International Journal of Engineering in Computer Science

E-ISSN: 2663-3590
P-ISSN: 2663-3582
www.computersciencejournals.com/ijecs
IJECS 2025; 7(1): 98-105
Received: 10-01-2025
Accepted: 13-02-2025

Sonia Mahesh Verma
Assistant Professor & Ph.D. Scholar, Chimanbhai Patel Institute of Computer Applications, Sardar Vallabhbhai Global University, Ahmedabad, Gujarat, India

Dr. Priyank Nahar
Associate Professor, Chimanbhai Patel Post Graduate Institute of Computer Applications, Sardar Vallabhbhai Global University, Ahmedabad - 380015, Gujarat, India

Corresponding Author:
Sonia Mahesh Verma
Assistant Professor & Ph.D. Scholar, Chimanbhai Patel Institute of Computer Applications, Sardar Vallabhbhai Global University, Ahmedabad, Gujarat, India

Machine learning and IOT Security: A review

Sonia Mahesh Verma and Priyank Nahar

DOI: https://www.doi.org/10.33545/26633582.2025.v7.i1b.161

Abstract

We are living in a connected world which encompasses various types of devices and networks. The most widely used network in modern age is Internet of Things (IoT). It is a type of network where IOT devices can communicate with each other without human intervention. The life of people is transformed with the evolution of smart city, smart home concepts. The areas are not limited. It ranges from healthcare, medical, transportation to banking, smart grid, agriculture. It has transformed the lives of people by providing ease and comfort of completing the tasks in less time with less overhead. However, there is a negative side of IoT and that is it lacks the implementation of sound and effective security measures. There are various risks associated with use of IoT. Some of them are risk of unauthorized access of data, node spoofing and different cyber-attacks like denial of service (DoS), Botnet, Ransom ware, eavesdropping, zero day attack. So, the main concern addressed here is how to detect intrusions into the network using Machine Learning. Machine Learning (ML) plays an important role in designing of intrusion detection systems to detect the anomalies in the network. The ML techniques are very efficient in detection of different types of cyber-attacks. This paper presents the review of different ML techniques in detection of intrusions into IoT networks. Different ML techniques like Random Forest, Neural Network, support vector machine etc. are useful in detection of anomalies in the network. Different datasets like UNSW-NB 15, CICIDS-2017, KDD-99, NSL-KDD, TON_IoT and ECU-IOHT are available which consists of data related to different types of cyber-attacks and are used to form intrusion detection systems. This paper presents the application of different ML algorithms in detection of abnormalities in the network, comparative analysis of different feature selection and machine learning techniques along with challenges and issues present in traditional intrusion detection systems(IDS) in IOT networks.

Keywords: Internet of things, security, machine learning, intrusion detection system

Introduction

IoT refers to network of physical objects which can communicate with each other and end user with the help of Internet. It can be defined as the interconnection of objects or things embedded with in the electronics software, sensors and connectivity to enable it to achieve greater value and service by exchanging data with manufacturer, operator and other connected devices. Each thing is uniquely identifiable through its embedded computing system. The conceptual framework of Internet of things can be described as physical object + controller, sensor & actuators +Internet = Internet of things [1]. The application of IoT involves many fields like agriculture, healthcare, transportation, homes etc. These fields are rapidly developing with the power of IoT [2-4].

There are three layers in IoT- sensing (perception) layer, network layer, data processing layer and application layer shown in Figure-1 [16]. The sensing (perception) layer is responsible for collection of data from different sources. The main components of this layer are sensors and actuators. Sensors are responsible for measuring physical parameters like humidity, temperature etc. and convert them to electrical signals. These are placed at input port of the system. Actuators use these electrical signals for creating force, motion, sound etc. Both sensors and actuators work simultaneously to complete a task. The network layer is responsible for transmission of data between networks. It forms the packets and transmit them. The data processing layer collect and analyse the data using which a meaningful insight and decisions can be made. This layer is a sensitive layer as the cyber attackers use this layer to perform cyber-attacks like Man in the Middle attack (MITM), Denial of Service (DOS), exploit etc. The hardware of IoT device consists of microcontroller, firmware, sensors, control unit, actuators. The network layer consists of device APIs and device interface for communication over network, middleware for creating communication stack, software for messages, information and commands which the devices receive and then output to actuators [1]. The data processing layer process the data for further analysis. The application layer is present at the top and it interact with the user directly. It provides access to users to control IOT devices.

