

Spectrum Occupancy Validation and Modeling Using Real-time Measurements

Chittabrata Ghosh &
Sumit Roy
Department of Electrical
Engineering
University of Washington
Seattle, WA 98195
ghoshc@u.washington.edu
sroy@u.washington.edu

Marepalli B. Rao
College of Medicine
University of Cincinnati
Cincinnati, OH 45221
marepalli.rao@uc.edu

Dharma P. Agrawal
Department of Computer
Science
University of Cincinnati
Cincinnati, OH 45221
dpa@cs.uc.edu

ABSTRACT

Existing research have considered Beta distribution in modeling channel occupancy of primary users in a licensed spectrum. This paper is the first initiative in validating this basic assumption in the cognitive radio paradigm using real-time measurements performed in Aachen, Germany on the 1500 MHz spectrum centered at 770 MHz. The Kolmogorov-Smirnov test is applied as a validation approach. The result of this test conforms to the validity of the Beta distribution modeling of spectrum occupancy. With this channel occupancy model, we have extended our research in modeling spectrum availability by proposing a new Poisson-normal approximation method. Further, we characterize available channels in a spectrum into five different types based on occupancy of its two adjacent channels. The probabilistic modeling of each of these types is determined using recurrence relations. Simulation results show that channel type classification and their modeling facilitate preferable selection of one section compared to another section over an operating spectrum.

Categories and Subject Descriptors

C.4 [Performance of Systems]: [modeling techniques, measurement techniques, performance attributes]

General Terms

Measurement, Verification

Keywords

Spectrum Occupancy, Beta distribution, Poisson-Normal Approximation, Kolmogorov-Smirnov Test, Cognitive Radio

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CoRoNet'10, September 20, 2010, Chicago, Illinois, USA.
Copyright 2010 ACM 978-1-4503-0141-1/10/09 ...\$10.00.

1. INTRODUCTION

Spectrum utilization of the primary (licensed) users over time does not follow a regular pattern. Certain portions of the spectrum are found to be highly utilized (*e.g.*, $\geq 75\%$) [1], by primary users, while some other sections of the spectrum are scarcely used ($\leq 5\%$) or unused for a long period of time. Cognitive radio [2] is capable of sensing idle frequencies (*i.e.*, white spaces), and usage of such frequencies can lead to enhanced spectral utilization. The primary users are granted the privilege of accessing their licensed channels according to their traffic demand. The secondary (unlicensed) users attempt to identify idle channels, unoccupied by the primary users, for their own use.

Modeling of spectrum occupancy or availability is challenging due to the stochastic arrival and departure of primary users in licensed bands. A two-fold model needs to be developed, first for each channel and second, over a spectrum (*i.e.*, a group of channels). A Beta distribution is an intuitive fit for channel occupancy due to the following facts: (i) broadcasting channels are *on-the-air* for majority of time with high probability, (ii) channels allocated for space navigation, telemetry, and governmental use are infrequently occupied (*i.e.*, low probability of usage), and (iii) moderate usage probability of channels for FM radio, paging, and text messaging. The author [3] assumed Beta distribution in his discussion on channel occupancy modeling and characterized the inherent parameters (α and β) of a Beta distribution substantiated with real-time measurements. However, it is essential to validate this assumption and to quantify the extent this distribution is valid for better comprehension of channel occupancy patterns of licensed users. To the authors' best knowledge, this paper is the first initiative in validating the Beta distribution for spectrum occupancy in the cognitive radio parlance.

Probabilistic modeling of spectrum occupancy was proposed by Ghosh *et al.* [4] assuming Poisson distribution of traffic arrival for primary users. Real-time measurements in the paging band (928-948 MHz) were used to substantiate their probabilistic model. Queuing theoretic modeling of spectrum occupancy by primary and secondary users simultaneously was proposed in [5] with the underlying assumption of Poisson traffic arrival rates for primary users. However, modeling of these occupancy probabilities is far more complex than just the wide assumption of Poisson distribution. Moreover, existing research have assumed an independent

and identically distributed (*i.i.d*) model for channel occupancy, and hence identical distributions for spectrum availability. This assumption does not reflect real-time temporal and spatial occupancy variations over a spectrum. In our proposed model, we have independence assumption but channel availability probabilities are stochastic and hence follow non-identical distributions.

