Tech Science Press

Spectral Vacancy Prediction Using Time Series Forecasting for Cognitive Radio Applications

Vineetha Mathai[*] and P. Indumathi

Department of Electronics Engineering, Madras Institute of Technology, Anna University, Chennai, 600044, India
*Corresponding Author: Vineetha Mathai. Email: vineethamathai@gmail.com

Abstract: An identification of unfilled primary user spectrum using a novel method is presented in this paper. Cooperation among users with the utilization of machine learning methods is analyzed. Learning methods are applied to construct the classifier, which selects the suitable fusion algorithm for the considered environment so that the out of band sensing is performed efficiently. Sensing performance is looked into with the existence of fading and it is observed that sensing performance degrades with fading which coincides with earlier findings. From the simulation, it can be inferred that Weibull fading outperforms all the other fading models considered. To accomplish missed detection probability of 1% in the Rayleigh channel, the false alarm probability obtained is almost 0.8 however to obtain the same missed detection probability in the Weibull channel, false alarm probability is less than 0.1 which is very favorable for both indoor and outdoor scenarios. Numerical analyses are carried out here to predict Primary User (PU) channel condition using Hidden Markov Model with the help of Time series forecasting learning method. It is evident that the prediction performance has reached 100% as the result of using the Weibull Fading Model for a period of 200 ms when compared to the Rayleigh model which is achieving only 84.5% accuracy in prediction.

Keywords: Co-operative spectrum sensing; machine learning; decision fusion; weibull fading; time series; HMM

1 Introduction

When the spectrum is underutilized, Cognitive Radio (CR) is applied to exploit spectrum holes. This permits the secondary users (SU) along with Primary Users (PU) to approach an unoccupied space in an adaptable style [1]. To observe radio frequencies incessantly Spectrum Sensing (SS) [2] is used. The detection method is used to recognize the availability of PU by computing the received signal energy by applying an Energy Detector (ED) [3]. To mitigate the hidden terminal problem and shadowing effects, the Cooperative Spectrum Sensing (CSS) [4] technique is introduced where numerous CRs in the system feel the presence of PU as shown in Fig. 1. In CSS, for creating decisions CR users interchange sensing outcomes to the Fusion Center (FC). The CR users interchange information to check whether the

acquired energy is more than the threshold. Estimated energy is transferred to the fusion center to create an improved decision by using several hard [5–8] and soft-decision algorithms. The FC gives a global result by merging the individual decisions from each CRs.
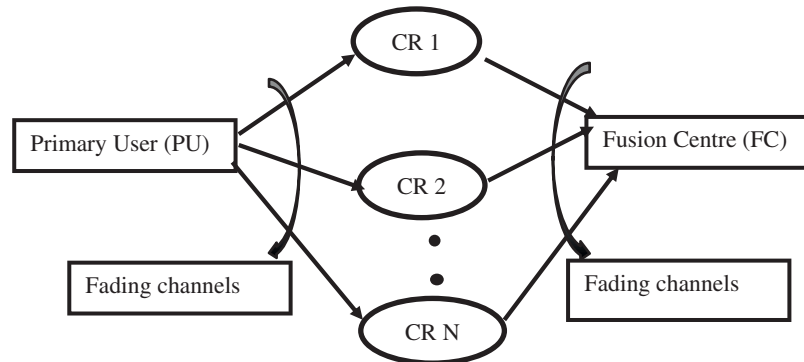

Figure 1: Proposed schematic for cooperative sensing

Using extended generalized-K distribution fading channels are modeled for spectrum sensing using ED in [9]. The CSS model with an Improved ED (IED) scheme is reviewed in [10] uses several antennas at each CR. The CSS model with an antenna at every CR over the Rician channel is explored in [11]. System parameter performance with an IED scheme is examined over channel models is done in [12]. The analogous study in [13–16] using Nakagami-m and Weibull fading channel with ED method is done since it does not require prior information about the channel. A classifier [16] with energy statistics of PU along with decisions about the existence of PU is used for training. Based on learning methods classifiers can be classified as supervised and unsupervised methods. The unsupervised machine learning technique we used is the K-means algorithm to study and find out the primary user's data and patterns. The supervised machine learning techniques we used here is SVM to train the model with the data labeled in the previous step. To forecast the decisions on undetected PU channel conditions, the trained classifier is used. Evaluating the PU channel state [17–21] in CR is a tedious job since it is time varying. A lot of works approach on earlier discussed problem [22–24] using Fast Fourier transform (FFT). It is having high complexity because of large computations. Here we introduce the time series forecasting method for identifying the occupancy of the PU channel state. A time series [25] is used to study the PU channel state (i.e., "idle" or "occupied"). Then Hidden Markov Model (HMM) model is used to capture PU channel status [26,27].

Here, we suggest a novel model based on Machine Learning (ML) techniques to detect the unused spectrum of primary users effectively. We consider "energy vector", as a feature vector in which the energy level is calculated at each CR. Then, the classifier classifies the vector into either the "channel available class" or the "channel unavailable class". We investigated impact of small-scale fading such as multi-path fading, delay spread and Doppler spread are considered over different fading channels. We examined the performance of an unsupervised machine learning algorithm as well as a supervised machine learning algorithm to construct the classifier for decision making. A qualitative performance evaluation of various fusion methods with learning is presented. To acquire the advantage of multi-path propagation, energy detection with receiver diversity is considered and analysis in terms of missed detection and false alarm probabilities over various fading channels is done. Using HMM, primary user channel state prediction is carried out. To our knowledge, no literature considers both sensing and reporting channels as fading models and includes ML methods for classification and further to foresee next state of PU channel using the time series forecasting method.

The rest of the paper is organized as below. In Section 2 we introduce the signal model and the proposed cooperative spectrum sensing scheme. Different classification and fusion methods are presented. Simulation results and discussions are given in Section 3 discusses the experimental outcomes. Finally, conclusion is given at Section 4.

2 System Model

CR network with a single PU and seven SUs is considered. Here the PU alternates between the idle and busy states. The SU knows the presence of the PU by sensing signals using an ED technique. The signal energy is used to predict the presence of PU. We consider a cooperative CR network with K nodes considering N test points for the ED and M frames for training the ML classifier, as illustrated in Fig. 2.
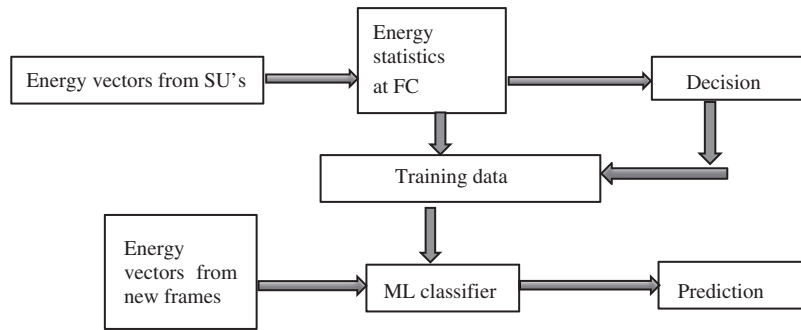

Figure 2: Proposed ML classifier model

The ith frame of the received signal at the jth cooperative node, $Y_{ij}(n)$ is given by

$$Y_{ij}(n) = \begin{cases} w_{ij}(n) & H_0 \\ \sqrt{\gamma_{ij}} S_{ij}(n) + w_{ij}(n) & H_1 \end{cases} \tag{1}$$

where $S_{ij}(n)$ is the primary user's signal, which is a Gaussian i.i.d random process with zero mean, $w_{ij}(n)$ is the noise. The nodes are sensing frame and obtain the statistics. For i^{th} frame at the fusion center, the energy samples Y_{ij}, is expressed as $\gamma_i = \frac{1}{N} \sum_{n=1}^{N} |Y_{ij}(n)|^2$, i lies between 1 and M, where γ_i is a random variable that has a pdf with chi-square distribution. γ_{ij} is the signal-to-noise ratio of the ith frame observed at the jth cooperative node. σ_j^2 is standard deviation of noise samples $w_{ij}(n)$. For all frames, noise variance and SNR remains unchanged during the training process is our assumption. For each frame in the training set a threshold, λ_j is selected then the probability of the false alarm, P_{fa} is obtained as

$$P_{fa}(\lambda_j) = Q\left(\left(\frac{\lambda_j}{\sigma_j^2} - 1\right)\right) \tag{2}$$

and probability of detection P_d is

$$p_d(\lambda_j) = Q\left(\left(\frac{\lambda_j}{\sigma_j^2(1+\gamma_j)} - 1\right)\sqrt{\frac{N}{2}}\right) \tag{3}$$

where Q (.) is the complementary distribution function.