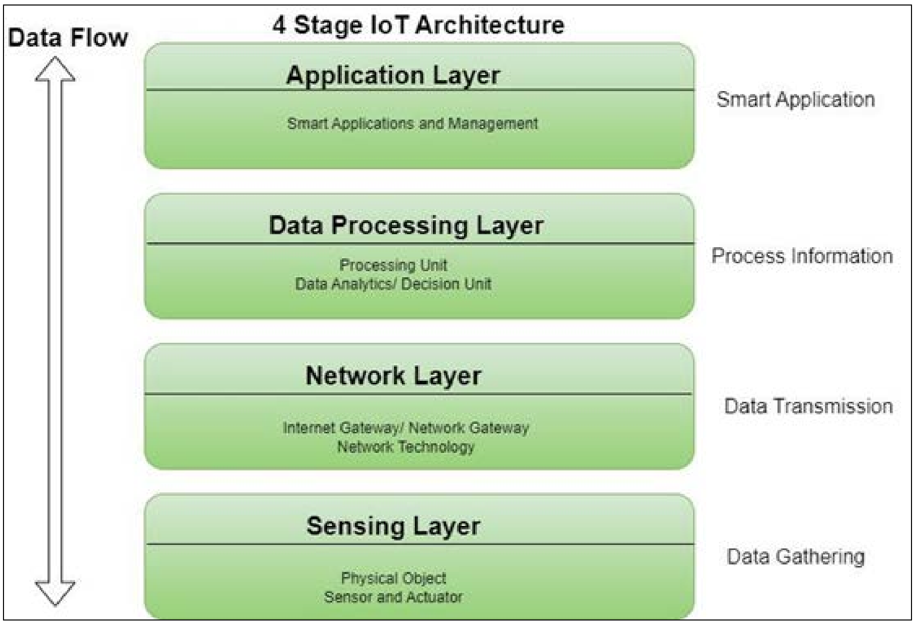


Fig 1: Layers in IOT

Limitations of IoT devices

The IoT devices are complex and heterogeneous in nature and have limited resources which makes them vulnerable to various security risks and threats. It is a serious matter of concern as the use of IoT devices is increasing day by day. With the huge growth cities, healthcare, transportation and other sectors will become smarter. There are more chances of expansion of IoT networks with the introduction of 5G network But as there are negative aspects of everything, It can open door to new types of cyber threats. So, the main concern is about security and privacy of IoT devices. The IoT devices lack strong security measures as they are resource constrained. They have less processing power and memory. They don't have robust design which is able to protect IoT devices from cyber-attack. Manufacturers of smart devices are paying very less attention to this aspect as it can increase their cost and there will be an additional overhead. So, the chances of intrusion into IoT devices are very high and it can cause a huge loss of important information. These devices are able to gather personal and other sensitive information of the users. This data can be intercepted by unknown hackers without the knowledge of the users and can be misused. An IOT device can become an entry point to get the access of a network. The increasing use of IOT devices can introduce new types of cyber-attacks including zero day attack as these devices are lacking implementation of high-level security measures. So, there is a need to provide security measures which will be able to protect IOT devices and can increase the trust of people in these devices.

Machine learning as a solution for cyber security of IOT Devices

An intrusion detection system can be implemented for IOT networks which will monitor the incoming traffic and will be able to detect the cyber-attacks efficiently. Machine learning plays an important role in the formation of intrusion detection system. There are three types of ML algorithms - supervised, unsupervised and reinforcement learning. The algorithms of machine learning can detect the patterns of network traffic and can help to detect the attacks. Different types of training models can be formed using different available databases which includes information regarding network data. These models can be applied to test the incoming traffic to detect any type of intrusion introduced in the network. The machine learning algorithms are of three types-supervised, unsupervised and reinforcement learning. These algorithms are employed in intrusion detection systems. The algorithms to detect cyber-attacks can be categorised as follows [5].

1. **Rule based algorithms:** These types of algorithms use prior explicit knowledge of attacks such as corresponding data distribution to create rule based systems and to perform detection. These algorithms are unable to handle noisy and incomplete data and it is very difficult to update them.
2. **Statistics based algorithms:** These types of algorithms build statistical model of intrusion patterns. They overcame the limitations of rule based systems and were able to handle noisy data but they are unable to handle large quantities of data.
3. **Machine learning algorithms:** These types of algorithms form training models of different types of attacks and use them to detect different types of anomalies in the network. The machine learning algorithms can be supervised, unsupervised or reinforced. They can handle large amount of data and noisy data as well. They can learn from experience. They can detect complex intrusions into the network by forming different complex models which are able to detect new types of cyber-attacks.

Traditional intrusion detection systems are not able to handle scalability issues and are unable to monitor the massive data amounts generated by excess of IoT devices [14]. The cyber security concerns are increasing in IOT devices with their increased use and it needs an immediate attention as very less work is done on this side. These devices are lacking the security and privacy measures. The rest of the paper is structured as follows. Different types of IDSs and available ML techniques for intrusion detection are explained in secion-4. The literature review is explained in section-5. Challenges and issues with the existing IDSs are explained in section-6. IoT security issues are described in section-7. Conclusion is presented in section 8.

Different types of IDS and Machine Learning Techniques employed

There are many types of intrusion detection systems available which makes use of ML techniques for detecting intrusions into the system. ML is a field of Artificial intelligence which can emulate human intelligence by learning. Different ML techniques are applied in different fields like pattern recognition, computer vision, finance, engineering etc. One of the important application of ML is in detection of cyber-attacks. The ML algorithm can be provided with the data of cyber-attacks to form the training model and this model can be used for detection of old and new cyber threats. They provides real time monitoring of network traffic.

This section gives an introduction of different types of IDS and different ML techniques used.

