International Journal of Communication Networks and Information Security
ISSN: 2073-607X, 2076-0930
Volume 14 Issue 02 Year 2022 Page 244:260

A Multi-dimensional Real World Spectrum Occupancy Data Measurement and Analysis for Spectrum Inference in Cognitive Radio Network

[1]Mudassar H. Naikwadi, [2]Kishor P. Patil

[1]Research Scholar, [2]Professor, Department of E&TC, Sinhgad Academy of Engineering, Kondhwa (Bk), Pune, Maharashtra, India
[1] mhnaikwadi@gmail.com, [2] kppatil1@gmail.com

Article History

Received: 26 June 2022
Revised: 28 July 2022
Accepted: 30 August 2022

Abstract

Spectrum Inference in contrast to Spectrum Sensing is an active technique for dynamically inferring radio spectrum state in Cognitive Radio Networks. Efficient spectrum inference demands real world multi-dimensional spectral data with distinct features. Spectrum bands exhibit varying noise floors; an effective band wise noise thresholding guarantees an accurate occupancy data. In this work, we have done an extensive real world spectrum occupancy data measurement in frequency range 0.7 GHz to 3 GHz for tele density wise varying locations at Pune, Solapur and Kalaburagi with time diversity ranging from 2 to 7 days. We have applied maximum noise (Max Noise), m-dB and probability of false alarm (PFA) noise thresholding for spectrum occupancy calculations in all bands and across all locations. Overall occupancy across these locations is 37.89 %, 18.90 % and 13.69 % respectively. We have studied signal to noise ratio (SNR), channel vacancy length durations (CVLD) and service congestion rates (SCR) as characteristic features of measured multi-dimensional spectrum data. The results reveal strong time, spectral and spatial correlations of these features across all locations. These features can be used for a multi-dimensional spectrum inference in cognitive radio based on machine learning.

Keywords: cognitive radio network, machine learning, spectrum measurement, spectrum inference, spectrum sensing

1. Introduction

5G systems deployment has already begun with the start of 2021. Many capabilities pertaining to it are still being standardized. To meet ever-increasing requirements future radio communication networks are being envisioned to give a global network coverage with seamless connectivity, spectrum and energy efficiency, enhanced cognitive capabilities and high security to name a few [1]. Till date the sporadic evolution of wireless communication systems has been essentially centered around the need for wide availability of spectrum bandwidth. This is because all emerging wireless solutions are spectrum hungry. Even though the demand for wireless capabilities is going to grow continuously new challenges pertaining to Internet of Everything (IoE) are slowly emerging. This has paved a paradigm shift from data rate centric applications and services to extremely reliable and low latency communication services [2]. Wireless radio spectrum since long has been a precious limited natural resource. Many communication services still are distributed over statically allocated frequency bands in this communication spectrum. This spectrum resource is still being allocated via a fixed auction policy which giving license to its owner to operate in its allocated band. This conventional management of spectrum bands has been inefficient to accommodate growing services with higher spectrum bandwidth demands. In addition to this studies across the globe for evaluating spectrum utilization have revealed time to time spectrum under-utilization and inefficiency. Cognitive Radio (CR) since then has come up as a prominent candidate with technologically feasible solution for realizing dynamic spectrum access. It adds cognitive ability to a radio by allowing real time use of unused spectrum bands. To achieve this goal of intelligent radio a CR is equipped with four major functionalities to deal with spectrum namely sensing, sharing, decision and mobility [3]. Initially spectrum sensing is a vital step for estimating spectrum occupancy and to understand the behaviour of licensed user called the primary user (PU) of spectrum. Efficient spectrum sensing is the key for making occupancy related future decisions. Accurate spectrum sensing has long being an active research topic. However, it suffers in terms of drawback related to sensing speed, energy utilization and sensing scope. Spectrum sensing seems to be passive decision making strategy which hinders spectrum performance. However if some sort of pro-activeness and dynamicity is added to this step then such intelligent decisions can impact the systems overall response. CRs require intelligent spectrum management techniques for dynamic spectrum management. The very important first step towards CR deployment is to learn features of spectrum characterization and explore utilization statistics. This is possible if we have measured data collected in the desired band of interest keeping in mind its variant nature. Spectrum Inference/Prediction in contrast to spectrum sensing is proposed as a technique of inferring the radio spectrum state from spectrum occupancy calculations by utilizing the existing correlations in the measured data. Since the advent of data mining and improved machine learning algorithms for prediction many studies have been addressed towards development of efficient spectrum inference techniques based on machine learning [4].

2. Literature survey

The Communication spectrum is a widely distributed band of frequencies and must accommodate a number range of wireless services with different technological specifications. Spectrum measurement is the first step towards spectrum inference. Simple measurement test bed involves a receiving wide band antenna, spectrum analyzer equipment and data processing terminal. The choice of equipment and design of measurement setup is highly influenced by target bands under consideration, objectives, and cost. Different antenna configurations and types have been used in measurement setups like vertically polarized antennas (Vert2450), bidirectional antenna (BiConiLog), Log periodic, circularly polarized wide band micro strip and omni-directional discone antenna etc. For spectrum analysis most popular choice has been handheld spectrum analyzers with built in pre-amplifiers which gives good portability and high performance to cost ratio. However, many campaigns have also used standalone spectrum analyzers with external low noise amplifiers and special sophisticated spectrum monitoring systems. The designed system should be able to measure weak signals from distant transmitters or faded signals due to nearby attenuation. It should also withstand the effect of strong signals which may otherwise affect the system. It should be able to sense along a wide frequency range encompassing narrow and wide bands. By measuring communication spectrum occupancy statistics, a useful insight into spectrum utilization of statically licensed vital communication services which includes Wi-Fi, mobile cellular operations, Broadcast (TV), Broadband Wireless Access (BWA) and ISM etc can be investigated. In the past a number of spectrum measurement campaigns have been done across globe in many western countries in Europe similarly US, UK, China and Germany, Spain, Singapore, New Zealand etc and recently in India, Turkey, Qatar, Philippines, Mexico etc [5-14] to extract simple signal statistics like Power Spectral Density (PSD), Occupancy (Occ), Duty Cycle (DC) as well as more rigorous and comprehensive statistical features like Cumulative Complementary Distribution Function (CCDF) and Probability Density Function (PDF) from the empirical measurements [15-17]. First order and second order statistics have been explored in terms of Channel availability durations and Service congestion rates to interpret spectrum usage pattern and use it for prediction in future [18].

The methodology used for analyzing the measured spectrum data has been quite similar in majority of campaigns. All measurement campaigns have considered mostly indoor/outdoor measurement setup. Normally outdoor locations are preferred to avoid signal fading and attenuation. Such locations are normally chosen as roof top of high-rise buildings in order to have a glimpse of surrounding spectrum activity around that area. High altitude also helps in combating signal strength attenuation due to propagation losses. Indoor locations are normally chosen as internal to office buildings having commercial, business or academic setup. Such a scenario readily mimics onboard application environment usually encountered in wireless communication applications. The number of spectrum opportunities indoor does exceed the number available outdoors. From practical point of view a simultaneous data gathering from different locations to address effects of location variation on spectrum occupancy characterization is difficult and costly. In general locations are normally fixed and meticulously surveyed and selected, the time and frequency parameters can then be varied and analyzed. However, spectrum utilization and characterization does vary with time, hence data measurements at different locations even though with same settings will vary with time [19].

