EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

A Cholesky decomposition and fusion clustering based spectrum sensing method ★

Jian Li[1], Yue Li[1,*], Xinpei Zhao[1], Xiaoxu Chen[1]

[1]Heilongjiang University, Harbin 150080, China

Abstract

Spectrum sensing is a key technology to detect unused frequency bands, and is widely applied in spectrum sharing and dynamic channel allocation. However, it is a challenge to provide high sensing accuracy under low signal-to-noise-ratio (SNR) environments. To address this issue, this paper proposes a novel method based on feature extraction and fusion clustering. First, the sampling matrix of the received signal is decomposed into two orthogonal components I and Q, and Cholesky decomposition is performed on the covariance matrices of I and Q components to extract their two-dimensional feature vectors. Then, the fusion clustering algorithm is proposed, where the GMM clustering algorithm is performed to classify the feature vectors, and the initial parameters of GMM, such as centroids, weights and covariance matrices, are generated by K-means clustering. Simulation results show that the proposed method accelerates the convergence speed of GMM and improves the classification accuracy. It effectively enhances the performance of spectrum sensing compared to other mainstream methods.

Received on 25 August 2024; accepted on 02 June 2025; published on 24 June 2025

Keywords: cooperative spectrum sensing, Cholesky, K-Means, Gaussian mixture model

doi:10.4108/eetsis.9552

1. Introduction

In recent years, the rapid development of wireless communications and the surge in spectrum demand have outpaced the available spectrum resources, which can no longer meet user needs [1]. Cognitive Radio (CR) technology is a key solution to alleviate spectrum resource scarcity. This technology intelligently detects unused frequency bands, efficiently allocates and fully utilizes spectrum resources, and enhances band utilization [2]. Traditional spectrum sensing techniques include Energy Detection (ED), Matched Filter (MF) detection, and Cyclostationary Feature (CF) detection [3, 4]. However, traditional spectrum sensing techniques have shortcomings. For example, ED cannot distinguish between signal and noise at low SNR [5]. MF requires prior knowledge of the primary user signal and the channel response [6]. CF is characterized by high complexity and latency [7]. Stochastic theory based and machine learning based methods are employed for spectrum sensing.

1.1. Random matrix theory based approaches

Spectrum sensing schemes based on random matrix theory have been applied [8–11], such as Ratio of Maximum and Minimum Eigenvalue (MME), Difference between the Maximum Eigenvalue and the Average Eigenvalues (DMEAE), Difference of Maximum and Minimum Eigenvalues (DMM)], and Ratio between Maximum Eigenvalue and the Trace (RMET). These blind spectrum sensing methods extract eigenvalues from the received signal covariance matrix to obtain statistical properties, without requiring any prior information about the PU signal and noise variance. However, these schemes require the calculation of precise judgment thresholds in practice. In addressing eigenvalue-related challenges, spectrum sensing schemes utilizing Cholesky decomposition have demonstrated promising performance. Reference [12] proposed a statistical approach to threshold determination, constructing a

★This work was supported in part by Heilongjiang provincial Internet of things perception layer and sensor network technology innovation service platform, and in part by the Basic scientific research project of Heilongjiang province [grant number 2020-KYYWF-1003].
*Corresponding author. Email: 2017021@hlju.edu.cn

test statistic based on the ratio between all elements and the sum of squared main diagonal elements from the decomposed lower triangular matrix. In contrast, Reference [13] adopted a different statistical formulation using the ratio of first-to-last elements along the main diagonal. While both blind spectrum sensing methods exhibit performance proportional to their respective statistics, this correlation may lead to detection instability. To overcome these limitations, Reference [14] introduced an improved blind spectrum sensing method by employing Cholesky decomposition as an RF band standard. However, its threshold derivation relies on empirical estimation without self-analytical capability, representing a significant constraint.

1.2. Machine learning based approaches

Machine Learning (ML) based spectrum sensing techniques avoid the issue of judgment threshold calculation [15, 16]. Recent studies have combined covariance based spectrum sensing techniques with ML by extracting eigenvalues from the covariance matrix to generate a training set, thus obtaining a spectrum sensing model. An unsupervised spectrum sensing technique based on K Means, using the MME of the covariance matrix as the training input for the classifier [17]. In reference [18], eigenvalues are extracted through the covariance matrix after I Q decomposition and a perception scheme based on K-Means clustering is analyzed with different statistical properties. In reference [19], eigenvalue computation are realized using the Decomposition and Reorganization (DAR) method for random matrices, increasing the amount of SUs from the theoretical derivation by constructing two covariance submatrices. In reference [20], the received signal is first preprocessed to generate a feature vector, which is then classified using a Support Vector Machine (SVM) , and this approach ultimately yields an effective spectrum sensing result.

1.3. Our contributions

In order to solve the problem of poor sensing ability and inaccurate threshold estimation at low SNR in traditional spectrum sensing system. This paper proposes a cooperative spectrum sensing method based on feature extraction and fusion clustering (FEFC). The main components are as follows:

(i) Construct the I and Q signals feature matrices, and use Cholesky to decompose the feature matrices to fully extract the features of the signals and construct two dimensional feature vectors.

(ii) Optimization of the GMM using the K-Means algorithm can effectively prevent the GMM from falling into local minima at low SNR, thus improving the performance of spectrum sensing under low SNR conditions.

