Machine Learning Based Attack Detection in Internet of Things Network

Md. Rumman Rafi, *Department of Computer Science, Southern University and A&M College, Baton Rouge, Louisiana, USA* mdrumman.rafi@sus.edu

Mohammad Abdus Salam, *Department of Computer Science, Southern University and A&M College, Baton Rouge, Louisiana, USA* md.salam@sus.edu

Osman Kandara, *Department of Computer Science, Southern University and A&M College, Baton Rouge, Louisiana, USA* osman.kandara@sus.edu

***Abstract*- In recent years, the Internet of Things (IoT) has grown up rapidly and tremendously. This growth has brought big and special problems. Two of the urgent topics of problems are security and privacy for IoT devices. Those devices are creating and gathering all data in their connections. For the security of IoT, detection of anomaly attacks is the first and crucial point for avoiding any interruption in the connection. Machine Learning algorithms have been rising and improving substantially year by year. Many classic tests can detect many attacks in the current time. However, those techniques are not enough for security since the types of attacks are changing and getting stronger frequently. In this study, we propose that how to detect a maximum number of attacks in IoT networks by applying machine learning techniques, especially K-Nearest Neighbors (KNN), Logistic Regression (LR), and Random Forest (RF) models. Dataset is presumably one of the most important starting points for the use of those techniques. UNSW-NB15 dataset which is publicly available has been used for this study. K-Nearest Neighbors algorithm shows 98.03% accuracy which is the best performance within the selected algorithms.**

***Keywords*- Internet of Things, Security, Attack detection, Machine Learning, Confusion matrix, Classification report**

I. INTRODUCTION

The Internet has connected computers and is spreading in every corner of the world at present. It provides many protocols which are used for communicating networks. Semi-Automatic Ground Environment (SAGE) and Semi-Automatic Business Research Environment (SABRE) have connected at the beginning of the 1950s due to the Internet network. The Advanced Research Projects Agency (ARPA) acted as design financing for ARPANET in the 1960s. In the late 1960s, usage of the internet increased in that time. From ARPANET, the modern internet comes which are now using by people.

Devices are needed to be used to communicate with each other due to the increase in Internet usage in recent years. The term named Internet of Things (IoT) has been taken place everywhere for this reason. IoT is forming connections among one smart device with many devices and building a communication network. This definition is used for the first time by Kevin Ashton in 1991. Internet of Things includes wide areas such as wearable technology, kitchen robots, electrical cars, mobile phones, etc.

To create better performance for the users of these systems, many researchers have been studying IoT for a long time. There are a lot of problems in the IoT environment. Cyberattacks are one of the most important problems of IoT devices. This study focuses on how to detect attacks on IoT devices using machine learning techniques, especially K-Nearest Neighbors (KNN), Logistic Regression (LR), and Random Forest (RF) Models to increase security layers of the network.

Review of related literature, types of attacks, dataset, machine learning methods, data analysis, and results and conclusion are discussed in sections II, III, IV, V, VI, and VII, respectively.

II. Review of related literature

Many research works related to the detection of malicious activities in IoT networks are going on. Monika et al. [1] proposed an advanced intrusion detection system for attacks detection in IoT networks using deep learning models named Convolutional Neural Network combined with LSTM applied on CICIDS2017 dataset, obtained an accuracy of almost 99.03%

Nilesh et al. [2] designed machine learning-based approaches for attacks detection in IoT networks using an Artificial Neural Network algorithm on the NSL-KDD99 dataset. Their model works well, and its accuracy is 99.4%. Machine learning approaches for anomaly detection in IoT networks have been proposed in [3] using K-Nearest Neighbors algorithm on the IoT network intrusion dataset. They showed 99% accuracy through their works.

Author et al. [4] a packet-level machine learning attacks detection for IoT devices has been proposed where K-Nearest Neighbors, Support Vector Machine, Neural Network, Random Forest and Decision Tree models are applied on KDD99 dataset. They got a 99% accuracy maximum in their research.

Abhinaya et al. [5] proposed high-level deep learning library for the detection of malicious activities in IoT networks where four deep learning models such as Convolutional Neural Network (CNN), Autoencoder, Multi-Layer Perceptron (MLP), and Deep Neural Network are used on UNSW-NB15 and NSL-KDD99 datasets. They showed 98% accuracy in their work with the model.