Spectrum occupancy statistics are highly influenced by time, frequency, space, threshold, and resolution frequency variations. Furthermore, there exist day wise, weekly as well as occasional festive changes. A more detailed and effective modeling taking into consideration all these factors need to be considered and derived. It is also required to assess the applicability of existing modeling methods for variety of bands under consideration of which GSM, ISM and HF being more popular. Existing communication services encompasses not only long-range telecommunications like cellular and satellite but also shorter range consisting of Bluetooth and wireless peer-based relaying etc. This gives a heterogeneous structured network topology hence only roof top measurement setups on high rise buildings would not suffice. There is also needed to capture spectrum activities at the local level comprising at the street levels to capture service hot spots and their influence over the measured data. For example, home based microwave ovens may also add to spectrum congestion in 2400-2500 MHz ISM band. There exist demands wherein operating devices can intelligently sense the availability of unused spectrum, adjust their operating parameters, and accommodate themselves in the unused bands. Such accessible opportunities have significantly enhanced the usage of unexplored and vacant spectrum bands. These developments have paved the way for not only qualitative understanding of spectrum behavior but also its micro level quantitative analysis. These studies and findings can help confirm the validation of existing spectrum behavior models and channel usage statistics. These studies will help built more real time spectrum models and channel models with enhanced prediction accuracy and minimum complexity.

Ultimately the goal of wise spectrum decisions can be practically realized and achieved then [20]. Table 1 gives a summary of statistical features studied in literature across three dimensions time, frequency, and space. With these objectives and goal, a detailed comprehensive spectrum occupancy data measurement campaign was carried out at four different demographically varying locations in Maharashtra and Karnataka state of India. The measurement settings have been kept in consistence with our past spectrum investigations using the same setup [21]. S. L. Bangare et al. [23-24] shown research in machine learning IOT. J. Surve et al. [25] worked in IOT for health. J. Alanya-Beltran et al. [26] have worked in man machine interface. Xu Wu et al. [27] worked in network security. The frequency band under study for investigation employed was 700 MHz to 2.7466 GHz. The important features of our measured data include,

1. Wide operating frequency range from 700 MHz to 2.7466 GHz for demo- graphically varying locations.
2. Frequency resolution of 200 KHz to extract the channel usage of all popular telecommunication services.
3. Weeklong usage statistics for comprehensive occupancy analysis and extracting the behavior pattern with temporal and spectral variations.

Table 1: *Summary of measurement campaigns under study Sample statistics under consideration*

Time	Frequency	Space
PSD [5-8][10,13]	[12]	[11,14,15]
Occ		[11]
DC [6,10,13]	[12]	[15]
CCDF [5]		[11]
PDF [9,12]		

3. Spectrum Occupancy Measurement Methodology

3.1 Measurement setup and specifications

The setup for spectrum data measurement campaign is shown in Figure 1. We have used a broadband discone antenna AOR DA5000, handheld spectrum analyzer Rhode & Schwartz R&S FSH3 and a desktop PC/Laptop for data processing and analysis [22]. A low loss coaxial cable is used to connect antenna to spectrum analyzer. The laptop is connected to spectrum analyzer using RS-232-C optical interface which is compatible with Windows operating systems. The optical connection prevents spurious measurements due to interference. Frequency range of Spectrum analyzer is from 100 kHz to 3GHz. The sensed signal power is measured in dBm by this instrument using Energy detection algorithm. This spectrum analyzer has inbuilt storage useful for storing the log files generated in real-time. Later these are processed using R & S FSH View and Remote. The detailed occupancy statistics was analyzed using Python and MATLAB.

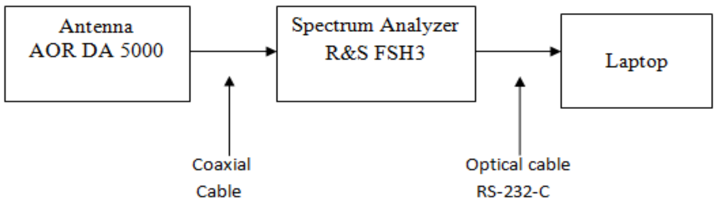

Figure 1. Block diagram of measurement setup

The antenna prototypes has been shown by placing a ruler on top of it, which shows its dimensions. The developed antenna's performance parameters are detailed in the next part of the results and analysis.

3.2 Measurement campaign sites and configurations

A detailed spectrum occupancy data measurement campaign for studying the spectrum occupancy statistics and analyzing the features of sensed data was conducted in three major cities, two in Maharashtra state and one in Karnataka state, India. The measurements were conducted at predefined locations fixed such that spectral activity for urban, semi urban and suburban scenarios is captured for study and analysis. The detail of these locations and the underlying setups is detailed in Table 2. The measurement location and sites along with configuration have been intentionally selected to get varied teledensity and set up configuration for distribution of different spectrum services as seen in Figures 2 to Figure 5. This arrangement is required to get an insight into the spectral behavior patterns for extracting different attributes and features of measured data required for predicting by learning. The possibility of getting very high degrees of temporal and spectral correlations can be expected amongst settings of similar and dissimilar arrangements. In order to capture time diversity the spectrum was sensed with varied time duration and taking into consideration the diversification of time slots and days. The details have been tabulated in Table 3.

Figure 2. Measurement site Pune (urban, semi urban

Figure 3. Google map site Pune (sub urban, urban)

Figure 4. Measurement site Solapur (sub urban, urban), Kalaburagi (urban)

Figure 5. Google map Solapur (sub urban, urban), Kalaburagi (urban)

Table 2: Summary of measurement campaigns under study

Measurement City/Tele density Classification	Location site	GPS coordinates/Latitudes-Longitudes	Setup Configu - ration	Dura tion
Pune Semi urban	Sinhgad Academy of Engineering, Kondhwa	(18° 28' 20.172" N, 73° 53' 4.704" E) (18.472270, 73.884640)	Indoor and Outdoor	2 days
Pune Urban	Konark Puram, Pune	(18° 29' 32.424" N, 73° 57' 3.06" E) (18.492340, 73.950850)	Outdoor	2 days
Solapur Sub urban	N. B. Navale Sinhgad College of Engineering, Kegaon	(17° 43' 41.009" N, 75° 51' 1.246" E) (17.728037, 75.850281)	Indoor and Outdoor	2 days
Solapur Urban	Seerat Nagar, Solapur	(17° 38' 53.591" N, 75° 55' 5.619" E) (17.648219, 75.918227)	Indoor and Outdoor	7 days
Kalaburagi Urban	Yadullah Colony, Kalaburagi	(17° 20' 26.674" N, 76° 51' 23.997" E) (17.340742, 76.856665)	Outdoor	2 days

Table 3: Summary of time and location diversity

	Outdoor	Indoor	Day time	Night time	Week end	Weekly holiday	National holiday	Complete week
Loc1 (Pune)	√	√	√	√				
Loc2 (Solapur)	√	√	√	√	√	√	√	√
Loc3 (Kalaburagi)	√	√	√	√	√	√		

4. Noise Signal Measurements and Analysis

The target spectrum from 700 MHz to 2.7466GHz was separated with 200 KHz frequency resolution into 34 sub bands. Here each sub band consists of 301 measurement points. The centre frequency for each sub band is dynamically computed through a software program. Total 10234 measured values are obtained in one sweep. One complete sweep consumes about 90 seconds. The program iterates in a loop to obtain data for complete 48 hours (2 days). The noise was measured for 2 hours at the same location by terminating the spectrum analyzer with 50Ω matching load. The stored data was further pre-processed and used for spectrum occupancy investigation and analysis [22] [28].

