Classification of Wireless Interference on 2.4GHz Spectrum

Weng, Z.; Orlik, P.; Kim, K.J.

TR2014-018 April 2014

Abstract

We propose two methods for the detection of RF interference. The first one is for the detection of the interferences from microwave ovens, and the second one is for Wi-Fi and Bluetooth signals. The motivation of this work is to design a system for reliable wireless communication. Specifically, the systems equipped with interference detectors will be able to choose the appropriate time intervals to transmit signals in the presence of other interferences, therefore avoid unnecessary collisions and retransmissions.

IEEE Wireless Communications and Networking Conference (WCNC)

Classification of Wireless Interference on 2.4GHz Spectrum

Zhiyuan Weng
Department of Electrical and Computer Engineering
Stony Brook University
Stony Brook, NY 11790, USA
Email: zhiyuan.weng@stonybrook.edu

Philip Orlik,Kyeong Jin Kim
Mitsubishi Electric Research Laboratories (MERL)
Cambridge, MA, 02139
Email: {porlik,kkim}@merl.com

***Abstract*—We[1] propose two methods for the detection of RF interference. The first one is for the detection of the interferences from microwave ovens, and the second one is for Wi-Fi and Bluetooth signals. The motivation of this work is to design a system for reliable wireless communication. Specifically, the systems equipped with interference detectors will be able to choose the appropriate time intervals to transmit signals in the presence of other interferences, therefore avoid unnecessary collisions and retransmissions.**

***Index Terms*—Bluetooth, hidden Markov model, microwave Oven, RF interference, signal classification, spectrogram, Wi-Fi.**

I. Introduction

There are a large number of devices operating on the industrial, scientific and medical (ISM) radio bands. To avoid collisions and wasting time on retransmissions, the system should be able to sense the spectrum for possible interferences and exploit the spectrum in an intelligent way. The first step toward achieving this objective is to detect and classify the interference present in the environment, which is the focus of this work. We note also that we seek low-complexity algorithms for our interference classification task as we envision these algorithms will operate on resource constrained wireless local area network (WLAN) hardware.

There is a large number of existing works related to the classification of wireless interferences. Some commercial systems include Spectrum XT [1], AirMaestro [2], and CleanAir [3]. These solutions utilize special customized hardware together with sophisticated software to detect interference. Many signal classification works deal with the classification of modulation type, e.g., AM, FM, QPSK, BPSK, FSK and MSK. Cyclostationary spectrum detection is the most common method for modulation detection and classification [4], [5]. Additionally, moments-based algorithms have been explored for modulation detection [6], [7]. The use of neural networks with traditional methods has also been proposed [8], [9]. Works that are closely related to this paper include [10]–[12]. In [10], the authors study the impact of interferences from Wi-Fi and microwave ovens (MWO) on wireless sensor networks. They use spectral signatures to identify the interferers. Since they have multiple sensors in the network, it is possible to measure the spectrum power on several different channels. Then the correlation between channels is exploited for interference detection. In [12], AirShark, a system that detects multiple non-Wi-Fi RF devices, is proposed. The authors use a commercial Wi-Fi card for the detection. However, its classification target set is very large. They extract a set of generic features, which include frequency, bandwidth, spectral signature, duty cycle, pulse signature, inter-pulse timing signature, pulse spread and device specific features like sweeps. Then a decision tree-based mechanism has been developed. The performance of the system is claimed to be comparable to that of commercial signal analyzers.

In this work, we focus on the development of simple yet efficient algorithms. We only assume that the system is equipped with a simple spectrum analyzer, like the on-chip spectrum analyzer embedded in the Atheros AR9280 chip. The resolution of the spectrum analyzer is less than 128 frequency bins. Also, we confine ourselves to simple algorithms that do not involve heavy computation. The classification set includes Wi-Fi, Bluetooth, MWO interference and noise as these are the most common interference sources. We process the data on the spectrogram. Although it is equally possible to do detection on time domain using the IQ baseband time series, we believe that it is easier for both human beings

[1] Zhiyuan Weng's work was done while he was an intern in MERL.