(iii) In the experimental simulation section, we compare and analyze the performance of FEFC and single antenna CSS method . These methods are simulated using the AWGN channel. The simulation results show that the FEFC can effectively improve the spectrum sensing performance.

The paper is structured as follows: Section 2 introduces the system model of CSS. Section 3 proposes a feature extraction method based on Cholesky decomposition. Section 4 presents a CSS approach utilizing K-means-optimized GMM, designated as FEFC . Section 5 conducts simulation verification of the FEFC sensing method, with results demonstrating significant improvement in spectrum sensing performance. Finally, Section 6 concludes the paper and outlines potential directions for future research.

2. System model

In cognitive radio, the secondary user (SU) senses the primary user (PU) signal, which is easily affected by multipath effects, shadowing, and channel fading, increasing the detection difficulty. To address this, a multiuser collaborative spectrum sensing model is proposed. Signal sensing through multiple users and paths reduces environmental influences and improves system performance. The model is illustrated in Figure 1.

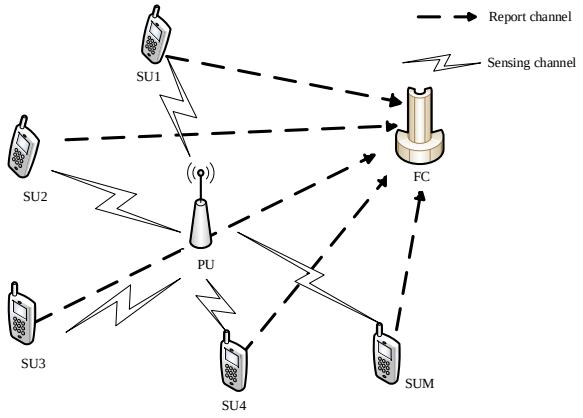

Figure 1. Cooperative spectrum sensing model

The cognitive radio network consists of 1 PU and M SUs. The detection of the PU signal by the SU can be expressed as a binary hypothesis model.

$$x(t) = \begin{cases} w(t), & H_0 \\ s(t) + w(t), & H_1 \end{cases} \quad (1)$$

In equation (1), $x(t)$ represents the received signal of the SU at time t, $s(t)$ represents the PU signal at time t,

and $w(t)$ represents additive Gaussian white noise with a mean of 0 and a variance of δ_x^2. The PU signal and the noise are independently distributed. H_0 indicates that the PU signal is absent, while H_1 indicates that the PU signal is present. Assuming S=0 and S=1 correspond to the channel's available states, respectively, they can be represented as:

$$S = \begin{cases} 0, & H_0 \\ 1, & H_1 \end{cases} \tag{2}$$

Thus, the false alarm probability (P_f) and the detection probability (P_d) are defined as:

$$P_f = p[S = 1|S = 0] \tag{3}$$

$$P_d = p[S = 1|S = 1] \tag{4}$$

In a spectrum sensing system, the signals sensed by M SUs form a vector matrix $\mathrm{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_M]^T$, where the signal sensed by the mth SU is $\boldsymbol{x}_m = [\boldsymbol{x}_m(1), \boldsymbol{x}_m(2), ..., \boldsymbol{x}_m(N)]$, and N is the number of samples. This results in an M×N dimensional signal matrix.

$$\mathbf{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_M]^T = \begin{bmatrix} x_1(1) & x_1(2) & \cdots & x_1(N) \\ x_2(1) & x_2(2) & \cdots & x_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ x_M(1) & x_M(2) & \cdots & x_M(N) \end{bmatrix} \tag{5}$$

3. Feature extraction

3.1. Signal decomposition

In signal analysis, the signal is typically decomposed into two components with the same peak amplitude and frequency but with a 90° phase difference, known as I Q decomposition. This method provides a comprehensive description of the signal's amplitude, frequency, and phase. To fully utilize the received signal information, the signal matrix X is vectorially decomposed:

$$\mathbf{X}^{\mathrm{I}} = \cos(\tfrac{2\pi f_c n}{f_s})\mathbf{X} \tag{6}$$

$$\mathbf{X}^{\mathrm{Q}} = \sin(\tfrac{2\pi f_c n}{f_s})\mathbf{X} \tag{7}$$

In equation (6) and (7), f_c and f_s denote the carrier frequency and sampling frequency, respectively. Thus, the signal matrix $\mathbf{X}$ is vectorially decomposed into two $M \times N$ signal matrices:

$$\mathbf{X}^{\mathrm{I}} = \begin{bmatrix} x_1^I(1) & x_1^I(2) & \cdots & x_1^I(N) \\ x_2^I(1) & x_2^I(2) & \cdots & x_2^I(N) \\ \vdots & \vdots & \ddots & \vdots \\ x_M^I(1) & x_M^I(2) & \cdots & x_M^I(N) \end{bmatrix} \tag{8}$$

$$\mathbf{X}^{\mathrm{Q}} = \begin{bmatrix} x_1^Q(1) & x_1^Q(2) & \cdots & x_1^Q(N) \\ x_2^Q(1) & x_2^Q(2) & \cdots & x_2^Q(N) \\ \vdots & \vdots & \ddots & \vdots \\ x_M^Q(1) & x_M^Q(2) & \cdots & x_M^Q(N) \end{bmatrix} \tag{9}$$

