



Big data theory based spectrum sensing algorithm for the satellite cognitive radio network

Mingchuan Yang^{1,2} · Xinye Shao¹ · Guanchang Xue¹ · Bingyu Xie¹

Accepted: 17 September 2021
© The Author(s) 2021

Abstract

In order to deal with the difficulty of spectrum sensing in cognitive satellite wireless networks, a large-scale cognitive network spectrum sensing algorithm based on big data analysis theory is studied, and a new algorithm using mean exponential eigenvalue is proposed. This new approach fully uses all the eigenvalues in sample covariance matrix of the sensing results to make the decision, which can effectively improve the detection performance without obtaining the prior information from licensed users. Through simulation, the performance of various large scale cognitive radio spectrum sensing algorithms based on big data analysis theory is compared, and the influence of satellite to ground channel conditions and the number of sensing nodes on the performance of the algorithm is discussed.

Keywords Big data theory · Dynamic spectrum access · Satellite communication network · Spectrum sensing

1 Introduction

The development of satellite communication technology makes it play an important role in satellite TV broadcasting, Internet, distance education, military applications and mobile communications. Satellite has become an indispensable way for realizing seamless global personal communication and high-speed Internet in the air because of its long distance and not being easily affected by time, region, and airspace [1]. As business demands increase, frequency resources become more and more precious. However, the unreasonable allocation and low utilization rate of spectrum resources in the satellite communication system have further caused the problem of spectrum shortage. At present, there are many ways to improve the utilization of unit spectrum resources, such as using multiple carrier frequency multiplexing or multi-antenna transmission technology [2]. But the starting point of this kind of technology is fixed spectrum allocation strategy, which cannot

fundamentally solve the unbalanced spectrum allocation [3]. In order to fundamentally solve this contradiction and improve the utilization of spectrum resources, people consider applying cognitive radio theory to satellite communication system, and then put forward the concept of satellite cognitive radio network [4].

In recent years, many scholars have conducted extensive research on the application scenarios, network architecture and key technologies of satellite cognitive wireless network. Traditional fast spectrum sensing algorithms, such as energy detection algorithm [5], cannot provide good sensing performance under such low SNR conditions. Other algorithms with good detection performance, such as cyclic stationary feature detection algorithm [6], have high complexity and long detection time. These algorithms are mainly applicable to a single sensor node, and are not applicable to scenarios where multiple sensor nodes are jointly detected. In 2013, European scientists launched the “Cognitive Radio for Satellite Communications” (CoRaSat) project, which aims to research, develop and build cognitive radio technology that can be applied to satellite communication systems [7]. Reference [8] analyzed the specific application scenarios, operating frequency bands and challenges of the satellite cognitive wireless network in the CoRaSat project. Reference [9] comprehensively analyzes the key technologies of satellite cognitive wireless

✉ Mingchuan Yang
mcyang@hit.edu.cn

¹ Communication Research Center, Harbin Institute of Technology, Harbin 150001, China

² Science and Technology on Communication Networks Laboratory, Shijiazhuang 050002, China

network, including spectrum sensing, interleaving and database technology, and matches the existing satellite multibeam pattern with the actual scene of European map to analyze the feasibility of satellite cognitive wireless network. The Institute of Microsystems, Chinese Academy of Sciences has studied the feasibility of applying cognitive radio technology to satellite communication system. The results show that cognitive radio technology can solve low spectrum utilization and low anti-interference ability. Reference [10] analyzes the application prospects of cognitive radio technologies in low earth orbit satellite communication systems. Although many scholars have done some theoretical research on related technologies, there is still no perfect application scheme for cognitive radio technology in satellite communication system.

Considering the particularity of satellite communication channel, the traditional single sensing node cognitive radio spectrum sensing algorithm is easily to be interfered, resulting in the performance degradation [11]. Therefore, the centralized multi-node spectrum sensing method is adopted to improve the spectrum sensing performance and realize the rapid and accurate sensing of authorized signals. Based on the traditional large-scale cognitive radio spectrum sensing algorithm, the eigenvalue exponential mean algorithm based on big data analysis theory is proposed.

The paper is organized as follows. In Sect. 2, we introduce the system model of satellite cognitive radio networks. In Sect. 3, the spectrum sensing algorithms based on big data analysis theory for large scale cognitive networks are investigated, and a new algorithm using mean exponential eigenvalue is proposed. The performance of the spectrum sensing algorithms is verified in Sect. 4. Finally, the conclusion is given in Sect. 5.

