International Journal of Future Generation Communication and Networking
Vol. 13, No. 3, (2020), pp. 1249–1257

Bandwidth Comparison Model for Future Internet Using Machine Learning

Swapnil Raghunath Pokharkar[1] , Dr. Sanjeev J. Wagh[2], Dr. Sachin N. Deshmukh[3]
[1] *Department of Computer Science and IT,Babasaheb Ambedkar Marathwada University, Aurangabad, India ,swapnil223469@gmail.com*
[2] *Department of Information Technology,Government College of Engineering, Karad, India, sjwagh1@yahoo.co.in*
[3] *Department of Computer Science and IT,Babasaheb Ambedkar Marathwada University, Aurangabad, India ,sndeshmukh@bamu.ac.in*

Abstract

Internet bandwidth is an element of communication system. To speed up any wireless communication system internet bandwidth plays a vital role. Internet bandwidth also effect on economy of nation. Various internet bandwidth techniques can be used to improve wireless communication for high data transmission. Bandwidth comparison model provide requirement of bandwidth in rural and urban area. Bandwidth comparison model can be implemented by using different machine learning algorithm. Wireless communication and internet bandwidth internally related to each other. Bandwidth prediction model also specify variation in rural bandwidth requirement and urban bandwidth requirement. The prediction of spectrum utilization has already been studied by some research groups. In this paper, we reviewed the India's most usable communication network for bandwidth comparison and bandwidth measurement for purpose of bandwidth prediction. In this paper we compared bandwidth requirement in 2018 and 2020.

Keywords: *Spectrum Management, Artificial Neural Network, Bandwidth prediction, Spectrum Utilization*

1. INTRODUCTION

Bandwidth prediction has a great impact in the field wireless communication. Bandwidth requirement varies in urban area and rural area. The word spectrum is also used for communication which indicates frequency. Frequency range 3 kHz to 300 GHz used for wireless communication. Need of connecting devices increases which effect on requirement of bandwidth. Necessity of bandwidth prediction increases due to innovation of 4G and 5G services. Cognitive radio can be used in wireless communication. [1] [2] Primary user include television band [3] [4] [5] [6] cellular band[7] [8] and ultra-high frequency band. . We need efficient model which define the bandwidth requirement. In this paper we focus on the machine learning techniques used for bandwidth prediction. We are using India's well-known network for bandwidth prediction. We analyze current bandwidth requirement based on that we predict future bandwidth requirement for both rural and urban areas

2. MACHINE LEARNING IN PREDICTION

Machine learning algorithms are divided into supervised learning and unsupervised learning. Regression and classification problems are in supervised learning. Classification and conclusion is done by using values which are obtained by observation. If we considered output variable as 'y' which is approximated by this problem using a mapping function say 'f' and considered input variable as 'x'. The output of classification is generally in discrete form but it can also continuous for every class label in the form of

ISSN: 2233-7857 IJFGCN

probability. A regression problem has output variable as a real value or continuous value. The output of regression is generally continuous but it can also be discrete for any class label. A problem with many output variables is referred to multivariate regression problem. In Artificial Intelligence, Machine learning plays essential part providing the capability of learning and improving by itself. Machine learning methods are mainly categorized in three common types such as supervised learning Reinforcement learning and Unsupervised learning [23].

Let us look at some of the standard ways to optimize a machine learning algorithm [22].

- Unconstrained Optimization:
  - Gradient Decent
  - Newton's Method
  - Batch Learning
  - Stochastic Gradient Decent