3.2. Feature extraction

Using the $\mathbf{X}^{\mathrm{I}}$ and $\mathbf{X}^{\mathrm{Q}}$ matrices from equation (8) and (9), we can calculate the corresponding covariance matrices $\mathbf{R}^{\mathrm{I}}$ and $\mathbf{R}^{\mathrm{Q}}$ as follows:

$$\mathbf{R}^{\mathrm{I}} = \tfrac{1}{N}\mathbf{X}^{\mathrm{I}}(\mathbf{X}^{\mathrm{I}})^{\mathrm{H}} \tag{10}$$

$$\mathbf{R}^{\mathrm{Q}} = \tfrac{1}{N}\mathbf{X}^{\mathrm{Q}}(\mathbf{X}^{\mathrm{Q}})^{\mathrm{H}} \tag{11}$$

In equation (10) and (11), $(\cdot)^{\mathrm{H}}$represents the conjugate transpose operation, and the covariance matrices $\mathbf{R}^{\mathrm{I}}$ and $\mathbf{R}^{\mathrm{Q}}$ are both $M \times M$ dimensional. Next, the two matrices undergo Cholesky decomposition, as shown in equation (12) and (13).

$$\mathbf{R}^{\mathrm{I}} = \mathbf{Y}^{\mathrm{I}}(\mathbf{Y}^{\mathrm{I}})^{\mathrm{T}} \tag{12}$$

$$\mathbf{R}^{\mathrm{Q}} = \mathbf{Y}^{\mathrm{Q}}(\mathbf{Y}^{\mathrm{Q}})^{\mathrm{T}} \tag{13}$$

Both $\mathbf{Y}^{\mathrm{I}}$ and $\mathbf{Y}^{\mathrm{Q}}$ are lower triangular matrices, expressed as:

$$\mathbf{Y}^{\mathrm{I}} = \begin{bmatrix} y_{11}^I & 0 & \cdots & 0 \\ y_{21}^I & y_{22}^I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ y_{M1}^I & y_{M2}^I & \cdots & y_{MM}^I \end{bmatrix} \tag{14}$$

$$\mathbf{Y}^{\mathrm{Q}} = \begin{bmatrix} y_{11}^Q & 0 & \cdots & 0 \\ y_{21}^Q & y_{22}^Q & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ y_{M1}^Q & y_{M2}^Q & \cdots & y_{MM}^Q \end{bmatrix} \tag{15}$$

The signal features G^I are extracted using the lower triangular matrix $\mathbf{Y}^{\mathrm{I}}$ after Cholesky decomposition, represented by equation (16) as follows:

$$G^I = \frac{\sum_{1 \le i \le j \le M} | y_{ij}^I |}{\sum_{1 \le i \le M} | y_{ii}^I |} \tag{16}$$

$y_{ij}^I \ge 0$ is calculated by equation (17).

$$\begin{cases} y_{jj}^I = \sqrt{\left(r_{jj}^I - \sum_{k=1}^{j-1}\left(y_{kj}^I\right)^2\right)} & i = j \\ y_{ij}^I = \left[r_{ij}^I - \sum_{k=1}^{i-1}\left(y_{ki}^I y_{kj}^I\right)\right]/y_{ii}^I & i < j \end{cases} \tag{17}$$

where r_{ij}^I represents the element in the ith row and jth column of matrix R^{I}. Similarly, Cholesky decomposition of the covariance matrix Y^{Q} is performed, and the signal feature G^{Q} is extracted, represented as:

$$G^Q = \frac{\sum_{1 \le i \le j \le M} | y_{ij}^Q |}{\sum_{1 \le i \le M} | y_{ii}^Q |} \tag{18}$$

Thus, based on G^I and G^Q , the two dimensional feature vector G of the signal can be constructed.

$$\boldsymbol{G} = [G^I, G^Q] \tag{19}$$

Since many training feature vectors are needed to train the clustering algorithm, all training feature vectors $\mathbf{G}$ must first be constructed as a training feature vector set $\tilde{\mathrm{G}}$:

$$\tilde{\mathrm{G}} = \{\boldsymbol{G}_1, \boldsymbol{G}_2, ..., \boldsymbol{G}_B\} \tag{20}$$

where $\boldsymbol{G}_b$ $(b = 1, 2, \ldots, B)$ is the two dimensional feature vector computed in equation (20), and B denotes the number of feature vectors.

4. A fusion clustering-based approach to spectrum sensing

In this paper, the K-Means clustering algorithm is used to optimize the initial parameters of the Gaussian Mixture Model (GMM), effectively preventing it from falling into local minima under low SNR. The optimized GMM model is then used to classify the constructed signal feature vectors, yielding the spectrum sensing results. Figure 2 shows the system modeling based on the fusion clustering algorithm.