2 System model

The spectrum sensing scenario based on big data analysis in satellite cognitive wireless network is shown in Fig. 1. In the cognitive network, there are multiple sensing nodes and cognitive users that perceive the signals of authorized users (users using authorized frequency bands). For the received signal $x(t)$, the sampling interval is T_s , and the received signal samples are:

$$x[n] = x(nT_s) \quad (1)$$

According to the spectrum sensing theory of cognitive radio, when the sensing node only receives noise, the decision is H_0 , which means that the authorized user does not occupy the spectrum. When both the signal and noise are received, the decision is H_1 , which means that the spectrum is occupied. Therefore, the received sampling signal can be expressed as:

$$H_0 : x[n] = w[n] \quad (2)$$

$$H_1 : x[n] = s[n] + w[n] \quad (3)$$

where $w[n]$ is the received independent and identically distributed gaussian white noise, the mean value is 0 and the variance is σ^2 . Assuming that the sensing node cannot predict the current channel condition, the mean value and variance of the received authorized user signal $s[n]$ are unknown.

Detection probability P_d and false alarm probability P_f are used to measure detection performance. P_d indicates that the sensing node correctly detects the signal when the authorized user occupies the spectrum. P_f indicates that when the authorized user does not occupy the spectrum, the sensing node misjudges that the signal of the authorized user exists, that is:

$$P_d = P(D_1|H_1) \quad (4)$$

$$P_f = P(D_1|H_0) \quad (5)$$

where D_1 indicates that the cognitive radio network determines that the authorized user is occupying the spectrum.

The sample space Γ_x collected by the sensing node are divided into several subspaces, and each subspace contains N_{tot} sampling data. The subspace $\Gamma_{x,i}$ is regarded as composed of N sample vectors \mathbf{X}_i of length L , that is:

$$\Gamma_{x,i} = \{\mathbf{X}_{(i-1)N+1}, \mathbf{X}_{(i-1)N+2}, \dots, \mathbf{X}_{(i-1)N+N}\} \quad (6)$$

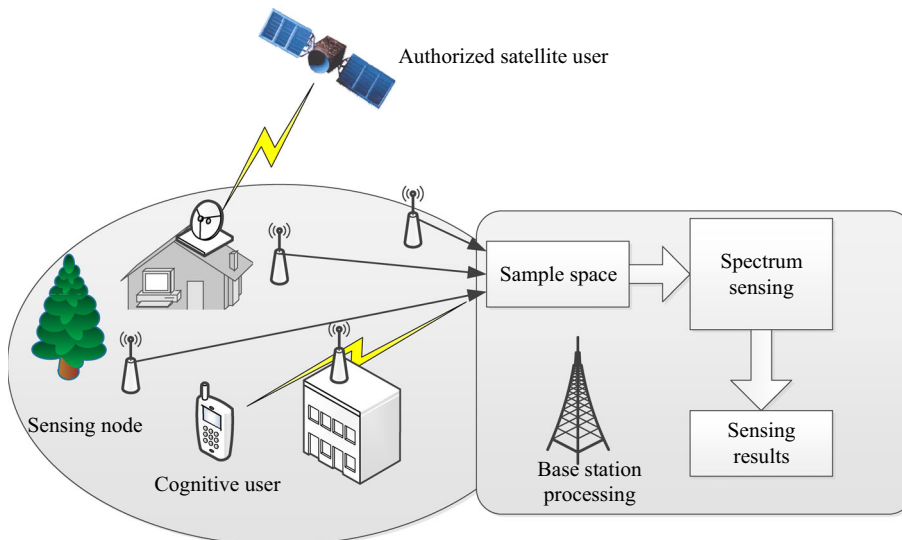
$$\mathbf{X}_i = [x[i], x[i+1], \dots, x[i+L-1]]^T \quad (7)$$

where L called smoothing factor, $()^T$ represents the transpose of matrix. Assume $\mathbf{X}_i \sim \mathbb{N}(0, \mathbf{R}_x)$, where \mathbf{R}_x is the covariance matrix of the received signal. In practical applications, \mathbf{R}_x is often unknown, but it can be approximated by the sample covariance matrix $\hat{\mathbf{R}}_x$.

$$\hat{\mathbf{R}}_x = \frac{1}{N} \sum_{i=1}^N \mathbf{X}_i \mathbf{X}_i^T \quad (8)$$

In large-scale cognitive radio networks, multiple sensing nodes are usually used to obtain large amounts of data. Multivariate analysis based on big data usually assumes that the sample of variables is large enough, that is, N/L is large enough. The satellite cognitive wireless network covers a wide range, there are a large number of sensing nodes and cognitive users, and it is easy to obtain a large amount of cognitive data. Centralized processing of the sensory data can obtain a sufficiently large sample subspace in the fusion center and improve the detection capability.

Fig. 1 Satellite cognitive wireless network spectrum sensing scenario



3 Spectrum sensing algorithms based on big data theory

In recent years, the use of big data analysis to extract statistical features of sample subspaces has become a hot spot in the study of spectrum sensing in large-scale cognitive wireless networks. The dimensionality of the observation space occupied by the authorized user signal is smaller than that of the noise, and the frequency spectrum of the authorized signal and Gaussian white noise is very different, so the spectrum sensing algorithm uses the sample covariance matrix for signal detection.

