# Analytical Study on the Estimation of Primary Activity Distribution Based on Spectrum Sensing

Ahmed Al-Tahmeesschi*, Miguel López-Benítez*, Kenta Umebayashi† and Janne Lehtomäki‡

*Dept. of Electrical Engineering and Electronics, University of Liverpool, United Kingdom
†Graduate School of Engineering, Tokyo University of Agriculture and Technology, Japan
‡Centre for Wireless Communications, University of Oulu, Finland

*Abstract*—**Cognitive Radio (CR) systems utilize spectrum sensing to decide transmission time in an opportunistic manner. Spectrum sensing can also be used not only to determine the instantaneous on/off state of the channel but also to monitor the statistics of primary user to gain information on occupancy pattern. This knowledge can be exploited in many ways to improve CR systems. In this paper, we propose an analytical model to link the sensing period with the observed spectrum occupancy. Moreover, the effect of spectrum sensing periods on the estimated primary activity pattern is analysed. Simulation results show that the proposed model captures with reasonable accuracy the spectrum occupancy observed at the CR.**

*Keywords*— *Cognitive radio, spectrum sensing, exponential distribution, primary activity statistics*

## I. INTRODUCTION

Dynamic spectrum access (DSA)/Cognitive Radio (CR) networks aim to solve the spectrum scarcity problem by increasing radio spectrum utilization. This can be achieved by allowing secondary (unlicensed) users (SUs) of DSA/CR to access the spectrum of primary (licensed) users (PUs) in an opportunistic and non-interfering manner during PU idle times (spectrum holes) [1, 2]. In essence, a CR is a smart device that is capable of changing its transmission parameters according to the surrounding environment. A detailed overview of CR systems can be found in [3]. Owing to the opportunistic nature of operation for DSA/CR, the SU behaviour is affected by the PU transmission pattern. It is important from theoretical [4] and practical points of view to have accurate knowledge of the PU activity (busy and idle) periods distribution. This knowledge can be exploited to access the spectrum more effectively by selecting the most appropriate channel for transmission [5] and enhancing the forecasting of PU occupancy patterns to minimize the interference [6].

DSA/CR users utilize spectrum sensing decisions to obtain information on PU channel activity. The PU channel is sensed periodically by DSA/CR users to decide the channel state (busy or idle) at every sensing event based on a signal detection algorithm [7]. These spectrum decisions can be used to estimate the durations of the idle and busy periods. Unfortunately, the estimation of PU activity periods and statistics by means of spectrum sensing (periodic channel observations) suffers from practical limitations. These limitations reduce the statistical estimation accuracy of PU parameters at DSA/CR users. The interest and focus of this work is on analysing the impact of spectrum sensing period on the accuracy of the estimated PU activity statistics (in particular, in the distribution of PU busy/idle periods). Despite being an elemental problem of crucial importance for CR systems, this has never been considered or analysed before in the existing literature. The impact of sensing errors (i.e., false alarms and missed detections) is out of the scope of this work and therefore a high signal-to-noise ratio (SNR) scenario with no sensing errors is here considered. The mathematical analysis of the low SNR scenario with sensing errors requires a significantly more complex study and will be addressed in future work.

The contribution of this work is threefold:

1) Analytical expressions are derived for both the probability density function (pdf) and the cumulative distribution function (cdf) observed at the SU taking into account the effect of the spectrum sensing period.
2) Analytical expressions are derived for the maximum error with different distributions taking into account the effect of the spectrum sensing period.
3) The effect of the spectrum sensing period on the distribution observed at the SU is studied.