4.1. Training process of K-Means clustering algorithm

For the input set of training feature vectors $\tilde{\mathrm{G}} = \{\boldsymbol{G}_1, \boldsymbol{G}_2, ..., \boldsymbol{G}_B\}$, the K-Means clustering algorithm divides the set into multiple clusters. The objective function of the K-Means clustering algorithm is as follows:

$$\min J = \sum_{i=1}^{k} \sum_{\boldsymbol{G}_b \in C_i} \|\mathbf{G}_b - \boldsymbol{\beta}_i\|^2 \tag{21}$$

Here, k is the number of clustering centers, which is set to 2. C_i represents the ith cluster, $\boldsymbol{G}_b$ denotes the sample points in the C_i cluster, and $\boldsymbol{\beta}_i$ is the center of mass of C_i. The final division of clusters is achieved when the iterative results of the algorithm no longer produce significant changes. At this point, the optimal center of mass $\boldsymbol{\beta}_k^{\psi}$ is calculated using Equation (22).

$$\boldsymbol{\beta}_k^{\psi} = \frac{1}{|C_k|} \sum_{\boldsymbol{G}_b \in C_k} \boldsymbol{G}_b \tag{22}$$

In the final clustering result, cluster k contains B_k data points out of a total of B data points. The weight of the cluster corresponding to the best quality center is calculated using equation (23):

$$\alpha_k^{\psi} = \frac{B_k}{B} \tag{23}$$

The corresponding covariance matrix of the corresponding K-Means is as follows:

$$\Sigma_k^{\psi} = \frac{1}{B_k - 1} \sum_{i=1}^{B_k} (\boldsymbol{G}_i - \boldsymbol{\beta}_k^{\psi})(\boldsymbol{G}_i - \boldsymbol{\beta}_k^{\psi})^T \tag{24}$$

4.2. Training process of fusion clustering algorithm

The GMM clustering algorithm is initialized using the optimal centroids, weights, and covariance matrices obtained from the convergence of the K-Means clustering algorithm. The GMM measures the affiliation category of each data point in terms of probability, with its probability density distribution given by the following form:

$$q(\boldsymbol{G}_b) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\boldsymbol{G}_b | \boldsymbol{\mu}_k, \Sigma_k), \sum_{k=1}^{K} \pi_k = 1 \tag{25}$$

where K is the number of models, π_k denotes the mixture weights, $\mathcal{N}(\boldsymbol{G}|\boldsymbol{\mu}_k, \Sigma_k)$ denotes the kth Gaussian distribution, $\boldsymbol{\mu}_k$ is the mean vector, Σ_k is the covariance matrix of the GMM, and the D-dimensional Gaussian distribution takes the following form:

$$\mathcal{N}(\boldsymbol{G}_b \mid \boldsymbol{\mu}_k, \Sigma_k) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma_k|^{1/2}} \cdot \exp\left\{-\frac{1}{2}(\boldsymbol{G}_b - \boldsymbol{\mu}_k)^T \Sigma_k^{-1} (\boldsymbol{G}_b - \boldsymbol{\mu}_k)\right\} \tag{26}$$

The GMM determines the parameters of each distribution using the maximum likelihood function, which is formulated as follows:

$$\ln q(\boldsymbol{G}|\pi, \boldsymbol{\mu}, \Sigma) = \sum_{b=1}^{B} \ln(\sum_{k=1}^{K} \pi_k \mathcal{N}(\boldsymbol{G}_b|\boldsymbol{\mu}_k, \Sigma_k)) \tag{27}$$

Since the membership of each data point to the mixture components is unknown (latent), the Expectation-Maximization (EM) algorithm is employed to iteratively estimate the parameters $\boldsymbol{\mu}_k$, Σ_k, and π_k. EM alternates between estimating the expected membership probabilities of data points to each Gaussian component (Expectation step) and maximizing the likelihood by updating the parameters accordingly (Maximization step). This iterative procedure continues until convergence, ensuring that the model parameters maximize the likelihood of the observed data under the mixture model.

The detailed implementation process of the EM algorithm is shown in Table 1:

4.3. Perceptual decision

The optimal solutions $\left(\boldsymbol{\mu}_k^*, \Sigma_k^*, \pi_k^*\right)$ for the relevant parameters of the $\tilde{\mathrm{G}}$ training vectors are obtained using Expectation-Maximization algorithm. The test vectors are then partitioned using the final clustering model, represented by the following mathematical model of perceptual judgment:

$$\omega = \ln \frac{\pi_1^* \mathcal{N}(\boldsymbol{G}_b^{test}|\boldsymbol{\mu}_1^*, \Sigma_1^*)}{\pi_2^* \mathcal{N}(\boldsymbol{G}_b^{test}|\boldsymbol{\mu}_2^*, \Sigma_2^*)} \tag{28}$$

