Estimating Transmitter Activity Patterns: an Empirical Study in the Indoor Environment

Elena Meshkova, Junaid Ansari, Janne Riihijärvi, Jad Nasreddine and Petri Mähönen
Institute for Networked Systems, RWTH Aachen University, Germany
email: {eme, jan, jar, jad, pma}@inets.rwth-aachen.de

***Abstract*—We empirically study problems of online estimation and modeling of radio activity (ON/OFF) patterns and their duty cycles. We consider performance of two algorithms that analyze power spectrum samples in time domain as sensed by WARP SDR boards placed at different indoor locations. The first algorithm utilizes density function of the marginal distribution of the received readings to perform clustering of the power samples into the classes that correspond to ON/OFF activity states. The second algorithm uses hidden Semi-Markov processes and the Viterbi algorithm to model and estimate the radio activity patterns of signal sources. Unlike the first algorithm it can also be used for both signal classification and estimation of the state duration that corresponds to the certain power level. We study signals obtained from one or two transmitters. Our results indicate the first technique can often provide accurate estimations. However, if sensor nodes receive low RSSI values or modeling of the temporal structure of the activity patterns is required then the second technique is preferable. This method also has a potential to benefit from cooperation between nodes.**

I. Introduction

The success of cognitive networking depends on *self-reconfigurability* [1] and *dynamic characterization of the radio environment* [2]. In this context, the concept of Radio Environment Maps (REMs) has been proposed to support Radio Resource Management (RRM) tools by providing information that characterizes the radio environment accurately. Compared to relatively static scenarios such as the TV white space reuse, the construction and maintenance of REMs becomes more challenging when the interference patterns are less predictable such as in scenarios with Dynamic Spectrum Access (DSA) enabled networks [3] and femtocells [4]. Constructing REMs for indoor environments is not straight-forward due to the complex propagation characteristics and interference patterns. Dynamic deterministic models based on real measurements are possible candidates in this type of environment. Using these models different nodes in a network can collaborate in order to create spatio-temporal maps of the environment that are valid for specific time and propagation conditions.

In this paper we empirically study performance of two algorithms that allow online modeling and estimation of radio activity patterns (ON/OFF states) based on the strength of the received power spectrum samples. The work is done on the power spectrum samples, gathered by the WARP SRD boards [5] located at different distances from the transmitters [6]. Information provided by the algorithms can be used, for example, to characterize temporal aspects of the REMs or MAC sensing applications [7]–[9].

The first of the proposed algorithms is computational efficient. It determines the thresholds to distinguish between different activity states and corresponding power levels. These thresholds are later used to classify the incoming samples. The algorithm can provide only basic estimates on the temporal structure of the signals: an average duration of a particular state and resulting duty cycle estimates. The second method is more complex, as it is based on Semi-Markov models and the Viterbi algorithm [10]. The algorithm learns statistical models for each activity state, along with the probability distribution of the holding times, i.e. how long the sensor is likely to sense corresponding power levels. Application of these models results in better signal classification. Both methods are applied online. In this work we report results for one and two transmitters, with activity patterns following the gamma distribution. We analyze both strong and weak power spectrum samples. Additionally we show possible benefits of cooperative sensing for correctly estimating the ON/OFF activity patterns with the hidden Semi-Markov processes using data obtained at different sampling rates.

The rest of the paper is structured as follows. Section II discusses the incentives for ON-OFF modeling and the chosen estimation algorithms. Section III describes the experimental setup. In Section IV we analyze the performance of the chosen algorithms on the gathered measurement results. Finally, Section V concludes the paper.

II. Algorithms for Estimation of Power Spectrum Usage

One may distinguish between two basic application scenarios for temporal power spectrum sensing data:

1) estimation of active (ON states) or inactive (OFF states) of different transmitters, including estimation of several ON power levels, i.e. power-level clustering;
2) modeling and prediction of the occupancy time for a particular frequency band.

