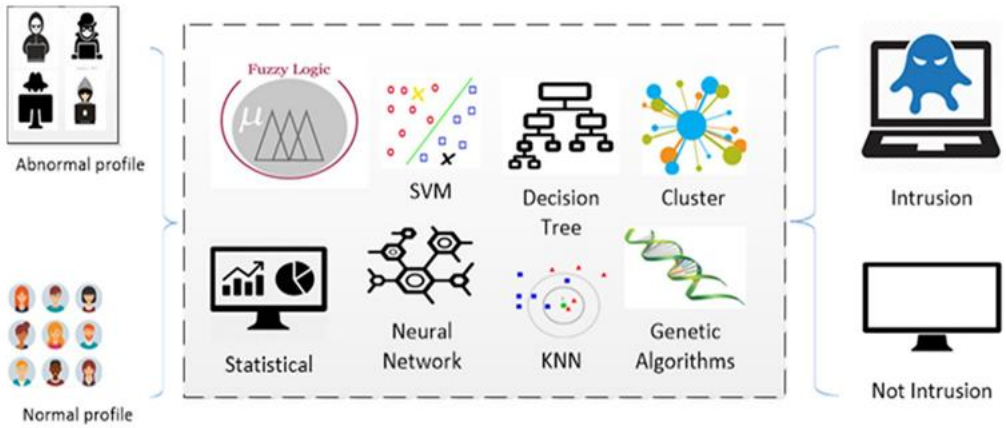

Figure 4: Working Of AIDS

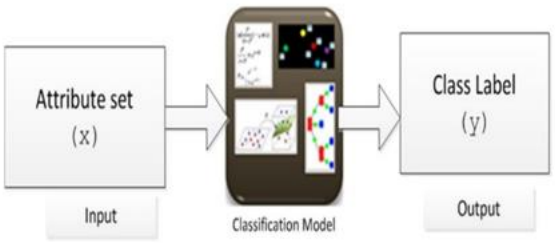

Figure 5: Classification Of Task

There are several classification techniques, including k-nearest neighbour, decision trees, rule-based systems, neural networks, and support vector machines. Each technique creates a classification model using a learning technique. Nevertheless, a successful classification method must be able to reliably identify the class of records it has never seen before in addition to handling the training data. The learning algorithm's main objective is to produce classification models with trustworthy generalizability.

Decision Tree: There are three essential components of a decision tree. A decision node, the initial element, serves as a test property by being utilised. The alternative two is a branch, with each branch stands in for a potential course of action depending on the test attribute's result. The third is a leaf made up of the class that the instance is a part of (Rutkowski et al., 2014). Decision tree algorithms come in a variety of forms, such as ID3 [85], C4.5 [85], and CART [29].

Naïve Bayes: The foundation of this strategy is the use of the Bayes principle with strong independence presumptions among the qualities. By using conditional probability equations, Nave Bayes responds to inquiries like "what is the likelihood that a specific type of assault is occuring, given the observed system activities?" The Naive Bayes algorithm depends on characteristics that have varying odds of occuring during assaults and in regular activity. The conditional independence assumption attribute of one of the most often utilised models in IDS is the naive Bayes classification model, due to its simplicity and computation efficiency [113].

Genetic Algorithms: Gene-based algorithms (GA) Based on the principles of evolution, A heuristic approach to optimisation is the use of genetic algorithms. A grouping of bits (genes) or chromosomes serves as a representation for each probable answer. Using the operators for selection and reproduction that are biased in favour of better answers, the quality of the solutions increases over time. There are typically two types of chromosomes encoding when using a genetic algorithm to solve the intrusion classification problem: one is based on clustering to produce a binary chromosome coding method, and the other is a coding chromosome that defines the cluster centre (clustering prototype matrix) as an integer, Murray et al. developed straightforward network traffic rules using GA [80].

Artificial Neural Networks (ANN): ANN, most distinguished ML algorithm in process of finding various assaults. Backpropagation (BP) algorithm is the most often used learning method for supervised learning. In relation to its adjustable weights, the BP algorithm evaluates the gradient of the error of the network. Nonetheless, there is still room for improvement in terms of detection accuracy for ANN-based IDS and detection precision, particularly for less frequent attacks. the training dataset for less common attacks is smaller, less than that for more frequent attacks, it is challenging for the ANN to accurately understand the characteristics of these attacks. As a result, fewer frequent assaults have poorer detection accuracy. As a result, fewer frequent assaults have poorer detection accuracy. If low-frequency assaults are not discovered in

the field of information security, severe harm may result. For instance, if User to Root (U2R) assaults manage to avoid detection, a cybercriminal may be able to acquire the root user's authorisation capabilities and so engage in destructive activity on the computer systems of the victim. Moreover, assaults that occur less frequently are often outliers [109].

Fuzzy logic: Instead of using the standard true or false Boolean logic that modern Computers are built on, this method is based on degrees of uncertainty. As a result, it offers a simple method for drawing a conclusion from input data that is murky, confusing, noisy, erroneous, or lacking. Fuzzy logic allows an instance to simultaneously belong, potentially partially, to numerous classes in a fuzzy domain. Fuzzy logic is an excellent classifier for IDS issues since the security itself contains ambiguity and it is difficult to distinguish between normal and abnormal conditions. The acquired data for the intrusion detection problem also includes a number of derived statistical metrics and different numerical properties.

Support Vector Machines (SVM): A splitting hyperplane defines SVM as a discriminative classifier. In order to linearly classify incursion, To convert the training data into a higher-dimensional space, SVMs use a kernel function. SVMs are renowned for their capacity to generalise and are most useful when there are a high number of characteristics and few data points. Applying a kernel, such as a Hyperbolic tangent, Gaussian Radial Basis Function (RBF), linear, polynomial, allows for the separation of various types of hyperplanes. Several characteristics in IDS datasets are redundant or have less of an impact on classifying data items into the appropriate groups. Hence, when training an SVM, feature selection should be taken into account. Multiple class classification is another use for SVM.

Hidden Markov's Model (HMM): A statistical Markov model called an HMM assumes the system under investigation is a hidden data Markov process. Prior research has shown that certain forms of malware may be recognised via HMM analysis [24]. This technique involves training a Hidden Markov Model against well-known malware.

K-Nearest Neighbours (KNN): A famous multivariate ML model is the k-Nearest Neighbor (k-NN) approach (Lin et al., 2015). A K-Nearest Neighbors classifier with k = 6 is shown in Figure 6. A specific instance of unlabelled data that has to be categorised is represented by the point X.

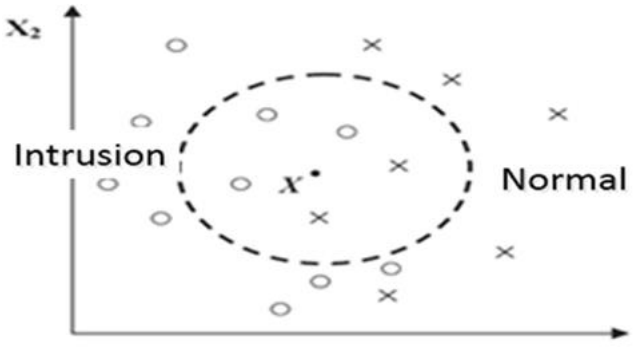

Fig. Classification k-NN for k = 6

Ensemble Methods: To achieve greater prediction performance than any one of the individual machine learning algorithms, several machine learning algorithms might be employed [104]. In order to enhance the detection rate, train several classifiers simultaneously to recognise various threats. The ensemble's ability is typically superior to that of a single classifier because it may

strengthen weak classifiers and give results that are superior to those of a single classifier [19]. Numerous distinct ensemble techniques, including boosting, bagging, and stacking, have been suggested.

Unsupervised Learning In Intrusion Detection System

Unsupervised ML methods take input datasets deprived of class tags to get back related information. The input facts are frequently observed as a consortium of random variables. The data collection is subsequently turned into a joint density model. In supervised learning, the computer is trained to provide the desired outputs for an unknown data point using the output labels that have been provided. In contrast, no labels are provided in unsupervised learning; instead, the learning process automatically divides the data into several groups. Unsupervised learning refers to the process of training a model to detect intrusions using unlabelled data in the context of creating an IDS. IoT network activity is grouped into categories without having to first define them, depending on how similar the traffic is.

When data are grouped, as in Fig. 7, all instances that appear in tiny clusters are classified as intrusions since typical occurrences should result in larger clusters than the anomalies. Also, because malicious incursions and regular instances differ from one another, they do not belong to the same cluster.

K-means One of the most popular clustering analysis approaches, the K-means methodology divides Each data object is picked in the cluster with the closest mean after grouping n data items into k clusters. K indicates that the clustering process is iterative and helps to find the greatest value for each iteration. As it uses a distance-based clustering method, it is not necessary to calculate the distances between every possible set of records. As a similarity metric, it uses a Euclidean metric. The user decides in advance how many clusters there will be. Before choosing the best one, multiple options will often be evaluated.