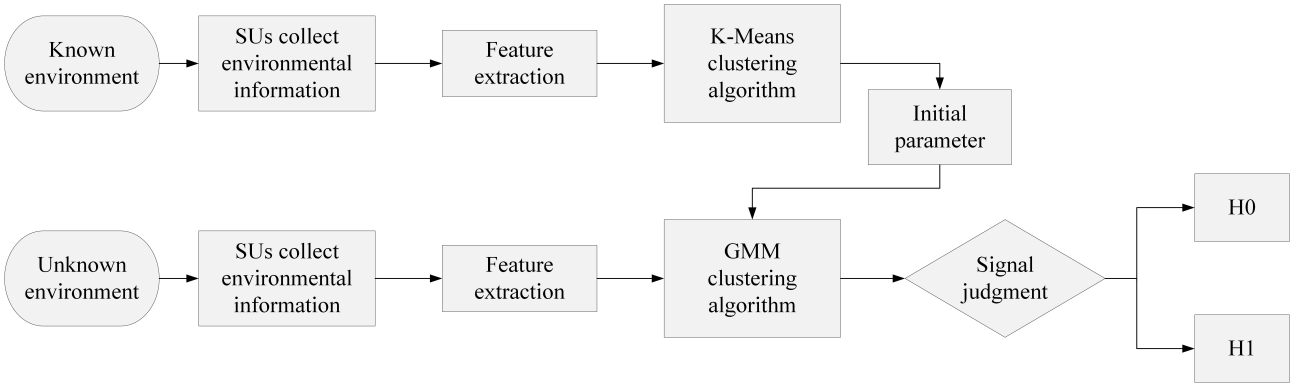

Figure 2. System modeling based on fusion clustering algorithm

Table 1. GMM EM algorithm

Expectation-Maximization algorithm steps	
Step 1	Use the K-Means clustering algorithm to initialize the GMM algorithm $\mu_k = \beta_k^{\psi},\ \Sigma_k = \Sigma_k^{\psi},\ \pi_k = \alpha_k^{\psi}$
Step 2	Desired Steps: $\gamma(z_k \mid \boldsymbol{G}_b) = \frac{\pi_k \mathcal{N}(\boldsymbol{G}_b \mid \mu_k, \Sigma_k)}{\sum_{j=1}^{k} \pi_j \mathcal{N}(\boldsymbol{G}_b \mid \mu_j, \Sigma_j)}$
Step 3	Maximization step: $\mu_k = \frac{1}{B}\sum_{b=1}^{B} \gamma(z_k \mid \boldsymbol{G}_b)\boldsymbol{G}_b$ $\Sigma_k = \frac{1}{B}\sum_{b=1}^{B} \gamma(z_k \mid \boldsymbol{G}_b)(\boldsymbol{G}_b - \mu_k)(\boldsymbol{G}_b - \mu_k)^{\mathrm{T}}$ $\pi_k = \frac{B_k}{B}$ $B_k = \sum_{b=1}^{B} \gamma(z_k \mid \boldsymbol{G}_b)$
Step 4	Until the parameters converge, otherwise return to Step 2.

In the above equation, G_b^{test} is the test vector, ω denotes the detection probability, and β is the threshold which controls the false alarm probability P_f. The detection and judgment process does not require retraining the model. If $\omega > \beta$, the judgment is H_1. Here, H_0 and H_1 denote channel unavailability and availability.

4.4. Complexity analysis

The model training time overhead of this paper's approach consists of the computation of the covariance matrix, the feature extraction and the computation of fusion clustering, the complexities are $O(M^2N)$, $O(M^3)$+$O(M^2)$ and $O(CdL) + O(2CdL)$, where d denotes the dimension, C is the number of clustering centers and L is the length of the training data.

We also compare the complexity of the method in this paper with other methods with better performance, as shown in Table 2. The clustering method in this paper has lower complexity compared to traditional GMM clustering, comparing with the other two algorithms, the complexity of this paper's method is higher, but the proposed method has better detection performance due to the use of GMM clustering.

Table 2. Comparison of complexity of different methods

Methods	**Complexity**
DMM+IQ	$O(N^2L + M^3 + CdL)$
DMEAE+DAR	$O(M^2L + M^3 + CdL)$
GMM	$O(M^2N + M^3 + M^2 + CdL + CL + 2CdL)$
FEFC	$O(M^2N + M^3 + M^2 + CdL + 2CdL)$

5. Simulation experiments

To demonstrate the spectrum sensing performance of the proposed method, its performance is analyzed through MATLAB simulations and compared with other spectrum sensing techniques. To ensure the accuracy and reliability of the experimental results, the simulated primary user signal used in the experiments is a BPSK signal, with a signal activity probability of 0.5. Additionally, a flat Rayleigh fading channel is assumed between the primary user (PU) and the secondary user (SU). The noise is an additive white Gaussian noise (AWGN) signal with a mean of 0 and a variance of 1, and it is independently distributed with ideal Gaussian white noise. Experimentally, 5000 signal feature vectors were extracted for clustering to obtain a classification model, and another 2500 signal feature vectors were extracted to analyze the perceptual performance of the model.

5.1. Classification effect based on FEFC clustering algorithm

This subsection presents a comparative visualization of raw feature vectors alongside their post-training classification outcomes processed through the FEFC clustering methodology.

Figure 3 shows 5000 feature vectors that have not been classified by the clustering algorithm, with 2500 being noise feature vectors and the remaining 2500 being signal plus noise feature vectors.