The first scenario basically allows to detect *instantaneous* channel occupancy. The information on the power levels and their average duty cycles can be used to perform run-time optimization of PHY- and MAC-layer parameters depending on the encountered amount of interference [11]. The second scenario allows to optimize Dynamic Spectrum Access (DSA) by utilizing temporal power spectrum data including *historical* information. Earlier work on the so-called MAC-layer sensing [7], [9] has shown that knowledge of either

978-1-4577-1348-4/11/$26.00 ©2011 IEEE

the *duty cycle* or the complete distribution of ON and OFF periods can be used to considerably enhance the efficiency of DSA algorithms. Ideally the second application scenario would allow to identify multiple signal sources in a network, predict their activity patterns and, therefore, effectively schedule transmissions of secondary users.

Depending on the application scenario, mechanisms of different complexity are to be applied for data processing. In this paper we consider two algorithms described below.

A. Density-based Threshold Detection Algorithm

This simple computationally efficient algorithm is applicable to fulfill the goals of the first application scenario. The density function of the marginal distribution of the collected data is used to determine thresholds to distinguish between several power levels, i.e. network *activity states*, that are later used for online classification of the power samples.

First, the algorithm determines number of power levels in which the samples are clustered. It requires the user to set the minimal signal duration of a transmitter, and the minimal expected difference between power levels. Using these inputs parameters for the moving average algorithm are determined in order to smooth the obtained samples. After the data is processed with this simple finite impulse response filter, the algorithm estimates the number of local maxima of a density function that are above certain level (e.g. $thr_{densityMax} = 0.02$) to determine the initial number of the activity states.

At the second stage the algorithm finds the *thresholds* between these states, that later will be used for classification of the samples. The algorithm cancels the data filtering. Operating on the raw data it finds local minima of a density function. These values should be less than a certain value, for e.g. $thr_{densityMin} = 0.05$, required for clear separation between the power levels. The number of these minima should be one less than the number of the clusters. If the above criteria on local minima are not satisfied then the algorithm gradually increases the number of samples the data is averaged over until the density function of the pre-processed samples does not yield the required results. Typically noisy or weak power spectrum samples require such an iterative filtering, which decreases data granularity, and in the extreme cases might prevent from detecting short ON activity bursts.

B. ON/OFF Patterns with HSMM and Viterbi Algorithm

This algorithm is more complex and is applicable in cases of severe propagation conditions, heavily sub-sampled data and the second application scenario. The algorithm (a) clusters received samples in the power levels corresponding to, in general case, the OFF and several ON network activity states, (b) provides their statistical characteristics also addressing the temporal aspects, and (c) use this data for online classification of the incoming readings. For example, for a single transmitter it estimates the distributions of both, power samples that correspond to the ON and OFF states. It also characterizes the lengths of the activity periods of the transmitter. This

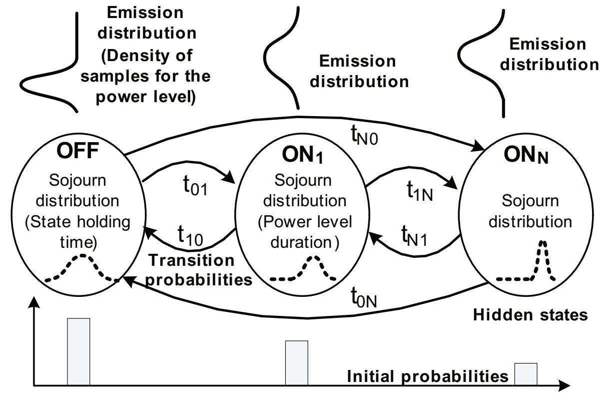


Fig. 1. Generalized hidden Semi-Markov model used. The figure shows three states of the HSMM corresponding to the OFF activity state of the network and two ON states that are characterized by different power levels.

information is then used to estimate if an incoming sample belongs to the ON or OFF network activity state.

The proposed algorithm models the activity patterns of transmitters by simple ON/OFF processes. It utilizes *Semi-Markov ON/OFF process* and uses the following transmission model. Each transmitter at any given time is either active (ON state) or inactive (OFF state). While being active the transmitter is assumed to send signals continuously with a constant power. The model relies on the further assumption that the durations of active and inactive periods for each transmitter are independent of each other. The earlier measurements in outdoor conditions have indicated that often for several different transmitter technologies the lengths of successive ON and OFF periods are independent [7], [8]. The durations of ON/OFF periods are presumed to follow general distributions $f_{\text{on}}(t)$ and $f_{\text{off}}(t)$ with mean values μ_{on} and μ_{off}, respectively.

