Modelling Primary System Activity in Dynamic Spectrum Access Networks by Aggregated ON/OFF-Processes

Matthias Wellens, Janne Riihijärvi, and Petri Mähönen

Department of Wireless Networks, RWTH Aachen University
Kackertstrasse 9, D-52072 Aachen, Germany
Email: {mwe, jar, pma}@mobnets.rwth-aachen.de

***Abstract*—Accurate modelling of primary system activity is key to reliable performance evaluation of dynamic spectrum access (DSA) technologies. Measurements on spectrum use show that in some cases the lengths of the vacant periods in a given frequency band can be correlated, a feature commonly used semi-Markov ON/OFF-models cannot reproduce. In this paper we study the stochastic properties of a simple primary system activity model constructed by aggregating the realisations of several semi-Markov ON/OFF-models. We show that despite its simplicity, our model does result in the desired correlations provided heavy-tailed distributions are used for the period length distributions of the individual processes. We also study in detail the statistical behaviour of our model as parameters such as number of processes being aggregated and distributions of the period lengths are changed. The results show that by suitable parameter selection wide range of behaviours can be modelled, making the model of interest in performance evaluation of algorithms and protocols for DSA networks.**

I. Introduction

Dynamic spectrum access (DSA) is emerging as a promising technology for reclaiming some of the licensed but currently unused spectrum [1], [2]. Reliable detection of the presence of the *primary user* (PU) whom the spectrum band is licensed to by means of spectrum sensing is important to guarantee good performance for both the primary as well as the *secondary user*, that is, the system performing DSA. Understanding the statistical behaviour of the spectrum use of the primary is key to reliable and efficient sensing.

Geirhofer *et al.* have shown in [3] empirically that the spectrum use in WLAN-systems can be modelled by means of a *semi-Markov ON/OFF-process*. In such a model a given channel is either occupied (ON) or vacant (OFF), with lengths of the ON- and OFF-periods being random variables following some specified distributions. Geirhofer *et al.* have fitted the Hyper-Erlang distribution to the WLAN-case. In earlier work [4] we investigated further primary user technologies using a more generic measurement approach.

Semi-Markov models have also been applied in theoretical work especially under the assumption of independent and geometrically distributed lengths of the ON- and OFF-periods (see, for example, [5]–[7]). In [4] we showed that the lengths of successive ON- and OFF-periods are sometimes independent, but this does not hold for all technologies. Indeed, for several bands correlations between successive period lengths were observed. Additionally, the distributions of the periods were often found to be *heavy-tailed* instead of geometric.

These results show that there is clear need for simple primary user activity models featuring correlated ON- and OFF-period lengths with heavy-tailed marginal distributions. In this paper we propose a simple model in which individual primary users are still modelled as binary ON/OFF-processes with i.i.d. period lengths. We then assume that we observe a superposition of such processes, that is, the channel is sensed as occupied if any of the individual processes is in the ON-state at sensing time. This is motivated by several spectrum sensing techniques, e.g., energy detection [8], which cannot provide the number of active primary users but determine a channel as being occupied if there is even a single active primary user in the sensing range. Our model also resembles closely commonly used techniques for generating *self-similar* traffic models for network performance evaluation [9]. We show that such a simple model can qualitatively reproduce correlations between successive OFF-period lengths (key statistic for reliability of spectrum sensing) while retaining the observed heavy-tailed period lengths.

The rest of the paper is structured as follows. In Section II we briefly recall results from our earlier work on the statistics of present-day primary user technologies. Our focus will specifically be on the features that earlier models were unable to replicate motivating the present work. We then describe our system model in more details in Section III, also drawing parallels to earlier applications of similar models in networking research. The statistics of the model are then studied by means of a simulation approach described in Section IV, with results given in Section V. Finally, we draw conclusions and outline future work in Section VI.