2.1 Fusion Center Threshold Calculation for Various Data Fusion Rules

In order to achieve optimal threshold, λ fusion rules are considered here. K nodes work together for computing the threshold to build the overall sensing decision. The following hard fusion rules are used in the decision center of the secondary users.

2.1.1 AND Fusion Rule

All nodes detect the signal and inform it to FC then, this rule decides that the signal is present so the threshold [2] is

$$\lambda_{AND} = \left(\sqrt{\frac{2}{N}} Q^{-1}\left(p_f^{\frac{1}{k}} \right) + 1 \right) \sigma_u^2 \tag{4}$$

the detection probability $P_{d\ AND}$ [7] can be expressed as

$$P_{dAND} = \left(Q\left(\left(\frac{\lambda_{AND}}{\sigma_u^2 + (1+\gamma_u)} - 1 \right) \sqrt{\frac{N}{2}} \right) \right)^K \tag{5}$$

2.1.2 OR Fusion Rule

If one among many users detect the signal, then this rule determines that signal is present so the threshold [2] is

$$\lambda_{OR} = \left(\frac{2}{N} Q^{-1}\left(1 - (1-p_f)^{\frac{1}{k}} \right) + 1 \right) \sigma_u^2 \tag{6}$$

the detection probability P_{dOR} [7] can be

$$P_{dOR} = \left(1 - \left(1 - Q\left(\frac{\lambda_{OR}}{\sigma_u^2 + (1+\gamma_u)} - 1 \right) \sqrt{\frac{N}{2}} \right) \right)^K \tag{7}$$

The soft fusion rules we considered here include maximum ratio combining (MRC) and square law selection (SLS).

2.1.3 Maximum Ratio Combining (MRC)

All CRs transmit their corresponding vectors to the FC. It gathers the data, combines them, creates a global decision by comparing the value with the detection threshold. The fusion center adds them after receiving these energy statistics $\gamma_{si} = \sum_{j=1}^{K} w_j \gamma_{ij}$ for i varies from 1 to M where γ_{ij} is i[th] frame energy test statistics. To maximize the detection probability, the optimum weight vector, w_j has to be obtained. The weighting coefficient vector w_j can be achieved as $w_j = sign(g^T w_0) w_0$ where $g^T = [\sigma_1^2 \gamma_1, \sigma_2^2 \gamma_2, \ldots, \sigma_K^2 \gamma_K]$ with instantaneous SNRs and noise variances. Then fusion threshold is

$$\lambda_{MRC} = \left(\sum_{j=1}^{K} w_j \sigma_j^2 \right) Q^{-1}[P_f] + \sum_{j=1}^{K} w_j \sigma_j^2 \tag{8}$$

the detection probability P_{dMRC} is given by

$$P_{dMRC} = Q\left(\left(\frac{\lambda_{MRC}}{\sum_{j=1}^{K}(1+\gamma_j)w_j\sigma_j^2} - 1 \right) \sqrt{\frac{N}{2}} \right) \tag{9}$$

2.1.4 Square Law Selection Rule (SLS)

FC chooses the user with maximum SNR, γ_{SLS} [7] with noise variance, σ_u^2. Then the threshold is given by

$$\lambda_{SLS} = \left(\frac{2}{N} Q^{-1} \left(1 - (1 - p_f)^{\frac{1}{k}} \right) + 1 \right) \sigma_u^2 \tag{10}$$

the detection probability P_{dSLS} is given by

$$P_{dSLS} = \left(1 - \left(1 - Q \left(\frac{\lambda_{SLS}}{\sigma_u^2 + (1 + \gamma_{SLS})} - 1 \right) \sqrt{\frac{N}{2}} \right) \right)^K \tag{11}$$

2.2 ML Based Classification Methods

From the energy vectors obtained is compared with the threshold to obtain the decision, d_i which is linked with i^{th} frame is mentioned as

$$d_i = \begin{cases} 1 & \gamma_c \geq \lambda \\ -1 & \gamma_c \leq \lambda \end{cases} \tag{12}$$

where $\gamma_c \in \{\gamma_i, \gamma_{si}\}$ and $\lambda \varepsilon \{\lambda_{AND}, \lambda_{OR}, \lambda_{MRC}, \lambda_{SLS}\}$. Here "−1" expresses the PU at rest, and "1" expresses busy PU communication on channel as mentioned in Eq. (12). For classification problem ML based classifier can be used to perceive the decision linked with a new frame. Initially, K-means is applied to find primary users' transmission patterns and statistics. Then, to differentiate between two states i.e., active or inactive state of the primary user signal, the Support Vector Machine (SVM) method is adopted. In the first step, unsupervised machine learning techniques are used to analyze the primary user's data and patterns. Further, supervised machine learning techniques are used to train the model with the data labeled in the previous step.

2.2.1 K-Means Clustering

This method divides a group of the training energy vectors into clusters. The set of those vectors that fit in to cluster k is denoted by C_k. It's centroid is α_k. It aims to select K clusters, $C_1, \ldots, C_K$, which minimize

$$\underset{C_1, \ldots, C_k}{\text{argmin}} \sum_{k=1}^{K} \sum_{y \in C_k} \| y^l - \alpha_k \|^2 \tag{13}$$

The steps involved in finding clusters that satisfy Eq. (13) are

1) The centroids for clusters are initialized first and its value in cluster 1 is fixed as $\alpha_1 = \mu_{\gamma|S}$. From the cluster set, one cluster is selected that PU is idle (i.e., S = 0) so it belongs to channel available, whereas further clusters belong to unavailable class.
2) From the energy vectors assigned earlier, centroids are calculated and are updated by obtaining the average of all vectors except for cluster 1.
3) The iterations are recalculated till it converges. By K-means algorithm centroid for cluster k is obtained as $\alpha_k^* = |C_k|^{-1} \sum_{y^l \in C_k} y^l, \forall k = 2, \ldots, K$. After the training it receives the test energy vector, y^* for further classification. Then classifier decides whether it is a part of cluster 1 or not, by computing distance from the test vector to the centroids.

$$\frac{\| y^* - \alpha_1^* \|}{\underset{k=1,\ldots,K}{\min} \| y^* - \alpha_k^* \|} \geq \beta \tag{14}$$

where parameter β is the threshold to limit the compromise between the misdetection and the false alarm probabilities.