3.1 Estimation-correlation (EC) algorithm

The EC algorithm performs sensing based on the prior information of authorized user signals. It is assumed that the sensed signal is a zero-mean Gaussian random process, and the covariance matrix is known. First, assume that the noise $w[n]$ in the channel is Gaussian white noise, which is not related to the signal, and the variance is σ^2 . The signal $s[n]$ is also zero mean. From Eqs. (2) and (3), the sample vector X_i of the sensing node obeys:

$$H_0 : X_i \sim \mathcal{N}(0, \sigma^2 \mathbf{I}) \tag{9}$$

$$H_1 : X_i \sim \mathcal{N}(0, \mathbf{R}_s + \sigma^2 \mathbf{I}) \tag{10}$$

where \mathbf{R}_s is the covariance matrix of the signal and \mathbf{I} is the unit matrix.

In the cognitive radio system, the given false alarm probability P_f . According to the Neyman-Pearson (NP) criterion [12], when Eq. (11) is established, it is considered that the authorized user is occupying the spectrum.

$$H_1 : L(X) = \frac{p(X; H_1)}{p(X; H_0)} > \gamma \tag{11}$$

where $p(X; H_0)$ and $p(X; H_1)$ are the probability density functions of random variables X in the case of H_0 and H_1 , respectively. The decision threshold γ is related to the expected false alarm probability α , which can be determined by:

$$P_f = \int_{\{X:L(X) > \gamma\}} p(X; H_0) dX = \alpha \tag{12}$$

In addition, the function $L(X)$ in Eq. (11) is called the likelihood ratio, which represents the similarity of variable x between H_0 and H_1 . Therefore, it is also called likelihood ratio detection or NP detection.

The probability density functions of the sample vector X_i at the sensing node satisfy the following expressions respectively:

$$p(X_i; H_1) = \frac{1}{(2\pi)^{\frac{N}{2}} \det^{\frac{1}{2}}(\mathbf{R}_s + \sigma^2 \mathbf{I})} \times \exp\left[-\frac{1}{2} X_i^T (\mathbf{R}_s + \sigma^2 \mathbf{I})^{-1} X_i\right] \tag{13}$$

$$p(X_i; H_0) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma^2} X_i^T X_i\right] \tag{14}$$

where $\det(\cdot)$ denotes the determinant of the matrix.

According to the NP criterion, when the likelihood ratio $L(X_i)$ satisfies the Eq. (15), the cognitive radio system will consider that the authorized user is occupying the spectrum.

$$L(\mathbf{X}_i) = \frac{\exp\left[-\frac{1}{2}\mathbf{X}_i^T(\mathbf{R}_s + \sigma^2\mathbf{I})^{-1}\mathbf{X}_i\right]}{(2\pi)^{\frac{N}{2}}\det^{\frac{1}{2}}(\mathbf{R}_s + \sigma^2\mathbf{I})} > \gamma \tag{15}$$

After taking the logarithm and simplifying, we get:

$$\begin{aligned} l(\mathbf{X}_i) &\triangleq \ln\{L(\mathbf{X}_i)\} \\ &= \frac{1}{2}\ln\left(\frac{\sigma^{2N}}{\det(\mathbf{R}_s + \sigma^2\mathbf{I})}\right) - \\ &\quad \frac{1}{2}\mathbf{X}_i^T\left[(\mathbf{R}_s + \sigma^2\mathbf{I})^{-1} - \frac{1}{\sigma^2}\mathbf{I}\right]\mathbf{X}_i \end{aligned} \tag{16}$$

At this time, when $l(\mathbf{X}_i) > \ln \gamma$, it is determined that the vector \mathbf{X}_i contains the signal of authorized user. For the sake of simplicity, only the part of Eq. (16) that contains the vector \mathbf{X}_i is considered, which is simplified to:

$$-\frac{1}{2}\mathbf{X}_i^T\left[(\mathbf{R}_s + \sigma^2\mathbf{I})^{-1} - \frac{1}{\sigma^2}\mathbf{I}\right]\mathbf{X}_i > \gamma' \tag{17}$$

Then we can get:

$$T(\mathbf{X}_i) \triangleq \sigma^2\mathbf{X}_i^T\left[\frac{1}{\sigma^2}\mathbf{I} - (\mathbf{R}_s + \sigma^2\mathbf{I})^{-1}\right]\mathbf{X}_i > 2\sigma^2\gamma' \tag{18}$$

For invertible matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ and \mathbf{D} , it has the following property:

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1} \tag{19}$$

Let $\mathbf{A} = \sigma^2\mathbf{I}$, $\mathbf{B} = \mathbf{D} = \mathbf{I}$, $\mathbf{C} = \mathbf{R}_s$, then:

$$(\sigma^2\mathbf{I} + \mathbf{R}_s)^{-1} = \frac{1}{\sigma^2}\mathbf{I} - \frac{1}{\sigma^4}\left(\frac{1}{\sigma^2}\mathbf{I} + \mathbf{R}_s^{-1}\right)^{-1} \tag{20}$$

