ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25nov173

A Review on Machine Learning and Deep Learning Based Modern Intrusion Detection Techniques for Mobile AdHoc Networks

P. Sreenivasulu[1]; Dr. M. Ussenaiah[2]

[1](Research Scholar); [2](Associate Professor)

[1;2]Department of Computer Science, Vikrama Simhapuri University Nellore, India

Publication Date: 2025/11/19

Abstract: Intrusion detection is very essential to secure any network communication. Different intrusion detection methods for MANET using Machine Learning techniques are reviewed. The main goal is to present the progress developments in intrusion detection techniques from traditional methods to the recent advanced Machine learning techniques. Usage of Machine leaning approach in intrusion detection is significantly increasing in recent years. Review on various Machine leaning techniques, types of intrusion, role of classification methods and data sets used are presented in this paper.

Keywords: *IDS, Attacks, Machine Learning, Classfication, Feature Selection, Data Sets.*

How to Cite: P. Sreenivasulu; Dr. M. Ussenaiah (2025) A Review on Machine Learning and Deep Learning Based Modern Intrusion Detection Techniques for Mobile AdHoc Networks. *International Journal of Innovative Science and Research Technology*, 10(11), 809-813. https://doi.org/10.38124/ijisrt/25nov173

I. INTRODUCTION

The tremendous growth of Internet and usage of applications accessed on the network has increased vastly [21]. However with the increase of applications accessed over the network, there is a possibility of cyber attacks which also increased these days. In many cases these types of attacks are new and not addressed earlier. Intrusion Detection System (IDS) analyzes the traffic, activities in the machine and identifies the malicious attacks. An intrusion detection system can be generally classified as anomaly based intrusion detection system and misuse based intrusion detection system. Misuse based IDS looks for the known attacks in the current traffic. If an attack is identified then an alarm is triggered. Whereas anomaly based IDS analyzes the traffic to detect any kind of deviations. If there is an abnormality then it reports it as an intrusion event. For a successful detection of new attacks we need huge amounts of data to build a model which should classify which is a normal activity and which is abnormal activity. This raised the importance of classification algorithms. For ages traditional classification methods are used to build IDS. However with the rapid growth of social media platforms and IOT devices which generates large volumes of data the focus of the researchers turned to Supervised Machine Learning algorithms.

II. BACKGROUND

➢ *Intrusion Detection [18]*

Intrusion Detection is an activity that determines whether a process or user is attempting something unexpected. It works on the basis of examining activity on a specific machine or network and deciding whether the activity is normal or suspicious. It can either compare current activity to known attack simply raise an alarm condition when specific measurements exceed preset values.

An Intrusion Detection System (IDS)[20] is a system that monitors network traffic for suspicious activity and issues alerts when such activity is discovered. It is a software application that scans a network or a system for harmful activity or policy breaching. The major goals of any IDS are.

- To monitor hosts and Networks.
- Analyse the behaviour of Networks.
- Generating the alerts.
- Responding suspicious behaviours.

Intrusion Detection Systems [18] can be categorized in to the following groups.

- *Signature Based IDS*

In signature-based IDS[20] the IDS monitor the network and compare actual behavior with known suspicious

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25nov173

patterns that are maintained in a database of attack signatures. If there is a similarity then an alert is generated. Most of the intrusion detection systems of this type are based on classification methods. But these systems are not efficient when the attack type is not known. The reason is the database does not contain the signature of new attack. This kind of systems is efficient for known type of attacks.

- *Anomaly Based IDS*

In Anomaly based IDS [20] the system behavior is compared with the base line that defines the normal state of the system which may include protocols, traffic load and packet size. Deviation from the base line indicates that there is an abnormal activity and raises an alert. Some times normal behavior can be misclassified as an attack.

- *Hybrid IDS*

Hybrid IDS [20] makes use of both signature based and anomaly based to gain the advantages of both the methods. This method tries to increase detection rates of known attacks and decreases false positive rates of new attacks.

➢ *Role of Classification in IDS:*

Audit records and log files plays an important role in intrusion detection. Intrusion Detection can be thought of as classifying these records in to normal or abnormal or particular type of intrusion.

Classification methods for intrusion detection works based on training sets of given data. A better classification algorithm builds the classifier within a less time and gives accurate results. We should consider the time and accuracy as two essential factors in determining the intrusion activities by using a classification algorithm.

