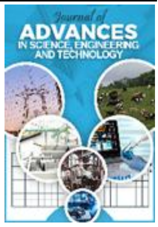
MMUST PUBLISHING PRESS

Journal of Advances in Science, Engineering and Technology



Machine Learning Based Spectrum Sensing for Interference Reduction in 5G Cognitive Radio Networks

Judith Ajega, Raymod Juma, Filbert Ombongi*

Department of Electrical and Communications Engineering, Masinde Muliro University of Science and Technology, Kakamega, Kenya

ABSTRACT

The rapid increase in global population has driven a surge in users of radio technology, leading to a shortage of available frequency spectrum for wireless systems. To optimize the use of limited spectrum, secondary (unlicensed) users can access the spectrum of primary (licensed) users when it is temporarily unused. These unused portions of spectrum are called spectrum holes or white spaces. Cognitive radios play a key role by performing spectrum sensing to detect when the spectrum is available for secondary users. Real-time spectrum detection is essential for allowing secondary users to access the spectrum without interfering with primary users. However, existing spectrum sensing methods often suffer from poor detection accuracy due to channel fading and noise. This research focuses on the design and evaluation of machine learning-based spectrum sensing algorithms for Cognitive Radio 5G networks. A hybrid sequential clustering algorithm, which combines Particle Swarm Optimization (PSO) with the K-means algorithm, is proposed. In this approach, PSO, a population-based optimization technique, determines the initial centroids and provides an optimal starting point for clustering. K-means then partitions the sensed spectrum into two clusters: occupied and unoccupied. Extensive simulations in Python were conducted to evaluate the performance of the PSO-K algorithm in various 5G network scenarios. Analysis of detection accuracy demonstrated a 9.3% improvement compared to traditional energy detection techniques.

ARTICLE INFO

Keywords:
Cognitive Radio,
5G,
Machine Learning,
Spectrum Sensing,
PSO-K,
Primary user,
Secondary user.

Article History:
Received 4 November 2024
Received in revised form 7 March 2025
Accepted 7 March 2025
Available online 8 April 2025

1. Introduction

The global population is increasing daily, resulting in a growing number of wireless communication users, as reported by Cisco (2020). As society becomes more reliant on wireless communication networks, there is a need for more efficient and dependable technology to accommodate the escalating demand for bandwidth. The valuable resource of the electromagnetic spectrum is encountering scarcity issues.

The emergence of accelerated telecommunication platforms has resulted in numerous users vying for this limited resource. This has led to a scarcity of accessible frequency slots for all users, necessitating innovative technology to address spectrum utilization issues and facilitate optimal usage by multiple users. The electromagnetic spectrum has been under intense pressure because of the rapid expansion of wireless communication technologies over the past few years (Walker, 2023; International Telecommunication Union, 2023). There is a real risk of congestion and shortage of the accessible spectrum owing to the increased use of this naturally occurring scarce resource.

Fifth-generation or 5G mobile networks, represent a significant advancement over previous generations of mobile communication technology that is 4G/LTE. It aims to provide faster speeds, lower latency, greater capacity, and more connected devices. The spectrum for 5G is divided into three main categories: low-band, mid-band, and high-band (Cadence, 2022). Frequency in the mid band range, specifically within the 3.2 GHz to 6 GHz range, are commonly used because of its balance between range and bandwidth. This band is normally crowded due to many devices operating in the same area. To solve the issue of spectrum scarcity, Mitola & Maguire (1999) initially proposed cognitive radio. CR posits that a communication device can modify its transmission parameters according to an evaluation of the state of the target frequency channel to satisfy certain performance requirements. According to (Singh,

* Corresponding author. e-mail: fombongi@mmust.ac.ke.
Editor: Francis Orata. Masinde Muliro University of Science and Technology, Kenya.
Citation: Judith A., Raymod J., & Filbert O. (2025). Machine Learning Based Spectrum Sensing for Interference Reduction in 5G Cognitive Radio Networks. Journal of Advances in Science, Engineering and Technology 1(2025), 60 – 71.

2017), the primary function of a CR is to analyze the spectrum to detect a spectrum hole and permit a secondary user, SU, to utilize it without causing interference. Spectrum holes denote the unoccupied frequency ranges within the radio spectrum (Sundrous & Alaa, 2014). If a spectrum hole exists, the CR can transmit inside the vacant frequency range until the Primary user, PU, resumes communication. A primary user is a licensed user and can use the spectrum exclusively. On the hand a secondary user is unlicensed and can only use the spectrum when it is vacant. In the work by Yucek & Arslan (2009), CR must consequently detect the spectrum and modify its broadcast settings accordingly.

The primary responsibility of the CR in a cognitive radio network, CRN, is spectrum sensing. Spectrum sensing constitutes the principal obligation of CRs within a CR network (Alhakami et al., 2014). Before initiating a transmission, it is essential to evaluate the existing spectrum occupancy and to identify any gaps within the spectrum in a certain area. The efficacy of the CR network depends on this critical role. Spectrum sensing is essential in CR networks, as it allows devices to detect and identify available spectrum bands for opportunistic utilization. A variety of strategies have been utilized to detect the presence of Primary users (PUs). Spectrum sensing based on energy detection was proposed by Gupta & Kumar (2019). In their work measuring the power of the received signal, the CR may determine whether a signal is present. By comparing the output of the energy detector with a specified threshold, spectral holes can be identified. Matched filtering detection is a spectrum sensing technique that senses the primary user when the transmitting power of the primary user is known. According to Dannana et al. (2018), when an unidentified signal corresponds to the transmission power of the signal, it is concluded that a PU is present in the spectrum. This provides a high signal-to-noise ratio for the designated input. To accomplish this, the CR must possess a comprehensive understanding of PU signal characteristics, including the operating frequency and bandwidth.

George & Prema (2019) suggested cyclostationary features as a method for feature detection. Signal statistics such as autocorrelation and mean show periodicity, which gives rise to cyclostationary properties. Periodicity is established to make spectrum detection easier. This technique uses a cyclic correlation function to detect the existence of a PU within a given spectrum. It isolates the signal from the PU and filters the background noise. The difference between the transmitted signals and noise can be used because noise signals do not have periodic features. This technique confirms the presence of the main users by analyzing the periodicity of the received signal. Chen & Nagaraj (2008) used entropy-based spectrum sensing. The entropy of a signal is a measure of the typical amount of information that is conveyed. The entropy of a signal was computed using a method based on histograms. The number of histogram bins is directly proportional to the number of energy levels in the PU signal. Spectrum occupancy is determined by computing the entropy at the CR and then comparing it with a threshold. Machine learning techniques have been used in spectrum sensing to improve the efficiency of cognitive radio networks. Muzaffar & Sharqi (2023) proposed an energy-based Machine Learning Spectrum Sensing in 5G Cognitive Radios. Ali & Hamouda (2016) proposed a naïve Bayes classifier to sense spectrum holes. Using signal eigenvalues and a clustering approach, (Wang et al., 2018) presented a spectrum-sensing method. This approach trains a classifier using signal eigenvalues and a K-means clustering technique and then uses the classifier to determine if the main user signal is there. A K-means clustering based blind multiband spectrum sensing algorithm for cognitive radio was proposed by Lei et al. (2018). Fouda et al. (2024) suggested a weighted joint likelihood ratio test for cooperative spectrum sensing using K-means clustering.