Combining Eqs. (18) and (20), we can get:

$$\begin{aligned} T(\mathbf{X}_i) &= \mathbf{X}_i^T\left[\frac{1}{\sigma^2}\left(\frac{1}{\sigma^2}\mathbf{I} + \mathbf{R}_s^{-1}\right)^{-1}\right]\mathbf{X}_i \\ &= \mathbf{X}_i^T\frac{1}{\sigma^2}\left[\frac{1}{\sigma^2}(\mathbf{R}_s + \sigma^2\mathbf{I})\mathbf{R}_s^{-1}\right]^{-1}\mathbf{X}_i \\ &= \mathbf{X}_i^T\mathbf{R}_s(\mathbf{R}_s + \sigma^2\mathbf{I})^{-1}\mathbf{X}_i \\ &\triangleq \mathbf{X}_i^T\hat{\mathbf{S}} \end{aligned} \tag{21}$$

Alternatively, $T(\mathbf{X}_i)$ can be written as the product of discrete signals:

$$T(\mathbf{X}_i) = \sum_{k=0}^{L-1} x[i+k]s[i+k] \tag{22}$$

Therefore, for the sample vector \mathbf{X}_i , the sensing node only needs to calculate $T(\mathbf{X}_i)$ according to the prior information \mathbf{R}_s and σ^2 of the authorized user signal, and compare it with the threshold γ'' under the specific false alarm probability α .

In satellite cognitive wireless networks, sufficient sample subspace can be obtained. Therefore, the N vectors $\mathbf{X}_i (i = 1, 2, \dots, N)$ in the subspace can be estimated and correlated, and then the detection results can be averaged to improve the detection performance. When authorized users occupy the spectrum, there is:

$$H_1 : T_{EC} = \frac{1}{N}\sum_{i=1}^N \mathbf{X}_i^T\mathbf{R}_s(\mathbf{R}_s + \sigma^2\mathbf{I})^{-1}\mathbf{X}_i > \gamma_{EC} \tag{23}$$

where γ_{EC} is the threshold. Equation (23) is the expression for estimation-correlation detection using the prior information \mathbf{R}_s and σ^2 of the authorized user signal.

3.2 Eigenvalue template matching (ETM) algorithm

In the actual spectrum sensing process of cognitive radio, the prior information \mathbf{R}_s and σ^2 are often unknown. In this case, the sensing node uses the previous sample subspace to extract features as a prior information, and compare with the features extracted in the current sample subspace to determine whether there is a spectrum hole.

The ETM algorithm uses the covariance matrix \mathbf{R}_x of the sample subspace $\Gamma_{x,i}$ for sensing, and extracts the first eigenvector of the matrix (that is, the eigenvector corresponding to the largest eigenvalue) as the feature of the subspace $\Gamma_{x,i}$, denoted as η_i . If the subspace contains only noise vectors, the feature is random, and if it contains signal, the feature is stable. Therefore, the ETM algorithm uses an adaptive feature learning algorithm (FLA) to automatically extract features from two consecutive sample subspaces, that is, to compare the similarity of the two subspace features.

$$T_{FLA} = \max_{l=1,2,\dots,L-k+1} \left| \sum_{k=1}^L \eta_i[k]\eta_{i+1}[k+l] \right| \tag{24}$$

where η_i and η_{i+1} are the feature vectors of two continuous sample subspaces $\Gamma_{x,i}$ and $\Gamma_{x,i+1}$ respectively. If the value of T_{FLA} is higher than γ_{FLA} , the feature learning process will end. The vector η_{i+1} at this time is regarded as the feature ϕ_s of the authorized signal.

Then use ϕ_s as the prior information, the vector $\eta_{current}$ corresponding to the current subspace $\Gamma_{x,current}$ with ϕ_s are calculated for similarity. If similarity T_{ETM} is higher than the threshold γ_{ETM} , it is determined that the current subspace contains the signal of the authorized user. The detection expression of ETM algorithm is:

$$H_1 : T_{ETM} = \max_{l=1,2,\dots,L-k+1} \left| \sum_{k=1}^L \phi_s[k]\eta_{current}[k+l] \right| > \gamma_{ETM} \tag{25}$$

The algorithm consists of two stages: First, the similarity function is used in feature learning to automatically extract the signal features of authorized users. Second, in the decision-making process, compare the similarity between the characteristics of authorized user signal and the current subspace to determine whether the authorized signal is included.

