Efficient time domain deterministic-stochastic model of spectrum usage

Kenta Umebayashi, *Member, IEEE,* Masanao Kobayashi, and Miguel López-Benítez, *Senior Member, IEEE,*

***Abstract*—For achieving efficient spectrum sharing in a context of dynamic spectrum access, understanding the spectrum usage by licensed users (primary user: PU), is important for secondary users (SU). Duty cycle (DC) has been used to express the deterministic and stochastic aspects of spectrum usage. Specifically, a deterministic model for the mean of the duty cycle (M-DC) has been proposed in a previous work. The deterministic aspect of M-DC is affected by social behavior, and common habits of users, which can be confirmed in cellular systems. On the other hand, the observed DC (O-DC) during short time duration has randomness and a stochastic model is more suitable, e.g. distribution of O-DC. In this paper, we extend the conventional approach, in which only either the deterministic or stochastic aspect is considered, to a combined deterministic-stochastic (DS) model which represents both the deterministic and stochastic aspects at once. For the distribution of the O-DC, the beta distribution has been used as stochastic model, but we employ a mixture of beta distributions. The mixture-beta distribution can achieve higher accuracy but requires more capacity for data storage in spectrum usage measurements since it has a higher number of parameters than the beta distribution. For this issue, we employ regression analysis in DS-model since this approach can reduce the number of parameters while retaining the accuracy. We show the validity of DS-model based on exhaustive spectrum measurements in IEEE 802.11-based wireless local area networks and Long-Term Evolution uplink.**

Keywords*—Dynamic spectrum access, Spectrum measurement, Cognitive radio, Smart spectrum access, Duty cycle, Deterministic model.

I. Introduction

In the wireless communication field, spectrum scarcity is a pressing issue. Since most of the spectrum has been exclusively assigned to the licensed wireless systems, there are not enough spectrum resources for emerging wireless services. On the other hand, spectrum utilization measurement reports have revealed that the utilization rate of most of the licensed spectrum is not very high in spatial and/or temporal domains [1]. One possible solution for this issue is *spectrum sharing* since it can utilize the underutilized spectrum resource. In fact, several types of spectrum sharing have been investigated, such as dynamic spectrum access (DSA) [2], [3], licensed shared access (LSA) [4], and citizens broadband radio service (CBRS) [5], and they differ with types of priorities and licensing. Several kinds of spectrum sharing are surveyed in [6], [7].

K. Umebayashi, and M. Kobayashi are with the Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology, Tokyo 183-8538, Japan.

M. López-Benítez is with the Dept. Electrical Eng. & Electronics, University of Liverpool, L693GJ, UK .

Manuscript received Month Day, Year; revised Month Day, Year.

In this paper, we focus on DSA in which a secondary user (SU) can utilize spectrum licensed to a primary user (PU), while the spectrum is not occupied by the PU [2], [3]. This approach is expected to achieve efficient spectrum utilization. In DSA there are mainly two key techniques: spectrum sensing to find vacant spectrum and spectrum management techniques, such as spectrum allocation and channel access, to utilize the vacant spectrum efficiently. The key techniques can be designed properly and enhanced by information of PU spectrum usage since this information indicates trends and aspects of the PU spectrum usage. For example, the knowledge of the duty cycle (DC) can enhance spectrum sensing performance [8], [9] and spectrum management [10]–[13].

In [14], a concept of smart spectrum access (SSA), which corresponds to sophisticated DSA, has been presented. Two-layered SSA has been presented as a practical configuration for SSA. This two-layered SSA concept consists of DSA as the first layer and a spectrum awareness system (SAS) as the second layer. SAS is dedicated for spectrum usage measurements to extract useful prior information so that SUs in DSA can be relieved of having to perform spectrum usage measurements. The spectrum usage information, based on long term, wide band, and broad area spectrum measurements, is stored in a data base in SAS. A large number of sensors have to be deployed for the measurements. At a sensor, the observed signal is just a sequence of I-Q samples which are converted into statistical information, such as DC. For storing the obtained statistical information efficiently, proper spectrum usage modeling is required.

There have been many investigations of spectrum usage modeling based on measurement campaigns and most of them are introduced in [13], [15], [16]. The measurement campaigns have shown the DC and measured power of various spectrum bands and various measurement sites [17]–[21]. This type of measurement campaigns has simply evaluated spectrum utilization and potential spectrum opportunities (i.e., how much spectrum in a certain dimension, such as time and space, is vacant). There are also detailed spectrum usage models in time, frequency and space domains, respectively [22]–[25]. In these investigations, not only the spectrum utilization ratio, but also the trends and variations of feature quantities in terms of spectrum utilization, such as DC [22], signal strength [24] and busy time (vacant time) [26], [27], have been considered. In addition, the extracted information can enhance the key technologies in the spectrum sharing. Specifically, a model considering statistics of signal strength and arrival rate of mobile service was proposed in [24] and the extracted information is used for an optimization of sensing strategy. In [27], statistics of channel busy time and idle time are

used to design channel selection algorithm in a multichannel cognitive radio network. In this paper, we focus on time-domain spectrum usage modeling in terms of DC since this information can enhance the spectrum sensing performance [9], [28].

Two-state Markov chains have been used to model the behavior of spectrum usage in the time domain [22], [29]. Specifically, the aspect of transition between "busy" and "idle" can be described by a set of transition probabilities in Markov chain. In fact, statistics of spectrum usage in the time domain, such as DC, transition probabilities, and state holding times vary over time. This time-varying aspect depends on the spectrum usage (i.e., social behavior and common habits). The time-varying aspect can involve a deterministic behavior, e.g., in the cellular system, the traffic load is typically high in daytime and low in nighttime and this provides the deterministic behavior of the DC. In this case, the time variation can be described by a deterministic model [22]. The deterministic behavior has a time period of one day for cellular systems. Therefore, several days are necessary for spectrum usage measurements to set proper parameters for the deterministic model.

In spectrum usage measurements with short time duration, such as one second, the observed duty cycle (O-DC) is random. The randomness of O-DC can be expressed by its probability density function (PDF), which corresponds to a stochastic model [22]. On the other hand, in the deterministic model, the mean of O-DC (M-DC) is used to describe the deterministic behavior. In previous works, the models characterize either the deterministic or stochastic behavior, but not both of them simultaneously.

In this paper, we extend the previous model in [22] to a deterministic-stochastic-DC (DS-DC) model in which both deterministic and stochastic behaviors can be expressed at once. We investigate a spectrum usage model which can describe not only the stochastic behavior, but also the deterministic behavior in the time domain. There are two issues in this investigation. First, we attempt to combine the stochastic and deterministic models, but how the models can be combined is not straightforward. Second, typically there is trade-off between accuracy of model and the number of parameters. Specifically, efficient models should reproduce the actual distribution, such as empirical distribution of O-DC, accurately while the number of parameters is relatively low. We investigate a method to achieve efficient modeling with a low number of parameters that can retain adequate accuracy. In this investigation, we need empirical data of O-DC, which are obtained by means of spectrum measurements.

Our main contributions are summarized as follows:

- For the PDF of O-DC, the beta and Kumaraswamy distributions have been used as stochastic models [22]. To achieve better accuracy, a mixture-beta distribution, in which two beta distributions are used, is employed as a stochastic DC model. We verify the accuracy of the mixture-beta distribution stochastic model by spectrum measurement campaigns.
- Two types of DS-DC models are presented in this paper. In the first DS-DC model, a deterministic model is used for each parameter of the stochastic DC model (i.e., the mixture-beta distribution used for the PDF of O-DC). While the mixture-beta distribution can be more accurate than the beta distribution, the mixture-beta distribution needs more parameters. For this issue, we use polynomial regression analysis to reduce the number of parameters, which is denoted by RDS-DC model. This reduction of parameters can reduce the data storage burden in the SAS.
- We verify the efficiency of the RDS-DC model based on spectrum measurement campaigns in a 2.4 GHz 802.11 wireless local area network (WLAN), and an 800 MHz Long Term Evolution (LTE) uplink. The RDS-DC model requires less parameters and can retain the same accuracy.