The remainder of this paper is organised as follows. First, Section II describes the system model and provides a formal description of the problem of estimating the PU activity statistics based on spectrum sensing. Section III provides closed-form expressions for the pdf and cdf of the periods observed at the SU as a function of the original distribution at the PU and the sensing period employed by the SU. Section IV provides a closed form expression for the maximum observed error as a function of sensing period and distribution parameters. Section V validates the proposed pdf/cdf models with simulations and analyses thoroughly the effects of the sensing period and distribution parameters on the distribution estimation. Finally, Section VI concludes the paper.
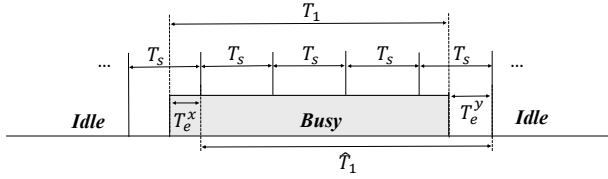
Fig. 1. Considered model. $T_s, T_1, \widehat{T}_1$ represent the sensing duration, original busy period duration and estimated busy period duration, respectively. $T_e^x$ and $T_e^y$ are the errors in period estimation.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

In this paper, an SU is considered to detect a single PU activity signal using spectrum sensing. The results of sensing decisions are introduced as a binary alternating state: busy when the PU signal is present and idle when the PU signal is absent. Based on the sequence of busy/idle spectrum sensing decisions, it is possible to estimate the durations of the idle ($T_0$) and busy ($T_1$) periods as shown in Fig. 1. The same model was considered in [8].

The average busy period is $\mathbb{E}\{T_1\}$ and the average idle time is $\mathbb{E}\{T_0\}$. The duty cycle ($\Psi$) can be estimated as:

$$\Psi = \frac{\mathbb{E}\{T_1\}}{\mathbb{E}\{T_0\} + \mathbb{E}\{T_1\}} \qquad (1)$$

As discussed in Section I, we assume a high SNR scenario with no sensing errors so that the only degrading effect considered in this study is the impact of the finite sensing period $T_s$, which is the aspect of interest in this work. The PU activity periods $T_i$ ($i \in \{0,1\}$, $i = 0$ for idle periods and $i = 1$ for busy periods) can be sensed accurately in case the channel is sensed exactly at the points of PU state change. In practice the SU is de-synchronised with the PU channel activity and the PU channel is sensed at arbitrary time instants every $T_s$ time units (t.u.). As a result, the estimated periods $\widehat{T}_i$ depend not only on the original periods $T_i$ but also on the employed sensing period $T_s$. The main objective of this work is to explore the relation between the original periods $T_i$ and the estimated periods $\widehat{T}_i$ as a function of the sensing period $T_s$. To this end, closed-form expressions are developed for the pdf/cdf of $\widehat{T}_i$ as a function of the pdf/cdf of $T_i$ and $T_s$.

## III. DISTRIBUTION OF THE ESTIMATED PERIODS

The estimated periods $\widehat{T}_i$ can be expressed as a function of the original periods $T_i$ as $\widehat{T}_i = T_i + T_e$, where $T_e$ is the error component, which according to the model of Fig. 1 is given by $T_e = T_e^y - T_e^x$. As it can be appreciated from Fig. 1, both $T_e^x$ and $T_e^y$ can take any value between 0 and $T_s$. A ressonable and intuitive assumption is that both of them follow a uniform distribution (i.e., $T_e^x$ and $T_e^y \sim U(0, T_s)$). This assumption can be verified from Fig. 2, which was obtained by simulating the sensing of a sufficiently high number of exponentially distributed periods $T_i$ using a sensing period $T_s = 5$ t.u., recording the error components $T_e^x$ and $T_e^y$, and computing their
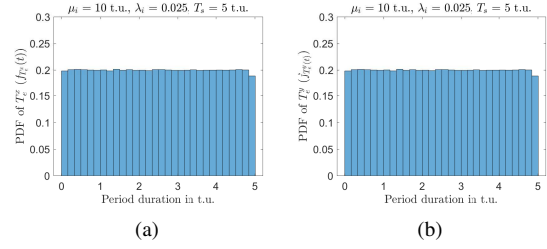


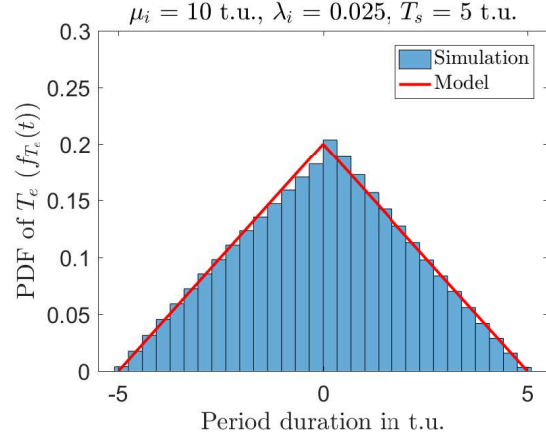Fig. 2. The pdf of the error components: (a) $f_{T_e^x}(t)$, (b) $f_{T_e^y}(t)$.