3.3 Matrix function detection (MFD) algorithm

The MFD algorithm uses the positive semi-definiteness of the covariance matrix for detection, and does not need to calculate the eigenvalues of matrix, but uses the trace of the matrix as the criterion for decision-making. Assume that the covariance matrices of authorized user signal and noise are \mathbf{R}_s and \mathbf{R}_n , respectively. When the signal contains authorized user signal and noise, there is:

$$\mathbf{R}_x = \mathbf{R}_s + \mathbf{R}_n + E(\mathbf{s}\mathbf{w}^T) + E(\mathbf{w}\mathbf{s}^T) \tag{26}$$

where $E(\cdot)$ means the average. When the sample value of the sensing node is large, it can be assumed that the signal is not correlated with the noise, then $E(\mathbf{s}\mathbf{w}^T) = 0, E(\mathbf{w}\mathbf{s}^T) = 0$.

$$H_1 : \mathbf{R}_x = \mathbf{R}_s + \mathbf{R}_n \tag{27}$$

Correspondingly, when authorized user do not occupy the spectrum, there is:

$$H_0 : \mathbf{R}_x = \mathbf{R}_n \tag{28}$$

In the MFD algorithm, the sample subspace $\Gamma_{x,i}$ is divided into K segments. Each segment contains $N_k = N/K$ sample vectors. Therefore, the covariance matrix of the k -th segment is:

$$\mathbf{R}_{x,k} = \frac{1}{N_k} \sum_{i=(k-1)N_k+1}^{(k-1)N_k+N_k} \mathbf{X}_i \mathbf{X}_i^T \tag{29}$$

From Eqs. (27) and (29), when the authorized user occupies the spectrum, the covariance matrix is:

$$H_1 : \mathbf{R}_{x,k} = \mathbf{R}_{s,k} + \mathbf{R}_{n,k} \tag{30}$$

Since the sample covariance matrix is positive semi-definite, for any $k = 1, 2, \dots, K$, we have:

$$\mathbf{R}_{s,k} + \mathbf{R}_{n,k} \succ \mathbf{R}_{n,k} \tag{31}$$

In the case of H_1 , the elements of the sample covariance matrix of any segment have larger values. We use this property to determine whether the subspace contains signal from authorized users. In order to improve the robustness, we average the sample covariance matrix of each segment. From Eq. (31), we can get:

$$\frac{1}{K} \sum_{k=1}^K (\mathbf{R}_{s,k} + \mathbf{R}_{n,k}) \succ \frac{1}{K} \sum_{k=1}^K \mathbf{R}_{n,k} \tag{32}$$

For $\mathbf{A} \prec \mathbf{B}$, we can usually find a monotonically increasing function $f(\cdot)$, such that $f(\mathbf{A}) \prec f(\mathbf{B})$. The output of the function $f(\cdot)$ is also a matrix, which is still difficult to compare. Therefore, we look for the trace of matrix to make the result a real number, and then compare.

For positive semi-definite matrix $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}$, suppose $\mathbf{A} \prec \mathbf{B}$, let $f : [0, \infty) \rightarrow [0, \infty)$ satisfy: ① $f(0) = 0$, ② $f(\cdot)$ is continuous increasing functions, then

$$Tr(f(\mathbf{A})) < Tr(f(\mathbf{B})) \tag{33}$$

Combining Eqs. (32) and (33), we can get:

$$Tr \left(f \left(\frac{1}{K} \sum_{k=1}^K (\mathbf{R}_{s,k} + \mathbf{R}_{n,k}) \right) \right) \succ Tr \left(f \left(\frac{1}{K} \sum_{k=1}^K \mathbf{R}_{n,k} \right) \right) \tag{34}$$

Equation (34) shows that when the authorized user occupies the spectrum, the left end of the formula is greater than the right end. The process of averaging the covariance matrix of each segment can stabilize the inequality. If T_{MFD} is higher than the threshold γ_{MFD} , it is assumed that the authorized user is occupying the spectrum. The detection expression of MFD algorithm is as follows:

$$H_1 : T_{MFD} = Tr \left(f \left(\frac{1}{K} \sum_{k=1}^K \mathbf{R}_{x,k} \right) \right) > \gamma_{MFD} \tag{35}$$

The monotonically increasing matrix function $f(\cdot)$ used by the MFD algorithm will affect the performance of the entire algorithm, so it is called the detection algorithm based on matrix function. In addition, F. Lin pointed out that monotonically increasing matrix function $f(\mathbf{X}) = \mathbf{X}$ can provide near optimal detection performance for MFD detection algorithm [13].

3.4 Maximum-minimum eigenvalue ratio (MME) algorithm

The MME detection algorithm is based on the generalized likelihood ratio, does not require any prior information of the detection signal, and is suitable for spectrum sensing scenarios in satellite cognitive wireless networks. Under the framework of big data, the MME algorithm makes decisions based on the statistical characteristics of the sample subspace. First, it calculates the maximum and minimum eigenvalues of the sample covariance matrix, denoted as λ_{\max} and λ_{\min} , respectively. If the sample subspace only contains noise, then $\lambda_{\max}/\lambda_{\min} = 1$. If it contains the detection signal, then $\lambda_{\max}/\lambda_{\min} > 1$. According to this property, we can get the detection expression of MME algorithm as follows:

$$H_1 : T_{MME} = \frac{\lambda_{\max}}{\lambda_{\min}} > \gamma_{MME} \tag{36}$$

where γ_{MME} is the threshold value.