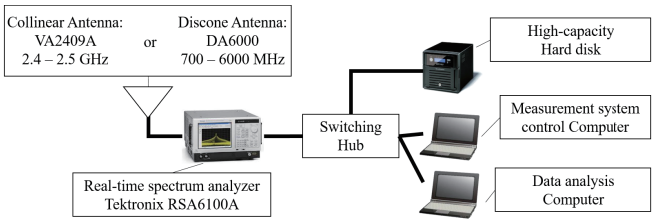


Fig. 1: Measurement system

The remainder of this paper is organized as follows. First, Section II describes the measurement setup (including signal processing) employed to obtain O-DC. Section III presents spectrum usage models proposed in previous works, namely a deterministic model for M-DC. In Section IV, stochastic models based on beta distribution and mixture-beta distribution for O-DC are shown. The stochastic model based on mixture-beta distribution corresponds to our proposal and we evaluate its validity by spectrum usage measurements. The proposed DS-DC and RDS-DC models are presented in Section V. In Section V-D, numerical evaluations based on two spectrum usage measurement campaigns are shown. Finally Section VI concludes this paper.

II. Measurement setup and methodology

The measurement setup and methodology to obtain the empirical data of O-DC are shown in this section. We performed two spectrum usage measurement campaigns in two frequency bands: the first frequency band W_1 is 2452 - 2472 MHz, mainly utilized by IEEE 802.11 WLAN, and the second frequency band W_2 is 835 - 845 MHz, utilized by LTE uplink. We use W_1 and W_2 to denote the two measurement campaigns. The measurement system is located in our laboratory on fourth floor of a building in Koganei-campus, Tokyo University of Agriculture and Technology, Tokyo, Japan (35°41'55.8"N 139°31'00.6"E).

The block diagram of the measurement system is shown in Fig. 1. The measurement system consists of antennas that can observe the target frequency bands W_1 and W_2, cables, a real-time spectrum analyzer (RSA) (Tektronix RSA6100A), a

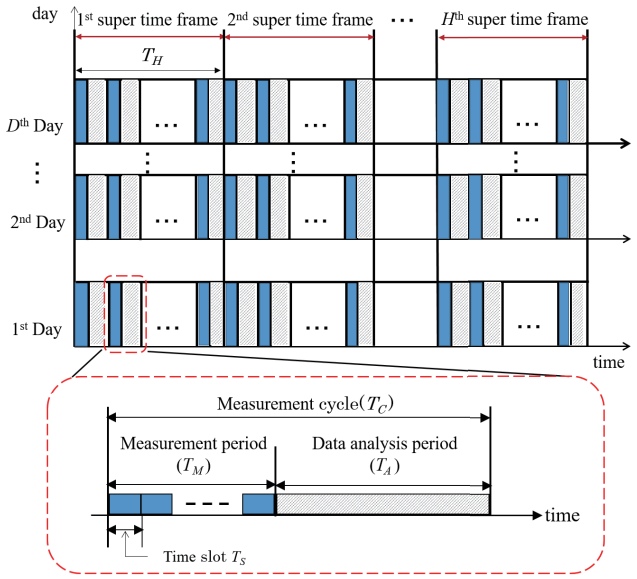


Fig. 2: Measurement time schedule

network hard disk, a measurement system control computer, and a data analysis computer.

The measurement system control computer takes care of the measurement time scheduling, which is shown in Fig. 2. The number of days for spectrum measurement is denoted by D. Since we focus on the daily deterministic behavior of the spectrum usage, the time duration of the deterministic model is set to one day as in [22]. While deterministic behaviors of spectrum usage in weekdays and weekends may be different, we only focus, without loss of generality, on the spectrum usage during weekdays. In the spectrum usage measurement campaigns, we set $D = 29$ days for W_1 and $D = 21$ days for W_2. The difference in terms of D in both measurements is due to the storage limitation of the high-capacity hard disk. One day (24 hours) is divided into H super time frames, each of which consists of M measurement cycles. The time durations for one super time frame, and one measurement cycle are denoted by T_H and T_C, respectively. In the two measurement campaigns, we set T_H and T_C to one hour and one minute, respectively. We assume that O-DC during T_H approximately maintains stationarity and the statistics of O-DC are estimated during each super time frame. To validate the stationarity of the observed O-DC series for one hour, we employ augmented Dickey-Fuller test with significance level 5% and we have confirmed that the measurement results pass the test.

One measurement cycle consists of a measurement period and a data analysis period, whose time durations are denoted by T_M and T_A, respectively. Typically T_M depends on the capability of the spectrum measurement devices, such as internal buffer size and sampling rate. During one measurement period, the RSA observes the target frequency (either W_1 or W_2) for T_M seconds. Note that the observation of the RSA is continuous unlike typical swept spectrum analyzers. T_M has to be much longer than one continuous spectrum usage cycle, such as data packet, for proper DC estimation in the target frequency bands (W_1 and W_2). The time durations for data packet in WLAN and LTE are at most about 0.87 ms, (corresponding to the time duration of the IEEE 802.11 PLCP (Physical Layer Convergence Procedure) protocol data unit), and 0.93 ms (corresponding to the time duration for the physical uplink shared channel in LTE), respectively. Based on this, we set $T_M = 100$ ms.

The observed data is first stored in the network hard disk and then transferred to the data analysis computer. The data analysis computer provides estimates of the DC by means of fast Fourier transform (FFT)-based energy detection and post processing to achieve accurate spectrum usage detection performance [30]. The estimated DC is denoted by $\Psi_E(c, h, d)$, where c, h, d, indicate the index numbers for the measurement cycle, super time frame, and day, respectively. This $\Psi_E(c, h, d)$ corresponds to O-DC. Since T_A has to be long enough to complete the data transfer and the data analysis, we set $T_A = 59.9$ sec, which was proven to be sufficient in our measurement system.

The measurement period is divided into N_T time slots and Welch FFT based power spectrum estimation is performed in each time slot. There are N_F frequency bins in one time slot. The time duration for one time slot is denoted by T_S and this time slot corresponds to one Welch FFT time duration. T_S has to be shorter than one continuous spectrum usage cycle (e.g., one data packet). The spectrum usage detections are performed at the data analysis computer based on the estimated power spectrum for $N_T \times N_F$, energy detection (ED), and signal area estimation with false alarm cancellation [30]. The parameters for the Welch FFT based ED are as follows. In Welch FFT, 1024 data samples are divided into 15 segments while the overlap ratio is set to 0.5 [31]. Therefore, the number of frequency bins are set to 128. We set the threshold based on constant false alarm rate criterion where the target false alarm rate is set to 0.01. In this criterion, we need noise floor information in order to set the threshold and we employ forward consecutive mean excision (FCME) algorithm for noise floor estimation [32], [33]. The displayed average noise level of Tektronix RSA6100A is -151 dBm/Hz.

The outputs of the spectrum usage detection are denoted by

$$D_{n_T,n_F} = \begin{cases} 1 & \text{(spectrum is occupied)} \\ 0 & \text{(spectrum is vacant)}, \end{cases} \tag{1}$$

where n_T is the time slot index number and n_F is the frequency bin index number. We define a set of index numbers of frequency bin, n_F, involved in W_i as $\mathbf{W}_i$. Now O-DC in the frequency band W_i can be obtained by

$$\Psi_E(c, h, d) = \frac{1}{N_T} \sum_{n_T} \left(1 - \prod_{n_F \in \mathbf{W}_i} (1 - D_{n_T,n_F}) \right). \tag{2}$$

This equation indicates that if a part of the target frequency band $\mathbf{B}_i$ is occupied, the state of the whole target frequency band is detected as occupied as well. In case of OFDM (Orthogonal Frequency Division Multiplexing) based communication, a few sub-carriers can be used at a particular time.