Canedo et al. [6] explained about security is one of the weakest areas in IoT. That is why they implemented the Artificial Neural Network (ANN) technique which is one of the Machine Learning techniques. They observed that can this technique detect anomaly attacks on the gateway. The authors also describing that IoT devices can divide into 2 primary groups. Edge devices and Gateway devices. Edge devices mean low-power and low-resource devices include sensors. Edge devices use for a single aim, kind of collecting humidity data and send to the gateway devices. Gateway devices mean they collect data from edge devices, and they connect with devices and gateways. The authors used Arduino Uno and Raspberry Pi 3. They used Arduino Uno to create an edge device. Raspberry Pi 3 was used to create a gateway device. Raspberry pi 3 works like a computer and it has 1 GB RAM for implement machine learning techniques to smaller data sets. They collected 4000 data samples from the edge device, and they stocked their data in the MySQL database. They implemented ANN to their real data and all data were not harmful. Then, they manipulated their sensors on edge devices and sent invalid data to the gateway device. They got 99% accuracy in their study.

Author et al. [7] used different ways for improving the IoT security level. He used cloud microservices for anomaly detection. The author used 3 different data sets and all of them include different information. For example, one of them includes credit card scores and credit risks. Another one includes temperature and humidity data. The author collected all data and sent it to the Raspberry Pi microcontroller. From this microcontroller, he uploads his data to the Microsoft Azure Machine Learning Studio. Finally, they are trying to detect malicious data and anomaly attacks. The author also used one of the most popular datasets which are publicly available. This dataset is generated from the University of California San Diego (UCSD). This dataset has many images of pedestrian walkways. If any non-pedestrian enters walkways, it looks like an anomaly attack. The author used 34 videos for training and 36 videos for testing. Each clip includes 200 frames.

Lawal et al. [8] propose different network anomaly reduction schemes in IoT networks. In this paper, they used 200,700 data and they implemented 3 different supervised machine learning algorithms in this dataset. Their result comes with different values. For example, in the KNN algorithm, they have 94% accuracy, Naïve Bayes has 85% accuracy. These results show to use more data increases accuracy. These results are not bad for 200,000 data but increased data might get better scores. The author used ROC (Receiver Operating Characteristic) curve method in their study. This curve works with two parameters. These are true positive rates and false-positive rates. When the curve reaches a peak, which refers to one decision threshold. The authors implemented the ROC curve to the same algorithms. Their new results increased well, such as KNN has 98% accuracy and Naïve Bayes has 96% accuracy. In the ROC curve, when the true positive rate and false positive rate reach the same value, it gives the best accuracy.

Zhang et al. [9] propose a unified model which is a combination of multiscale convolutional neural network (MSCNN) and long short-term memory (LSTM). This model is used for analyzing spatial-temporal features performance in the article. The author used 175,341 data for the training set and 82,332 data for the test set in model 1. The author also selected a dataset, which is classified by attack types, in this style, the author used 1601 normal and attack data in model 2. When it comes to the result section, the MSCNN-LSTM model has 95.6% accuracy for model 1, which is good but also the false-negative rate is 1.6% that means algorithm divided attack or normal mostly correct. When model 2 result comes on the article, MSCNN-LSTM model 89.8% accuracy which means algorithm did not detect attacks correctly in the study. Also, the false-negative rate goes up to 8.6% that is normal, because classification and correlation matrix has grown up, and it makes complicated mathematical problems. The

interesting point of this study is algorithm got 99.1% accuracy for worm attacks. Even worm attack is the lowest number of attack type in the dataset, that sound is good.
The existing proposals for attacks detection in IoT networks result in low accuracy. Also, these schemes used datasets which are available in public for a decade. So, the chances for anomaly detection of newly released IoT devices are low. For this reason, this study focuses on multiple number machine learning approaches on a newly available dataset for obtaining more than 95% accuracy.

III. TYPES OF ATTACKS

The usage of IoT is increasing day by day. IoT devices do not have effective antivirus systems. Although the malware things, Scientists are producing useful firewalls for IoT. Even though this huge security market, many attacks are happening, and they are giving more damage to devices. This section offers to explain some of the attack types which is used in this study, all attack types have different details.

A. Backdoor

Backdoor is a secure illegal remote Access to a device which is attacker circumvents a normal furtiveness authentication method. It is very tough for detection in the system because it runs background and is covered. That is why users can not recognize this attack type. If attackers reach to device from the back door, it is possible to steal information, install malware applications, and easily take it under the control of computer systems.

