American Journal of Artificial Intelligence
2021; 5(2): 46-55
http://www.sciencepublishinggroup.com/j/ajai
doi: 10.11648/j.ajai.20210502.11
ISSN: 2639-9717 (Print); ISSN: 2639-9733 (Online)

Classification of Spectrum Scheduling Using Conditional Probability and Decision Tree Supervised Learning Approach

Imeh Umoren[*], Esther Polycarp, Godwin Ansa

Department of Computer Science, Faculty of Physical Sciences, Akwa Ibom State University, Mkpat Enin, Nigeria

Email address:
imehumoren@aksu.edu.ng (I. Umoren), estherpolycarp@aksu.edu.ng (E. Polycarp), godwinansa@aksu.edu.ng (G. Ansa)
[*]Corresponding author

To cite this article:
Imeh Umoren, Esther Polycarp, Godwin Ansa. Classification of Spectrum Scheduling Using Conditional Probability and Decision Tree Supervised Learning Approach. *American Journal of Artificial Intelligence.* Vol. 5, No. 2, 2021, pp. 46-55. doi: 10.11648/j.ajai.20210502.11

Received: June 18, 2021; **Accepted**: July 21, 2021; **Published**: August 27, 2021

Abstract: Spectrum Scheduling is an efficient scheme of improving spectrum utilization for faster communications, higher definition media (HDM) and data transmission. Radio spectrum is very limited in supply resulting in enormous problems related to scarcity. It owes the physical support for wireless communication, both fixed applications and mobile broadband. Basically, effective use of the spectrum depends on the channel settings, sensing performance, detection of spectrum prospect as well as effective transmission of both Primary Users (PUs) and Secondary Users (SUs) packets at a specific time slot. In order to improve spectrum utilization this paper adopted quantitative method which employs Probability Theorem to identify the probabilities of both primary Users (PUs) and secondary users (SUs) in the spectrum datasets allocation and further used conditional probability to compare two Frequency Bands i.e., High Frequency (HF) and Very High Frequency (VHF). The result indicates available spectrum holes (SH) left unutilized in the Secondary User (SU) resulting in the need for spectrum scheduling for the SU. The procedure makes the secondary users occupy a probability of 0.002mhz compared to the primary users on 0.00004mhz utilization. This further indicates that some spectrum holes were left unutilized by the license users (Primary Users). However, spectrum allocation is one of the major issues of improving spectrum efficiency and has become a considerable tool in cognitive wireless networks (CWN). Consequently, the goal of spectrum allocation is to assign leisure spectrum resources efficiently to achieve the optimal Quality of Service (QOS and cognitive user requirements of wireless network. Again, classification of spectrum allocation was carried out through difference methods. Firstly, we employ a probability theorem to identify the probability of both Primary Users (PUs) and Secondary Users (SUs) in the allocated spectrum data sets. Secondly, conditional probability was used to compare two frequency band based on primary and secondary allocation policies designed to identify the specific allocation of each band. Thirdly, Machine Learning (ML) Algorithm based on Decision Tree - Supervised Learning (DTSL) approach was adopted to classified our data sets. The result yielded 68% which correctly classified instances based on the total records of sixty-nine (69) data sets. Research findings demonstrate a highly optimized spectrum scheduling for efficient networks service provisions.

Keywords: High Frequency (HF), Very High Frequency (VHF), Primary Users (PUs), Secondary Users (SUs), Decision Tree Supervised Learning (DTSL), Algorithm and Cognitive Wireless Networks (CWN)

1. Introduction

Presently, most organizations in the world are embracing wireless communication to improve their businesses. Some of the basic digital technology involves mobile, cloud computing, security and privacy as well as blockchains. This increasingly dependent on the evolving range of wireless communication services and the demands being placed on the scarce supply of usable radio spectrum are rapidly increasing. Mobile communication is a broad term that incorporates all procedures and forms of connecting and communicating between two or more devices using a radio waves or wireless signal through range of wireless communication technologies and devices. Fundamentally, the public needs for more information, faster communications and higher definition media (HDM) entails that the demand for radio spectrum exceeds supply. The concepts of spectrum refer to the invisible radio frequencies that wireless signals travel through. These signals are what enable us to make calls from our mobile devices, pull up directions to a destination, and carry out everything on our mobile devices. The frequencies we use for wireless communication are portion of what is called the electromagnetic spectrum. The entire electromagnetic spectrum encompasses other frequencies we interact with daily. Other parts of spectrum carry broadcast radio, television and serve other everyday functions. Portions of electromagnetic spectrum are grouped in "bands" depending on their wavelengths, the distance over which the wave's shape repeats. The development in wireless devices, and applications has led to increase in the demand for efficient radio frequency spectrum. This is expected to expand even further based on the given projection that the global traffic per year may reach 4.8 zettabytes by 2022 Cisco [3]. Moreover, it is expected that the number of Internet users may reach 4.8 billion and the number of connected devices is expected to be close to 28.5 billion devices as presented by Cisco [3]. Nevertheless, radio spectrum is often allocated and divided accordingly. FCC [4], proposed provision of more spectrum to increase or expand existing services to offer quality of service (QoS) delivery which has become more challenging. However, the recent research has shown that the issue of expanding spectrum for most users on demand for optimal services is not as a result of lack of spectrum but rather based on spectrum access Tilghman. P [18]. This means that, spectrums capacity is not well utilized and fully exploited. Mainly, this is due to the exclusive use licensing framework adopted by spectrum policy regulators and management strategies globally. Basically, operators under-utilize the spectrum license they allocated or hold, Opendata [15]. Indeed, numerous Service Providers (SPs) are searching for innovative ways to satisfy the exponential growth in data requirements for services provision. The increase in demand for new bandwidth-intensive services and applications such as video and music streaming simultaneously improving the average gain per user. Fundamentally, critical requirements for generation of networks which are extremely stringent centered on User Quality of experience (UQoE) and Quality of service (QoS) provisions. Hence, this paper proposed machine learning approach for classification and optimization of spectrum scheduling based on Decision Tree - Supervised Learning (DTSL) approach.