In this case, subset of the channel is physically occupied by PU at that time, but the whole channel is reserved for the PU. Therefore, (2) is a convenient and reasonable way to define O-DC for PU protection.

The estimated M-DC at the hth super time frame, $\Psi_E(c,h,d)$, is obtained by averaging over c and d as

$$\Psi_M(h) = \frac{1}{D \cdot M} \sum_d \sum_c \Psi_E(c,h,d), \tag{3}$$

The deterministic behavior of spectrum usage, which depends on social behavior and common habits, is determined in this work by the time schedule in the laboratory: the laboratory members arrive in the office at 9:00 and leave anytime between 17:00 and 22:00.

III. Deterministic model for DC

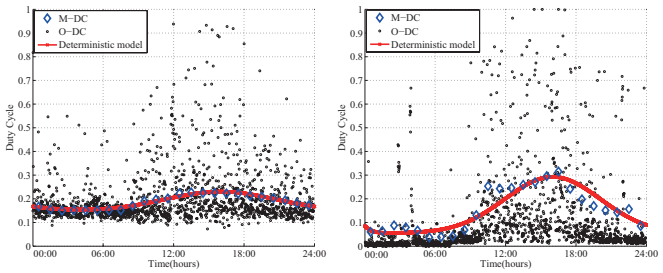
In this section, we introduce the deterministic model proposed in [22]. The deterministic model shows the deterministic behavior of M-DC, Ψ_M, and it is defined by

$$F_{\Psi_M}(t) = F_{min} + \sum_{k=0}^{K-1} A_k e^{-(\frac{t-\tau_k}{\sigma_k})^2} \ (0 \le t \le T) \tag{4}$$

where F_{min} is the minimum of $F_{\Psi_M}(t)$, the term $A_k e^{-(\frac{t-\tau_k}{\sigma_k})^2}$ is the kth bell shaped exponential term, A_k is the maximum amplitude for the kth bell shaped exponential term, τ_k is the central time to determine the location of the kth bell shaped exponential term, and σ_k determines the width of the kth bell shaped exponential term. The number of parameters involved in this deterministic model is $N(F_{\Psi_M}|K) = 1 + 3K$. This model corresponds to the deterministic model for low-medium loads proposed in [22].

Now we confirm the validity of the model in (4) based on the two measurement campaigns W_1 and W_2. In Figs. 3a and 3b, $\Psi_M(h)$ (M-DC), $\Psi_E(c,h,d)$ (O-DC), and $F_{\Psi_M}(t)$ for WLAN (W_1) and LTE (W_2) are plotted, respectively. The ratio of the size of one frequency bin n_F to the total bandwidth W_i is $1/128$ in both results. Each point corresponding to O-DC is obtained by averaging the measurements over one measurement cycle with duration $T_C = 1$ minute (see Fig. 2). Each point corresponding to M-DC is obtained by averaging the measurements over one super time frame with a duration of $T_H = 1$ hour (see Fig. 2). In $F_{\Psi_M}(t)$, the least square error criterion is used for parameter fitting. In both cases (W_1 and W_2), the deterministic model $F_{\Psi_M}(t)$ is fitted to the M-DC points $\Psi_M(h)$. In WLAN, the minimum $\Psi_E(c,h,d)$ is around 0.1, which is caused by periodic beacon signals from WLAN access points (APs). On the other hand, the minimum $\Psi_E(c,h,d)$ of LTE can be as low as zero. In W_2, even though the fit is not perfect for every single point, which is not uncommon with empirical data as observed in related studies, e.g., [22], the overall fit is satisfactory.

In both measurements results, the DC is relatively high during the day time as a result of human presence in the laboratory. In addition, in the case of relatively high M-DC, $\Psi_E(c,h,d)$ can have a significantly greater variance and the maximum $\Psi_E(c,h,d)$ can be larger than 0.9. This suggests that the deterministic model may be applicable for the variance of $\Psi_E(c,h,d)$, which will be discussed in detail in Section V-A.



(a) Measurement campaign W_1. (b) Measurement campaign W_2.

Fig. 3: Measured DC (M-DC and O-DC) and M-DC based on model ($\Psi_D^{\mathrm{L/M}}(t)$) as a function of time.

IV. Stochastic models for DC

A. Stochastic model based on Beta distribution for O-DC

If the spectrum measurement time duration is not long enough, such as one hour, the deterministic aspect may not be noticeable and only the stochastic behavior component may be present. For example, in Figs. 3a and 3b, $\Psi_E(c,h,d)$ is approximately stationary when observed in one-hour periods but not when observed across different super frames. In such a case, a stochastic model can express the stochastic behaviors of O-DC.

In stochastic DC models, typically a basic PDF is employed. The beta distribution is one of the strong candidates to describe the stochastic aspect of O-DC $\Psi_E(c,h,d)$ in the hth super frame [22], [34]–[36]. The beta distribution has two preferable properties. The first one is that its domain is the interval $[0,1]$ which is the same domain as that of the DC. The second one is that the beta distribution is a conjugate prior distribution, therefore it is useful for Bayesian estimation.

Based on the beta distribution, a stochastic model for O-DC $\Psi_E(c,h,d)$ is given by

$$f_{\Psi_E,B}(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \tag{5}$$

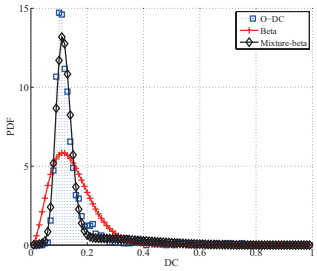
where $0 \le x \le 1$, $\alpha > 0$ and $\beta > 0$ are shape parameters, and $B(\alpha,\beta)$ is the Beta function defined by:

$$B(\alpha,\beta) = \int_0^1 z^{\alpha-1}(1-z)^{\beta-1} dz. \tag{6}$$

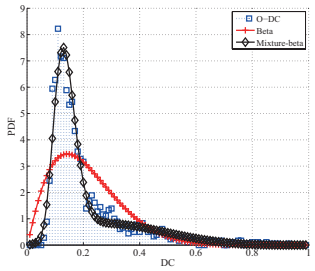
The mean and variance of the beta distribution, μ_B and σ_B^2 are given by

$$\mu_B = \frac{\alpha}{\alpha+\beta} \tag{7}$$

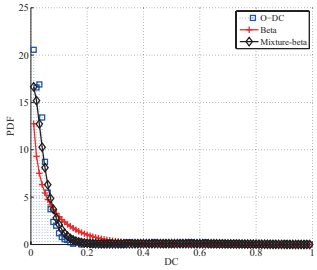
$$\sigma_B^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}, \tag{8}$$



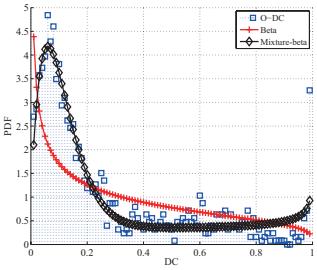
(a) Measurement campaign W_1, Time: 1:00 - 2:00. DKL for beta distribution and mixture-beta distribution are 0.74 and 0.09, respectively.



(b) Measurement campaign W_1, Time: 14:00 - 15:00. DKL for beta distribution and mixture-beta distribution are 0.43 and 0.04, respectively.



(c) Measurement campaign W_2, Time: 1:00 - 2:00. DKL for beta distribution and mixture-beta distribution are 0.77 and 0.09, respectively.



(d) Measurement campaign W_2, Time: 14:00 - 15:00. DKL for beta distribution and mixture-beta distribution are 0.41 and 0.14, respectively.

Fig. 4: Empiric distribution of the measured DC (O-DC) and PDF based on the stochastic DC models.

respectively. The mean, μ_B, from the hth super frame, is equivalent to $\Psi_M(h)$. We employ the Metropolis-Hastings curve-fitting algorithm, which is a Markov chain Monte Carlo method, to set the model parameters [37].