B. Denial of Service

Assume that a data center has 500 Mbps upload that it is giving service. The malicious bots who organized from attackers, if they create a 500 Mb traffic in a second, the application cannot give service. This attack calls DoS. DoS attack sends packets such as unnecessary messages, and they have an invalid return address. Computer systems are not available for finding certain return addresses while attackers sending packets, and computer systems start to wait before the connection is interrupted. After the connection is interrupted, attackers are sending more and more packets to the system, these invalid return addresses cut all connections and users cannot reach such as browsers, emails, and applications.

C. Shellcode

This attack uses a payload of software susceptibility. The attack is a limited piece of code that is inside an exploit attack. It Works on command Shell in software. This malignant code is implemented in computer memory (RAM) thoroughly exploiting the susceptibilities of stack buffer overflow.

D. Reconnaissance

This attack could also be described as a probe. This attack collects data from the network to avoid its security controls. This attack type works from ping. Attack follows ping to detect real IP address.

E. Exploit

This attack causes malicious behavior on the network with the array instructions such as bugs or vulnerable things. These behaviors include take the control of computer system or giving access to the system easily for DoS attacks.

IV. DATASET

As we explained in the related literature reviews section, all machine learning algorithms need the dataset for implementation. Science has been creating many dataset types such as DARPA98, KDD, NSL-KDD, CAIDA, UNSW-NB15, etc. All these datasets have different implementation areas, for example, information technologies, autonomous cars, education, Fintech, human resources, etc. All implementation areas have similar issues in their fields. The most significant point in these issues is security. UNSW-NB15 dataset has been used in this study. This dataset has created in the Cyber Range Lab of the Australian Centre for Cyber Security (ACCS) through the IXIA

Perfect Storm tool [10]. This dataset includes a hybrid of normal traffic and attack traffic. Twelve algorithms and tools such as Bro-IDS, Argus have been used to produce UNSW-NB15. This dataset includes 39 features with a class label. This dataset has over 2.5 million data which is good for prediction. Table I shows the categorization of features of the dataset.

TABLE I
FEATURES CATEGORIZATION OF UNSW-NB15 DATASET

Number	Name of category	Description
1	Flow features	It contains the identifier attributes between hosts such as client to server or server to client
2	Basic features	It includes the attributes that characterize the connections of protocols
3	Content features	It contains the attributes of TCP/IP and also includes some attributes of http services
4	Time features	It includes the attributes of time such as round-trip time of TCP protocol start/end packet time arrival time between packets etc.
5	Additional generated features General purpose features	Own purpose features which to care for the protocols service
	Connection features	Built based on the chronological order of the last time feature
6	Labelled features	It shows the label of recorded data

Table II shows how much data UNSW-NB15 has in the dataset. In this study, all attack data have been used for getting the best result. This dataset has five different attack types and all of them are created in the dataset.

TABLE II
NUMBER OF DATA IN UNSW-NB15 DATASET

Name	Count
Total Number of Data	2,540,044
Normal	2,218,761
Attacks	321,283

Table III shows the number of attack data and attack category in the dataset.

TABLE III
ATTACKS CATEGORIZATION IN UNSW-NB15 DATASET

Attack Category	Number of Data
Backdoor	63294
Denial of Service	56353
Shellcode	75118
Reconnaissance	71993
Exploits	54525

V. MACHINE LEARNING METHODS

Machine Learning is the subfield of Artificial Intelligence. It is one of the best methods that can provide a result without being explicitly programmed. Machine learning has four big different techniques. Supervised Learning, Unsupervised Learning, Reinforcement Learning, and Semi-Supervised Learning.

A. K-Nearest Neighbor Model

K-Nearest Neighbor (KNN) is the subfield of the supervised learning method. This algorithm works for regression and classification problems. The rule of this algorithm is similar things should have similar characteristics. When the algorithm classifies all characteristic things, it should be close to proximity. This algorithm estimates the distances between input value and classified value. K refers to when we classify the new data, algorithm looks to classified data for examine the closeness between these data. KNN is extremely easy to implement in its most basic form, and yet performs quite complex classification tasks. It is a lazy learning algorithm since it does not have a specialized training phase. Rather, it uses all data for training while classifying a new data point or instance. KNN is a non-parametric learning algorithm, which means that it does not assume anything about the underlying data. This is an extremely useful feature since most of the real-world data does not really follow any theoretical assumption e.g., linear separability, uniform distribution, etc. The intuition behind the KNN algorithm is one of the simplest of all the supervised machine learning algorithms. It simply calculates the distance of a new data point to all other training data points. The distance can be of any type e.g., Euclidean or Manhattan, etc. It then selects the K-nearest data points, where K can be any integer. Finally, it assigns the data point to the class to which most of the K data points belong. Figure 1 shows KNN algorithm.