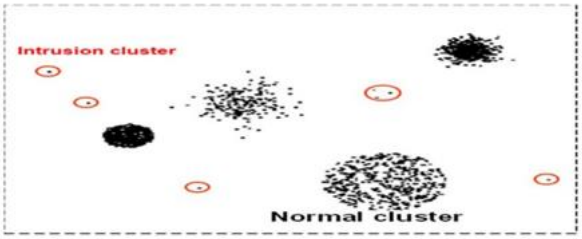

Fig. 7: Using Clustering for Intrusion Detection

A popular technique for extracting a group of low dimensional characteristics from the greatest set of data is principal component analysis.

Using a hierarchy to cluster This method of clustering seeks in order to create a cluster hierarchy. Two categories of hierarchical clustering techniques are frequently used:

(i) **Agglomerative,** bottom-up clustering approaches, whereby cluster pairings are connected as one moves up the hierarchy; clusters contain sub-clusters, which in turn have sub-clusters.

(ii) **Divisive,** hierarchical clustering techniques, where the greatest diameter cluster in the feature space is iteratively chosen and divided into binary sub-clusters with a lower range.

Analysis of independent components It is used to reveal underlying causes that lay behind collections of seemingly random traits.

Reinforcement Learning

For the construction of IDS, deep reinforcement learning makes use of these ideas. An agent interacts with the environment during reinforcement learning. The agent is making some sort of attempt to carry out a task inside the setting. The agent's objective is to discover how to communicate with its surroundings in a way that will help it accomplish its objectives.

Deep Q-network: Deep neural networks and reinforcement learning are coupled at scale. The algorithm was created by using deep neural networks to improve the Q-Learning conventional RL technique.

Double-Q learning: In order to overcome overestimation issues with standard Q-learning, this off-policy reinforcement learning method uses double estimation.

Deep Learning

A computer uses an experience-based hierarchy of data to create several layers as an output in deep learning, a type of machine learning. Both supervised and unsupervised deep learning are possible. Although unsupervised deep learning analyses data patterns, supervised deep learning allows for the classification of data. Deep learning is closely connected to artificial intelligence, where robots will take the role of human intelligence and learn via experience. Deep learning uses algorithms created by human intellect to analyse vast volumes of data on the artificial neural network platform.

Each neural node of every hidden layer in neural networks calculates the weighted values received from the layer before it and transmits the results to the layer after it. The final results that the neural networks produced from the raw data may be viewed as being represented by the result value of the last layer.

Fully Connected Neural Networks (FCNN): The standard network design used in the majority of fundamental neural network applications is fully connected feedforward neural networks. Completely connected means that every neuron in the succeeding layer is connected every each and every single neuron.

Recurrent Neural Networks (RNN): The recurrent neural network is capable of processing a sequence of data with varying input lengths with efficiency. To put it another way, RNNs use the creation of their previous state as an input for making predictions about the present. This process might be repeated for a variety of steps to allow the network to propagate information over time using its hidden state. This is analogous to providing a short-term memory to a neural network. RNNs are incredibly useful for dealing with data sequences that occur throughout time because of this property. It is the best option for creating IDS with high accuracy, for both binary and multiclass classification, its performance is superior to that of conventional machine learning approaches.

Generative Adversarial Networks (GAN): The Generator Network and a Discriminator Network, two deep learning neural networks, are combined to form the Genera- tive Adversarial Network. The Discriminator Network seeks to determine if the data it is viewing is genuine or synthetic by using the Generator Network to produce synthetic data. In that they are both vying to outperform one another, these two networks are rivals.

Convolutional Neural Networks (CNN): Like in a typical multilayer, more than one neural network connected in multi layers forms CNN [105]. A neural network using convolutions has many hidden layers in addition to input and output layers. A series of convolutional layers

commonly seen in a CNN's hidden layers convolve with a multiplication. Using a series of hidden layers, a CNN abstracts high-level information from a 2-D input. Spatial characteristics are advantageous to CNNs, which improve on the standard neural network design [106]. In the IDS region, spatial characteristics are frequently used as different kinds of traffic features. When network traffic is transformed into traffic pictures using spatial characteristics; as a result, the goal of identifying intrusion traffic is also finally met by classifying the traffic images using an image classification approach. Although this method is relatively new, multiple recent study findings demonstrate its enormous potential.

Autoencoder: A trained autoencoder restructures its inputs. Online IoT IDS have been created using autoencoders [76]. The capacity to restructure unseen instances from the same data distribution as X often comes with an auto-encoder that has been trained on the example X. It is anticipated that the restructure will have a high error rate if an instance does not fit the model discovered from X.

IOT IDS DEPLOYMENT STRATEGIES

The deployment used to identify IoT threats may also be used to categorise IDS. IDS can be categorised as distributed, centralised, or hybrid in IDS deployment methodologies.

Distributed IDS

In a dispersed setting, IoT devices can be in charge of inspecting other IoT devices. Advanced intrusion detection systems, packet analysis, and incident response are supported by a central server that is accessible by several distributed IDS spread throughout a large IoT ecosystem.

Several IDS use scattered architectural designs. A portion of the network's other nodes are checked as part of this. The incident analyst has various benefits from distributed IDS versus centralised IDS. The ability to recognise different attack types throughout the whole IoT ecosystem is the main advantage. This may result in quicker IoT attack prevention and detection. Allowing for early identification of an IoT Botnet that is making its way via corporate IoT devices is the extra supported feature.

Centralized IDS

The IDS is installed in central devices, such as the boundary switch or a designated device, at the centralised IDS site. The network border switch receives all the data that the Internet of Things devices gather and deliver to it [27]. As a result, the packets moved between the IoT devices and the network may be checked by the IDS installed in a boundary switch. Even then, monitoring the network packets that travel through the border switch is insufficient to spot abnormalities that interfere with IoT devices. IDS is used to centrally monitor network traffic. This network traffic is retrieved from many network data sources, including packet capture, NetFlow, etc. The linked computers Network-based IDS can keep an eye on a network. Moreover, NIDS is capable of keeping an eye on any harmful activity that may have been started earlier as a result of an external attack, before such dangers spread to other computer systems. Nevertheless, because to the volume of data travelling through current high-speed communication networks, NIDS has several limitations, such as its limited capacity to check the entirety of the data in a high bandwidth network [28].

Hierarchical IDS

Clusters are created within the network in hierarchical IDS. Usually, the sensor nodes that are next to one another are members of the same group. The so-called cluster head, who serves as the leader of each cluster, who oversees network-wide analysis and filters the member nodes.

IDS VALIDATION STRATEGIES

The process of IDS validation determines if the IoT IDS model is a sufficient representation of the system for identifying IoT assaults. Researchers have employed a variety of methodologies, including theoretical, empirical, and hypothetical strategies, to validate the efficacy of IDSs.
IDS are frequently assessed using the following common performance metrics:
If one sample is an anomaly and the predicted label also stands anomaly, then it are called as **true positive (TP).**

If one sample is an anomaly, but the predicted label stands normal, then it is called as **false negative (FN).**

If one sample are a normal and the predicted label also stands normal, then it are **true negative (TN).**

If one sample is normal, but the predicted label stands anomaly, then it are termed as **false positive (FP).**

TP stands the number of true positive samples, FN stands the number of false negative samples, The letters FP and TN stand for the number of false positive and true negative samples, respectively.

From equation (1) and (5), the F1 score, True positive rate(TPR), False Negative rate (FNR), False Positive Rate (FPR) and False Alarm rate (FAR) are calculated.

$$F1\ score = (2TP)/(2TP+FP+FN) \quad (1)$$

$$TPR = TP/(TP+FN) \quad (2)$$

$$FNR = FN/(FN+TP) * 100 \quad (3)$$

$$FPR = FP/(FP+TN) * 100 \quad (4)$$

$$FAR = (FPR+FNR)/2 \quad (5)$$

Where TPR = True Positive Rate, FNR= False Negative Rate, FPR= False Positive Rate, FAR= False Alarm Rate

AUC stands for Area under the ROC curve whose values lies between 0 to 1 and the ROC curve plots between TPR and FPR.

State-of-the-art intrusion detection in IoT

A method for examining border router packets for communication between physical and network devices was put forth by Cho et al. Their approach is based on botnet assaults that measure packet size [37]. Nevertheless, nothing is written about how a typical behavior profile was created. Furthermore, it is not apparent how the recommended IDS methodologies would work on IoT nodes with limited resources.

Framework for IoT distributed threat detection using semi-supervised fuzzy learning was proposed by Rathore et al [87]. Due to the evaluation's use of the NSL-KDD dataset, it was subject to the same dataset's above-mentioned constraints.