and machines to distinguish different interferences by the look of the signals on the spectrogram. To detect MWO interference, we use its 60Hz periodicity feature, which is the most distinguishing feature. The exact shapes of the signals on the spectrogram may vary depending on the maker, the model and the power of the MWO. However, the 60Hz periodicity does not change as long as the frequency of the AC power does not change. We introduce another method to detect Wi-Fi or Bluetooth signals, using the fact that they have different bandwidths. Wi-Fi signals are wideband (20MHz or 40MHz) and Bluetooth signals are narrowband (1MHz). Specifically, we use templates to measure the bandwidth of the signals. By doing so, we can reduce the high-dimensional space down to a two-dimensional space. Then we use Hidden Markov Model (HMM) to model the sequence. The Expectation Maximization (EM) algorithm is used to learn the parameters of the HMM. After the HMM is learned, we can use the model to classify Wi-Fi and Bluetooth signals.

The paper is organized as follows. In Section 2, we introduce the algorithm for the detection of MWO interferences. In Section 3, we discuss a method for the classification of Wi-Fi and Bluetooth signals. In Section 4, we present our experimental results. We conclude our work in Section 5.

II. Classification of Microwave Oven Interference

Radiation leakage from commercial MWOs is one of the most significant interference sources in the 2.4GHz spectrum. A typical spectrogram of MWO interference is shown in Fig. 1, where red and green color indicate large and small amplitude, respectively. The two curved lines pointed to by blue arrows are the MWO interferences. The challenging part is that the frequencies, the shapes and the duty cycles all depend on the make and the model of the MWO. The only two features that are invariant with respect to different models are the 60Hz periodicity and the narrowband property. Note that some references might describe MWO interference as wideband. It depends on how we look at them. It is called wideband because the signature on the spectrogram sweeps over a wide spectrum. However, if we look at a specific slice on the spectrogram, it occupies narrow space. There are also ”transients” at the beginning and end of each frequency sweep which tend to be much wider band, however, we chose to focus on the detection of the narrowband sweep portion of the MWO signal. The periodicity of 60Hz is due to the frequency of the AC power. We use the two features to detect MWO interferences. Basically, our algorithm checks whether there is a narrowband signature that appears periodically with the frequency of 60Hz.

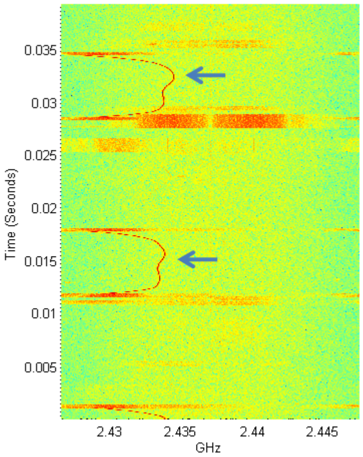

Fig. 1. A typical MWO spectrogram.

A. Proposed algorithm

In this subsection, we describe our algorithm in detail. It consists of two parts. In the first part, we calculate a value that reflects the bandwidth of the signal. In the second part, we check if similar patterns appear with the frequency of 60Hz. The steps of the algorithm are as follows.

- Step 1: Take a J-sample subsequence from every M-samples, where M determines how often we sense the spectrum. Then we calculate the FFT of the subsequence. Denote the magnitude of the FFT samples by $|Z_{j,n}|$ the j-th element of the n-th FFT vector generated by the n-th subsequence. We normalize $Z_{j,n}$ by the sum:

$$X_{j,n} = |Z_{j,n}| / \Big(\sum_{l=1}^{J} |Z_{l,n}| \Big). \tag{1}$$

  The first step is summarized in Fig. 2.
- Step 2: We calculate $y(n)$ from $X_{j,n}$ as follows:

$$y(n) = \sum_{j=1}^{J} (X_{j,n} - X_{\text{threshold}})_{+} \tag{2}$$

  where operator $(x)_+$ is defined as

$$(x)_+ = \begin{cases} x & x > 0 \\ 0 & x \leq 0 \end{cases} . \tag{3}$$

  For example, in Fig. 3, the value of $y(n)$ is equal to the length of the dashed line segments summed together. The value of $y(n)$ tells us how narrow the bandwidth is. To see this, note that the slice is normalized. If the bandwidth is wide, it spreads out and the values become smaller after normalization. Few of them would be larger than a threshold. On the other hand, if it is narrowband, the values are concentrated. Some of the bins would be larger than the threshold after normalization. Therefore, it is not difficult to see that the larger the value, the narrower the signal is.