- Constrained Optimization:
  - Lagrange Duality
  - Constrained Methods

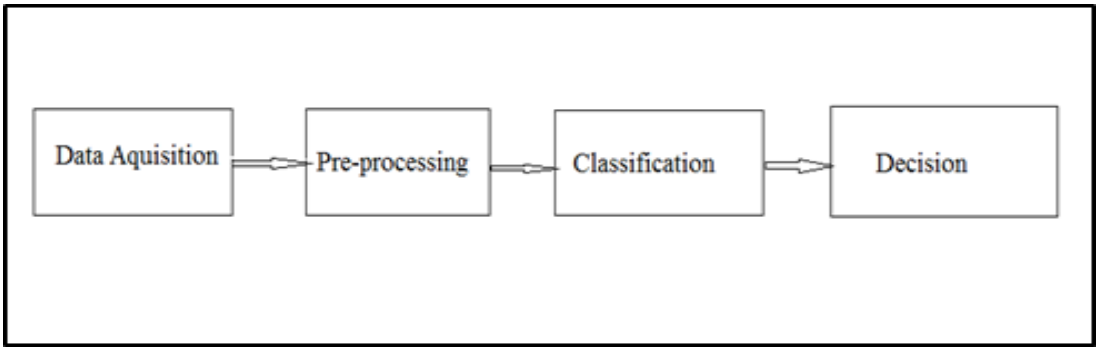

Fig.1 Task Performed for Prediction in ANN

Fig.1 shows various tasks performed in ANN. Data acquisition represents selection of dataset. Pre-processing contains data cleaning and data processing. Classification refers to applying algorithm. Decision contains final steps of prediction. This paper focuses on problem regarding bandwidth prediction. Internet bandwidth predicted using machine learning method. It can be classified as regression problem and problem can be comes under supervised learning.

3. RELATED WORKS

There are various types and variants of Computational Intelligence and machine learning algorithms that can be used in cognitive radio such as genetic algorithms for optimization of transmission parameters [9], swarm intelligence for development of radio resource allocation [10], fuzzy logic system for decision making [11] [12], neural network and hidden Markov model for prediction of spectrum holes game theory, linear regression and linear predictors for spectrum occupancy prediction [13] etc. Some of the methods are used for learning and prediction, some for optimization of certain transmission parameters while others for decision making [14]. TV idle channels prediction using ANN[15], however, data were collected only for two hours everyday day (5pm to 7pm) within a period of four weeks, this is not sufficient to capture all the various trends associated with TV broadcast. Also, recognize the idle channels does not represent any spatial or temporary data of the anticipate noise and level of interference depends on the channels history which is vital in selecting the channels. Spectrum hole prediction using Elman

recurrent artificial neural network was present in cognitive radio [16]. It uses the cyclo-stationary features of modulated signals to decide the presence or absence of primary signals while the input of the Elman recurrent artificial neural network contain of time instances. The input and target output used in training of the Elman recurrent artificial neural network [17] and prediction were modeled using ideal multivariate time series equations, which are often different from day to day life.

Techniques	Response delay	Energy consumption	Inference	Network Throughput
Artificial Neural Network	Less	Less	Minimum	Average
Hidden Markov Model	Minimum	Low	Minimum	Average
Multilayer Perceptron Neural Network Based Prediction	High	Very low	Minimum	Average
Bayesian Inference based Prediction	Low	Low	Minimum	Very high
Moving Average Based Prediction	Low	Low	Minimum	Good
Static Neighbor Graph Based Prediction	High due to network	High	High	Average
Channel State Prediction	High	High	High	Low

TABLE. I COMPARISON OF VARIOUS PREDICTION MODELS

Table.1 [21] shows comparison of various prediction models which are used in recent time. The comparison based on response delay, energy consumption, inference and network throughput.

4. SURVEY ON INTERNET TRAFFIC IN INDIA

Number of telephone user in India increased from 1186 M at end of June-2019 to 1195 M at end of September-2019.Where growth rate become 0.73%. This indicate year on year growth rate of 0.32% over last year in same quarter. Telephone user in urban area increased from 675 M at end of June-2019 to 677 M at end of September-2019.Rural telephone user increased from 511 M at end of June-2019 to 517 M at end of September-2019. Rural subscription increased from 43% to 43.3% at end of September-2019.

Total wireless subscriber increased at end of September-2019 with growth rate of 0.71% over previous quarter. Wireless user increased year one year at rate 0.38%.wireline user decrease by 2.8% in quarter September 2019.total internet user increased at end of September-2019 with growth rate 3.35%.[18]