4.1 Noise Signal and Decision Threshold

In order to get the occupancy data signal needs to be separated from noise by a right decision threshold. A simple method to arrive at the decision threshold is by visually inspecting noise and signal curves and placing it midway between these curves of the plotted empirical data. In literature different methods are used to decide the decision thresholds. These techniques are used in many spectra measurement campaigns worldwide [19].

4.1.1 Max Noise method

This method makes use of $X_{max}(f)$ which is the maximum value of noise amongst different frequencies. This value is obtained by terminating a matched load to the antenna. This method ensures that there is no overestimation as no noise sample lies above decision threshold.

This maximum value is selected as the decision threshold $\gamma(f)$ given by Eq. (1),

$$\gamma(f) = X_{max}(f) \tag{1}$$

4.1.2 m-dB method

In case of MaxNoise method there is a possibility of underestimation as some weak signals may lie below the maximum noise value selected. In order to overcome this problem normally the decision threshold is fixed above the average noise level $X_{mean}(f)$ by m decibels given by Eq. (2),

$$\gamma(f) = X_{mean}(f) + m \tag{2}$$

4.1.3 PFA method

The noise parameters like variance and maximum level vary band wise and is a function of configuration parameters like frequency span. Hence, a fixed decision threshold may not be a better choice when we are switching configuration parameters from one band to another. Probability of False Alarm (PFA) gives a measure of declaring a channel busy when actually it is free. Taking due consideration of desired PFA for a CR network say P_{fa} at each measured value of frequency a decision threshold $\gamma(f)$ is fixed. This method ensures that only a certain percentage P_{fa} of the measured noise samples $X(f)$ are above the decision threshold $\gamma(f)$ given by Eq. (3),

$$\gamma(f) = F_{X(f)}^{-1}\left(1 - P_{fa}\right) \tag{3}$$

where $F_{X(f)}^{-1}(.)$ represents the inverse of $F_{X(f)}(.)$ i e the CDF of measured noise samples $X(f)$. This is an intermediate approach. The value of P_{fa} for this criterion gives the maximum overestimation error.

4.1.3 Competitive analysis of different decision threshold techniques

When MaxNoise method is used there are chances of spectrum occupancy underestimation since some weak signals may lie below the maximum noise value selected as decision threshold. Since noise variance and maximum noise level vary with respect to bands a constant m-dB threshold for entire band is not suitable. As an optimum choice Probability of False Alarm (PFA) criterion conciliates the above two methods. PFA method is a optimum choice considering the underestimation due to MaxNoise method and static nature of m-dB method [19]. After experimentally assessing carefully the effect of applying these three methods a PFA criterion of 1% was selected as a decision threshold for computing the spectrum bands occupancy statistics as in Figure 6. Different decision thresholds for these three techniques have been tabulated in Table 4.

Table 4: Decision thresholds with different techniques at different locations

MaxNoise (dBm)	**Mean Noise (dBm)**	**mdB (dBm)**	**PFA (dBm)**
Loc1 -91.4863	-93.2044	-90.2044	-91.3430
Loc2 -93.4598	-93.7650	-90.7650	-91.9030
Loc3 -93.3430	-93.6456	-90.6456	-91.6160

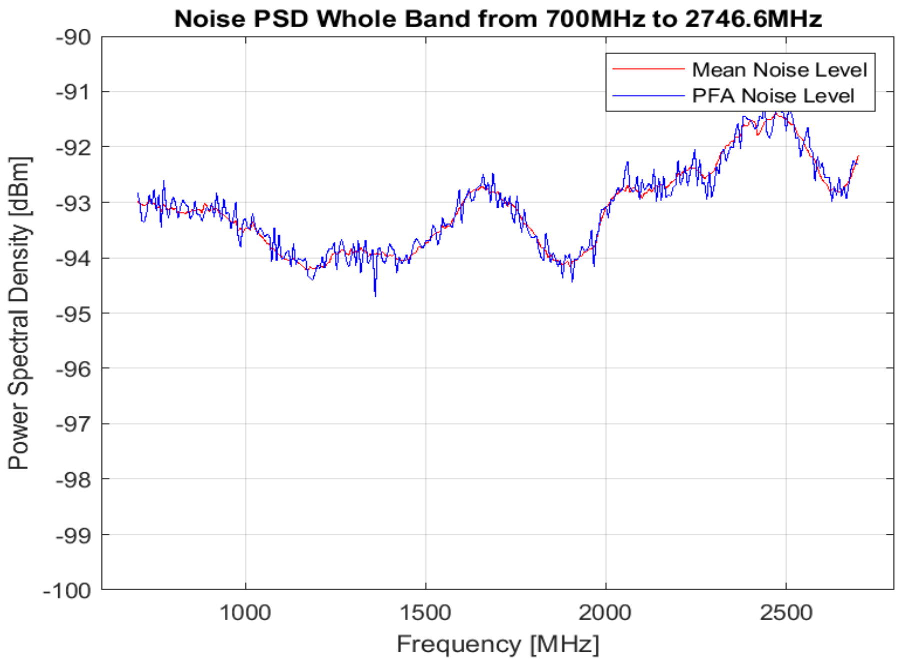

Figure 6. Mean Noise vs PFA noise level variation at Pune location

4.2 Mathematical formulation of Spectrum Occupancy

If $O_{t(k),f(n)}$ denotes the spectrum occupancy of a channel at time instant $t(k)$ and frequency $f(n)$ then comparing the PSD measured P_γ the frequency point with the decision threshold γ we can get the binary occupancy data as in Eq. (4),

$$O_{t(k),f(n)} = \begin{cases} 0, & P_\gamma < \gamma \\ 1, & P_\gamma > \gamma \end{cases} \qquad (4)$$

Thus$O_{t(k),f(n)} = 1$ indicates channel is occupied. If PSD value P_γ is below γ is an indication of empty/free channel. If we consider a service band to be a collection of several channels then the average occupancy of each such band can be computed by using Eq. (5),

$$O = \frac{1}{KN} \sum_{k=1}^{K} \sum_{n=1}^{N} O_{t(k),f(n)} \qquad (5)$$

Here, N being the total number of the measured frequencies in complete band and K denotes the number of corresponding time samples at this frequency [7]. In this work there are total 10234 frequency points in each sweep and total number of such sweeps is 1920.

5. Spectrum Occupancy Measurement Results and Discussion

5.1 Location and spectral diversity in spectrum occupancy

Power Spectral Density (PSD) plot gives maximum power of the received signal at every individual frequency. In Figure 7 the PSD plot for selected three different locations is depicted. It can be readily inferred from plots that utilization pattern of spectrum matches for two days. Major usage is in cellular bands namely GSM (900, 1800), CDMA, LTE (2100, 2300) as well as ISM band as seen in Figure 7. This shows that a major chunk of communication spectrum is being consumed by the telecommunication services. This means there is significant overcrowding for these bands and less traffic in other service bands like TV broadcasting etc.
The frequency ranges under consideration covers almost all major telecommunication services, wireless applications, broadcast services etc. The spectrum utilization seems to be quite high for

GSM900, GSM1800, IMT-3G, Broadband Wireless Access, ISM band and LTE irrespective of location under consideration. The measured occupancy is very sensitive to threshold setting; a minor change in threshold affects occupancy drastically.