To identify DDoS and DoS assaults against genuine IoT network traffic, Hodo et al. deploy an Artificial Neural Network (ANN). A simulated IoT network was used to evaluate the suggested ANN model. IoT threat analysis using ANN to identify DDoS/DoS threats was proposed by Hodo et al. Using internet packet traces, a multi-level perceptron, a sort of supervised ANN, is trained before being evaluated for its capacity to prevent (DDoS/DoS) assaults [49]. Hodo et al. neglected to take efficacy into account when deploying the suggested IDS on low-capacity devices in the IoT environment. Their testing revealed that the technology has a 99.4% accuracy rate for DDoS/DoS. Nevertheless, no information on the dataset is given.

A distributed deep learning-based IoT network threat detection system was created by Diro et al. Their research shown that distributed attack detection has a 96% detection rate, outperforming centralised attack detection in the identification of IoT assaults. The NLS-KDD dataset was used to assess their strategy. Although this dataset is a different form of the KDD data set, it still has a number of problems that McHugh has examined [73]. We believe this dataset shouldn't be used as a meaningful benchmark in the IoT because it was obtained using a traditional network [42]. This prompts the development of IDSs that account for the unique needs of IoT protocols like (Low-power Wireless Personal Area Networks) 6Low- PAN.

In order to detect anomalous activity in particular botnet assaults against Hypertext Transfer Protocol (HTTP), Message Queue Telemetry Transport (MQTT), and Domain Name System (DNS), according to Moustafa et al. suggested an ensemble of IDSs [79]. They employed three machine learning approaches to assess their methodology: Artificial Neural Networks (ANN), Decision Tree (DT), and Naive Bayes (NB). Their ensemble methods are based on the AdaBoost learning algorithm [79]. The suggested IDS has a noticeable overhead that lowers its performance.

A One-Class Support Vector Machine and C5 classifiers are used in a hybrid intrusion detection system (HIDS) has been suggested by Khraisat et al. [60]. Well-known intrusions are detected using the C5 classifier. To find a fresh assault, a one-class support vector machine classifier is utilised.

Attacks on IoT ecosystem

The purpose of connecting to other networks and exchanging the data has been effectively

achieved since IoT technology uses several devices, including sensors, CPUs, and many other technologies. The shared data may not be safe because of the numerous related vices, which poses security concerns. IoT security refers to the safeguarding of data sent between various networks by means of IoT devices and IoT technologies. These gadgets are linked to other devices over the internet, which creates weaknesses and makes it possible for data to be stolen. Data without security will cause several problems and cause enormous loss for many sectors and even for individuals, ultimately leading to the destruction of their systems' data [60].

IoT caught the interest of individuals and organizations from a wide range of industries by offering enormous benefits to them. Along with its rapid expansion, significant security concerns have emerged, making it difficult for individuals to exploit many of the IoT's anticipated uses. As a result, this portion of the paper addresses the idea of IoT security, the challenges associated with it, their effects, and IoT attack types. On a reliable network, IoT devices may be accessed from anywhere. Thus, there is a high likelihood of numerous hostile assaults on the IoT network. To prevent hostile assaults on the IoT, security, privacy, and confidentiality problems must be properly handled.

Figure 8 depicts the levels of the IoT system architecture where attacks may take place. The perception layer, network layer, and application layer are the three main layers that an IoT system can have [67].

The conventional IoT design has a minimum layer known as the perception layer. Devices, sensors, and controls make up this layer. The primary function of this layer is to collect useful data from IoT sensor devices.

IoT encompasses a wide range of varied networks at the network layer, including WLAN, wireless mesh networks, and WSNs. These networks facilitate information flow between IoT sensors. Several sensors can communicate more easily over the network with the help of a gateway. Thus, a gateway may be useful to manage many complicated elements of network communication. While the application layer is the uppermost layer and analyses the data for viewing, the network layer facilitates the successful transfer of data.

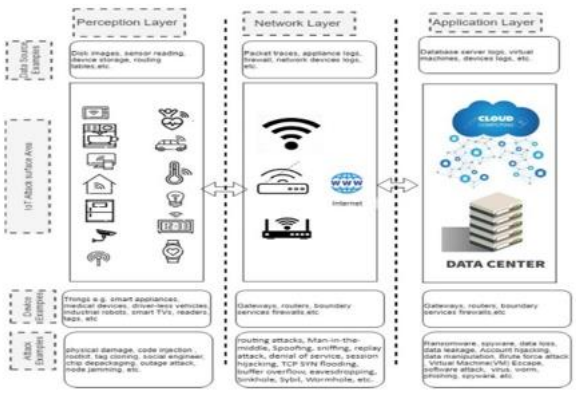

Fig 8: IoT Architecture and Layer

The following is a summary of the main reasons why IoT is a target for malware:

- All of the equipment and devices in an IoT must be always active, and attackers may easily analyse equipment whose power mode is active at any one time.
- Another reason for malware targets in the Internet of Things is the absence of adequate encryption mechanisms in the connected devices and weak passwords.
- In comparison to using a single device, the IoT requires significantly less expertise and is simpler to use.
- Another reason why IoT is a target for malware is because the equipment and gadgets have been exposed to the internet for twenty-four hours. The gadgets will accept incoming traffic signals as a result of the limitless internet connection, making them open to assaults.

There is a description of the many forms of assaults, their effects on the IoT network, and their significance.

PHYSICAL/PERCEPTION LAYER

Attacks are based on tools' hidden features and machinery. By messing with the hardware, these attackers are able to take over the target device. When an assault is close to a network or an IoT device, a physical IoT attack is initiated. At the physical/perception layer, some of the major dangers include:

Node tampering

Hacking the system to discover the secret keys needed to decode the data is known as node tampering.

Radio frequency (RF) Interface

The Internet of Things (IoT) uses Radio Frequency (RF) for wireless communication. The Internet of Things (IoT) devices are easily exposed to certain assaults due to the wireless technologies used for data transmission between devices.

Node jamming

A sort of denial-of-service assault known as jamming involves the adversary sending a long-range signal to disrupt the communication. In jamming attacks, a rogue sensor node broadcasts a jamming signal using frequencies that are identical to those of the sensor nodes. By generating noise in the IoT network and rendering the services inaccessible, this jamming attack prevents the sensor nodes from transmitting or accepting data.

Node attack

The sensor nodes might be fully controlled by the cybercriminal. Tags are vulnerable to physical assaults since IoT devices are positioned in various areas. To take advantage of an RFID system, a cybercriminal may easily steal these tags and duplicate them.

Physical damage

To modify the data or steal sensitive information, the attacker actively takes part in the attack.

Social engineering attacks

Social engineering methods are used by the attacker to gain unauthorised access to a system and covertly install harmful software. In order to provide its users with a customised experience, IoT devices, in particular wearables, collect enormous amounts of personally identifiable information (PII). Such Internet of Things (IoT) devices also make use of customers' personal information to provide user-friendly amenities, such ordering things online with voice control. Cybercriminals, however, may target PII in order to get unauthorised access to sensitive data such user passwords, purchase histories, and personal information.

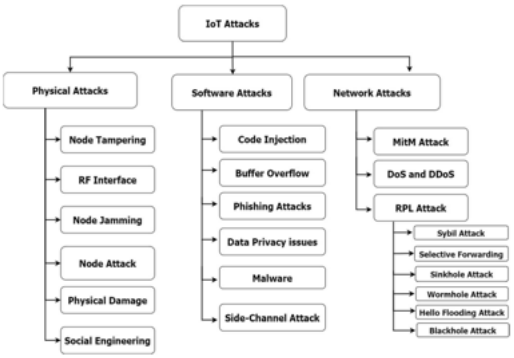

Fig 9: Taxonomy Of Security Attacks

SOFTWARE/ APPLICATION LAYER

The apps used in IoT technology are online applications that require the installation of software in order to function. Attacks against software are carried out by utilising phishing scams, trojan horses, ransomware, worms, viruses, or other malicious software, including spyware and adware.

Code injection

Change the execution by injecting code into a sensor node that is weak. For instance, the inaudible attack programme Dolphin Attack uses the ultrasonic channel to insert inaudible speech commands into voice-activated devices [118]. Another illustration is a When the victim is conversing with the VPA service, the attacker manipulates it by adding malicious code, which is known as a voice squatting attack in order to obtain her personal data.

Buffer overflow

Due to insufficient boundaries verifying, When data are written to a sensor node buffer, a buffer overflow occurs, corrupting data values in memory near to the destination buffer as well.

Data privacy issue

Several home objects might have RFID tags added by the attackers. RFID tag tracking IoT devices may be used to track users' movements and create user profiles in order to violate their privacy.

Malware

Any malicious software created with the intention of harming or damaging IoT infrastructure is referred to as malware. Malware comes in a wide variety ransomware, spyware, adware, Trojan horses, infections, and other types of malware.

