



Multi Feature Based Classifier for Spectrum Sensing in Cognitive Radio

V.C.S Kaushik^{1*}, S. Kolangiammal², B.E Manoj Kumar³

^{1, 2, 3}Department of Electronics and Communications Engineering
SRM Institute of Science and Technology, Chennai, India

*Corresponding Author E-mail: ¹vcaskaushik1@gmail.com, ²kolangiammal.s@ktr.srmuniv.ac.in
³manojkumarbe.1997@gmail.com

Abstract

Cognitive Radio (CR) is an important technology which can enable the implementation of Dynamic Spectrum Access, which is a paradigm shift from the static spectrum access model. It is an intelligent wireless communication system which can sense the environment and can take decisions to effectively use the available radio resource without creating any interference to the Licensed Primary Users. Hence sensing of the spectrum plays a very important role in the effective implementation of this technology. We propose a new spectrum sensing algorithm in this paper which is based on machine learning and uses a Multi Feature based Classifier (MFC) model for classification of the spectrum.

Keywords: Dynamic Spectrum Access, Cognitive Radio, Machine Learning, Spectrum Sensing, Multi Feature Based Classifier

1. Introduction

The vast number of wireless devices and technologies has increased the load on the available wireless spectrum. In the upcoming years we can definitely expect newer technologies which will bring in more devices which will utilize the spectrum. Thus the RF spectrum is bound to become much more crowded than it is today. While research is going on to explore newer spectrums for communications, it is also important that we ensure that the available spectrum is used efficiently. Therefore, spectrum sensing is becoming increasingly important for both the existing systems and future ones to identify the underused spectrum and operate reliably and efficiently. The CR is an intelligent wireless communication technology which is aware of the environment it is in and can learn and change its behavior to ensure efficient utilization of the spectrum.

In CR there are two types of users namely the primary users (PU) and the secondary users (SU). The PU's are those users who are licensed to use a particular part of the spectrum, and the SU's are the unlicensed users. Hence the PU's are given a higher priority and only when the PU's are inactive the SU's can utilize the spectrum. Hence the SU's need to have the ability to correctly sense the spectrum and abstain from causing any interference to the PU.

2. System Model and Basic Concepts

In this project we are opting for Support Vector Machine (SVM) as the classifier and use the Multiple Feature based Classifier model for classification. In order to generate data sets, which consists of signals for both testing and training we opt for using Rayleigh fading and Rician Fading models.

These generated signals will then be processed for extracting multiple feature vectors which are i) Energy vector ii) Eigen values of covariance matrix iii) Ratio of sum of non-diagonal elements to sum of diagonal elements of covariance matrix iv) Eigen vector corresponding to the maximum Eigen value of covariance matrix. These feature vectors will then be used in MFC model in which every feature vector is used to train a single SVM model. There are two decision making models which have been used in this project, one being hard decision making model and another being SVM based decision making model. These models then determine the final decision.

A. Multi Feature Based Classifier

A conventional classifier uses all the N classes of features which describe the data, together as input to the trainer and classifier. When N sets of features which consist of distinct individual features are combined and used, the distinct features generally get assigned to the same maximum and minimum magnitudes (typically 0 and 1). Because of this unnatural treatment of the features for training and the normalization process on features with different characteristics, the characteristics of individual features are not exploited to the maximum. As a result, the classification may tend to be inaccurate.

MFC unlike a conventional classifier does not combine the multiple classes of features which are obtained from the data and take them as input. Instead it uses attributes related to each feature vector separately to classify each datum [3]. During the training for each classifier which is trained using a particular feature vector, a confusion table is calculated. From the confusion table of the classifier its accuracy is calculated. In the testing, using an appropriate decision making model the final classification output is obtained.