Fig. 3. The pdf of the combined error component $f_{T_e}(t)$.

normalized histograms (i.e., pdfs). As it can be observed, the assumption of uniform distribution for the $T_e^x$ and $T_e^y$ error components is correct.

Assuming $T_e^x$ and $T_e^y$ are independent, the pdf of the combined error $T_e$ (which is the sum of $T_e^y$ and $-T_e^x$) can be obtained as the convolution of the distributions of $T_e^y \sim U(0, T_s)$ and $-T_e^x \sim U(-T_s, 0)$ [9], which leads to a symmetric triangular distribution with width $2T_s$, $T_e \sim \Delta(-T_s, T_s)$. The pdf for the triangular distribution of $T_e$ is:

$$f_{T_e}(t) = \begin{cases} 0 & t < -T_s \\ \frac{T_s + t}{T_s^2} & -T_s \leq t \leq 0 \\ \frac{T_s - t}{T_s^2} & 0 \leq t \leq T_s \\ 0 & t > T_s \end{cases}, \qquad (2)$$

This model can be verified from simulation results as shown in Fig. 3.

The PU state holding times ($T_0$ and $T_1$) are random variables assumed to be independent and exponentially distributed [10]. The exponential distribution is the most common model used to describe the periods of the on/off states in the literature [11, 12] even though it is proven not to be the most accurate since other distributions provide better fit for real scenarios such as the generalized Pareto, Gamma or even more complicated distributions [13]. We use the exponential distribution because it is a special case of the generalized Pareto distribution with a simpler mathematical form. The pdf and cdf for the exponential distribution are given as [14]:

$$f_{\widehat{T}_i}(t) = \begin{cases} 0 & t < \mu_i - T_s \\ \dfrac{T_s + t - \mu_i}{T_s^2} - \dfrac{1}{\lambda_i T_s^2}\left[1 - \dfrac{1}{\lambda_i}f_{T_i}(t + T_s)\right] & \mu_i - T_s \le t < \mu_i \\ \dfrac{T_s - t + \mu_i}{T_s^2} + \dfrac{1}{\lambda_i T_s^2}\left[1 + \dfrac{1}{\lambda_i}f_{T_i}(t + T_s) - \dfrac{2}{\lambda_i}f_{T_i}(t)\right] & \mu_i \le t \le \mu_i + T_s \\ \dfrac{1}{(\lambda_i T_s)^2}\left[f_{T_i}(t + T_s) - 2f_{T_i}(t) + f_{T_i}(t - T_s)\right] & t > \mu_i + T_s \end{cases} \tag{6}$$

$$F_{\widehat{T}_i}(t) = \begin{cases} 0 & t < \mu_i - T_s \\ \dfrac{t^2 - (\mu_i - T_s)^2}{2T_s^2} - \dfrac{\left[1 - \lambda_i(T_s - \mu_i)\right](t + T_s - \mu_i)}{\lambda_i T_s^2} + \dfrac{1}{(\lambda_i T_s)^2}F_{T_i}(t + T_s) & \mu_i - T_s \le t < \mu_i \\ \dfrac{\mu_i^2 - t^2}{2T_s^2} - \dfrac{2 - \lambda_i T_s}{2\lambda_i T_s} + \dfrac{\left[1 + \lambda_i(T_s + \mu_i)\right](t - \mu_i)}{\lambda_i T_s^2} + \dfrac{1}{(\lambda_i T_s)^2}\left[F_{T_i}(t + T_s) - 2F_{T_i}(t)\right] & \mu_i \le t \le \mu_i + T_s \\ 1 + \dfrac{1}{(\lambda_i T_s)^2}\left[F_{T_i}(t + T_s) - 2F_{T_i}(t) + F_{T_i}(t - T_s)\right] & t > \mu_i + T_s \end{cases} \tag{8}$$

