Spectrum Prediction via Temporal Conditional Gaussian Random Field Model in Wideband Cognitive Radio Networks

Zhenghao Zhang[1,2], Husheng Li[1], Hannan Ma[1], Kun Zheng[1], Depeng Yang[1], and Changxing Pei[2]

[1] Department of EECS, The University of Tennessee, Knoxville TN 37996, USA
chancejack@gmail.com, husheng@eecs.utk.edu
http://www.ece.utk.edu/ husheng/
[2] Xidian University,
No.2 Taibai Road, Xi'an, China, 710071

Abstract. Wideband spectrum sensing remains an open challenge for cognitive radio networks due to the insufficient wideband sensing capability. This paper introduces the theory of Gaussian Markov Random Field to estimate the un-sensed sub-channel status. We set up a measurement system to capture the WiFi spectrum data. With the measurement data, we verify that the proposed model of Temporal Conditional Gaussian Random Field can efficient estimate the sub-channel status.

Keywords: cognitive radio networks, net components, dynamic software architecture, modeling, agents, software development approach.

1 Introduction

Cognitive radio (CR) [1][3] has been considered to be a promising technology to tackle the problem of spectrum resource scarcity. As one of the most important requirements for CR, perceiving the activities of primary users (PU) has been widely studied [9]. However, in wideband CR networks, wideband spectrum sensing still remains an open challenge [2][6]. Due to the radio frequency front-end hardware limitation, it is difficult to capture the change of spectrum with a limited sensing time in a wide spectrum range, subject to the contamination of noise. With limited wideband sampling capacity, [10] has proposed an opportunistic spectrum sensing scheme, in which the authors offered an optimal scheme to select sub-channels to sense.

Since the shared frequency band may have different division scheme for the PU network and the CR network, two adjacent sub-channels in the CR network may be affected by the same channel in the PU network. Fig. 1 illustrates the different spectrum division schemes of PU network and CR network. It is observed that, when the sub-channel 1 of the PU network is active, both sub-channels 1 and 2 of the CR network cannot be utilized. Thus, the channel usage statuses of sub-channel 1 and sub-channel 2 of the CR network are highly correlated since they are affected by the same PU sub-channel. However, the CR

X. Zhang and D. Qiao (Eds.): QShine 2010, LNICST 74, pp. 16–27, 2011.

network typically has no priori knowledge of the sub-channel division scheme of the PU network. In addition, from the temporal viewpoint, two adjacent sensing periods may also have correlation. Based on the binary time series analysis, [8] proposed a spectrum occupancy prediction algorithm where the subchannel occupancy status is modeled as time series and the model parameters are learned from previous known subchannel occupancy status using the logistic regression algorithm. However, it is assumed that the time series representing subchannel occupancy status is stationary, which is a very strong assumption. Our measurement data is verified not to satisfy this requirement. According to the correlation existing among sub-channels and sensing periods, in this paper, we propose to use the Temporal Conditional Gaussian Random Field (TCGRF) [4][7] to model the wideband spectrum activity.

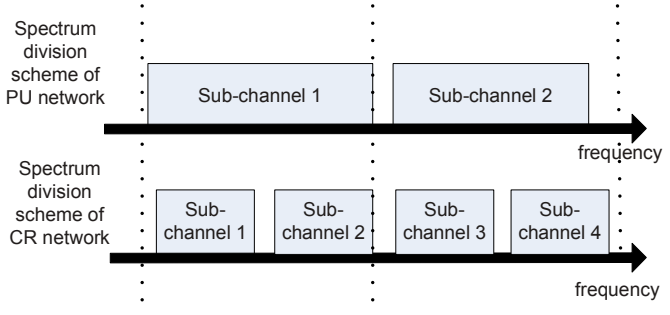

Fig. 1. An illustration of different sub-channel division scheme between PU network and CR network

As Fig. 2 illustrates, each pixel[1] represents a channel in a certain sensing period. Pixels within a certain neighborhood have correlation with each other. These pixels can be considered as the realization of a Gaussian Markov random field (GMRF). In each sensing period, not all the sub-channels are sensed. In this paper, we propose a TCGRF model to depict the sub-channel usage status within certain sensing periods. Based on the TCGRF model, a fraction of sub-channel usage status in current sensing period can be predicted without actual spectrum sensing.