To analyze the data and to classify we can use supervised machine learning method or unsupervised machine learning methods. These two types can be applied for the purpose of automatic detection of intrusions. The main focus of this survey is the application of supervised machine learning techniques for IDS. In supervised learning or classification, there must be labeled data that can be used to train a model for detection purposes. The process of classification can be summarized in the following steps: -

- Preprocess the data.
- Obtain training data from the processed data.
- Apply a classification algorithm.

➢ *Role of Machine Learning Algorithms in Intrusion Detection – Related Work.*

Huge data set has to be analyzed to produce accurate results [22]. Machine learning algorithms play a vital role in analyzing large amounts of data for improved accuracy. In proposing solutions for intrusion detection, recent research work was done in the application of supervised machine learning techniques for developing intrusion detection systems.

So many researchers worked on classification algorithms for the use of intrusion detection. Few contributions done in this area are summarized below.

Vinay kumar R,Soman KP etl[19] compared classical machine learning classifiers with that of Convolution Neural Networks (CNN). The authors used KDDcup 99 data set for their experiments. They implemented CNN as the first layer with a RNN. The Authors claimed deep learning based methods are suitable for modeling network traffic based on TCP/IP.

Pooyan Azizi doost, Sadegh Sarhani etl [1] tried combining CNN and RF to get better results. In the proposed model CNN is used for feature selection and Random Forest (RF) is used for the classification purpose. This Hybrid method resulted, 97% accuracy. The authors claimed that the execution time is less when compared to c4.5, NBTree and NBFS. The proposed Hybrid method also achieved 99.3% precision.

Malhotra, H., & Sharma, P [2] investigated the efficacy of ten classification algorithms on the NSL-KDD dataset, employing a different feature selection method than that of [12]. The method of feature selection relies on utilizing attribute evaluators and filtering techniques. The authors executed the subsequent algorithms such as Naive Bayes, Bayes-Net, Logistics, Random tree, Random forest, J48, Bagging, OneR, PART, ZERO. Among all these algorithms random forest algorithm achieved an accuracy of 99.9% and maintains a minimal false alarm rate of 0.001. The classifier with the second-best performance is bagging, achieving an accuracy of 99.8%, with the PART algorithm showing similar results.

Belavagi, M. C., & Muniyal, B [3] focused on the NSL-KDD data set and utilized a hold-out testing method without employing any feature selection technique. The four classification methods being evaluated are random forest, SVM, logistic regression, and Gaussian mixture model. The random forest demonstrated the highest algorithm performance with an accuracy reported at 99%. The classifier with the second-best performance is logistic regression, which has an accuracy of 84%.

Taher, K. A., Jisan, B. M. Y., & Rahman, M.[4] conducted experiments on the effectiveness of artificial neural networks (ANN) and SVM using a sample from the NSL-KDD data set. The sample constitutes 20% of the entire data collection. Two techniques for feature selection are employed: methods based on correlation and chi-square. The first produced a choice of 17 features, while the second produced 35 features. Following feature selection, the data is input into ANN and SVM classifiers. The outcomes from correlation-based feature selection and ANN showed the best performance with an accuracy of 94.0%

El Mourabit, Y., Bouirden, A., Toumanari, A., & Moussaid, N. E.[5] utilized the KDD'99 dataset. The categories include: normal, Prob, DoS, U2R, and R2L. The feature selection algorithms CFSSubSet Eval and Best First

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25nov173

are utilized alongside four techniques: one unsupervised technique, k-means, and three supervised techniques, SVM, Naïve Bayes, and random forest. The three supervised approaches exceeded the performance of the unsupervised method and focus on the eight top features. The highest reported accuracy belongs to the random forest, achieving an accuracy of 99.0%

Li, Y., Xia, J., Zhang, S., Yan, J., Ai, X., & Dai, K. [6] initially, the ant colony algorithm was used to choose an adequate representative subset from the original dataset, which included 550 samples. The authors then used a novel feature reduction technique called the Gradual Feature Removal (GFR) method to lower the dimensionality of the feature space to 19 features. The reduced features are then used with Support Vector Machine for classification. The stated accuracy is 98. 67% before feature selection, and the results are 98. 62% after feature selection, showing no appreciable improvement in accuracy.