5. ARTIFICIAL NEURAL NETWORKS (ANN'S)

An ANN is set of inter-connected units or nodes called artificial neurons, which like as model neurons in a brain. Each connection, like the synapses in a brain, can transmit a signal to other neurons. Artificial neurons that receive a signal then proceed on it. Artificial Neural Networks are collection of simple elements operating in parallel. These elements are inspired by biological nervous systems [19], [20]. Human beings have a capacity of understand patterns within the accessible details with high degree of accuracy. Artificial neural networks are learning systems that offer ability like human learning capabilities via a difficult architecture that resembles the human nervous system. Artificial neural networks are inspired by the human neural network structure. The simplest neural network consists of only one neuron and it is called a perceptron.
Fig.2 shows the different types of layers in ANN. Input layers represent the various input features taken from dataset. Hidden layers represent internal working of algorithm. Prediction layer is responsible to generate predicted output. Output layer represents the output for the prediction model. We have these units of calculation called neurons. These neurons are inter-connected by synapses which are really just weighted values. Neuron performs some sort of calculations, and then the result of this calculation will be multiplied by a weight. The weighted result can sometimes be the output of your neural network, or you can have more neurons configured in layers, this is the basic idea that we call deep learning.

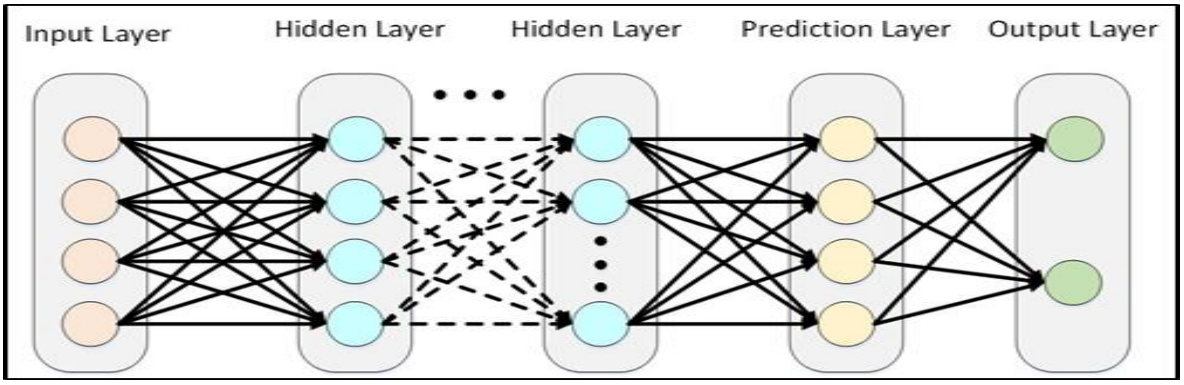

Fig.2 Layers in Artificial Neural Network

6. EXPERIMENT DETAILS AND NUMERICAL SIMULATIONS

Spectrum.csv is used as a dataset where each row provides information on India's most usable Cellular Network. it also contains uploaded and downloaded data with CellName, LAC, CellID, Totalul, Totaldn, totaldata etc. Some of the attributes are explained as follows:

A. Cell Name:

The cell name attribute is used to describe cell name in Pune region. It includes various rural and urban area cells which describes unique identity of particular cell and differentiate it from others.

B. LAC:

The LAC attribute is used to describe Location Area Code in Pune region. It includes various rural and urban location area codes which describes unique identity of particular area and differentiate it from others

C. Cell ID:
The cell ID attribute is used to describe identity of cell in Pune region. It includes various rural and urban area cells which describes the unique identity of particular cell and differentiate it from others.

D. Total Download:
The total download attribute is used to describe the total downloaded data of particular cell. It includes downloaded data for various cells in Pune region for both urban and rural areas.

E. Total Upload:
The total upload attribute is used to describe total uploaded data of particular cell. It includes uploaded data of various cells in Pune region for both urban and rural areas.

F. Total Data:
The total data attribute is used to describe downloaded and uploaded data of cell. It includes the summation of uploaded and downloaded data of various cells in Pune region for both urban and rural areas.
We have performed the experiment in python programming language. We have used Pandas, NumPy, matplotlib, seaborn, sklearn python libraries for solving the problem.

- Machine Learning Packages Used for Prediction:
  - Pandas: It is an open source popular library which allows data analysis in python.
  - NumPy: It is a python package which is popular library for scientific computing.
  - Scikit-learn: It is a free machine learning package of python. It uses various machine learning algorithms.
  - Matplotlib: It is a Python plotting library which generates quality figures in different hardcopy formats.