Kumar et al. (2016) suggested an approach to spectrum sensing that relies on k-means clustering. His approach used a K-means clustering technique to categorize features such as signal energy into two groups: those whose channels were available and those whose channels were unavailable. Arjoune & Kaabouch (2019) proposed an approach. This approach proposed labelling a dataset using K-means clustering, which divides the received PU signal into two categories: present and absent. In their work, Several ML algorithms, including K-Nearest Neighbors (KNN), support vector machines (SVM), logistic regression (LR), decision tree (DT), and random forest (RF), were used to classify the received signals into one of two classes after they were split into them. Gupta et al. (2021) suggested evolutionary algorithm for spectrum sensing in 5G cognitive radio networks. Besides, the particle swarm optimization (PSO) algorithm has shown very effective performance in CRN. Authors in Gul et al. (2021) suggested a robust spectrum sensing against malicious users using particle swarm optimization.

2. Materials and Methods

In this study, PSO-K means algorithm was employed to detect spectrum availability. The PSO-K Means technique is a hybrid method that integrates PSO with K-Means clustering to improve the clustering efficacy. This study employed the Particle Swarm Optimization-based K-means (PSO-K) technique to ascertain the existence or absence of a PU signal in spectrum sensing. The methodology has two principal elements: an optimization algorithm and a clustering algorithm. Energy-based feature extraction was employed to improve the detection procedure. Fig. 1 presents a detailed flow chart describing the operation of the PSO-K Means algorithm, accompanied by examples for each scenario.

This hybrid strategy utilizes the global optimization strengths of PSO to identify effective initial cluster centroids, which are subsequently refined by the local search skills of the K-means algorithm. K-means is a common clustering approach that partitions a dataset into K clusters. Each data point was allocated to the cluster corresponding to the nearest mean (centroid), and the centroids were successively refined until convergence was achieved. The objective of k-means clustering is to minimize the aggregate of squared distances between each data point and its respective cluster centroid, referred to as the within-cluster sum of squares (WCSS).

2.1. Data Generation

First, the process begins by defining the spectrum environment. The frequency band for spectrum sensing is specified. In this study, a 5G mid-band frequency was selected, with frequencies ranging from 3.8 GHz 7 GHz. Then, the signal characteristics were determined. The PU transmission power level was observed as 20dBm. Orthogonal frequency-division multiplexing (OFDM) was used as the modulation technique, and Additive White Gaussian Noise was assumed in the channels. Second, PU Activities were simulated. This was accomplished by creating models for spectrum utilization by PUs. The model involved periodic usage patterns in the spectrum. The patterns provided details of the periods of activity and inactivity (spectrum holes). Synthetic signals were then generated for the PUs considering their modulation schemes, power levels, and signal characteristics. Python was used to generate the signals because it has rich libraries and packages integrated with signal processing. Simulations were performed to introduce SUs to potentially transmit simultaneously with PUs.

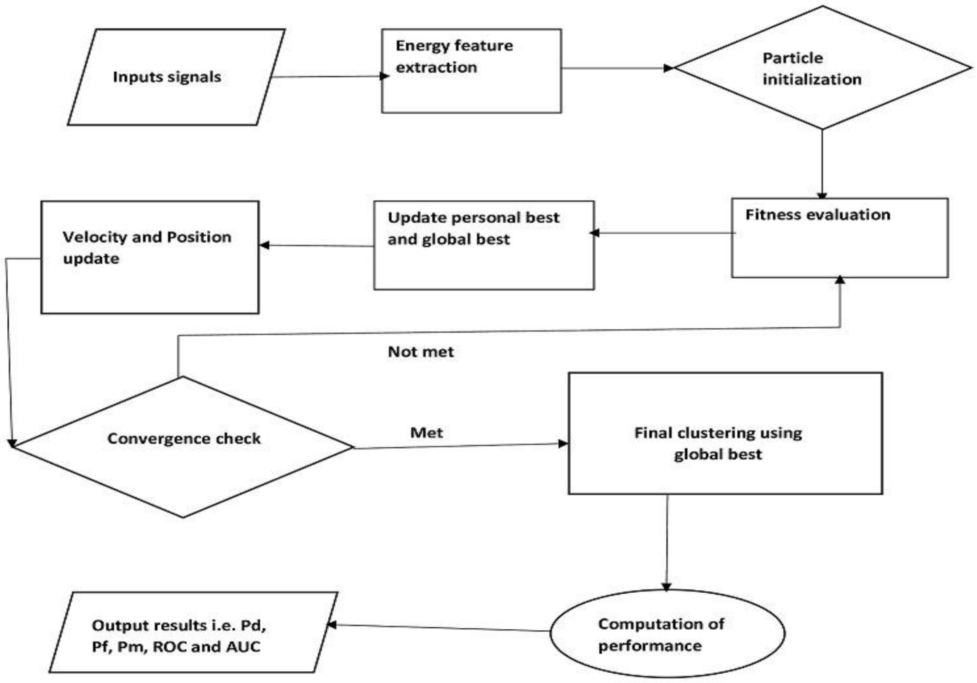


Fig. 1: Methodology flow chart

2.2. Data Collection

Information about spectrum usage, signal characteristics, and network performance was acquired. Parameters such as the signal strength, frequency bands in GHz, and modulation types were considered. Simulations were performed in Python to model the spectrum-usage scenarios. Parameters for data collection were then set up, and Python was configured to run multiple scenarios with varying signal-to-noise ratios (SNR) and different threshold levels. Simulations were performed as per the design, and the data were collected and recorded. The simulated results were then compared with theoretical values to ensure the validity of the simulation.

2.3. Feature Extraction

ML-based feature extraction was also employed. First, data collected was taken, and it included both PU's transmission and any noise present in the spectrum. Relevant features were identified and extracted from the dataset. Energy levels were used as features derived from the various characteristics of each frequency range. These features were then refined. Subsequently, the average energy levels were computed. Eq. 1 shows the computation of these features, where $(Zn(k))^2$ is the energy of the k^{th} sample.