---

$$f_{T_i}(t) = \begin{cases} 0 & t < \mu_i \\ \lambda_i e^{-\lambda_i(t - \mu_i)} & t \ge \mu_i \end{cases} \tag{3}$$

$$F_{T_i}(t) = \begin{cases} 0 & t < \mu_i \\ 1 - e^{-\lambda_i(t - \mu_i)} & t \ge \mu_i \end{cases} \tag{4}$$

where $\lambda_i$ is the distribution scale parameter and $\mu_i$ is the distribution location parameter (also the smallest value for the PU activity period).

Since $\widehat{T}_i = T_i + T_e$, the pdf of the estimated periods can be obtained as [9]:

$$f_{\widehat{T}_i}(t) = f_{T_i}(t) * f_{T_e}(t) = \int_{-\infty}^{\infty} f_{T_i}(\tau) \cdot f_{T_e}(t - \tau)d\tau \tag{5}$$

where $f_{T_i}(t)$ and $f_{T_e}(t)$ are given by (3) and (2) respectively. The operator $*$ refers to the convolution operation. The resulting expression for the pdf $f_{\widehat{T}_i}(t)$ is shown in (6) while the cdf $F_{\widehat{T}_i}(t)$ can be obtained through the direct integration of $f_{\widehat{T}_i}(t)$ as shown below:

$$F_{\widehat{T}_i}(t) = \int_{-\infty}^{t} f_{\widehat{T}_i}(\tau)d\tau \tag{7}$$

The final cdf expression can be seen in (8).

Note that the distributions in (6) and (8) have a continuous domain, while the actual distributions of the periods observed at a SU are discrete since the periods estimated from spectrum sensing as shown in Fig. 1 are integer multiples of the employed sensing period (i.e., $\widehat{T}_i = kT_s, k = 1, 2, 3 \ldots$). Such discrete distribution can be obtained by evaluating (6) and (8) at the right points of each interval/bin of the pdf and cdf, respectively, as:

$$g_{\widehat{T}_i}(k) = f_{\widehat{T}_i}(kT_s) \tag{9}$$

$$G_{\widehat{T}_i}(k) = F_{\widehat{T}_i}((k + 1/2)T_s) \tag{10}$$

The set of obtained expressions provide closed-form relations between the distributions of the original periods $T_i$ resulting from the PU transmission (and its parameters $\mu_i, \lambda_i$), the distribution of the estimated periods $\widehat{T}_i$ as observed by the SU based on spectrum sensing decisions, and the employed sensing period $T_s$. These mathematical results are useful to evaluate the impact of the employed sensing period on the accuracy of the distributions estimated by the SU and can find many practical applications such as mathematical analysis, simulation or system design (e.g., determine the maximum value of $T_s$ required for a given level of estimation accuracy).

## IV. ERROR OF THE ESTIMATED DISTRIBUTION

To better understand the sensing period effect on the observed distribution cdf, we utilize the well-known Kolmogorov-Smirnov (KS) distance. This is the most commonly used metric to quantify the error between two distributions. The KS distance is defined as the largest absolute error between two continuous cdfs and given as follows [15]:

$$D_{KS} = \sup_t \left| F_{T_i}(t) - F_{\widehat{T}_i}(t) \right| \tag{11}$$

To find the value of $t$ that returns the maximum distance ($D_{KS}$), the partial derivative of the absolute difference in the KS distance is taken and equated to zero as follows:

$$\frac{\partial[F_{T_i}(t) - F_{\widehat{T}_i}(t)]}{\partial t} = 0 \tag{12}$$

The largest difference occurs at $t = \mu_i$. Since $F_{T_i}(\mu_i)$ is zero at $t = \mu_i$, the final expression of KS distance will be:

$$\begin{aligned} D_{KS} &= F_{\widehat{T}_i}(\mu_i) \\ &= \frac{1}{2} - \frac{1}{\lambda_i T_s} + \frac{1 - e^{-\lambda_i T_s}}{(\lambda_i T_s)^2} \end{aligned} \tag{13}$$
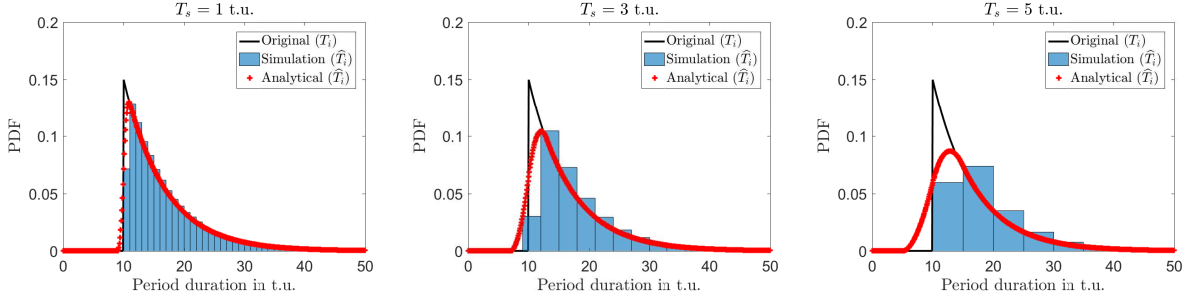
Fig. 4. Validation of the pdf of the estimated periods ($\lambda_1 = 0.15, \mu_1 = 10\ t.u, \mathbb{E}\{T_1\} = 16.66\ t.u.$ and $\Psi = 0.5$.)
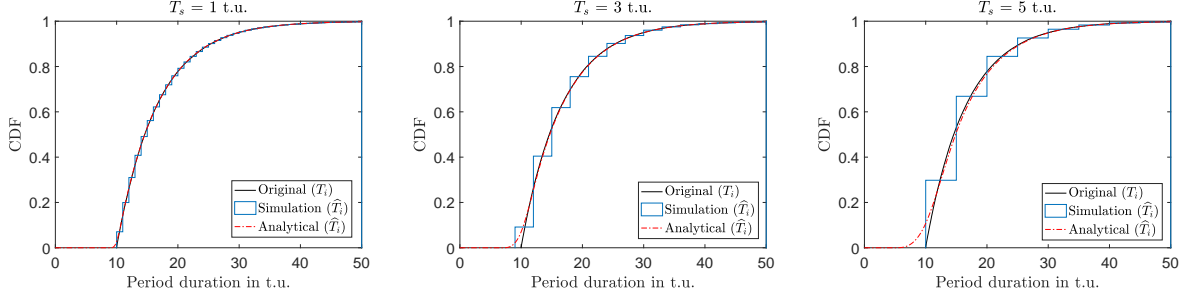


Fig. 5. Validation of the cdf of the estimated periods ($\lambda_1 = 0.15, \mu_1 = 10\ t.u, \mathbb{E}\{T_1\} = 16.66\ t.u.$ and $\Psi = 0.5$.)

Expression (13) provides an easy and accurate tool to mathematically calculate the KS distance between the estimated and original cdfs as a function of the employed sensing period. Moreover, expression (13) can be used to calculate the $T_s$ required for a given target estimation error.

## V. NUMERICAL RESULTS

In this section, first we will assess the accuracy of the proposed model for both pdf and cdf, then the effect of the sensing period on the distribution estimation. For all the considered cases the sensing period is lower than the minimum PU activity time ($T_s < \mu_i$). This is required to ensure that no activity periods are missed in the sensing process (the shortest detectable period is $T_s$), which would otherwise lead to significant estimation errors. Notice that this consideration implicitly assumes that the minimum PU activity time $\mu_i$ is known to the SU so that the value of $T_s$ can be configured not to exceed $\mu_i$. This assumption is realistic since the value of $\mu_i$ is available for some well-known standardised radio technologies (e.g., the time-slot duration of GSM or LTE) or can be obtained with other methods such as blind recognition/estimation [16] or from PU beacon signals [17].