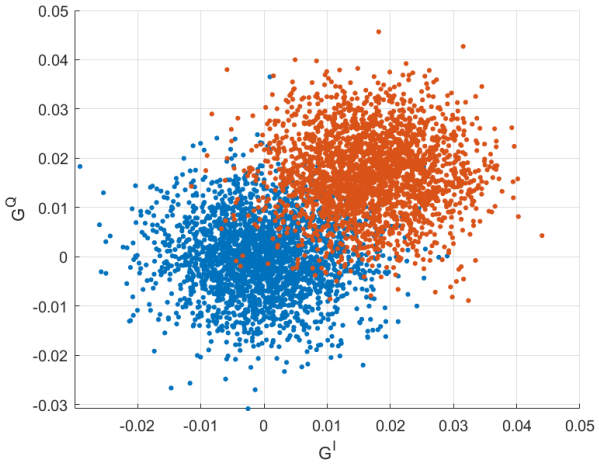

Figure 3. Raw signal feature vector

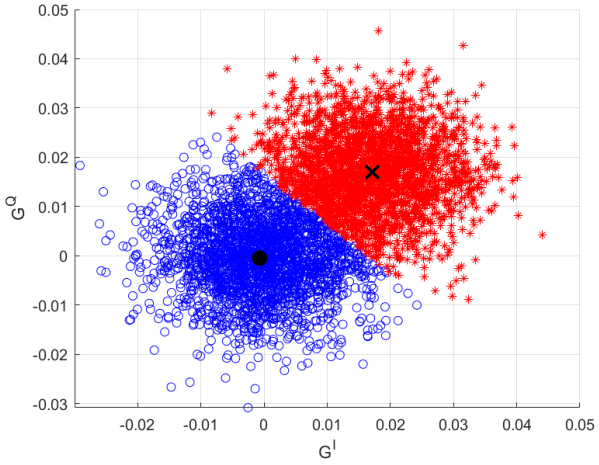

Figure 4. Clustered signal feature vector

Figure 4 shows the results of clustering 5000 feature vectors using the proposed method. In the figure, red indicates the channel available class, while blue represents the channel unavailable class.

5.2. FEFC clustering algorithm under the detection performance

Figure 5 compares the channel classification accuracy of the proposed fusion clustering method with traditional GMM clustering at different SNR.The SNR was varied from -20 dB to -6 dB, with the number of SUs M=5 and the number of sampling points N=2000. From the figure, it can be seen that the classification accuracy of this paper's algorithm is close to 100% when SNR=-8dB, 97.6% when SNR=-10dB, and 62.75% in the worse environment with SNR of -20dB. When the SNR is higher than -10 dB, the accuracy of the two algorithms in the figure is comparable. However, when the SNR is lower than -12 dB, the classification accuracy of the proposed method is significantly better than that of traditional GMM clustering.

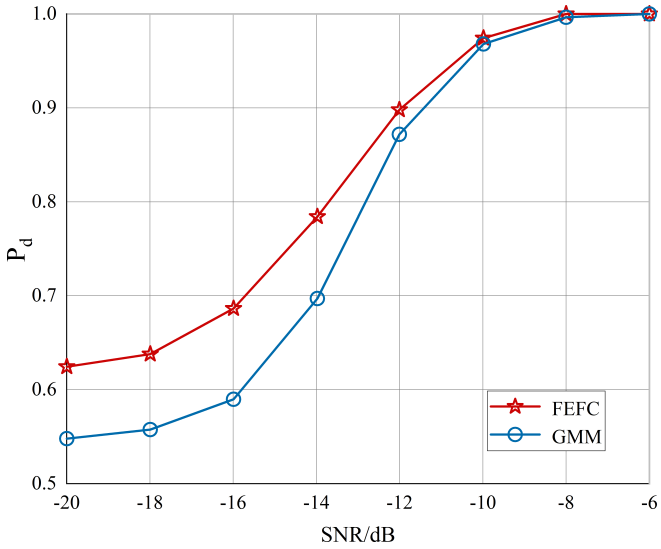

Figure 5. Comparison of classification accuracy

To evaluate the performance of our proposed method, we compare it with other mainstream spectrum sensing techniques, including the Maximum to Minimum Eigenvalue ratio (MME), the Difference between Maximum and Minimum Eigenvalues (DMM), and the Ratio of Maximum Eigenvalue to Trace (RMET).

Figure 6 shows the spectrum sensing performance of various algorithms with the number of SUs M=5, the SNR is set to be -12 dB and the number of sampling points N=2000. When the false alarm probability P_f is 0.1, the detection probability P_d reaches 89%, representing a 9% improvement over the DMM based K-Means clustering method.

Figure 7 shows the spectrum sensing performance of various algorithms at an SNR of -14 dB. When the false alarm probability P_f is 0.1, the detection probability P_d reaches 78.5%, representing an 8% improvement over the DMM based K-Means clustering method.