Types of IDS

Host Based IDS

This type of IDS system is installed on every host and it monitors each and every incoming and outgoing network traffic in the host. If any anomaly is detected then an alert is raised. It can detect configuration changes, file modification, system logs, and incorrect client server requests. It follows the rule of string or pattern matching for detection of a threat.

Network Based IDS: This type of IDS system is installed in the subnets and it monitors the incoming and outgoing traffic and in case if there is any anomaly is detected then an alert is raised. It can detect activities like port scanning, denial of service attacks, sudden increase in network traffic.

Protocol Based IDS: This type of IDS analyses the different protocols of the system. Here the servers consist of various types of agents on the servers which scrutinize different protocols. This IDS is installed on front end of the server and monitor the HTTP stream.

IDS can also be classified on the basis of their detection behaviour. It is of three types-signature based IDS, anomaly-based IDS, hybrid IDS.

Signature based IDS-This type of IDS is based on the database consists of signature of known attacks. It can detect only known cyber threats and unable to detect unknown cyber threats.

Anomaly based IDS

This type of IDS detect deviation from normal flow of network data. It can detect unknown type of cyber threat.

Hybrid IDS: It is combination of signature based and anomaly based IDS. It is more effective as it covers the strengths of both types.

Different ML techniques for intrusion detection-

Decision Tree: This algorithm is expressed as a recursive partition of the instance space [17]. It is a distributed tree with a basic node called root without any incoming edge. Other nodes have exactly one incoming edge. The nodes which as outgoing edges are called internal nodes or test node. The rest of the nodes are called leaves. Root node consists of the feature value on the basis of which branching is done. Decision nodes or internal nodes contains the different values of the root node feature on the basis of which branching is done. Each leaf is assigned to one class and it shows final prediction. It is suitable for both classification and regression. The Decision Tree is shown in the figure-2

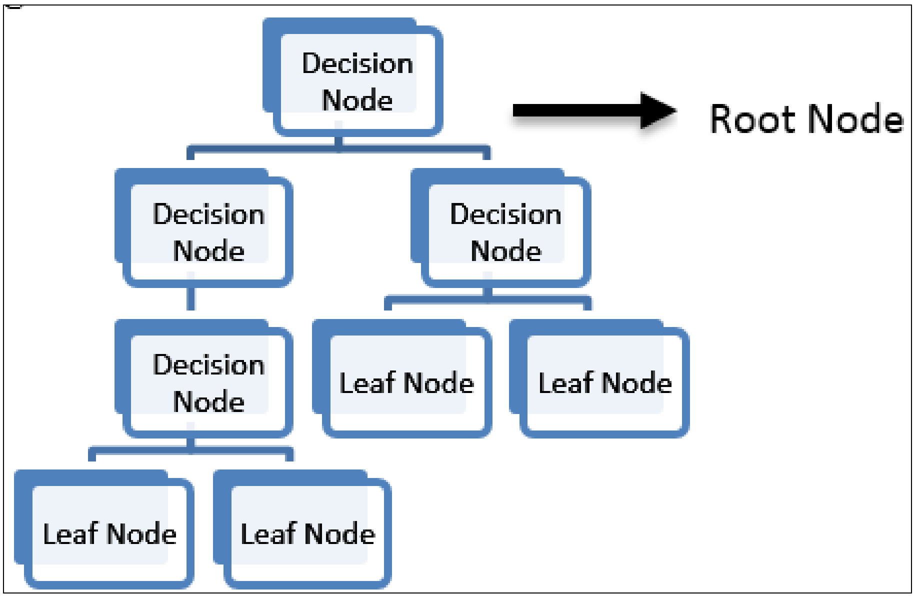
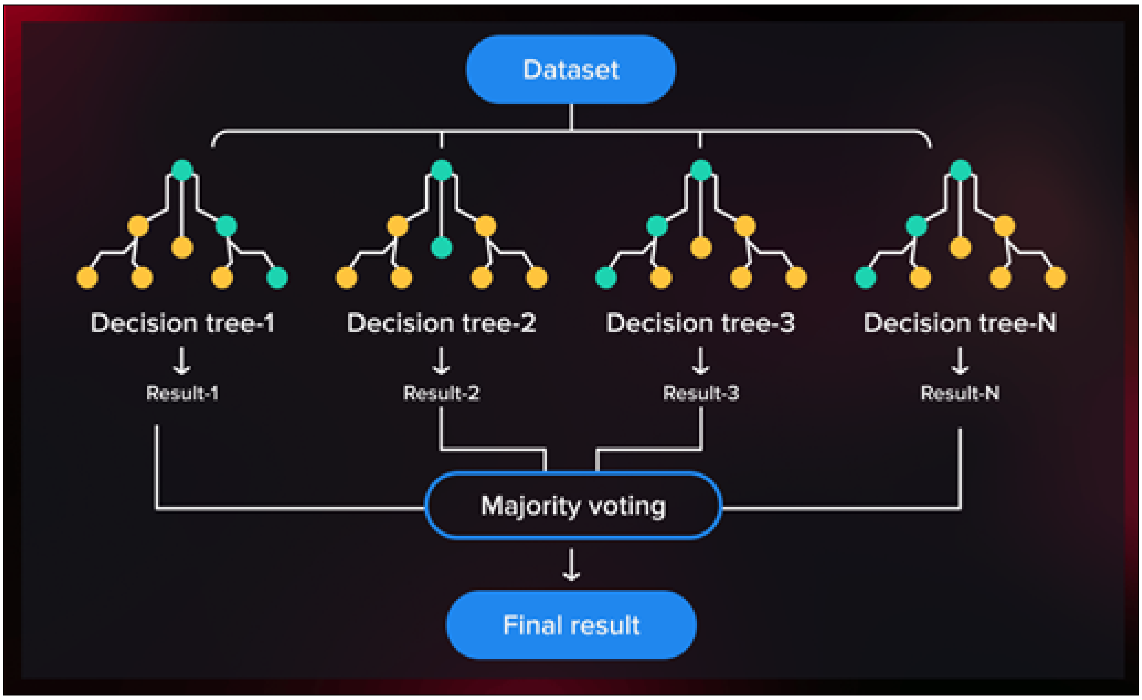


