**Unauthorized Access Point Detection Using Machine Learning Algorithms for Information Protection**

**Abstract**

With the frequent use of Wi-Fi and hotspots that provide a wireless Internet environment, awareness and threats to wireless AP (Access Point) security are steadily increasing. Especially when using unauthorized APs in company, government and military facilities, there is a high possibility of being subjected to various viruses and hacking attacks. It is necessary to detect unauthorized Aps for protection of information. In this paper, we use RTT (Round Trip Time) value data set to detect authorized and unauthorized APs in wired / wireless integrated environment, analyze them using machine learning algorithms including SVM (Support Vector Machine), C4.5, KNN (K Nearest Neighbors), and MLP (Multilayer Perceptron). Overall, KNN shows the highest accuracy.

**Keywords**—detection, unauthorized AP, machine learning, protection

**Existing system:**

In a wireless local area network (WLAN), an access point is a station that transmits and receives data (sometimes referred to as a transceiver). An access point connects users to other users within the network and can also serve as the point of interconnection between the WLAN and a fixed wire network. Research to detect log AP and its risks have been actively studied until recently. Various methods of research are currently in progress, addressing various aspects of the issue. In order to prevent such damage, it is necessary to ascertain which AP is an illegal AP.

**Disadvantages:**

However, due to various smart devices, the existence of unauthorized AP has become unavoidable. Usage is also irrelevant, because there are no regulations or provisions relating to unauthorized APs, such as hotspots, as well as public places. This provides a very weak point to wireless networks. The network can be harmed by stealing or gleaming information of other users who have access to unauthorized APs, and because PCs can also be hacked.

**Proposed system:**

In this, a dataset was created using KDDkup values. 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. For detection of unauthorized AP, we developed the “Intelligent Wireless AP Detection System” as shown in Architecture. Data are collected from the authorized APs and unauthorized Aps and constructed as data sets. Data sets are analyzed by applying machine-learning algorithms including SVM (Support Vector Machine), C4.5, KNN (K Nearest Neighbors), and MLP (Multilayer Perceptron). The unauthorized AP has built a new AP using the LG XNOTE P210-GE30P and the iptime N500 connected to it.

**Advantages:**

The protocol used in network experiments can affect the results depending on which one is used. There may be a big difference in the communication protocol for each protocol, and the bandwidth and channel can also cause errors in the experiment. 802.11n is used for wireless communication protocol that is the most widely used in real world.

**Architecture:**

****

**Modules:**

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 (Iterative Dichotomizer 3) 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.

**System Specifications:**

**Hardware Requirements:**

• RAM: 4GB and Higher

• Processor: Intel i3 and above

• Hard Disk: 500GB: Minimum

**Software Requirements:**

* OS: Windows or Linux
* Python IDE : python 2.7.x and above
* Pycharm IDE Required, jupyter notebook
* Language : Python Scripting