B. Stochastic model based on mixture Beta distribution for O-DC

As an alternative model, we propose a mixture-beta distribution in which two beta distributions are used. The stochastic model based on the mixture-beta distribution is defined by

$$\begin{aligned} f_{\Psi_E,B_m}(x) &= \sum_{b=0}^{1} w_b f_{\Psi_E,B_b}(x) \\ &= \sum_{b=0}^{1} w_b \frac{1}{B(\alpha_b,\beta_b)} x^{\alpha_b-1}(1-x)^{\beta_b-1}, \end{aligned} \quad (9)$$

where w_b is the weighting coefficient for the bth beta distribution ($b \in \{0,1\}$), $f_{B_b}(x)$ is the bth beta distribution as defined in (5), and $\alpha_b > 0$ and $\beta_b > 0$ are the shape parameters. Since we consider two beta distributions, $w_0 + w_1 = 1$, $w_0 \geq 0$ and $w_1 \geq 0$. The mean and variance for the bth beta distribution are denoted by μ_{B_b} and $\sigma^2_{B_b}$. They are available from (7) and (8), respectively. The parameters are related as follows:

$$\mu_{B_m} = \sum_{b=0}^{1} w_b \mu_{B_b}, \quad (10)$$

The number of parameters of the beta and mixture-beta distribution are $N(f_{\Psi_E,B}(x)) = 2$ and $N(f_{\Psi_E,B_m}(x)) = 5$, respectively. Notice that the mixture-beta model could be extended to more than two beta distributions, however the evaluations will indicate that two beta distributions already provide a reasonable level of accuracy. This fact indicates that there is no need to use more than two beta distributions since this would unnecessarily increase the number of parameters and the data storage burden in the spectrum measurement system without providing noticeable accuracy improvement. Therefore, we only consider two beta distributions in the mixture-beta model.

C. Comparison of stochastic models

We will compare the validity of the stochastic models by means of the Kullback-Leibler divergence (DKL). For discrete probability distributions, $g(x_i)$ and $f(x_i)$, DKL is defined by

$$\mathrm{DKL} = \sum_i g(x_i) \log\left(\frac{g(x_i)}{f(x_i)}\right). \quad (11)$$

This is a measure of the non-symmetric difference between two probability distributions. The DKL which has been commonly used to measure the difference between two probability distributions, i.e., evaluation of O-DC model in [22], [34], takes non-negative values. Specifically, smaller DKL indicates that two probability distributions are more similar, e.g., $\mathrm{DKL} = 0$ indicates that $g(x_i)$ and $f(x_i)$ are identical. It is worth noting that, while the DC is by definition a continuous parameter with a continuous distribution, the empirical distribution obtained from experimental measurements is discrete as a result of the binning required to calculate the histogram (i.e., the empirical PDF). Therefore, a discretized version of the continuous beta distribution model needs to be defined for comparison with the experimental results in order to calculate DKL. Since the domain of the beta distribution is $[0,1]$, $0 \leq x_i \leq 1$, we define $\Delta x = 1/N_D$ where N_D is the number of partitions for the interval (in this work we set $N_D = 100$). Specifically, $x_i = (i + 1/2) \times \Delta x$, where $i = 0, 1, \cdots N_D - 1$. Finally, $f(x_i)$ for the beta distribution $f_{\Psi_E,B}(x)$ in (5) of the paper is given by

$$f(x_i) = \int_{x=x_i-\Delta x}^{x=x_i+\Delta x} f_{\Psi_E,B}(x)dx.$$

Figs. 4a and 4b show the measured O-DC at W_1 (WLAN) and the fitted distributions based on the beta (5) and the mixture-beta (9) models, respectively, for different time periods (1:00-2:00 and 14:00-15:00, respectively). The time periods 1:00-2:00 and 14:00-15:00 correspond to typical examples of relatively low M-DC case and relatively high M-DC case, respectively. In all cases, the model based on the mixture-beta

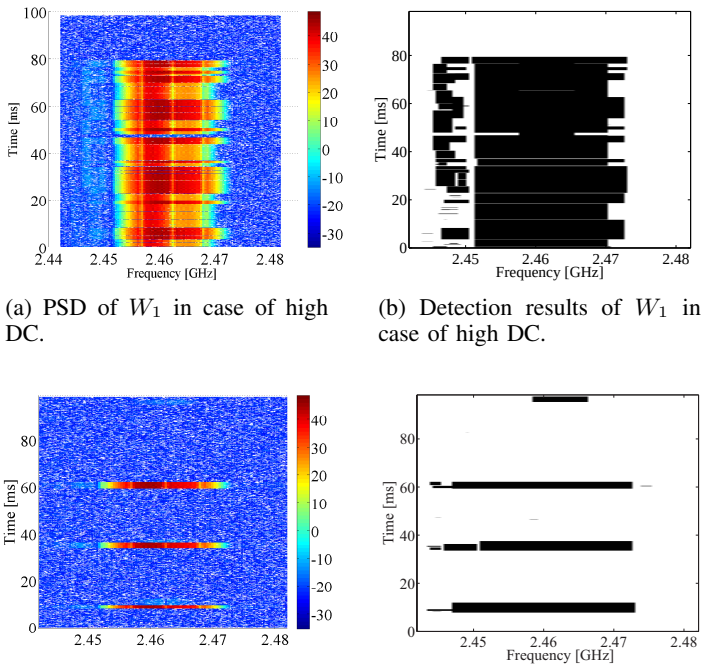
distribution agrees better with the measured O-DC compared to the beta distribution. In WLANs the beacon signals provide a narrow peak at DC=0.1 in Figs. 4a and 4b for W_1. The beta distribution cannot express this narrow peak adequately. In the case of daytime (Fig. 4b: 14:00-15:00), the traffic provides a heavier right tail in the region where the DC is greater than 0.2. This aspect can be well described by the mixture-beta distribution, but not by the beta distribution.

Figs. 4c and 4d show the measured O-DC at W_2 (LTE) and the fitted distributions based on the beta (5) and mixture-beta (9) models, respectively, for different time periods (1:00-2:00, and 14:00-15:00). Again the model based on the mixture-beta distribution agrees well with the measured O-DC. In Fig. 4c, the curve is monotonically decreasing and this aspect can be described by both models. However, the model based on the mixture-beta distribution provide a more accurate fit to the measured O-DC. In Fig. 4d, it is observed that there is a narrow peak at DC=0.1 but this shape cannot be fitted by the beta distribution while the mixture-beta distribution provides a good fit. In the region where the DC is greater than 0.3, the curve provided by the beta distribution is above the curve provided by the measured O-DC in most DC, on the other hand the mixture-beta distribution again provides a better fit. DKL performances are shown in the captions of Figs. 4a, 4b, 4c and 4d, which also shows the benefit of the mixture-beta distribution.

There are two reasons why the mixture distribution is a more convenient model. Firstly, the mixture-beta distribution can provide more accurate fits to empirical data since it has more the parameters (i.e., degrees of freedom) than the beta distribution, i.e., $N(f_{\Psi_E,B}(x)) = 2 < N(f_{\Psi_E,B_m}(x)) = 5$. Secondly, measured O-DC during one super time frame at a given location may include various types of traffic such as periodic beacon signals and random data packets. In this case, a mixture distribution may be more appropriate since each traffic follows a different distribution of O-DC and the model distribution has to describe the mixture of the different distributions. The existence of different types of traffic during one super time frame at the same location can be confirmed in the actual measurement results shown in Figs. 5. In Fig. 5a, measured PSD (power spectrum density) at 13:10 is shown while Fig. 5b shows the corresponding detection results (white and black correspond to $D_{n_T,n_F} = 0$ and $D_{n_T,n_F} = 1$, respectively). There are both bursty data packets and beacon signals. In Fig. 5c and Fig. 5d, PSD and detection results at 13:13 are shown and there are only periodic beacon signals. As it can be appreciated, there may be cases where the observed spectrum occupancy is the combination of different types of traffic, each of which may have a different distribution, and therefore a mixture distribution is a more appropriate model.