2.2.2 Support Vector Machine

These are supervised learning models and is used to produce the best decision boundary which can separate space into classes. This boundary is called hyperplane. The way of choosing vectors that aid in creating the boundary. These extreme vectors are called support vectors. The dataset of n points is given in the form of $(x_1, y_1),\ldots,(x_n, y_n)$ and the final choice is either 1 or -1 that denotes the class to which x_i belongs [16] based on the hyperplane equation as

$$\mathbf{w.x} - b = 0 \tag{15}$$

where w is the normal vector to the hyperplane and b is constant. For the training set (γ_i, d_i) where i varies from 1 to M, the optimization is done to maximize the border between both classes (i.e., $w\,\gamma_i + b = \pm 1$) corresponding hyper-plane equation $w\,\gamma_i + b = 0$ can be found out using this classifier.

$$\min_{w,b} \left(\frac{\parallel w \parallel^2}{2} \right) \text{where} \parallel w \parallel^2 = w^T w \tag{16}$$

Using Lagrangian function the solution for quadratic optimization is mentioned as

$$L(w, \; b, \; \alpha) = \frac{\parallel w \parallel^2}{2} - \sum_{i=1}^{M} \alpha_i(d_i(w\gamma_i + b) - 1), \; \alpha_i \geq 0 \tag{17}$$

where $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots \alpha_M)$ is the Lagrangian multipliers. If LHS of Eq. (17) is equal to zero then we get $w = \sum_{i=1}^{M} \alpha_i d_i \gamma_i$ and $\sum_{i=1}^{M} \alpha_i d_i = 0$ and by putting it into Eq. (17) it can be reduced as

$$\min_{\alpha} \left(\frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} d_i d_j (\gamma_i \gamma_j) \alpha_i \alpha_j - \sum_{i=1}^{M} \alpha_j \right), \; \alpha_j \geq 0 \tag{18}$$

We can find α from Eq. (18) and we can be computed using $w = \sum_{i=1}^{M} \alpha_i d_i \gamma_i$. By selecting α_i is greater than 0 further, find b from the expression $b = d_j - \sum_{i=1}^{M} \alpha_i d_i (\gamma_i \gamma_j)$ then the new decision γ_x using the given classification function is

$$\text{class}(\gamma_x) = \text{sign} \left(\sum_{i=1}^{M} \alpha_j d_j (\gamma_i \gamma_x) + b \right) \tag{19}$$

2.2.3 Channel Classification

Based on the energy vectors received, the fusion rules are applied and the energy vectors are given as inputs to the above-mentioned algorithms. The ML algorithms classify the channel as available and unavailable class. The former defines that the PU is idle in the channel so SU can get an admission. The latter defines that the PU is present in the channel so the SU cannot access the channel and are redirected to other available channels by cognitive radio. Based on priority and observations at the decision center, it assigns the channel to the SU.

2.3 Fading Models

2.3.1 Log Normal Shadowing

This model tells the path loss a signal experiences inside a building or tightly inhabited areas over distance. The Probability Density Function (PDF) expression is specified as

$$f_\gamma^\gamma = \frac{10}{\ln(10)\gamma\sigma\sqrt{2\pi}} \exp\left(\frac{-10log_{10}\gamma - \mu^2}{2\sigma^2}\right) \tag{20}$$

where μ is the mean and σ indicates standard deviation. Then approximated detection probability [14] is given by

$$P_{d,ln} = 2\sqrt{\frac{\alpha}{2\pi}} \exp\left(\frac{\alpha}{\beta}\right) \sum_{n=0}^{\infty} \frac{\Gamma(u+n, \lambda/2)}{\Gamma(u+n)n!} \left(\sqrt{\frac{\alpha\beta^2}{\alpha+2\beta^2}}\right)^{n-1/2} K_{n-1/2}\left(\sqrt{\frac{\alpha(\alpha+2\beta^2)}{\beta^2}}\right) \tag{21}$$

where α is the shape parameter and β denotes the expectation of γ. $K_v(.)$ signifies a modified Bessel function of second kind, $\Gamma(.)$ is gamma function and its complementary incomplete function be $\Gamma(\cdot,\cdot)$.

2.3.2 Rayleigh Fading

This model assumes signal to noise ratio, γ follows an exponential PDF given as

$$f_y^x = \frac{1}{\bar{\gamma}}\exp\left(\frac{-\gamma}{\bar{\gamma}}\right), \ \gamma \geq 0 \tag{22}$$

The approximated detection probability [14] is given by

$$P_{d,Ray} = \frac{1}{1+\bar{\gamma}}\sum_{n=1}^{u-1}\left(\frac{\lambda}{2}\right)^n \exp\left(\frac{-\lambda/2}{n!}\right) {}_1F_1\left[1; \ n+1; \ \frac{\lambda\bar{\gamma}}{2(1+\bar{\gamma})}\right] + \exp\left(\frac{\lambda}{2(1+\bar{\gamma})}\right) \tag{23}$$

where λ denotes the threshold of the energy detector, $\bar{\gamma}$ denotes average Signal to Noise Ration (SNR) and ${}_1F_1(.;.;.)$ denotes confluent hypergeometric function.

2.3.3 Rician Fading/Nakagami-n Fading Model

In Rician fading, a strong line-of-sight wave is prevailing. The corresponding PDF is given by

$$f_y^x = \frac{1+K}{\bar{\gamma}}\exp\left(\frac{-K(1+K)\gamma}{\bar{\gamma}}\right) * I_0\left[2\sqrt{\frac{K(1+K)\gamma}{\bar{\gamma}}}\right] \tag{24}$$

where K denotes Rician factor, n^{th} order modified Bessel function of the first kind be $I_n(\cdot)$. The approximated probability for detection of this model is

$$P_{d,Ric} = \frac{1+K}{\bar{\gamma}} \exp(-K) \sum_{n=0}^{\infty} \frac{\Gamma(u+n, \lambda/2)}{\Gamma(u+n)} \left(\frac{\bar{\gamma}}{1+K+\bar{\gamma}}\right)^{n+1} \cdot {}_1F_1\left[n+1; \ 1; \ \frac{K(1+K)}{1+K+\bar{\gamma}}\right] \tag{25}$$

where u is the time-bandwidth product.

2.3.4 Weibull Fading

Weibull fading model has been introduced to analyze rapid signal fluctuations in non-line-of-sight scenario. The PDF expression of the instantaneous SNR γ is specified as

$$f_\gamma^\gamma = c\left[\frac{\Gamma(p)}{\bar{\gamma}}\right]^c \gamma^{c-1} \exp\left[-\left\{\frac{\gamma\Gamma(p)}{\bar{\gamma}}\right\}^c\right]; \quad \gamma \geq 0 \tag{26}$$

where c = v/2 and p = 1 + 1/c. The factor v, always greater than zero, is the Weibull fading parameter shows fading severity and if v = 2 Eq. (26) reduces to Rayleigh PDF. The approximated detection probability for Weibull fading channel [14] is

$$P_{d,Wei} = (2\pi)^{\frac{1-c}{2}} \left[\frac{\Gamma(p)}{\bar{\gamma}}\right]^c \sum_{n=0}^{\infty} \frac{\Gamma(u+n,\,\lambda/2)}{\Gamma(u+n)n!} c^{n+c+1/2} . G_{1,c}^{c,1}\left[\frac{1}{\left[\frac{\Gamma(p)}{\bar{\gamma}}\right]^c} \Big| \frac{n+c}{c}, \;\ldots, \; \frac{n+2c-1}{c}\right] \tag{27}$$

where $G_{p,q}^{m,n}\left(x\Big| \begin{matrix} a_1 \ldots & a_p \\ b_1 \ldots & b_p \end{matrix}\right)$ denotes Meijer's G function