2. Review of Related Literature

The rapid growth of wireless standards and bandwidth technologies has led to a perceived spectrum scarcity. In order to satisfy the growing demand for efficient spectrum utilization. The spectrum management policy needs to be re-organized. The administrative or command and control approach to spectrum management has proven to be ineffective as presented by Linda E. [11]. The idea of sharing spectrum between a primary user (PU, the entity the spectrum was assigned to) and a secondary user (SU, the user that uses the spectrum opportunistically without interference) by using dynamic spectrum access techniques is what mainly cognitive radio is all about as recommended by Salim A. Hanna [17]. In a conventional method each user is assigned a license to operate in a certain frequency band. Most of the time spectrum remains unused and it is also difficult to find it. The allocated spectrum has not been utilized properly; it varies with time, frequency and geographical locations. Thus, to overcome the spectrum scarcity and unutilized frequency band, a new communication technique called Cognitive Radio (CR) and Dynamic Spectrum Access (DSA) is introduced. Cognitive radio was first introduced by Mitola J. and Maguire G. [12], Cognitive Radio (CR) network provides efficient utilization of the radio spectrum and highly reliable communication to users whenever and wherever needed. Dynamic Spectrum Access (DSA) technology allows unlicensed secondary system to share the spectrum with licensed primary system as proposed by Anita G., Partha P., Bhattacharya [1], and Mansi S. and Gajanan B. [13]. Basically, Cognitive Radio Networks (CRNs) are based on Cognitive Radio Devices (CRDs) which are able to configure different parameters such as Frequency Band, Waveform, Transmit Power, based on the surrounding environment. Consequently, exploiting under-utilized spectrum portions and avoiding irritating bottlenecks. Cognitive Process (CP) is used for purposes like collecting relevant information, Machine Learning and, based on that, reasoning and decision-making. Configurable radio platforms like Software Defined Radio (SDR) can be used to execute the decisions of Cognitive Process (CP). Measurement's operations and research studies has demonstrated that, some parts of spectrum are not used in term of time and space. These section of the spectrum, are named white spaces depicted figure 1, which have no active primary users. Secondary users in these parts of the spectrum may be able to detect and communicate with each other freely. This method is called the overlay manner, which is considered in many radio systems (2005). Imeh Umoren and Saviour Inyang [8] proposed a Fuzzy Knowledge-Based (FKB) approach with Triangular Membership Function (TMF) for evaluation of input parameters on selected three (3) mobile network operators in Niger Delta region and recordings were made for the periods of 21 days to ascertain service capacity. The model for optimizing mobile broadband networks was based on the test data. Results demonstrate that, the selected network operators vary in Quality of Service (QoS). Comparison in terms of signal strength, packet loss and data rate were observed at the instance of six (6) scenarios, Operator x provides reasonable Data rate of about 51.93mbps (lowest download speed), Operator y performed efficiently on packet loss with about 0.01% loss of packet and Operator z performed excellently well on signal strength of 98.23% for networks QoS and user Quality of Experience (uQoE) provisioning, Imeh Umoren and Samuel Okon, [9] measured wireless networks inherent variances from wireline networks based on network traffic metrics; such as *latency, packet Loss* *and packet delay* in certain wireless environments experienced some challenges in networks performance. To overcome these challenges, the paper focus on network performance optimization techniques and proposed a framework using *Type 1 Fuzzy knowledge-based* approach for efficient WLAN performance.