If the states of the different transmitters were directly observable, the estimation of the ON and OFF distributions would easily be solved using usual parametric or non-parametric statistical techniques. In practice this occurs if there is a single transmitter of high enough power that it can be reliably distinguished from ambient interference and noise [8]. However, especially in indoor scenarios, fast fading and hostile propagation environment often cause the measured received power from an active transmitter to fluctuate heavily. Signal can often disappear into the noise and interference for short periods of time. We overcome these difficulties by considering a *Hidden Semi-Markov Models* (HSMM) [12][1]. The true state of the system is still assumed to follow a Semi-Markov ON/OFF process (or in the case of multiple transmitters, a product of such processes). But instead of making observations on that process directly, we assume that we observe at given times a received power that is random with distribution depending on the actual state.

The general structure of the employed HSMM is shown

[1]This method was first applied on the empirical data for the offline activity patter estimation in [8]. However, in this paper the reported difference between the spectrum power samples corresponding to ON/OFF levels was high and, therefore, the challenge for the states estimation was minimal.

in Fig. 1. The model requires specification of a) *number of states* in the HSMM graph, b) *initial probabilities* of states, c) *transition probabilities* between states, d) *sojourn times* spent in each of the states (corresponding to $f_{\text{on}}(t)$ and $f_{\text{off}}(t)$ for a single transmitter case), e) the *emission distributions* of the received power conditional on the state of the system [10].

For the initial classification of the power levels, estimation of their empirical averages and standard deviations, we use the density-based threshold detection algorithm described above. The number of states of the HSMM is set to be equal to the number of resulting power classes. The statistical characteristics of these power classes are used to set initial values the emission distributions of the HSMM. We use equal initial state probabilities for the HSMM. The transition probabilities between all HSMM states are initially set to be even as well. However during the model training, which we describe later they get automatically adjusted. We applied normal emission distributions throughout, but these can easily be replaced by more general distributions to cover the effects of, for example, non-Gaussian interference. The first zero of the autocorrelation function of the received power was used to determine initial values for the parameters of the sojourn distributions whenever parametric model was considered. We chose the gamma distribution as parametric models for the sojourn time distributions due to its flexibility. The scale parameter θ was chosen to have initial value ranging from 1 to 5, while the shape parameter was set as $k = \mu\theta$, where μ is the time scale obtained from the analysis of the autocorrelation function as discussed above.

To refine the initial values of emission and sojourn distributions and adjust state transition probabilities, the method of finding the maximum likelihood iteratively using the *Expectation Maximization (EM) algorithm* [12] is typically applied. In particular, we use the implementation of [13] specifically developed for estimating hidden Semi-Markov models.

As said the HSMMs provide distributions of expected power levels and ON/OFF period durations. We utilize these models by employing the Viterbi algorithm, as proposed in [10], to estimate the most likely sequence of hidden states from measurements. This information is used to classify the incoming power readings in one of the states of the system (in general case the OFF or one of the ON power levels).

III. Experiments Setup

In our measurements we used 8 WARP SDR boards [5]. They were located in five different rooms covering over $240\,\text{m}^2$, with one semi-concrete wall dividing the space. (The detailed description of the testbed is given in [6]). These boards were continuously sampling the spectrum of a bandwidth of 22 MHz with frequency of $740\,\mu$s during the periodic transmissions generated by TelosB nodes [14]. The MAX2829 [15] radio transceiver of the WARP board provides a 10-bit (Received Signal Strength Indicator) RSSI samples, which are converted into dBm power readings.