Researchers have collected data on spectrum utilization by primary users over an appreciable length of time. McHenry et al. [1], [6] have monitored spectrum occupancy for different channels in Chicago as well as Vienna, Virginia over a certain length of time. We can exploit the vast trove of spectrum measurement data [1] for better real-time characterization of spectrum occupancy. The following entities are then instrumental in identifying channel availability for the secondary users:

Total number of free channels (N_{free}): At any given time instant, let N_{free} be the total number of channels that are not used by primary users. The entity N_{free} indicates how many secondary users can simultaneously access the network. The probability distribution, average, and variance of N_{free} indicate the extent of spectrum utilization by primary users.

Occupancy of a free channel's neighborhood: Given a free channel, the information about adjacent channels' occupancy is critical in mitigating interference to its adjacent primary users. For example, if two neighboring channels, *i.e.*, $(i-1)^{st}$ and $(i+1)^{st}$, of the i^{th} free channel are occupied by primary users, the transmission power from the i^{th} channel need to be adjusted in order not to interfere with its adjacent neighbors.

The knowledge of N_{free} facilitates in selection of preferable sections of the spectrum for further sensing, which can then be accordingly performed only on those selective channels with high probabilities of being available. Additionally, modeling of spectrum availability based on occupancy of its adjacent neighbors is especially important for networks defined by IEEE 802.22 standard, since the Federal Communications Commission (FCC) has restricted fixed devices from transmitting on adjacent channels of an active broadcasting channel. Therefore, analytical modeling of such classified channel types will be of special interest to the IEEE 802.22 networks. These models can be utilized in extracting information of selective sections in a spectrum that have available channels not adjacent to active broadcasting channels.

Our contribution in spectrum occupancy modeling is three fold:

- Validation of Beta distribution assumption for channel occupancy in a spectrum using real-time measurements on the 1500 MHz band inside a modern office building at Aachen, Germany;
- Proposition of Poisson-normal approximation of spectrum availability with independent, non-identical distributions of channel availability; and
- Classification of available channels based on occupancy of its adjacent channels and proposition of analytical models of such available channels.

The rest of the paper is organized as follows. Section 2 provides a brief introduction to Beta distribution and furnishes the results of the Kolmogorov-Smirnov test used for validating Beta distribution in spectrum occupancy modeling. Sec-

tion 3 deals with the theoretical analysis of spectrum availability, an useful information to the secondary users. Section 4 gives detailed probabilistic modeling of available channels based on occupancy of its adjacent channels. Simulation results are provided in Section 5 to illustrate preferable selection of spectrum based on models discussed in Section 4. Finally, Section 6 concludes the paper.

2. BACKGROUND ON BETA DISTRIBUTION

Beta distribution is used to generate p_i 's, $i = 1, 2, \dots, N$. The probability density function of a Beta distribution has two parameters $\alpha > 0$ and $\beta > 0$ and is given by:

$$f(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, 0 < x < 1, \quad (1)$$

where $B(\alpha, \beta)$ is the Beta function defined as:

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt. \quad (2)$$

The entity $B(\alpha, \beta)$ is a normalization constant to ensure that the function in Eq. (1) integrates to unity. Any continuous probability distribution on $(0, 1)$ can reasonably be approximated by a Beta distribution [7]. This is the main reason we have chosen Beta distribution for generating a set of channel free probabilities randomly.

2.1 Validation of Beta Distribution Assumption for Spectrum Occupancy

The Beta distribution assumption for primary user occupancy in a spectrum of channels is validated using real-time indoor measurements on a band (1500 MHz) with center frequency 770 MHz in Aachen, Germany [8] during December 2006 and January 2007. The measurements were conducted inside a modern office building (Latitude: $50^\circ 47' 24.01''$ North and Longitude: $6^\circ 3' 47.42''$ East) over 8192 frequency points with 200 KHz resolution bandwidth. Average sweep time is considered to be 1.8 s. The details of the measurement setup and spectrum analyzer configurations are specified in [8]. We have partitioned the obtained data set over seven days period into four intervals of interest: (i) morning 7:00 to 8:00 am, (ii) noon 12:00 to 1:00 pm, (iii) afternoon 3:00 to 4:00 pm, and (iv) midnight 11:00 pm to 12:00 am. The validation approach used in our paper is the Kolmogorov-Smirnov (K-S) test.