The percentage of occupancy is significantly determined by teledensity of the respective location as depicted by Figure 8. The spectrum activity patterns are influenced by usage which varies with time. The occupancy pattern shows a relatively similar graph irrespective of change in location. This is true since it is the telecom circle which decides the frequencies being utilized. In order to accommodate variation with respect to telecom circle we considered Kalaburagi city in Karnataka telecom circle. Unless otherwise specified with slight variation in special services allocated in certain bands, normally if we consider the pan India structure there should be minimum interference in movement across different telecom circles with same telecom operators. Hence, we see that the pattern of occupancy is largely same across all locations. To confirm this band wise, we plotted a similar graph for SNR, PSD and Occupancy for GSM band and found a similar variation across all three locations as depicted in Figure 9.

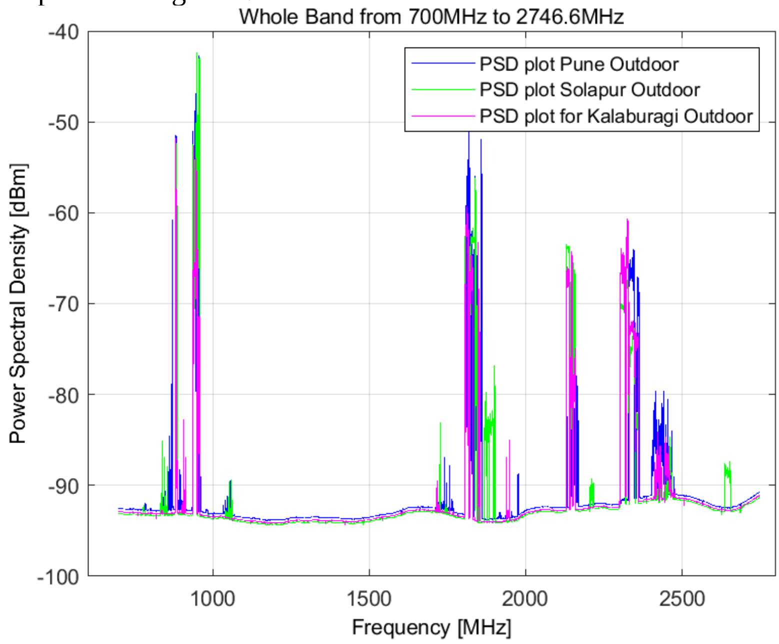

Figure 7. Measured Signal Strengths across all locations

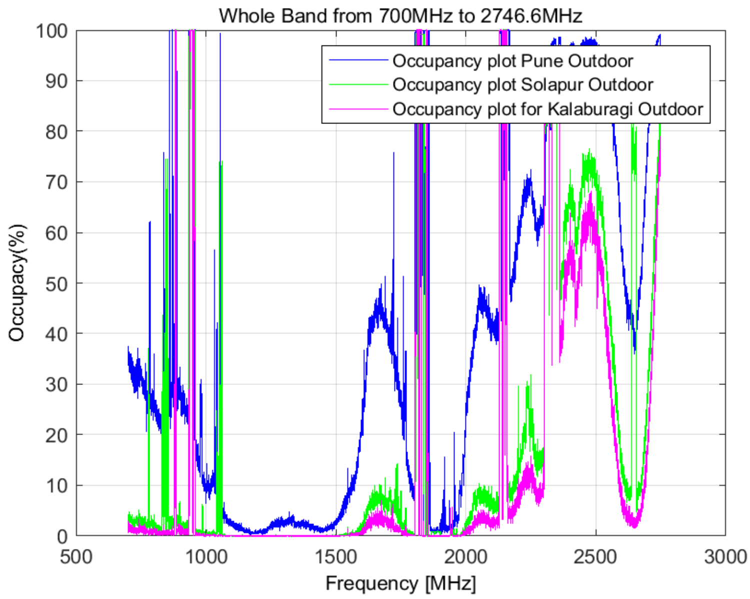

Figure 8. Occupancy plots across different locations

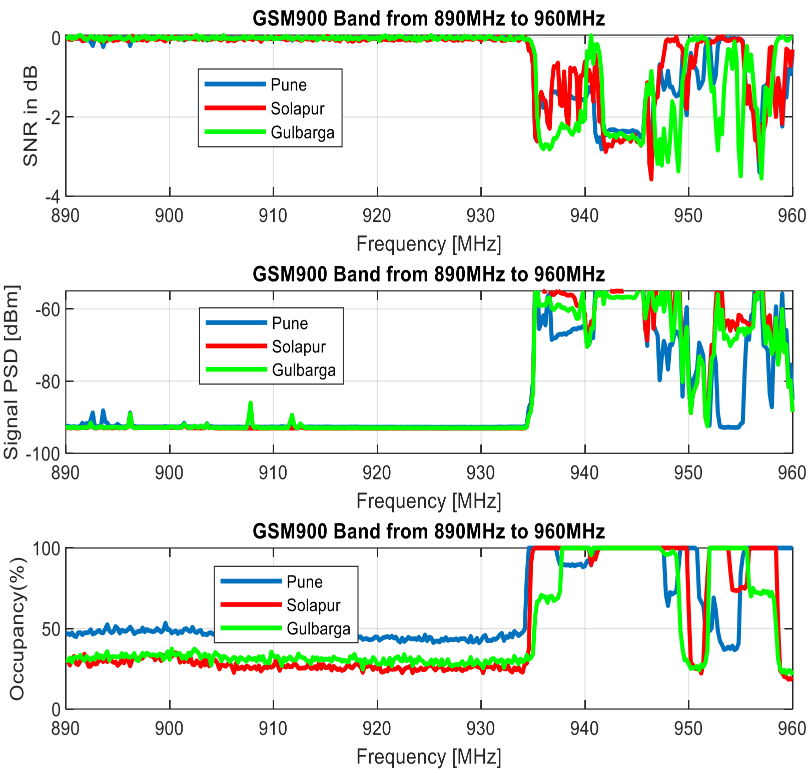

Figure 9. GSM Band plots across different locations

5.2 Time Diversity in Spectrum Occupancy Measurement

To investigate occupancy pattern utilization with respect to time diversity, we have plotted and studied the hourly and weekly occupancy plot. It can be inferred that the spectrum activity shows a gradual ascent at the beginning of the day and reaches peak with increase in communication activities as day progresses. After a high peak noted at the end of working day a gradual descent is being observed to depict the decay in spectral activities as is evident in Figure 10. Unless for seasonal variations this pattern seems to be highly intact. In order to study the effect of variation in days of week the data for entire week was processed and analyzed as shown in Figure 11. Weekly occupancy trend shows consistency in occupancy percentage unless there is a national holiday or weekend.