Fig. 4 shows the busy periods pdfs $f_{\widehat{T}_i}(t)$ obtained from simulation and analytical expression versus the original distribution $f_{T_i}(t)$ for multiple values of sensing periods ($T_s = 1$, 3 and 5 t.u.). The discrete expression $g_{\widehat{T}_i}(k)$ has not been included for clarity but its corresponding values can be easily obtained as the values of the analytical expression $f_{\widehat{T}_i}(t)$ at $kT_s$. It can be

appreciated that the closed form analysis provides an excellent fit with the simulation results for all the considered scenarios, which verifies the validity of of the mathematical expression obtained for the pdf. Moreover, Fig. 4 shows the effect of sensing period $T_s$ on the discrete estimated pdf $g_{\widehat{T}_i}(k)$. High sensing periods give higher estimation errors and vice versa.

Fig. 5 shows the busy periods cdfs obtained from simulation $G_{\widehat{T}_i}(t)$ (discrete) and analytical expression $F_{\widehat{T}_i}(t)$ versus the original distribution $F_{T_i}(t)$ for multiple values of sensing periods ($T_s = 1$, 3 and 5 t.u.). The closed form analysis provides an excellent fit with the simulation results for all the considered scenarios, which verifies the validity of the mathematical expression obtained for the cdf. The stair shape of the observed cdf $G_{\widehat{T}_i}(t)$ represents the effect of the spectrum sensing operation and the resulting discrete observed periods.

Fig. 6 shows the KS distance for the simulated and analytical cdf with respect to the original distribution. The x-axis represents the duration of sensing period in time units and the y-axis represents the KS distance. Since the sensed cdf is a discrete distribution $G_{\widehat{T}_i}(t)$, it is to be transformed to a continuous form for comparison purposes. To this end, we utilize the cdf frequency polygons [18], where mid points of the discrete cdf are joined together and extended to include the zero frequency cases from left of the normalised histogram and hence obtain the continuous form of the cdf. As it can be appreciated from Fig. 6, the analytical expression (13) gives an excellent prediction of the estimation error. High $T_s$ values will result in larger errors in the estimation of the PU activity pattern, however the resulting estimation
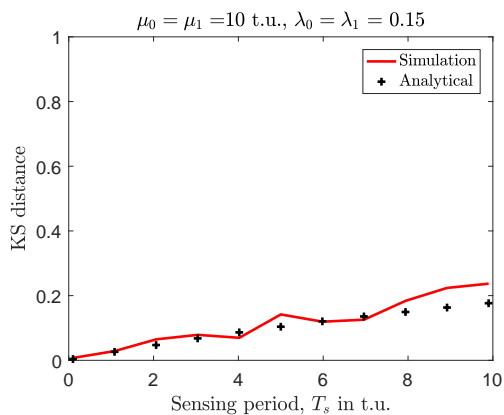
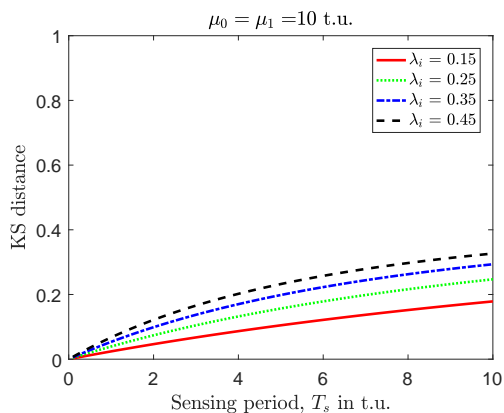Fig. 6. KS distance for the observed and analytical model cdfs.



Fig. 7. KS distance for the observed and analytical model cdfs.

error can be reduced by decreasing $T_s$.

Fig. 7 analyses the impact of different $\lambda_i$ values ($\lambda_i = 0.15, 0.25, 0.35$ and $0.45$) on the KS distance based on (13). Fig. 7 implies that not only the value of $T_s$ has an impact on the estimated error (KS distance) but also the value of $\lambda_i$ (distribution scale). The KS distance increases with higher values of $\lambda_i$. The analytical result in (13) can be used as shown in Fig. 7 to determine the maximum value of $T_s$ required for a given level of estimation accuracy of the distribution.