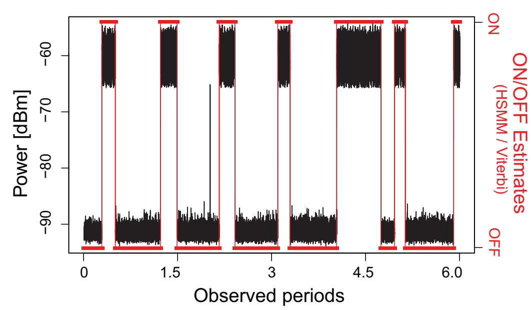
Signals were produced by a one or two active transmitters using the test mode of CC2420 radio transceiver chip on TelosB [16]. The chip provides a continues 5 MHz wide signal with most of the power concentrated in a bandwidth of 2 MHz [6]. The transmit power was set to 0 dBm. The duty cycle of the first transmitter was 25%, with average ON period of 1.25 seconds and variance of $0.16\,s^2$. The second transmitter generated a signal with 50% fixed duty cycle, the ON and OFF durations were on average equal to 2.5 s.

After the measurement traces were collected we have chosen for the analyzes the two traces that correspond to the sensor boards receiving the highest and the lowest power readings from the transmitters. To estimate the performance of the proposed methods we fed the traces as incoming online samples to our algorithms realized in R [17].

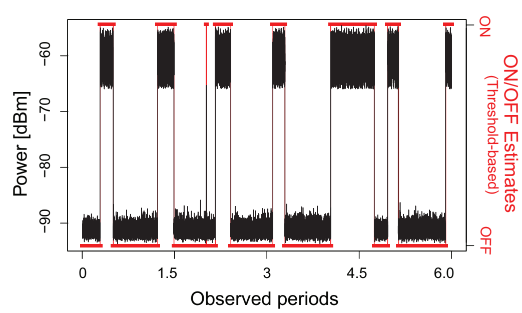
IV. Analysis of Measurement Results

We apply the two proposed algorithms for the estimation of the ON/OFF activity patterns produced by one or two TelosB nodes. Figs. 2 – 4 show the basic characteristics of the power readings received by the WARP boards, including the sample time series and the corresponding density functions of the marginal distributions of the gathered samples. For the single transmitter scenario we analyze the readings for two WARP boards that received the strongest and the weakest signals. For the scenario with two transmitters we process the measurement traces only from the weakest receiver.

We annotate the obtained results in terms of the number of periods observed by the receivers rather than in seconds. One period is defined as a sum of the average ON and OFF durations and basically corresponds to one duty cycle. In this work one period is equal on average to 5 s.

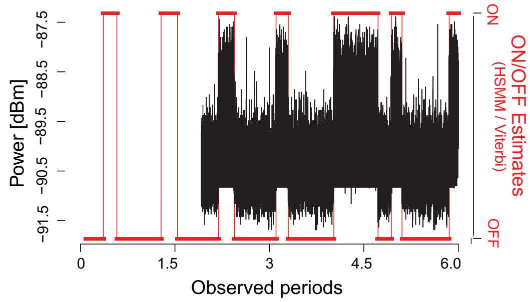


(a) Received power levels and the predicted states with the HSMM/Viterbi algorithm.

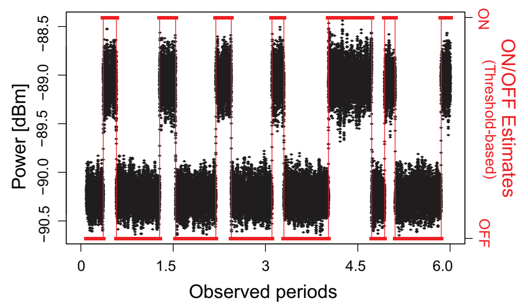


(b) Received power levels and the predicted states with the threshold-based algorithm.

Fig. 2. Single transmitter scenario. The WARP board receives a strong signal. The receiver is located at the distance of 0.5 meters from the transmitter. Both algorithms for training use sample gathered over 16 periods.



(a) Received power levels and the predicted states with the HSMM/Viterbi algorithm.



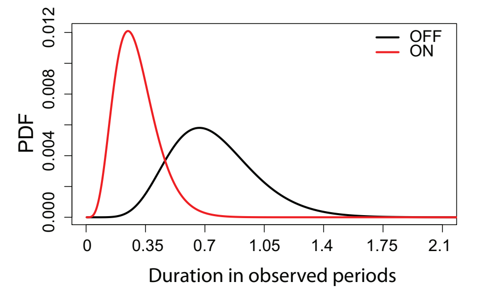
(b) Received power levels and the predicted states with the threshold-based algorithm that uses the moving window of 12 samples.