2.3.1. Initialization

The number of particles used was 1000. The number of clusters (k) was then determined and two clusters were obtained. The first cluster represents PU presence, and the second cluster represents PU absence. The number of iterations was also determined and 300 iterations were used. The iterations represent the maximum number of generations performed before converging to the optimal solution. Position Initialization was performed, where each particle in the swarm was assigned an initial random position in the search space. This position represents a potential set of centroids in the k-means clustering algorithm. The positions were defined randomly based on the range of the signal`s energy levels. In addition to these positions, each particle was assigned an initial velocity.

$$En = \frac{1}{K}\sum_{K=1}^{K}(Zn(k))^2 \quad (1)$$

Velocities were initialized randomly but within a range to ensure that the particles effectively explored the search space without exceeding the bounds. The position and velocity of the particles were determined using Eq. 2 and Eq.3 respectively.

$$V_{id}^{t+1} = w\ x\ v_{id}^t + c_1 r_1 (p_{id}^t - x_{id}^t) + c_2 r_2 (p_{gd}^t - x_{1d}^t) \quad (2)$$

$$x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1} \quad (3)$$

where w is the inertia weight, x_{id}^t and v_{id}^t represents the position and velocity of the particle i in the d dimensional space, at time t. c_1 is the self-learning factor, while c_2 is the group learning factor, r_1 and r_2 are random numbers ranging from 0 to 1. p_{id}^t is the particle best value, or local best, and p_{gd}^t is the swarm's best value or global best.

2.3.2. Fitness Evaluation

This process evaluates the extent to which a particular solution solves a problem. The fitness of each particle was evaluated based on the extent to which centroids were represented in the cluster of the dataset. The goal of K-means clustering was to minimize the WCSS, leading to tightly packed clusters with low variance. The WCSS for each particle in the clusters was calculated using Eq. 4.

$$WCSS = \sum_{k=1}^{k}\sum_{x_1 \in k}\left[\frac{(x_1 - \mu_k)^2}{1}\right] \quad (4)$$

where k is the number of clusters; x_1 represents the data points; μ_k is the centroid of the cluster c_k, and $x_1 - \mu_k$, is the Euclidean distance between the data point x_1 and centroid μ_k. A lower WCSS indicates that the data points within a cluster are close to each other and to the centroid, implying better-defined clusters. Euclidean distance is a distance metric used in the K-means algorithm to calculate the distance between data points and centroids. The Euclidean distance between two points $x = (x_1, x_2, \ldots., x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ in n-dimensional space is given by Eq.5. where x_1 and y_1 represent the coordinates of the two points in the i-th dimension and $(x_1 - y_1)^2$ is the squared difference between the coordinates of the two points. The K-means algorithm determines the optimal clustering strategy by minimizing the Within-Cluster Sum of Squares (WCSS) for each data point. Minimizing the WCSS distance, which is calculated as the sum of the squared distances between the points and their respective centroids, requires decreasing the Euclidean distances within each cluster.

$$d(x,y)\sqrt{\sum_{i=1}^{n}(x_1 - \mu_k)^2} \quad (5)$$

The fitness of each particle was then calculated based on the energy levels of the corresponding frequency bands per work by Ratanavilisagul (2020). Eq. 6 shows the fitness function.

$$Fitness = \frac{max\ d_1 y_i}{min\ d_2 y_i} \quad (6)$$

where $max\ d_1 y_i$ is the maximum value of the average values of distances within the same classes in the classification plan shown by particle y_i. $Min\ d_2 y_i$ is the minimum value of the distances between classes in the classification plan shown by particle y_i, m_j is class j and, m_i is class i. For clusters considering the Euclidean distance and WCSS, fitness was expressed in Eq. 7.

$$Fitness = \frac{1}{WCSS} = \sum_{k=1}^{k}\sum_{x_1 \in k}\left[\frac{(x_1 - \mu_k)^2}{1}\right] \quad (7)$$

where k denotes the number of clusters, x_1 denotes the data point, and μ_k denotes the centroid. The algorithm maximizes the fitness value by minimizing WCSS. Particles with higher fitness values were required. The algorithm adjusts the positions and velocities of the particles to explore the search space and find the best centroids that minimize the WCSS.

2.3.3. Initialization of Personal Best

Personal best (p_{best}) refers to the best solution that an individual particle finds during the search process. The personal best fitness for each particle was set as the fitness value obtained from the initial fitness evaluation. The personal best represents the best set of cluster centroids that the particle identified for classifying the data. The personal best is used in the velocity update equation, which determines how each particle moves in the search space.

2.3.4. Initialization of Global Best

The global best (g_{best}) is the best position found by any particle in the swarm during the optimization process. Initially, g_{best} is set to the position of the particle with the best fitness value among the initial p_{best} values. The fitness of g_{best} was then evaluated using the objective function, and it was updated as the algorithm was iterated. The global best served as a reference point for all the particles in the swarm. Each particle is influenced by both its personal best and global best when updating its position.

The effect of the personal best and global best helped particles converge towards the optimal regions of the solution space. The global best corresponded to the set of cluster centroids that had produced the best clustering results at that time, meaning the most accurate separation of occupied and unoccupied spectrum bands, were achieved, leading to optimal spectrum sensing performance

2.3.5. Position and Velocity Update

The velocity update mechanism governs how particles in the swarm move through the solution space in search of the optimal cluster centroids for spectrum sensing. Each particle represents a potential solution. The velocity was influenced by both the personal best position of the particle and the global best position found by the swarm. The velocity of a particle was updated using Eq. 8.

$$v_i^{t+1} = w \; x \; v_i^t + r_1 \; x \; c_1 \; x \; (p_{best} - x_i^t) + c_2 \; x \; r_2 \; x \; (g_{best} - x_i^t) \quad (8)$$

where v_i^{t+1} denotes the velocity of particle i at iteration $t + 1$. v_i^t is the velocity of particle i at iteration t. w is the inertia weight, which controls the effect of the previous velocity on the current velocity; c_1 and c_2 are the acceleration coefficients, controlling the influence of the personal best and global best elements; and $r_1$and $r_2$are random numbers that are distributed uniformly between 0 and 1. They introduced randomness into the movement of particles. p_{best} refers to the personal best position of particle i and g_{best} is the global best position of the particle found by the algorithm. Where, $x_i t$ is the current position of particle i at iteration t.

Once the velocities of the particles are updated, the positions of the particles are updated. The position of each particle was updated by adding the updated velocity to the current position. Eq. (9) was used to calculate the positions of the particles.

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (9)$$

Where x_i^{t+1} is new position of the particle i. Each particle's position represents a potential set of cluster centroids that define the boundaries between the occupied and unoccupied spectrum states. The velocity update mechanism allows particles to explore various configurations of these centroids, guided by their personal best and the swarm's global best, leading to the optimal detection of spectrum holes or occupied bands.

The velocity update mechanism in the PSO-K algorithm plays a critical role in guiding the swarm of particles towards the optimal solution. It balances the new solutions and the known good solutions, ensuring that the particles converge to an optimal set of cluster centroids for better spectrum sensing. This process is key to enhancing the accuracy and efficiency of spectrum detection, ultimately leading to improved spectrum utilization.