II. Measured Statistical Properties

We measured spectrum use in various spectrum bands at several locations in Germany and Netherlands over longer time durations of multiple days up to two weeks [10], [11]. During

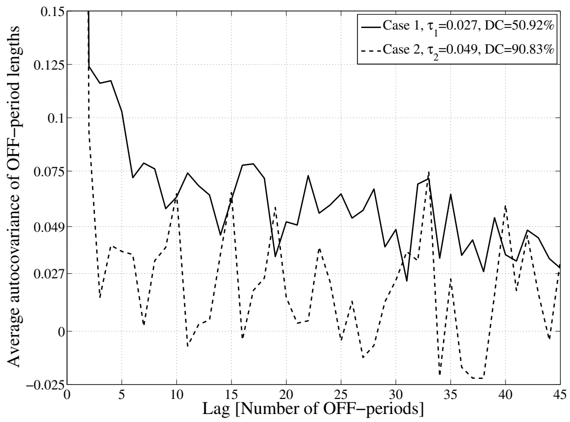

Fig. 1. Average autocovariance of the time series of consecutive OFF-period lengths $\overline{\varphi_L}$ of a selected GSM1800 uplink channel as measured at a location in Maastricht, Netherlands.

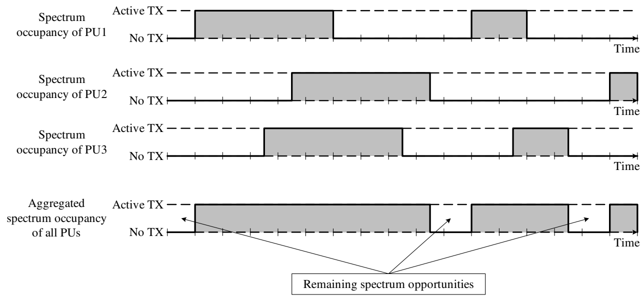

Fig. 2. Process aggregated from multiple binary ON/OFF processes of individual primary users based on a binary OR-operation (TX abbreviates transmission).

the detailed analysis of the gathered data we also evaluated the length of the ON- and OFF-periods of each measurement channel. Since we measured the received signal power and applied energy detection we cannot differentiate between individual primary users and focused on the aggregated process.

In [4] we reported that ON- as well as OFF-period lengths may be correlated over time. Here, we focus on the OFF-periods since they are more interesting in a DSA scenario. Let φ_L denote the autocovariance function (ACV) of consecutive OFF-period lengths:

$$\varphi_L(m) = \frac{1}{N-|m|} \sum_{n=0}^{N-|m|-1} \left(L_{n+m} - \overline{L}\right)\left(L_n - \overline{L}\right), \quad m \geq 0. \tag{1}$$

where L_n is the length of the n-th OFF-period. In total, the trace under study consists of N OFF-periods with average length $\overline{L}$.

We examined multiple different technologies in [4] but show results for two uplink channels of the Global System for Mobile Communication in its flavour working at 1800 MHz (GSM1800) as examples in Figure 1. In order to decide if a autocovariance result describes a significant deviation from random behaviour it is often compared to a significance threshold τ, which depends on the number of available samples N [12]:

$$\tau = \frac{1.96}{\sqrt{N}}. \tag{2}$$

We also included τ to Figure 1. Additionally, we list the duty cycle (*DC*) for both cases, which is the fraction of time during which the received power was above the selected energy detection threshold [8]. In our scenario, we applied the detection threshold $\gamma =$-107 dBm/200 kHz, as it has been proposed at the beginning of the IEEE 802.22 standardization for the detection of wireless microphones [13].

The ACV for trace two with high *DC* does rarely pass the significance threshold showing that consecutive OFF-periods could be modelled as independent and identically distributed (i.i.d.) random variates. In contrast, in case one, an i.i.d. modelling would be less appropriate due to the rather clear correlations between adjacent OFF-period lengths. We found similar results especially for the OFF-period lengths in cases with low *DC* and only few and short ON-periods [4].

For most of these traces the distribution of OFF-period lengths could be fitted well to log-normal distributions with significantly higher probability for very long OFF-periods compared to the commonly used exponential or geometric distributions.