The contribution of this paper is as follows. *We propose a TCGRF model to describe the sub-channel usage status, and provide the prediction of current sub-channel usage which has not been assigned for spectrum sensing. Furthermore, we set up a spectrum measurement to measure the spectrum activities in a WiFi network. The TCGRF model based prediction is verified using the measurement data.*

The remainder of the paper is organized as follows. The system model is defined in Section 2. The concept of GMRF is briefly introduced and the sub-channel usage

[1] Here we use the terminology of image processing since the 2-dimensional spectrum occupancy is similar to an image.

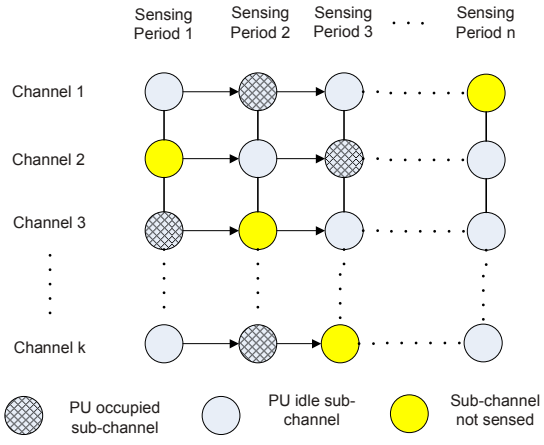

Fig. 2. An illustration of different sub-channel division schemes of PU network and CR network

status prediction based on TCGRF model is proposed in Section 3. We present the WiFi data measurement setup in Section 4 and provide the simulation results in Section 5. Conclusions and future works are provided in Section 6.

2 System Model

In this section, we introduce the system model for the wideband spectrum sensing in CR network. We assume that the CR network has different sub-channel division schemes from the PU network and the CR network does not have the priori information about how the PU network divides the entire shared frequency band. We assume that the CR network divides the shared spectrum into K sub-channels. In sensing period j, we define a sub-channel sensing set χ^j, which includes all the selected sub-channels to be sensed at sensing period j. For the i^{th} $(i \in \chi^j)$ sub-channel, the received signal is represented by $r_{i,j}(t)$. Applying the energy detection to each sub-channel, we obtain the sub-channel energy $y_{i,j}$ on the i^{th} sub-channel at sensing period j, which is given by

$$y_{i,j} = \int_{\mathcal{T}_j} (r_{i,j}(t))^2 dt, \tag{1}$$

where $\mathcal{T}_j$ denotes the sensing interval of the j^{th} sensing period. Thus, the secondary user carries out the binary hypotheses testing based on the sub-channel energy $y_{i,j}$. And we define an $K \times N$ spectrum usage indicator matrix, denoted by $\mathbf{X}$ where its element $x_{i,j}$, $(i = 1, 2, ..., K \ j = 1, 2, ..., N)$ represents the primary users' spectrum usage status on the i^{th} sub-channel in the j^{th} sensing period. The binary hypotheses testing is denoted by

$$x_{i,j} = \begin{cases} 1, & \text{if } y_{i,j} \geq \tau; \\ -1, & \text{if } y_{i,j} < \tau. \end{cases}, \tag{2}$$

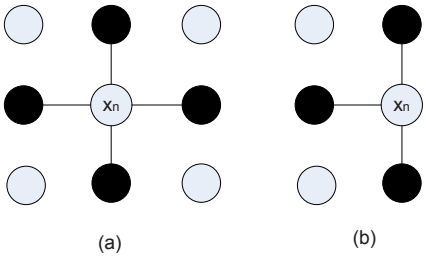

Fig. 3. Two kinds of neighborhood of pixel x_n

where τ is the energy detection threshold.

Therefore, the spectrum usage indicator matrix $\mathbf{X}$ can be considered as a binary-value image and its element $x_{i,j}$ represents the pixel in this image. The double index for each pixel $x_{i,j}$ can be transformed into a univariate index, which is given by

$$(i, j) \mapsto n = i + (j - 1)K. \tag{3}$$

Since we try to apply the GMRF, we define a neighborhood N_n for element x_n. We define two types of neighborhoods in this paper, which is illustrated in Fig. 3.

3 Gaussian Markov Random Field and Sub-channel Status Prediction by TCGRF

In this section, we briefly introduce the GMRF [5], which has been widely used in image processing, biologic information processing, etc. Then we will apply TCGRF to predict the sub-channel status.