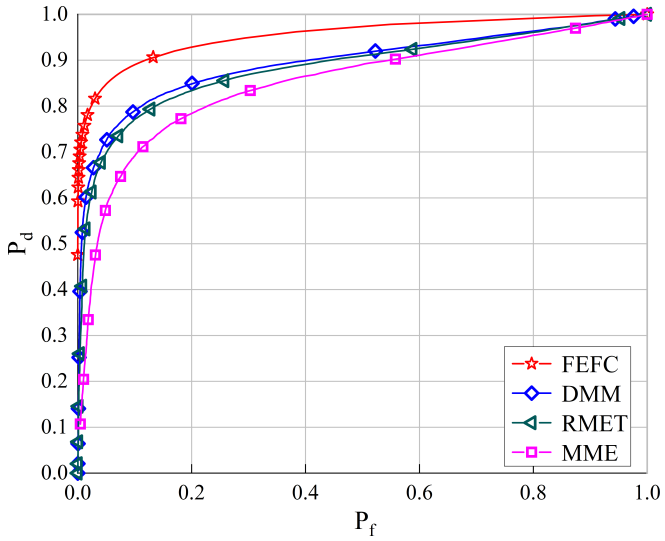

Figure 6. ROC curves for each algorithm with SNR = -12 dB and M = 5

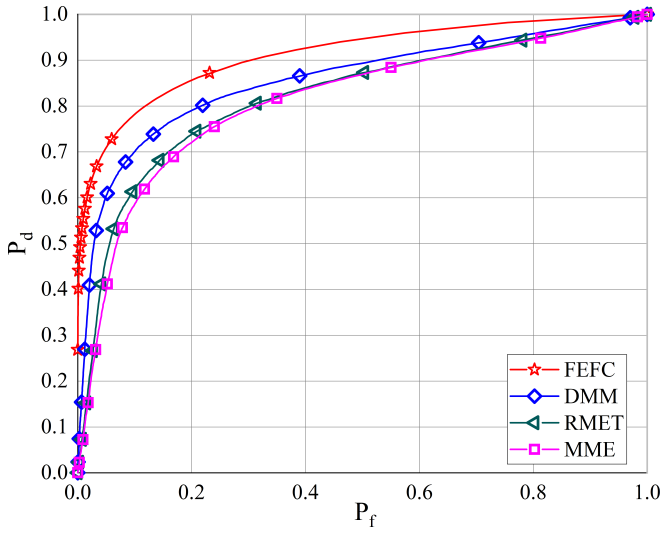

Figure 7. ROC curves for each algorithm with SNR = -14 dB and M = 5

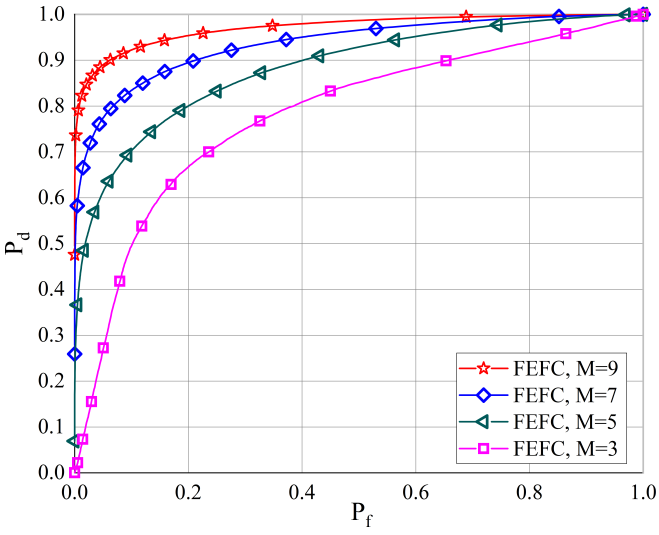

Figure 8. Performance analysis of different number of SU with SNR = -16dB

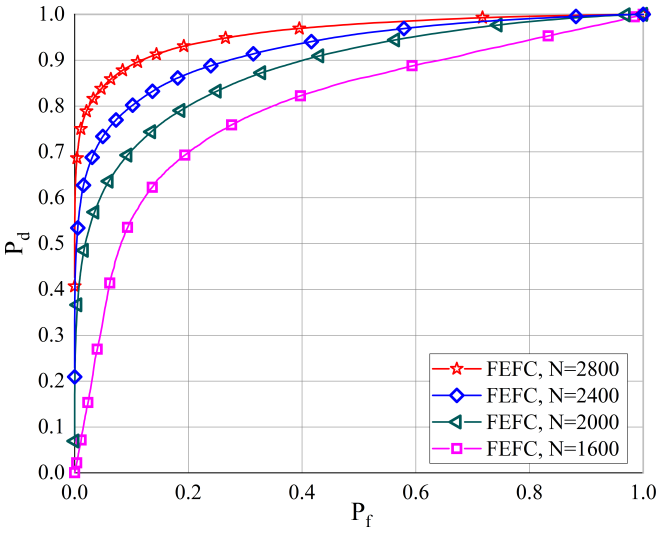

Figure 9. Performance analysis of different sampling points with SNR = -16dB

Figure 8 shows the performance analysis of different number of SU with SNR=-16dB and N=2000. From Figure 8, it can be seen that the number of secondary SUs M is highly related to the detection probability of the algorithm, and the perceived performance of the FEFC algorithm improves with the increase of M.

Figure 9 shows the performance analysis of different sampling points with SNR=-16dB and M=5. From Figure 9, it can be seen that the perceptual performance of the FEFC algorithm improves with the increase in the number of sampling points.

The experimental results indicate that the detection performance of the method proposed in this paper is significantly improved compared to other spectrum sensing methods across a range of SNR.

6. Conclusions

This paper proposes a spectrum sensing method based on feature extraction and fusion clustering to address spectrum sensing problem under low SNR conditions. First, the sampling matrix of the received signal is vectorially decomposed to obtain the I and Q component signals. Then, the information of the signal matrix is fully extracted using the cholesky decomposition of the covariance matrix of the I and Q signals. Finally, the GMM model, optimized by the K-Means algorithm, is used to classify the signal samples

and determine the channel state. Simulation results indicate that the algorithm exhibits good perceptual performance. In this paper, we have not considered the problem of forging data by malicious users in the system, which can generate outliers and thus affect the performance of clustering, and further research will be conducted on this issue in the subsequent work.