2.3.6. Iteration

Iteration refers to one complete cycle through the algorithm process, where the positions and velocities of all particles in the swarm are updated. The application of the PSO-K algorithm steps was repeated until convergence was achieved. During each iteration, the particles move through the solution space based on their current velocity, which is influenced by their personal best positions and the global best position found by the swarm. After all particles were initialized with positions, velocities, p_{best}, and g_{best}, the algorithm began the iterative process. The iteration counter was initially set to zero. The velocity of each particle was updated based on its current velocity, the difference between its current position and its personal best, and the difference between its current position and the global best. This step guides the particles toward better solutions by balancing exploration and exploitation in the search space. The iterative process involved updating the particles' positions and velocities, adjusting their p_{best}, and potentially updating the g_{best}, all aimed at converging at an optimal convergence point. Once the velocities are updated, the particle positions (which correspond to the cluster centers) are also updated. This new position represents new potential clustering of data. The algorithm is iterated until convergence is achieved. During each iteration, the particle positions and velocities were updated, their fitness was evaluated, and the best solution found by the swarm was also updated.

Table 1: Simulation parameters

S/N	Parameters	Description	Values
1	Algorithm	Algorithm type	PSO-K means Algorithm
2	Swarm size	Number of particles in PSO	1000
3	Clusters	Number of clusters in k-means	2
4	Iterations	Maximum number of PSO iterations	300
5	Distance Metrics	Type of distance metric	Euclidean
6	Channel type	Type of channel model	AWGN
7	Signal to Noise ratio range	Signal to noise ratio	0-20dB
8	Detection threshold	Variable	01-09
9	Frequency band	Frequency range	FR-1 frequency range
10	Probability of detection	Ideal value of detection probability	1
11	Probability of false alarm	Ideal value of false alarm probability	0
12	Probability of missed detection	Ideal value of missed detection probability	0
13	AUC	Area under the curve for ideal system	1

2.3.7. Convergence

Convergence is the point at which the algorithm's particles stop making significant improvements or changes in their positions, indicating that it is optimal. Convergence in the PSO-K Means algorithm occurred when the particles in the swarm found an optimal solution, and further iterations could not significantly improve the solution. Several criteria were used to determine convergence. Fitness Stabilization means that there is an improvement in fitness across successive iterations, which becomes negligible. This indicates that the particles had converged to a solution and were no longer making significant progress in finding a better solution. As the particles converged towards the global best, their velocities tended to decrease. The reduction in velocity across the swarm indicates that the particles converged to a common solution and that the search space exploration was near completion.

After convergence, the final positions of the centroids after iterations are used to classify the data points into clusters. Clustering was then performed on the set of frequency bands to group them into clusters based on their energy level similarities. The K-means algorithm clusters the particles into two groups based on their positions in the search space. Eq.10 was used.

$$d(z_p, m_j) = \sqrt{\sum_{k=1}^{n_d} (Zpk - mjk^2)} \quad (10)$$

where nd is the number of attributes, z_p is the object, Zpk is attribute k in object p, mj is class j, mjk is attribute k in class j, and $d(z_p, m_j)$ is the distance between object p and class j. Each class size was calculated using Eq. 11. where $z_p$is the object.

$$m_j = \frac{1}{n_j} \sum_{zp \in cj} Zp \quad (11)$$

2.4. Spectrum Decision and Performance Measurement

The final clustering result was then evaluated to make decision to determine the clusters that represent the occupied frequency bands and those that represent unoccupied frequency bands. After the convergence of the PSO-K Means algorithm and clustering, performance metrics were measured. These metrics typically include the probability of detection (Pd), probability of false alarms (Pf), and probability of missed detection (Pm). Cluster labelling was performed, where one cluster represented the occupied state and the other represented a vacant state. The algorithm labels these clusters based on centroid positions, where one cluster typically represents the presence of a PU (occupied spectrum) and the other represents the absence (vacant spectrum). The independent variables were used to achieve resignation. Table 1 presents the key variables used in the simulation of the algorithm.

To evaluate the performance, the results from the algorithm were compared with a ground-truth dataset, which contained the actual states of the spectrum bands (whether they were occupied or vacant). This ground truth is essential for calculating the performance metrics. True Positives (TP) refer to the number of correctly detected occupied spectrum bands. False Positives (FP) refer to the number of incorrectly detected occupied spectrum bands (actually vacant). True Negatives (TN) refer to the number of correctly detected vacant spectrum bands. False Negatives (FN) refer to the number of missed detections where the spectrum is occupied, but the algorithm classifies it as vacant. These metrics were then used to measure the performance of the algorithm in terms of the probability of detection, probability of false alarms, probability of missed detection, and detection accuracy. The probability of detection is also known as the True Positive Rate (TPR) or the sensitivity, calculated using Eq. 12.

$$Pd = \frac{TP}{TP + FN} \quad (12)$$

The probability of a false alarm is also known as specificity (false positive rate). Mathematically, this is represented by Eq. 13 which represents the ratio of falsely detected occupied bands to all truly vacant bands.

$$Pf = \frac{FP}{FP + TN} \quad (13)$$

The probability of missed detection values were calculated using Eq. 14.

$$Pm = \frac{FN}{TP + FN} \quad (14)$$

Probability of missed detection can also be given as $Pm = 1 - Pd$.

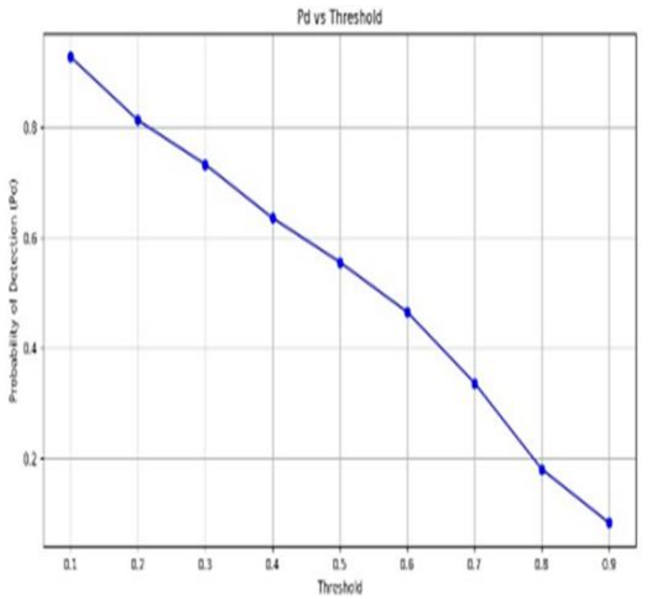


Fig. 2: Detection probability vs threshold

3. RESULTS AND DISCUSSION

Clustering was performed, and performance metrics were measured to analyze the performance of the algorithm. The probability of detection is a performance metric used to measure the performance of PSO-K means to reduce inter-user interference.

3.1. Probability of Detection

Fig. 2 shows a graph of Pd against the detection threshold. Increasing the detection threshold lowers Pd, and vice versa. A lower detection threshold makes the system more sensitive, meaning that it is more likely to detect even weak signals from the PU. A higher threshold makes the system less sensitive, making it possible to miss weaker signals from the PU and leading to a lower detection probability. It is usually very important to maintain the detection threshold at a value that gives a higher Pd; in this study, the detection threshold is at a value that gives a high Pd to allow accurate sensing while minimizing interference between users.