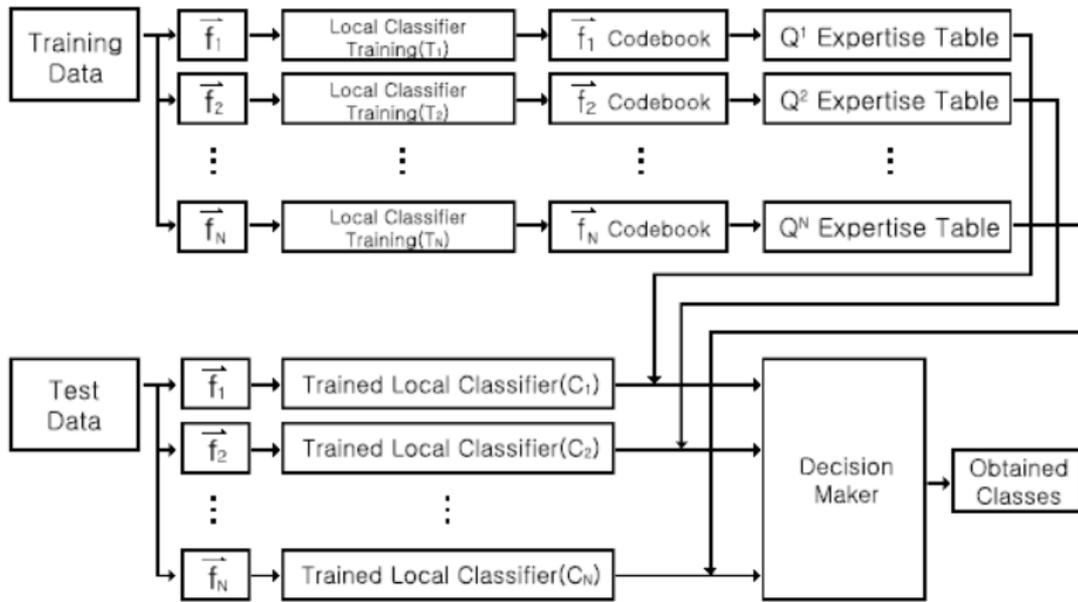


Fig. 1: Multi Feature Based Classifier

B. Support Vector Machine

The Support Vector Machine is a non-parametric, learning technique that has been used in many real world problems which involve classification of data. It is based on concepts like convex optimization and geometry and has the ability to do non-linear classifications. SVM is commonly used for machine learning and data mining application. For both spatio-temporal and temporal spectrum sensing in CR networks under both single PU and multiple PU’s scenarios we can use SVM.

SVM constructs a single hyperplane or multiple hyperplanes in an infinite-dimensional space, which can be used for tasks like classification, outlier’s detection, regression etc. The hyperplane which has the largest distance to the nearest training-data point (known as functional margin) of any class is considered to have achieved a good separation.

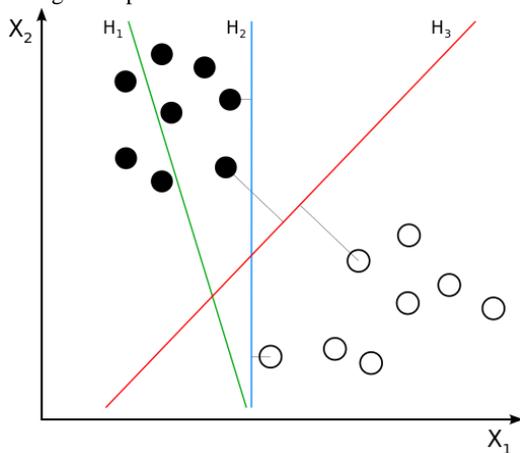


Fig. 2: Hyperplanes separating data

C. Rayleigh Fading Channel

Rayleigh fading is a statistical model for radio propagation. In this model it is assumed that when a signal passes through such a medium its amplitude will vary according to the Rayleigh Distribution. Signal propagation in troposphere and ionosphere layers is observed to be similar to that in a Rayleigh channel. Rayleigh fading channels also reasonably model radio signals propagation in dense urban environments.

In scenarios where a particular signal is not very dominant i.e.) there is no line of sight (LOS) signal; the Rayleigh fading model is more applicable. The Probability Distribution Function (PDF) of strength of the signal received from a Rayleigh fading channel is given as

$$f(x) = \frac{x}{\sigma^2 \alpha} \exp\left(-\frac{x^2}{2\sigma^2 \alpha^2}\right) \quad \forall x > 0. \tag{1}$$

Where σ^2 represents the variance of the received signal strength and α represents the fading parameter.