3.1 Gaussian Markov Random Field

A GMRF is defined as a random vector $\mathbf{z} = (z_1, z_2, ..., z_n)^T \in \mathbb{R}^n$, whose conditional probability of each element z_n satisfies Markov property, which is given by

$$\pi(z_n|\mathbf{z}_{-n}) = \pi(z_n|z_i \in N_n), \tag{4}$$

where $\mathbf{z}_{-n}$ represents the set of all the elements in $\mathbf{z}$ excluding z_n. Furthermore, the probability density function (pdf) of $\mathbf{z}$ is given by

$$\begin{aligned}\pi(\mathbf{z}) &= (2\pi)^{-\frac{n}{2}}|\Sigma|^{-1/2}\exp -\frac{1}{2}(\mathbf{z}-\mu)^T\Sigma^{-1}(\mathbf{z}-\mu) \\ &= (2\pi)^{-\frac{n}{2}}|Q|^{1/2}\exp -\frac{1}{2}(\mathbf{z}-\mu)^T Q(\mathbf{z}-\mu),\end{aligned} \tag{5}$$

where μ and Σ are the expected trend and the covariance matrix of the GMRF, respectively. Due to the fact that z_i and z_j are conditionally independent for $z_i \notin$

$\mathcal{N}_j$, Σ^{-1} has a sparse representation [5]. Thus Σ^{-1} is defined as precision matrix Q for mathematical convenience. Furthermore, the elements of $\mathbf{z}$ is regarded as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Graph $\mathcal{G}$ depicts the neighborhood for each element z_n, each edge in the graph represents a pair elements in a neighborhood. And the precision matrix Q has the following property:

$$Q_{i,j} \begin{cases} \neq 0, & \text{if } i,j \text{ and } z_j \text{ is neighbor} \\ = 0, & \text{if } z_i \text{ and } z_j \text{ isn't neighbor} \end{cases}. \tag{6}$$

In this paper, we adopt the neighborhood defined in Fig. 3; hence the precision matrix Q is a sparse matrix.

3.2 Intrinsic GMRF

Given the precision matrix Q, we can attain the conditional distribution for some elements in the GMRF $\mathbf{z}$ given the remaining elements, which is given by

$$\mathbf{X}|\mathbf{Y} \sim \mathcal{N}(\mu_X - Q_{X,X}^{-1} Q_{X,Y}(Y - \mu_Y), Q_{X,X}^{-1}), \tag{7}$$

where $\mathbf{X}$ and $\mathbf{Y}$ represent the unknown elements and the known elements in the GMRF, respectively. Meanwhile, $Q_{X,X}$ and $Q_{X,Y}$ represent the inverse of the covariance matrix of $\Sigma_{X,X}$ and $\Sigma_{X,Y}$, respectively. We apply the Cholesky factorization of $Q_{X,X}$ and attain the conditional expectation of $\mathbf{X}$ given $\mathbf{Y}$, which is given by

$$E(X|Y) = \mu_X - R_{X,X}^{-1}((R_{X,X}^{-1})^T Q_{X,Y}(Y - \mu_Y)), \tag{8}$$

where

$$Q_{X,X} = R_{X,X}^T R_{X,X}. \tag{9}$$

For attaining the construction of Q according to the known data, we introduce the normal increment which is given by

$$z_i - z_j \sim \mathcal{N}(0, \kappa^{-1}). \tag{10}$$

According to the conditional independence provided by the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, we attain the intrinsic GMRF model, which is given by

$$\pi(\mathbf{z}) \propto \kappa^{(n-1)/2} \exp(-\frac{\kappa}{2} \sum_{i \sim j} (z_i - z_j)^2), \tag{11}$$

where the element pair $i \sim j$ represents an edge in the graph. We define the number of neighbors of z_i as n_i. The precision matrix Q has elements given by

$$Q_{i,j} = \begin{cases} \kappa n_i, & \text{if } i = j \\ -\kappa, & \text{if } i \sim j \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

Given the precision matrix Q, we can apply (8) to estimate the mean value of the unknown data in the intrinsic GMRF.