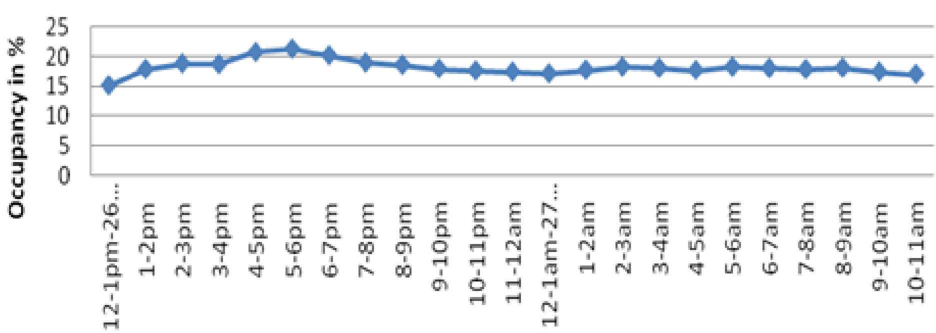

Figure 10. Hourly Occupancy plot at Solapur

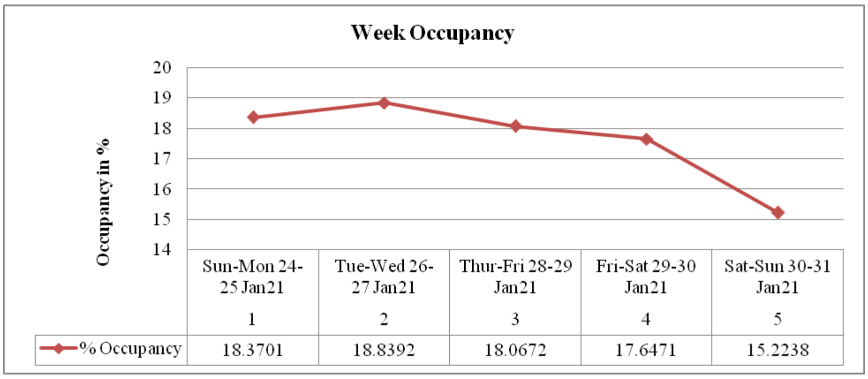

Figure 11. Weekly Occupancy plot at Solapur

5.3 Spectrum Occupancy signal statistics

The measured data has been analysed for varying levels of statistics which include [18],

1. Channel availability duration with regards to time,
   for different channels and multiple wireless services. It is an indication of long term static usage of channels in a service.
2. Service blockage analysis and congestion rates. It is a short term dynamic measure for service under consideration.

5.3.1 Channel C_H

It is our smallest scale for differentiating measured power levels. It is a chunk of 200 KHz evenly spaced at measurement points for the entire band from (700 MHz to 2.7466 GHz). This gives us a total of 10234 channels in our spectrum database. Channel C_H is defined as the frequency interval

$$[(x+200(C_H - 1))\ \text{KHz},\ (x+200\,C_H)\ \text{KHz}],\ C_H > 0$$

5.3.2 Service

It is essentially a service rendered for use of a particular operation in the spectrum. It may be a cellular service for mobile application, WIFI service or broadcast application service. Essentially many channels together form a service. Depending upon allocated bandwidths and application this number varies within the spectrum.

5.3.3 Channel State Information

It is an important indication of channel occupancy and availability. It depends on both time and the channel under consideration. The value of 0 for CSI (t, C_H) gives an indication that Channel C_H is idle at time slot t, and a value of 1 for CSI (t, C_H) gives indicates otherwise[18].

5.3.4 Channel Occupancy and Availability durations

This can be characterized from CSI time series. In channel availability duration a channel remains idle which can be identified by a stream of continuous 0s. In channel occupancy condition a CSI time series would produce a stream of continuous 1s. The procedure for obtaining CSI time series data is as follows; At a time slot t, measured power level of a channel (PSD_{CH}) is compared with threshold (γ)

$$\begin{gathered} \text{If } PSD_{CH} > \gamma \\ \text{then } CSI(t, CH) = 0; \\ \text{else } CSI(t, CH) = 1 \end{gathered}$$

The threshold is selected using Probability of False Alarm (PFA) method with 1% criteria given by Eq. (6)

$$\gamma(f) = F_{X(f)}^{-1}\left(1 - P_{fa}\right) \qquad (6)$$

Here, $F_{X(f)}^{-1}(.)$ is inverse function of $F_{X(f)}(.)$ i.e the CDF of measured noise samples$X(f)$. This is an intermediate approach. The value of P_{fa} gives the maximum overestimation error for PFA criterion. There is an underlying assumption noted here that noise behaves similarly across channels. Threshold value is inversely proportional to signal detection probability and vice versa. In earlier studies a same threshold mechanism was employed in a measurement campaign at Sinhgad Academy of Engineering, Pune, India [21]. The outcome of this task yields a bit stream of (1s and 0s CSIs) for every single channel with 200 KHz bandwidth. It represents its occupancy with approximately 1.5 minutes time resolution.

5.4 Channel Availability and Occupancy Duration Distributions

In this section the sample statistics related to channel usage and services is discussed. In channel availability duration a channel remains idle which can be identified by a stream of continuous 0s. After the process of thresholding, we get a stream of 1s and 0s over the entire band. This CSI time series sequence contains of the order of more than hundreds of channel availability intervals across total 10234 time slots of all channels. If we want to have a statistical measure of this kind of distribution it is very difficult to obtain since this sample size seems to be very small. Therefore, we increase this sample size by considering channel availability duration across all channels within the same service across entire measurement duration. In normal case the spectrum usage pattern of channels within the same service is similar. Thus, spectrum availability and vacancy intervals in all channels within any service say GSM900 up link service (885-915MHz) is almost the same. We use this fact to obtain the statistical empirical distribution of channel availability and vacancy.

We can plot the probability distribution (histogram) of the CAD for all services across all locations and observe their nature. The same analysis was done for all the cellular bands GSM900/1800 uplink/downlink, CDMA as well as broadcasting TV, ISM etc across all locations. The empirical distribution of CVLD for GSM 800DL, GSM 800UL, GSM 900DL, GSM 900UL, GSM 1800UL, GSM 1800DL is being analyzed and depicted in Figure 12, 13 and 14. This distribution has exponential nature and has been verified for all services at different locations. The channel availability and busy durations are not independently distributed over time. It can be inferred that spectrum usage is prone to be predictable and better algorithms and models can be applied to get reasonable predictive power.

5.5 Service Blockage Analysis and Congestion Rate

CAD measures are for a single channel giving its long-term average information. If we wish to dynamically track the activity of spectrum channel wise then these measures are not much useful. Hence we need to define new measures to dynamically track real-time status of channels and services. A statistical measure for such a provision is Service Blockage Rate (SBR)/Service Congestion Rate (SCR) given by Eq. (7) and Eq. (8).

$$SCR(t,S) = \frac{no\ of\ busy\ channels\ in\ service\ S\ at\ time\ t}{Service\ S\ total\ number\ of\ channels} \qquad (7)$$

Thus,

$$SCR\ (t,S) \ = \sum\nolimits_{(c \epsilon S)} \frac{CSI(t,c)}{n} \qquad (8)$$

here the number of channels in service S is n.