## VI. CONCLUSION

CR utilizes spectrum sensing to periodically monitor PU channel activity states (idle/busy). A CR bene-fits from this knowledge to improve the general sys-tem/device performance. However spectrum sensing uses a finite sensing period which imposes limitations on the measured durations of busy/idle periods and hence the resulting distribution for PU activity. This work focuses on the analytical perspective of how this limitation affects the estimation of PU distribution. Closed form expressions are derived to show the relationship between the employed sensing period and the resulting estimated distribution under finite sensing periods, as well as the corresponding estimation error in terms of the KS distance. The analytical results showed a good agreement with simulation results and can be used in the design and analysis of CR systems.

## REFERENCES

[1] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Comms.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[2] M. López-Benítez and F. Casadevall, *Spectrum Usage Models for the Analysis, Design and Simulation of Cognitive Radio Networks*. Dordrecht: Springer Netherlands, 2012, pp. 27–73.

[3] Y. C. Liang, K. C. Chen, G. Y. Li, and P. Mahonen, "Cognitive radio networking and communications: an overview," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 3386–3407, Sept 2011.

[4] W. Gabran, C. H. Liu, P. Pawelczak, and D. Cabric, "Primary user traffic estimation for dynamic spectrum access," *IEEE J. Sel. Areas Comms.*, vol. 31, no. 3, pp. 544–558, March 2013.

[5] X. Liu, B. Krishnamachari, and H. Liu, "Channel selection in multi-channel opportunistic spectrum access networks with perfect sensing," in *Proc. 2010 IEEE Int'l. Symp. Dyn. Spect. Access Networks (DySPAN 2010)*, Apr. 2010, pp. 1–8.

[6] C. H. Liu, W. Gabran, and D. Cabric, "Prediction of exponentially distributed primary user traffic for dynamic spectrum access," in *2012 IEEE Global Communications Conference (GLOBECOM)*, Dec 2012, pp. 1441–1446.

[7] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys Tutorials*, vol. 11, no. 1, pp. 116–130, First 2009.

[8] M. López-Benítez, "Can primary activity statistics in cognitive radio be estimated under imperfect spectrum sensing?" in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept 2013, pp. 750–755.

[9] R. J. Larsen and M. L. Marx, *An introduction to mathematical statistics and its applications; 4th ed.* Boston, MA: Prentice Hall, 2006.

[10] H. Kim and K. G. Shin, "Efficient discovery of spectrum opportunities with mac-layer sensing in cognitive radio networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 5, pp. 533–545, May 2008.

[11] W. Gabran, C. H. Liu, P. Pawelczak, and D. Cabric, "Primary user traffic estimation for dynamic spectrum access," *IEEE J. Sel. Areas Comms.*, vol. 31, no. 3, pp. 544–558, March 2013.

[12] 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.

[13] M. López-Benítez and F. Casadevall, "Time-dimension models of spectrum usage for the analysis, design, and simulation of cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 5, pp. 2091–2104, Jun 2013.

[14] N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous univariate distributions*, ser. Wiley series in probability and mathematical statistics: Applied probability and statistics. Wiley & Sons, 1995, no. v. 2.

[15] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes: The art of scientific computing*, 3rd ed. Cambridge University Press, 2007.

[16] M. López-Benítez, F. Casadevall, A. Umbert, J. Perez-Romero, R. Hachemani, J. Palicot, and C. Moy, "Spectral occupation measurements and blind standard recognition sensor for cognitive radio networks," in *2009 4th International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, June 2009, pp. 1–9.

[17] M. López-Benítez, "Cognitive radio," in *Heterogeneous cellular networks: Theory, simulation and deployment*. Cambridge University Press, 2013, ch. 13.

[18] K. A. Stroud and D. J. Booth, *Engineering mathematics*. Basingstoke : Palgrave Macmillan, 2013., 2013.