3.3 Temporal Conditional Gaussian Random Field

In the learning stage, we apply the neighborhood structure of (a) in Fig. 3. Suppose that we assign N_0 spectrum sensing periods for learning, which means that we try to reconstruct an $K \times N_0$ spectrum usage indicator matrix. Due to the limitation of the wideband spectrum sensing, we can obtain only a fraction of the spectrum usage information within this N_0 sensing period. We denote the k_0 known spectrum status vector by $\mathbf{z}_{k0}$ and the unknown spectrum status vector by $\mathbf{z}_{\tilde{k0}}$. We define the sensing ratio as

$$\rho = \frac{k_0}{K \times N_0}. \tag{13}$$

According to (8), we attain the learning spectrum status matrix $\mathbf{z}_0$, which is given by

$$\mathbf{z}_0 = \mathbf{z}_{k0} + \mathbf{z}_{\tilde{k0}}. \tag{14}$$

After we attain the learning spectrum status matrix $\mathbf{z}_0$, we further apply the neighborhood structure (b) in Fig. 3 to estimate the sub-channel usage status in $(N_0 + 1)$ sensing period. Implementing spectrum sensing on m sub-channels, we attain $K \times N_0 + m$ known spectrum usage status. The unknown sub-channel usage status $z_i, (i = 1, 2, ..., K - m)$ can be estimated by

$$\begin{aligned} z_i &= E(z_i | \mathbf{z}_{-i}) \\ &= -\frac{1}{Q_{ii}} \sum_{j:j\sim i} Q_{ij} x_j. \end{aligned} \tag{15}$$

Then, z_i can be added to the known set of the sub-channel usage status to estimate the unknown sub-channel status in next sensing period.

4 Measuring WiFi Spectrum

In this section, we will introduce our spectrum measurement for WiFi signal which is considered as a PU. Fig. 4 shows the experiment setup. A photo of the equipment is shown in Fig. 5. The measurement is carried out inside or outside the Ferris Hall of the University of Tennessee.

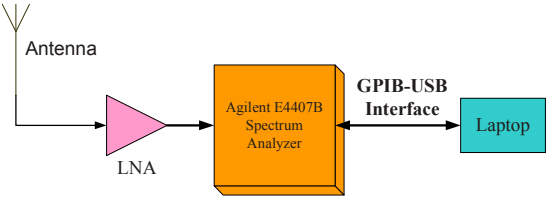

Fig. 4. An illustration of the setup for measuring WiFi spectrum signal

Fig. 5. A photo of the equipment for measuring the WiFi spectrum signal

The key instrument is the Agilent E4407B (9kHz to 26.5GHz). Through the GPIB-USB cable, the laptop can control the spectrum measurement by setting the configurations. The GPIB-USB communication bus is fast and instant. Hence the spectrum activity data can be quickly updated from the spectrum analyzer to the laptop. Also through the GPIB-USB, the spectrum measurement is handled by Matlab, which provides convenience and flexibility for realtime operation and online data storage.

The antenna in the measurement system is the Vivaldi wideband antenna. The frequency range of the antenna is from 1GHz-4GHz with the return loss lower than -10dB. The wideband antenna provides more flexibility for sensing different bands than a narrow band one. Although the antenna is not omni-directional, the testing environment is on campus where dense buildings, trees and other objects make the far-field electrical-magnetic wave reflect and refract. Hence the antenna has a capability to sense the spectrum activities from all directions.

Note that, in order to improve the sensitivity of the experiment, we add an external LNA. Generally, modern complicated spectrum analyzer is able to sense a very wide band spectrum. However, the wideband spectrum sensing incurs a lot of thermal noise, which buries the weak signal. Therefore, we utilize an external LNA close to the antenna to further lower the noise floor for sensing the weak signal activity. The LNA has a 20dB gain with 1.7dB Noise Figure (NF). The LNA can amplify the signal from 1GHz-3GHz.

Since we focus on the WiFi band, we set frequency measurement range from $2.4GHz$ to $2.5GHz$, applying 1024 points sampling for each spectrum measurement. Each sensing period is $20ms$ including spectrum measurement and data recording.