Fig. 3. Single transmitter scenario. The board receives a weak signal. It is located at the distance of 11 meters from the transmitter across the semi-concrete wall. Both algorithms for training use sample gathered over 16 periods.

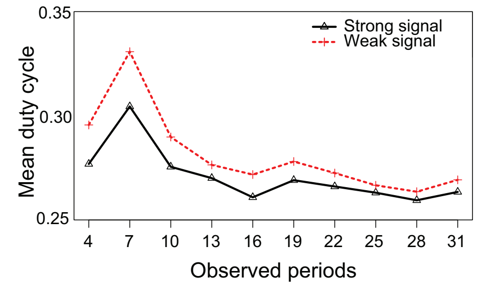
A. Single Transmitter

First we apply the proposed algorithms to process the power spectrum samples estimated from RSSI readings of a WARP board as discussed in Section III. The chosen WARP board receives strong signals. The results are shown in Fig. 2. The density-based threshold detection algorithm correctly identifies two classes for the power readings corresponding to the ON and OFF states, and uses the threshold level set to -69.95 dBm to correctly classify the incoming samples (see Fig. 2b). The algorithm does not require any data smoothing. As shown in Fig. 2a the HSMM/Viterbi algorithm also correctly estimates the ON/OFF states of the incoming samples based on the learned Semi-Markov model. Moreover, this algorithm is capable to correctly process the erroneous/noise sample received around the second period and classify it as the OFF state (see Fig. 2), as it takes the temporal structure of the activity states into the account. The first algorithm operates only on the thresholds, which leads to the signal misclassification. Overall, both algorithms successfully classify the samples, when the difference between the noise level and the power received from an active transmitter is high.

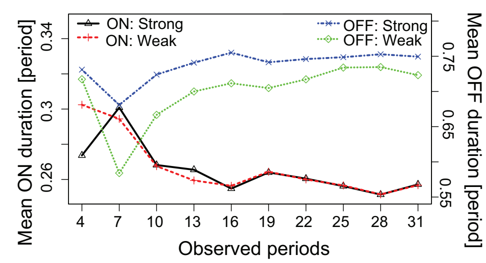
However, the threshold-based algorithm requires additional processing when this difference becomes small (in our study the signal averages are less than 2 dB apart). If no data filtering is applied then samples are misclassified as the two power levels are very close to each other. The algorithm could achieve correct classification of the samples with application of the



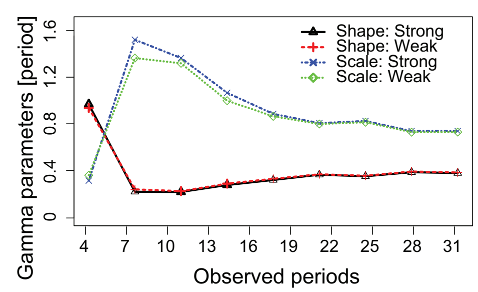
(a) The sojourn distribution for the HSMM for the WARP board receiving a weak signal.



(b) Mean duty cycles estimated after the models' training as the function of the duration of the input time series given in periods.



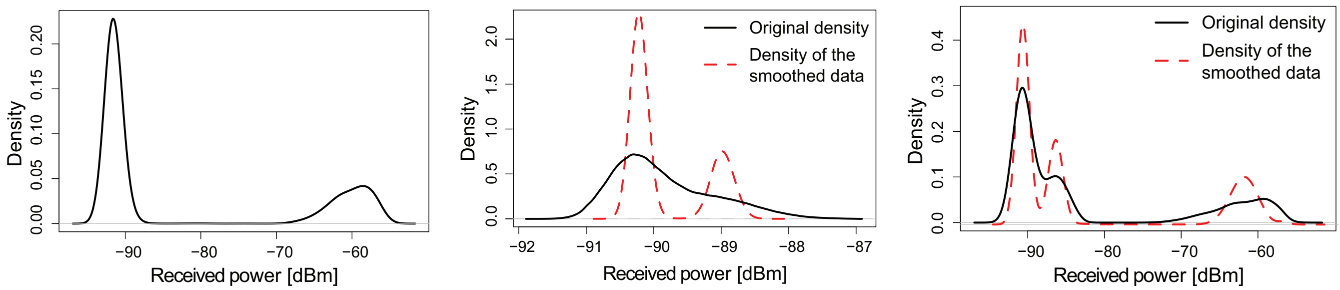
(c) Mean ON/OFF periods estimated after the models' training as the function of the duration of the input time series given in periods.