III. System Model

Our system model closely reflects the measurement setup discussed in the previous section. We assume that spectrum occupancy is measured in binary fashion (ON/OFF) in *discrete time*. We further assume the presence of n primary users, each modelled as statistically identical semi-Markov ON/OFF-process. If any of the primary users is in the ON-state at a given time instant, spectrum is taken as being occupied, that is, to be in the ON-state. This way the overall spectrum occupancy is modelled as an ON/OFF-process obtained as a logical OR of the individual processes as illustrated in Figure 2. The parameters of the model are the total number of primary users, and the distributions of the ON- and OFF-periods of the individual primary user activity models, which we assume to be identical for all primaries for simplicity.

The model studied in this paper is heavily motivated by a well-known network traffic model. Numerous measurements have shown that the total amount of traffic observed at a fixed point in a fixed network is often *self-similar*, that is, the autocovariance function of the traffic volume follows a power law [14], [15]. As shown in [9], such a self-similar behaviour can be reproduced by assuming individual nodes as following a semi-Markov ON/OFF-process with heavy-tailed distributions for the ON- and OFF-periods. When in ON-state a node generates traffic at constant rate, while in OFF-state no traffic is produced. The observed process is the arithmetic sum of the traffic volumes produced by the individual nodes,

TABLE I
SIMULATION PARAMETERS.

Scenario	Number of primary users	*DC* [%]	OFF-period lengths k	x_m	ON-period lengths k	x_m
DL	1…25	24	1.1…2.5	30	2.3	adapted to match
AH	1…70	8.5	1.1…2.5	65	2.3	the selected *DC*