2.3.5 Nakagami-m Fading Model

Here the random variable γ obeys a Gamma PDF is mentioned as

$$f_\gamma^\gamma = \left(\frac{m}{\bar{\gamma}}\right)^m \frac{\gamma^{m-1}}{\Gamma(m)} \exp\left(\frac{-m\gamma}{\bar{\gamma}}\right); \; \gamma \geq 0 \tag{28}$$

where m is the fading severity parameter which lies between 0.5 and ∞. The approximated detection probability [14] is given by

$$\begin{aligned} P_{d,Nakm} = &\left(\frac{m}{m+\bar{\gamma}}\right)^m \sum_{n=1}^{u-1} \left(\frac{\lambda}{2}\right)^n \exp\left(\frac{-\lambda/2}{n!}\right) {}_1F_1\left[m;\; n+1;\; \frac{\lambda\bar{\gamma}}{2(m+\bar{\gamma})}\right] + \frac{\bar{\gamma}}{(m+\bar{\gamma})} \exp \\ &\left(\frac{\lambda m}{2(m+\bar{\gamma})}\right) . \left\{\sum_{k=0}^{m-2} \left(\frac{m}{m+\bar{\gamma}}\right)^k L_k\left[\frac{-\lambda\bar{\gamma}}{2(m+\bar{\gamma})}\right] + \left(\frac{m+\bar{\gamma}}{\bar{\gamma}}\right)\left(\frac{m}{m+\bar{\gamma}}\right)^{m-1} L_{m-1}\left[\frac{-\lambda\bar{\gamma}}{2(m+\bar{\gamma})}\right]\right\} \end{aligned} \tag{29}$$

If we put m = 1 this distribution becomes Rayleigh distribution and reduces to Eq. (23).

2.3.6 Hoyt Fading / Nakagami-q Fading Model

This model portrays fading severely than Rayleigh fading where q $(0\leq q \leq 1)$ is the severity parameter. The PDF is mentioned as

$$f_\gamma^\gamma = \frac{1}{\bar{\gamma}\sqrt{p}} \exp\left(\frac{-\gamma}{p\bar{\gamma}}\right) I_0\left[\frac{\gamma\sqrt{1-p}}{p\bar{\gamma}}\right]; \; \gamma \geq 0 \tag{30}$$

where $p = 4q^2/(1+q^2)^2$, p lies between 0 and 1. When p and q are equal to one, then the distribution becomes Rayleigh PDF. The approximated detection probability [14] is given by

$$P_{d,Hoyt} = \frac{1}{\bar{\gamma}\sqrt{P}} \sum_{n=0}^{\infty} \frac{\Gamma(u+n,\,\lambda/2)}{\Gamma(u+n)} \left(\frac{p\bar{\gamma}}{1+p\bar{\gamma}}\right)^{n+1} . 2F_1\left[\frac{n+2}{2}, \frac{n+1}{2}; \; 1; \; \frac{1-p}{(1+p\bar{\gamma})^2}\right] \tag{31}$$

2.4 Spectrum Prediction Model

2.4.1 Time Series Generation

Over time the state of PU is to be predicted whether it is in "Idle state" or "Occupied state" and to obtain the transition from one state to another. For that time series is created to assign every state of the detection series into a different sample space with random variables using Autoregressive Integrated Moving Average (ARIMA) model as shown in Fig. 3. The time series g_t is expressed in (32)

$$g_t \in \begin{cases} u_1, u_2, \ldots u_l & \gamma_t < \lambda \\ u_{l+1} \ldots u_m & \gamma_t \geq \lambda \end{cases} \tag{32}$$

where $g_t \in u_1, u_2, \ldots u_l$ represents PU absent and $g_t \in u_{l+1} \ldots u_m$ represents PU present. We want to estimate the next state X_{t+1} of a PU channel state currently at X_t.

Figure 3: Block diagram for prediction of channel condition

2.4.2 Hidden Markov Model for Primary User Channel Condition Forecasting

In the Hidden Markov Model the system outcome is known and state change probabilities are unknown. Let $X = \{X_1, X_t, \ldots X_T\}$ represents the hidden state series where $X_t \in s_i$, $i = 1, 2, \ldots K$, s_i represents states of PU can take a value of 0 and 1. The number of hidden states is denoted by K. Assume $O = \{O_1, O_t, \ldots O_T\}$ exhibits the observation sequence where $O_t \in u_1, u_2, \ldots u_M$ and a number of the observations is mentioned as M. The transition and the emission probabilities matrix respectively are A and B also the initial state probability vector is represented by π. HMM is represented as $\theta = (\pi, A, B)$ and considered two unknown states $i = 2$, $\pi = (\pi_1, \pi_2)$ is used for calculating above matrices.

$$A = (a_{ij})_{K \times K} \tag{33}$$

$$a_{ij} = P(X_{t+1} = s_j | X_t = s_i) \tag{34}$$

where a_{ij} is the probability that for the present state s_i, the next state is s_j. The emission probabilities matrix is given as

$$B = (b_{jm})_{K \times M} \tag{35}$$

$$b_{jm} = P(O_t = u_M | X_t = s_j) \qquad j = 1 \text{ to } K, \ m = 1 \text{ to } M \tag{36}$$

here b_{jm} denotes probability for current observation u_M, the current state is s_j. Let the probability of state series t which ends at state i be

$$\delta_t(i) = \max_{X_1, \ldots, X_{t-1}} P(X_1, \ldots, X_{t-1} = s_i; \ O_1, O_t | \theta) \tag{37}$$

To maximize Eq. (37) we use a vector $\varphi_t(i)$ which stores the argument values. Following are the steps used in the Viterbi Algorithm to estimate the channel state.

Step 1: Initializes $\delta_t(i)$ and $\varphi_t(i)$

$$\delta_t(i) = \pi_i b_i(O_1) \tag{38}$$

$$\varphi_t(i) = 0 \quad i = 1, \ldots, K \tag{39}$$

Step 2: Repeats to upgrade $\delta_t(i)$ and $\varphi_t(i)$

$$\delta_t(j) = \max_{1 \leq i \leq K} [\delta_{t-1}(i) a_{ij}] b_j(O_t) \qquad j = 1, \ldots K; \ t = 2, \ldots, T \tag{40}$$

$$\varphi_t(j) = \underset{1 \le i \le K}{\mathrm{argmax}}[\delta_{t-1}(i)a_{ij}] \qquad j = 1, \ldots K; \quad t = 2, \ldots, T \tag{41}$$

Step 3: Computes the likelihood probability P*, then estimated state q_T^* by

$$P^* = \max_{1 \le i \le K}[\delta_T(i)] \tag{42}$$

$$q_T^* = \underset{1 \le i \le K}{\mathrm{argmax}}[\delta_T(i)] \tag{43}$$

to obtain HMM parameters $\theta = (\pi, \text{A}, \text{B})$ statistically Baum-Welch Algorithm and maximum likelihood estimation is used. This estimation is done these ways.