D. Rician Fading Channel

Rician fading channel is a statistical model for radio propagation in which a radio signal which takes multiple paths can partially cancel itself. The signal propagating through several variant paths (hence demonstrating multipath interference) arrives at the receiver, and having one or more paths changing in length. In a scenario where the signal from one path is much stronger than the signals from other paths, Rician fading occurs. The stronger signal is usually the line of sight (LOS) signal. Rician distribution characterizes the amplitude gain in this type of fading. The Probability Distribution Function (PDF) of strength of the signal received from a Rician fading channel is given as

$$f(x|v, \sigma) = \frac{x}{\sigma^2} \exp\left(-\frac{(x^2 + v^2)}{2\sigma^2}\right) I_0\left(\frac{xv}{\sigma^2}\right) \tag{2}$$

In the above equation, x follows the Rician Distribution, σ^2 stands for the variance of received signal strength and with

$$v^2 = \frac{K}{1 + K} \Omega \tag{3}$$

And

$$\sigma^2 = \frac{\Omega}{2(1 + K)} \tag{4}$$

Where Ω is the total power from both the direct path and scattered paths and K is the ratio between the powers in both the paths.

3. Feature Vectors

A. Energy Vector

Energy detection (ED) is also referred to as radiometry or periodogram. It is a non-coherent technique. Based on the sensed energy one can determine whether the PU signal is present or absent. The energy vector is the most used feature vector due to its low computational complexity. To obtain this feature vector it is not required to have prior information about the PU signal's characteristics. The energy of the received signal at a particular time can be calculated as the mean squared values of the amplitude of the received signal. The obtained energy values of the received signal over a particular time can be used as a feature vector (En). The summation of the mean squared values of the amplitude of the received signal in a certain time interval can also be used as a feature. This feature can be calculated by using the following equation.

$$Y(x) = \sum_{n=1}^{N_s} |x(n)|^2 \quad (5)$$

Energy detection is very pragmatic because we do not require any information about the PU. However, in scenarios where the noise is highly uncertain the accuracy of the classifier using this feature reduces. Besides, below a certain SNR threshold known as the SNR wall, it is not possible to achieve a reliable detection even by increasing the sensing duration. Moreover, in such a scenario by using the energy vector a classifier will not be able differentiate the PU's signal and the noise and other interference signals, which can cause an increase in the false alarm probability.

B. Sample Covariance Matrix

Generally the statistical covariance matrix of a signal differs from that of the noise. This difference of the sample covariance matrices (for N_s samples) can be used to perform spectrum sensing. It is assumed that the Primary User's signal is correlated. If we denote the statistical covariance matrix as R_x , and the sample auto-correlations of the received signal is computed as

$$r(l) = \sum_{n=0}^{N_s-1-l} x(n)x(n-l), \quad l = 0, 1, \dots, N-l \quad (6)$$

N is the smoothing factor. The approximation of the statistical covariance matrix is the sample covariance matrix, $\hat{R}_x(N_s)$, which can be defined as

$$\hat{R}_x(N_s) = \begin{bmatrix} r(0) & r(1) & \dots & r(N-1) \\ r(1) & r(2) & \dots & r(N-2) \\ \vdots & \vdots & \ddots & \vdots \\ r(N-1) & r(N-2) & \dots & r(0) \end{bmatrix} \quad (7)$$

Under the scenario where no signal is transmitted, the non-diagonal elements of $\hat{R}_x(N_s)$ are zero in theory as noise is generally considered to be uncorrelated. The elements in the diagonal of the matrix contain the noise power. Whereas, under the scenario where a signal is transmitted, the non-diagonal elements of $\hat{R}_x(N_s)$ should be non-zeros as the primary signal will be correlated. Therefore we have two terms of interest which are the sum of non-diagonal elements and the sum of diagonal elements and they are computed as

$$T_1(N_s) = \frac{1}{N} \sum_i^L \sum_j^N |r_{ij}(N_s)| \quad (8)$$

And

$$T_2(N_s) = \frac{1}{N} \sum_i^N |r_{ii}(N_s)| \quad (9)$$

Where $r_{ij}(N_s)$ are the elements of the matrix $\hat{R}_x(N_s)$. The test statistics for determining whether the PU signal is present or absent can be given by their ratio.

$$Y(x) = \frac{T_1(N_s)}{T_2(N_s)} \quad (10)$$

Apart from this there are other test statistics which have been extracted from the covariance matrix and used as feature vectors. These include the Eigen vector which corresponds to the maximum Eigen value of the sample covariance matrix and the Eigen values of the matrix.