For each time interval under scrutiny, a random sample of channels is taken with occupancy probabilities between 0 and 1. The data consists of occupancy probabilities of the channels in the sample. The following table gives a summary of data analyzed. The various mean and occupancy over the four different data sets are:

7:00 - 8:00 am	: Sample size: 42, $\mu_{occ} = 0.4176$, $\sigma_{occ} = 0.1703$,
12:00 - 1:00 pm	: Sample size: 43, $\mu_{occ} = 0.3756$, $\sigma_{occ} = 0.1377$,
3:00 - 4:00 pm	: Sample size: 42, $\mu_{occ} = 0.4654$, $\sigma_{occ} = 0.1617$,
11:00 pm - 12:00 am	: Sample size: 43, $\mu_{occ} = 0.4571$, $\sigma_{occ} = 0.1396$,

where μ_{occ} and σ_{occ} are the mean and variance of the channel occupancy over 100 channels chosen randomly over the spectrum for a specified duration of an hour.

Table 1: Observed and Expected Frequencies of Spectrum Occupancy

Time of the day	Frequencies				
	0.0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1.0
<i>Observed</i> (7 – 8am)	21	5	5	1	15
<i>Expected</i> (7 – 8am)	22.27	3.72	3.11	3.49	14.4
<i>Observed</i> (12 – 1pm)	23	4	5	8	9
<i>Expected</i> (12 – 1pm)	23.3	5.56	4.55	4.76	10.83
<i>Observed</i> (3 – 4pm)	16	6	3	6	11
<i>Expected</i> (3 – 4pm)	16.86	3.98	3.41	3.85	13.9
<i>Observed</i> (11p – 12a)	15	10	1	4	12
<i>Expected</i> (11p – 12a)	15.6	5.16	4.47	4.87	11.9

The null hypothesis of interest is that the occupancy probabilities come from a Beta distribution. The Beta distribution has two parameters α and β and these parameters are estimated using the method of moments. The estimated $\hat{\alpha}$ and $\hat{\beta}$ parameters for the Beta distribution are computed from the obtained data using the following expressions [9]:

$$\begin{aligned}\hat{\alpha} &= \mu_{occ} \left(\frac{\mu_{occ}(1 - \mu_{occ})}{\sigma_{occ}} - 1 \right) \\ \hat{\beta} &= (1 - \mu_{occ}) \left(\frac{\mu_{occ}(1 - \mu_{occ})}{\sigma_{occ}} - 1 \right),\end{aligned}\quad (3)$$

The null hypothesis is tested using the K-S test [9]. The following are the results. The D-statistic and corresponding p-values [9] from the K-S test are as follows:

7:00 - 8:00 am : D = 0.1585, p-value = 0.1882,

12:00 - 1:00 pm: D = 0.0952, p-value = 0.766,

3:00 - 4:00 pm : D = 0.0867, p-value = 0.9103,

11:00 pm - 12:00 am: D = 0.1076, p-value = 0.7156. The

null hypothesis should be rejected if the p-value is ≤ 0.05 . It is evident that the p-value for all the four cases are < 0.05 , hence the null hypothesis of Beta distribution assumption for spectrum occupancy is accepted based on the KS test.

Estimated parameters $\hat{\alpha}$ and $\hat{\beta}$ for the Beta distribution are as follows:

7:00 - 8:00 am : $\hat{\alpha} = 0.1786$, $\hat{\beta} = 0.2492$,

12:00 - 1:00 pm: $\hat{\alpha} = 0.2642$, $\hat{\beta} = 0.4392$,

3:00 - 4:00 pm : $\hat{\alpha} = 0.2506$, $\hat{\beta} = 0.2879$,

11:00 pm - 12:00 am: $\hat{\alpha} = 0.3554$, $\hat{\beta} = 0.422$.

Based on the occupancy probabilities over 100 channels, we have classified these probabilities into five intervals of equal widths namely, 0.0 to 0.2, 0.2 to 0.4, 0.4 to 0.6, 0.6 to 0.8, and 0.8 to 1.0. The channels with zero probability and those with unit probability are ignored for the restriction imposed by the KS test validation approach. The number of occupied channels in each interval are specified as ‘Observed frequencies’ in Table 1. The ‘expected frequencies’ in Table 1 are obtained using the $\hat{\alpha}$ and $\hat{\beta}$ parameters for the Beta distribution. The observed frequencies and the Beta distribution with the estimated parameters is shown in Figure 1. The figure indicates that the Beta distribution is a good fit to the spectrum occupancy patterns in the observed 1500 MHz band under investigation.