Step 1: Obtain the training observation vectors $O_1, O_2 O_t, \ldots O_L$ with length L

Step 2: Obtain $\theta(1)$ when k = 1 and update k = k + 1

Step 3: Calculate $\theta(k)$ based on $O_1, O_2 O_t, \ldots O_L$. Let the probability of current state s_i be $\gamma_t(i)$ and its transition to next state s_j be $\zeta_t(i, j)$

$$\gamma_t(i) = P(X_t = s_i | O_1, \ldots, O_L, \theta) \qquad t = 1, 2 \ldots L \tag{44}$$

$$\zeta_t(i, j) = P(X_t = s_i; \ X_{t+1} = s_j | OO_1, \ldots, O_L, \theta) \tag{45}$$

Step 4: Obtain expected values of state s_i and its transitions to s_j as $E(\gamma_t(i))$ and $E(\zeta_t(i, j))$

$$E(\gamma_t(i)) = \sum_{t=1}^{L-1} \gamma_t(i) \tag{46}$$

$$E(\zeta_t(i, j)) = \sum_{t=1}^{L-1} \zeta_t(i, j) \tag{47}$$

the parameters calculation from the training sequence in [21] is used.

$$\pi_i = \gamma_1(i) \qquad \text{for } i = 1, 2, 3 \ldots K, \text{ at } t = 1 \tag{48}$$

Calculate $\widehat{a_{ij}}$ by varying i and j from 1 to K

$$\widehat{a_{ij}} = \frac{\text{Expected number of transitions from state } s_i \text{ to state } s_j}{\text{Expected number of transitions from state } s_i} = = \frac{\sum_{t=1}^{L-1} \zeta_t(i, j)}{\sum_{t=1}^{L-1} \gamma_t(i)} \tag{49}$$

Calculate $\widehat{b_i}(m)$ by varying i from 1 to K

$$\widehat{b_i}(m) = \frac{\text{Expected number of times in state } s_j \text{ and observing } u_m}{\text{Expected number of times in state } s_j} = \frac{\sum\limits_{\substack{t=1 \\ O_{t = u_m}}}^{L} \gamma_t(i)}{\sum_{t=1}^{L} \gamma_t(i)} \tag{50}$$

Step 5: Get the new approximation of $\widehat{a_{ij}}$, $b_i(k)$ and π_i then label it as $\theta(k+1)$

Step 6: If not converging then go to step 3.

3 Results and Discussions

We assume that PU and SUs are stationary. In our scenario, we considered two PUs and seven co-operative SUs. The simulation is carried out with the assumption of the following parameters as shown in Tab. 1.

Table 1: Simulation parameters

Parameters	Values
Bandwidth	5 MHz
Sensing duration	100 μs
Noise spectral density	−170 dBm
Path loss exponent	4
Transmit power of PU	100 mW

3.1 Training Duration of Classifiers

500 energy vectors were chosen for training each classifier. For unsupervised learning details about channel availability are not required while for supervised learning the channel availability states are required for training. It involves dividing the vectors into clusters and its centroid calculation is done in K-means clustering. But in SVM, locating the hyper-plane which splits the training energy vectors distinctly is the primary task. After successful training of the classifier is completed, the test energy vector is provided for classification. The training duration of classifiers for the size of the energy vector is used for training is shown in Tab. 2. As we increase the number of samples up to 1000 for training the supervised learning technique K-Means technique shows a high training duration (0.117 s) for 1000 samples. Also, the time taken to decide the channel availability is calculated for 1000 samples and obtained as 1.8×10^{-5} for K-Means and 5.5×10^{-5} for SVM.

Table 2: Training duration required (second)

Classification schemes	Samples used for training classifier		
	200	400	1000
K-Means	0.093	0.094	0.117
SVM-Linear	0.014	0.021	0.062

3.2 Machine Learning Classification

We represent the scatter plot of energy vectors here. In Fig. 4a the K-means method defines a threshold margin between the energy vector points by calculating centroids and classify them as the channel available and unavailable class. Fig. 4b shows SVM algorithm using Linear kernel separates the energy points based on channel availability and draws a decision boundary to determine the threshold for classification.
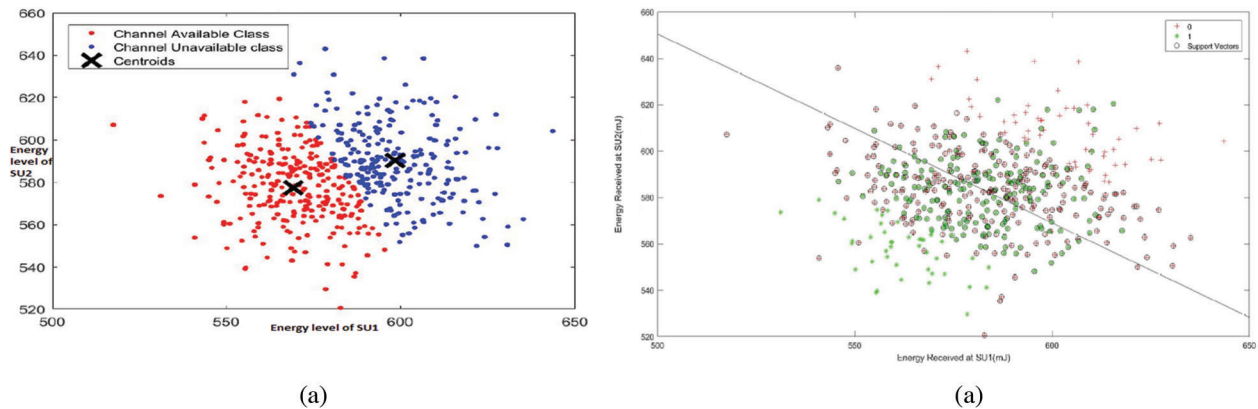
(a)   (a)

Figure 4: Scatter plot of energy vectors using various classification methods (a) K-Means classification (b) SVM classification

We consider SUs located at different places that are 1 km apart. Fig. 5 presents various CSS methods accuracy in terms of Receiver Operating Characteristic (ROC) curves. It portrays that SVM method is better than the other methods. So this method is suited for multiple primary user cases with high accuracy. The SVM with Linear kernel classifier achieves high detection probability even in the presence of seven secondary users. Also, the unsupervised learning K-Means method is having a comparable performance with the SVM method. By bringing intelligence to the system upgrades the classification process. Even though complexity is associated with SVM, the CR scenario is more prone to lose connectivity the supervised learning is more suitable. AND and OR methods will not produce exact outcomes since outcome depends on all and one decision by CRs.
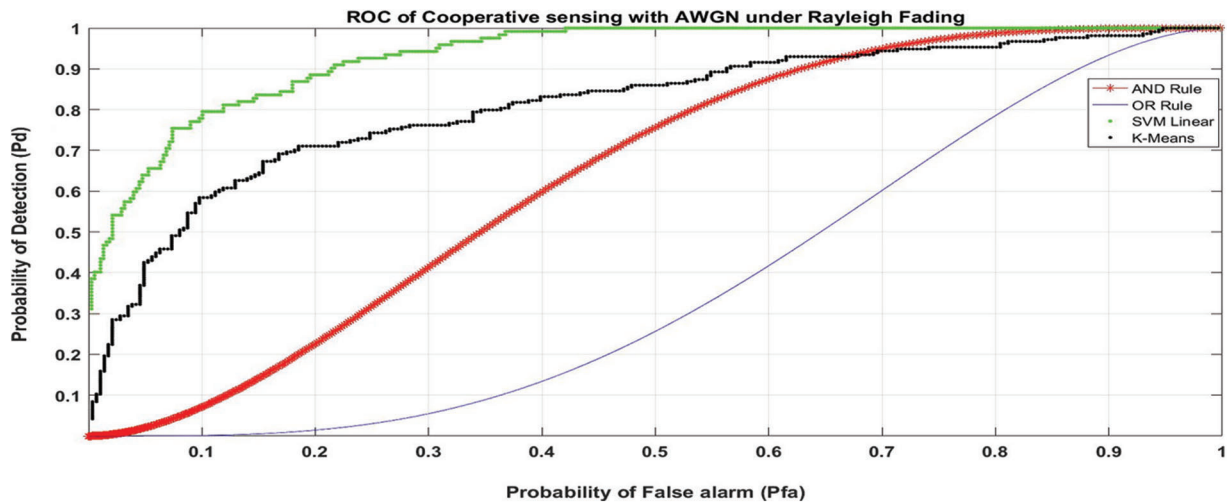