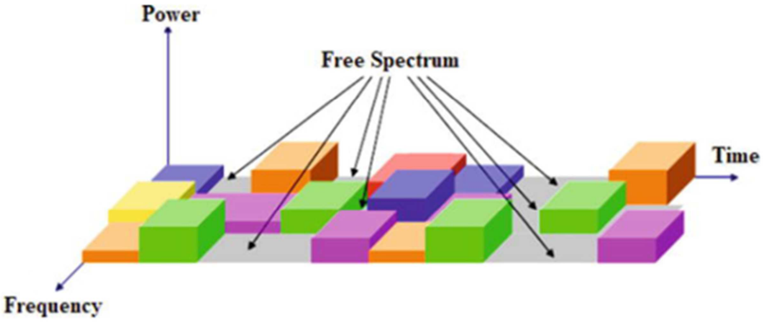

Figure 1. *The Free and occupied spectrums by Habibzadeh A. [6].*

Habibzadeh A. [7] Proposed that, using the capabilities of the cognitive radio, the spectrum can be used in the underlay manner, if the transmit power of the secondary users can be controlled below a predefined threshold level and it does not make any harmful impact on the primary users. Akyildiz, Won-Yeol L, Vuran M., Mohanty S. [2] carried out a study on new communication paradigm to exploit the existing wireless spectrum opportunistically. This new networking paradigm it was referred to as Next Generation Networks as well as Dynamic Spectrum Access (DSA) and cognitive radio networks. Li Zhang, Kai Z, Prasant M. [10] in their study Opportunistic Spectrum Scheduling for Mobile Cognitive Radio Networks in White Space. With fast growing number of secondary users, carefully scheduling the spectrum allocation in cognitive radio networks operating on white space becomes vital. Their studied focus on how to schedule the spectrum assignment for mobile cognitive radio devices. With the mobility information, their work formally defined the related problem as the Maximum Throughput Channel Scheduling problem (MTCS) which seeks a channel assignment schedule for each cognitive radio device such that the maximum expected throughput can be achieved. It was also, a general scheduling framework for solving the maximum throughput channel scheduling (MTCS). Raouia M., E. V. Belmega, Inbar F. [16] carried out a study on the centralized spectrum access and power management for several opportunistic users, secondary users (SUs), without hurting the primary users (PUs). The radio resource focus was to minimize the overall power consumption of the opportunistic system over several orthogonal frequency bands under constraints on the minimum quality of service (QoS) and maximized peak and average interference to the PUs. GSMA P. [5] carried out analysis of spectrum management nationally and internationally. It considered how to handle data growth, densification and new technologies to add. The study specified the need to provide access to more spectrum to further reduce the demand of this scarce commodity (Spectrum). The existing literature indicated that there is no existing research work currently on classification of spectrum scheduling using machine learning. This paper considers the classification of spectrum allocation based on primary and secondary users using conditional probability and decision tree Algorithm. The approach is to carry out spectrum classification based on primary and secondary users for effective spectrum utilization.

3. Conceptual Framework for Spectrum Classification

The Spectrum classification model is depicted in figure 2, which comprises of a step-by-step process. This step-by-step process method will aid us on how to carried out classification and optimization of spectrum using machine learning through a data mining approach to efficient schedule spectrum from the data set gathered.

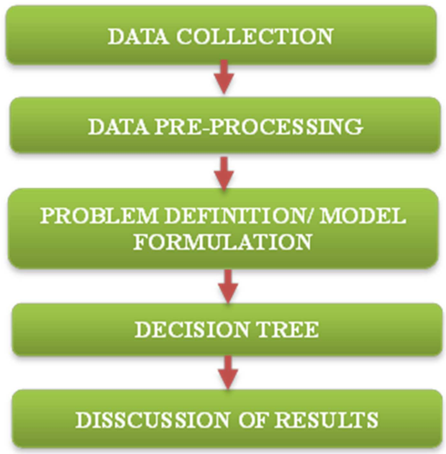

Figure 2. *Spectrum Classification Framework.*

4. Data Collection

Data collection is the process of gathering quantitative and qualitative information on specific variables with the aim of evaluating outcomes or gleaning actionable insights, in this research seminar, we have collected Strategy data for space spectrum, covering the satellite and space science sectors, and including Meteorological and Earth observation satellites in Uk recorded in 2016 online spectrum data repository. The spectrum allocation data obtained through a secondary method of data collection is depicted in table 2.

Table 1. *Sample Data for Spectrum Allocation (n=69) Source: Spectrum Opendata [14].*

Band	Allocation	AL Freq from (MHz)	AL Freq to (MHz)	AL Freq Range (MHz)	Rank
HF Band	Space Research	2.501	2.502	0.001	Secondary
HF Band	Earth Exploration-Satellite	432	438	6	Secondary
HF Band	Space Research	5.003	5.005	0.002	Secondary
HF Band	Space Research	10.003	10.005	0.002	Secondary
HF Band	Radio Astronomy	13.36	13.41	0.05	Primary
VHF Band	Space Operation	30.005	30.01	0.005	Primary
VHF Band	Space Research	30.005	30.01	0.005	Primary
VHF Band	Radio Astronomy	37.5	38.25	0.75	Secondary
VHF Band	Space Research	39.986	40.02	0.034	Secondary
VHF Band	Space Research	40.98	41.015	0.035	Secondary
P Band	Radionavigation-Satellite	399.9	400.05	0.15	Primary
P Band	Mobile-Satellite	399.9	400.05	0.15	Primary
P Band	Mobile-Satellite	400.15	401	0.85	Primary
P Band	Space Research	400.15	401	0.85	Primary
P Band	Space Operation	400.15	401	0.85	Secondary
L Band	Radionavigation-Satellite	1164	1215	51	Primary
L Band	Radionavigation-Satellite	1215	1240	25	Primary
L Band	Radionavigation-Satellite	1215	1240	25	Primary
L Band	Earth Exploration-Satellite	1215	1240	25	Primary
L Band	Space Research	1215	1240	25	Primary
S Band	Space Operation	1805	1880	75	Secondary
S Band	Mobile-Satellite	1980	2010	30	Primary
S Band	Mobile-Satellite	1980	2010	30	Primary
S Band	Space Operation	2025	2110	85	Primary
S Band	Earth Exploration-Satellite	2025	2110	85	Primary
C Band	Fixed-Satellite	3600	4200	600	Primary
C Band	Fixed-Satellite	3600	4200	600	Primary
C Band	Fixed-Satellite	3600	4200	600	Primary
C Band	Fixed-Satellite	4500	4800	300	Primary
C Band	Radio Astronomy	4800	4990	190	Secondary
X Band	Mobile-Satellite	7250	7300	50	Primary
X Band	Fixed-Satellite	7250	7300	50	Primary
X Band	Fixed-Satellite	7300	7450	150	Primary
X Band	Fixed-Satellite	7450	7550	100	Primary
X Band	Meteorological-Satellite	7450	7550	100	Primary