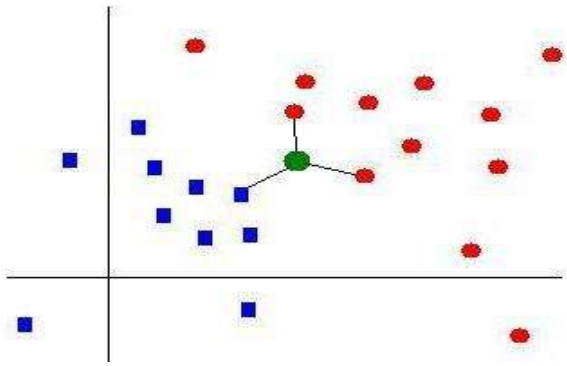


Figure 1. K-Nearest Neighbor Algorithm

B. Logistic Regression Model

Logistic Regression is a supervised machine learning algorithm from the field of Statistics. It is the method for binary classification problems. Logistic Regression is named for the function used at the core of the method known as logistic function. This function is also called sigmoid function. Logistic Regression is known as logit regression, maximum-entropy classification, or the log-linear classifier. In this model the probabilities describing the possible outcomes of a single trial are modeled using logistic function or sigmoid function. A logistic function f(x) or logistic curve is a common “S” shape (sigmoid curve) in Figure 2. Where f(x) is the prediction probability and x is the input data, f(x)=1/1+e^-(x)

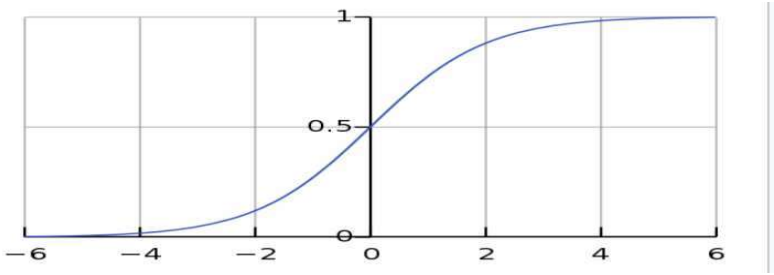


Figure 2. Sigmoid Curve

C. *Random Forest Model*

Random Forest is a supervised machine learning method that operates by constructing multiple decision trees. A decision tree is a tree-shaped diagram used to determine the output. Figure 3 shows that each branch gives the output of individual possible outcomes. The base estimator of this model is the decision tree. Each estimator is trained on a different bootstrap sample having the same size as the training set. The Random Forest model introduces further randomization on features in the training of individual trees. Random Forest algorithm is used for both classification and regression problems. Classification aggregates predictions by the majority votes. And regression aggregates predictions by averaging.

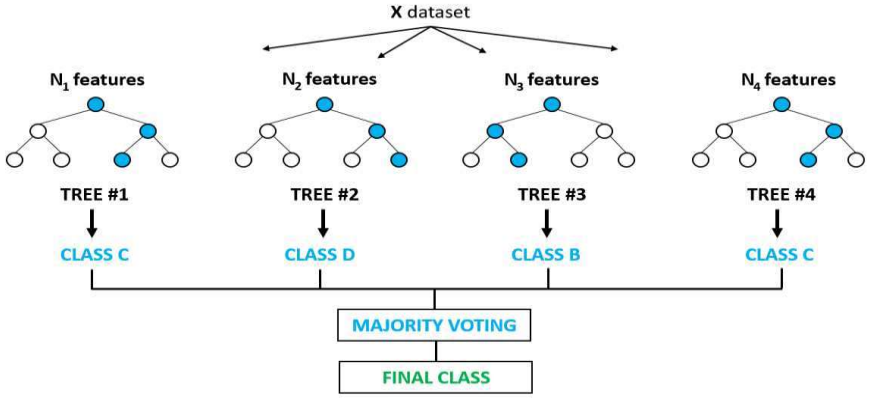


Figure 3. Random Forest Model

VI. DATA ANALYSIS AND RESULTS

The evaluation of the results of the implementation of the algorithms has been calculated according to the confusion matrix. Size of the matrix is depending on the number of the class in the testing dataset. There are a couple of criteria terms produced by the confusion matrix. Those criteria's names are respectively precision (Pr), recall (Rc), f-measure (F1), and accuracy.