Figure 5: ROC curve of different cooperative spectrum sensing methods

3.3 Numerical Analysis of Spectrum Sensing Using Multiple CR

Numerical outcomes for decision fusion methods are illustrated here. The average snr $\bar{\gamma}$ is assumed to be 15 dB. Fig. 6 manifests the probability of missed detection (P_{md}) *vs.* the probability of false alarm (P_{fa}) under several fading scenarios. Fading channels like Rician, Nakagami-*m*, Hoyt, Weibull and log-normal shadowing are considered with parameters chosen to be K = 3, m = 3, q = 0.25, v = 6, and σ = 5 dB after

running simulations. These parameters are found to be optimal values and it is also coinciding with previous literature. Fading introduces missed detection. We also observe $v = 2$, Weibull fading channel performance is identical to Rayleigh channel. Apparently, Weibull fading shows the best result among all fading models as the slope of the curve drops quickly. To achieve a P_{md} of 1% in a Rayleigh channel, P_{fa} is found to be 0.8 while in the Weibull channel to obtain the same P_{md}, P_{fa} of even lesser value is obtained, which is highly desirable. It is observed that P_{md} value reduces as P_{fa} increases, hence increase detection probability P_d. This reduction in false alarm probability is required in cognitive radar applications to increase detection probability in cluttered environment. So spectrum utilization can be enhanced. From Fig. 7 it is understood that the probability of detection P_d of 75% is achieved even for very less P_{fa} in the Weibull fading model when compared with the Rayleigh model which can produce upto 55% only.
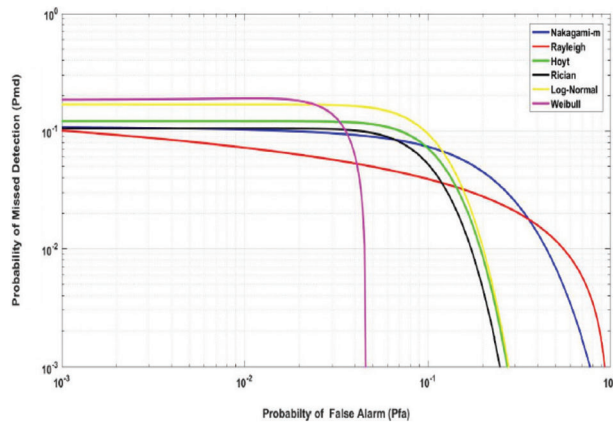

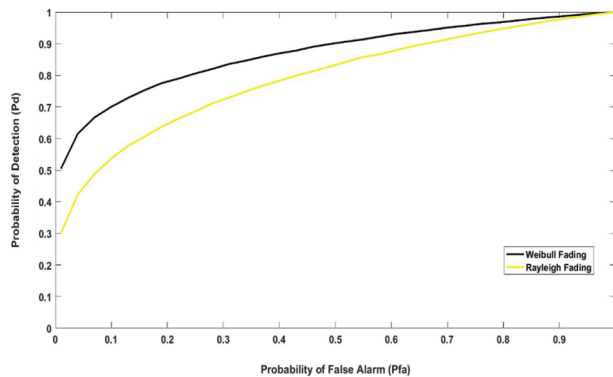**Figure 6:** Impact of CR sensing in different fading channels

Figure 7: ROC curve for Rayleigh and Weibull fading models (P_d $v_s.$ P_{fa})

As the Weibull Fading has shown better performance, we utilize this to improve the channel sensing and the fusion decision approach of the cognitive radio network. The Weibull fading has followed a chi-square distribution function. Fig. 8 represents ROC of various decision fusion methods. We consider seven cooperative nodes and for soft fusion rules the assumed SNR values vary from −25 to −15 dB.

We keep the probability of false alarm also low. It is evident that MRC outperforms the other schemes. It is due to the right selection of fading channels and learning methods. Detection is done with greater accuracy because the weight vector is chosen optimally. It is achieving a P_d of 70% for a P_{fa} of 0.5. Even in less SNR, it works well.
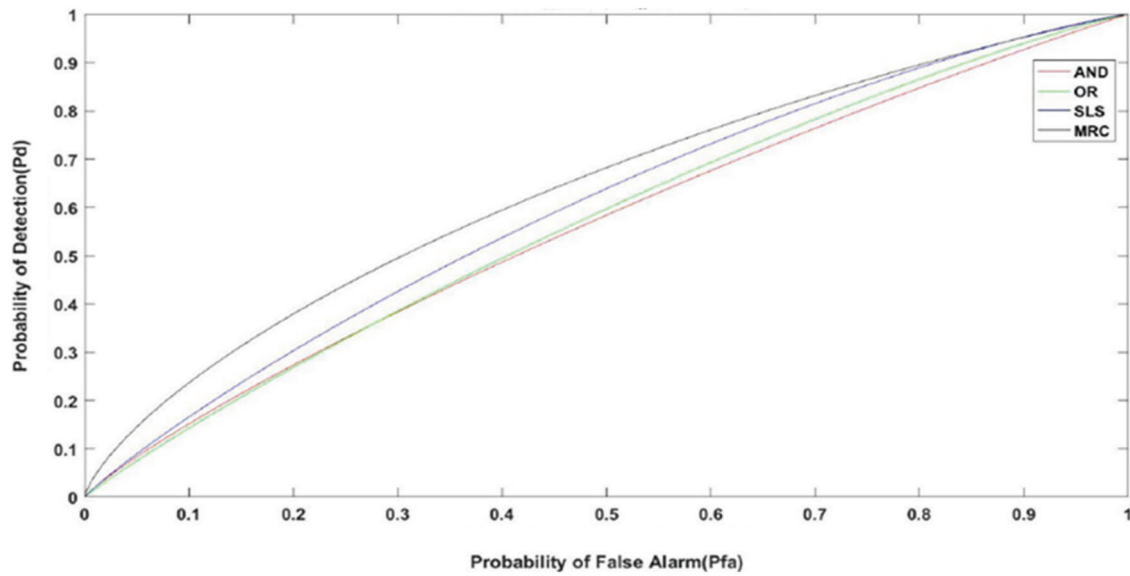
Figure 8: ROC curves for various fusion rules