Shah, B., & Trivedi, B. H. [7], studied the impact of decreasing the dimensionality of KDD'99 on overall classification results. They analyzed the feature selection algorithm; information gain derived from entropy (IG), and merged the resulting feature vector (22 features) with Back Propagation Neural Network (BBNN). The findings indicated that accuracy did not change following the feature reduction, remaining at 91%. Another dataset that garnered interest in this field is the CICIDS2017 dataset.

Arif Yulianto , Parman Sukarno and Novian Anggis Suwastika[8] addressed the problem of imbalance of training data. The authors used Synthetic Minority Oversampling Technique (SMOTE) to overcome the imbalance problem. For feature selection they used Principal Component Analysis (PCA) and Ensemble Feature Selection (EFS). They worked on only one data set CICIDS2017 to carry out experiments. According to the authors Adaboost based intrusion detection system performance can be improved by using SMOTE, PCA and Ensemble Feature Selection (EFS).

Abdul hammed Miad Faezipour Hassan Musafer Abdelshakour Abuzneid [9] aimed in reducing high dimensional features of the CICIDS2017 data set to low dimensional features using principal component analysis(PCA). High dimensional features may take longer classification times than lower dimension features. The authors used these low dimensional features to test the performance of various classifiers.

Pelletier, Z., & Abualkibash, M. [10], examined the challenge of categorizing attacks employing random forest and ANN methods utilizing the CICIDS2017 dataset. They utilized a package called Boruta for selecting features and identified the 10 most significant features. The classifiers are then provided with the feature set. They reported an average precision of 96% with ANN and 96.4% with random forest.

Hammad, M., El-medany, W., & Ismail, Y. [11] utilized a different dataset that gained interest in this field, which is the UNSW-NB15. The researchers analyzed this data set by employing an approach that incorporates k-means clustering, CFS feature selection, and four distinct methods: SVM, RF, J48, and Zero. The suggested method has proven successful in enhancing the performance of most classifiers. The highest documented accuracy is achieved with J48, showing an accuracy of 96.7% through the use of 10-fold cross validation.

Faker, O., & Dogdu, E[12] worked on integrating big data and deep learning techniques to improve the performance of intrusion detection techniques. To classify the network traffic data sets a Deep Feed forward Neural Network and two ensemble techniques, Random Forest and Gradient Boosting Tree (GBT) are used. The experimental results are carried on the data sets UNSW NB15 and CICIDS2017.

Sara AI-Emadi[13] focused on developing an intelligent detection system to detect different attacks. The authors have chosen deep learning techniques for Network Intrusion Detection. The experiments are carried using NSL-KDD data set. The proposed system uses two Deep Neutral Networks architectures, Convolution Neural Networks and Recurrent Neural Networks. According to the authors CNN outperformed RNN in terms of F1 score, precision and accuracy, whereas recall metric is better in RNN.

The experimental results are carried only on NSL-KDD data set, also for fewer attacks. The authors also specified that, limited computational resources caused longer training time.

Pramita Sree Muhuri[14] worked on deep learning techniques and developed a new method. The authors combined a genetic algorithm for optimal feature selection and long short –term memory (LSTM) with a recurrent neural Network. The authors used NSL-KDD datasets which were mapped into 122 features for their experimental setup. The experimental results showed high detection rate while using only five neurons in the hidden layer. An accuracy of 96.51% for KDD Test data. They applied Genetic Algorithm on the same data set and obtained optimal set of 99 features with an increased accuracy of 99.91%.In this case 40 neurons are used in the hidden layer. However in case of a binary classification Random Forest technique is the better one compared to the LSTM-RNN method. Also by using LSTM-RNN with GA is more complex when compared to existing RF algorithm.

Asmaa Halbouni[15] developed a hybrid model for intrusion detection that combines convolution neural networks (CNN) and Long-Short Term memory (LSTM). CNN can extract the spatial features whereas LSTM can extract temporal features. To increase the performance of the model they added batch normalization and dropout layers. The experiments are carried out on three data sets such as CIC-IDS 2017, UNSW-NB15, and WSN-DS. The data sets are divided into two classes for binary classification such as benign and attack. Initially the authors compared the performance of the data sets independently i.e, CNN only, LSTM only and then the combination of LSTM-CNN

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25nov173

separately and CNN-LSTM separately. Among these models CNN-LSTM hybrid model produced better accuracy and detection rate than other models.