4. Decision Making Models

A. Hard Decision Making Model

In this decision model we use a fixed threshold to arrive at the decision. The labels for a set of signals are obtained from different SVM's which have been trained using specific feature vector. Then a weighted average is performed over the obtained labels using the accuracies of the SVM's. The weighted average can be obtained can be computed as:

$$L_{avg} = \frac{L_{S1}Acc_{S1} + L_{S2}Acc_{S2} + \dots + L_{Sn}Acc_{Sn}}{Acc_{S1} + Acc_{S2} + \dots + Acc_{Sn}} \quad (9)$$

Where L_{avg} is the averaged set of Labels, L_{Sn} is the set of Labels obtained from n^{th} SVM and Acc_{Sn} is the accuracy of n^{th} SVM. Then using a fixed threshold the set of averaged Labels can be classified.

B. SVM Based Decision Making Model

In this decision model we use a Support Vector Machine to arrive at the decision. From the SVM's which are trained by specific feature vectors we obtain the scores for each class. The scores are such that if a data point is closer to particular cluster center then the score is closer to 1 and if the scores are farther from a cluster center then the score is closer to 0. In case of a binary classification like ours, the score of a data point for 1 class can be considered as the distance of the data point from the other class. Unlike the labels the score can give us the distance from the cluster points. The scores/distances from multiple SVM's are combined by using class wise accuracies of each classifier to obtain a new score/distance. These new scores are computed as:

$$D_{new}^i(p) = \left[\left(\frac{Acc_{S1}^i}{D_{S1}^i(p)} \right) + \left(\frac{Acc_{S2}^i}{D_{S2}^i(p)} \right) + \dots + \left(\frac{Acc_{Sn}^i}{D_{Sn}^i(p)} \right) \right]^{-1} \quad (10)$$

Where $D_{new}^i(p)$ is the new Distance/score from the cluster center of class i for that point p , Acc_{Sn}^i is the accuracy of n^{th} SVM for class i and $D_{Sn}^i(p)$ is the distance/score of the data point p from the cluster center of class i of the n^{th} SVM. These new distance/score are then sent to the SVM for training the SVM. Even for the testing data the same computation has to be followed in order to get the data classified.

5. Results

Table 1: Accuracy Table

| S.No | SVM | Confusion Matrix (true, predicted) | | | | Accuracy |
|--|--|-------------------------------------|-------|-------|-------|----------|
| | | (0,0) | (0,1) | (1,0) | (1,1) | |
| Specific Feature Vectors | | | | | | |
| 1 | Energy vector (En) | 760 | 2 | 2 | 736 | 99.7 |
| 2 | Sum of the Energy vector (S) | 727 | 59 | 35 | 679 | 93.7 |
| 3 | Eigen Values of Covariance Matrix (Eig) | 731 | 64 | 31 | 674 | 93.7 |
| 4 | Ratio of Sum of Non Diagonal elements to Sum of Diagonal elements of Covariance Matrix (R) | 754 | 4 | 8 | 734 | 99.2 |
| 5 | Eigen Vector corresponding to the Maximum Eigen Value of the Co-Variance matrix (V) | 385 | 123 | 377 | 615 | 66.7 |
| MFC with Hard Decision Making model | | | | | | |
| 6 | En + Eig | 998 | 2 | 0 | 1000 | 99.9 |
| 7 | S + R | 985 | 15 | 10 | 990 | 98.75 |
| 8 | S + R + V | 985 | 15 | 8 | 992 | 98.85 |
| MFC with SVM based Decision Making model | | | | | | |
| 9 | En + Eig | 998 | 2 | 0 | 1000 | 99.9 |
| 10 | S + R | 991 | 9 | 9 | 991 | 99.1 |
| 11 | S + V | 962 | 38 | 4 | 996 | 97.9 |
| 12 | R + V | 986 | 14 | 10 | 990 | 98.8 |
| 13 | S + R + V | 986 | 14 | 0 | 1000 | 99.3 |

From the results obtained in TABLE I we can take the following inferences

- The energy vector and the Ratio vector are very good feature vectors
- The MFC with hard decision making model does significantly improve the accuracy of classification
- MFC with SVM based decision making model further improves the accuracy
- The combination of energy vector and Eig vector results in an almost ideal classifier

It is to be noted that the Energy vector and Eigen values vector are large vectors which have 500 attributes and 999 attributes respectively per signal. This makes using these vectors heavily memory intensive for very high frequency signals. Whereas the other feature vectors are single attribute vectors hence are very less memory intensive in comparison with the Energy and Eigen values vector.