Once the energy vectors are obtained they are classified as the channel available and unavailable classes and the SVM-Linear algorithm is used for the classification. The ROC plot shows that the performance of the decision Rules is highly improved and almost reached an accuracy of ~100% in Fig. 9. MRC reaches the best accuracy of 99.9% where the AND rule has the least accuracy of 97.8%. Over thousand frames this SVM method predicted efficiently when various thresholds are fed into the classifier. The true positive rate obtained for various methods is shown here. Hence it is evident that the Weibull fading along with Chi-Square Distribution has increased spectrum sensing and SVM classifier with MRC can classify the channels with great accuracy. We use 1000 testing frames for classification. Fig. 10 shows thresholds of the above rules is fed into SVM classifier and it achieves detection rate of 99.9%. It depicts the identification of holes in the spectrum which can be effectively utilized. Among various methods, MRC is showing the best performance.
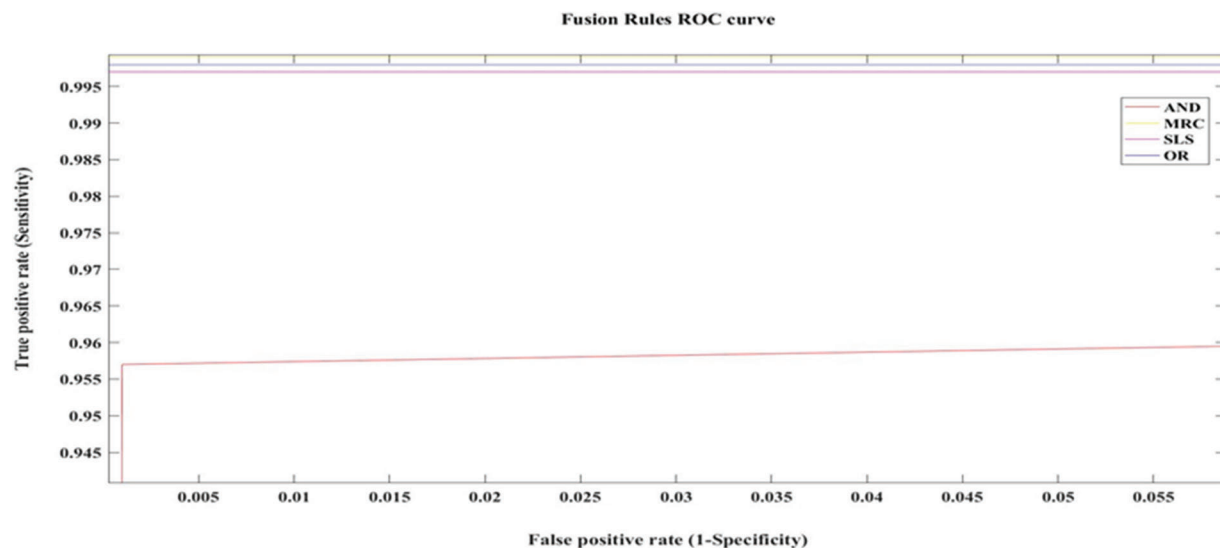

Figure 9: SVM classifier model predicts the output for various fusion thresholds
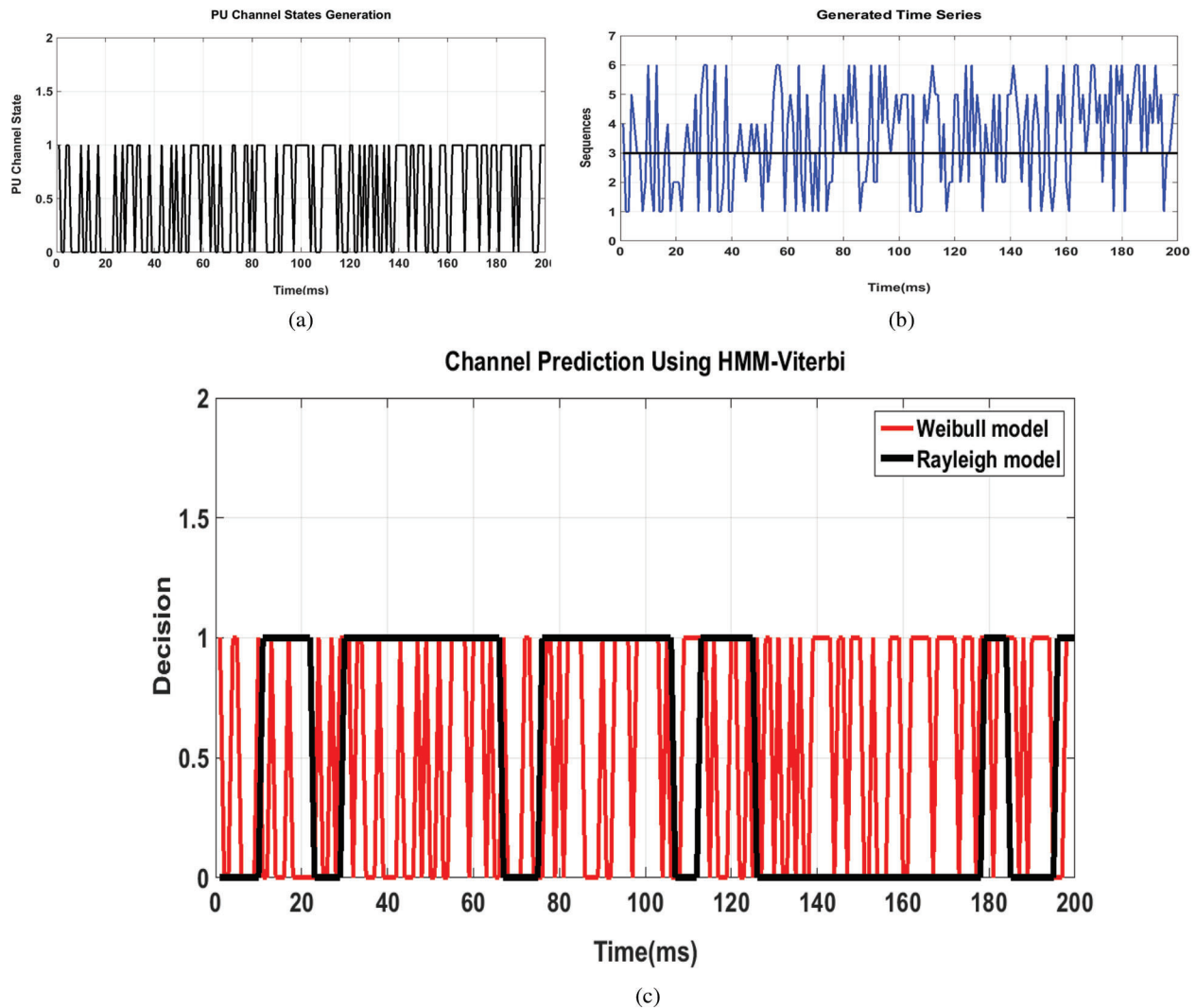
(a)

(b)

(c)

Figure 10: HMM algorithm to capture PU channel conditions (a) Generation of channel states of PU (b) Time series generated for prediction (c) Channel vacancy prediction for various channel models

3.4 Prediction of Primary User Channel State Using HMM Model

To estimate the next state of PU we examine a method and is divided into three steps. First, we find the PU channel state then obtain a model based on the detection that creates a time series to learn about its present condition and finally propose a model for predicting its next state using HMM. For a time period, T = 200 ms Fig. 10a represents the random distribution of PU channel conditions. It presents the state of the PU in a particular channel. It has both idle state (channel available state) and occupied state (channel unavailable state). These idle states are termed as the spectrum holes and can be effectively used for the SU communication in CRN. In Fig. 10b to differentiate the state of activity of the PU channel a margin is drawn. The points below the margin show the amount of the spectrum holes in a channel. While the points above the margin show the instances at which the channel is occupied by the PU. This sequenced time series is used as the input parameter for the HMM prediction model. Once the sequenced time series is generated, it is used along with emission and transmission probability for the prediction of the channel. The Hidden Markov Model along with Viterbi and Baum Welsh algorithm is used for the Spectrum

Prediction. It is evident that the prediction performance has reached 100% as the result of using the Weibull Fading Model for a period of 200 ms from Fig. 10c when compared to the Rayleigh model which is achieving only 84.5% accuracy in prediction. Hence the spectrum prediction of the PU channel activity is done with maximum accuracy.