5. Problem Definition and Model Formulation

In Cognitive Radio (CR), a transceiver can intelligently detect which communication channels are in use and which ones are not in use. It instantly moves into vacant channels while avoiding occupied ones. It does not cause any interference to the licensed user. The users in cognitive radio are the Primary Users (PUs) and the secondary users (Sus). Primary Users (PUs) are users who has higher priority or. legacy rights on the usage of a specific part of spectrum, while Secondary Users are users (SUs) who has a lower priority. This research paper model Primary users and secondary users of spectrum based on the spectrum allocation data obtained (n=69), which was recorded in a cognitive radio enabled environment, in order for us to have efficient spectrum scheduling. The total number of spectrum data obtained were 69 records. Spectrum data as recorded. The parameters considered in the model from the data set in table 2, are given as Band-Freq, Allocations, Freq-Range, and Rank (Primary and Secondary).

This paper adopted probability theorem in modelling our spectrum schedule based on our data sets in table 2. Probability is a measure for calculating the chances or the possibilities of the occurrence of a random event. It calculates the chance of the favorable outcome amongst the entire possible outcome.

In mathematical terms probability is defined as the ratio of the number of favorable events to the total number of possible outcomes of a random experiment. It is denoted by ‘p’. The probability of an event, say ‘A’,

$$P(A) = \frac{number\ of\ favourable\ cases\ or\ outcomes}{total\ number\ of\ outcomes} \tag{1}$$

The probability of the events based on the band frequencies allocation are given below;

For HF

$$P(HF - SR - S) = \frac{0.036}{350.405} = 0.0001mhz$$

$$P(HF - SR - P) = 0$$

$$P(HF - RA - P) = \frac{0.17}{4468.205} = 0.00004mhz$$

$$P(HF - ES - S) = \frac{6}{350.405} = 0.0017mhz$$

For VHF

$$P(VHF - SO - P) = \frac{0.005}{4468.205}=0.000001\text{mhz}$$

$$P(VHF - SO - S) = \frac{0.175}{350.405} = 0.0005mhz$$

$$P(VHF - SR - P) = \frac{0.005}{4468.205} = 0.000001mhz$$

$$P(VHF - SR - S) = \frac{0.094}{350.405} = 0.003mhz$$

$$P(VHF - RA - S) = \frac{0.75}{350.405} = 0.002mhz$$

$$P(VHF - MS - P) = \frac{0.025}{4468.205} = 0.000006mhz$$

For P Band:

$$P(PBand - MS - S) = \frac{64.5}{350.405} = 0.1840mhz$$

$$P(PBand - MS - P) = \frac{1.1}{4468.205} = 0.0002mhz$$

$$P(PBand - RN - P) = \frac{0.15}{4468.205} = 0.00003mhz$$

$$P(PBand - SR - P) = \frac{0.85}{4468.205} = 0.00002mhz$$

$$P(PBand - SO - S) = \frac{0.85}{350.405} = 0.002\ mhz$$

$$P(PBand - SO - P) = \frac{5}{4468.205} = 0.001mhz$$

$$P(PBand - MT - S) = \frac{5}{350.405} = 0.01mhz$$

$$P(PBand - RA - P) = \frac{3.9}{4468.205} = 0.0009mhz$$

For L Band:

$$P(LBand - RN - P) = \frac{152}{4468.205} = 0.03mhz$$

$$P(LBand - EE - P) = \frac{85}{4468.205} = 0.02mhz$$

$$P(LBand - SR - P) = \frac{45}{4468.205} = 0.03mhz$$

$$P(LBand - EE - S) = \frac{3}{350.405} = 0.009mhz$$

For S Band:

$$P(SBand - SO - S) = \frac{80}{350.405} = 0.2mhz$$

$$P(SBand - SO - P) = \frac{85}{4468.205} = 0.02mhz$$

$$P(SBand - MS - P) = \frac{120}{4468.205} = 0.03mhz$$

$$P(SBand - EE - P) = \frac{85}{4468.205} = 0.02mhz$$

$$P(SBand - SR - P) = \frac{95}{4468.205} = 0.02mhz$$

For C Band:

$$P(CBand - FS - P) = \frac{2700}{4468.205} = 0.6mhz$$

$$P(CBand - RA - S) = \frac{190}{350.405} = 0.5mhz$$

$$P(CBand - RA - P) = \frac{10}{4468.205} = 0.002mhz$$

$$P(CBand - RN - P) = \frac{30}{4468.205} = 0.007mhz$$

$$P(CBand - AM - P) = \frac{10}{4468.205} = 0.002mhz$$

For X Band:

$$P(XBand - FS - P) = \frac{475}{4468.205} = 0.1mhz$$

$$P(XBand - MS - P) = \frac{125}{4468.205} = 0.03mhz$$

$$P(XBand - MT - P) = \frac{250}{4468.205} = 0.06mhz$$

$$P(XBand - SR - P) = \frac{90}{4468.205} = 0.02mhz$$

Hence, the total probability of primary user's spectrum allocated=0.993mhzs which the secondary users=0.917mhz, this shows that primary users had larger allocation in the spectrum utilization.