Phishing attack

IoT edge node is used by the attacker as a trap. The objective is to gather data, including usernames, passwords, etc.

Side-Channel attack

By using data that cryptography has revealed, a side-channel attack defeats encryption.

NETWORK LAYER

Data transmission happens network layer, where security issues might develop and pave the way for attackers. Eavesdropping, man-in-the-middle, denial-of-service, storage, exploit, spoofing, and other tactics may be used in these attacks. IoT assaults encompass a variety of information security assaults that may be directed at particular systems, networks, or data sets. Physical security attacks may be carried out on the IoT networks' devices as a target. Most Internet of Things (IoT) attacks are network-based or targeted at specific information attributes. Often, they are malicious assaults intended to harm the IoT application's availability or the data's confidentiality. Significant risks at the network layer include some of the following:

Man-in-the-middle (MITM) Attack

The confidentiality, integrity, and availability of Internet of Things communications may be threatened by man-in-the-middle attacks during wireless sensor connections [82]. Wireless attacks can include packet sniffing, eavesdropping, MAC spoofing, rogue wireless devices, and encryption breaking. When an attacker modifies communications between two parties that believe they are secretly speaking with one other without the authenticating user's consent, this is known as an MITM attack. It is quite similar to an eavesdropping attack in which the attacker may interject into two participants' conversations. Email spoofing, WiFi eavesdropping, Session Hijacking, DNS spoofing, and IP spoofing are some examples of MITM attacks. An attacker may, for instance, install network spyware. In order to undertake a spy operation and intercept the packet while it is in transit, (a sniffer) may be installed on a computer or server.

Denial of service (DoS) attack

The steady availability of the provided services offered by a system is prohibited by a denial of service attack. The system's legitimate users are denied access to its resources. A distributed denial of service attack is one that is launched by several malicious nodes (DDoS). Instead of losing information as a result of service holders transferring the services from the original provider owing to security concerns, a DOS attack would cost the victim time and money. DoS attacks can impact CPU use, bandwidth, and network resources [96]. There is a heightened risk of an assault since IoT equipment and gadgets are always in power-on mode and linked to the IoT system. Malware payloads may be transmitted at any moment through an IoT network in a house or workplace. For instance, the botnet "Mirai" launched a Distributed Denial of Service (DDoS) assault that rendered a large portion of the network unavailable [60].

Distributed denial of service (DDoS)

In a DDoS Cyberattack, an attacker temporarily seizes control of numerous IoT devices to create a botnet and then makes synchronised wishes to one or more servers for a particular service, congestion on the server and compelling it to fulfil actual requests from end users. That frequently happens when IoT devices overwhelm every device with messages, which is primarily done to create congestion on devices.

Attacks on RPL (routing protocol for low-power and Lossynetworks)

Routing Protocol for Power-Limited and Lossy Devices, the transmitter transmits the DODAG Information Object while the Destination Oriented Directed Acyclic Graph (DODAG) is being created (DIO). The receiver delivers its changed sibling list, parent list, ring message, and DAO message with route information after receiving the DIO [72]. The errant nodes never update after receiving the DIO message; instead, they always broadcast a false rank. The malicious node sends the DIO message to the other, non-malicious node, which receives it and modifies its rank based on the false rank. If a malicious node is the preferred parent after the formation of DODAG, the transmitting node will send the packet to the malicious node rather than the intended recipient,its parent just discards the packet, yielding a throughput of 0.

RPL builds and updates its graph topology and route table using three different forms of control signals. The DODAG Advertising Object (DAO), DODAG Information Solicitations, and the DODAG Information Object (DIO) are among the control messages (DIS). DIO is used for the development, upkeep, and discovery of the DODAG topology. As the RPL network is being launched by DIO, nodes exchange DODAG messages. With DIO, the nodes choose their preferred parents. For downward routing purposes, RPL employs DAO messages to send a node's prefix to its progenitor nodes. The DIS message is used by any unattached node in the network to find possible parents. After a certain amount of time, DIS is initiated by a node when it is unable to get DIO. The construction of an RPL network in a DODAG is known as an RPL instance. These RPL instances are capable of including a DODAG and having object functions. Assaults on the RPL topology include:

Sybil Attack: Sybil attack is when many nodes impersonate different peers in order to undermine an IoT ecosystem. It is employed to transmit fake data from an unreliable network. In the context of an e-health system, Sybil attacks—where a sensor node makes numerous false identities—could be quite harmful. A hacker might convey fraudulent information using these assaults by using bogus identities. As a result, either a true emergency situation is overlooked. A malicious node within a network has numerous identities in this attack. In a peer-to-peer network, a rogue node can influence the routing protocol, detection method, and routing mechanism.

Select Forwarding Attack: Attacks using selective forwarding include a malicious node acting as a legitimate node while deliberately dropping particular a node's or group of nodes' data packets [56]. A rogue node halts the data transmission that is arriving via it and refuses to advance it. A malicious and infected node might send the message over the incorrect network path.

Sinkhole attack: This type of attack targets data transmission from nearby nodes. A routing algorithm serves as the key tool for carrying out this. A sinkhole attack is an internal assault where a rogue node attempts to draw network traffic to it by promoting phoney routing changes. An attacker starts an attack by inserting phoney nodes into a network [32]. A sinkhole attack's primary goal is to divert traffic away from a targeted location through a hacked node that stands out as particularly alluring to other nodes [98].

Wormhole attack: Unfriendly nodes attack using wormholes, always provide the sender device and the recipient device with an illusion. In order to trick the base station into sending data through it and being lost on the route, a virtual tunnel is constructed that falsely pretends to be the shortest path between the two endpoints, which are the malicious nodes. Data that is being

transferred locally is intercepted by the attacker node and forwarded to a remote location. The attack might happen in either a hidden mode or a participation mode [55].

Hello Flooding Attack: One of the most frequent assaults on the network layer, the "Hello Flood Attack," forces Internet of Things (IoT) devices to broadcast themselves to their neighbours by sending Hello packets. The network node broadcasts the initial message as a Welcome packet to establish a connection. By sending a Welcome message, the cybercriminal can pose as a neighbour node to several nodes. A node will presume it is within radio range of the node that transmitted the Welcome packet if it gets one.

Blackhole Attack: In blackhole attack is a malicious apparatus that displays the network with a blackhole by falsely claiming to be the fastest route to the target location.

INTRUSION DETECTION DATASETS

The assessment datasets are essential to the validation of any IDS technique because they let us gauge how well the suggested method can identify invasive behaviour. Due to privacy concerns, The network packet analysis datasets utilised by commercial solutions are not easily accessible. Yet, a few datasets are available to the general public that are frequently used as benchmarks, including DARPA, KDD, NSL-KDD, and ADFA-LD. This section discusses the features and limitations of the existing datasets that are used for the construction and comparative assessment of IDS.

DARPA / KDD Cup99

The KDD98 (Knowledge Discovery and Data Mining (KDD)) dataset was the first IDS dataset to be produced by DARPA (Defence Advanced Research Project Agency) in 1998. DARPA launched a programme at MIT Lincoln Laboratories in 1998 to provide a complete and accurate environment for IDS benchmarking (Lincoln Laboratory, 1999). Nonetheless, this dataset was a crucial addition to the study of IDS, many people have disputed its accuracy and capacity to take into account real-world circumstances [38].

The network packets that were gathered were around four gigabytes in size and contained about 4,900,000 data. Each of the 2 million connection records in the test set of two weeks included 41 attributes and was assigned to one of two categories: normal or abnormal.

The data that was retrieved consists of a series of TCP sessions that begin and terminate at predetermined periods and are used to transfer data between a source IP address and a target IP address. These sessions feature a wide range of assaults that were simulated in a military network environment. The KDD Cup99 dataset, which was utilised in the Third International Knowledge Discovery and Data Mining Tools Competition, was created using the 1998 DARPA Dataset as its foundation (KDD, 1999).

CAIDA

This dataset, which was compiled in 2007, contains network traffic traces from Distributed Denial-of-Service (DDoS) assaults [47]. By overwhelming the target with a torrent of network packets, this kind of denial-of-service attack tries to stop routine traffic on a targeted computer or network from getting to its intended destination computer. The absence of a variety of assaults in the CAIDA dataset is one of its drawbacks.

NSL-KDD

The prior KDD cup99 dataset was used to produce the public dataset NSL-KDD [102]. an analytical study of the cup99 dataset revealed significant flaws that have a significant negative impact on the quality of intrusion detection and lead to an inaccurate assessment of AIDS [102]. The KDD data set has a significant amount of duplicate packets, which is the major problem. Due to the large number of duplicate instances in the training set, machine learning algorithms would be biased towards learning about regular instances and would not be able to learn about irregular cases, which are often more harmful to the computer system. To address the issues raised above, modified the dataset which has test dataset has 22,544 records, whereas the NSL-KDD train dataset has 125,973 records.