4 Conclusion

Thus an efficient machine learning algorithm for spectrum sensing and prediction is developed and simulated in the wireless cognitive environment. The spectrum sensing was initially done with the Rayleigh fading model and the channel was classified with the help of machine learning but it has shown a poor act in terms of detection. Hence an improvisation by considering the other channel fading models and tested them in the wireless environment is found. As a result, the Weibull fading model is the best fading model to use in both indoor and outdoor wireless environments. Further to find the best fusion rule algorithm to suit this scenario, we conducted the performance comparison of soft and hard fusion rules and found maximal ratio combining algorithm is performing well by making use of receiver diversity in multiple CRs. Thus we combined the Weibull fading model with the MRC fusion rule and tested it in the cognitive environment under the chi-square distribution for multiple secondary users and this resulted in better detection characteristics. Then HMM with Baum Welsh and Viterbi Algorithm is used for the spectrum prediction by generating a time series to capture the states of PU. This has given a high accuracy in forecasting the next state of the channel.

The chosen energy vectors correctly train the classifier to find the occupancy of the primary user. The duration of training is found to be very less. So SVM method can be a good choice to get a decision on the channel conditions ambiguity. Among various fading models, the Rayleigh model is due to multipath reception and all the other models are its variants. By doing simulations we obtain the best possible values of severity factor of all the methods and found out that detection is good in the Weibull method. It is observed that the P_{md} value reduces as P_{fa} increases, hence increase detection probability. This reduction in missed detection is required in cognitive radar applications to identify targets in a cluttered environment. Thus spectrum utilization can be enhanced. The simulation results revealed that the classifier trained with the ML algorithm helps the fusion center to attain well in terms of sensing accuracy. From the received statistics MRC method gives its best because of the accurate choice of weight parameter based on energy data obtained from all CR users. The fusion center can conclude channel status accurately. Then Hidden Markov model is used to study the occupancy of the primary user channel and predicts its near future to make use of spectrum is efficiently and found out that high accuracy in predicting the occupancy of the channel. Using this model the prediction in various fading models is done and found out that using time series-forecasting model completely fits on past data and using them to predict future interpretation. This work finds application in the field of 5G technology where the wireless spectrum is very important to be shared with lots of IoT machines and smart devices for communication.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] J. Mitola and G. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, 1999.

[2] Y. C. Liang, K. C. Chen, G. Y. Li and P. Mahonen, "Cognitive radio networking and communications: An overview," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 3386–3407, 2011.

[3]   H. Urkowitz, "Energy detection of unknown deterministic signals," in *Proc. of the IEEE*, Philadelphia, vol. 55, no. 4, pp. 523–531, 1967.

[4]   I. F. Akyildiz, B. F. Lo and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Physical Communications*, vol. 4, no. 1, pp. 40–62, 2011.

[5]   S. Kyperountas, N. Correal and Q. Shi, "A comparison of fusion rules for cooperative spectrum sensing in fading channels," *EMS Research*, Motorola, USA, vol. 2010, pp. 1–28, 2010.

[6]   Q. Qin, Z. Zhimin and G. Caili, "A study of data fusion and decision algorithms based on cooperative spectrum sensing," in *Proc. Sixth Int. Conf. on Fuzzy Systems and Knowledge Discovery*, Tianjin, China, pp. 76–80, 2009.

[7]   J. Ma, G. Zhao and Y. Li, "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4502–4507, 2008.

[8]   G. Ding, Q. Wu, J. Wang and X. Zhang, "Joint cooperative spectrum sensing and resource scheduling for cognitive radio networks with soft sensing information," in *Proc. IEEE Youth Conf. on Information Computing and Telecommunications (YC-ICT)*, Beijing, China, pp. 291–294, 2010.

[9]   M. Ranjeeth and S. Anuradha, "Performance of fading channels on energy detection based spectrum sensing," in *Proc. of Procedia Materials Science*, India, vol. 10, pp. 361–370, 2016.

[10]   A. S. Nallagonda, A. Chandra, S. D. Roy, S. Kundu, P. Kukolev *et al.,* "The detection performance of cooperative spectrum sensing with hard decision fusion in fading channels," *International Journal of Electronics*, vol. 103, no. 2, pp. 297–321, 2016.

[11]   I. E. Atawi, O. S. Badarneh, M. S. Aloqlah and R. Mesleh, "Spectrum-sensing in cognitive radio networks over composite multipath/shadowed fading channels," *Computers and Electrical Engineering*, vol. 52, no. C, pp. 337–348, 2016.

[12]   A. M. Mikaeil, B. Guo, X. Bai and Z. Wang, "Machine learning to data fusion approach for cooperative spectrum sensing," in *Proc. Int. Conf. on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*, Shanghai, China, pp. 429–434, 2014.

[13]   M. Ranjeeth and S. Anuradha, "Threshold-based censoring of cognitive radios in Rician fading channel," *Wireless Personal Communications*, vol. 93, no. 2, pp. 409–430, 2016.

[14]   M. Ranjeeth, S. Behera, S. Nallagonda and S. Anuradha, "Optimization of cooperative spectrum sensing based on improved energy detector with selection diversity in AWGN and Rayleigh fading," in *Proc. of Int. Conf. on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, Chennai, India, pp. 1–5, 2016.

[15]   M. Ranjeeth, S. Nallagonda and S. Anuradha, "Optimization analysis of improved energy detection based cooperative spectrum sensing network in Nakagami-m and Weibull fading channels," *Journal of Engineering Science and Technology Review*, vol. 10, no. 2, pp. 114–121, 2017.

[16]   C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning Journal*, vol. 20, no. 3, pp. 273–297, 1995.

[17]   K. M. Thilina, K. W. Choi, N. Saquib and E. Hossain, "Machine learning techniques for cooperative spectrum sensing in cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 11, pp. 2209–2221, 2013.

[18]   N. Reisi, M. Ahmadian, V. Jamali and S. Salari, "Cluster-based cooperative spectrum sensing over correlated log-normal channels with noise uncertainty in cognitive radio networks," *IET Communications*, vol. 6, no. 16, pp. 2725–2733, 2012.

[19]   Z. Chen and R. C. Qiu, "Prediction of channel state for cognitive radio using higher-order hidden Markov model," in *Proc. IEEE SoutheastCon 2010 (SoutheastCon)*, Concord, NC, USA, pp. 1–21, 2010.

[20]   Z. Chen, N. Guo, Z. Hu and R. C. Qiu, "Channel state prediction in cognitive radio, Part II: Single-user prediction," in *Proc. of IEEE Southeastcon*, Nashville, TN, USA, pp. 50–54, 2011.

[21]   C. Park, S. Kim, S. Lim and M. Song, "HMM based channel status predictor for cognitive radio," in *Proc. Asia-Pacific Microwave Conf.*, Bangkok, Thailand, pp. 1–4, 2007.

[22]   L. R. Welch, "Hidden Markov models and the Baum-Welch algorithm," *IEEE Information Theory Society Newsletter*, vol. 53, no. 4, pp. 10–13, 2003.

[23] Z. Chen, N. Guo, Z. Hu and R. C. Qiu, "Experimental validation of channel state prediction considering delays in practical cognitive radio," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 4, pp. 1314–1325, 2011.

[24] H. Y. Fa and J. W. Wang, "Cooperative spectrum sensing in cognitive radio using Bayesian updating with multiple observations," *Journal of Electronic Science and Technology*, vol. 17, no. 3, pp. 252–259, 2019.

[25] J. D. Hamilton, "Time Series Analysis," Princeton, New Jersey, United States: Princeton university press, vol. 2, 1994.

[26] X. Xing, T. Jing, W. Cheng, Y. Huo and X. Cheng, "Spectrum prediction in cognitive radio networks," *IEEE Wireless Communications*, vol. 20, no. 2, pp. 90–96, 2013.

[27] G. M. F. Elias, E. M. G. Fernandez and V. A. Reguera, "Multi-step-ahead spectrum prediction for cognitive radio in fading scenarios," *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, vol. 19, no. 4, pp. 457–484, 2020.