Additionally, a Conditional probability was adopted to access different proportions of band frequency in order to identify the proportions of each frequency band in the spectrum allocation data in table 1.

A conditional probability is the likelihood of an event or outcome occurring, based on the occurrence of a previous event or outcome. Conditional probability is calculated by multiplying the probability of the preceding event by the updated probability of the succeeding, or conditional, event.

Mathematically conditional probability is;

$$p(A/B) = \frac{p(A \cap B)}{p(B)} \tag{2}$$

Where $p(A \cap B) =$

$$p(A).P(B) \tag{3}$$

Substituting $p(A \cap B)$ into eqn. (2)

We then have

$$p(A/B) = \frac{p(A).P(B)}{p(B)} \tag{4}$$

Furthermore, we want to use the conditional probability to determine the space or hole each band here HF (High frequency) and VHF (Very High Frequency) its occupy in the total spectrum space allocated based on the primary users.

From eqn. (2), let A represents, probability of primary in HF (High Frequency) and B represents probability of primary in VHF (Very High Frequency).

Where

A=0.00004mhz and

B=0.000008mhz

From conditional probability, we let A and B be two events defined on a sample space say *S.* the conditional probability of A given B is given

$$p\left(\frac{A}{B}\right) = \frac{0.00004*0.000008}{0.000008} = 0.00004mhz$$

Therefore, the probability of primary space allocated for HF and VHF in spectrum allocation table 1 which gives=0.00004mhz.

Again, considering secondary under secondary allocations under High Frequency (HF) and Very High Frequency (VHF) such that;

A represent Secondary in HF=0.002mhz

B represent Secondary in VHF=0.006

From eqn. (2) we determine their conditional probabilities for the secondary users in the spectrum allocation.

$$p\left(\frac{A}{B}\right) = \frac{0.002*0.006}{0.006} = 0.002mhz$$

Therefore, the probability of secondary space allocated for HF and VHF in spectrum allocation table 1 which also gives=0.002mhz

Nevertheless, based on the conditional probability of primary and secondary users for both HF band and VHF, results shows that there were numerous holes that was left unutilized in the primary allocation which was reschedule for the secondary users which makes the secondary users to occupy a probability of 0.002mhz has compared to the primary which was 0.00004mhz utilization. This work adopted Decision Tree Algorithm for the classification technique of the spectrum data sets based on the ranks.

5.1. Decision Tree Algorithm in Classification

A decision tree is a structure that includes a root node, branches, and leaf nodes. Each internal node denotes a test on an attribute, each branch denotes the outcome of a test, and each leaf node holds a class label.

Basically, Decision Tree is a supervised Machine learning algorithm. The Decision tree algorithm is typically used for solving classification and regression problems**.** The algorithm is developed using a heuristic process known as recursive partitioning as well as divide and conquer approach. It splits the data into components or subsets, subsequently split repeatedly into even smaller subsets, until the procedures ends when the algorithm determines the data within the subsets to be very sufficiently homogenous. Decision Tree is required to develop a training method to be used in prediction of the class or value of the target variable through learning simple decision rules inferred from training data. Again, classification is commonly use approach in machine learning usually considered as a two-step procedure for learning and prediction. Often, the algorithm for classification is usually developed based on the given training data sets. The basic assumptions of Decision tree Algorithm are:

i. In the beginning, the whole training set is considered as the root.
ii. Feature values are preferred to be categorical. If the values are continuous then they are discretized prior to building the model.
iii. Records are distributed recursively on the basis of attribute values.
iv. Order to placing attributes as root or internal node of the tree is done by using some statistical approach.

Basic components of a Decision Tree

The list of elements that constitute Decision tree are as follows:

a) Nodes: This indicates the point where the tree splits according to the value of some attribute/feature of the dataset.
b) Edges: It usually directs the outcome of a split to the next node we can see in the figure above that there are nodes for features like outlook, humidity and windy. There is an edge for each potential value of each of those attributes/features.
c) Root: This is actually the node where the first split takes place.
d) Leaves: It consist of the terminal nodes that predict the outcome of the decision tree.

Consequently, once a decision tree is developed, many nodes represent outliers or noisy data. Pre-processing approach is applied such as Tree pruning method to remove unwanted data. This, in turn, improves the accuracy of the classification model. To determine the accuracy of the model, a test set consisting of test tuples and class labels is adopted. The percentages of the test set tuples are correctly classified by the model to identify the accuracy of the model. Therefore, if the model is found to be optimally accurate then it is used to classify the data tuples for which the class labels are not known.

5.2. Training Set

Training data set in machine learning is the actual dataset used to train model for performing various actions. This is the actual data the ongoing development process models learn with various API and algorithm to train the machine to work automatically. In machine learning, one task is the studying and construction of algorithms that can learn from and make predictions on data. Such algorithms work by making data-driven predictions or decisions, through building a mathematical model from input data. In this research work our training set is derived from a spectrum allocation (Spectrum open data, 2016). Table 3 shows the training sets.