3. PROBABILISTIC MODELING OF SPECTRUM AVAILABILITY

In this section, we present our approach of Poisson-normal approximation to spectrum availability. The availability

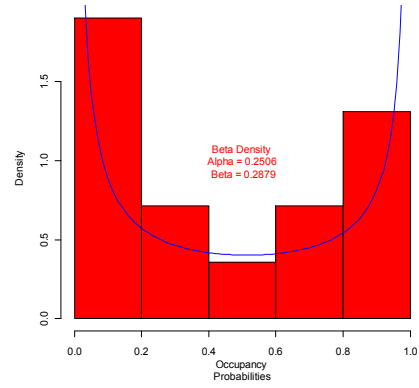


Figure 1: Histogram of occupancy probabilities of channels and plot with Beta density over real time measurements from 3 to 4 pm.

probability p_i of i^{th} channel can be small, moderate, or large. For this, we define a lower threshold P_{th1} and an upper threshold P_{th2} , which alternatively can be written as $0 < P_{th1} < P_{th2} < 1$. All p_i 's within the range $0 < p_i \leq P_{th1}$ relate to channels with small availability probabilities. For p_i 's $\geq P_{th2}$, channels are classified into a group with large probabilities of being available. Otherwise, channels within the range $P_{th1} < p_i < P_{th2}$ are categorized into a group with moderate probabilities. With this, we have the following definitions.

Definition 1 (Ch_{small}): This is a set of all channels having $0 < p_i \leq P_{th1}$. Let m be the size of Ch_{small} and $N_{free_{small}}$ be the number of free channels in Ch_{small} .

Definition 2 (Ch_{mod}): This is a set of all channels having $P_{th1} < p_i < P_{th2}$. Let n be the size of Ch_{mod} and $N_{free_{mod}}$ be the number of free channels in Ch_{mod} .

Definition 3 (Ch_{large}): This is a set of all channels having $p_i \geq P_{th2}$. The size of Ch_{large} is $(N - m - n)$. Let $N_{free_{large}}$ be the number of free channels in Ch_{large} .

Note that P_{th1} is typically close to zero and P_{th2} is close to one. In order to find the distribution of N_{free} , we first compute the approximate distributions for each of $N_{free_{small}}$, $N_{free_{mod}}$, and $N_{free_{large}}$. Therefore, the distribution of N_{free} can be computed by using the relation

$$N_{free} = N_{free_{small}} + N_{free_{mod}} + N_{free_{large}}. \quad (4)$$

Therefore, the probability of having k available channels is as follows:

$$\begin{aligned}Pr(N_{free} = k) &= \sum Pr(N_{free_{small}} = k_1, N_{free_{mod}} = k_2, N_{free_{large}} = k_3) \\ &\quad Pr(N_{free_{large}} = k_3) \\ &\simeq Pr_{Poi}(N_{free_{small}} = k_1) Pr_{Normal}(N_{free_{mod}} = k_2) \times \\ &\quad Pr_{Poi}(N_{free_{large}} = k_3) \\ &= Pr_{Poi-Normal}(N_{free} = k),\end{aligned}\quad (5)$$

where the summation is taken over all $k_1 \geq 0$, $k_2 \geq 0$, and $k_3 \geq 0$ with $k_1 + k_2 + k_3 = k$. The distribution of $N_{free_{small}}$ can be approximated by a Poisson distribution

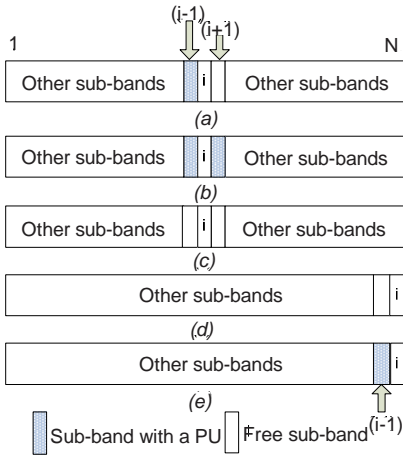


Figure 2: Types of free channels: (a) Type I, (b) Type II, (c) Type III, (d) Type IV, and (e) Type V; channels are also referred to as sub-bands here

and the probability that there are such k free channels is

$$\begin{aligned} Pr(N_{free_{small}} = k) &\simeq \frac{\lambda_s^k e^{-\lambda_s}}{k!} \\ &= Pr_{Poi}(N_{free_{small}} = k), \end{aligned} \quad (6)$$

where $\lambda_s = \sum_{i \in Ch_{small}} p_i$. This approximation follows a so-called Law of Rare Events. The distribution of $N_{free_{large}}$ has similar expression as in Eq. 6 with little algebraic simplifications [10]. The distribution of $N_{free_{mod}}$ in Ch_{mod} can be approximated by a normal distribution and the probability that there are k free channels is:

$$\begin{aligned} Pr(N_{free_{mod}} = k) &\simeq \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} \frac{1}{\sqrt{2\pi C_n}} e^{-\left(\frac{(x-\bar{N}_{mod})^2}{2C_n}\right)} dx \\ &= Pr_{Normal}(N_{free_{mod}} = k), \end{aligned} \quad (7)$$

where n is the size of Ch_{mod} , $k = 0, 1, \dots, n$, $\bar{N}_{mod} = E[N_{free_{mod}}] = \sum_{i \in Ch_{mod}} p_i$, and $C_n = \sum_{i \in Ch_{mod}} p_i(1-p_i)$ represents the variance of $N_{free_{mod}}$. Detailed analysis of the Poisson-normal approximation is omitted for space limitations. Please refer to [10] for further interest. Substituting the expressions from Eqs. 6 and 7 in Eq. 5 will result in the probability distribution of N_{free} .

Now, with the information of total number of available channels in a spectrum, we will focus our attention in classifying the channels based on its adjacent neighbors in the next section.

4. PROBABILISTIC MODELING OF AVAILABLE CHANNEL TYPES

Consider any interior free channel i , i.e., $i \neq 1$ and $i \neq N$. In this case, it has two adjacent neighbors. Each adjacent neighbor is either occupied by a primary user or free. This results in three different types of free channels based on the number of free adjacent neighbors. The number of possible free adjacent channels is either 0, 1, or 2. The three types are depicted in Figure 2 (b), (a), and (c), respectively. If the free channel is not interior, i.e., $i = 1$ or $i = N$, then two possible types arise. Figures 2 (d) and (e) depict the case $i = N$. The five types of free channels are defined below.

Type I channel: It is a free channel i having a primary user as the $(i-1)^{st}$ neighbor and a free channel as the $(i+1)^{st}$ neighbor, or vice versa as shown in Figure 2(a). Let $X_I(N)$ be the random variable that represents the total number of Type I channels in the N -spectrum. Possible values of $X_I(N)$ are $0, 1, \dots, N-2$.

Type II channel: It is a free channel i with two primary users as $(i-1)^{st}$ and $(i+1)^{st}$ neighbors as shown in Figure 2(b). Let $X_{II}(N)$ be the random variable that represents the total number of Type II channels in the N -spectrum. Possible values of $X_{II}(N)$ are $0, 1, \dots, N-2$.

Type III channel: It is a free channel i with two free channels as $(i-1)^{st}$ and $(i+1)^{st}$ neighbors, resulting in three contiguous free channels as shown in Figure 2(c). Let $X_{III}(N)$ be the random variable that represents the total number of Type III channels in the N -spectrum. Possible values of $X_{III}(N)$ are $0, 1, \dots, N-2$.

Type IV channel: It is a free channel i at the left or right edge of the spectrum with a free channel as its neighbor, resulting in two contiguous free channel on the edge of the spectrum as shown in Figure 2(d). Let $X_{IV}(N)$ be the random variable that represents the total number of Type IV channels in the N -spectrum. Possible values of $X_{IV}(N)$ are 0, 1, and 2.

Type V channel: It is a free channel i at the left or right edge of the spectrum with a primary user as its neighbor as shown in Figure 2(e). Let $X_V(N)$ be the random variable that represents the total number of Type V channels in the N -spectrum. Possible values of $X_V(N)$ are 0, 1, and 2.

Following the above definitions, we determine the probability distribution of $X_i(N)$, $i = I, II, III, IV$, and V . Recurrence relations are used to compute the probability distribution of $X_i(N)$.

4.1 Probability Distribution of $X_I(N)$

The approach described in Section 3 is not applicable for $X_I(N)$, since the probability of an available channel depends on its two neighbors. In view of this, we develop a recurrence relation to compute $X_I(N)$.

Proposition 1: Let $X_I(m+1)$ be the total number of Type I channels in the $(m+1)$ -spectrum represented by $\{1, 2, \dots, (m+1)\}$ with channel free probabilities $p_1, p_2, \dots, p_{(m+1)}$. Let $X_I(m)$ be the total number of Type I channels in the m -spectrum represented by $\{1, 2, \dots, m\}$ with channel free probabilities p_1, p_2, \dots, p_m . Then,

$$X_I(m+1) = \begin{cases} [X_I(m) + 1] \text{ with probability } p_I(m) \\ X_I(m) \text{ with probability } (1 - p_I(m)), \end{cases} \quad (8)$$

where $p_I(m) = (1-p_{(m-1)})p_m p_{(m+1)} + p_{(m-1)}p_m(1-p_{(m+1)})$, $m = (N-1), (N-2), \dots, 3$.