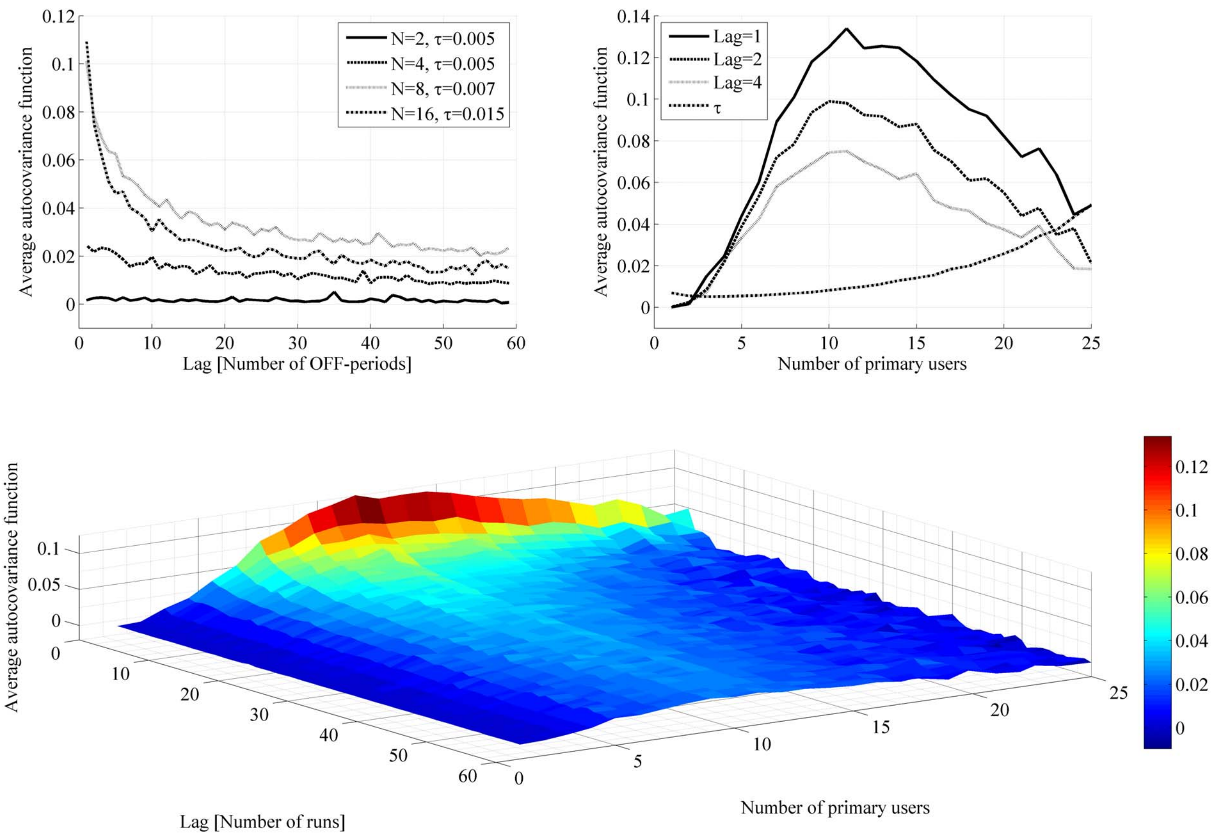

Fig. 3. Average autocovariance of the time series of consecutive OFF-period lengths simulated for an aggregated ON/OFF-process. The individual primary users have each $DC \approx 24\%$ (DL-scenario). The Pareto shape parameter k for the OFF-periods of the individual primary users is $k = 1.1$.

which can be shown to be self-similar by analytic means. The key difference between our model and the network traffic model is in the way the aggregation is performed. In our case logical OR is used, whereas simple summation is carried out in the latter case. We shall study in the following by means of Monte Carlo simulations whether this change affects the desired properties of the model, that is, the capability of introducing correlations between successive OFF-period lengths of the superimposed process while retaining heavy-tailed distributions for the period lengths.

IV. Simulation Model

We ran extensive simulations in order to evaluate the autocovariance properties of a superimposed process and the impact of selected system parameters. The binary activity pattern of each primary user is generated separately while assuming that all have the same statistical characteristics. We use the Pareto-distribution with probability distribution function $f(x)$ to model ON- and OFF-period lengths:

$$f(x) = k\frac{x_m^k}{x^{k+1}}, \quad x \geq x_m, \tag{3}$$

where k is usually referred to as the *shape* parameter and x_m is known as the *scale* parameter of the Pareto-distribution. We did additional tests using log-normal and geometric distributions. However, in the following we will focus on the case with Pareto-distributions applied to both ON- and OFF-period lengths.

We differentiate two scenarios, namely the *downlink* (DL) and the *adhoc* (AH) case. Table I summarizes the simulation