SCR gives an indication to the extent of blockage in a service [18]. Its value ranges from 0 to 1. High value of SCR of a particular service indicates fewer unoccupied channels and vice versa. It is thus a measure for a real- time statistic on spectrum occupancy for deciding dynamical access adaptive. This series is periodic and has a repetition frequency of one day which can be readily verified. Another important and interesting behavior of this series is that for two different services there seem to be higher degree of similarity: the rises and drops in the graphs seem to be aligned in symphony.

CVLD measure gives a long-term average duration distribution for channels/services. This measure cannot be used for real time decision making in case of shorter durations. Hence another useful statistic called Service Congestion Rate (SCR) is used which is a real-time short-term measure for dynamic decision for dynamic spectrum access. This statistical measure seems to be periodic in nature day wise with respect to specific wireless service under consideration. This can be understood considering user activity patterns being consistent within service with respect to time.

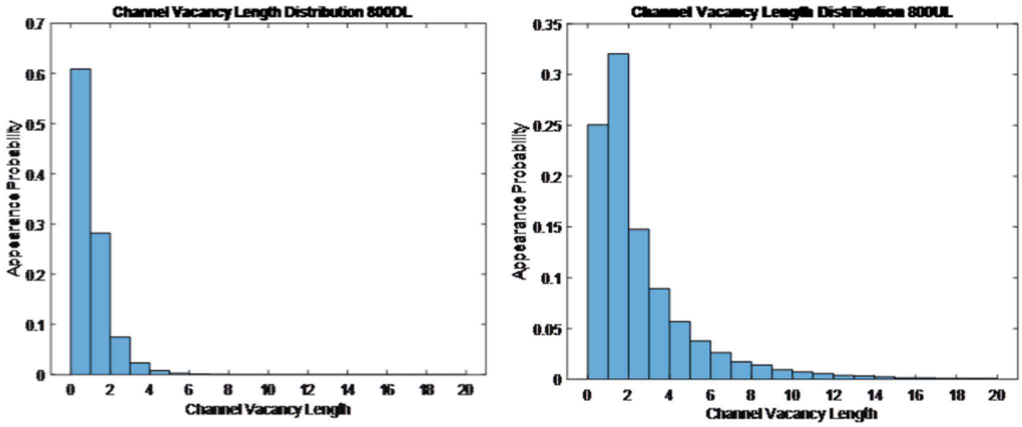

Figure 12. CVLD distribution for GSM800 Downlink and Uplink

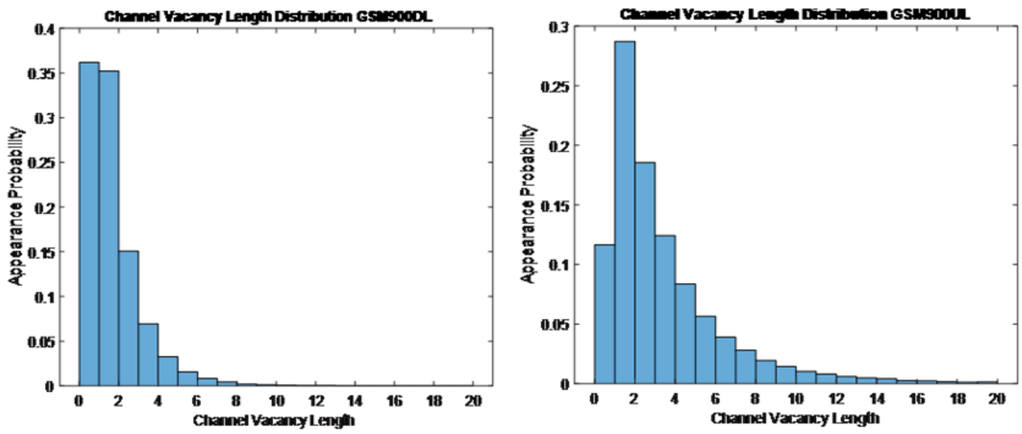

Figure 13. CVLD distribution for GSM900 Downlink and Uplink

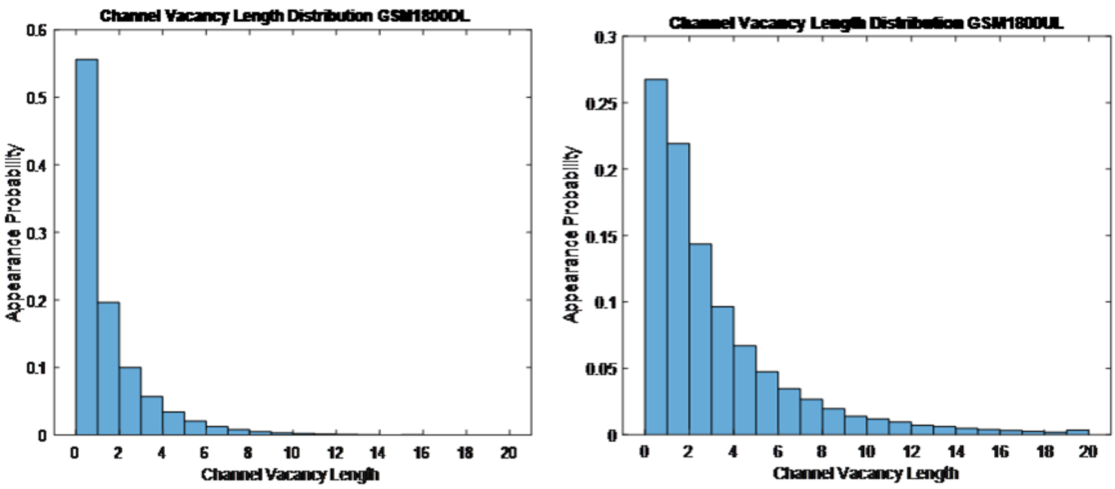

Figure 14. CVLD distribution for GSM1800 Downlink and Uplink

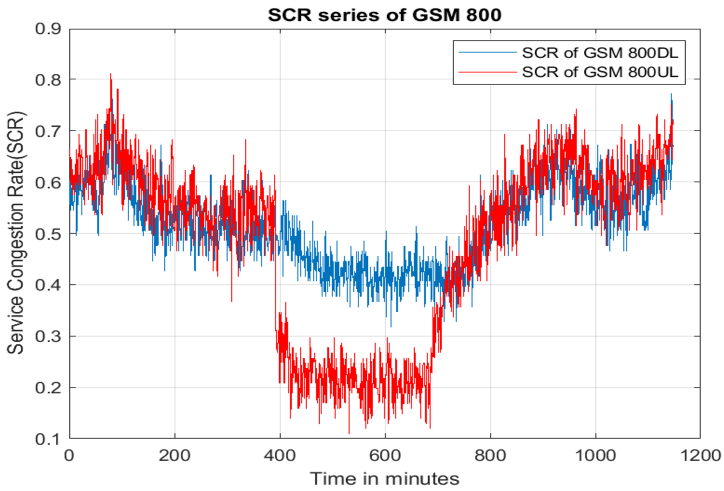

Figure 15. SCR plot for GSM800 Uplink and Downlink

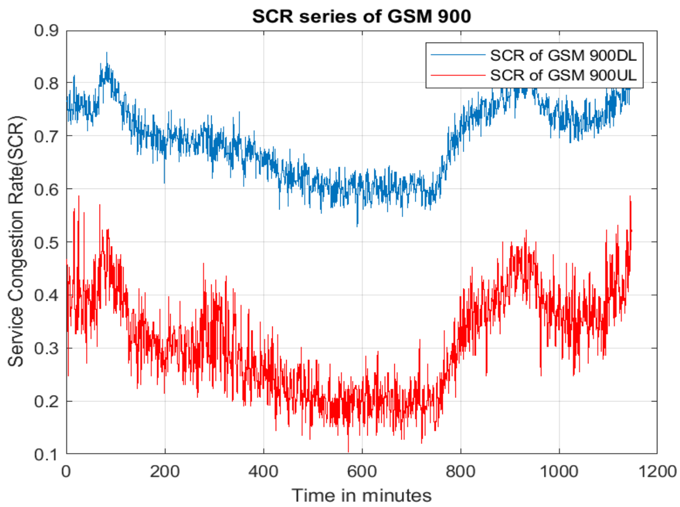

Figure 16. SCR plot for GSM900 Uplink and Downlink

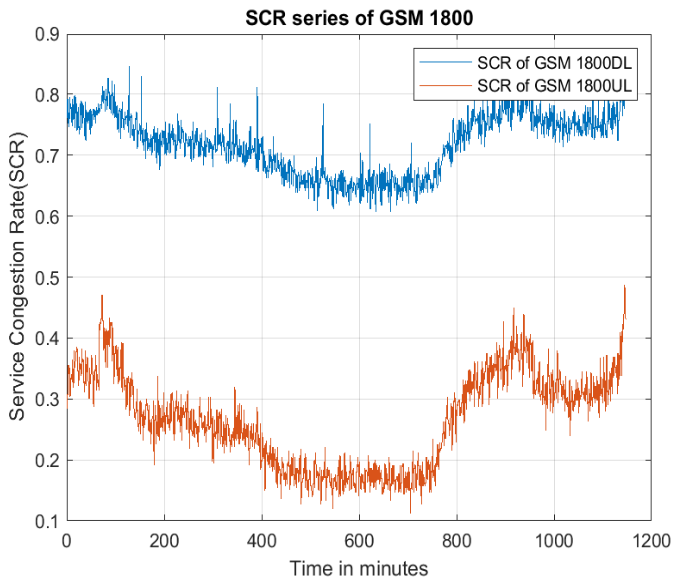

Figure 17. SCR plot for GSM1800 Uplink and Downlink

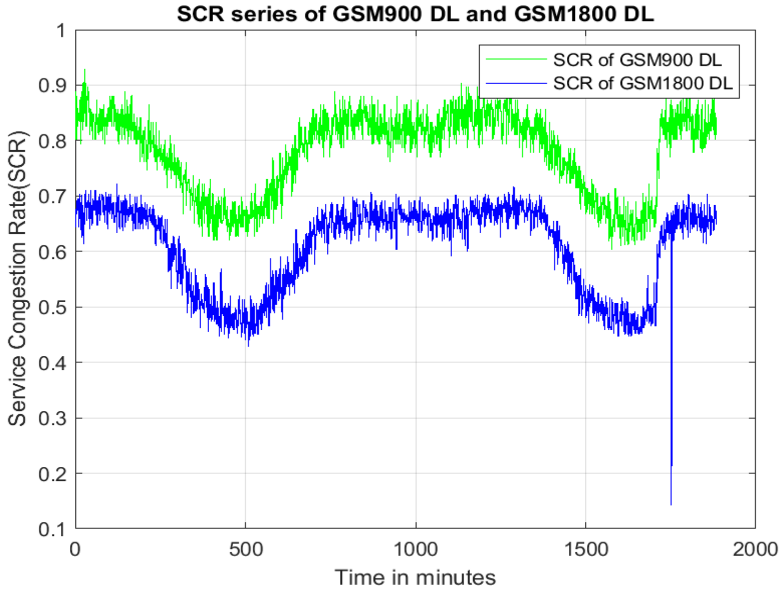

Figure 18. SCR plot for GSM900 and GSM1800 Downlink

6. Conclusion

This study details spectrum data statistics and occupancy analysis for spectrum inference in cognitive radio network. We have selected and configured a measurement set up to accommodate different teledensity locations for varied services. Barring the same equipment setups and technical specifications it can be concluded that the occupancy pattern is largely independent with respect to locations. However, the percentage occupancy is highly influenced by the usage pattern and teledensity of area under consideration. Spectrum usage pattern for hourly and weekly occupancy reveal variations with respect to time of the day as well as day of the week. The statistics of the collected data viz; CVLD and SCR give a long term and short-term usage statistics. Though CVLD essentially reveals the exponential nature of distribution of intervals