The proof of the proposition is omitted for space limitations. For further interest, please refer to [10]. Proposition 1 is used to compute the distribution of $X_I(m+1)$ once we know the distribution of $X_I(m)$ as detailed below. The possible values of $X_I(m)$ are $0, 1, 2, \dots, (m-2)$. From Eq. (8), it is clear that the number of possible values of $X_I(m+1)$ are $0, 1, 2, \dots, (m-1)$. A possible value of $X_I(m+1)$ is represented by r . For the computation of $Pr(X_I(m+1) = r)$, we identify three cases: $1 \leq r \leq (m-2)$; $r=0$; and $r=(m-1)$. If $1 \leq r \leq (m-2)$, the events $X_I(m+1)=r$ can arise from $X_I(m)=(r-1)$ or $X_I(m)=r$. Consequently, by Eq. (8), the

probability of r Type I channels is given by:

$$Pr(X_I(m+1) = r)$$

$$= p_I(m)Pr(X_I(m) = (r-1)) + (1-p_I(m))Pr(X_I(m) = r). \quad (9)$$

The event $X_I(m+1)=0$ can occur if and only if $X_I(m)=0$. Therefore,

$$Pr(X_I(m+1) = 0) = (1-p_I(m))Pr(X_I(m) = 0). \quad (10)$$

Finally,

$$Pr(X_I(m+1) = (m-1)) = p_I(m)Pr(X_I(m) = (m-2)). \quad (11)$$

Thus, once we know the distribution of $X_I(m)$, the distribution of $X_I(m+1)$ can be determined using Eqs. (9), (10), and (11). The ultimate goal is to find the distribution of $X_I(N)$. For this, we need the distribution of $X_I(3)$ from which we can find the distribution of $X_I(4)$ using Eqs. (9), (10), and (11). We continue this process until we reach $X_I(N)$.

4.2 Probability Distribution of $X_{II}(N)$

As evident from the analysis performed for Type I channels, we develop a similar recurrence relation for computing total number of Type II channels. The methodology is similar to that of Type I.

Proposition 2: Let $X_{II}(m+1)$ be the total number of Type II channels in the $(m+1)$ -spectrum represented by $\{1, 2, \dots, (m+1)\}$ with channel free probabilities $p_1, p_2, \dots, p_{(m+1)}$. Let $X_{II}(m)$ be the total number of Type II channels in the m -spectrum represented by $\{1, 2, \dots, m\}$ with channel free probabilities p_1, p_2, \dots, p_m . Then,

$$X_{II}(m+1) = \begin{cases} [X_{II}(m) + 1] \text{ with probability } p_{II}(m) \\ X_{II}(m) \text{ with probability } (1 - p_{II}(m)), \end{cases} \quad (12)$$

where $p_{II}(m) = (1 - p_{(m-1)})p_m(1 - p_{(m+1)})$, $m = (N-1), (N-2), \dots, 3$. The proof of the proposition is omitted for space limitations.

4.3 Probability Distribution of $X_{III}(N)$

The recurrence relation for computing the total number of channels of Type III is given by the following proposition.

Proposition 3: Let $X_{III}(m+1)$ be the total number of Type III channels in the $(m+1)$ -spectrum represented by $(1, 2, \dots, (m+1))$ with channel free probabilities $p_1, p_2, \dots, p_{(m+1)}$. Let $X_{III}(m)$ be the total number of Type III channels in the m -spectrum represented by $\{1, 2, \dots, m\}$ with channel free probabilities p_1, p_2, \dots, p_m . Then,

$$X_{III}(m+1) = \begin{cases} [X_{III}(m) + 1] \text{ with probability } p_{III}(m) \\ X_{III}(m) \text{ with probability } (1 - p_{III}(m)), \end{cases} \quad (13)$$

where $p_{III}(m) = p_{(m-1)}p_m p_{(m+1)}$, $m = (N-1), (N-2), \dots, 3$. The proof of the proposition is omitted for space limitations.