Table 2. *Training Sets for Spectrum, n=69 (Source: Opendata [14].*

Band	Allocation	AL Freq from (MHz)	AL Freq to (MHz)	AL Freq Range (MHz)	Rank
HF	SR	2.501	2.502	0.001	Secondary
HF	EES	432	438	6	Secondary
HF	SR	5.003	5.005	0.002	Secondary
HF	SR	10.003	10.005	0.002	Secondary

Band	Allocation	AL Freq from (MHz)	AL Freq to (MHz)	AL Freq Range (MHz)	Rank
HF	RA	13.36	13.41	0.05	Primary
VHF	SR	30.005	30.01	0.005	Primary
VHF	RA	37.5	38.25	0.75	Secondary
VHF	SR	39.986	40.02	0.034	Secondary
VHF	SR	40.98	41.015	0.035	Secondary
VHF	SR	137	137.025	0.025	Secondary
P	MS	399.9	400.05	0.15	Primary
P	MS	400.15	401	0.85	Primary
P	SR	400.15	401	0.85	Primary
P	SR	400.15	401	0.85	Secondary
P	SR	401	406	5	Primary
L	RS	1215	1240	25	Primary
L	RS	1215	1240	25	Primary
L	EES	1215	1240	25	Primary
L	SR	1215	1240	25	Primary
L	EES	1240	1260	20	Primary
S	MS	1980	2010	30	Primary
S	MS	1980	2010	30	Primary
S	SO	2025	2110	85	Primary
S	EES	2025	2110	85	Primary
S	SR	2025	2110	85	Primary
C	FS	3600	4200	600	Primary
C	FS	3600	4200	600	Primary
C	FS	4500	4800	300	Primary
C	RA	4800	4990	190	Secondary
C	RA	4990	5000	10	Primary
X	FS	7250	7300	50	Primary
X	FS	7300	7450	150	Primary
X	FS	7450	7550	100	Primary
X	MS	7450	7550	100	Primary
X	FS	7550	7750	200	Primary

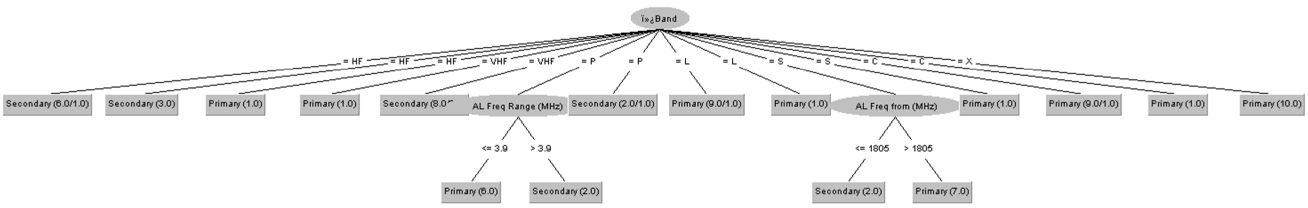

Figure 3. *Weka J48 Decision Tree classifier.*

Table 3. *Build Model Information.*

===Run information===	
Scheme:	weka. classifiers. trees. J48 -C 0.25 -M 2
Relation:	training data
Instances:	69
Attributes:	6
	ï»¿Band
	Allocation
	AL Freq from (MHz)
	AL Freq to (MHz)
	AL Freq Range (MHz)
	Rank
Test mode:	10-fold cross-validation

===Classifier model (full training set)===
J48 pruned tree

ï»¿Band=HF: Secondary (6.0/1.0)
ï»¿Band=HF: Secondary (3.0)
ï»¿Band=HF: Primary (1.0)
ï»¿Band=VHF: Primary (1.0)
ï»¿Band=VHF: Secondary (8.0/2.0)
ï»¿Band=P

===Classifier model (full training set)===
J48 pruned tree

AL Freq Range (MHz) <=3.9: Primary (6.0)
AL Freq Range (MHz) > 3.9: Secondary (2.0)
ï»¿Band=P: Secondary (2.0/1.0)
ï»¿Band=L: Primary (9.0/1.0)
ï»¿Band=L: Primary (1.0)
ï»¿Band=S
AL Freq from (MHz) <=1805: Secondary (2.0)
AL Freq from (MHz) > 1805: Primary (7.0)
ï»¿Band=S: Primary (1.0)
ï»¿Band=C: Primary (9.0/1.0)
ï»¿Band=C: Primary (1.0)
ï»¿Band=X: Primary (10.0)
Number of Leaves: 16
Size of the tree: 19

Time taken to build model: 0.06 seconds		
===Stratified cross-validation===		
===Summary===		
Correctly Classified Instances	47	68.1159%
Incorrectly Classified Instances	22	31.8841%
Kappa statistic		0.1812
Mean absolute error		0.3747
Root mean squared error		0.494
Relative absolute error		87.901%
Root relative squared error		107.2341%
Total Number of Instances		69

===Detailed Accuracy by Class===								
TP Rate	**FP Rate**	**Precision Recall**	**F-Measure**	**MCC**	**ROC Area**	**PRC Area**	**Class**	
0.333	0.167	0.467	0.333	0.389	0.186	0.610	0.397	Secondary
0.833	0.667	0.741	0.833	0.784	0.186	0.610	0.761	Primary
Weighted Avg.	0.681	0.514	0.657	0.681	0.664	0.186	0.610	0.650

===Confusion Matrix===		
a	b	<-- classified as
7 |	14	a=Secondary
8 |	40	b=Primary

6. Model Performance Evaluation

Basically, there are two approaches to examine the performance of classifiers: Confusion matrix, and to use a ROC graph.