Fig 2: Decision Tree

Random Forest- This algorithm forms decision trees by random node splitting and resampling. We can say that it is an ensemble of various trees. The final classification result is voted by multiple trees [18]. The voting technique or averaging technique is applied to reach to the final result. This algorithm provides more accurate results. It is widely used in solving complicated problems. Random Forest algorithm has high accuracy, very less over fitting, handle large datasets, and handle missing values. The RF is shown in figure-3 [23].



Neural Network: It is a single layer perceptron where the weights are multiplied by a series of inputs before they reach the layer. Then, total is calculated by adding weighted input data together [19]. We can say it is a network of interconnected nodes with number of layers and it functions like human brain and able to reform itself from the feedback received. The diagram is shown in figure-4 [24].

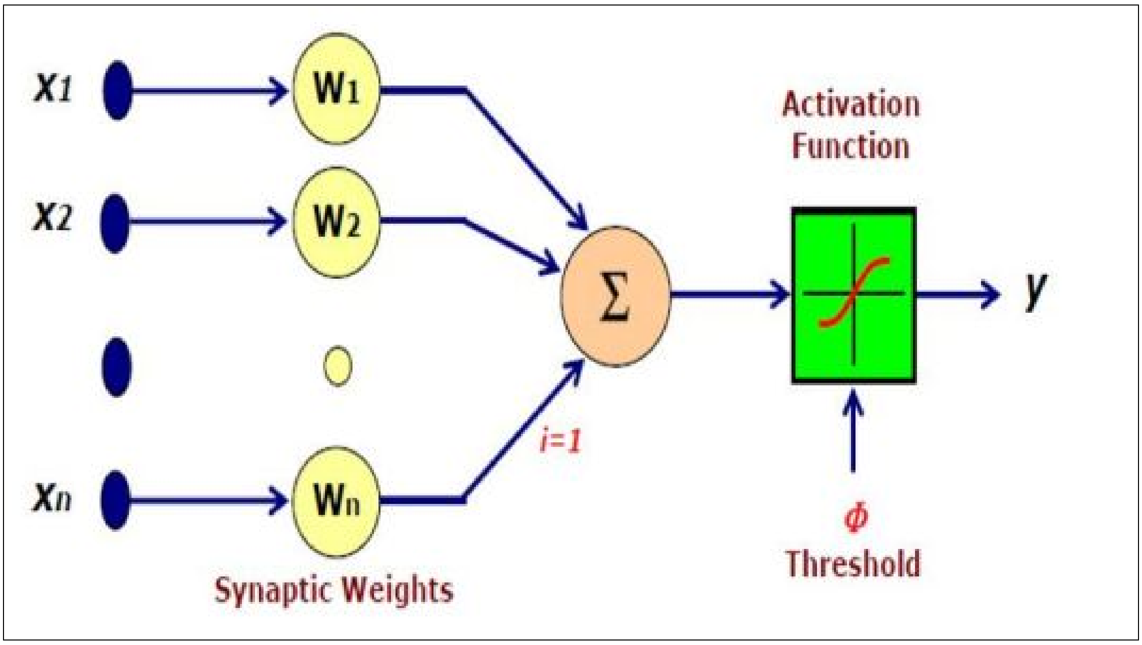


Fig 4: Neural Network

Support Vector Machine

This algorithm searches for optimal separating surface known as hyper-plane which is equidistant from the classes. It can be used for both linear and nonlinear data [20]. It is supervised machine learning algorithm and it tries to find linearly separable data points. It is shown in figure-5