➢ *Data Sets used in Building Models for Intrusion Detection*

The collection of data is not an easy task, and hence, several benchmark data sets exist such as KDD'99 and NSL-KDD, and UNSW-Nb15 and CICIDS2017[23].

- NSL-KDD dataset: in this dataset, the application of feature selection has consistently enhanced classification performance. Additionally, the random forest algorithm performs exceptionally well on this dataset and shows impressive results with various validation techniques. ANN appears to be doing well on this dataset; however, it was evaluated using just a 20% sample of the data.

- KDD'99 dataset: in this dataset, the use of feature selection has not proven effective in enhancing classification performance. Various validation techniques are utilized on the data set, and no definitive conclusion can be made regarding the top-performing classification algorithm; however RF and SVM exhibit strong performance on this dataset.

- CICIDS2017 dataset: definite conclusions regarding the impact of feature selection cannot be established; nonetheless, it proved advantageous in enhancing classification performance in a single instance. The dataset presents a notable imbalance issue, which can be addressed through sampling methods. QDA is the top classification algorithm.

- UNSW-NB15 dataset: in this dataset, whenever feature selection is applied, it has proven to enhance classification performance. DNN is a widely used and effective method for this dataset.
- Apart from the above data sets, other data sets are also available with different features and to identify different attacks [16].

- UAVIDS-2025 dataset: This dataset is published in May 2025 especially designed for UAV networks. Useful for experimenting with Sybil, blackhole, wormhole and flooding attacks in mobile adhoc networks.

Table 1 The Following Table Shows the Experimental Results of Various Algorithms Carried Out by Various Authors. We Considered Only Accuracy.

Ref	DataSet	Algorithms	Best Method	Accuracy
[1]	NSL KDD	C4.5,NBTree,NBFS,Hybrid	Proposed Hybrid Method	97%
[2]	NSL-KDD	Naive Bayes, Bayes-Net, Logistics, Random forest, J48, Bagging, OneR, PART, ZERO	Random forest	99.9%
[3]	NSL-KDD	SVM, GMM, Random forest, logistic regression	Random forest	99%
[4]	NSL-KDD	Artificial neural networks (ANN), SVM	ANN	94.0%
[5]	KDD'99	k-means, random forest, naïve bayes, SVM	Random forest	99.0%
[6]	KDD'99	SVM	SVM	98.7%
[7]	KDD'99	BBNN	BBNN	91.0%
[8]	CICIDS2017	Adaboost	Adaboost	81.83%
[9]	UNSW-NB15	SVM, J48, RF, Zero	J48	96.7%
[10]	CICIDS2017	LDA, QDA, BN, RF	QDA	98.8%
[11]	CICIDS2017	ANN, RF	RF	96.4
[12]	UNSW-NB15	DNN, RF	DNN	97.0%
[13]	NSL-KDD	CNN,RNN-LSTM,RNN-GRU	CNN	97.01%
[14]	KDD	LSTM-RNN,RF,SVM	RF	99.99%
[15]	CIC-IDS 2017	CNN-LSTM,KNN,REP TREE,MLP	CNN-LSTM	99.64%

III. CONCLUSIONS

With the growth of internet, social media and mobile applications, cyber crimes also increased.

Use of Intrusion detection systems is the initial step in detecting and reporting such issues. Most of the intrusion detection systems are based on the classification algorithms. In this paper we reviewed some of such machine learning classification algorithms along with the data sets used in the process of building the classifiers. Most of the experiments are carried out majorly NSL-KDD data set and other on fewer. It is observed that in many cases Random Forest (RF) showed better accuracy when compared to other methods.

REFERENCES

[1]. Ali Basem, Edris Khezri, Sadegh Sarhani Moghadam Pooyan Azizi doost & Mohammad Trik-(2025)"A new intrusion detection method using ensemble classification and feature selection"-ScientificReports-| https://doi.org/10.1038/s41598-025-98604-w

[2]. Malhotra, H., & Sharma, P. (2019). Intrusion Detection using Machine Learning and Feature Selection. International Journal of Computer Network & Information Security, 11(4).