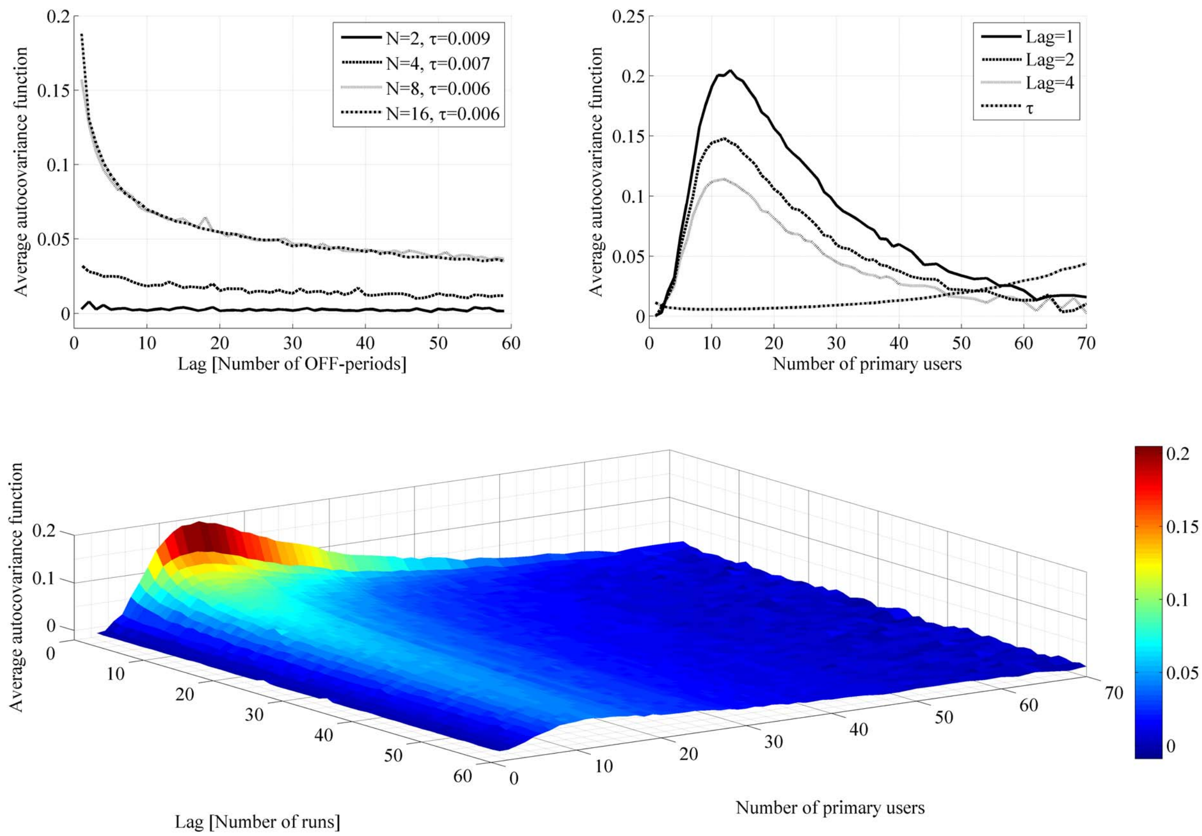

Fig. 4. Average autocovariance of the time series of consecutive OFF-period lengths simulated for an aggregated ON/OFF-process. The individual primary users have each $DC \approx 8.5\%$ (AH-scenario). The Pareto shape parameter k for the OFF-periods of the individual primary users is $k = 1.1$.

parameters for both. The major difference is the lower *DC* in the AH-scenario that allows to add more primary users before the primary system fully loads the channel and no opportunities for secondary access are left.

The binary ON/OFF-processes generated per primary user are aggregated using an OR-operation as discussed in Section III. In the DL-simulations we used traces of 2.5 million binary samples and traces of 1.5 million samples for the AH scenarios, respectively. We ran $S = 25$ simulations per parameter set and averaged the results across all simulation runs, e.g.:

$$\overline{\varphi_L}(m) = \frac{1}{S}\sum_{i=1}^{S}\varphi_{L,i}(m). \tag{4}$$

V. Results

We investigate $\overline{\varphi_L}$ for the OFF-periods and its dependence on the number of primary users. Figure 3 shows that consecutive OFF-period lengths are more or less independent if only very few primary users are active. This result is expected because single OFF-period lengths and also the OFF-periods of different primary users are generated independently. However, correlations will emerge if the number of primary users is further increased. These are especially clear for short lags of few OFF-period lengths.

If the number of primary users increases further $\overline{\varphi_L}$ starts to drop again. At the same time the significance threshold τ increases. The latter result is caused by the very high *DC* of the trace aggregated from more than approximately 12 primary user traces. The channel is free only for very short OFF-periods and also the number of these OFF-periods decreases significantly. Additionally, the variance of the OFF-period lengths that occur in these aggregated traces is low limiting the ACV due to its inherent mean-correction.

If we lower the *DC* per primary user the effect is slightly clearer as can be seen from Figure 4. More primary users can be accommodated due to the lower load per user but the underlying effect is the same. Additionally, the highest correlation occurs again for approximately 12 primary users. Thus, the emergence of the correlation seems to be rather independent of the primary user *DC* as long as some capacity is left for secondary operation.