Given a class, C_j, and a tuple, t_i, that tuple may or may not be assigned to that class while its actual membership may or may not be in that class. With two classes, there are four possible outcomes with the classification as:

True Positives: this is the rate instances correctly classified as a given class.

False Positives: this is the rate of instances falsely classified as a given class.

True Negatives: is an outcome where the model correctly predicts the positive class.

False Negatives: is an outcome where the model incorrectly predicts the negative class.

A confusion matrix contains information about actual and predicted classifications. Performance is evaluated using the data in the matrix. In this research work, a total of 69 instances roughly split evenly between two classes (i.e Primary and Secondary) was use as our training sets in order to formulate the model.

Confusion Matrix

A confusion matrix is a summary of prediction results on a classification problem. The number of correct and incorrect predictions are summarized with count values and broken down by each class. This is key to the confusion matrix. The confusion matrix shows the ways in which our classification model is confused when it makes predictions. It gives us insight not only into the errors being made by a classifier but more importantly the types of errors that are being made

Table 4. *Confusion Matrix Observed.*

	Primary	**Secondary**
Primary	TP	TN
Secondary	FP	FN

Table 5. *Confusion Matrix of our model.*

	Primary	**Secondary**
Primary	7	14
Secondary	8	40

Furthermore, Figure 4 depicts the classification of primary users and secondary users of spectrum, primary users were given much allocation than the secondary users where the secondary users are the opportunistic users who are not permanently given space, but rather the occupy space in the spectrum when there are available holes in the primary allocations.

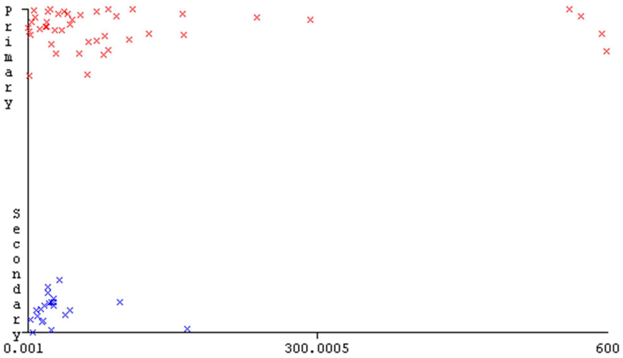

Figure 4. *Primary vs Secondary Users.*

Again figure 5 depicts allocations based on different space satellite allocation based on primary and secondary users Space Research (SR) satellite seems to be allocated with the highest secondary space.

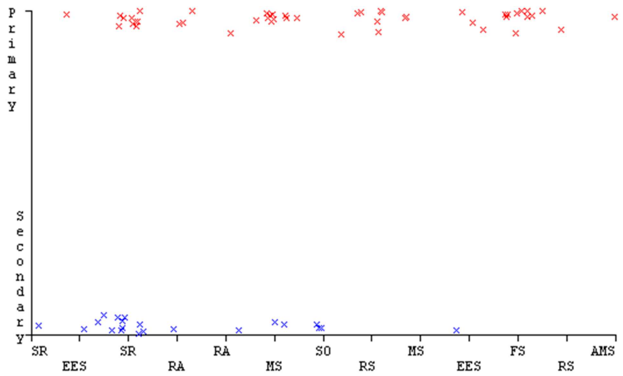

Figure 5. *Satellite vs Users.*

Also, figure 6, depicts frequency band and users, only X band in the figure that was not allocated and space for the secondary users meaning that all the space that was allocated to x band was utilized.

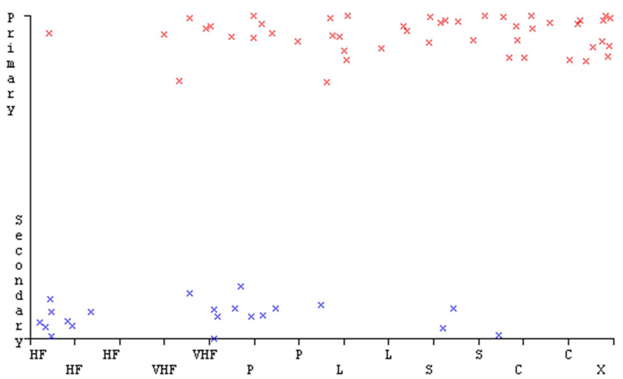

Figure 6. *Band vs Users.*

Furthermore, figure 7 below in the X plane represents predicted classifier results, the Y plane represents actual classifier results. Squares represent wrongly classified samples. Stars represent true classified samples, so the figure 4.8 illustrate that, 8 instances of primary and 14 instances of secondary users were incorrectly classified on our test data sets.

Figure 7. *Band vs Users.*

Figure 8, shows the probabilities of each individual frequency users in mega hertz's C-band had the highest probability of space allocated to it based on primary usage has compared to all other frequency bands.