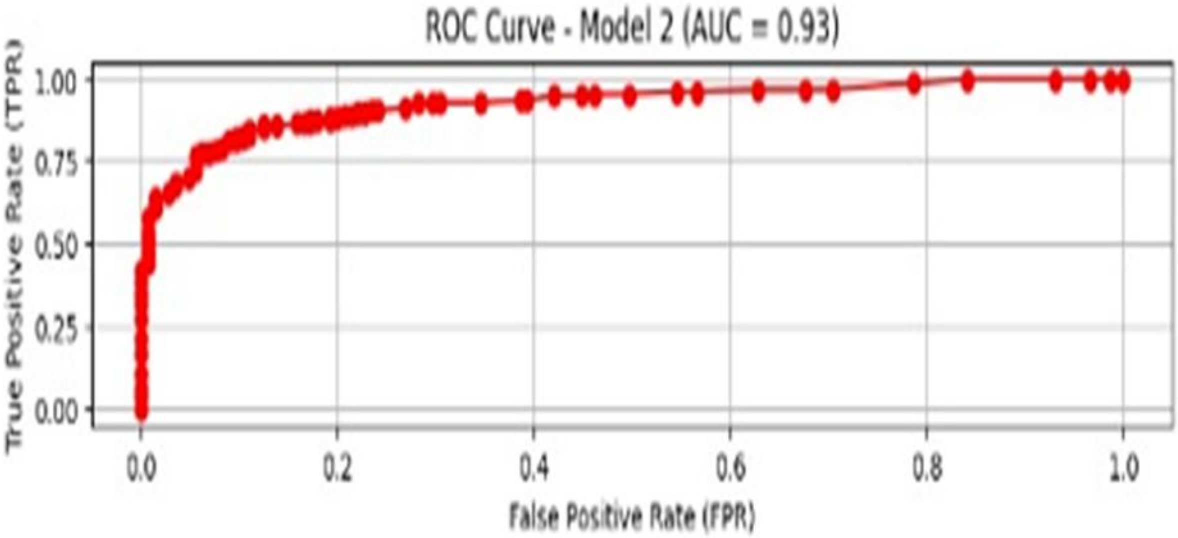


Fig. 3: Receiver operating characteristics curve for PSO-K

3.2. Receiver Operating Characteristic Curve

Several previous studies (Aloqlah, 2014; Musuvathi et al., 2024) suggested that the detection probability, which is also known as the true positive rate, shows the likelihood of CR correctly detecting the presence of a PU in the spectrum. The PSO-K means algorithm achieved a high detection probability. In Fig. 3, the receiver operating characteristic curves show the true positive rate on the x-axis and the false positive rate on the y-axis. It is clear that Pd has a value of 0.93 or 93%, which shows that the proposed technique detects the presence of PU by 93% and indicates a robust spectrum sensing capability. This is an improvement over traditional energy detection, which typically yields a Pd value of approximately 0.7. An ideal ROC curve for spectrum sensing has an AUC of 1 or 100%. In previous studies, the typical values for Pd range between 0.7 and 0.9. The improvement of 0.93 in this study suggests that the proposed method is more effective in detecting PU signals.

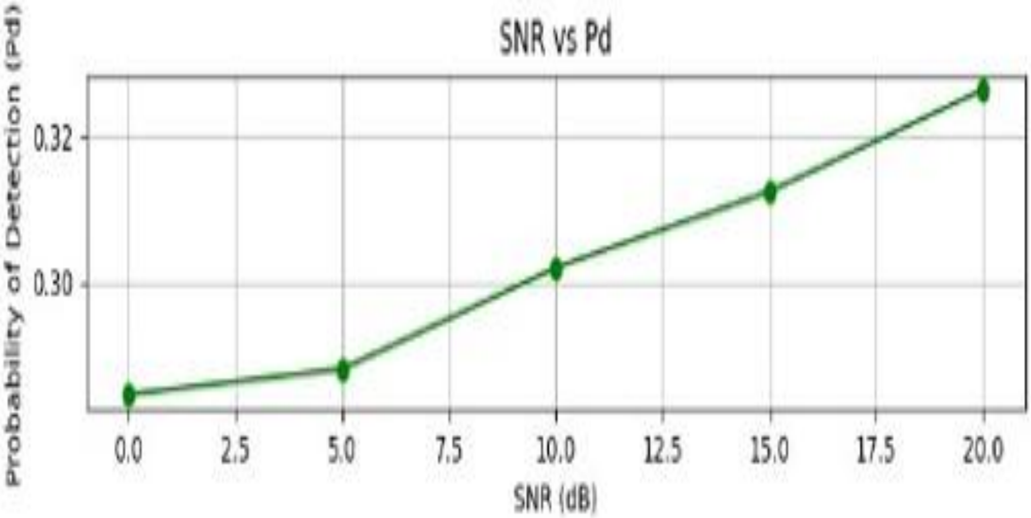


Fig. 4: Graph of detection probability vs SNR

ROC curve representation was used to evaluate the performance of the detection system. This shows the trade-off between the system's sensitivity and its specificity (false positive rate) with varying detection thresholds. The TPR or Pd is the proportion of actual positives (presence of a signal) correctly identified by the system. Normally, it is plotted on the y-axis of the ROC curve. The Pf also known as Specificity, is the proportion of actual negatives (absence of a signal) incorrectly identified as positive by the system. It is plotted on the x-axis of the ROC curve.

Noisy environments affect spectrum sensing. The signal-to-noise ratio (SNR) is a vital metric in communication networks. An evaluation of the channel quality is provided. When properly implemented, spectrum sensing can reliably provide accurate results. Noise, as well as the effects of fading and shadowing on the channel in question, make this practically impossible. A considerable detection probability was achieved across various SNR levels using the PSO-K means approach. Figure 4 shows the correlation between the signal-to-noise ratio and the detection chance. Gains in the signal-to-noise ratio are directly proportional to the increase in the detection probability. From 0 dB to 5 dB, the detection probability (shown in Fig. 4) is very low; nevertheless, it increases exponentially from 5 dB to 20 dB, as shown in Fig. 4. This trend demonstrates that PSO-K indicates the algorithm's sensitivity to signal quality, exhibiting the best performance in settings with an elevated SNR. The suboptimal detection performance in a low-SNR channel indicates that this approach is adversely influenced by the channel conditions, preventing the SUs from capitalizing on all available transmission chances.

This high Pd minimized the risk of interference between the secondary and the PU. In a 5G CR network, the SU seldom interferes with the primary network, leading to better performance. This suggests that the algorithm can effectively identify spectrum holes, thereby allowing for the optimal use of the available spectrum.

3.3. Probability of False Alarm

Table 2 shows the trade-off between the probabilities of detection and false alarm. There is a relationship between the probability of detection, probability of false alarms, and threshold. When the detection threshold was adjusted from 0.2 to 0.5, there was a noticeable decrease in the false alarm rate from 0.79 to 0.5, at the expense of a slight reduction in the detection probability. This trade-off highlights the importance of carefully tuning the detection threshold to balance detection accuracy and false alarms.