TP = Number of True Positive predicted values. (Correct classified)
TN= Number of the True Negative predicted values. (Incorrect classified)
FP= Number of the False Positive predicted values. (Correct classified)
FN= Number of the False Negative predictive values. (Incorrect classifier)

Table IV, Table V, and Table VI show all type confusion matrix which is used in this study.

TABLE IV
RESULT TYPES OF IMPLEMENTED ALGORITHMS

Precision	TP/(TP+FP)
Recall	TP/(TP+FN)
F-measure	(TP+TN)/(TP+TN+FP+FN)
Accuracy	$2*\frac{Precision * Recall}{Precision + Recall}$

TABLE V
2X2 CONFUSION MATRIX

Actual Value	Predicted Value	
	True Positive	False Negative
	False Positive	True Negative

TABLE VI
NXN CONFUSION MATRIX

		Predicted Value			
		Class 1	Class 2	…	Class n
Actual Value	Class 1	X_{11}	X_{12}	…	X_{1n}
	Class 2	X_{21}	X_{22}	…	X_{2n}
	. .	. . .	. . .	. . .	. . .
	Class n	X_{n1}	X_{n2}	…	X_{nn}

Due to imbalanced data, we have used 210,000 datasets. It includes 61,987 attack data and 148,013 normal data. When we implement our first machine learning algorithm, which is the KNN algorithm, we got 98.03% accuracy in this algorithm. 896 benign data have been predicted as attack traffic, and 3245 attack data have been predicted as benign data. Table VII shows the confusion matrix of the KNN algorithm. Figure 4 and Figure 5 show the classification report and ROC curve of the model, respectively.

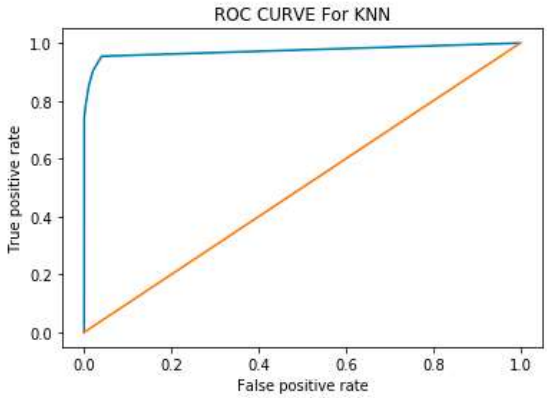
TABLE VII
CONFUSION MATRIX OF KNN MODEL

KNN	Benign	Attack
Benign	193168	896
Attack	3245	12691

https://sites.google.com/site/ijcsis/
ISSN 1947-5500
https://doi.org/10.5281/zenodo.5504352

	precision	recall	f1-score	support
0	0.98	1.00	0.99	194064
1	0.93	0.80	0.86	15936
accuracy			0.98	210000
macro avg	0.96	0.90	0.92	210000
weighted avg	0.98	0.98	0.98	210000

Figure 4. Classification Report of KNN Model



ROC_AUC_Score of KNN Model : 0.9732975151155989
Accuracy of KNN Model : 0.9802809523809524

Figure 5. ROC Curve of KNN Model

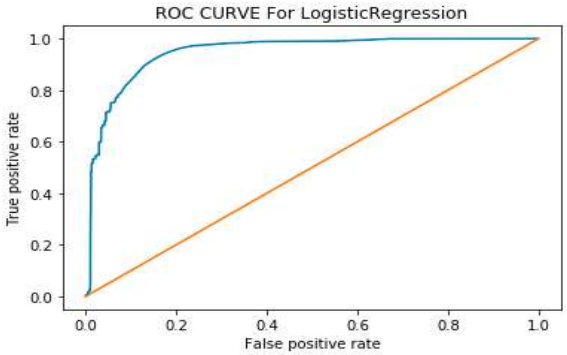
When we implement the Logistic Regression method, we got 92.4% accuracy in attack detection. Table VIII shows the confusion matrix of this algorithm where 15936 attack data is predicted as benign data. Figure 6 and Figure 7 show the classification report and ROC curve of the model, respectively.

TABLE VIII
CONFUSION MATRIX OF LOGISTIC REGRESSION MODEL

Logistic Regression	Benign	Attack
Benign	194064	0
Attack	15936	0

	precision	recall	f1-score	support
0	0.92	1.00	0.96	194064
1	0.00	0.00	0.00	15936
accuracy			0.92	210000
macro avg	0.46	0.50	0.48	210000
weighted avg	0.85	0.92	0.89	210000

Figure 6. Classification Report of Logistic Regression Model




```
ROC_AUC_Score of LogisticRegression Model :  0.9476864807061953
Accuracy of LogisticRegression Model :  0.9241142857142857
```

Figure 7. ROC Curve of Logistic Regression Model

When we implement the Random Forest model, we got 92.41% accuracy in attack detection. Table IX shows the confusion matrix of this algorithm where 15936 attack data is predicted as benign data. Figure 8 and Figure 9 show the classification report and ROC curve of the model, respectively.