After we attained the spectrum data, an energy detection threshold is set to $-65dBm$ and $-70dBm$ for outdoor and indoor measurements, respectively. We divide the whole spectrum measurement band into 16 equal bandwidth sub-channels. Thus, if the energy of a sub-channel exceeds the threshold, we decide that this sub-channel is currently occupied by primary users, which is represented by a black strip in a time-frequency chart. If the energy of a sub-channel is

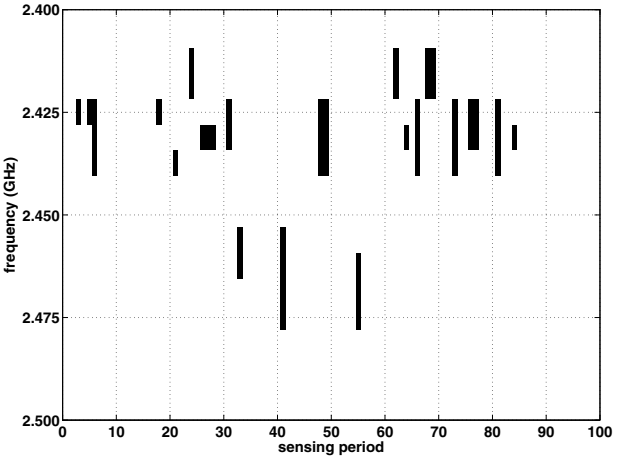

Fig. 6. Time-frequency chart of WiFi spectrum measurement in an indoor environment

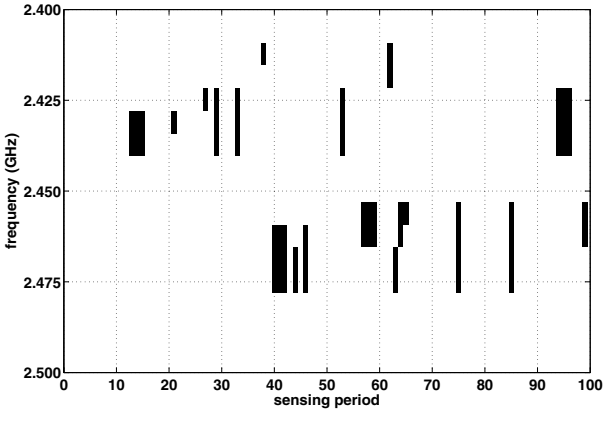

Fig. 7. Time-frequency chart of WiFi spectrum measurement in an outdoor environment

smaller than the threshold, we decide that this sub-channel is idle, represented by a white strip. Fig. 6 illustrates one of the spectrum measurement in an indoor environment and Fig. 7 illustrates the case in an outdoor environment.

5 Simulation Results

In this section, we demonstrate the performance of sub-channel usage status prediction by comparing the true spectrum activity measured in the WiFi frequency band. According to the neighborhood structure depicted in Fig. 3, we

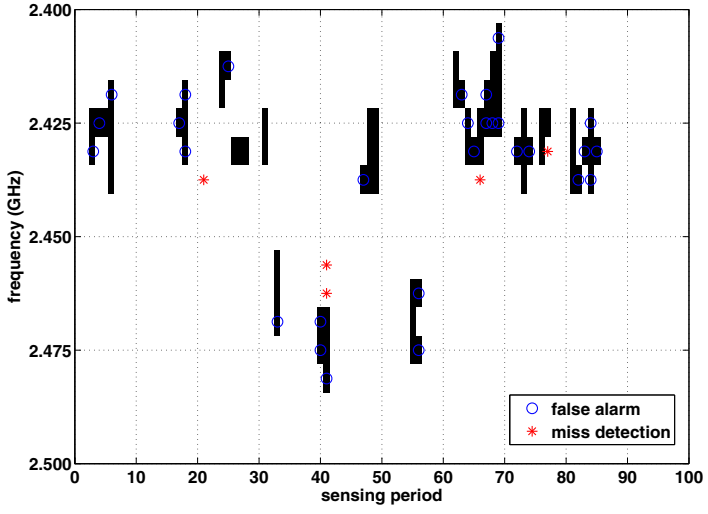

Fig. 8. Predicted time-frequency chart of WiFi spectrum usage status in an indoor environment

set the sensing ratio $\rho = 0.6$ and apply the TCGRF algorithm to estimate the status of the unsensed sub-channels. Fig. 8 and Fig. 9 demonstrate the estimated time-frequency chart of the WiFi spectrum usage status in indoor and outdoor environments, respectively.

It can be observed that the false alarm[2] probability of an outdoor measurement is higher than that of indoor environment. This may be due to the higher background noise level of an outdoor environment.

Fig. 10 illustrates the prediction error rate along with the sensing period. We apply an adaptive learning scheme, when the prediction error rate exceeds 0.02, we will implement the learning procedure according to the neighborhood structure of Fig. 3 (a). The spikes in the graph denote the sensing period when the prediction error exceeds 0.02; hence a learning procedure will be carried out, and the prediction error will be controlled within the threshold again.