(d) Estimates of shape and scale parameters of the sojourn time distributions for two WARP boards as the function of the duration of the input time series.

Fig. 5. The summarizing plots on the parameter estimates of the HSMM for the single transmitter scenario following the gamma signal distribution.

moving average algorithm with the window size of 12 samples, setting the threshold to -89.64 dBm. High window size means that the heavy smoothing of the data is required, which can potentially lead to the signal losses. The resulting density function is shown in Fig. 4b and the results of the classification



(a) The WARP board receiving the strong signal in the single transmitter scenario. (b) The WARP board receiving the weak signal in the single transmitter scenario. (c) The WARP board receiving the weak signal in the two transmitter scenario.

Fig. 4. Probability density functions obtained for the power spectrum samples received by the WARP board in the two considered scenarios.

are displayed in Fig. 3b.

The results show that HSMM/Viterbi algorithm performs well without applying additional filtering in this scenario. The Viterbi algorithm achieves accurate state estimation based on the learned Semi-Markov model, as shown in Fig. 3a. The sojourn distribution for the HSMM for both strong and weak signal sources are the almost identical. This indicates the ability of HSMM to be successfully applied also for the receivers observing only very weak signals. The sojourn distribution for the weak signal is displayed in Fig. 5a. Of course, a certain performance degradation in terms of the required training samples exists for the weaker signal as shown in Fig. 5b-d. The difference in the estimated ON and OFF periods and duty cycles does not exceed 5 % from the true values. The figure also shows the fast convergence of the parameters of the Semi-Markov models to the almost true values of the original signal distribution.

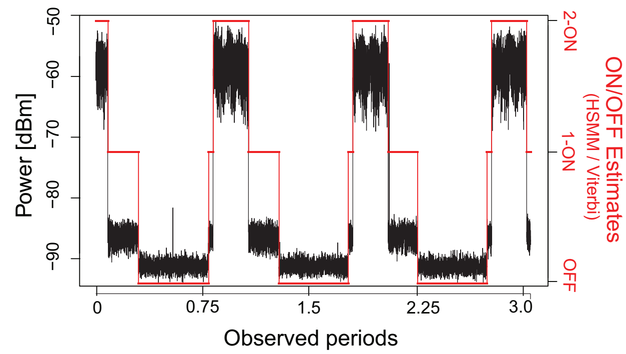
B. Two Transmitters

For the two transmitter scenario we have chosen the WARP board that receives very weak signals from one of the TelosB transmitters, as shown in Fig. 4c and 6. The weak signal from the furtherest transmitter can be easily misclassified into the OFF state. The average difference to the noise level for this transmitter's power does not exceed 1.5 dBm.

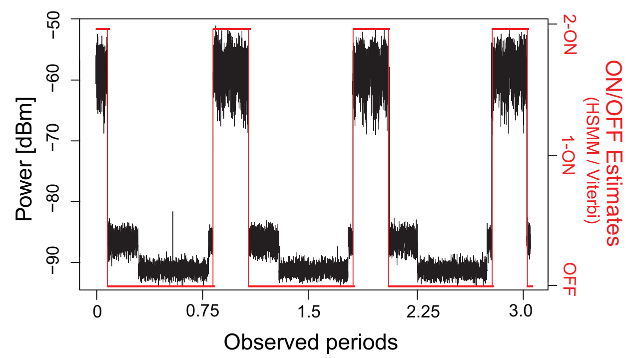
To obtain acceptable classification results for this scenario the density-based threshold detection algorithm had to apply a moving average filter with the window size of 10 samples. The algorithm correctly determined that the samples have classified into three power levels (OFF, 1-ON and 2-ON states) and deduced the classification thresholds of 89.3 dBm and 78.4 dBm. The results of the online state estimation with these thresholds are shown in Fig. 4c and 6c.