References

[1] Kaur J. and Kaur K. (2015) *A fuzzy approach for an IoT-based automated employee performance appraisal*, Computers Materials and Continua, Vol. 53, No. 1, pp. 24–38.

[2] Alemseged Y.D. and Harada H. (2009) *Spectrum sensing for cognitive radio*, Proceedings of the IEEE, Vol. 97, No. 5, pp. 849–877.

[3] Yu L., Wu Q. and Wang J. (2014) *Optimal adaptive multiband spectrum sensing in cognitive radio networks*, KSII Transactions on Internet and Information Systems, Vol. 8, pp. 984–996.

[4] Awe O.P., Deligiannis A. and Lambotharan S. (2018) *Spatio-temporal spectrum sensing in cognitive radio networks using beamformer-aided SVM algorithms*, IEEE Access, Vol. 6, pp. 25377–25388.

[5] Luo J., Zhang G., Xu H., Zhou Y. and Yang W. (2015) *An energy detection based signal detection method with multiple antennas in cognitive radio*, Proc. IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), pp. 187–191.

[6] Surampudi A. and Kalimuthu K. (2016) *An adaptive decision threshold scheme for the matched filter method of spectrum sensing in cognitive radio using artificial neural networks*, Proc. 1st India International Conference on Information Processing (IICIP), pp. 1–5.

[7] Du J., Huang H., Jing X.J. and Chen X. (2016) *Cyclostationary feature based spectrum sensing via low-rank and sparse decomposition in cognitive radio networks*, Proc. 16th International Symposium on Communications and Information Technologies (ISCIT), pp. 615–619.

[8] Liu C.; Yang M. (2011) *A Distance-Weighed Algorithm Based on Maximum-Minimum Eigenvalues for Cooperative Spectrum Sensing*, Proc. IEEE 7th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM), Wuhan, China, pp. 1–4.

[9] Zeng Y.; Liang Y. (2009) *Eigenvalue-based spectrum sensing algorithms for cognitive radio*, IEEE Transactions on Communications, Vol. 57, No. 6, pp. 1784–1793.

[10] Abed V.; Shahzadi A. (2014) *A Maximum-Minimum Eigenvalue Detection Simpler Method Based on Secondary Users Locations for Cooperative Spectrum Sensing*, Proc. 2014 International Congress on Technology, Communication and Knowledge (ICTCK), Mashhad, Iran, pp. 1–4.

[11] Miah M.S.; Yu H.J.; Godder T.K.; Rahman M.M. (2015) *A Cluster-Based Cooperative Spectrum Sensing in Cognitive Radio Network Using Eigenvalue Detection Technique with Superposition Approach*, International Journal of Distributed Sensor Networks, Vol. 2015, pp. 1–11.

[12] Yang X.; Lei K.; Peng S.; Cao X. (2011) *Blind detection for primary user based on the sample covariance matrix in cognitive radio*, IEEE Communications Letters, Vol. 15, No. 1, pp. 40–42.

[13] Zhou F.; Li Z.; Si J.; Guan L. (2014) *An efficient spectrum sensing algorithm for cognitive radio based on finite random matrix*, Proc. IEEE 25th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), pp. 1223–1227.

[14] Li Z.; Zhou F.; Si J.; Qi P.; Guan L. (2016) *Feasibly efficient cooperative spectrum sensing scheme based on Cholesky decomposition of the correlation matrix*, IET Communications, Vol. 10, No. 9, pp. 1003–1011.

[15] Bkassiny M.; Li Y.; Jayaweera S.K. (2013) *A survey on machine-learning techniques in cognitive radios*, IEEE Communications Surveys & Tutorials, Vol. 15, No. 3, pp. 1136–1159.

[16] Lu Y.; Zhu P.; Wang D.; Fattouche M. (2016) *Machine learning techniques with probability vector for cooperative spectrum sensing in cognitive radio networks*, Proc. IEEE Wireless Communications and Networking Conference (WCNC), pp. 1–6.

[17] Sobabe G.C.; Song Y.; Bai X. *et al.* (2018) *A cooperative spectrum sensing algorithm based on unsupervised learning*, Proc. 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), pp. 1–6.

[18] Zhang Y.; Wan P.; Zhang S.; Wang Y.; Li N. (2017) *A spectrum sensing method based on signal feature and clustering algorithm in cognitive wireless multimedia sensor networks*, Advances in Multimedia, Vol. 2017, Art. ID 2895680, 10 pages.

[19] Wang Y.; Zhang Y.; Zhang S. *et al.* (2018) *A cooperative spectrum sensing method based on a feature and clustering algorithm*, Proc. 2018 Chinese Automation Congress (CAC), pp. 1029–1033.

[20] Bao J.; Nie J.; Liu C.; Jiang B.; Zhu F.; He J. (2019) *Improved blind spectrum sensing by covariance matrix Cholesky decomposition and RBF-SVM decision classification at low SNRs*, IEEE Access, Vol. 7, pp. 97117–97129.