As shown in Table 2, lowering the threshold lowers the probability of detection and the probability of false alarm, whereas increasing the threshold increases the probability of detection and the probability of false alarm.

Table 2: Probability of detection, false alarm rate and detection of threshold

Threshold	Pd	Pf
0.1	0.93	0.88
0.2	0.81	0.79
0.3	0.73	0.69
0.4	0.64	0.61
0.5	0.56	0.50
0.6	0.47	0.42
0.7	0.34	0.29
0.8	0.18	0.21
0.9	0.08	0.10

A lower detection threshold makes the detector very sensitive, implying that even small signals are detected. However, this increases the likelihood of false alarms when the noise is mistaken for a signal. This leads to a higher probability of false alarms. However, a higher threshold reduces the sensitivity of the detector, which implies that only stronger signals are detected. This reduction in the probability of false alarms is critical, as it minimizes unnecessary spectrum sensing interference and improves the overall efficiency of spectrum utilization.

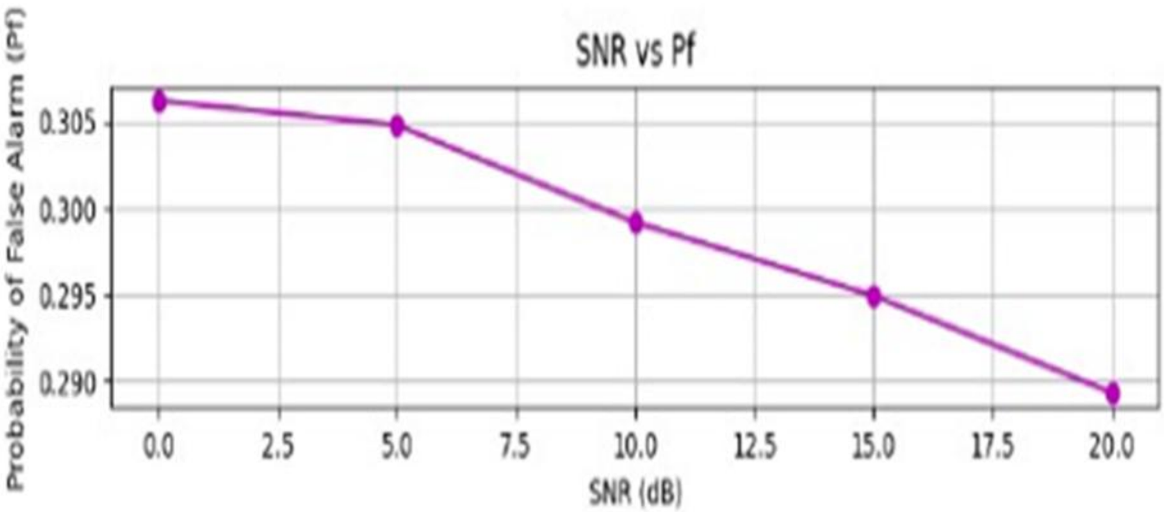


Fig. 5: Graph of False alarms probability vs SNR

3.4. Signal to Noise Ratio

Fig. 5 is a graph of signal to noise ratio versus false alarm probability. It illustrates that the chance of a false alarm is elevated at low signal-to-noise ratio (SNR) values and diminished at higher SNR values. Between SNR 0.0 and 2.5, the Pf is higher, but the vale starts to decrease at snr 5 Db. At high signal-to-noise ratio, the PSO-K algorithm accurately identifies the presence of a principal user. The performance of the spectrum-sensing algorithm is evaluated at different SNR levels. The ML-based approach maintained a high Pd value across a wide range of SNRs, demonstrating better resilience in noisy environments. According to the literature, traditional methods often struggle at low SNRs, with significant reductions in Pd. The improved performance at low SNRs in this study indicates that the ML approach is more effective in challenging environments, making it suitable for real-world applications in 5G networks.

3.5. Probability of Missed Detection

Missed detection is a barrier to spectrum sensing because it allows the SU to interfere with the primary signal. Fig. 6 illustrates that the chance of missed detection is elevated at low SNR values; however, as the SNR value increases, the Pm decreases. Generally, the value of PM decreases as the SNR value increases.

A plot of the probability of missed detection versus threshold is presented in Fig. 7. It is seen that raising the detection threshold increases the probability of missed detection, whereas lowering it decreases the probability of missed detection. Lowering the detection threshold increases the sensitivity of the system to detect weak signals and leads to a reduction in the probability of missed detection, because the system is more likely to detect signals that are weak or close to the noise level, which in turn reduces the probability of missed detection. Conversely, a higher detection threshold decreases detector sensitivity, meaning that only stronger signals are detected. This increases the probability of missed detection because the system is less likely to detect weaker signals. In this study, the detection probability is high and ensures that the probability of missed detection is kept low.