We introduce the receiver operating characteristic (ROC) curve to demonstrate the detection performance of various spectrum occupancy prediction algorithms. We compare the proposed TCGRF algorithm with the logistic regression algorithm [8] and a traditional image filter algorithm. The logistic regression algorithm assumes that the subchannel occupancy is a stationary time series and the logistic regression model parameter can be learned from the previous subchannel occupancy status. However, in our practical spectrum measurements for

[2] Here we define the false alarm as the event that the spectrum sensor claims that an idle sub-channel is detected while the sub-channel is actually occupied by primary users.

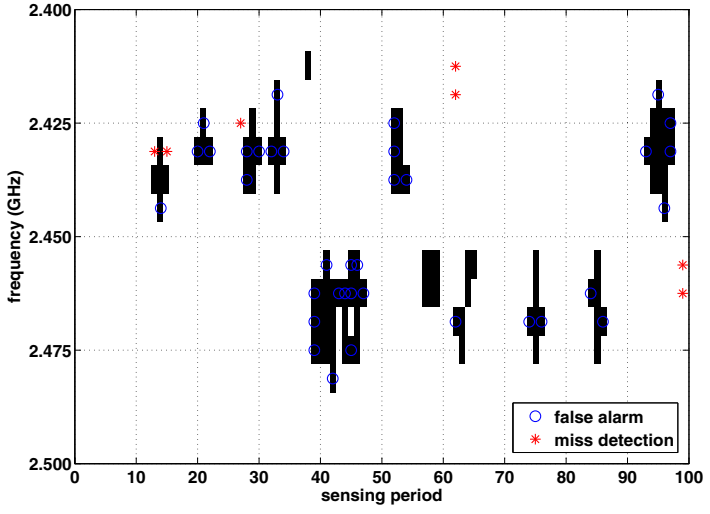

Fig. 9. Predicted time-frequency chart of WiFi spectrum usage status in an outdoor environment

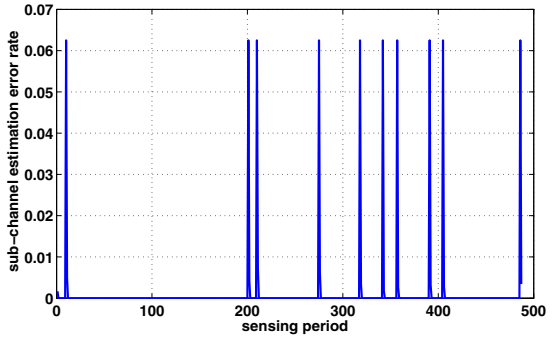

Fig. 10. Sub-channel status prediction error rate along with different sensing period

WiFi band, we found that it is difficult to meet this assumption. Consequently, the subchannel occupancy prediction performance of logistic regression degrades severely, which is illustrated in the following simulation results. Fig. 11 and Fig. 12 show the ROC curves of the indoor environment measurement and the outdoor environment, respectively.

The probability of false alarm is defined as the probability of the subchannel is idle but report busy. And the probability of detection is defined as the probability of the busy subchannels are successfully detected. It can observed that, in the low false alarm area, TCGRF based spectrum detection performance is significantly outperform the logistic algorithm and the image filter method. Under the same