When the prior information of the authorized user signal is unknown, the MME algorithm can use the information obtained in the signal sample subspace. The signal and noise are distinguished by the ratio of the maximum and minimum eigenvalues of the sample covariance matrix. Therefore, it can provide better detection performance in low signal-to-noise ratio (SNR) and complex electromagnetic environment.

3.5 Absolute value of covariance (CAV) algorithm

Both MFD and MME algorithms do not need the signal of authorized user as prior information for detection, but they need to find the eigenvalue of sample covariance matrix. When the dimension of covariance matrix is large, the computational complexity is very high, and the detection complexity and detection time will increase accordingly. CAV algorithm has low complexity and does not need prior information. First, the covariance matrix \mathbf{R}_x of the current subspace $\Gamma_{x,i}$ is calculated. Second, the variables T_1 and T_2 are used for detection.

$$T_1 = \frac{1}{L} \sum_{i=1}^L \sum_{j=1}^L |r_{ij}| \tag{37}$$

$$T_2 = \frac{1}{L} \sum_{i=1}^L |r_{ii}| \tag{38}$$

where r_{ij} , $1 \leq i \leq L$, $1 \leq j \leq L$ are the elements in covariance matrix. The detection expression of CAV algorithm is:

$$H_1 : T_{CAV} = \frac{T_1}{T_2} > \gamma_{CAV} \tag{39}$$

3.6 Eigenvalue exponential mean (EEM) algorithm

The spectrum sensing algorithm based on big data analysis theory essentially describes the statistical characteristics of subspace in a way that is easy to quantify and compare under the framework of big data. In satellite cognitive radio networks, cognitive base stations and cognitive users often do not know the prior information of the authorized user signal. In order to better distinguish between signal and noise in this “blind detection” state, the spectrum sensing algorithm needs to make full use of the subspace containing the current received signal. Based on this, an eigenvalue exponential mean (EEM) algorithm is proposed to provide

better detection performance without knowing the prior information of authorized user signals.

Similar to MME algorithm, the proposed EEM algorithm also uses the eigenvalues of the covariance matrix to distinguish the signal and noise. However, MME algorithm only uses the ratio of the maximum and minimum eigenvalues to distinguish. In the case of low SNR, the size of eigenvalues will be easily affected by noise, which will reduce the stability and performance of the algorithm.

The EEM algorithm uses all the eigenvalues $\{\lambda_i(\mathbf{R}_x)\}$ to make decision, and improve the robustness and detection performance. Specifically, the EEM algorithm takes the exponential mean of L eigenvalues of the sample covariance matrix as the judgment basis. The detection expression of the EEM algorithm is as follows:

$$H_1 : T_{EEM} = \ln\left(\frac{1}{L} \sum_{i=1}^L e^{\lambda_i}\right) > \gamma_{EEM} \tag{40}$$

where γ_{EEM} is the threshold. If the sample subspace only contains noise, then $T_{EEM} = \ln(L \times e/L) = 1$. If it contains authorized signal, then $T_{EEM} > 1$.

4 Simulation results and analysis

4.1 Algorithm performance comparison

The false alarm probability is set to 1%, and the Rice factor K of Rice channel is set to 6 dB. The cognitive radio spectrum sensing algorithms based on big data discussed above are all applied to satellite cognitive radio networks, and the detection probabilities under different SNR are compared, as shown in Fig. 2. In order to improve the

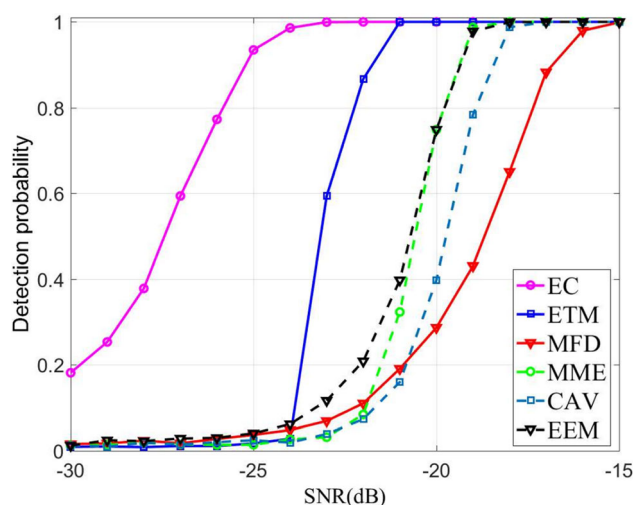


Fig. 2 Detection probability of spectrum sensing algorithm with SNR (false alarm probability is 1%, 10 sensing nodes, rice factor $K = 6$ dB)

robustness of the simulation, the algorithm uses Monte Carlo simulation 2000 times in each SNR.