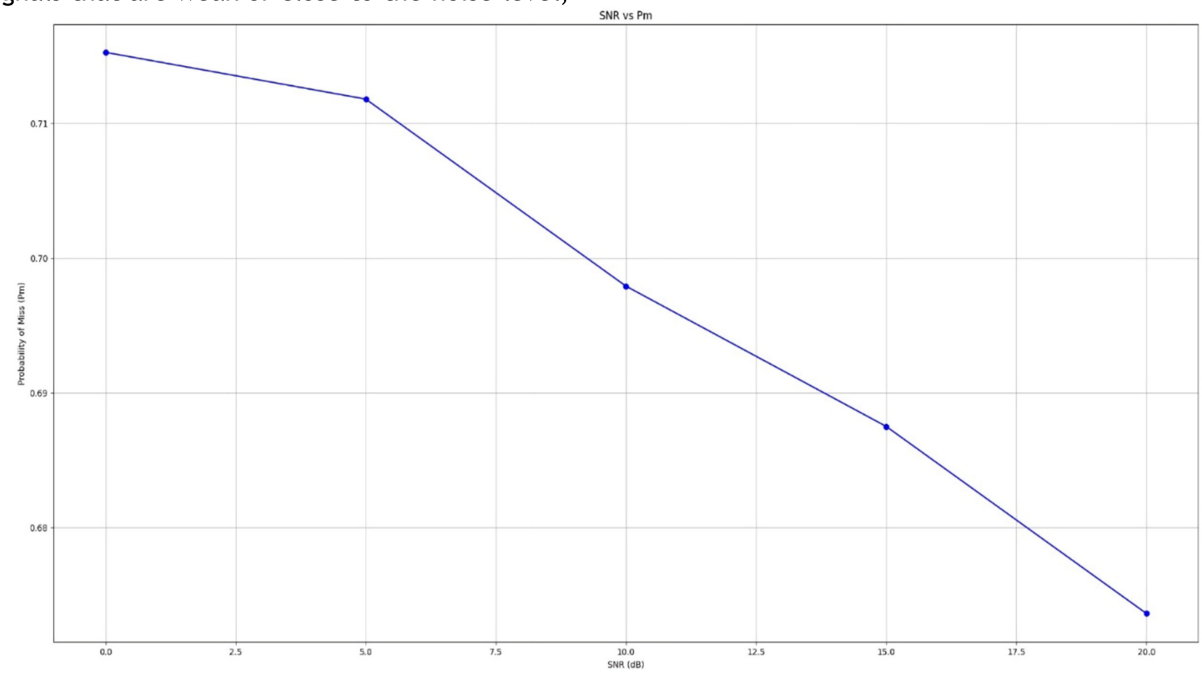


Fig. 6: Probability of Missed detection vs SNR

Traditional methods have a higher probability of missed detection, particularly for low SNRs. The improved performance at low SNRs in this study indicates that the ML approach is more effective in challenging environments, making it more suitable for real-world applications in 5G networks. In 5G CR networks, the proposed algorithm will ensure interference with PUs is highly minimized. The proposed technique achieved a probability of missed detection of 0.08, indicating an improvement over traditional methods, where the probability of missed detection typically gives a value of 0.2. This lower probability of missed detection indicates that the ML-based approach is more reliable for identifying spectrum occupancy, which is crucial for avoiding harmful interference with PUs in 5G CR networks.

3.6. Sensing accuracy

Sensing accuracy in CR networks refers to the ability of a CR system to correctly detect the presence or absence of a PU in a spectrum band. Accurate spectrum sensing is very important in CRNs because it directly affects the efficiency of the network, reduction of interference between the primary and SU, and overall utilization of the available spectrum. To compare the performance of the proposed detectors, we evaluated the detection accuracies of the different detectors in a 5G CR network environment.

Sensing accuracy is important for the following reasons. Firstly, it protects the PU. One of the main aims of the CR network is to ensure that SUs do not interfere with PUs. High sensing accuracy helps protect PUs by reliably detecting their presence and vacating the spectrum when PUs want to transmit. Second, it increases the efficiency of the spectrum utilization. Accurate spectrum sensing allows for better utilization of the spectrum by SUs.

Table 3: Accuracy comparison table

Algorithm	Detection accuracy
Logistic Regression	0.85
Random Forest	0.86
Energy Detection	0.55
PSO-K means	0.94

The key components of sensing accuracy involve the performance metrics for spectrum sensing, that is, the probability of detection, probability of false alarm, probability of missed detection, and receiver operating characteristic curve, which provides a graphical representation of the sensing accuracy. The roc curve also summarizes the overall sensing accuracy, where an AUC of 1 indicates perfect sensing accuracy. When spectrum vacancies are accurately detected, SUs can opportunistically access them without causing interference to the PU, thereby improving the overall spectrum efficiency. Finally, the accuracy improves network performance. In a 5G CR environment, where the demand for spectrum is high and the available spectrum is limited, accurate sensing is crucial for maintaining the network performance. It ensures that SUs can access the spectrum only when it is available.

Kumar (2018) analyzed spectrum sensing using an energy detection method. Also, in Atapattu et al. (2010), it was analyzed in area under the receiver operating curve. From their analysis we see that the accuracy values for energy detection method was 0.5, or 50%. The authors performed a comprehensive analysis of the performance of the logic regression algorithm in spectrum sensing based on the detection accuracy. It was observed that the logic regression gave a detection accuracy of 0.85 or 85% when the analysis was performed on 5G CR networks. In a recent work Gao & Wang (2021), a spectrum-sensing algorithm was proposed, based on random forest. The performance of the algorithm was analyzed, and the energy levels of the received signals were used as features. In their study, the detection rate was found to be 0.86, which was 86%.

In this study, the PSO-K algorithm produced high detection accuracy values. The detection accuracy when measuring true negatives and false positives gives the PSO-K means algorithm an accuracy of 0.94, or 94%. These results show an improvement of 9.3% in the detection accuracy. This implies that the application of the PSO-K means algorithm in spectrum sensing in 5G CR networks will lead to reduced interference on primary and secondary user in the CR network, as well as enhance network performance and improve spectrum utilization in CR networks. Table 3 shows a comparison between PSO-K-means and other methods of spectrum sensing. The PSO-K means algorithm demonstrated higher accuracy and better performance than the other spectrum detection methods. Table 3 presents the results.

3.7. Implications for 5G Cognitive Radio Networks

The results obtained from the proposed ML-based spectrum sensing algorithm show clear improvements over traditional methods in several key performance metrics, including the detection probability, false alarm rate, and missed detection probability. The ROC curve and AUC further confirmed the superiority of the proposed method in accurately detecting PU signals while minimizing interference in 5G CR networks. These improvements are crucial for enhancing spectrum efficiency and reducing interference, making the proposed approach highly suitable for deployment in 5G networks. The enhanced spectrum sensing technique provided by the PSO-K means algorithm contributes to a more efficient spectrum utilization and reduced interference in 5G CR networks. This could lead to more reliable communication and better service quality in places where spectrum resources are heavily contested.

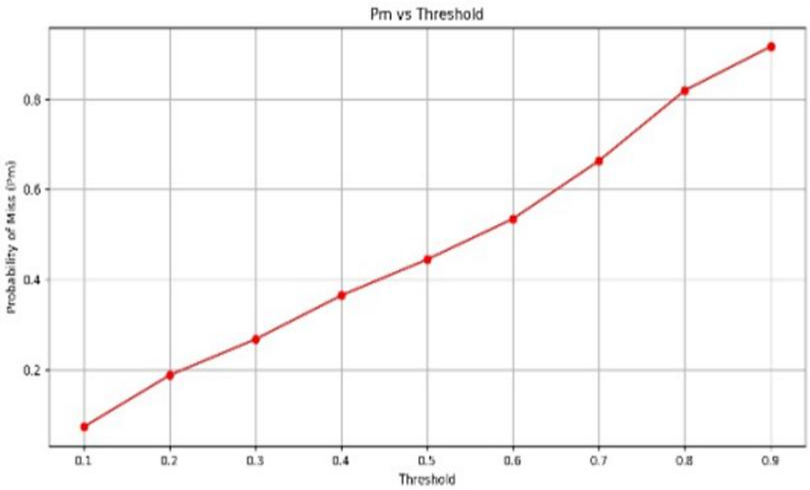


Fig. 7: Probability of Missed detection vs threshold

4. Conclusion

This study introduces a hybrid method that combines particle swarm optimization with the k-means clustering algorithm to enhance the spectrum sensing accuracy in CR networks and mitigate interference between the primary and SUs. This methodology, which integrates the characteristics of the PSO and K-means algorithms, demonstrates a robust and efficient solution for spectrum sensing in CR networks, ensuring accurate detection performance and maximizing spectrum use. The principal aim of this research was to devise and assess an ML-based methodology to enhance spectrum sensing in 5G CR networks, concentrating on minimizing interference and augmenting the 5G spectrum usage. The incorporation of the PSO-K Means algorithm significantly enhances the detection probability while decreasing both the false alarm and missed detection probabilities compared with conventional energy detection methods. Analysis of the detection accuracy of the PSO-K means revealed an enhancement of 9.3% in accuracy. The findings indicate that utilizing ML techniques, specifically the PSO-K Means algorithm, can improve the spectrum utilization efficiency in 5G CR networks. This may result in enhanced communication reliability, diminished interference, and augmented network capacity, thereby rendering the proposed technology a viable solution for next-generation wireless networks. This study enhances the existing knowledge on CR networks by presenting an optimal ML strategy that addresses the issues associated with dynamic spectrum availability. The proposed technique improves detection accuracy and provides a scalable solution for real-time spectrum sensing in 5G scenarios.