Mathematical Model
X indicate Location Area Code and Y indicate values as uploaded or downloaded data in GB.
Passing X-axis values and Y-axis values to plot ( ) function
Plot ('X-axis values', 'Y-axis values', 'data=obj')
Line equations for downloaded data
Y= -7.39X+7801.1 for year 2020
Y= -6.44X+6684.1 for year 2018
Line equations for uploaded data
Y= -4.55X+4752.9 for year 2020
Y= -6.44X+6684.1 for year 2018
Apply show ( ) function to draw graph
Displaying comparison of graph

Algorithm:
Step-1: Pick a Network Architecture (i.e. Connectivity Pattern between the neurons)
Step-2: Random Initialization of the Weights
Step-3: Implement Forward Propagation method to get initial prediction value for any x(i)
Step-4: Implement Cost Function for Artificial Neural Network
Step-5: Implement back-propagation function to calculate partial derivatives
Step-6: Use gradient checking to ensure about your back-propagation works. Then disable gradient checking.
Step-7: Use gradient descent function or a built-in optimization function to reduce cost function with weights in theta.
Step-8: Final Decision.
Step-9: Checking accuracy of prediction model.

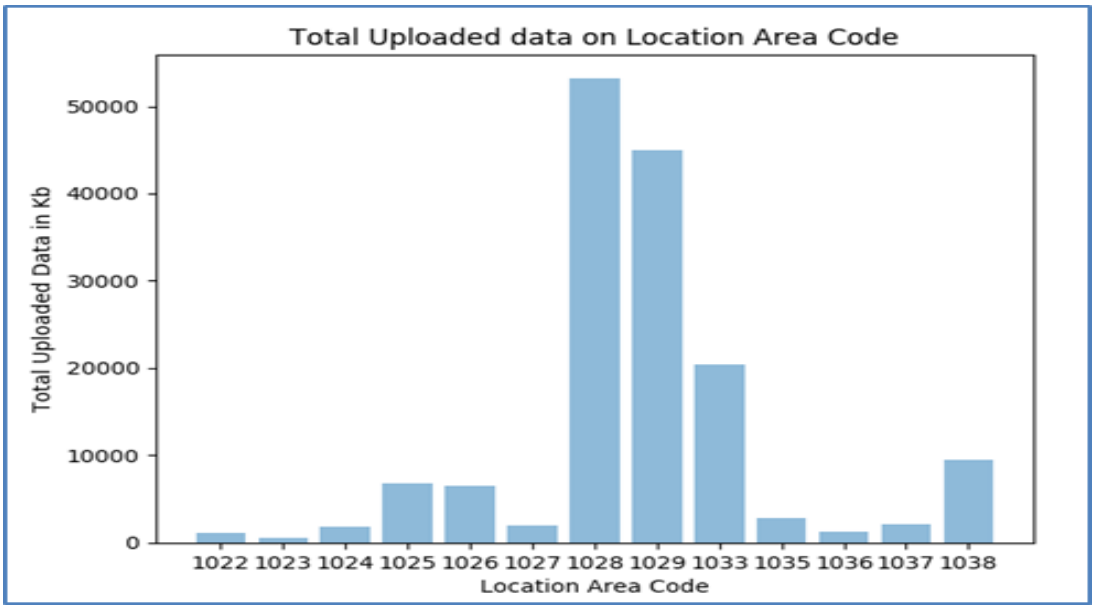

Fig.3 Uploaded Data on Location Area Code

Fig.3 shows total uploaded data on each Location Area Code. These are some Location Area Code in datasets and show the quantity of uploaded data on Location Area Code. X-axis shows Location Area Code and y-axis shows total data uploaded in KB. We can see there is a huge variation in uploaded data of each Location Area Code. Because some Location Area Code are used in rural areas and some are used in urban areas. The bandwidth of Location Area Code used in urban areas is higher than the bandwidth of Location Area Code used in the rural area. So, it proves that necessary bandwidth of Location Area Code in urban areas is higher than the bandwidth of Location Area Code in rural areas due to higher use of network in urban areas.