TABLE IX
CONFUSION MATRIX OF RANDOM FOREST MODEL

Random Forest	Benign	Attack
Benign	194064	0
Attack	15936	0

	precision	recall	f1-score	support
0	0.92	1.00	0.96	194064
1	0.00	0.00	0.00	15936
accuracy			0.92	210000
macro avg	0.46	0.50	0.48	210000
weighted avg	0.85	0.92	0.89	210000

Figure 8. Classification Report of Random Forest Model

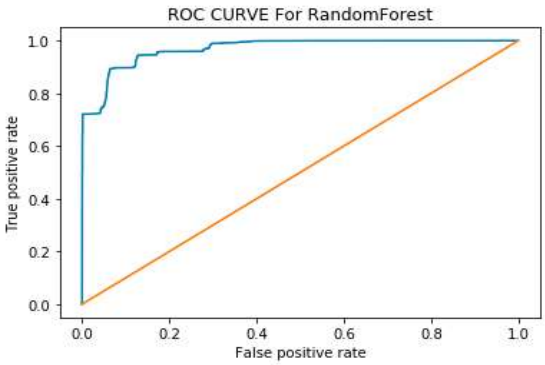


Figure 9. ROC Curve of Random Forest Model

VII. CONCLUSION

In this technology age, it shows us the electronic devices which relate to the internet, it covers human life. The data science world is developing new techniques for the security of those devices. When we look for the collection of data in networks, IoT devices are closing to the computer systems. This study aims for improving the security level of IoT networks. With the machine learning algorithms, we reached at least 92% accuracy with all our algorithms that are implemented to the dataset. We got the highest 98.03% accuracy from the K-Nearest Neighbors algorithm which is a good contribution in respect to the state-of-the-art of research. This study reached our goals, we detected most of the attacks in our dataset. Data science will grow up with our next research and articles. Data is very important, so we always try to secure all types of data. In the future, we will research to apply Deep learning and Reinforcement learning models for bigger datasets to get better accuracy result in detecting attacks.

REFERENCES

[1] M. Roopak, G. Y. Tian and J. Chambers, "An Intrusion Detection System against DDoS Attacks in IoT Networks," 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), 2020.
[2] N. K. Sahu and I. Mukherjee, "Machine Learning based anomaly detection for IoT Network," International Conference on Trends in Electronics and Informatics (ICOEI 2020), 2020
[3] Z. Liu, N. Thapa, A. Shaver, K. Roy, X. Yuan and S. Khorsandroo, "Anomaly Detection on IoT Network Intrusion Using Machine Learning," International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD), 2020.
[4] R. Doshi, N. Apthorpe and N. Feamster, "Machine Learning DDoS Detection for Consumer Internet of Things Devices," 2018 IEEE Symposium on Security and Privacy Workshops, 2018.
[5] A. Nagisetty and G. P. Gupta, "Framework for Detection of Malicious Activities in IoT Networks using Keras Deep Learning Library," International Conference on Computing Methodologies and Communication (ICCMC), 2019
[6] J. Canedo, A. Skjellum, "Using Machine Learning to Secure IoT Systems," 14th Annual Conference on Privacy, Security and Trust (PST), 2016.
[7] N. Rakesh, "Performance analysis of Anomaly detection of different IoT datasets using cloud micro services," International Conference on Intensive Computation Technologies, 2016.
[8] M.A. Lawal, R.A. Shaikh, S.R. Hassan, "Security Analysis of Network Anomalies Mitigation Schemes in IoT Networks," IEEE Access Vol. 8, Pages 43355-43374, February 27, 2020, ISSN: 2169-3536.
[9] J. Zhang, Y. Ling, X. Fu, X. Yang, G. Xiong, R. Zhang, "Model of The Intrusion Detection System Based On The Integration of Spatial-Temporal Features," Computers & Security, Volume 89, February 2020, 101681.
[10] N. Moustafa, J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection," in Proceedings of the MilCIS-IEEE Stream, Military Communications and Information Systems Conference, IEEE publication, Canberra, Australia, November 2015