across all services with respect to all locations, it is still a long-term average which is not much useful for real time decision making. SCR gives a short-term real time estimate of channel vacancy and occupancy and can be utilized to take fruitful decisions. SCR series of all services are periodical with a period equal to one day. SCR series of similar services across different locations exhibit a strong pattern similarity. The inherent relations between the observed statistics across different services and locations motivate us and lay the foundation for regression-based analysis to determine the strong correlations in time and frequency as well as space. These features will be used to formulate a multi-dimensional spectrum inference scheme to predict available spectrum occupancy.

References

[1] W. Saad, M. Bennis and M. Chen, "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems," in IEEE Network, vol. 34, no. 3, pp. 134-142, May/June 2020, https://doi: 10.1109/MNET.001.1900287.

[2] You, X., Wang, CX., Huang, J. et al. Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts. Sci. China Inf. Sci. 64, 110301 (2021). https://doi.org/10.1007/s11432-020-2 955-6

[3] F. Akyildiz, W. Lee, M. C. Vuran and S. Mohanty, "A survey on spectrum management in cognitive radio networks," in IEEE Com- munications Magazine, vol. 46, no. 4, pp. 40-48, April 2008, https://doi: 10.1109/MCOM.2008.4481339.

[4] G. Ding et al., "Spectrum Inference in Cognitive Radio Networks: Algorithms and Applications," in IEEE Communications Surveys & Tutorials, vol. 20, no. 1, pp. 150-182, Firstquarter 2018, https://doi: 10.1109/COMST.2017.2751058.

[5] Khoja, Nurali et al. "Opportunistic Spectrum Access Measurements and Analysis in Urban Area for a better cognitive approach." 2017 International Conference on Computing, Communication, Control and Automation (IC- CUBEA) (2017): 1-6.

[6] Shirgan, S., Bombale, U. Analysis of Wideband Microstrip Antenna based Spectrum Occupancy Measurement Campaign for Cognitive Radio Appli- cation. Radioelectron.Commun.Syst. 61, 55–63 (2018). https://doi.org/103103/S0735272718020024

[7] Agarwal, A. S. Sengar, R. Gangopadhyay and S. Debnath, "A real time measurement based spectrum occupancy investigation in north- western india for Cognitive Radio applications," 2016 International Con- ference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 2016, pp. 2035-2039, https://doi: 10.1109/WiSP- NET.2016.7566499.

[8] J. Jacob and B. R. Jose, "Spectrum occupancy measurement and analysis in Kochi-India from a cognitive radio perspective," 2016 Sixth International Symposium on Embedded Computing and System Design (ISED), Patna, 2016, pp. 328-333, https://doi: 10.1109/ISED.2016.7977107.