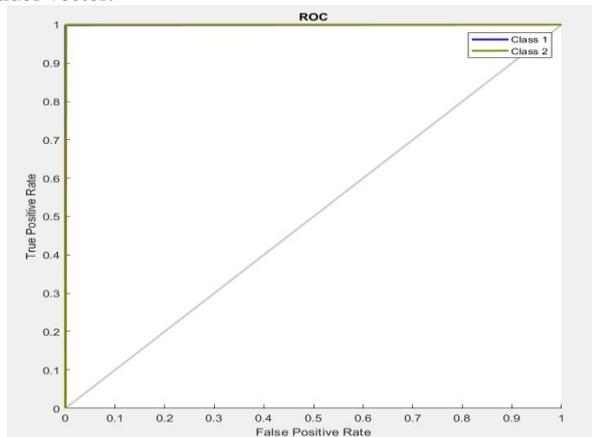


Fig. 3: Receiver Operating Characteristics for MFC (En + Eig)

The Receiver Operating Characteristics (ROC) curve shown in Fig.2 is very close to the ideal ROC curve when considering MFC which combines Energy and Eig vector. But is not a feasible option as the feature vectors are heavily memory intensive

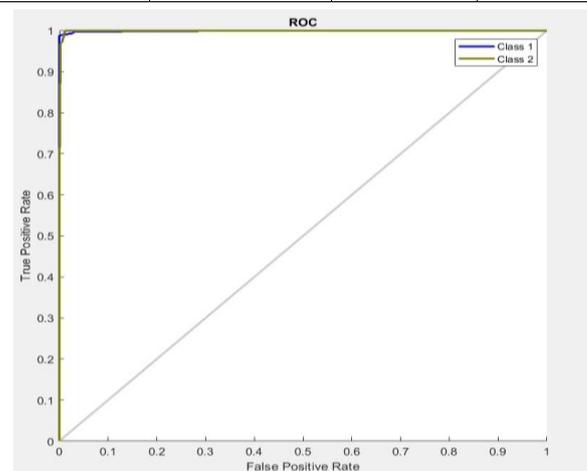


Fig. 4: Receiver Operating Characteristics for MFC (S + R + V)

The Receiver Operating Characteristics of the MFC which combines S vector, R vector and V vector shown in Fig.3 is also close to the ideal. We propose the MFC which combines Sum of Energy (S), Ratio of sum of non diagonal elements to sum of diagonal elements of covariance matrix (R) and Eigen vector corresponding to the maximum Eigen value of the Covariance matrix (V), based models. This MFC is very less memory intensive and has good accuracy.

6. Future Enhancements

The Multi Feature based Classifier is a very good option for the multiple Primary user scenario. Because of the usage of class accuracies in the SVM based decision making model, feature vectors whose class accuracies outperform the overall accuracy can also be incorporated in improving the accuracy of decision making. This makes it very compelling to use it under multiple class scenarios like presence of multiple primary users. This algorithm can also be tested and optimized for MIMO channels.

7. Conclusion

We proposed a machine learning based spectrum sensing algorithm which uses a Multi Feature based Classifier (MFC) model. Energy vector and the sample covariance matrix were used for deriving the required feature vectors. Support Vector Machine was then used in the MFC model. Two decision making models were used in the MFC model. The results show that MFC with

SVM based decision making model showed great accuracy. Even though the MFC which combines En and Eig vector displays close to ideal accuracy, it is highly memory intensive. We propose to use MFC which combines S vector, R vector and V vector which is very less memory intensive. On the whole the MFC proves to be an effective algorithm for improvement of the performance of the spectrum sensing algorithm.

References

- [1] Karaputugala Madushan Thilina, Kae Won Choi, Nazmus Saquib, and Ekram Hossain, "Machine Learning Techniques for Cooperative Spectrum Sensing in Cognitive Radio Networks", *IEEE Journal on Selected Areas in Communications* (Volume: 31, Issue: 11, November 2013).
- [2] Haozhou Xue and Feifei GAO, "A Machine Learning based Spectrum-Sensing Algorithm Using Sample Covariance Matrix", *10th International Conference on Communications and Networking in China (China Com)*, 2015.
- [3] Dong-Chul Park, "Multiple Feature-based Classifier and Its Application to Image Classification", *2010 IEEE International Conference on Data Mining Workshops*
- [4] Ethem Alpaydin "Introduction to Machine Learning ", Third Edition