Figure 5 evaluates the effect of the Pareto-shape parameter k used to generate OFF-period lengths. For very low parameter values more users are required to induce the correlation-

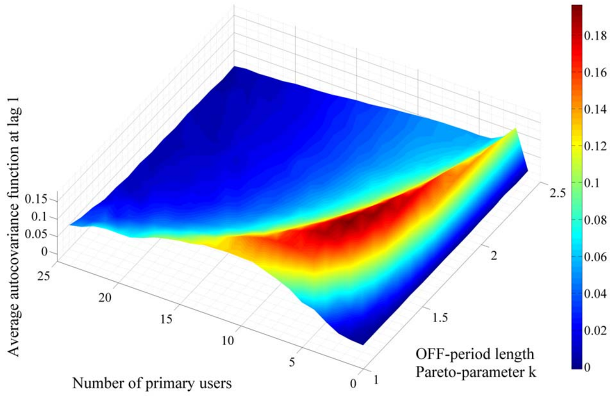

Fig. 5. Average autocovariance at lag 1 of the time series of consecutive OFF-period lengths simulated for an aggregated ON/OFF-process. The impact of the number of nodes and the shape parameter k of the Pareto-distribution of the OFF-period lengths are compared. The individual primary users have each $DC \approx 24\%$ (DL-scenario).

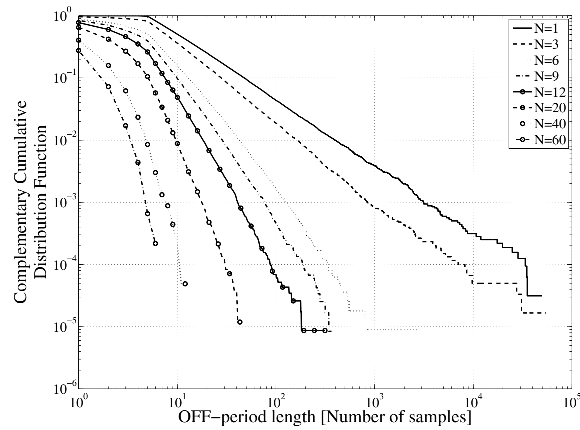

Fig. 6. Length distribution of the OFF-periods of the aggregated, simulated trace. The individual primary users have each $DC \approx 8.5\%$ (DL-scenario).

effect. This is probably caused by the infinite variance of the distribution and, thus, a not negligible probability to get very long OFF-periods. The impact of these extremely long periods is strong enough that a higher number of primary users is required to change the characteristics of the aggregated process.

The effect is also present if the Pareto- is replaced by a log-normal distribution although it is slightly less explicit. The log-normal distribution is more heavy-tailed than a geometric distribution but its tail is shorter compared to a Pareto-distribution with any of the examined k-values. Thus, the log-normal case could be seen as a continuation of Figure 5 for a higher value for the shape parameter k. Since the basic effect of emerging correlation is still present it may be an answer to the question for the causes of the correlations that we detected in our measurement traces.

The correlation-effect is controlled mostly by the distribution of the OFF-period lengths. Replacing the Pareto-distribution for the ON-periods by, e.g., a geometric distribution does hardly change the result.

A. Distribution of OFF-period lengths of the aggregated process

The second main statistical characteristic that we want to investigate is the OFF-period length distribution of the aggregated process. Figure 6 shows the complementary cumulative distribution function (CCDF) of the distribution based on the DL-scenario. The heavy-tailed nature of the input distribution for one primary user is clearly visible. However, the straight line in the double-logarithmic plot, typical for power-law distributions, is present also for traces aggregated from more and more sources. Since the maximum OFF-period length decreases due to the increased load of the channel, the slope of the straight line considerably decreases making it very difficult to accurately decide on the presence of power-law behaviour [16]. Thus, our simulation results are not sufficient to conclude that the correlation-effect and the heavy-tailed characteristic of the period length distribution do or do not occur concurrently.