- Step 3: We calculate auto-correlation of the sequence $y(n)$ as,

$$r(n) = E[y(m)y(m+n)]. \tag{4}$$

- Step 4: We check the position of the maximum value of $r(n)$, of course, $r(n)$ will have its maximum value at $r(0)$. However, since $r(n)$ derived from a cyclostationary process we expect that it is also periodic when MWO intference is present in the measurement data. Thus we need only look for a maximum in specific interval, say $[a, b]$ that includes the location of the expected maximum, and since $r(n)$ is periodic we need to exclude the peaks at $r(0)$ and higher order harmonics. We check if a maximum value appears between interval $[a, b]$, where $[a, b]$ itself is contained within an interval $[n_1, n_2]$ which excludes higher order harmonics. The intervals are shown in Fig. 4. If the peak lies in $[a, b]$, we decide it is MWO interference, The computation of the expected period depends on the frequency of the AC power and the value of M and is discussed in further detail below.

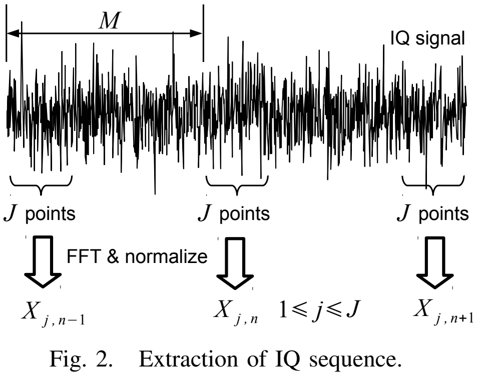

Fig. 2. Extraction of IQ sequence.

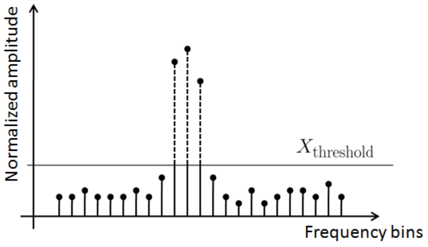

Fig. 3. Calculation of $y(n)$.

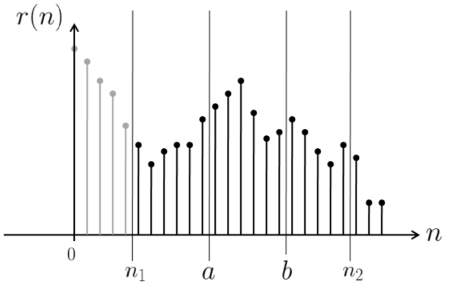

Fig. 4. Check the location of the peak of $r(n)$.

B. Choices of parameters

In this subsection, we discuss how to choose the appropriate values for the parameters in the algorithm.

1) Determine the threshold $X_{threshold}$: It is obvious that the range of $y(n)$ is from 0 to 1. Ideally we would like $y(n)$ to be the value that can help us to differentiate different bandwidth signals. Naturally we would like $y(n)$ to be zero when there is only noise and Wi-Fi traffic on the channel. We can see that the larger the $X_{\text{threshold}}$, the more likely that $y(n)$ becomes zero. On the other hand, we shall also make $y(n)$ large when narrowband signals appear. In the noise-only or WLAN traffic case, it is unlikely that the normalized FFT values $X_{j,n}$ have a large peak. If we assume that the $X_{j,n}$ is uniformly distributed on $[0, u_{max}]$ when no MWO interference is present, we can derive the distribution of $y(n)$. Note that the value of u_{max} does not affect the distribution since there is a normalization step. The distribution of the statistics $y(n)$ is equivalent to the probability of covering a circle with random arcs [13], which can be expressed as

$$\begin{aligned} P(y < Y) =& 1 - \sum_{l=1}^{J} \sum_{k=0}^{l-1} (-1)^{k+l+1} \binom{J}{l} \binom{l-1}{k} \binom{J-1}{k} \\ & \cdot Y^k (1 - lX_{\text{threshold}} - Y)_+^{J-k-1}. \end{aligned} \tag{5}$$