4.4 Probability Distribution of $X_{IV}(N)$

As every spectrum has two edges, the number of Type IV channels is 0, 1, or 2. The probability distribution of the

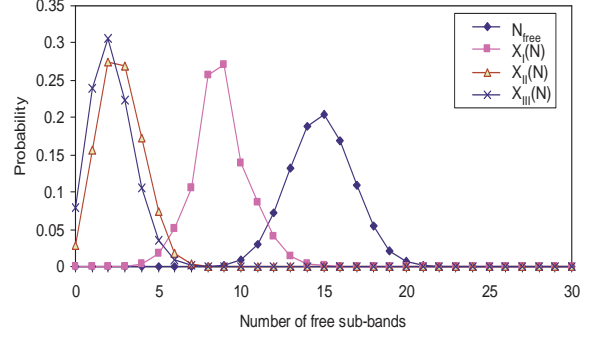


Figure 3: Comparison of probability distributions of N_{free} and $X_i(N)$, $i = I, II, III$ in a spectrum of 30 channels with 5 small and 5 large channel free probabilities

number of Type IV channels is given by:

$$\begin{aligned} P(X_{IV}(N) = 0) &= 1 - P(X_{IV}(N) = 1) - P(X_{IV}(N) = 2) \\ P(X_{IV}(N) = 1) &= p_1 p_2 [p_{N-1}(1 - p_N) + (1 - p_{N-1})p_N + \\ &\quad (1 - p_{N-1})(1 - p_N)] + \\ &\quad p_{N-1} p_N [(1 - p_1)p_2 + \\ &\quad p_1(1 - p_2) + (1 - p_1)(1 - p_2)] \\ P(X_{IV}(N) = 2) &= p_1 p_2 p_{N-1} p_N. \end{aligned} \quad (14)$$

The distribution of $X_V(N)$ is computed using the same approach. The analysis is omitted for space limitations.

5. SIMULATION RESULTS

By varying the values of the parameters α and β , we have generated two sets of probabilities resulting in two following scenarios:

Scenario-1: In this scenario we consider a spectrum with $N=30$ channels. A set of 30 channel availability probabilities p_i 's are generated from $Beta(0.5, 0.5)$ distribution. The output of the distribution has five small, five large, and 20 moderate p_i 's.

Scenario-2: This scenario also consists of $N=30$ channels. For simulation, a different Beta distribution is used to create an imbalance in small and large channel availability probabilities, unlike *Scenario-1*. A set of 30 channel availability probabilities p_i 's are generated from $Beta(0.6, 0.1)$ distribution. The output of the distribution has two small, 16 large, and 12 moderate p_i 's.

5.1 Probability Distribution of $X_i(N)$

Analytical model of the types of channels based on its neighbors is described in Section 4. In this sub-section, we show the simulation results of the distributions of $X_i(N)$, $i = I, II, III$. The distribution of $X_{IV}(N)$ and $X_V(N)$ are not included since each of them takes only three values, namely 0, 1, and 2 and their distributions have no significant contribution in our analysis for preferable selection of spectrum enabling efficient sensing.

The distributions of $X_i(N)$ depend on the spatial and temporal variations of the channel occupancy by the primary users as well on channel availability probabilities. For example, the distribution of $X_I(N)$ for Type I channels based

on p_1, p_2, \dots, p_N is different from the distribution of $X_I(N)$ when p_1, p_2, \dots, p_N are permuted. On the other hand, the distribution of N_{free} is invariant under permutations.

Figure 3 compares the distribution of $X_i(N)$'s with that of N_{free} under *Scenario-1*, in which Ch_{small} and Ch_{small} have the same cardinalities. Figure 4 deals with *Scenario-2* with an imbalance in the cardinality of Ch_{small} and Ch_{small} . The distributions of N_{free} , $X_I(N)$, $X_{II}(N)$, and $X_{III}(N)$ reflect the temporal and spatial occupancy of channels in each scenario. We compare some of the properties of distributions under *Scenarios-1* and *2*, in view of the fact that they have the same number of channels, i.e., $N = 30$. The numbers of small p_i 's and large p_i 's have an impact on the modes of the distributions. In the balanced case of 5 small and 5 large p_i 's (i.e., *Scenario-1*), the mode is 15, which is located at the center of the range of the distribution of N_{free} . In the unbalanced case of 2 small and 16 large p_i 's (i.e., *Scenario-2*), the mode of the distribution of N_{free} is 24. The distributions, on the other hand, of $X_i(N)$'s depend on the order of the probabilities of Type i channel and its neighbors.

The distributional analysis indicates the extent of proliferation of channels of each type. If we compare the distributional results from Figures 3 and 4, the spectrum associated with Figure 4 is preferable. This conclusion is based on the following observations:

Number of available channels (N_{free}): The mode of the distribution of N_{free} from Figure 4, namely 24, is much larger than the mode (i.e., 15) of the distribution of N_{free} from Figure 3. A secondary user is more likely to get a higher number of available channels from the spectrum of Figure 4.