The spectrum sensing algorithm based on big data analysis can provide good spectrum sensing performance in the low SNR environment of -15 dB. The EC algorithm has the best detection performance and can accurately detect the authorized signals with the SNR of -23 dB. In the actual sensing process, it is often impossible to obtain the prior information of authorized users, so the EC algorithm is not applicable in actual engineering. The performance of ETM algorithm is second only to EC algorithm, because it can automatically extract the prior information. However, when the SNR is lower than -24 dB, the sensing node is affected by noise and it is difficult to extract the authorized user signal from the subspace, and the detection probability is very low. When the SNR increases, the signal features of authorized users are easier to extract. Once the prior information is accurately extracted, the detection probability will be greatly improved.

For FMD, MME, EEM and CAV algorithms, which do not depend on the prior information, their detection performance is lower than the former two algorithms. However, these algorithms do not require prior information and have lower complexity, are easier to implement in engineering, and have more practical significance. Among these “blind detection” algorithms, the EEM algorithm has the highest detection probability. However, the EEM algorithm needs to calculate eigenvalues, which has higher requirements for the computing power of the fusion center. The MFD algorithm has the lowest detection probability when the SNR is higher than -21 dB, but it only needs to calculate the trace of the sample covariance matrix, so it is easier to implement.

4.2 Influence of channel conditions on algorithm performance

The Rician factor is increased to 10 dB, and other system parameters remain unchanged. For each SNR, the detection probability of the spectrum sensing algorithm based on big data analysis is simulated by Monte Carlo 2000 times, and the simulation result is shown in Fig. 3.

Comparing Figs. 2 and 3, when the size of the subspace is unchanged, the detection probability of various spectrum sensing algorithms changes similarly. In addition, increasing the Rician factor can improve the detection probability to a certain extent. Therefore, in the satellite cognitive wireless network, the sensor node should be located in an open area with good channel conditions, and try to avoid the shelter of trees and tall buildings.

On the other hand, the increase of the Rician factor can greatly improve the detection performance of EC, ETM,

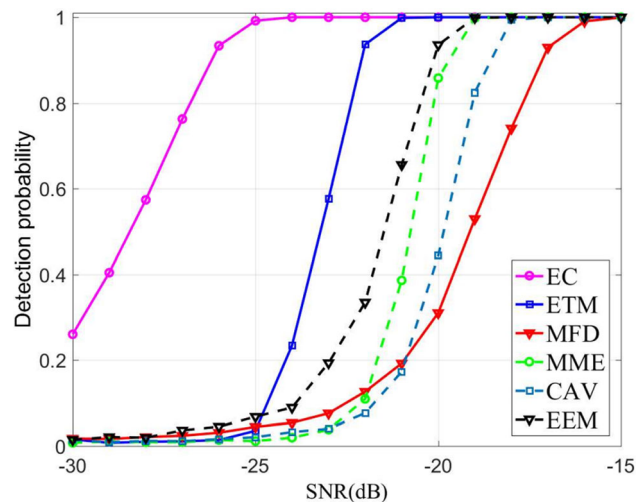


Fig. 3 Detection probability of spectrum sensing algorithm with SNR (false alarm probability is 1%, 10 sensing nodes, rice factor $K = 10$ dB)

MME and EEM algorithms, enabling the ETM algorithm to obtain the signal characteristics of authorized users from the sample subspace with a lower SNR. However, the detection performance of MFD and CAV algorithms are not sensitive to changes in channel conditions.

4.3 Influence of sensing nodes number on algorithm performance

The number of sensor nodes in the satellite coverage area is set to 20, the false alarm probability is 1%, and the Rice coefficient is 10 dB. In a sensing process, the sensing node samples the signals of authorized satellite users 50 times, and the sampling frequency remain unchanged. Therefore, 128,000 data points are collected for each sensing process and converge to form a subspace at the fusion center. The curve of the detection probability is shown in Fig. 4.

Comparing Figs. 3 and 4, the detection performance can be improved by increasing the number of sensing nodes under the same channel conditions. As the subspace becomes larger, it is easier for the fusion center to extract statistical features under the framework of big data, and to better distinguish between signal and noise.

In order to obtain better detection performance in satellite cognitive wireless networks, the number of sensing nodes and sampling time can be increased. However, this will increase the cost of data transmission between the sensing node and the fusion center, the amount of calculation of the fusion center, and the sensing time. Therefore, it is necessary to make a trade-off between perception performance and system complexity according to the actual situation.