We would like to emphasize that for noise-only case, the value of $y(n)$ does not depend on the power of the noise due to the normalization effect. The above analysis applies to noise-only cases. If there is interference, the effect of the threshold depends on the signal to noise ratio (SNR). In the higher SNR, we can expect $y(n)$ be able to correctly reflect the bandwidth of the interference. However, at lower SNRs, the interference, even if it is narrowband, is likely to be buried in the noise, which makes $y(n)$ close to zero. Therefore, it is difficult to determine the threshold based on SNR.

The calculation of $y(n)$ itself can be viewed as a detection problem, with the output being soft values. If we make decision by comparing the threshold and the value of $y(n)$, it becomes a standard hypothesis testing problem. In hypothesis testing, if the probability of detection is hard to derive, one can simply choose the threshold according to the false alarm ratio. Similarly in our problem, we can select $X_{\text{threshold}}$ simply based on the noise-only case by using (5).

In Fig. 5, we plot the cumulative distribution function (CDF) of $y(n)$ in noise-only case as expressed in (5) for $J = 64$ with different values of $X_{\text{threshold}}$. We can see that when $X_{\text{threshold}} \geq 0.1$, $y(n)$ would be very close to zero with high probability.

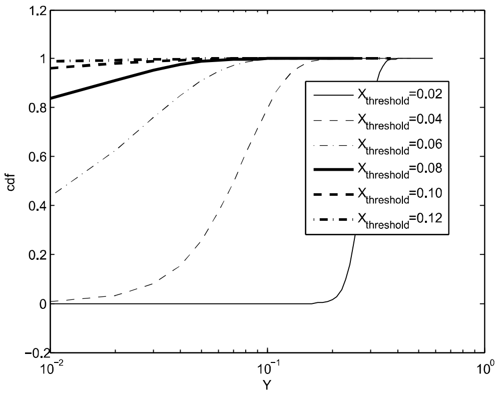

Fig. 5. CDF of y for different values of $X_{\text{threshold}}$.

2) Determine the intervals $[a, b]$ *and* $[n_1, n_2]$*:* The periodicity of the MWO interferences is due to the AC power. Given the frequency of the AC power f_{ac}, we can calculate the period of the MWO interference, say T_0, in terms of samples as follows:

$$T_0 = \frac{f_{\text{s}}}{f_{\text{ac}}} \cdot \frac{1}{M}. \tag{6}$$

Note that T_0 is not necessary an integer. We shall choose integer values of a, b such that T_0 lies between $[a, b]$. Also we know that $r(0)$ must be the maximum, which we should exclude in the range. We should also exclude $r(2T_0)$ since any integer multiple of T_0 also corresponds to a peak in $r(n)$. The purpose of interval $[n_1, n_2]$ is to exclude both $r(2T_0)$ and $r(0)$.

III. Classification of Wi-Fi and Bluetooth

In this section, we introduce an algorithm to detect Wi-Fi and Bluetooth signals. The main feature we use for Wi-Fi and Bluetooth classification is the bandwidth information. In the left subfigure in Fig. 7, both Wi-Fi and Bluetooth signals are present. We can see that the major difference is that the bandwidth of Wi-Fi signals is wider than that of Bluetooth signals. Obviously we should use the difference in bandwidth to classify the two signals. This is the motivation for the preprocessing.