ISCX 2012

Real network traffic traces for the HTTP, SMTP, SSH, IMAP, POP3, and FTP protocols were examined in this dataset to detect typical computer behaviour [97]. This dataset is built on actual network traffic that has been labelled and includes a variety of assault cases.

ADFA-LD and ADFA-WD

Two datasets (ADFA-LD and ADFA-WD) were generated by researchers at the Australian Defence Force Academy as open-source datasets to illustrate the organisation and methodology of previous assaults [39]. The datasets, which were produced through the examination of system-call-based HIDS, include information from both Linux and Windows operating systems. To create ADFA- LD, the host operating system Ubuntu Linux version 11.04 was utilised (Creech and Hu, 2014). This dataset is appropriate for demonstrating differences between SIDS and AIDS techniques to intrusion detection since some of the assault cases in ADFA-LD were originated from fresh zero-day malware. It consists of three distinct types of data, each of which contains raw system call traces. The host was used to collect each training dataset, which included user actions like online browsing and creating LATEX documents.

IoT botnet

To assess our suggested approach, the Bot-IoT dataset—which contains both regular IoT network traffic and a range of attacks—is employed. DDoS, DoS, OS and Service Scan, Keylogging, and Data Exfiltration assaults are all included in the dataset.

Comparison Of Public IDS Datasets

The datasets that are utilised for machine learning techniques are crucial for a meaningful evaluation of these approaches since they are employed in AIDS research. The characteristics of the datasets are given in Table 10. We discovered that the popular KDD'99 or comparable sets designed for a wired network environment won't result in the development of optimal IDS aimed at the IoT ecosystem.

Table 1: Comparison Of Datasets

Table 1: The comparison of datasets (✓ = True, ✗ = False)

Dataset	Real Traffic	Label data	IoT traces	Zero-day attacks	Full packet captured	Year
DARPA 98	✓	✓	✗	✗	✓	1998
KDDCUP 99	✓	✓	✗	✗	✓	1999
CAIDA	✓	✗	✗	✗	✗	2007
NSL-KDD	✓	✓	✗	✗	✓	2009
ISCX 2012	✓	✓	✗	✗	✓	2012
ADFA-WD	✓	✓	✗	✓	✓	2014
ADFA-LD	✓	✓	✗	✓	✓	2014
CICIDS2017	✓	✓	✗	✓	✓	2017
Bot-IoT	✓	✓	✓	✓	✓	2018

CHALLENGES OF IOT IDS

The number of massive communicated devices is fast increasing in the era of IoT (Internet of Things). Using the aforementioned IDSs to secure communications in an IoT environment presents challenges and interesting research avenues.

IDSs have obviously been the subject of extensive research, but there are still a number of important problems that need to be solved. IDSs must be expanded, precise, capable of identifying a variety of intrusions with fewer false alarms, and capable of overcoming other difficulties.

Feature -Engineer extraction

The design of the traffic characteristics utilised in training has a significant impact on the method's ability to recognise objects. When alternative feature sets of network traffic are employed, the IDS accuracy frequently performs differently.

IoT Device Limitations

IoT devices have small memory space, which makes it difficult to keep track of things because the system runs continuously and can be overwritten owing to the memory capacity being so low, increasing the chance that crucial evidence could be lost. Because IoT devices have limited capacity, data may be easily erased or not stored at all in some IoT devices. Transferring the data to the storage device could be a technique to save it, however this option isn't always effective because data can be readily changed while being transferred to the local storage device. The computational power of the other IoT device is its restriction. A cyberthief may use the stored energy to send out a torrent of good or bad signals, exposing the sensors inaccessible to authorised users [82].

Because certain IoT devices are transported in environments where charging is not possible, they only have a limited amount of energy the devices' resources to run the IDS designer and rigorous IDS analysis. For the Internet of Things, a Lightweight Intrusion Detection System must be developed, and thus fewest amount of security requirements on the IoT device as possible, this is necessary. By simplifying the difficult features extraction and features, a lightweight IDS system might be created. To accurately identify an intrusion in the IoT ecosystem, a small number of attributes need be derived from raw data. Feature selection aids in lowering computational complexity, eradicating redundant data, increasing reducing false alarm rates, simulating data, and increasing the detection rate of machine learning systems. Several methods have been employed in this field of study to develop a small IoT IDS.

Problems Of Smart Devices

A poorly setup IoT device or one that releases firmware updates for smart devices slowly might lead to security problems. IoT gadgets, for instance, might be utilised for illegal purposes, alternatively, a hacker with access to an IoT device might spy. passwords chosen in advance by the manufacturer are another issue. For instance, the authentication login is easily accessible online. The fact that many IoT devices have communication ports accessible to the external network complicates cybercriminals' activity in another way.

Overhead traffic

The performance of traffic-based trust computation in identifying insider assaults in conventional network contexts is good. Huge packets have become a problem with the high-speed network connection, though, since the traffic may seriously exceed an IDS's finite processing capacity.

Heterogeneity device type

Different heterogeneous methods exist. The IoT connects numerous sorts of devices to enable communication between the real and virtual worlds. Smart phones, watches, microwave, ACs, lights, automated home systems and other gadgets in general may all be connected.

The fact that the numerous heterogeneous devices operate on different platforms and frameworks makes connecting them to one another a particularly difficult task. The development of the IDS will be a very difficult task due to the characteristics of the Internet of Things are in abundance of varied devices, complexity at the network level, communication between different communication protocols, heterogeneity at the device and network levels, and the vast number of activities that these sensors naturally create.

Privacy

The vast bulk of IoT datasets are held by large organisations who are reluctant to release them openly. Access to datasets with copyright restrictions or privacy issues. In the domain of personal data, including healthcare and education, these are more generic.

Feature extraction

The aim of feature extraction is to obtain the network traffic from the communication of IoT gadgets. The context and purpose of each packet moving across the network must be extracted by IDS as attributes. The package might be one of the billions of malicious packets sent with the intent to source harmful operations, or it could be a regular connection to communicate with a server [76]. Because packets from distinct subnet networks overlap, several networks may be connected at once, and a fast connection, extracting these kinds of data from IoT network traffic can be difficult [93].

Big IoT data

Growth in the quantity, diversity, and speed of IoT data as well as a sharp increase in the the quantity of linked devices. When more and more physical items are connected to the internet, scaling issues usually arise [101]. several levels, such as data networking and transit, data processing and administration, and service supply, scalability is difficult when there are many distinct things. Large amounts of data being transmitted simultaneously throughout the IoT ecosystem can also result in recurring delays, conflicts, and communication issues. Creating networking technologies and standards that enable data collected by many different devices to travel quickly across IoT networks is a difficult challenge.

Immaturity of communication protocol

To identify IoT threats, IDS is typically included in IoT protocols. The maturing of stable IDS has an effect on the inexperience of security procedures. Hence, IDS borrowed characteristics from network protocols [82]. IoT protocols are many, and devices present problems that are undoubtedly worthwhile to be met in order to build IoT IDS and robustness. The wireless networking protocols that operate at the physical and data connection levels, as well as other protocols and standards specifically designed for IoT applications, power the IoT ecosystem. Bluetooth and ZigBee are examples of wireless personal area networks (WPANs) that are used for short-range communication. Near-Field Communication is yet another short-range wireless protocol utilised by many Internet of Things sensors (NFC). Cellular networks are mostly used for a greater range.

Data collection

Every IoT sensor presents a hurdle for data collecting. As the data is being processed, it may be updated, changed, or even disappear entirely. Also, it is difficult to locate evidence because to decentralised data, unknown or inaccessible physical locations, big and dynamic systems, and data erasure upon IoT reboot. In order to secure customer data, cloud service providers do not provide any information on the underlying workings of the cloud. This presents another data collecting problem. For instance, since information might be encrypted before being saved in the cloud, the data gathered from IoT devices may be in a different format from the data stored there. Moreover, archival data on tolerance obtained by professional groups has comparable challenges. A cyber-security specialist is also required to gather a dataset that includes both regular traffic and network assaults.

Unavailability of training datasets

Deep learning and machine learning require substantial datasets, which are currently absent, for effective use. Also, it is important to look at the rules and setups needed for describing the learning procedures in spite of anything. More accurate datasets from real world needs to be collected and analyzed with various combinations of DL and RL calculations. There have been attempts up to this point to adjust to this test. Nonetheless, further study in this area is now needed.