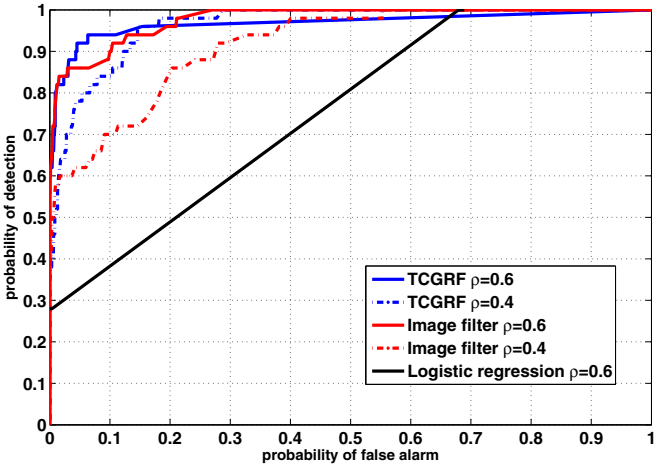

Fig. 11. ROC curve with various sensing ratios in an indoor environment

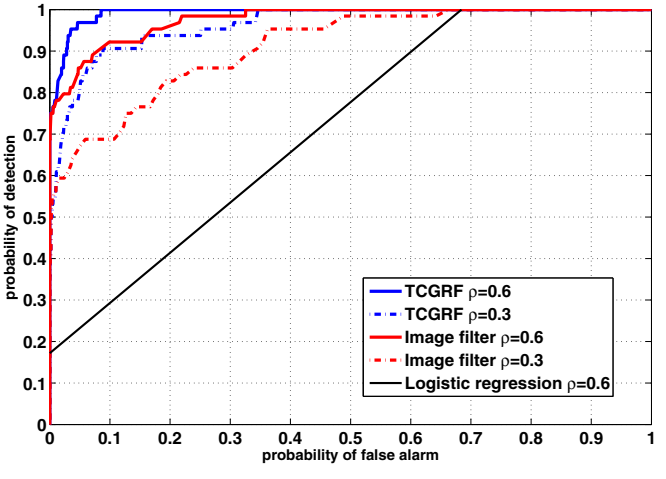

Fig. 12. ROC curve with various sensing ratios in an outdoor environment

probability of false alarm and sensing ratio ρ, the TCGRF based method can achieve substantially higher probability of detection. It is noticed that, when the sensing rate $\rho = 0.6$, the probability of detection can is higher than 90% while the probability of false alarm is only 0.1. This implies that the TCGRF algorithm can offer an efficient sub-channel status prediction. The insight of the TCGRF algorithm is that it utilizes the temporal correlation of the the sub-channel status between two adjacent sensing periods.

6 Conclusion

In this paper, we have used the theory of GMRF to predict the sub-channel usage status when not all sub-channel statuses can be sensed due to the limited wideband spectrum sensing capability. We have proposed a TCGRF model to predict the statuses of the unsensed sub-channels. In addition, we have carried out a measurement experiment over the WiFi band. The practical spectrum activity measured in the WiFi band is used to verify the TCGRF model. By applying the correlation between two adjacent sensing periods, the TCGRF algorithm provides an efficient sub-channel status prediction method, which can be used for channel selection in wideband spectrum sensing.

Acknowledgements. This work is supported by National Science Foundation under grants CCF-0830451 and ECCS-0901425, the "973" program of China under Grant 2007CB311201, China scholarship Council (CSC), National Natural Science Foundation of China (61072067, 60970119), and State Key Laboratory of Integrated Service Networks Project Grant (ISN02080002).

References

1. Akyildiz, I., Lee, W., Vuran, M., Mohanty, S.: Next generation dynamic spectrum access/cognitive radio wireless networks: A survey. Elsevier Computer Networks 50, 2127–2159 (2006)
2. Cabric, D., Mishra, S.M., Brodersen, R.W.: Implementation issues in spectrum sensing for Cognitive Radio. In: Proc. of Asilomar Conference on Signals, Systems, and Computers (2004)
3. Haykin, S.: Cognitive radio: Brain-empowered wireless communications. IEEE J. Sel. Areas Commun. 23(2), 201–220 (2005)
4. Lafferty, J., McCallum, A., Pereira, F.: Conditional random Fields: Probabilistic models for segmenting and labeling sequence data. In: In Proc. ICML, pp. 282–289 (2001)
5. Rue, H., Held, L.: Gaussian Markov random fields: theory and applications. Chapman & Hall, Boca Raton (2005)
6. Tang, H.: Some physical layer issues of wide-band cognitive radio systems. In: Proc. IEEE Dynamic Spectrum Access Networks (DySPAN), pp. 151–159 (November 2005)
7. Winkler, G.: Image Analysis, Random Fields and Markov Chain Mone Carlo Methods. Springer, Heidelberg (2003)
8. Yarkan, S., Arslan, H.: Binary time series approach to spectrum prediction for cognitive radio. In: IEEE 66th Vehicular Technology Conference, VTC 2007 Fall, pp. 1563–1567 (2007)
9. Zhao, Q., Sadler, B.: A survey of dynamic spectrum access. IEEE Signal Processing Mag. 24, 79–89 (2007)
10. Zhao, Q., Krishnamachari, B., Liu, K.: On myopic sensing for multi-channel opportunistic access: Structure, optimality, and performance. IEEE Trans. Wireless Communications 7, 5431–5440 (2008)