International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488|

||Volume 8, Issue 5, May 2020||

Enhanced Security in Internet Environment using Machine Learning Algorithms

Amar Patil[1], Vaidehi M[2]

Department of ISE, Dayananda Sagar College of Engineering, VTU, Bengaluru, Karanataka, India[1,2]

ABSTRACT: In the current scenario the Internet access using Wi-Fi is common in institution, government and many more MNC companies there is a possibility of being malicious attack. To detect an unauthorized Access node for the protection of information. In this paper we use data set of authorized and unauthorized Access nodes in the Wi-Fi environment, analyse them using Machine learning Algorithms include KNN (K Nearest Neighbors), SVM (Support Vector Machine), MLP (Multilayer Perceptron), C4.5.

I.INTRODUCTION

The wireless network devices have been using rapidly so it's hard to find places without Wi-Fi in one's lives. Wi-Fi is available in cafes, companies, schools and military facilities. Wi-Fi has been using by many specified users by making very much difficult to check everyone. For authorized Wi-Fi identification is very much difficult unless you look directly at the Access node. In a wireless local area network (WLAN) a thread called access node has emerged an important security problem. An access node is a station that receives and transmits the data which is known as transceiver. Two types of access node can be set with different equipment's. The first type uses a wireless router connected directly into an Ethernet jack on a wall .The second type are set on a portable laptop with two wireless cards one connected to a real Access node and the other configured as an Access node to provide internet access to WLAN Station Due to the various smart devices the existence of unauthorized Access node has become unavoidable. Usage is also irrelevant as there is no provisions or regulations related to unauthorized Access node such as hotspots as well as public places. This actually provides a weak point to wireless networks. Due to this the network can easily be harmed by gleaming information of other users who have access to unauthorized Access node.

II.RELATED STUDIES

The growing use of Wi-Fi network creates so many threats that access data from networks which can be damage for a particular organization network.

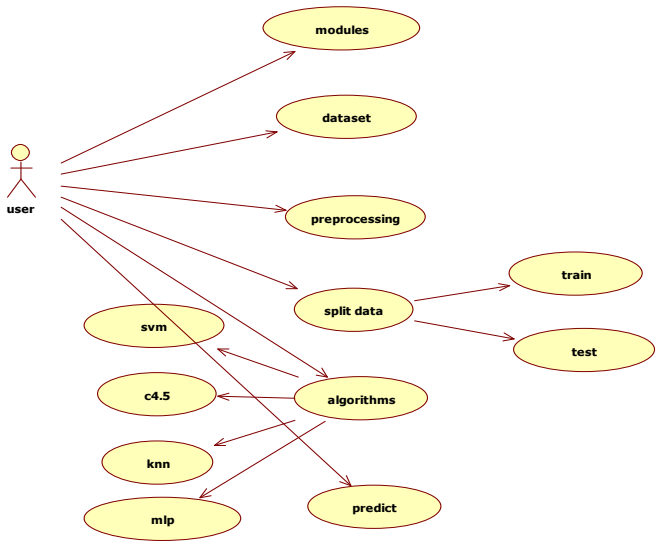

Fig. 1. Use Case Diagram

So, the data set thus constructed is applied to the machine learning algorithm to obtain the result, and then the results obtained are compared, to show which algorithm is more accurate. The Use Case Diagram As show in Figure 1.

The machine-learning algorithms and their features used in this paper for classification are as follows:

- SVM (support vector machine): Based on a given set of data, we create a non-probabilistic binary linear classification model that determines which classifications of new data should be broken down and used to represent boundaries in the space in which data is mapped. The SVM algorithm is the algorithm that finds the boundary with the largest width.
- C4.5: It is one of the algorithms for classifying and predicting data by making a decision tree. It is an algorithm that complements the limit of the existing ID3 algorithm. The C4.5 algorithm uses the concept of information entropy to create a decision criterion and uses it to classify the sample set most effectively.
- KNN(k-nearest neighbors algorithm) : As a type of map learning, the input consists of the k closest training data in the feature space, and if used for classification purposes, the object is the object assigned to the most common item among the k nearest neighbors and classified by majority vote.
- MLP (multilayer perceptron): The hidden layer is added between the input layer and the output layer, and supervisory learning is performed using the back-propagation algorithm, so that data that cannot be linearly separated can be classified.

III.SYSTEM CONFIGURATION AND DATASET EXTRACTION

For the Detection of unauthorized Access node, the given data can be made by Data Pre-processing and Data Sampling, Data Pre-processing is the step-in which data get transformed, the features of the data can now be easily interpreted by the algorithm.
We find the Four Major attack class attack in Data Pre-processing.

- Denial of Service (DoS): is an attack in which an adversary directed a deluge of traffic requests to a system in order to make the computing or memory resource too busy or too full to handle legitimate requests and in the process, denies legitimate users access to a machine.
- Probing Attack (Probe): probing network of computers to gather information to be used to compromise its security controls.
- User to Root Attack (U2R): a class of exploit in which the adversary starts out with access to a normal user account on the system (gained either by sniffing passwords, a dictionary attack, or social engineering) and is able to exploit some vulnerability to gain root access to the system.
- Remote to Local Attack (R2L): occurs when an attacker who has the ability to send packets to a machine over a network but who does not have an account on that machine exploits some vulnerability to gain local access as a user of that machine.