The HSMM/Viterbi algorithm has also correctly determined that the Semi-Markov model requires three states. Further training of the model was successful and the Viterbi algorithm performed correct clustering of the incoming samples without the need for data pre-processing, see Fig 6a.

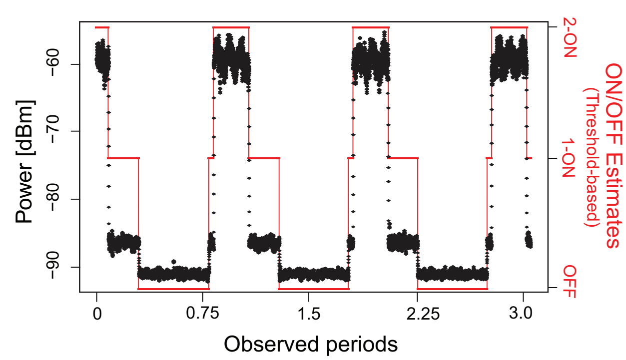
For comparison, we trained the two-state HSMM on the same set of data. The Viterbi algorithm completely misclassified 1-ON state into OFF state for the weak power spectrum samples, as shown in Fig 6b. This indicates that the HSMM/Viterbi approach is useful for modeling and estimating



(a) Received power levels and the predicted states with the HSMM/Viterbi algorithm that used three-state Semi-Markov model.

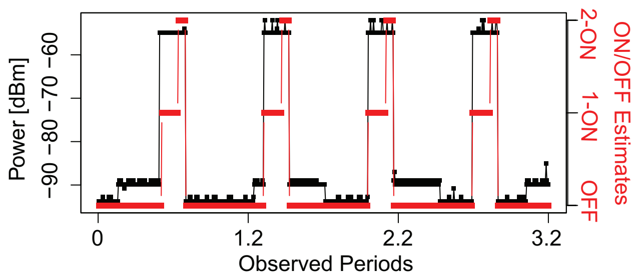


(b) Received power levels and the predicted states with the HSMM/Viterbi algorithm that used two-state Semi-Markov model.

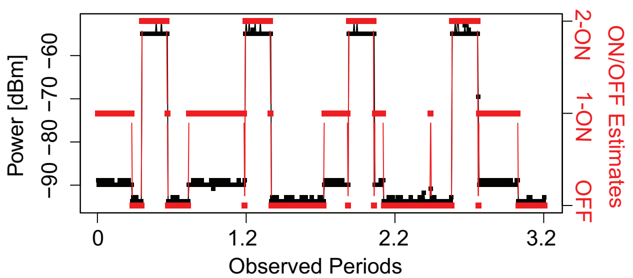


(c) Received power levels and the predicted states with the threshold-based algorithm that uses the moving window of 10 samples.

Fig. 6. Two transmitters scenario. The WARP board receiving a weak signal. The receiver is located at the distance of 9 and 11.5 meters from the transmitters across the semi-concrete wall. Both algorithms use for training samples gathered over 16 periods.



(a) Received power samples and the state prediction on the three state gamma model trained by the same with slow sampling.



(b) Received samples and the state prediction on the three state gamma model obtained from the nearby node, which performs fast sampling.

Fig. 7. Illustration of benefits of information sharing between the two receives for the two TelosB nodes transmitting scenario, sub-sampling case.

the state of complex power spectrum time series, if the number of states for the Semi-Markov models increases correspondingly. However, this leads to an increase in processing time and computational overhead.

We also used the two transmitter scenario to demonstrate possible benefits from cooperative nodes behavior. Two nearby WARP nodes performed the sensing at different rates, one 50 times slower than the other. If we estimate the three-state HSMM directly from the slower sensing device then we may arrive to the wrong estimations, as shown in Fig. 7a. Here the OFF estimate includes the low-power level of the ON signal. However, if we import the trained three-state model from the faster sampling sensor node to the slower sampling device then the states estimations made by this device become much more accurate, as shown in Fig. 7b.