[3]. Belavagi, M. C., & Muniyal, B. (2016). Performance evaluation of supervised machine learning algorithms for intrusion detection. Procedia Computer Science, 89, 117-123.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25nov173

[4]. Taher, K. A., Jisan, B. M. Y., & Rahman, M. M. (2019, January). Network intrusion detection using supervised machine learning technique with feature selection. In 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST) (pp. 643-646). IEEE.
[5]. El Mourabit, Y., Bouirden, A., Toumanari, A., & Moussaid, N. E. (2015). Intrusion detection techniques in wireless sensor network using data mining algorithms: comparative evaluation based on attacks detection. International Journal of Advanced Computer Science and Applications, 6(9), 164-172.
[6]. 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 systems with applications, 39(1), 424-430.
[7]. Shah, B., & Trivedi, B. H. (2015, February). Reducing features of KDD CUP 1999 dataset for anomaly detection using back propagation neural network. In 2015 Fifth International Conference on Advanced Computing & Communication Technologies (pp. 247-251). IEEE.
[8]. Yulianto, A., Sukarno, P., & Suwastika, N. A. (2019, March). Improving Adaboost-based intrusion detection system (IDS) performance on CIC IDS 2017 dataset. In Journal of Physics: Conference Series (Vol. 1192, No. 1, p. 012018). IOP Publishing.
[9]. Abdulhammed, R., Faezipour, M., Musafer, H., & Abuzneid, A. (2019, June). Efficient network intrusion detection using pca-based dimensionality reduction of features. In 2019 International Symposium on Networks, Computers and Communications (ISNCC) (pp. 1-6). IEEE.
[10]. Pelletier, Z., & Abualkibash, M. (2020). Evaluating the CIC IDS-2017 Dataset Using Machine Learning Methods and Creating Multiple Predictive Models in the Statistical Computing Language R. Science, 5(2), 187-191.
[11]. Hammad, M., El-medany, W., & Ismail, Y. (2020, December). Intrusion Detection System using Feature Selection With Clustering and Classification Machine Learning Algorithms on the UNSW-NB15 dataset. In 2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies (3ICT) (pp. 1-6). IEEE.
[12]. Faker, O., & Dogdu, E. (2019, April). Intrusion detection using big data and deep learning techniques. In Proceedings of the 2019 ACM Southeast Conference (pp. 86-93).
[13]. Sara A1-Emadi, Aisha A1-Mohannadi,Felwa A1-Senaid(2020).Using Deep Learning Techniques for Network Intrusion Detection-IEEE.
[14]. Pratima Sree Muhuri,Prosenjit Chatterjee,Xiaohong,Kaushik Roy and Albert Esterline(2020,March) .Using a Long Short Memory Recurrent Neural Network to Classify Network Attacks.
[15]. Asmaa Halbouni,Teddy Surya Gunawan,et.al(2022 August). CNN-LSTM:Hybrid Deep Neural Netork for Network Intrusion Detection System.-IEE Access,volume 10,2022.
[16]. Daniela Pinto, Ivone Amorim , Eva Maia, Isabel Praça (2025,May) A review on intrusion detection datasets: tools, processes, and features-Volume 262, May 2025, 111177-computer networks, Elsevier.
[17]. Emad E. Abdallah*, Wafa' Eleisah, Ahmed Fawzi Otoom(2022) Procedia Computer Science 201 (2022) 205–212.
[18]. https://www.ibm.com/think/topics/intrusion-detection-system.
[19]. VinayKumar R,Soman KP and Prabaharan Poornachandran."Applying Convolutional Neural Network for Network Intrusion Detection" .IEEE Xplore December 2017.
[20]. https://securityjournaluk.com/intrusion-detection/.
[21]. Julian Jang-Jaccard, Surya Nepal-" A survey of emerging threats in cybersecurity". Journal of Computer and System Sciences 80 (2014) 973–993.
[22]. Md. Alamin Talukder, Md. Manowarul Islam,etal-"Machine learning-based network intrusion detection for big and imbalanced data using oversampling, stacking feature embedding and feature extraction."-Talukder et al. Journal of Big Data-Journal of Big Data(2024).11-33.
[23]. Ankit Thakkar, Ritika Lohiya – "A Review of the Advancement in Intrusion Detection Datasets".Science Direct- Procedia Computer Science 167 (2020) 636–645.