[9] Şeflek, I. and E. Yaldiz. "Spectrum Occupancy Measurements at University Campus in Turkey." International Journal of Electronics and Electrical Engineering (2017): 1-6.

[10] B. KorunurEngiz and Y. A. Rajab, "Investigation of Spectrum Occupancy in GSM Band in Samsun, Turkey," 2019 6th International Conference on Electrical and Electronics Engineering (ICEEE), Istanbul, Turkey, 2019, pp. 158-161, https://doi: 10.1109/ICEEE2019.2019.00038.

[11] F. Kiftaro, M. El-Tarhuni and K. Assaleh, "UHF Spectrum Occu- pancy Measurements in Sharjah - UAE," 2017 9th IEEE-GCC Con- ference and Exhibition (GCCCE), Manama, Bahrain, 2017, pp. 1-5, https://doi: 10.1109/IEEEGCC.2017.8447900.

[12] M. Mehdawi, N. G. Riley, M. Ammar, A. Fanan and M. Zolfaghari, "Spectrum occupancy measurements and lessons learned in the context of cognitive radio," 2015 23rd Telecommunications Forum Telfor (TELFOR), Belgrade, Serbia, 2015, pp. 196-199, https://doi: 10.1109/TELFOR.2015.7377446.

[13] H. Maloku, Z. L. Fazliu, M. Ibrani, A. Mekuli, E. Sela and M. Rajarajan, "Measurement of Frequency Occupancy Levels in TV Bands in Urban Environment in Kosovo," 2018 18th Mediterranean Microwave Symposium (MMS), Istanbul, 2018, pp. 268-271, https://doi: 10.1109/MMS.2018.8612014.

[14] D. A. Arista Ramirez, M. Cardenas-Juarez, U. Pineda-Rico, A. Arce and E. Stevens-Navarro, "Spectrum Occupancy Measurements in the Sub-6 GHz Band for Smart Spectrum Applications," 2018 IEEE 10th Latin- American Conference on Communications (LATINCOM), Guadalajara, Mexico, 2018, pp. 1-6, https://doi: 10.1109/LATINCOM.2018.8613211.

[15] Y. Chen and H. Oh, "A Survey of Measurement-Based Spectrum Occu- pancy Modeling for Cognitive Radios," in IEEE Communications Surveys & Tutorials, vol. 18, no. 1, pp. 848-859, Firstquarter 2016, https://doi: 10.1109/COMST.2014.2364316.

[16] M. Höyhtyä et al., "Spectrum Occupancy Measurements: A Survey and Use of Interference Maps," in IEEE Communications Surveys & Tutorials, vol. 18, no. 4, pp. 2386-2414, Fourthquarter 2016, https://doi: 10.1109/COMST.2016.2559525.

[17] Yasir Saleem, Mubashir Husain Rehmani, Primary radio user activity models for cognitive radio networks: A survey, Journal of Network and Computer Applications,Volume 43,2014,Pages 1-16,ISSN 1084-8045,https://doi.org/10.1016/j.jnca.2014.04.001.

[18] S. Yin, D. Chen, Q. Zhang, M. Liu and S. Li, "Mining Spectrum Usage Data: A Large-Scale Spectrum Measurement Study," in IEEE Transactions on Mobile Computing, vol. 11, no. 6, pp. 1033-1046, June 2012, https://doi: 10.1109/TMC.2011.128.

[19] M. López-Benítez and F. Casadevall, "Methodological aspects of spectrum occupancy evaluation in the context of cognitive radio," 2009 European Wireless Conference, Aalborg, 2009, pp. 199-204, https://doi: 10.1109/EW.2009.5357973.

[20] Patil, K. P. (2014). Spectrum Utilisation and Management in Cognitive Radio Networks. Department of Electronic Systems, Aalborg University.

[21] N. Khoja et al., "Opportunistic Spectrum Access Measurements and Anal- ysis in Urban Area for a better cognitive approach," 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA), 2017, pp. 1-6, https://doi: 10.1109/ICCUBEA.2017.8463736.

[22] Naikwadi M., Patil K. (2022) A Real Time Radio Spectrum Measurement Campaign for Machine Learning Based Spectrum Inference in Cognitive Radio Network. In: Kumar A., Senatore S., Gunjan V.K. (eds) ICDSMLA 2020. Lecture Notes in Electrical Engineering, vol 783. Springer, Singapore. https://doi.org/10.1007/978-981-16-3690-538.

[23] Bangare S. L., Prakash S., Gulati K., Veeru B., Dhiman G. and Jaiswal S., "The Architecture, Classification, and Unsolved Research Issues of Big Data extraction as well as decomposing the Internet of Vehicles (IoV)," 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC), 07-09 October 2021, pp. 566-571, doi: 10.1109/ISPCC53510.2021.9609451.

[24] Bangare S. L., Virmani Deepali, Karetla Girija Rani, Chaudhary Pankaj, Kaur Harveen, Hussain Bukhari Syed Nisar, Miah Shahajan, "Forecasting the Applied Deep Learning Tools in Enhancing Food Quality for Heart Related Diseases Effectively: A Study Using Structural Equation Model Analysis", Journal of Food Quality, vol. 2022, Article ID 6987569, 8 pages, 2022. https://doi.org/10.1155/2022/6987569

[25] Surve J., Umrao D., Madhavi M., Rajeswari T. S., Bangare S. L. and Chakravarthi M. K., "Machine Learning Applications For Protecting The Information Of Health Care Department Using Smart Internet Of Things Appliances -A REVIEW," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2022, pp. 893-898, doi: 10.1109/ICACITE53722.2022.9823642.

[26] Alanya-Beltran J., Narang P., Taufikin, Bangare S. L., Valderrama-Zapata C. and Jaiswal S., "An Empirical Analysis of 3D Image Processing by using Machine Learning-Based Input Processing for Man-Machine Interaction," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 28-29 April 2022, pp. 2478-2482, doi: 10.1109/ICACITE53722.2022.9823699.

[27] Wu Xu, Wei Dezhi, Vasgi Bharati P., Oleiwi Ahmed Kareem, Bangare S. L., Asenso Evans, "Research on Network Security Situational Awareness Based on Crawler Algorithm", Security and Communication Networks, vol. 2022, Article ID 3639174, 9 pages, 2022. https://doi.org/10.1155/2022/3639174

[28] Mr. Dharmesh Dhabliya, Mr. Rahul Sharma. (2012). Efficient Cluster Formation Protocol in WSN. International Journal of New Practices in Management and Engineering, 1(03), 08 - 17. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/7

[29] Sandeep Pande and Manna Sheela Rani Chetty, “Linear Bezier Curve Geometrical Feature Descriptor for Image Recognition”, Recent Advances in Computer Science and Communications, Vol. 13, No. 5, pp. 930-941, 2020.

[30] Dwarkanath Pande, S.., & Hasane Ahammad, D. S. . (2022). Cognitive Computing-Based Network Access Control System in Secure Physical Layer. Research Journal of Computer Systems and Engineering, 3(1), 14–20.

[31] Borgohain, U., Borkotokey, S., & Deka, S. K. (2021). A coalition formation game for cooperative spectrum sensing in cognitive radio network under the constraint of overhead. International Journal of Communication Networks and Information Security, 13(3), 431-438. doi:10.54039/IJCNIS.V13I3.5077