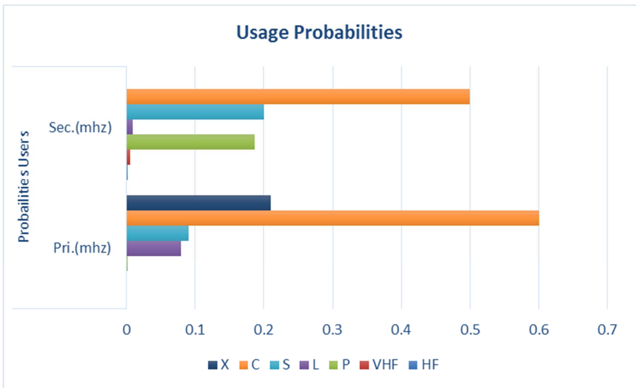

Figure 8. *Usage Probabilities.*

7. Conclusion

Spectrum Scheduling is an efficient scheme of improving spectrum utilization for faster communications, higher definition media (HDM) and data transmission. Radio spectrum is very limited in supply resulting in enormous problems related to scarcity. It owes the physical support for wireless communication, both fixed applications and mobile broadband. Nevertheless, effective use of the spectrum depends on the channel settings, sensing performance, detection of spectrum prospect as well as effective transmission of both Primary Users (PUs) and Secondary Users (SUs) packets at a specific time slot. Spectrum scarcity can be addressed from many different perspectives. The main goal of spectrum allocation is to assign leisure spectrum resources efficiently to achieve the optimal state under the requirements of cognitive user and network Quality of

Service (QOS). In this paper, classification of spectrum allocation was carried out using probability theorem to identify the probability of both primary and secondary users in the allocated spectrum data sets. Again, conditional probability was used to compare the two-frequency band based on primary and secondary allocation policies to identify the specific allocation of each band. Consequently, Machine Learning (ML) Algorithm based on Decision Tree - Supervised Learning approach was adopted to classified the data sets which yielded 68% results. This evidence indicated correctly classified instances based on the total records of 69. The results demonstrate that the improved algorithm achieves better performance and optimized spectrum scheduling for efficient networks service provisioning.

Author Contributions

All authors have contributed substantially to the work reported.

Conflicts of Interest

The authors declare that they have no competing interests.

References

[1] Anita G., Partha P., Bhattacharya (2011). Dynamic Spectrum Access in Cognitive Radio: a brief review, International Journal of Computer Application in Engineering Sciences, Special Issue on Computer Networks & Security, pp 149-153.

[2] Akyildiz, Won-Yeol L, Vuran M., Mohanty S. (2006). Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey. Elsevier B. V.

[3] Cisco (2018). Cisco Predicts More IP Traffic in the Next Five Years Than in the History of the Internet, Nov. 2018.

[4] Fcc, (2020), Federal communication commission "FCC online table of frequency allocations," Tech. Rep., 2020.

[5] GSMA P. (2017). Introducing spectrum management (Spectrum primer series). Floor 2, The Walbrook Building, 25 Walbrook, London EC4N 8AF 020 7356 0600.

[6] Habibzadeh A. (2017). Improvement of handover in femtocell-based cellular cognitive radio network [Ph.D. Dissertation]. Iran: Shahid Rajaee Teacher Training University; (In Farsi).

[7] Habibzadeh A, Shirvani Moghaddam S (2015). Noise calibrated GLRT-based spectrum sensing method for cognitive radio applications. In: 15th IEEE International Symposium on Signal Processing and Information Technology; Abu Dhabi, UAE; 7-10 Dec. 2015. pp. 174-179.

[8] Imeh J. Umoren and Saviour J. Inyang (2021). Methodical Performance Modelling of Mobile Broadband Networks with Soft Computing Model. *International Journal of Computer Applications* 174 (25): 7-21, NY, USA.

[9] Imeh J. Umoren and Samuel B. Okon, (2021). A Multidimensional Fuzzy Knowledge-based System for Optimizing Wireless Local Area Networks Performance, International Journal of Computer Applications *(0975–8887),* 183 (1): 8 19, New York, USA.

[10] Li Zhang, Kai Z, Prasant M. (2010). Opportunistic Spectrum Scheduling for Mobile Cognitive Radio Networks in White Space. Computer Science Department University of California, Davis, CA, USA.

[11] Linda E. (2009). "Essentials of Cognitive Radio", New York: Cambridge University Press.

[12] Mitola J, Maguire G. (1999). Cognitive radio: Making software radios more personal. IEEE Personal Communications, vol 6 (4): pp 13-18.

[13] Mansi S. and Gajanan B. (2011), "Spectrum Sensing Techniques in Cognitive Radio Networks: A Survey", International Journal of Next Generation Networks, Vol. 3, No. 2.

[14] Opendata, (2016). https://www.ofcom.org.uk/research-and data/data/opendata.

[15] Rosston G (2014). Increasing the efficiency of spectrum allocation, Springer's Review of Industrial Organization, vol. 45, no. 3, pp. 221–243.

[16] Raouia M., E. V. Belmega, Inbar F. (2016). EURASIP Journal on Wireless Communications and Networking Efficient spectrum scheduling and power management for opportunistic users. Springer Nature.

[17] Salim A. Hanna (2011), Spectrum metrics for 2.4GHz ISM Band Cognitive Radio Applications, IEEE 22nd International Symposium on Personal, Indoor, and Mobile Radio Communications.

[18] Tilghman. P (2019). Will rule the airwaves: A darpa grand challenge seeks autonomous radios to manage the wireless spectrum, IEEE Spectrum, vol. 56, no. 6, pp. 28–33.