V. Deterministic-Stochastic model for DC

As confirmed in the previous section, the deterministic model can express the deterministic behavior of one statistic, such as M-DC, and the stochastic model provides whole statistical information at a given time. To express both deterministic and stochastic aspects at once, a DS model for DC is proposed in this section.



(a) PSD of W_1 in case of high DC.

(b) Detection results of W_1 in case of high DC.

(c) PSD of W_1 in case of low DC.

(d) Detection results of W_1 in case of low DC.

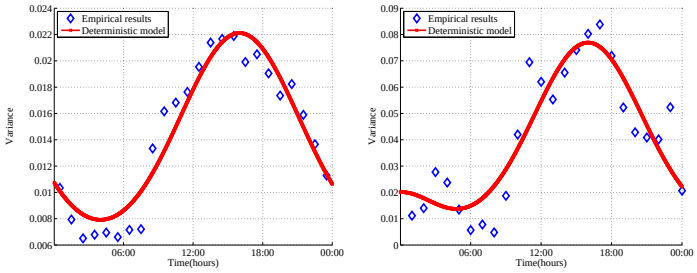
Fig. 5: PSD and detection results of W_1 in cases of high and low DC.

First, we will show that the variance of O-DC can be also described by the deterministic model. If it is possible, it implies that the whole statistical information may be expressed by a deterministic model and this corresponds to the DS-DC model. We will show two DS-DC models based on the mixture-beta distribution. The first model is based on a straightforward approach while the second model is based on regression analysis.

A. Deterministic model for the variance of the DC

As confirmed in Figs. 3a and 3b, the variability of O-DC is relatively low in the night, e.g., 3:00, and relatively high in the daytime, e.g., 15:00. During night, the number of users is relatively low and it leads to that low variance on the O-DC. On the other hand, during day-time, the number of users increases and there are various types of data traffic, which includes high O-DC and low O-DC (see Fig. 5). This leads to high variance on the O-DC. This fact suggests that the deterministic model in (4) may be applicable to express the deterministic behavior of the variance as well.

Comparisons between the measured variance and the deterministic model for variance in W_1 and W_2 are shown in Figs 6a and 6b. In both cases, the deterministic model can express the deterministic aspect of the variance. This extension should be available for all parameters in the stochastic models and it



(a) Measurement campaign W_1 (b) Measurement campaign W_2

Fig. 6: Measurement of O-DC variance and deterministic model of O-DC variance as a function of time.

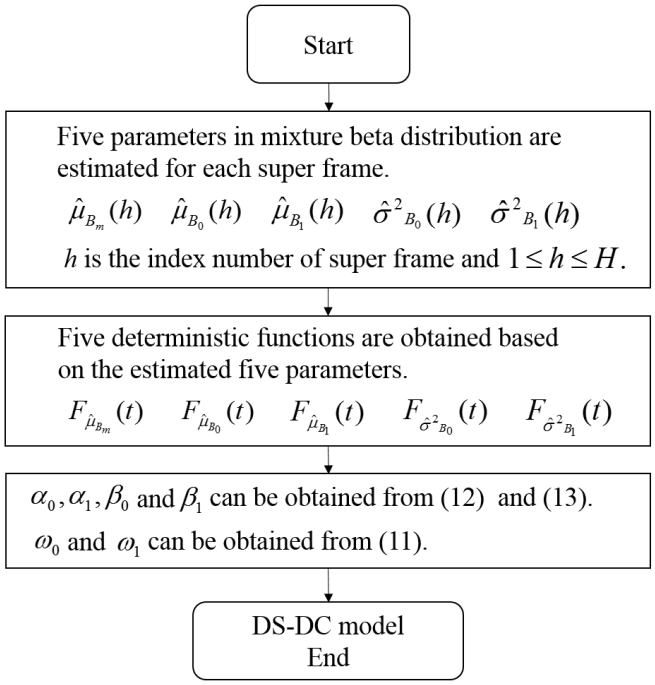


Fig. 7: Flow chart for derivation of DS-DC model based distribution parameters.

is used in the first DS-DC model, which is presented in the following subsection.

B. Deterministic-stochastic model for DC

The DS-DC model is a straightforward approach to combine the deterministic and the stochastic (mixture-beta distribution) models. In the DS-DC model, the deterministic model is used to describe the deterministic behavior of each parameter in the stochastic model. In the mixture-beta distribution, five parameters are necessary to determine the distribution (9). We apply the deterministic model (4) to the five parameters, μ_{B_m}, μ_{B_0}, μ_{B_1}, $\sigma^2_{B_0}$, and $\sigma^2_{B_1}$. The derivation process for parameters of a distribution based on the DS-DC model is shown below and the flow chart of the derivation is shown in Fig. 7. Specifically, during one super frame, the five parameters can be estimated; such estimates are denoted by $\hat{\mu}_{B_m}(h)$, $\hat{\mu}_{B_0}(h)$, $\hat{\mu}_{B_1}(h)$, $\hat{\sigma}^2_{B_0}(h)$, and $\hat{\sigma}^2_{B_1}(h)$. By curve fitting, such as the least-square method, each parameter in the deterministic model can be specified and the obtained deterministic functions are denoted by $F_{\hat{\mu}_{B_m}}(t)$, $F_{\hat{\mu}_{B_0}}(t)$, $F_{\hat{\mu}_{B_1}}(t)$, $F_{\hat{\sigma}^2_{B_0}}(t)$, and $F_{\hat{\sigma}^2_{B_1}}(t)$. Note that $F_{\hat{\mu}_{B_m}}(t)$ is equivalent to $F_{\Psi_M}(t)$. A set of functions can provide $\mu_{B_m}(t)$, $\mu_{B_0}(t)$, $\mu_{B_1}(t)$, $\sigma^2_{B_0}(t)$, and $\sigma^2_{B_1}(t)$ at time t.

Based on the provided parameters by the set of functions, the PDF of O-DC is available. Therefore, this set of functions constitutes the DS-DC model. ω_0 and ω_1 can be obtained from (10), while α_b can be obtained as

$$\alpha_b = \frac{\mu^2_{B_b} - \mu^3_{B_b} - \mu_{B_b}\sigma^2_{B_0}}{\sigma^2_{B_0}} \tag{12}$$

and β_b can be obtained as

$$\beta_b = \frac{\alpha_b(1-\mu_{B_b})}{\mu_{B_b}}. \tag{13}$$

In this DS-DC model based on the mixture-beta distribution, the number of parameters is given by $N(f_{\Psi_E,B_m}(x)) \times N(F_{\Psi_M}|K) = 5(1+3K)$. In the case of the beta distribution for DS-DC model, the number of parameters is given by $N(f_{\Psi_E,B}(x)) \times N(F_{\Psi_M}|K) = 2(1+3K)$.