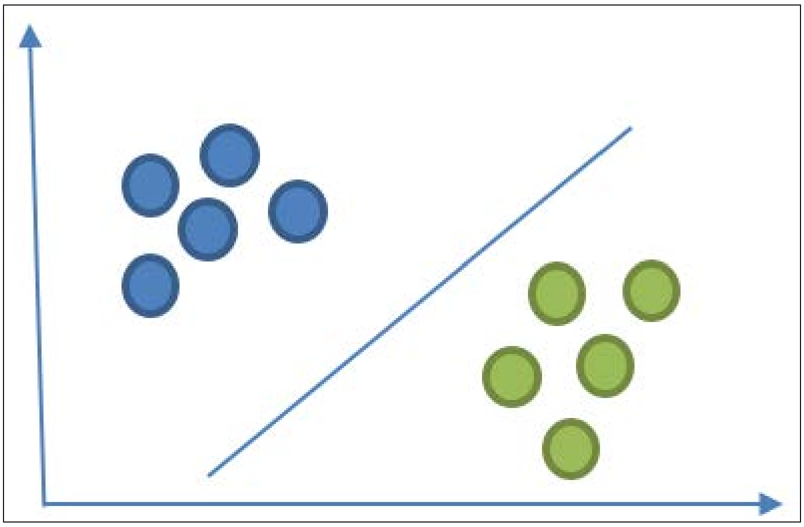


Figure-5 Support Vector Machine

K-Nearest Neighbour

It is simple and widely used ML technique for classification and regression problems. It is supervised and non-parametric learning classifier. It involves two types of learnings- Instance based and proximity based learnings. In instance based learning it memorizes the training data and makes predictions based on its similarity to the stored instances. In proximity based prediction it assumes that the similar data points (neighbours) exist in close proximity within the feature space. The algorithm involves the steps - Choosing value of K, calculate distance (Euclidean/Manhattan/Minkowsky), find K nearest neighbours and find vote or average. The voting technique is applied to classification problem and average technique is applied to regression problems. KNN method memorizes the whole dataset which makes training fast but prediction becomes slower as the size of data grows. It can handle multi class classification tasks.

Naïve Bayes: This algorithm is based on Bayes theorem. It is a probabilistic machine learning algorithm. It is a good method for high-dimensional data. Here the Bayes' theorem is applied with naïve assumption that the features are independent of each other. This assumption is not possible for all the cases but it still works well for some sort of problems.

$$P (C|X) = (P (X|C)-P(C))/ P(X)$$

Related Work

Many researchers have done efforts in this direction to build a robust cyber-attack detection system which is also known as intrusion detection system. This section presented the literature review of different researches done earlier in this direction. Here number of machine learning algorithms are employed to detect the anomalies in the network.

In [7], the authors emphasize on the need of Security solutions requirement for IOT environment. They have addressed the performance and class imbalance issues which makes it difficult to detect the intrusions accurately. They have proposed an automated network IDS system which used random forest algorithm on UNSW15 NB dataset and achieved accuracy of 90.17%. Data is collected from UNSW-NB 15 dataset. Four steps are performed- Pre-processing to check for missing values, exclusion of redundant packets, and samples of different classes are gathered. Symbolic data is encoded into numeric values. Then data balancing is performed in which training data is resampled by oversampling the minority classes and under sampling the majority classes. Feature selection is done by using three methods-filter method as it leveraged due to its superior speed. Data set is separated into train and test set. RF model is applied for training. Feature selection is done using Pearson's correlation coefficient. The features with significant correlation of 0.98 are retained while other are excluded. It improves accuracy by 3%. The training data consists of 80% of data and testing data consists of 20% of data. RF has high tolerance for outliers and noise. It is less prone to over fitting. The proposed method has achieved accuracy of 90.17% for unbalanced dataset and 98.77% for balanced dataset which surpasses SOTA (state of the art) approach by 7.34% and. 53% respectively. RF improves multiclass classification evaluation metric both for balanced and unbalanced datasets. RF is compared with other algorithms and it is concluded that RF outperforms other algorithms. Introduction of filter reduces the training time. So, combination of RF and Pearson's coefficient of correlation gives more efficient training process and gives robust performance. This work is for only six types of traffic-Exploits, Fizzers, Generic, Normal, Reconnaissance, DoS. So, this work can be extended for other types of traffic. A real time RF model can be prepared to detect intrusions in IOT networks.

In [8], the authors address the problem of noise and irrelevant features into traditional datasets. They proposed a new optimized feature selection method for accurate cyber-attack detection. This method targets features with high impact on target variables to optimize feature selection and reduction. The CICIDS-2017 data set is used and this method achieves 51% reduction in irrelevant features and increase in detection accuracy to 99.9%, 50% reduction in model computation time. Data is collected from CICIDS-2017 dataset. Four steps are performed- Data pre-processing, feature selection, model training and optimization, performance evaluation. In pre-processing redundant values are removed and missing values are filled by using binning method to obtain 70 features. All values like null, NaN, infinite are replaced by NaN. The min-max normalization technique is used for error mitigation. Only unique value features are considered for normalization. Integer and float values are unchanged and categorical values are encoded. Irrelevant features are reduced using chi-square test. It determines independence of two events by calculating chi-square co-relation. A modified chi-rev method is used for feature selection. The comparison is done for binary, multi class and all attack classification. The authors proposed chi-rev method for feature reduction. It is tuned with various algorithms but it outperforms with random forest classifier. It achieved an accuracy of 99.90% combined with almost 51% feature reduction and a 50% reduction in training time as compared with state-of-the-art methods.