Challenges of IoT IDS for ICS

In recent years, a wide range of industrial IoT systems have been employed in infrastructures related to transportation, manufacturing, retail, and smart cities. Day by day devices are increasing rapidly with IoT, thanks to advancements in sensor network technologies, wireless communication, smartphones, healthcare (like remote patient 24-hour care), smart grid, home automation (like security, heating, and lighting management), and smart cities (like distributed pollution monitoring). Cyber-Physical Systems (CPS) are composed of physical sensors and actuators that are networked with computer-based control systems. As a result, CPS rely on the IoT ecosystem.

A hacked ICS may have catastrophic effects for national security, the economy, and public health and safety. Explosions, hazardous toxic chemical discharges, and huge cascading power outages have all been caused by compromised ICS systems. Use of secure ICSs is necessary for performance that is dependable, safe, and adaptable.

Microsoft no longer releases security updates for outdated systems, making them vulnerable to ransomware and zero-day malware attacks.

Challenge Of IoT IDS On Intrusion Evasion Detection

In order to prevent SIDS and AIDS, it is crucial to identify attacks that are concealed by evasion strategies. The effectiveness of evasion strategies would depend on the IDS's capacity to recover the assaults' original signature or produce new signatures to mask their alteration. More research is still needed to determine how resilient IDS is to different evasion techniques. SIDS in regular expressions, for instance, can spot minor changes like rearranging spaces, but they are still worthless against a number of obfuscation methods that hackers employ to breach the security.

CONCLUSION & FUTURE SCOPE

We have provided a comprehensive study of IoT intrusion detection system methodologies, deployment strategy, validation methodology, dataset, and technology, as well as their advantages and disadvantages, drawbacks, in this study. To identify IoT threats, a number of intrusion detection systems have been developed. Due to IoT design, these approaches could struggle to identify all IoT assaults. In order to address IoT security concerns, we reviewed previous research findings and looked at current models for IoT IDS performance enhancement. We also clarified the limitations of the conventional IoT intrusion detection method. The obstacles and potential avenues for future research were then outlined, and we talked about the current IDS.

A unique IDS must be created in order to create reliable IoT IDS based on heterogeneous device categories. We have identified four components that are essential to the development of dependable IDS for the IoT. Due to the vast amount of data, you should first try to minimise false alarms. As a result of unexpected behaviour in IoT sensors that earlier seemed normal, attacks may start to be considered, it is important to second, be extremely adaptable to extreme IoT communication systems. Finally, as new vulnerabilities are discovered, be able to recognise zero-day attacks. Fourth, employ modern machine learning and deep learning algorithms that can learn from massive IoT data and be autonomous IDS.

In conclusion, we think that by examining the current state of this significant and extremely dynamic area of research, thereby providing a platform for scientists to create and design a complete IDS to remove IoT security concerns dealing with IoT devices management and connection, this review may make a significant contribution to security researchers.

REFERENCES

1. Lee, I. The Internet of things for enterprises: An ecosystem, architecture, and IoT service business model. Internet Things Eng. Cyber Phys. Hum. Syst. 2019, 7, 100078.
2. Chakraborty, A.; Chandru, V.; Rao, M.R. A linear programming primer: From Fourier to Karmarkar. Ann. Oper. Res. 2020, 287, 593–616.
3. Narayan, Vipul, and A. K. Daniel. "CHHP: coverage optimization and hole healing protocol using sleep and wake-up concept for wireless sensor network." International Journal of System Assurance Engineering and Management 13.Suppl 1 (2022): 546-556.
4. Nurse, J.R.C.; Creese, S.; de Roure, D. Security risk assessment in Internet of Things systems. IT Prof. 2017,19, 20–26.

5. Malik, V.; Singh, S. Security risk management in IoT environment. J. Discret. Math. Sci. Cryptogr. 2019, 22,697–709.
6. Markets and Markets. IoT Security Market Worth $35.2 Billion by 2023. 2019. Available online: https://www.marketsandmarkets.com/PressReleases/iot-security.asp (accessed on 17 September 2020).
7. Narayan, Vipul, and A. K. Daniel. "CHOP: Maximum coverage optimization and resolve hole healing problem using sleep and wake-up technique for WSN." ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal 11.2 (2022): 159-178.
8. Irdeto. New 2019 Global Survey: IoT-Focused Cyberattacks Are the New Normal. 2019. Available online: https://resources.irdeto.com/global-connected-industries-cybersecurity-survey/new-2019-globalsurvey- iot-focused-cyberattacks-are-the-new-normal (accessed on 17 September 2020).
9. Narayan, Vipul, and A. K. Daniel. "IOT based sensor monitoring system for smart complex and shopping malls." Mobile Networks and Management: 11th EAI International Conference, MONAMI 2021, Virtual Event, October 27-29, 2021, Proceedings. Cham: Springer International Publishing, 2022.
10. Narayan, Vipul, and A. K. Daniel. "A novel approach for cluster head selection using trust function in WSN." Scalable Computing: Practice and Experience 22.1 (2021): 1-13.Rao, A.; Carreón, N.; Lysecky, R.; Rozenblit, J. Probabilistic threat detection for risk management in cyber-physical medical systems. IEEE Softw. 2018, 35, 38–43.
11. Narayan, Vipul, and A. K. Daniel. "Multi-tier cluster based smart farming using wireless sensor network." 2020 5th international conference on computing, communication and security (ICCCS). IEEE, 2020.
12. Bendavid, Y.; Bagheri, N.; Safkhani, M.; Rostampour, S. IoT Device Security: Challenging "A Lightweight RFID Mutual Authentication Protocol Based on Physical Unclonable Function". Sensors 2018, 18, 4444.
13. Hejazi, D.; Liu, S.; Farnoosh, A.; Ostadabbas, S.; Kar, S. Development of use-specific high-performance cyber-nanomaterial optical detectors by effective choice of machine learning algorithms. Mach. Learn. Sci. Technol. 2020, 1, 025007.
14. Mollah, M.B.; Azad, M.A.; Vasilakos, A. Security and privacy challenges in mobile cloud computing: Survey and way ahead. J. Netw. Comput. Appl. 2017, 84, 38–54.
15. Sha, K.; Wei,W.; Yang, T.A.; Wang, Z.; Shi, W. On security challenges and open issues in Internet of Things. Future Gener. Comput. Syst. 2018, 83, 326–337.
16. Yu, R.; Xue, G.; Kilari, V.T.; Zhang, X. Deploying Robust Security in Internet of Things. In Proceedings of the 2018 IEEE Conference on Communications and Network Security (CNS), Beijing, China, 30 May–1 June 2018;pp. 1–9.
17. Abbasi, J. Wetzels, W. Bokslag, E. Zambon, S. Etalle, "On emulation-based network intrusion detection systems," in Research in attacks, Intrusions and Defenses: 17th International Symposium, RAID 2014, Gothenburg, Sweden, September 17–19, 2014. Proceedings, A. Stavrou, H. Bos, G. Portokalidis, Cham: Springer International Publishing, 2014, pp. 384–404
18. Aburomman AA, Ibne Reaz MB (2016) A novel SVM-kNN-PSO ensemble method for intrusion detection system. Appl Soft Comput 38:360–372
19. Narayan, Vipul, and A. K. Daniel. "Multi-tier cluster based smart farming using wireless sensor network." 2020 5th international conference on computing, communication and security (ICCCS). IEEE, 2020.
20. Agrawal S, Agrawal J (2015) Survey on anomaly detection using data mining techniques. Procedia Computer Science 60:708–713