ISSN: 2233-7857 IJFGCN

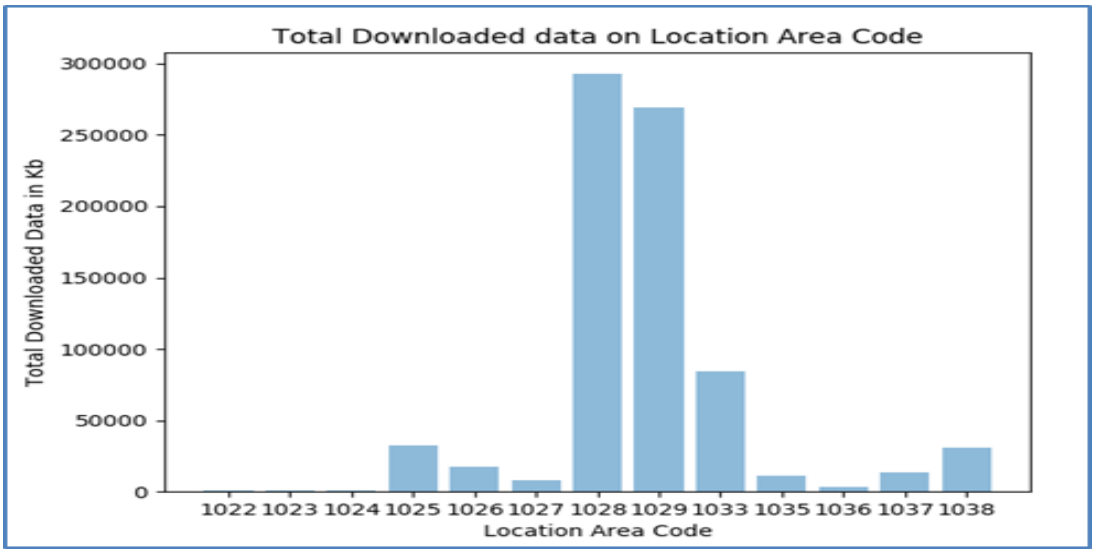

Fig.4 Downloaded Data on Location Area Code

Fig.4 shows total downloaded data on each Location Area Code. These are some Location Area Code in datasets and show the quantity of downloaded data on Location Area Code. X-axis shows a location area code and y-axis shows total downloaded data in KB. We can see there is a huge variation in downloaded data of each Location Area Code. Because some Location Area Code are used in rural areas and some are used in urban areas. The bandwidth of Location Area Code used in urban areas is higher than the bandwidth of Location Area Code used in the rural area. So, it proves that the necessary bandwidth of Location Area Code in urban areas is higher than the bandwidth of Location Area Code in rural areas due to higher use of network in urban areas.

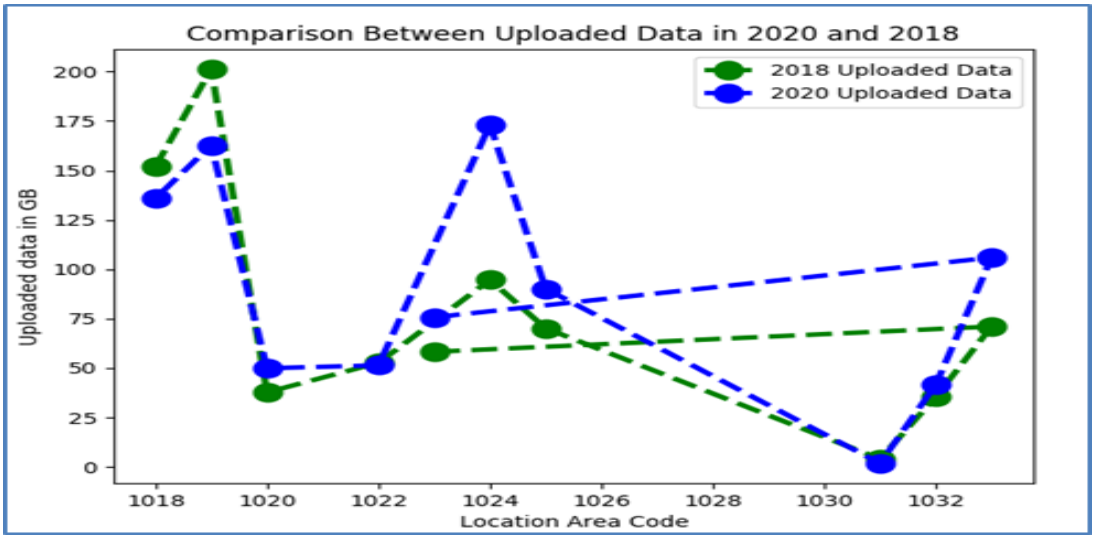

Fig.5 Comparison between Uploaded Data 2020 and 2018

Fig.5 shows comparison between uploaded data in 2020 and uploaded data in 2018. Blue line shows 2020 uploaded data and green line shows 2018 uploaded data. The comparison is done based on the Location

Area Code and Uploaded Data in GB. We can see the amount of uploaded data in 2020 is higher than the amount of uploaded data in 2018. In this scenario no of user for particular network plays vital role while predicting bandwidth requirement.