C. Deterministic-stochastic DC model with regression analysis

The number of parameters in a model is related to the required storage capacity of the data server in SAS and therefore a smaller number of parameters is desirable. For this issue, the RDS-DC model is here proposed, which employs regression analysis to reduce the number of parameters. This idea comes from the fact that the variance of O-DC (Figs. 6a and 6b) and the mean of O-DC (Figs. 3a and 3b) have similar deterministic behaviors and this suggests that there may be some correlation between mean and variance. Specifically, a larger M-DC leads to a larger variance as confirmed in Fig. 3a and Fig. 6a. In the deterministic model, it may require five parameters for each statistical parameter, but a polynomial regression analysis can reduce the number of parameters. In RDS-DC, we use the deterministic M-DC $F_{\hat{\mu}_{B_m}}(t)$. Then, the other parameters, μ_{B_m}, μ_{B_0}, μ_{B_1}, $\sigma^2_{B_0}$, and $\sigma^2_{B_1}$, are obtained by N_Rth order polynomial regression analysis between μ_{B_m} and each parameter, p, as

$$p = \sum_{n=0}^{N_R} B_{n,p} \cdot \mu^n_{B_m}. \tag{14}$$

where $B_{n,p}$ is a coefficient for the regression analysis. The coefficients can be obtained by the least-squares method with the measurement results. The estimated parameter p and μ_{B_m} in dth super time frame are denoted by $\hat{\mu}_{B_m}(d)$ and $\hat{p}(d)$, respectively. In this paper, we consider cases of $N_R = 2$ and

$N_R = 1$. When $N_R = 2$, according to the least-squares method the regression analysis parameters can be given by

$$B_{1,p} = \frac{\sum_{d=1}^{D}(\hat{\mu}_{B_m}(d) - \bar{\hat{\mu}}_{B_m})(\hat{p}(d) - \bar{\hat{p}})}{\sum_{d=1}^{D}(\hat{p}(d) - \bar{\hat{p}})^2}, \quad (15)$$

where $\bar{\hat{\mu}}_{B_m}$ and $\bar{\hat{p}}$ are the averages of $\hat{\mu}_{B_m}(d)$ and $\hat{p}(d)$, respectively, and

$$B_{0,p} = \bar{\hat{p}} - B_{1,p}\bar{\hat{\mu}}_{B_m}. \quad (16)$$

When $N_R = 1$, by setting $B_{1,p} = 0$, $B_{0,p}$ can be obtained from (16).

The number of parameters for RDS-DC is $(N(f_{\Psi_E,B_m}(x)) - 1) \times (N_R + 1) + N(F_{\Psi_M}|K) = 4 \times (N_R + 1) + 1 + 3K$.

D. Model verification based on measurement results

1) O-DC distributions: To confirm the validity of the DS-DC models, several distributions and models of O-DC are shown in Fig. 8. In Fig. 8a, the empiric distribution of O-DC in each time (super time frame) is shown. This is a baseline result that the other models attempt to reproduce. In Fig. 8b, the stochastic model with mixture-beta distribution is used to express the distribution of O-DC at each super time frame. This distribution is denoted by Stochastic model based Distribution (SD). SD does not consider the behavior in time domain and corresponds to a conventional approach. In this case, the number of parameters is $N(f_{\Psi_E,B_m}(x)) \times H = 5H = 120$.

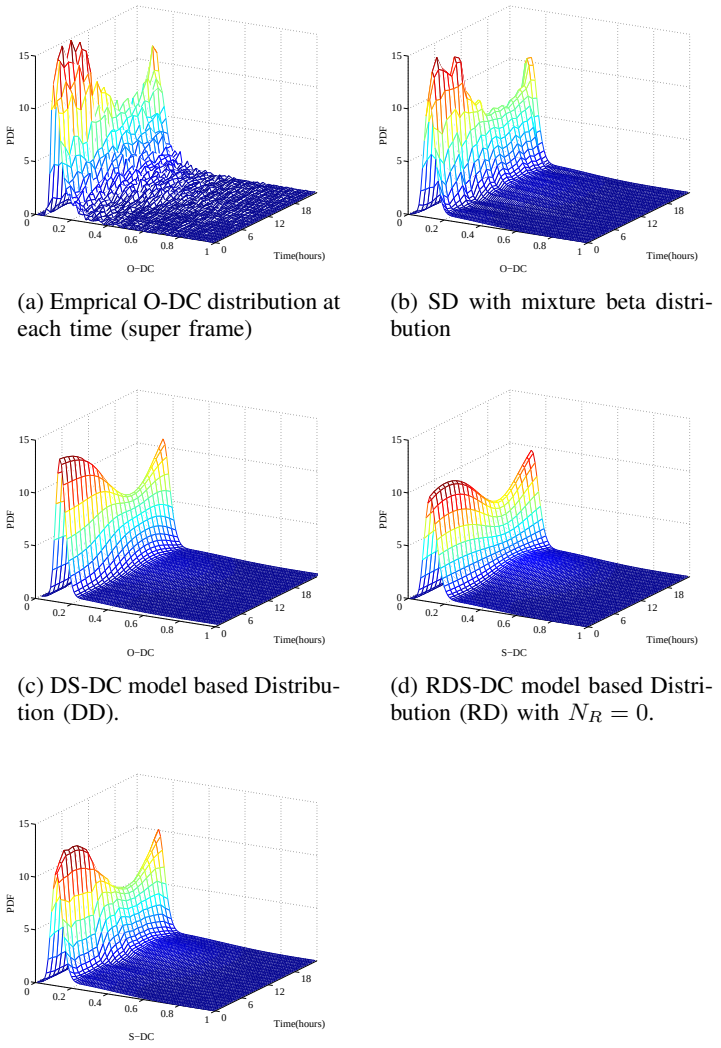
The distribution in Fig. 8c is obtained by DS-DC model in which the number of parameters is 35. This distribution is denoted by DS-DC model based Distribution (DD). In DD, the parameters in the stochastic model are modeled by the deterministic model, i.e., the deterministic model and stochastic model are combined. The results in Figs. 8d and 8e are obtained by RDS-DC models with $N_R = 0$ and $N_R = 1$ in (14), respectively. This distribution is denoted by RDS-DC model based Distribution (RD) with N_R and RD corresponds to our proposed approach. The number of parameters in RDS-DC models are 11 and 15 when $N_R = 0$ and 1, respectively. The numbers of parameters for each distribution are summarized in Table. I. We will evaluate accuracy of SD, DD and RD in the following section by DKL performance.

Three points can be confirmed by the results in Fig. 8. First, all of them have almost analogous shape, but there are a few differences as follows. Second, SD with mixture beta distribution reproduces the empirical distribution (Fig. 8a) adequately, especially the SD can reproduce non-smooth aspects in time domain since it does not have constraints in the time domain. Third, RD with $N_R = 1$ and DD result in similar distributions, however RD with $N_R = 0$ is slightly different. Specifically, the peak of the distribution in time interval from 0:00 to 6:00 in RD with $N_R = 0$ (Fig. 8d) is smaller than the others.

2) DKL performances: We also evaluate the accuracy of the models numerically by means of DKL between the empirical distribution and the model based distributions. DKL can indicate the difference between different probability distributions,

TABLE I: Number of parameters for each distribution.

Distribution	Number of parameters
SD: Stochastic model based Distribution	120
DD: DS-DC model based Distribution	35
RD: RDS-DC model based Distribution with $N_R = 0$	11
RD: RDS-DC model based Distribution with $N_R = 1$	15



(a) Emprical O-DC distribution at each time (super frame)

(b) SD with mixture beta distribution

(c) DS-DC model based Distribution (DD).

(d) RDS-DC model based Distribution (RD) with $N_R = 0$.

(e) RDS-DC model based Distribution (RD) with $N_R = 1$.

Fig. 8: Distributions of O-DC in W_1

with smaller DKL values indicating a better accuracy. DKL performance as a function of time for W_1 is shown in Fig. 9a, where SD with beta distribution is also evaluated.

First, we can confirm that SD with beta distribution achieves the poorest DKL performance. On the other hand, SD with mixture-beta distribution can achieve the best DKL per-

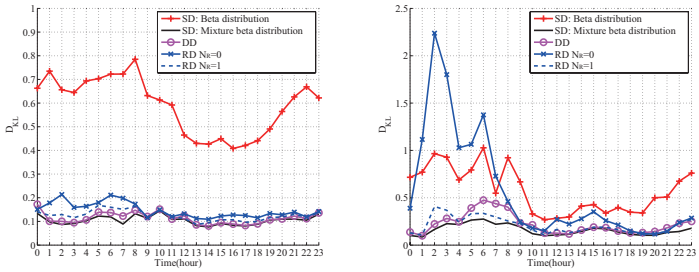
formance. This indicates the benefit of the mixture beta-distribution. Second, it can be confirmed that RD with $N_R = 1$ and DD can achieve a similar DKL performance compared to SD with mixture-beta distribution. RD with $N_R = 0$ based distribution is slightly worse than RD with $N_R = 1$ and DD. The trends observed in Fig. 9a can be understood based on the results shown in Fig. 8.