A. Preprocessing

To extract the bandwidth information, we use two templates to filter the spectrogram. One is for Wi-Fi signals; the other is for Bluetooth signals. The shape of the templates is shown in Fig. 6. We use three parameters E, W, and L to specify each template. The underlying principle is similar to matched filters used in communication receivers. When the same shape appears, the response to the filter reaches a peak. The parameter, L, is the length in time domain. L should be larger than the minimum length of the signal we expect on the spectrogram. The same rule applies for E and W. The number of -1s, E, controls how thin the mainlobe of the template response is. The more -1s you put on the edge of the template, the thinner the shape of the response. The effects of the filtering on the filtered spectrogram can be seen in Fig. 7. The first left figure is the original spectrogram. The second left one is the filtered spectrogram by the template for Wi-Fi signal. The second right one is the filtered spectrogram by Bluetooth template.

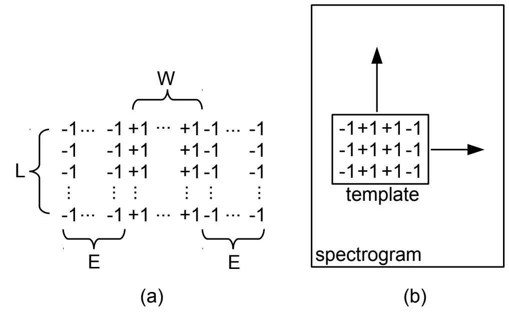

Fig. 6. Illustration of the template and filtering.

B. Dimensionality reduction

Before we reduce the dimensions, our state space has J dimensions. However, we do not care about what happens on a particular frequency bin. We care about what happens on a particular timeslot. Therefore, we can take the maximum value in each time slot for the two filtered spectrograms. By doing so, we effectively reduce J dimensions to two dimensions. This can be seen in the right subfigure in Fig.7.

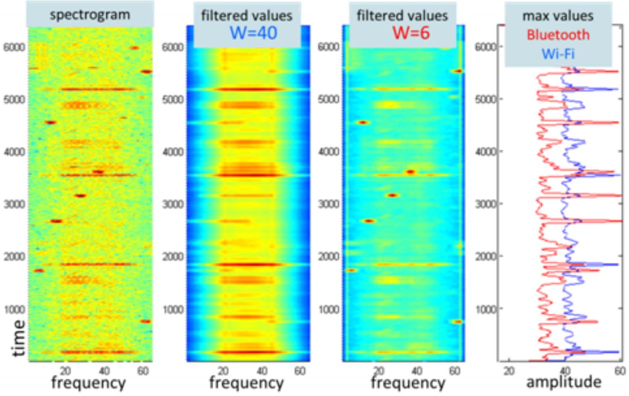

Fig. 7. The effect of 2D filtering and dimension reduction.

C. Hidden Markov Model

A hidden Markov model (HMM) is a statistical Markov model in which the system is assumed to have the Markov property with hidden states [14]. Denote by s_t and $\mathbf{y}_t$ the unknown state of the system and the observation, respectively, at time t. Note that, $\mathbf{y}_t$ depends only on s_t, and s_t depends on s_{t-1}. In our model, $s_t \in \{S_1, S_2, S_3\}$, where S_1,S_2, and S_3, respectively, stand for noise, Wi-Fi, and Bluetooth. Given the sequence of observation $\mathbf{y}_0, \mathbf{y}_1, \cdots, \mathbf{y}_T$, we would like to infer the states of the system $s_0, s_1, \cdots, s_T$. In our case, $\mathbf{y}_t \in \mathbb{R}^2$. The probabilistic graph can be represented in Fig. 8. The dependency between states is given by $P(s_n = S_j | s_{n-1} = S_i) = a_{ij}$, for $i, j \in \{1, 2, 3\}$. A matrix A with its elements $(a_{i,j})$ is called a transition matrix. We assume that noise is Gaussian, i.e., the dependency between s_t and $\mathbf{y}_t$ is given by

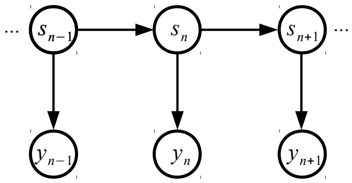

Fig. 8. The probabilistic graph for hidden Markov model.