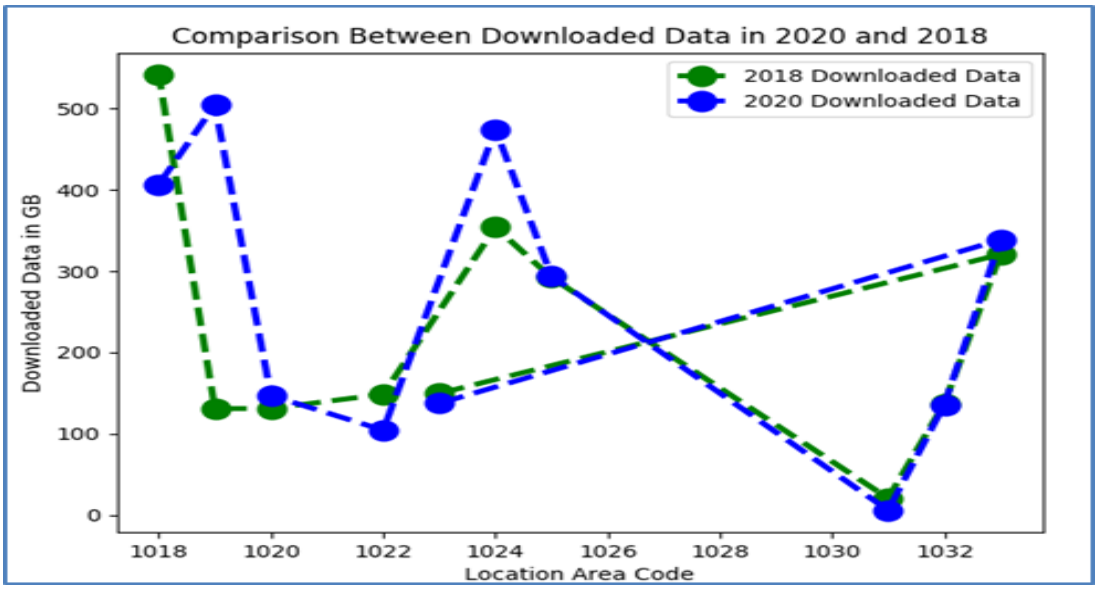

Fig.6 Comparison between Downloaded Data in 2020 and 2018

Fig.6 shows comparison between downloaded data in 2020 and downloaded data in 2018. Blue line shows 2020 downloaded data and green line shows 2018 downloaded data. The comparison is done based on the Location Area Code and Downloaded Data in GB. We can see the amount of downloaded data in 2020 is higher than the amount of downloaded data in 2018. In this scenario no of user for particular network plays vital role while predicting bandwidth requirement.

7. CONCLUSION

In this paper, we reviewed how spectrum gets managed and various types of wireless infrastructure prediction carried out by different research groups. For spectrum prediction artificial neural network is one of the most suitable machine learning prediction algorithm. We reviewed some statistical functions used for prediction of spectrum. We analyze dataset by using graphs and conclude requirement of bandwidth in urban and rural area. According to comparison between uploaded data and downloaded data in 2020 and 2018 we can see bandwidth requirement increases in 2020. So, we can conclude that bandwidth requirement is higher in 2020 than 2018. Network plays vital role while predicting bandwidth requirement. There are still many difficulties and challenges for the prediction method, which will be the direction of our future research.