In [9], the authors have used a lightweight deep neural network with principal component analysis (PCA), expansion and compression structure, inverse residual structure and channel shuffle operation. This model outperforms with low complexity, small model size and suitable for IOT traffic. Two datasets were used - First is UNSW NB15 dataset is used as it overcomes the shortcomings of KDD99 dataset. Raw network packets created by IXIA perfect storm tool for creating hybrid of real mode normal activities and synthetic contemporary attacks. Second is Bot-IoT dataset which is the latest NID data set for IOT which has normal IOT traffic and four attack scenarios-Dos, DDoS, Reconnaissance, theft. Data pre-processing is performed. PCA algorithm is employed for transforming original high dimensional traffic features into new low dimensional features through linear transformation. The features of training and testing dataset are reduced using PCA algorithms and a new k-dimensional feature space is formed. For binary classification detection, binary cross entropy is used as loss function otherwise use NID loss function.

In [10], the authors have stated that the security issues are increasing with the expansion of IoT in worldwide. Their study presents the model for enhancing security of IoT system using machine learning. A cyber-attack detection solution is developed for IoT devices using ML. The study used seven ML algorithms to identify most accurate classifier which can detect attack activities and patterns in networks connected to IoT. Proposed approach achieved 99.9% accuracy, 99.8% detection rate, 99.9% FI score and AUC score of 1. So, it achieved overall high execution speed and accuracy. Two datasets were used - First is UNSW NB15 dataset as it overcomes the shortcomings of KDD99 dataset. Raw network packets created by IXIA perfect storm tool for creating hybrid of real modes normal activities and synthetic contemporary attacks. Seven ML algorithms are applied in the datasets and it is found that RF, Boost, AdaBoost, and Ensembled RF-BPNN classifiers did the best overall. They achieved an accuracy of 99.9%, an AUC of 1, and an F1 score of 99.9%.

In [11], the authors stated that an Intrusion Detection System plays an important role in security and prevents unauthorized users to access network resources by analysing network patterns. They introduce a hybrid intrusion detection system based on Support Vector Machine and Grey Wolf optimization (GWO) algorithms. Support Vector Machine (SVM) is used to train and differentiate anomaly records from normal records and GWO is used to find kernel function, feature selection and to adjust optimal parameters for SVM to improve the classification. The datasets used were - NSL-KDD and TON_IoT datasets. The proposed IDS has been validated through python language on 2 datasets - NSL-KDD and TON_IoT. It outperforms in terms of detection accuracy, precision, recall and F-score. Proposed approach is independent of datasets and has acceptable performance on both datasets. The false positive rate of the proposed approach is 0.12 and 1.27 on NSL-KDD and TON_IoT datasets respectively.

In [12], the authors have stated that conventional IDS have poor detection capabilities and has high communication and device overhead. They introduced a teaching learning-based optimization IDS which ensure low overhead and effectively protect IoT networks. This method detects analysis attack, fuzzing attacks, shell code attacks, worms, denial of service, exploits, backdoor intrusion attack. Its performance is better than state of the art algorithms and it outperforms bat and genetic algorithm by 22.2% and 40%. This method has excellent accuracy and detection rate. The proposed approach can be enhanced by incorporating various encryption standards.

In [13], the authors emphasized the need to develop strong cyber-attack detection system for internet of healthcare things as IOHT devices are vulnerable to cyber-attacks due to lack of security procedure implementation in these devices. They presented an AI based cyber-attack detection system using deep neural network. The dataset used is ECU-IOHT. The proposed system achieved an accuracy of 99.85%. False positive rate is 0.01. This method has achieved higher detection rate as compared to existing methods. ECU-IOHT is a new dataset developed and reflected various cyber-attacks like ARP spoofing, DOS attacks, Nmap portscan, and Smurf attacks. There are two phases- Data preparation phase and DNN based attack detection phase. Five features are extracted from ECU-IOHT dataset and one hot encoding is used for encoding categorical features. The dataset is labelled as normal, ARP spoofing, DOS, Nmap and smurf attack. The dataset is split into training and testing (80% and 20%). The DNN is trained on the training dataset using multiclass classification. The trained model is tested using test dataset for predicting the attacks. Five features are selected from eleven features - type, source packets, destination packets, type of protocol, length. The extracted features are numerical and categorical. Categorical features are converted into numerical one using one hot encoding algorithm. In the proposed model input layer entails five neurons followed by two dense layers each of eight neurons with ReLU activation function and output layer with softmax activation function for categorization. Performance is evaluated using accuracy, precision, recall, FI score, true positive rate and false positive rate. Python is used as an implementation tool with libraries- matplotlib, NumPy, pandas, scikit learn, keras and tensor flow. The average precision, Recall and F1 Score values for the proposed system were 99.3%, 96.8 and 90.3%, respectively. This clearly underlines that the proposed system evidently outperforms the contemporary works.

In [25], the authors stated that detecting zero day attack is always a challenging task. They reviewed various AI based methods for zero day attack detection. They have reported both strengths and challenges. They have concluded that supervised learning methods give high false alarm rate in detecting zero day attack. Deep Learning based methods are time consuming and computationally expensive due to complex algorithms. DL model does not meet real world requirements.