The distribution of the ON-period lengths changes its shape much faster and looses its heavy-tailed nature due to the specific characteristics of the OR-operation applied during the trace aggregation.

VI. Conclusions

In this paper we studied a conceptually simple model for activity of a primary system constructed by aggregating realisations of semi-Markov ON/OFF-models. Based on extensive simulations we showed that our model is capable of reproducing qualitatively correlations between lengths of successive vacant periods observed in measurements. In earlier work binary semi-Markov models have been used directly, resulting in absence of such correlations. We believe our model can be used as a useful starting point for future studies on performance for DSA-related algorithms and protocols, especially in studying the sensitivity of the results on the behaviour of the primary system.

Interesting items for future work include developing analytical approaches towards analysing the statistics of the proposed model, as well as further clarifying the relationship between the correlations and heavy tails in the period length distributions as the number of aggregated processes is increased. It would also be very interesting to carry out further spectrum use measurements to study whether cases in which correlations between OFF-period lengths were observed in earlier measurements could indeed be *quantitatively* explained by our model. This would require refinement of the measurement approach to enable studying the activity of the individual transceivers of the primary system instead of measuring only the aggregated system behaviour.

Acknowledgment

The authors would like to thank the RWTH Aachen University and the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) for providing financial support through the UMIC research center. We would also like to thank the European Union for providing partial funding of this work through the ARAGORN project.

References

[1] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Elsevier Computer Networks Journal*, vol. 50, no. 13, pp. 2127–2159, September 2006.

[2] Q. Zhao and B. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, May 2007.

[3] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic spectrum access in the time domain: modeling and exploiting white space," *IEEE Communications Magazine*, vol. 45, no. 5, pp. 66–72, May 2007.

[4] M. Wellens, J. Riihijärvi, and P. Mähönen, "Empirical Time and Frequency Domain Models of Spectrum Use," *Accepted for publication in Elsevier Physical Communication Journal, Special Issue on Cognitive Radio: Algorithms & System Design*, May 2009.

[5] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 589–600, April 2007.

[6] 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.

[7] L. Yang, L. Cao, and H. Zheng, "Proactive Channel Access in Dynamic Spectrum Networks," in *Proc. of International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom)*, Orlando, FL, USA, August 2007, pp. 487–491.

[8] H. Urkowitz, "Energy Detection of Unknown Deterministic Signals," *Proceedings of the IEEE*, vol. 55, no. 4, pp. 523–531, April 1967.

[9] M. S. Taqqu, W. Willinger, and R. Sherman, "Proof of a fundamental result in self-similar traffic modeling," *ACM SIGCOMM Computer Communication Review*, vol. 27, no. 2, pp. 5–23, April 1997.

[10] M. Wellens, J. Wu, and P. Mähönen, "Evaluation of spectrum occupancy in indoor and outdoor scenario in the context of cognitive radio," in *Proc. of International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom)*, Orlando, FL, USA, August 2007, pp. 420–427.

[11] M. Wellens and P. Mähönen, "Lessons Learned from an Extensive Spectrum Occupancy Measurement Campaign and a Stochastic Duty Cycle Model," in *Proc. of International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (TridentCom)*, Washington D.C., USA, April 2009.

[12] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*, 2nd ed. Springer, 2002.

[13] S. Shellhammer and G. Chouinard, "Spectrum sensing requirements summary," IEEE 802.22-05/22-06-0089-05-0000, July 2006.

[14] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, no. 1, pp. 1–15, February 1994.

[15] A. Erramilli, M. Roughan, D. Veitch, and W. Willinger, "Self-similar traffic and network dynamics," *Proceedings of the IEEE*, vol. 90, no. 5, pp. 800–819, May 2002.

[16] A. Clauset, C. R. Shalizi, and M. E. J. Newman, "Power-law distributions in empirical data," *accepted for publication in Society for Industrial and Applied Mathematics (SIAM) Review (SIREV)*, available online at http://arxiv.org/abs/0706.1062v2 [Cited on: 5th March 2009], revised in February 2009.