$$p(\mathbf{y}_t | s_t = S_i) = \frac{1}{\sqrt{(2\pi)^2 |\boldsymbol{\Sigma}_i|}} \exp\left(-\frac{1}{2}(\mathbf{y}_t - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{y}_t - \boldsymbol{\mu}_i)\right) \quad (7)$$

where $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ are the mean and the covariance matrix of the normal distribution associated with state S_i. The hidden Markov model has the following parameters (1) The transition matrix A. (2) The mean $\boldsymbol{\mu}_i$ and the covariance matrix $\boldsymbol{\Sigma}_i$ for $i \in \{1, 2, 3\}$. (3) The prior for each state $P(S_i)$ for $i \in \{1, 2, 3\}$. Given a dataset, these parameters can be trained using the Expectation Maximization (EM) algorithm. After the model is trained, we can use the trained model to classify a given sequence.

IV. Experiment

In our experiment, we collect wireless signal using a special device and then process the data on computers. We use a ThinkRF WSA4000 receiver to collect IQ baseband signals at the MERL office. Although it can sample the signal at a rate up to 125MHz, we only use 20.83MHz[2] sampling rate to make it consistent with the device on which we would be implementing the algorithms. The center frequency is set to be 2.437GHz, which is the center of Channel 6 according to IEEE 802.11b/g/n standard.

A. Detection of MWO interference

We placed the WSA4000 receiver at an office that was approximately 10 meters from a commercial MWO. We collected 1154 sequences a total of 53 had MWO interference present. Note that Wi-Fi signals are commonly present due to the office environment. Therefore, most of the time we are in fact detecting MWO signatures buried in Wi-Fi signals rather than in pure noises. Each sequence has 4.096×10^6 samples, duration of 0.1966 seconds.

We have tested various values of $M = 640, 1280, 1920$ and 2560. Let $X_{\text{threshold}} = 0.1$, $J = 64$. Using (6), we have obtained the period T_0 for each values of M. Based on the values of T_0, we selected the corresponding n_1, n_2, a, and b that give the best performance. For all the cases, the probability of detection (PD) and the false alarm ratio (FA) are shown in Table 1. We can see that the performance is insensitive to value of M. We still achieve reasonalbe performance for as large as $M = 2560$.

B. Detection of Wi-Fi and Bluetooth signals

As noted previously, Wi-Fi signals are always in the air. To generate Bluetooth signals, we have streamed music between two bluetooth devices. We placed the interferer and the receiver at two corners of an office room, where we collected 40 sequences for training. We use the following parameters for the templates: L=45, W=40, E=2, for Wi-Fi signals; L=45, W=6, E=1, for Bluetooth signals. We have reduced the dimensions to two and plotted them on the plane as shown in Fig. 9. Afterwards, we used the EM algorithm to train the

[2] We set the decimation rate to be six.

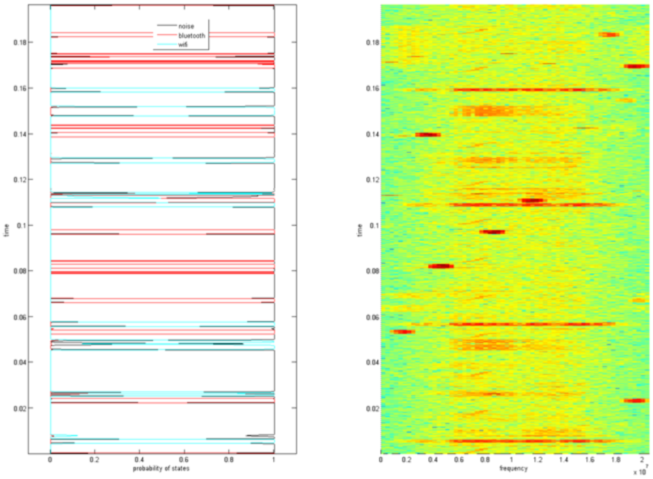

Fig. 10. The result of the classification.

M	640	1280	1920	2560
T_0	542.53	271.26	180.84	135.63
n_1	100	50	50	30
n_2	700	500	300	230
a	540	265	178	133
b	545	276	183	138
PD	48/53	48/53	47/53	47/53
FA	24/1154	33/1154	30/1154	30/1154

TABLE I
THE RESULT OF THE DETECTION OF MWO INTERFERENCES FOR DIFFERENT VALUES OF M.