Comparative Analysis: The given table shows the summary of the literature review in tabular form. It shows feature selection method, ML technique and accuracy achieved in table-1.

Table 1: Comparative Analysis of different Machine Learning techniques using different feature selection methods

Related Work	Feature Selection method	ML Technique used	Dataset Used	Accuracy
7	PCA(Principle Component Analysis	Random Forest	UNSW-15 NB	90.17%
8	Chi-Rev	Random Forest	CICIDS-2017	99.9%
9	PCA(Principle Component Analysis	Deep Learning	UNSW NB15, Bot-IoT	86.11%
10	PCA(Principle Component Analysis	Ensembled RF with Neural Network	UNSW-15 NB	99.2
11	Grey Wolf Optimization	SVM	NSL-KDD and TON_IoT	98%
12	Teaching Learning Based	TLBOIDS	UNSW-15 NB	99.8%
13	Only Five features selected	Deep Neural Network	ECU-IOHT	99.855

This table is presenting the comparative analysis of different feature selection and ML techniques. It can be seen from the table that if we combine chi-rev method for feature selection and Random Forest Machine Learning technique for classification, we got the highest accuracy 99.9%.

Security Issues with IoT systems

The use of IoT is expanding and it is making the life of people easier. It covers almost every sector. But security is one of the biggest challenges of this technology. It is a serious issue as it can be the major reason for downfall of IoT technology in future. There is an immediate attention required to safeguard the connected devices. The complexity of the IoT systems has been increased with the expansion and spreading of IoT networks [15]. There exists issues of scalability and heterogeneity. Apart from this, IoT devices are lacking implementation of sound security standards into the devices due to the limitation of the technology, limited power and high cost of implementing it. The present security measures cannot fit into this new technology. Some of the challenges available with IOT devices are listed below [21].

- IOT devices have outdated firmware.
- IOT devices are resource constrained.
- Use of weak and default credentials to log in
- Lack of encryption
- Malware and Ransom ware can affect IoT systems easily.

A review of different case studies is given here which proves that the security of IOT devices is a critical issue which must be addressed.

Case Study 1

Electric Vehicle Charger from Charge Point Inc.

It is related with device software failure. The vulnerabilities are lying with firmware. The author in [22] demonstrated how electric vehicle charger's password authentication phase is bypassed by changing BEQ assembly instruction to BNE in debug mode. So, attacker can input wrong password and can log in into device's system code. It is also shown that attacker can modify or create any file.

Case Study 2

Unauthorised access

An attacker can gain access to IOT system by exploiting hardware/software vulnerabilities as well as he can also do illegal login attempts. Here a case study of Tesla Model [23] is presented which shows such type of attack. Tesla service stations and charging stations are equipped with Wi-Fi SSID. The credentials are stored in Tesla's web browser which is used for auto connecting. The hackers redirected traffic from web browser to their domain by faking the SSID. There are number of bugs in Tesla's browser. The attackers used this weak point to read and write arbitrary memory addresses and to execute arbitrary code for gaining access to shell. After gaining access, they disabled the Linux module. They got the privileged access to Gateway shell and misused it by writing customised firmware.

Challenges and issues with the existing IDS systems in Iot Networks

The existing intrusion detection systems are lacking following points-

1. An IDS system is designed using different ML approaches. Many of the IDSs are implemented using traditional old datasets. In [7], the authors have used a very old dataset UNSW-NB-15 to improve the efficiency. So the old IDS systems are not effective to detect new types of cyber-attacks.
2. Apart from this the number of records for training dataset taken was 175341 and for testing model 82332 [7]. Bot-IoT dataset is also used where no. of training dataset is 364562 and testing dataset is 243043.The result will vary if we increase number of training and testing records and will affect the efficiency. The applied techniques can be used on new datasets with increased number of records and it will lead to a perfect IDS. The number of training records should be increased and should be updated with the data of new cyber-attacks.
3. These systems are not able to detect new cyber-attacks in IOT environment. The dataset used to create training set should have data of new cyber threats like zero day threat. The available traditional datasets consist of very old data which is a major cause of failure of available intrusion detection systems in present scenario of cyber world. In [8, 9, 11], the authors have used the dataset CICIDS2017, NSL-KDD which are not suitable to detect new types of cyber threats in IoT as the IoT network has become complicated and the existing datasets of IDSs does not consider it. There is a need of accurate dataset which will be able to address new types of threats.
4. Zero-day attack in IOT network is an emerging threat which needs an immediate attention as use of IOT devices is increasing exponentially. It is not easy to recover the systems from zero-day attacks as their type is unknown and there is no solution available. These attacks are planned by taking into account any vulnerability present in the system and there is no recovery time.