21. Alazab A, Hobbs M, Abawajy J, Alazab M (2012) Using feature selection for intrusion detection system. In: 2012 International Symposium on Communications and Information Technologies (ISCIT), pp 296–301
22. Narayan, Vipul, and A. K. Daniel. "Design consideration and issues in wireless sensor network deployment." (2020): 101-109.Alcaraz C (2018) Cloud-assisted dynamic resilience for cyber-physical control systems. IEEE Wirel Commun 25(1):76–82
23. Annachhatre C, Austin TH, Stamp M (2015) Hidden Markov models for malware classification. J Comput Virol Hack Technique 11(2):59–73
24. Axelsson S (2000) "Intrusion detection systems: A survey and taxonomy," Technical report on Security Systems.
25. Narayan, Vipul, et al. "E-Commerce recommendation method based on collaborative filtering technology." International Journal of Current Engineering and Technology 7.3 (2017): 974-982.
26. Benkhelifa E, Welsh T, Hamouda W (2018) A critical review of practices and challenges in intrusion detection systems for IoT: toward universal and resilient systems. IEEE Commun Survey Tutor 20(4):3496–3509
27. Bhuyan MH, Bhattacharyya DK, Kalita JK (2014) Network anomaly detection: methods, systems and tools. IEEE Commun Survey Tutorial 16(1):303–336
28. Narayan, Vipul, and A. K. Daniel. "Energy Efficient Protocol for Lifetime Prediction of Wireless Sensor Network using Multivariate Polynomial Regression Model." Journal of Scientific & Industrial Research 81.12 (2022): 1297-1309.
29. Choudhary, Shubham, et al. "Fuzzy approach-based stable energy-efficient AODV routing protocol in mobile ad hoc networks." Software Defined Networking for Ad Hoc Networks. Cham: Springer International Publishing, 2022. 125-139.
30. Breiman L (1996) Bagging predictors. Machine Learn 24(2):123–140 Buczak AL, Guven E (2016) A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Commun Surveys Tutorial 18(2):1153–1176
31. Narayan, Vipul, A. K. Daniel, and Ashok Kumar Rai. "Energy efficient two tier cluster based protocol for wireless sensor network." 2020 international conference on electrical and electronics engineering (ICE3). IEEE, 2020.
32. Can O, Sahingoz OK (2015) A survey of intrusion detection systems in wireless sensor networks. In: 2015 6th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO), pp 1–6 IEEE
33. Cervantes C, Poplade D, Nogueira M, Santos A (2015) Detection of sinkhole attacks for supporting secure routing on 6 LoWPAN for Internet of Things. In: 2015 IFIP/IEEE International Symposium on Integrated Network Management (IM), pp 606–611 IEEE
34. Chaabouni N, Mosbah M, Zemmari A, Sauvignac C, Faruki P (2019) Network Intrusion Detection for IoT Security Based on Learning Techniques, in IEEE Communications Surveys & Tutorials, vol. 21, no. 3, pp. 2671–2701, thirdquarter 2019. https://doi.org/10.1109/COMST.2019.2896380
35. Awasthi, Shashank, et al. "A Comparative Study of Various CAPTCHA Methods for Securing Web Pages." 2019 International Conference on Automation, Computational and Technology Management (ICACTM). IEEE, 2019.
36. Chebrolu S, Abraham A, Thomas JP (2005) Feature deduction and ensemble design of intrusion detection systems. Comput Security 24(4):295–307
37. Cho EJ, Kim JH, Hong CS (2009) Attack model and detection scheme for botnet on 6LoWPAN. Springer Berlin Heidelberg, Berlin, pp 515–518

38. Narayan, Vipul, and A. K. Daniel. "FBCHS: Fuzzy Based Cluster Head Selection Protocol to Enhance Network Lifetime of WSN." ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal 11.3 (2022): 285-307.
39. Narayan, Vipul, and A. K. Daniel. "Novel protocol for detection and optimization of overlapping coverage in wireless sensor networks." Int. J. Eng. Adv. Technol 8 (2019).
40. Creech G (2014) Developing a high-accuracy cross platform host-based intrusion detection system capable of reliably detecting zero-day attacks. University of New South Wales, Canberra
41. da Costa KAP, Papa JP, Lisboa CO, Munoz R, de Albuquerque VHC (2019) Internet of Things: A survey on machine learning-based intrusion detection approaches. Comput Network 151:147–157
42. Narayan, Vipul, et al. "To Implement a Web Page using Thread in Java." (2017).Diro AA, Chilamkurti N (2018) Distributed attack detection scheme using deep learning approach for internet of things. Futur Gener Comput Syst 82: 761–768
43. Dua S, Du X (2016) Data Mining and Machine Learning in Cybersecurity Publishers Auerbach. Publications Location UK
44. Irfan, Daniyal, et al. "Prediction of Quality Food Sale in Mart Using the AI-Based TOR Method." Journal of Food Quality 2022 (2022).
45. Granjal J, Monteiro E, Silva JS (2015) Security for the internet of things: a survey of existing protocols and open research issues. IEEE Commun Survey Tutor 17(3):1294–1312
46. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The WEKA data mining software: an update. ACM SIGKDD explorations newsletter 11(1):10–18
47. Pramanik, Sabyasachi, et al. "A novel approach using steganography and cryptography in business intelligence." Integration Challenges for Analytics, Business Intelligence, and Data Mining. IGI Global, 2021. 192-217.
48. H. Hindy et al., "A taxonomy and survey of intrusion detection system design techniques, network threats and datasets," arXiv preprint arXiv: 1806.03517, 2018
49. Hodo E et al (2016) Threat analysis of IoT networks using artificial neural network intrusion detection system. In: 2016 International Symposium on Networks, Computers and Communications (ISNCC), pp 1–6
50. Hoque MAM, Bikas AN (2012) An implementation of intrusion detection system using genetic algorithm. arXiv preprint arXiv:1204.1336. Chicago; 109–120
51. Smiti, Puja, Swapnita Srivastava, and Nitin Rakesh. "Video and audio streaming issues in multimedia application." 2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2018.
52. Ji S-Y, Jeong B-K, Choi S, Jeong DH (2016) A multi-level intrusion detection method for abnormal network behaviors. J Network Comput Application 62(Supplement C):9–17
53. KDD. (1999). The 1999 KDD intrusion detection. Available: http://kdd.ics.uci.edu/ databases/kddcup99/task.html
54. Kenkre PS, Pai A, Colaco L (2015) Real time intrusion detection and prevention system. In: Satapathy SC, Biswal BN, Udgata SK, Mandal JK (eds) Proceedings of the 3rd international conference on Frontiers of intelligent computing: theory and applications (FICTA) 2014: volume 1. Springer International Publishing, Cham, pp 405–411
55. Khabbazian M, Mercier H, Bhargava VK (2006) Nis02–1: Wormhole attack in wireless ad hoc networks: Analysis and countermeasure. In: IEEE Globecom 2006, pp 1–6 IEEE

56. Smiti, Puja, Swapnita Srivastava, and Nitin Rakesh. "Video and audio streaming issues in multimedia application." 2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2018.
57. Khraisat A, Gondal I, Vamplew P (2018) An Anomaly Intrusion Detection System Using C5 Decision Tree Classifier. In: Trends and Applications in
58. Srivastava, Swapnita, and P. K. Singh. "Proof of Optimality based on Greedy Algorithm for Offline Cache Replacement Algorithm." International Journal of Next-Generation Computing 13.3 (2022).
59. Khraisat A, Gondal I, Vamplew P, Kamruzzaman J (2019a) "Survey of intrusion detection systems: techniques, datasets and challenges," Cybersecurity. J Article 2(1):20
60. Khraisat A, Gondal I, Vamplew P, Kamruzzaman J, Alazab A (2019b) A Novel Ensemble of Hybrid Intrusion Detection System for Detecting Internet of Things Attacks. Electronics 8(11):1210
61. Khraisat A, Gondal I, Vamplew P, Kamruzzaman J, Alazab A (2020) Hybrid Intrusion Detection System Based on the Stacking Ensemble of C5 Decision Tree Classifier and One Class Support Vector Machine. Electronics 9(1):173
62. Koc L, Mazzuchi TA, Sarkani S (2012) A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier. Expert Syst Appl 39(18):13492–13500
63. Koroniotis N, Moustafa N, Sitnikova E, Turnbull B (2018) "Towards the Development of Realistic Botnet Dataset in the Internet of Things for Network Forensic Analytics: Bot-IoT Dataset," arXiv preprint arXiv:1811.00701
64. Kreibich C, Crowcroft J (2004) Honeycomb: creating intrusion detection signatures using honeypots. SIGCOMM Comput Commun Rev 34(1):51–56
65. Li Y, Xia J, Zhang S, Yan J, Ai X, Dai K (2012) An efficient intrusion detection system based on support vector machines and gradually feature removal method. Expert Syst Appl 39(1):424–430
66. Liao H-J, Lin C-HR, Lin Y-C, Tung K-Y (2013b) Intrusion detection system: a comprehensive review. J Netw Comput Appl 36(1):16–24
67. Liao H-J, Richard Lin C-H, Lin Y-C, Tung K-Y (2013a) Intrusion detection system: A comprehensive review. J Network Comput Appl 36(1):16–24
68. Lin C, Lin Y-D, Lai Y-C (2011) A hybrid algorithm of backward hashing and automaton tracking for virus scanning. IEEE Trans Comput 60(4):594–601
69. Lin W-C, Ke S-W, Tsai C-F (2015) CANN: An intrusion detection system based on combining cluster centers and nearest neighbors. Knowledge Based Syst 78(Supplement C):13–21
70. MIT Lincoln Laboratory. (1999). DARPA Intrusion Detection Data Sets. Available: https://www.ll.mit.edu/ideval/data/
71. Srivastava, Swapnita, and P. K. Singh. "HCIP: Hybrid Short Long History Table-based Cache Instruction Prefetcher." International Journal of Next-Generation Computing 13.3 (2022).Mayzaud A, Badonnel R, Chrisment I (2016) A taxonomy of attacks in RPL-based internet of things. Int J Network Security 18(3):459–473
72. Srivastava, Swapnita, and Shilpi Sharma. "Analysis of cyber related issues by implementing data mining Algorithm." 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2019.
73. Awasthi, Shashank, Naresh Kumar, and Pramod Kumar Srivastava. "A study of epidemic approach for worm propagation in wireless sensor network." Intelligent Computing in Engineering: Select Proceedings of RICE 2019. Springer Singapore, 2020.