Number of Type III channels: A Type III channel is preferable to Type I or Type II channel. The spectrum of Figure 4 gives a mode of 14 for the distribution of $X_{III}(N)$ where as the mode is 2 for the underlying spectrum of Figure 3. Therefore, the spectrum of Figure 4 is preferable in terms of Type III channels.

Number of Type II channels: A Type II channel is the least preferable since both the neighbors are occupied by primary users. The mode of distribution of $X_{II}(N)$ for the spectrum depicted in Figure 4 is zero where as the mode in Figure 3 is 2. On this count, the spectrum of Figure 4 is again preferable.

Hence, this preferential selection of one N-spectrum over the other can help in adaptive spectrum sensing, i.e., selecting a set of preferred channels for further sensing.

6. CONCLUSION

In this paper, the conventional assumption of Beta distribution for channel occupancy is validated extensively using real-time measurements in the 1500 MHz band conducted in Aachen, Germany. Further, we have proposed an approximation method for spectrum availability modeling with independent and non-identical distributions of channel availability over the spectrum. Additionally, with theoretical analysis of spectrum availability, we extend our research to classify available channels based on occupancy of its adjacent channels. Probabilistic modeling of such channel types are analyzed using recurrence relations. The simulation results indicate that these analytical models of spectrum availability and channel type classification will facilitate in efficient sensing over preferable sections of an operating spectrum.

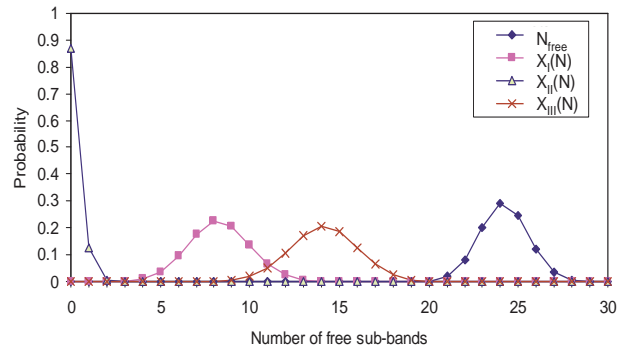


Figure 4: Comparison of probability distributions of N_{free} and $X_i(N)$, $i = I, II, III$ in a spectrum of 30 channels with 2 small and 16 large channel free probabilities

7. REFERENCES

- [1] M. A. McHenry and K. Steadman, "Spectrum occupancy measurements, location 2 of 6: Tyson's Square center, Vienna, Virginia, April 9, 2004," *Shared Spectrum Company Report*, August, 2005.
- [2] J. Mitola, III and G. Q. Maguire, Jr., "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, Vol. 6, Issue 4, Aug. 1999, pp. 13 - 18.
- [3] P. F. Marshall, "Closed-form analysis of spectrum characteristics for cognitive radio performance analysis," *IEEE DySpan*, Chicago, IL, 2008, pp. 1-12.
- [4] C. Ghosh, S. Pagadarai, D. P. Agrawal, and A. M. Wyglinski, "Statistical Spectrum Occupancy Modeling and Validation against real-time Measurements," *IEEE Transactions on Wireless Communications*, Vol. 9, No. 1, Jan. 2010, pp. 38-44.
- [5] L. Luo and S. Roy, "Analysis of Dynamic Spectrum Access with Heterogeneous Networks: Benefits of Channel Packing Scheme," *Proc. of IEEE Globecom*, Hawaii, 2009.
- [6] M. A. McHenry, P. A. Tenhula, D. McCluskey, D. A. Roberson, and C. S. Hood, "Chicago spectrum occupancy measurements & analysis and a long-term studies proposal," *Proceedings of the First International Workshop on Technology and Policy for Accessing Spectrum, TAPAS'06*, ACM International Conference Proceeding Series, Vol. 222.
- [7] P. M. Lee. *Bayesian Statistics: An Introduction*. Oxford University Press, Third Ed., 2004.
- [8] M. Wellens and P. Mahonen, "Lessons learned from an extensive spectrum occupancy measurement campaign and a stochastic duty cycle model," *Springer Mobile Networks and Applications*, August 2009. Available online: [http://dx.doi.org/10.1007/s11036-009-0199-9].
- [9] A. Stuart, K. Ord, and S. Arnold. *Kendall's Advanced Theory of Statistics, 2A*. London, pp. 25.37-25.43, 1999.
- [10] C. Ghosh and D. P. Agrawal *Innovative Approaches to Spectrum Selection, Sensing, and Sharing in Cognitive Radio Networks*. LAP Lambert Publishers, Saarbrucken, Germany, July 2010.