Conclusion

Cyber-attacks are increasing to a large extent due to increase in the use of Internet and IoT devices. There is a need to identify and stop these attacks. The solution to this problem is Intrusion Detection System (IDS). To detect new and complicated cyber-attacks, the weak points of existing IDS need to be addressed and the existing IDS should be upgraded so that they will be able to tackle new types of attacks in new technology. In this paper first of all the overview of the IoT technology is presented. Then limitations are addressed followed by the solution to cyber security risks using Machine Learning. Then related work is presented in the form of literature review. The challenges and issues of the available IDS are explored. It is concluded that there is a need to develop new intrusion detection systems which will be capable to detect new types of cyber-attacks for IoT network. Moreover, there is an urgent need to detect zero-day attacks in IOT network because they can cause high loss as they are unknown and there are no previous signature is available related to these attacks. The datasets taken for forming different models should be upgraded to address the new types of threats. It tis concluded that the existing ML techniques can be applied to new datasets of IoT along with improved feature selection and classification techniques. Moreover, feature selection methods should be chosen carefully after analysing the problem. Feature selection is very much helpful in reducing dimensions, to improve the speed and increase the comprehensibility of the result. The most commonly used techniques of feature selection are filter method, wrapper method and hybrid methods. Hybrid methods of feature selection can be proven the best methods by complementing limitations of both the methods. In addition, different ML techniques can also be combined to get higher and perfect accuracy of the detection result specifically in case of zero-day attack where unsupervised machine learning can be proven more effective as they do not require any training data.

References-

1. Kamal R. Internet of Things: Architecture and design principles.
2. Wang J, Jiang C, Zhang H, Ren Y, Chen K-C, Hanzo L. Thirty years of machine learning: The road to Pareto-optimal wireless networks. IEEE Commun Surv Tutorials. 2020;22(3):1472-1514.
3. Wang Z, Liu Y, Ma Z, Liu X, Ma J. LiPSG: Lightweight privacy-preserving Q-learning-based energy management for the IoT-enabled smart grid. IEEE Internet Things J. 2020;7(5):4480-4489.
4. Cai Z, Shi T. Distributed query processing in the edge-assisted IoT data monitoring system. IEEE Internet Things J. 2021;8(16):12679-12693.
5. Gümü¸sba¸ D. A comprehensive survey of databases and deep learning methods for cybersecurity and intrusion detection systems. IEEE Syst J. 2021;15(2):1950-1962.
6. Rivera, Goasduff L. Gartner says a thirty-fold increase in Internet-connected physical devices by 2020 will significantly alter how the supply chain operates. Stamford, CT: Gartner; 2020.
7. Maghrabi LA. Automated network intrusion detection for Internet of Things: Security enhancements. Jeddah, Saudi Arabia: Department of Software Engineering, College of Engineering, University of Business and Technology.
8. Tripathi G, Singh VK, Sharma V, Vinod Bhai MV. Weighted feature selection for machine learning based accurate intrusion detection in communication networks.
9. Zhao R, Gui G, Xue Z, Yin J, Ohtsuki T, Adebisi B, Gacanin H. A novel intrusion detection method based on lightweight neural network for Internet of Things.
10. El-Sofany H, El-Seoud SA, Karam OH, Bouallegue B. Using machine learning algorithms to enhance IoT system security.
11. Ghasemi H, Babaie S. A new intrusion detection system based on SVM-GWO algorithms for Internet of Things.
12. Kaushik A, Raweshidy HAI. A novel intrusion detection system for Internet of Things devices and data.
13. Vijayakumar KP, Pradeep K, Balasundaram A, Prusty MR. Enhanced cyber attack detection process for Internet of Health Things (IoHT) devices using deep neural network.
14. Alhasavi Y, Alghamdi S. Federated learning for decentralized DDoS attack detection in IoT networks. IEEE Access. 2024;12:1-14.
15. Jan S. Applications and challenges faced by Internet of Things-a survey. Int J Eng Trends Appl. 2016;2393-9516.
16. Nasteski V. An overview of the supervised machine learning methods. Available from: https://www.geeksforgeeks.org/architecture-of-internet-of-things-iot.
17. Dhanabal L, Shantharajah SP. A study on NSL-KDD dataset for intrusion detection system based on classification algorithms. Int J Adv Res Comput Commun Eng. 2015;4(6):446-452.
18. Muruganandam S, Joshi R, Suresh P, Balakrishna N, Hari Kishore K, Manikanthan SV. A deep learning-based feed-forward artificial neural network to predict the K-barriers for intrusion detection using a wireless sensor network.
19. Bhavsar H, Panchal MH. A review on support vector machine for data classification.
20. Ambimat. Challenges to IoT security. Available from: https://ambimat.com/challenges-to-iot-security-1/.
21. Kaspersky Lab security services. ChargePoint Home security research. 2018. Available from: https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/13084354/ChargePont-Home-Security-research_final.pdf.
22. Nie S, Liu L, Du Y. Free-fall: hacking Tesla from wireless to CAN bus. Black Hat USA. 2017;1-16.
23. Serokell. Random forest classification. Available from: https://serokell.io/blog/random-forest-classification.
24. Sathyabama University. Course material on IoT. Available from: https://sist.sathyabama.ac.in/sist_coursematerial/uploads/SEC1609.pdf.
25. Por LY, Dai Z, Leem SJ, Chen Y, Yang J, Binbeshr F, Phan KY, Ku CS. A systematic literature review on AI-based methods and challenges in detecting zero-day attacks. IEEE Access. 2024;12:144150-144163.