References

[1] Hossain E, Niyato D, Kim DI (2015) Evolution and future trends of research in cognitive radio: a contemporary survey. Wirel Commun Mobile Comput 15(11):1530–1564

[2] Liang Q, Leung VCM, Meng W, Adachi F (2014) Cooperative communications and sensing. Wirel Commun Mobile Comput 14(13):1201–1203

[3] Nekovee M (2009) Quantifying the availability of TV white spaces for cognitive radio operation in the UK. Communications workshops, 2009. In: ICC Workshops 2009. IEEE International Conference 2009; pp 1–5
[4] Jayavalan S, Mohamad H, Aripin NM, Ismail A, Ramli N, Yaacob A, Ng MA (2014) Measurements and analysis of spec-trum occupancy in the cellular and TV bands. Lect Notes Softw Eng 2(2):133–138
[5] Martin FL, Correal NS, Ekl RL, Gorday P, O'Dea R (2008) Early opportunities for commercialization of TV whitespace in the US. In: Cognitive Radio Oriented Wireless Networks and Communications, 2008. CrownCom, pp 1–5
[6] Petty VR, Rajbanshi R, Datla D, Weidling F, DePardo D, Kolodzy PJ, Marcus MJ, Wyglinski AM, Evans JB, Minden GJ, Roberts JA (2007) Feasibility of dynamic spectrum access in underutilized television bands. In: New frontiers in dynamic spectrum access networks, pp 331–339
[7] López-Benítez M, Umbert A, Casadevall F (2009) Evaluation of spectrum occupancy in Spain for cognitive radio applications. In: Vehicular Technology Conference, 2009. VTC Spring, pp 1–5
[8] Xue J, Feng Z, Zhang P (2013) Spectrum occupancy measurements and analysis in Beijing. IERI Procedia 4:295–302
[9] T. W. Rondeau, B. Le, C. J. Rieser, and C. W. Bostian, "Cognitive radios with genetic algorithms: Intelligent control of software defined radios," in c 2004 SDR Forum, Proceeding of the SDR 2004 Technical Conference and Product Exposition, 2004.
[10] S. K. Udgata, K. P. Kumar, and S. L. Sabat, "Swarm intelligence-based resource allocation algorithm for cognitive radio network," in Parallel Distributed and Grid Computing (PDGC), 2010 1st International Conference on, Oct 2010, pp. 324–329.
[11] M. Matinmikko, J. Del Ser, T. Rauma, and M. Mustonen, "Fuzzy-logic based framework for spectrum availability assessment in cognitive radio systems," Selected Areas in Communications, IEEE Journal on, vol. 31, no. 11, pp. 2173–2184, November 2013.
[12] L. Giupponi and A. Perez, "Fuzzy-based spectrum handoff in cognitive radio networks," 2008.
[13] Y. Chen and H.-S. Oh, "A survey of measurement-based spectrum occupancy modelling for cognitive radios," in 1553-877X 2013 IEEE IEEE Communications Surveys and Tutorials, 2013.
[14] R. Azmi, "Support vector machine based white space predictors for cognitive radio," Master's thesis, 2011.
[15] O. Winston, A. Thomas, and W. OkelloOdongo, "Optimizing neural network for tv idle channel prediction in cognitive radio using particle swarm optimization," in Computational Intelligence, Communication Systems and Networks (CICSyN), 2013 Fifth International Conference on, June 2013, pp. 25–29.
[16] M. I. Taj and M. Akil, "Cognitive radio spectrum evolution prediction using artificial neural networks based mutivariate time series modelling," in European Wireless, Vienna Austria, April 2011.
[17] X. Li and S. A. Zekavat, "Traffic pattern prediction and performance investigation for cognitive radio systems," in IEEE Communication Society, WCNC Proceedings, 2008.
[18] "The Indian Telecom Services Performance Indicators", July-September 2019.
[19] S. Haykin, Neural Networks and Learning Machines, 3rd ed., 2008.
[20] Z. Jianli, "Based on neural network spectrum prediction of cognitive radio," in 978-1-4577-0321-8/11/$26.00 2009 IEEE, 2009.
[21] Swapnil Pokharkar, Dr. Sanjeev Wagh, Dr.. Sachin Deshmukh "comparison of statistical prediction model for spectrum utilization" IJCNWC vol7, No 5 Sep-Oct 2017.
[22] Kunal Pahwa, Neha Agarwal "Stock Market Analysis using Supervised Machine Learning" International Conference on Machine Learning, 2019.
[23] Nan Zheng, Pinaki Mazumder, "Fundamentals and Learning of Artificial Neural Networks", wileyIEEEpress 2020.