74. Meiners CR, Patel J, Norige E, Torng E, Liu AX (2010) "Fast regular expression matching using small TCAMs for network intrusion detection and prevention systems," presented at the proceedings of the 19th USENIX conference on security, Washington, DC
75. Meshram A, Haas C (2017) Anomaly Detection in Industrial Networks using Machine Learning: A Roadmap. In: Beyerer J, Niggemann O, Kühnert C (eds) Machine Learning for Cyber Physical Systems: Selected papers from the International Conference ML4CPS 2016. Springer Berlin Heidelberg, Berlin, pp 65–72
76. Mirsky Y, Doitshman T, Elovici Y, Shabtai A (2018) "Kitsune: an ensemble of autoencoders for online network intrusion detection," arXiv preprint arXiv: 1802.09089
77. Modi C, Patel D, Borisaniya B, Patel H, Patel A, Rajarajan M (2013) A survey of intrusion detection techniques in Cloud. J Network Comput Appl 36(1):42–57
78. Smriti, Puja, Swapnita Srivastava, and Saurabh Singh. "Keyboard invariant biometric authentication." 2018 4th International Conference on Computational Intelligence & Communication Technology (CICT). IEEE, 2018.
79. Moustafa N, Turnbull B,Choo KR (2019) "An Ensemble Intrusion Detection Technique Based on Proposed Statistical Flow Features for Protecting Network Traffic of Internet of Things," IEEE Internet of Things Journal, vol. 6, pp. 4815-4830
80. Murray SN, Walsh BP, Kelliher D, O'Sullivan DTJ (2014) Multi-variable optimization of thermal energy efficiency retrofitting of buildings using static modelling and genetic algorithms – A case study. Build Environ 75(Supplement C):98–107
81. Nourian A, Madnick S (2018) A systems theoretic approach to the security threats in cyber physical systems applied to Stuxnet. IEEE Transact Dependable Secure Comput 15(1):2–13
82. Neshenko N, Bou-Harb E, Crichigno J, Kaddoum G, Ghani N (2019) Demystifying IoT security: an exhaustive survey on IoT vulnerabilities and a first empirical look on internet-scale IoT exploitations. IEEE Commun Survey Tutorial 21(3): 2702–2733
83. Nourian A, Madnick S (2018) A systems theoretic approach to the security threats in cyber physical systems applied to Stuxnet. IEEE Transact Dependable Secure Comput 15(1):2–13
84. Pretorius B, van Niekerk B (2016) Cyber-security for ICS/SCADA: a south African perspective. Int J Cyber Warfare Terrorism (IJCWT) 6(3):1–16
85. Quinlan JR (1986) Induction of decision trees. Mach Learn 1(1):81–106 Quinlan JR (2014) C4. 5: Programs for Machine Learning; Morgan Kaufmann Publishers Inc.: San Francisco; 2014;8
86. Pretorius B, van Niekerk B (2016) Cyber-security for ICS/SCADA: a south African perspective. Int J Cyber Warfare Terrorism (IJCWT) 6(3):1–16
87. Rathore S, Park JH (2018) Semi-supervised learning based distributed attack detection framework for IoT. Appl Soft Comput 72:79–89
88. Rege-Patwardhan A (2009) Cybercrimes against critical infrastructures: a study of online criminal organization and techniques. Crim Justice Stud 22(3): 261–271
89. K. Riesen, H. Bunke, "IAM Graph Database Repository for Graph Based Pattern Recognition and Machine Learning," in Structural, Syntactic, and Statistical Pattern Recognition: Joint IAPR International Workshop, SSPR & SPR 2008, Orlando, USA, December 4–6, 2008. Proceedings, N. da Vitoria Lobo et al., Berlin: Springer Berlin Heidelberg, 2008, pp. 287–297
90. Roesch M (1999) Snort-lightweight intrusion detection for networks. In: Proceedings of LISA '99: 13th Systems Administration Conference Seattle, Seattle, pp 229–238

91. Rutkowski L, Jaworski M, Pietruczuk L, Duda P (2014) Decision trees for mining data streams based on the Gaussian approximation. IEEE Trans Knowl Data Eng 26(1):108–119
92. S. Duque and M. N. b. Omar (2015) Using Data Mining Algorithms for Developing a Model for Intrusion Detection System (IDS). Procedia Comput Sci 61(Supplement C):46–51
93. S. P. R. M et al (2020) An effective feature engineering for DNN using hybrid PCA- GWO for intrusion detection in IoMT architecture. Comput Commun 160: 139–149
94. Sharafaldin I, Lashkari AH, Ghorbani AA (2018) Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. In: ICISSP, pp 108–116
95. Shen C, Liu C, Tan H, Wang Z, Xu D, Su X (2018) Hybrid-augmented device fingerprinting for intrusion detection in industrial control system networks. IEEE Wirel Commun 25(6):26–31
96. Sherasiya T, Upadhyay H, Patel HB (2016) A survey: Intrusion detection system for internet of things. Int J Comput Sci Eng (IJCSE) 5(2):91–98
97. Shiravi A, Shiravi H, Tavallaee M, Ghorbani AA (2012) Toward developing a systematic approach to generate benchmark datasets for intrusion detection. Comput Security 31(3):357–374
98. Singh AP, Singh P, Kumar R (2015) A Review on Impact of Sinkhole Attack in Wireless Sensor Networks. Int J 5(8)
99. Subramanian S, Srinivasan VB, Ramasa C (2012) Study on classification algorithms for network intrusion systems. J Commun Comput 9(11):1242–1246
100. Symantec (2017) Internet Security Threat Report 2017, vol 22
101. Tang M, Alazab M, Luo Y (2019) Big data for Cybersecurity: vulnerability disclosure trends and dependencies. IEEE Transact Big Data 5(3):317–329
102. Tavallaee M, Bagheri E, Lu W, Ghorbani AA (2009) "A detailed analysis of the KDD CUP 99 data set," in 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, pp 1–6
103. Thaseen S, Kumar CA (2013) An analysis of supervised tree based classifiers for intrusion detection system. In: 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, pp 294–299
104. Vasan D, Alazab M, Venkatraman S, Akram J, Qin Z (2020a) MTHAEL: cross- architecture IoT malware detection based on neural network advanced ensemble learning. IEEE Trans Comput 69(11):1654–1667
105. Vasan D, Alazab M, Wassan S, Naeem H, Safaei B, Zheng Q (2020b) IMCFN: Image-based malware classification using fine-tuned convolutional neural network architecture. Computer Networks 171:107138
106. Vasan D, Alazab M, Wassan S, Safaei B, Zheng Q (2020c) Image-Based malware classification using ensemble of CNN architectures (IMCEC). Comput Security 92:101748
107. Venkatraman S, Alazab M (2018) Use of Data Visualisation for Zero-Day Malware Detection. Security Commun Network 2018:1728303
108. Vigna G, Kemmerer RA (1999) NetSTAT: a network-based intrusion detection system. J Comput Secur 7:37–72
109. Wang G, Hao J, Ma J, Huang L (2010) A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering. Expert Syst Application 37(9):6225–6232
110. Wang W et al (2018) HAST-IDS: learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection. IEEE Access 6: 1792–1806
111. Wang X, Han Y, Leung VC, Niyato D, Yan X, Chen X (2020) Convergence of edge computing and deep learning: a comprehensive survey. IEEE Commun Survey Tutorial
112. Xiao L, Wan X, Lu X, Zhang Y, Wu D (2018) IoT security techniques based on machine learning: how do IoT devices use AI to enhance security? IEEE Signal Process Mag 35(5):41–49

113. Yang X, Tian YL (2012) EigenJoints-based action recognition using Naïve-Bayes- Nearest-Neighbor. In: 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 14–19
114. Yang Y, Wu L, Yin G, Li L, Zhao H (2017) A survey on security and privacy issues in internet-of-things. IEEE Internet Things J 4(5):1250–1258
115. Yar M, Steinmetz KF (2019) Cybercrime and society. SAGE Publications Limited Yin C, Zhu Y, Fei J, He X (2017) A deep learning approach for intrusion detection using recurrent neural networks. IEEE Access 5:21954–21961
116. Knowledge Discovery and Data Mining, Cham. Springer International Publishing, pp 149–155
117. Zarpelao BB, Miani RS, Kawakani CT, de Alvarenga SC (2017) A survey of intrusion detection in internet of things. J Netw Comput Appl 84:25–37
118. Zhang G, Yan C, Ji X, Zhang T, Zhang T, Xu W (2017) Dolphin attack: Inaudible voice commands. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, pp 103–117
119. 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, 1-27.