HMM. We have plotted $\boldsymbol{\mu}_i$ and $\boldsymbol{\Gamma}_i$ in Fig. 9 as well. We can see that those points can be roughly divided into three clusters. Also, for HMM, the dependencies between two consecutive states have been considered.

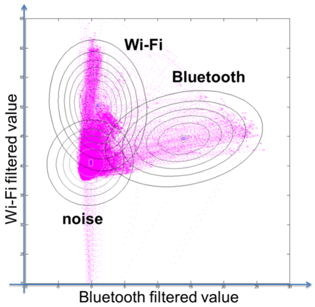

Fig. 9. The training result of the HMM.

After the model has been trained, we can use the trained model to classify signals. Fig. 10 shows the classification result of a specific data sequence. The subfigure on the right is the spectrogram of a collected data sequence. The one on the left is the classification result. X axis is the probability of the state, and Y axis is time. Red line indicates Bluetooth signals; Cyan line indicates the Wi-Fi signals; Black line indicates noises. The two subfigures are aligned in time domain, so that we can easily compare the result. We can see that Wi-Fi and Bluetooth signals are correctly classified.

V. Conclusion

We have proposed two algorithms. The first algorithm is for the detection of MWO interference. The features we exploit are the narrowband width and the periodicity. The second algorithm is used to detect Wi-Fi and Bluetooth signals. We differentiate them by their bandwidths and smooth the result using the hidden Markov model. We have shown that they work well for real data collected in the office building.

References

[1] A. SpectrumXT, “http://www.flukenetworks.com/,” 2013.

[2] B. A. S. A. Solutions, “http://www.bandspeed.com/,” 2013.

[3] C. C. Technology, “http://www.cisco.com/,” 2013.

[4] K. Kim, I. A. Akbar, K. K. Bae, J.-S. Urn, C. M. Spooner, and J. H. Reed, “Cyclostationary approaches to signal detection and classification in cognitive radio,” in *IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 212–215, 2007.

[5] J. Chen, A. Gibson, and J. Zafar, “Cyclostationary spectrum detection in cognitive radios,” in *IET Seminar on Cognitive Radio and Software Defined Radios: Technologies and Techniques*, pp. 1–5, 2008.

[6] B. Wang and L. Ge, “A novel algorithm for identification of OFDM signal,” in *Proc. Int. Conf. on Wireless Commun., Networking and Mobile Computing*, vol. 1, pp. 261–264, 2005.

[7] S. S. Soliman and S. Z. Hsue, “Signal classification using statistical moments,” *IEEE Trans. on Commun.*, vol. 40, no. 5, pp. 908–916, 1992.

[8] A. Fehske, J. Gaeddert, and J. H. Reed, “A new approach to signal classification using spectral correlation and neural networks,” in *IEEE Int. Symp. on New Frontiers in Dynamic Spectrum Access Networks*, pp. 144–150, 2005.

[9] A. K. Nandi and E. E. Azzouz, “Algorithms for automatic modulation recognition of communication signals,” *IEEE Trans. on Commun.*, vol. 46, no. 4, pp. 431–436, 1998.

[10] K. R. Chowdhury and I. F. Akyildiz, “Interferer classification, channel selection and transmission adaptation for wireless sensor networks,” in *IEEE Int. Conf. on Commun.*, pp. 1–5, 2009.

[11] B. Bloessl, S. Joerer, N. Nordin, C. Sommer, and F. Dressler, “SaFIC: A spectrum analysis framework for interferer classification in the 2.4 GHz band,” in *IEEE INFOCOM*.

[12] S. Rayanchu, A. Patro, and S. Banerjee, “Airshark: detecting non-WiFi RF devices using commodity WiFi hardware,” in *Proc. 2011 ACM SIGCOMM conf. on Internet measurement conference*, pp. 137–154, 2011.

[13] A. F. Siegel, “Random arcs on the circle,” *Journal of Applied Probability*, pp. 774–789, 1978.

[14] C. M. Bishop, *Pattern recognition and machine learning*. Springer New York, 2006.