The Output of Attack Class Distribution As show in Figure 2.

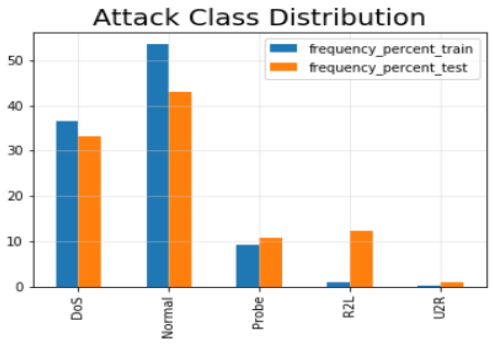

Fig. 2. Attack Class Distribution

The Data Sampling is a statistical analysis technique used to select, manipulate a representative subset of data points to identify patterns and trends in the larger data set being examined. The Output of Data Sampling feature and importance as shown in Figure 3.

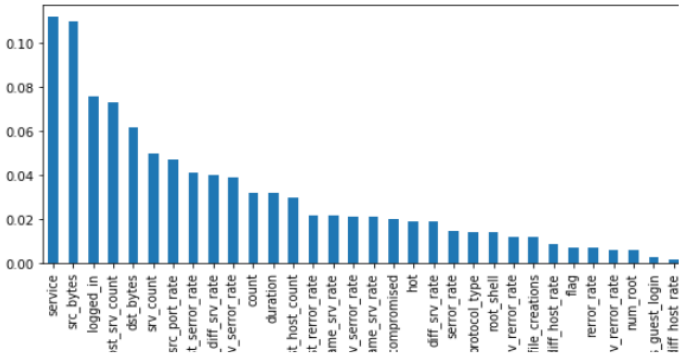

Fig. 3. Feature and importance of Data Sampling

IV.EXPERIMENTAL RESULTS

In the experiments, the algorithm to be compared were selected from the classification. The algorithms are SVM (Support Vector Machine), C4.5, KNN (K nearest neighbors) and MLP (Multilayer Perceptron). The experimental results for each classification algorithm are shown in Table 1.

TABLE I.EXPERIMENTAL RESULT FOR ALGORITHMS

Algorithms	**SVM**	**C 4.5**	**KNN**	**MLP**
Accuracy	0.9891	0.9999	0.9977	0.9989

V.CONCLUSION

In this, Presented the simple Intrusion Detection and Prevention Approach for unauthorized access node in Wireless LAN. The difference between authorized and unauthorized access node can be classified by Machine Learning algorithms. If we detect the attacks from an unauthorized access node, can be disconnect it for protection of the system. The method will be applied to the protection of information, including personal lifelog data.

ACKNOWLEDGMENT

I would like to thank Mrs.Vaidehi M, Assistant Professor in Department of ISE, Dayananda Sagar College of Engineering, Bengaluru, Karanataka, India

REFERENCES

[1] S. Jana and S.K. Kaseya. "On fast and accurate detection of unauthorized wireless access points using clock skews," IEEE Transactions on Mobile Computing, Vol. 9, No. 3, pp. 449-462,2010

[2] Hahn, et al. "A timing-based scheme for rogue AP detection," IEEE Transactions on parallel and distributed Systems, Vol. 22, No.11, pp. 1912-1925, 2011.

[3] F. Awad, M. Al-Refai, and A. Al-Qerem. "Rogue access point localization using particle swarm optimization," in 8th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan. May 2017. doi: 10. 11 09/ IA CS. 2017.7921985

[4] M. Agarwal, S. Biswas, and S. Nandi. "An Efficient Scheme to Detect Evil Twin Rogue Access Point Attack in 802.11 Wi-Fi Networks," International Journal of Wireless Information Networks, Vol 25. No. 3, pp 120-135, 2018.

[5] V. Modi, and C. Parekh. "Detection & Analysis of Evil Twin Attack in Wireless Network," International Journal of Advanced Research in Computer Science Vol. 8, No. 5, pp. 774-777, 2017.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488|

||Volume 8, Issue 5, May 2020||

[6] V. Vapnik, “Support Vector Machine,” in The nature of statistical learning theory, Springer science & business media, 2013.
[7] J. R. Quinlan, C4. 5: Programs for Machine Learning, Morgan Kaufmann Publishers Inc., 1993.
[8] Kuo-Fong Kao, I-En Liao, Yueh-Chia Li,"Detecting rogue access points using client-Side bottleneck bandwidth analysis”.
[9] Qu, G., Nefey M.M.,“RAPid an Indirect Rogue Access point Detection System”, IEEE 2010.
[10] Chao Yang, Yimin Song, Guofei Gu,“Active User-side Evil Twin Access Point Detection Using Statistical Techniques