V. Conclusions

In this paper, we have empirically studied the temporal aspect of the problem of transmitter activity modeling and estimation for the indoor environment. We have performed spectrum power sampling using a number of WARP boards to detect signals from one or two transmitters. Based on this data-set, we have then studied the differences between the measurements carried out by different receivers. We have shown that the density-based threshold detection algorithm can be successfully applied online to classify power samples and determine the threshold to distinguish between the ON and OFF states for both strong and weak signals. However, for the weak signals the algorithm requires strong smoothing of the data which might prevent it from detecting short transmission sequences. It has been also shown that the ON/OFF activity patterns can be successfully modeled using hidden Semi-Markov models without significant loss in data granularity. These models can be further used to accurately classify online the incoming power samples with the Viterbi algorithm. Furthermore, we have shown possible benefits from intra-node information sharing for more accurate spectrum activity prediction on the example of the sub-sampled data.

We plan to further extend our study to multiple interference sources and irregular ON/OFF activity patterns, as well as to dynamic environments. We want to study applicability of our algorithms to the cases where the assumption on independent durations of ON and OFF periods is violated.

Acknowledgment

The authors would like to thank DFG for partial funding through UMIC Research Center. We thank D. Denkovski, M. Pavloski and L. Gavrilovska for useful discussions.

References

[1] F. J. Mullany *et al.*, "Self-deployment, self-configuration:critical future paradigms for wireless access networks," *Lecture Notes in Computer Science, Springer*, vol. 3457/2005, pp. 58–68, 2005.
[2] Y. Zhao, B. Le, and J. H. Reed, *Cognitive Radio Technology*. Elsevier, ch. Network Support: The Radio Environment Map, pp. 337–363, 2006.
[3] R. Tandra, M. Mishra, and A. Sahai, "What is a spectrum hole and what does it take to recognize one?" *Proceedings of the IEEE*, vol. 97, no. 5, pp. 824–848, 2009.
[4] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: A survey," *Communications Magazine, IEEE*, vol. 46, no. 9, pp. 59–67, September 2008.
[5] A. Khattab *et al.*, "WARP: a flexible platform for clean-slate wireless medium access protocol design," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 12, no. 1, pp. 56–58, 2008.
[6] E. Meshkova *et al.*, "Experimental Spectrum Sensor Testbed for Constructing Indoor Radio Environmental Maps," in *Poster paper, IEEE DySPAN*, May 2011.
[7] M. Wellens, J. Riihijärvi, and P. Mähönen, "Evaluation of Adaptive MAC-Layer Sensing in Realistic Spectrum Occupancy Scenarios," in *IEEE DySPAN*, May 2010, pp. 1–12.
[8] S. Geirhofer, L. Tong, and B. Sadler, "Dynamic spectrum access in the time domain: Modeling and exploiting white space," *IEEE Communications Magazine*, vol. 45, no. 5, p. 66, 2007.
[9] H. Kim and K. 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, 2007.
[10] Y. Guédon, "Estimating hidden semi-Markov chains from discrete sequences," *Journal of Computational and Graphical Statistics*, vol. 12, no. 3, pp. 604–639, 2003.
[11] V. Mhatre, K. Papagiannaki, and F. Baccelli, "Interference mitigation through power control in high density 802.11 WLANs," in *IEEE INFOCOM*, pp. 535–543, 2007.
[12] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
[13] J. O'Connell and S. Højsgaard, *mhsmm: Parameter Estimation and Prediction for Hidden Markov and Semi-Markov models for Data with Multiple Observation Sequences*, R package v. 0.3.6, 2010, http://CRAN.R-project.org/package=mhsmm.
[14] *TelosB Node*, Crossbow Inc., http://www.willow.co.uk/TelosB_Datasheet.pdf, Last visited: 11.06.2010.
[15] *MAX 2829 datasheet*, MAXIM, http://datasheets.maxim-ic.com/en/ds/MAX2828-MAX2829.pdf, February 2010, last visited: 11.06.2010.
[16] *Chipcon CC2420 Datasheet*, Last visited: 15.06.2010, http://www.chipcon.com/files/CC2420_Data_Sheet_1_3.pdf.
[17] R. Ihaka and R. Gentleman, "R: A language for data analysis and graphics," *Journal of computational and graphical statistics*, vol. 5, no. 3, pp. 299–314, 1996.