The DKL performance is also evaluated in W_2 to confirm the validity of the models in a different case and the result is shown in Fig. 9b. The stochastic model with beta based distribution again achieve a relatively worse DKL performance. In W_2, there is one aspect different from W_1, i.e., RD with $N_R = 0$ achieves the poorest DKL performance in the time interval from 01:00 to 09:00. As shown in Fig. 3b, distribution of O-DC between midnight and morning, such as 01:00 to 09:00, is significantly different from distribution of O-DC in day-time in W_2. In this case, when $N_R = 0$, there is only a constant term in (14) and it is not sufficient to express the deterministic behavior appropriately, which explains the observed result. The performance for the rest of models is similar as for W_1.

3) Final remarks: The results in Table I and the DKL performances in Figs. 9a and 9b indicate the existence of a trade-off between the accuracy of the model and the required number of parameters (which determines the data storage and complexity requirements). SD with mixture-beta distribution can achieve the best DKL performance, however it also requires the highest number of parameters. On the other hand, RD with $N_R = 1$ can achieve a slightly worse DKL performance compared to SD with mixture-beta distribution, e.g., average gap regarding DKL performance between RD with $N_R = 1$ and SD with mixture-beta distribution is 0.01, with a significantly lower number of parameters, which in fact provides the best trade-off between accuracy and storage/complexity requirements. A user of the spectrum usage models can select a proper model by considering such trade-off. A comparison in terms of the number of parameters between RD with with $N_R = 1$ and DD indicates that the reduction rate by RD with $N_R = 1$ is about 42%. Actual values of storage/complexity requirements are determined by not only the number of parameters, but also the measurement points, such as frequency bands and measurement locations. Therefore, the improvement regarding the number of parameters is appreciated when the number of measurement points is large.

VI. Conclusion

In this paper we have investigated models for spectrum usage in the time domain. The spectrum usage in the time domain is typically expressed in terms of the DC. In previous works, models for expressing the stochastic and deterministic behaviors of the DC have been investigated separately but not jointly. To overcome this drawback we have proposed several joint deterministic-stochastic models that can express both the deterministic and stochastic behaviors simultaneously. The improved accuracy of the proposed models has been corroborated with empiric data obtained from two long-term spectrum measurement campaigns performed in the WLAN and LTE uplink bands. Moreover, we have also shown that



(a) Measurement campaign W_1 (b) Measurement campaign W_2

Fig. 9: Kullback-Leibler divergences

by means of a regression analysis it is possible to reduce the number of parameters required by the proposed deterministic-stochastic model while preserving a similar level of accuracy. This improved modeling approach not only provides an equally remarkable level of accuracy but also reduces the data storage and complexity requirements in the data server of a SAS and therefore constitutes an excellent candidate for spectrum usage modeling in future spectrum-aware communication systems.

Acknowledgment

Kenta Umebayashi would like to thank the supports received from the MIC/SCOPE #165003006, and JSPS KAKENHI Grant Numbers JP15K06053, JP15KK0200. Miguel López-Benítez would like to thank the support received from the British Council under UKIERI DST Thematic Partnerships 2016-17 (ref. DST-198/2017)

References

[1] FCC, "Spectrum policy task force report," FCC, Tech. Rep. Docket No. 02-135, Nov. 2002.

[2] Q. Zhao, "A survey of dynamic spectrum access: signal processing, networking, and regulatory policy," *IEEE Signal Processing Mag.*, vol. 24, pp. 79–89, May 2007.

[3] I. F. Akyildiz, W. Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks: The International Journal of Computer and Telecommunications Networking*, vol. 50, pp. 2127–2159, Sep. 2006.

[4] M. Palolo, T. Rautio, M. Matinmikko, J. Prokkola, M. Mustonen, M. Heikkila, T. Kippola, S. Yrjola, V. Hartikainen, L. Tudose, A. Kivinen, J. Paavola, J. Okkonen, M. Makelainen, T. Hanninen, and H. Kokkinen, "Licensed shared access (LSA) trial demonstration using real LTE network," in *2014 9th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Oulu, Finland, Jun. 2014, pp. 498–502.

[5] M. M. Sohul, M. Yao, T. Yang, and J. H. Reed, "Spectrum access system for the citizen broadband radio service," *IEEE Communications Magazine*, vol. 53, no. 7, pp. 18–25, Jul. 2015.

[6] S. Bhattarai, J. M. J. Park, B. Gao, K. Bian, and W. Lehr, "An overview of dynamic spectrum sharing: Ongoing initiatives, challenges, and a roadmap for future research," *IEEE Transactions on Cognitive Communications and Networking*, vol. 2, no. 2, pp. 110–128, Jun. 2016.

[7] R. H. Tehrani, S. Vahid, D. Triantafyllopoulou, H. Lee, and K. Moessner, "Licensed spectrum sharing schemes for mobile operators: A survey and outlook," *IEEE Communications Surveys Tutorials*, vol. 18, no. 4, pp. 2591–2623, Fourthquarter 2016.

[8] N. Wang, Y. Gao, and X. Zhang, "Adaptive spectrum sensing algorithm under different primary user utilization," *IEEE Commun. Lett.*, vol. 17, no. 9, pp. 1838–1841, Sep. 2013.

[9] T. Nguyen, B. L. Mark, and Y. Ephraim, "Spectrum sensing using a hidden bivariate Markov model," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4582–4591, Sep. 2013.

[10] C. Cormio and K. R. Chowdhury, "A survey on MAC protocols for cognitive radio networks," *Ad Hoc Netw.*, vol. 7, no. 7, pp. 1315–1329, Sep. 2009.

[11] K. Umebayashi, K. Kasahara, Y. Kamiya, and Y. Suzuki, "A novel spectrum sharing technique based on channel occupancy rate information," in *Proc. IEEE global communications conference (GLOBECOM)*, Honolulu, HI, USA, Dec. 2009, pp. 1–6.

[12] M. Hoyhtya, S. Pollin, and A. Mammela, "Classification-based predictive channel selection for cognitive radios," in *2010 IEEE International Conference on Communications*, Cape Town, South Africa, May 2010, pp. 1–6.

[13] Y. Chen and H. S. Oh, "A survey of measurement-based spectrum occupancy modeling for cognitive radios," *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 848–859, Firstquarter 2016.

[14] K. Umebayashi, S. Tiiro, and J. J. Lehtomaki, "Development of a measurement system for spectrum awareness," in *Proc. of 1st International Conference on 5G for Ubiquitous Connectivity*, Akaslompolo, Finland, Nov. 2014.

[15] M. López-Benítez and F. Casadevall, "Spectrum usage in cognitive radio networks: From field measurements to empirical models," *IEICE Trans. Commun.,*, vol. E97-B, no. 2, pp. 242–250, Feb. 2014.

[16] M. Hoyhtya, A. Mammela, M. Eskola, M. Matinmikko, J. Kalliovaara, J. Ojaniemi, J. Suutala, R. Ekman, R. Bacchus, and D. Roberson, "Spectrum occupancy measurements: A survey and use of interference maps," *IEEE Communications Surveys Tutorials*, vol. 18, no. 4, pp. 2386–2414, Fourthquarter 2016.

[17] M. A. McHenry, P. A. Tenhula, D. McCloskey, D. A. Roberson, and C. S. Hood, "Chicago spectrum occupancy measurements & analysis and a long-term studies proposal," in *Proceedings of the First International Workshop on Technology and Policy for Accessing Spectrum*, ser. TAPAS '06. Boston, MA, USA: ACM, 2006.