References

Alhakami, W., Mansour, A., & A, G. (2014). Spectrum Sharing Security and Attacks in CRNs: a Review. International *Journal of Advanced Computer Science and Applications*, 5(1). https://doi.org/10.14569/ijacsa.2014.050111.

Ali, A., & Hamouda, W. (2016). Advances on Spectrum Sensing for Cognitive Radio Networks: Theory and Applications. *IEEE Communications Surveys & Tutorials*, 19(2), 1277–1304. https://doi.org/10.1109/comst.2016.2631080.

Aloqlah, M. (2014). Performance analysis of energy detection-based spectrum sensing in κ-μ shadowed fading. *Electronics Letters*, 50(25), 1944–1946. https://doi.org/10.1049/el.2014.2944.

Arjoune, Y., & Kaabouch, N. (2019). A comprehensive survey on spectrum sensing in cognitive radio networks: recent advances, new challenges, and future research directions. *Sensors*, 19(1), 126. https://doi.org/10.3390/s19010126

Atapattu, S., Tellambura, C., & Jiang, H. (2010). Analysis of area under the ROC curve of energy detection. *IEEE Transactions on Wireless Communications*, 9(3), 1216–1225. https://doi.org/10.1109/twc.2010.03.091085

Cadence. (2022). An Overview of Frequency Bands and Their Applications. Cadence PCB solutions.

Chen, N. X., & Nagaraj, S. (2008). Entropy based `spectrum sensing in cognitive radio. *Wireless Telecommunications Symposium*, 57–61. https://doi.org/10.1109/wts.2008.4547544

Cisco. (2020). Cisco Annual Internet Report (2018–2023) White Paper.

Dannana, S., Chapa, B. P., & Rao, G. S. (2018). Spectrum sensing using matched filter detection. In *Advances in intelligent systems and computing* (pp. 497–503). https://doi.org/10.1007/978-981-10-7566-7_49.

Fouda, S., Farghaly, I., & Dawood, S. (2024). Weighted joint LRTs for cooperative spectrum sensing using K-means clustering, Vol. 67. *Physical Communication*. https://doi.org/10.1016/j.phycom.2024.102528.

Gao, Z., & Wang, X. (2021). Spectrum sensing algorithm based on random forest in dynamic fading channel. In *Advances in intelligent systems and computing* (pp. 39–46). https://doi.org/10.1007/978-981-33-4575-1_5.

George, G. R., & Prema, S. C. (2019). Cyclostationary feature detection based blind approach for spectrum sensing and classification. *Radioengineering*, 27(1), 298–303. https://doi.org/10.13164/re.2019.0298.

Gul, N., Ahmed, S., Najeebullah, N., Kim, S. M., & Kim, J. (2021). Robust Spectrum Sensing Employing PSO. *Twelfth International Conference on Ubiquitous and Future Networks* (ICUFN), 58, 461–466. https://doi.org/10.1109/icufn49451.2021.9528553.

Gupta, M. S., & Kumar, K. (2019). Progression on spectrum sensing for cognitive radio networks: A survey, classification, challenges and future research issues. *Journal of Network and Computer Applications*, 143, 47–76. https://doi.org/10.1016/j.jnca.2019.06.005

Gupta, V., Beniwal, N. S., Singh, K. K., Sharan, S. N., & Singh, A. (2021). Optimal cooperative spectrum sensing for 5G cognitive networks using evolutionary algorithms. *Peer-to-Peer Networking and Applications*, 14(5), 3213-3224. https://doi.org/10.1007/s12083-021-01159-6

International Telecommunication Union, I. (2023). Assessment of spectrum efficiency and economic

Kumar, S. (2018). Performance of ED Based Spectrum Sensing Over α-η-μ Fading Channel. *Wireless Personal Communications*, 100(4), 1845–1857. https://doi.org/10.1007/s11277-018-5677-6

Kumar, V., Kandpal, D. C., Jain, M., Gangopadhyay, R., & Debnath, S. (2016). K-mean clustering based cooperative spectrum sensing in generalized κ-μ fading channels. 2016 *Twenty Second National Conference on Communication* (NCC). https://doi.org/10.1109/ncc.2016.7561130

Lei, K., Tan, Y., Yang, X., & Wang, H. (2018). A K-means clustering based blind multiband spectrum sensing algorithm for cognitive radio. *Journal of Central South University*, 25(10), 2451–2461. https://doi.org/10.1007/s11771-018-3928-z

Mitola, J., & Maguire, G. (1999). Cognitive radio: making software radios more personal. *IEEE Personal Communications*, 6(4), 13–18. https://doi.org/10.1109/98.788210

Musuvathi, A. S. S., Archbald, J. F., Velmurugan, T., Sumathi, D., Devi, S. R., & Preetha, K. S. (2024). Efficient improvement of energy detection technique in cognitive radio networks using K-nearest neighbour (KNN) algorithm. *EURASIP Journal on Wireless Communications and Networking*, 2024(1). https://doi.org/10.1186/s13638-024-02338-8

Muzaffar, M. U., & Sharqi, R. (2023). Energy based Machine Learning Spectrum Sensing in 5G Cognitive Radios. *International Conference on Computational Intelligence and Knowledge Economy*, 326–330. https://doi.org/10.1109/iccike58312.2023.10131872

Ratanavilisagul, C. (2020). Modification fitness function of particle swarm optimization to improve the cluster centroid. *IOP Conference Series Materials Science and Engineering*, 965(1), 012038. https://doi.org/10.1088/1757-899x/965/1/012038

Singh, S. (2017). COGNITIVE RADIO - Overview, Benefits, Challenges and Applications. *Technology Connect*.

Sundrous, K., & Alaa, H. (2014). Cooperative Spectrum Sensing in Cognitive Radio Networks: A Survey on Machine Learning-based Methods. *Journal of Telecommunication and Information Technology*.

Walker, J. (2023). Spectrum Overview. NASA. gov.

Wang, Y., Zhang, Y., Wan, P., Zhang, S., & Yang, J. (2018). A spectrum sensing method based on empirical mode decomposition and K-Means clustering algorithm.

Wireless Communications and Mobile Computing, 2018(1). https://doi.org/10.1155/2018/6104502.

Yucek, T., & Arslan, H. (2009). A survey of spectrum sensing algorithms for cognitive radio applications. *IEEE Communications Surveys & Tutorials*, 11(1), 116–130. https://doi.org/10.1109/surv.2009.090109.