[18] M. Lopez-Benitez, A. Umbert, and F. Casadevall, "Evaluation of spectrum occupancy in spain for cognitive radio applications," in *VTC Spring 2009 - IEEE 69th Vehicular Technology Conference*, Barcelona, Spain, Apr. 2009, pp. 1–5.

[19] V. Valenta, R. Marsalek, G. Baudoin, M. Villegas, M. Suarez, and F. Robert, "Survey on spectrum utilization in Europe: Measurements, analyses and observations," in *2010 Proceedings of the Fifth International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, Cannes, France, Jun. 2010, pp. 1–5.

[20] R. Schiphorst and C. H. Slump, "Evaluation of spectrum occupancy in Amsterdam using mobile monitoring vehicles," in *2010 IEEE 71st Vehicular Technology Conference*, Taipei, Taiwan, May 2010, pp. 1–5.

[21] N. Q. B. Vo, Q. C. Le, Q. P. Le, D. T. Tran, T. Q. Nguyen, and M. T. Lam, "Vietnam spectrum occupancy measurements and analysis for cognitive radio applications," in *The 2011 International Conference on Advanced Technologies for Communications (ATC 2011)*, Da Nang, Vietnam, Aug. 2011, pp. 135–143.

[22] M. Lopez-Benitez and F. Casadevall, "Empirical time-dimension model of spectrum use based on a discrete-time Markov chain with deterministic and stochastic duty cycle models," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 6, pp. 2519–2533, Jul. 2011.

[23] M. Lopez-Benitez, F. Casadevall, D. Lopez-Perez, and A. Vasilakos, "Modeling and simulation of joint time-frequency properties of spectrum usage in cognitive radio," in *Proc. 4th Int'l. Conf. Cognitive Radio and Adv. Spect. Management (CogART 2011)*, Barcelona, Spain, Oct. 2011, pp. 1–5.

[24] S. Yin, Q. Zhang, E. Zhang, L. Yin, and S. Li, "Statistical modeling for spectrum usage characterizing wireless fading channels and mobile service dynamics," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 8, pp. 3800–3812, Oct. 2013.

[25] M. Lopez-Benitez and F. Casadevall, "Space-dimension models of spectrum usage for cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 306–320, Jan. 2017.

[26] A. Gupta, S. Agarwal, and S. De, "A new spectrum occupancy model for 802.11 WLAN traffic," *IEEE Communications Letters*, vol. 20, no. 12, pp. 2550–2553, Dec. 2016.

[27] S. Sengottuvelan, J. Ansari, P. Mahonen, T. G. Venkatesh, and M. Petrova, "Channel selection algorithm for cognitive radio networks with heavy-tailed idle times," *IEEE Transactions on Mobile Computing*, vol. 16, no. 5, pp. 1258–1271, May 2017.

[28] K. Umebayashi, K. Hayashi, and J. J. Lehtomaki, "Threshold-setting for spectrum sensing based on statistical information," *IEEE Communications Letters*, vol. 21, no. 7, pp. 1585–1588, Jul. 2017.

[29] S. Geirhofer, L. Tong, and B. M. Sadler, "Cognitive radios for dynamic spectrum access - dynamic spectrum access in the time domain: Modeling and exploiting white space," *IEEE Communications Magazine*, vol. 45, no. 5, pp. 66–72, May 2007.

[30] K. Umebayashi, K. Moriwaki, R. Mizuchi, H. Iwata, S. Tiiro, J. Lehtomäki, M. Lopez-Benitez, and Y. Suzuki, "Simple primary user signal area estimation for spectrum measurement," *IEICE Trans. Commun.*, vol. E99-B, no. 2, pp. 523–532, Feb. 2016.

[31] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. 15, pp. 70–73, Jun. 1967.

[32] J. J. Lehtomaki, R. Vuohtoniemi, and K. Umebayashi, "On the measurement of duty cycle and channel occupancy rate," *IEEE J. Select. Areas Commun.*, vol. 31, no. 1, pp. 2555 – 2565, Nov. 2013.

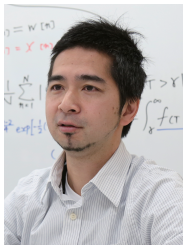
[33] K. Umebayashi, R. Takagi, N. Ioroi, Y. Suzuki, and J. J. Lehtomaki, "Duty cycle and noise floor estimation with Welch fft for spectrum usage measurements," in *Proc. of Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Oulu, Finland, Jun. 2014, pp. 73–78.

[34] M. Wellens and P. Mahonen, "Lessons learned from an extensive spectrum occupancy measurement campaign and a stochastic duty cycle model," in *2009 5th International Conference on Testbeds and Research Infrastructures for the Development of Networks Communities and Workshops*, Washington, DC, USA, Apr. 2009, pp. 1–9.

[35] C. Ghosh, S. Roy, M. Rao, and D. Agrawal, "Spectrum occupancy validation and modeling using real-time measurements," in *Proc. CoRoNet' 10*, New York, NY, USA, Sep. 2010, pp. 25–30.

[36] K. Patil, "Validation of Beta distribution for spectrum usage using Kolmogorov-Smirnov test," *International Journal of Computer Applications*, vol. 144, no. 9, pp. 23–26, Jun. 2016.

[37] S. Chib and E. Greenberg, "Understanding the Metropolis-Hastings algorithm," *The American Statistician*, vol. 49, no. 4, pp. 327–335, Nov. 1995.



Kenta Umebayashi (S ' 00, M ' 04) received the LL.B. degree from Ritsumeikan University in 1996 and the B.E., M.E., and Ph.D. degrees from the Yokohama National University in 1999, 2001, and 2004, respectively. From 2004 to 2006, he was a research scientist at the University of Oulu, Centre for Wireless Communications (CWC). He is currently an associate professor at the Tokyo University of Agriculture and Technology. He was an associate editor of IEICE Transaction on Communications from May 2015 to May 2017. He was a principal investigator in the three Grants-in-Aid for Scientific Research projects and two Strategic Information and Communications R&D Promotion Programme projects. His research interests lie in the areas of signal detection and estimation theories for wireless communication, signal processing for multiple antenna based wireless communication, cognitive radio networks, and terahertz band wireless communication. He received the Best Paper Award at IEEE WCNC 2012 for a paper he authored, and the Best Paper Award at IEEE WCNC workshop IWSS 2015 for a paper he co-authored.

Masanao Kobayashi received B.E. and M.E. degrees from Tokyo University of Agriculture and Technology, Tokyo, Japan, in 2015 and 2017. His research interests are cognitive radio and wireless communication systems.

Miguel López-Benítez (S ' 08, M ' 12, SM ' 17) has been a Lecturer (Assistant Professor) in the Department of Electrical Engineering and Electronics of the University of Liverpool, UK since 2013. Prior to this, he was a Research Fellow in the Centre for Communication Systems Research of the University of Surrey, Guildford, UK from 2011 to 2013. He received the B.Sc. and M.Sc. degrees in Telecommunications Engineering (both with Distinction/First-Class Honours) from Miguel Hernández University, Elche, Spain in 2003 and 2006, respectively, and a Ph.D. degree in Telecommunications Engineering (2011 Outstanding Ph.D. Thesis Award) from the Technical University of Catalonia, Barcelona, Spain in 2011.

His research interests include the field of wireless communications and networking, with special emphasis on cellular mobile communications and dynamic spectrum access in cognitive radio systems. He is/has been the principal investigator or co-investigator of research projects funded by EPSRC, the British Council and the Royal Society, and has been involved in the European-funded projects AROMA, NEWCOM++, FARAMIR, QoSMOS and CoRaSat. He is Associate Editor of IEEE Access and IET Communications, and has been a member of the Organising Committee for the IEEE WCNC International Workshop on Smart Spectrum (IWSS 2015, 2016 and 2017). Please visit http://www.lopezbenitez.es for more details.