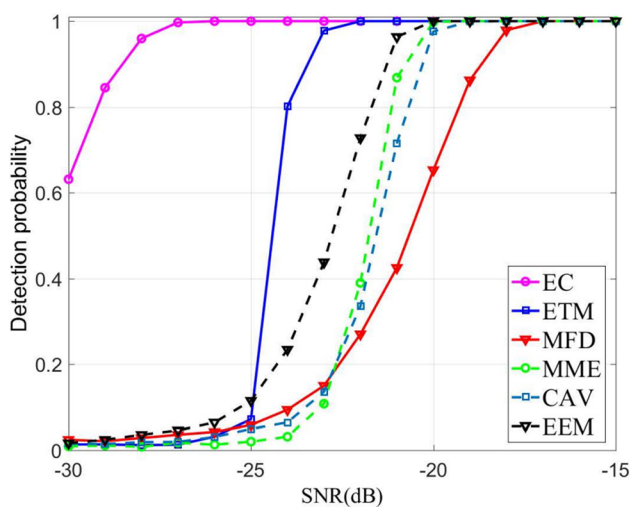


Fig. 4 Detection probability of spectrum sensing algorithm with SNR (false alarm probability is 1%, 20 sensing nodes, rice factor $K = 10$ dB)

5 Conclusion

In the context of satellite cognitive radio networks, traditional spectrum sensing algorithms based on big data analysis theory are studied, including EC, ETM, FMD, MME and CAV algorithms. These algorithms utilize the wide coverage of satellite network and the large number of sensing nodes, and transmit the sampling data of the sensing nodes to the fusion center to form a sample subspace. Make judgments based on different statistical rules of sample data under the framework of big data. This paper proposes the EEM algorithm based on big data analysis theory, which uses all the eigenvalues of the sample covariance matrix to make decision, and the detection performance is better than traditional algorithms. The simulation results show that under the condition of SNR above -18 dB, different algorithms can provide good detection performance. Based on big data analysis theory, the influence of channel conditions and the number of sensing nodes on the spectrum sensing algorithm is verified. In addition, compared with traditional algorithms, EEM algorithm provides higher detection probability without knowing the prior information of authorized users.

Acknowledgements The paper is sponsored by National Natural Science Foundation of China (No. 62071146), Projects of Science and Technology on Communication Networks Laboratory, Grant/Award Number: 6142104190113.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless

indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Liu, X., Zhai, X. B., Lu, W., & Wu, C. (March 2021). QoS-guarantee resource allocation for multibeam satellite industrial internet of things with NOMA. *IEEE Transactions on Industrial Informatics*, 17(3), 2052–2061. <https://doi.org/10.1109/TII.2019.2951728>
- Li, F., Lam, K., Liu, X., Wang, J., Zhao, K., & Wang, L. (2018). Joint pricing and power allocation for multibeam satellite systems with dynamic game model. *IEEE Transactions on Vehicular Technology*, 67(3), 2398–2408. <https://doi.org/10.1109/TVT.2017.2771770>
- Bayhan, S., Gur, G., & Alagoz, F. (2007). *Satellite assisted spectrum agility concept* (pp. 1–7). Military Communications Conference, Milcom.
- Celandroni, N., Ferro, E., Gotta, A., & Oligieri, G. (2013). A survey of architectures and scenarios in satellite-based wireless sensor networks: System design aspects. *International Journal of Satellite Communications & Networking*, 31(1), 1–38.
- Zhang, W., Yang, M., Wei, W., & Guo, Q. (2019). *Channel states information based energy detection algorithm in dual satellite systems*.
- Yu, C., Wan, P., Wang, Y., Wang, Y., Liang, T., & Wan, P. (2016). Spectrum sensing algorithm based on improved MME-Cyclic stationary feature. In *12th international conference on natural computation* (pp. 857–861).
- CoRaSat (Cognitive Radio for Satellite Communications), FP7 ICT STREP Grant Agreement n. 316779[EB/OL], <http://www.ict-corasat.eu/>
- Liolis, K., Schlueter, G., Krause, J., Zimmer, F., Combelles, L., Grotz, J., Chatzinotas, S., Evans, B., Guidotti, A., Tarchi, D., & Vanelli-Coralli, A. (2013). *Cognitive radio scenarios for satellite communications: The CoRaSat approach* (pp. 1–10). Future Network & Mobile Summit.
- Sharma, S. K., Chatzinotas, S., & Ottersten, B. (2012). Satellite cognitive communications: interference modeling and techniques selection. In *6th advanced satellite multimedia systems conference (ASMS) and 12th signal processing for space communications workshop (SPSC)* (pp. 111–118).
- Chatzinotas, S., Evans, B., Guidotti, A., Icolari, V., Lagunas, E., Maleki, S., Sharma, S. K., Tarchi, D., Thompson, P., & Vanelli-Coralli, A. (2016). Cognitive approaches to enhance spectrum availability for satellite systems. *International Journal of Satellite Communications & Networking*, 35.
- Liu, X., & Zhang, X. (Aug. 2020). NOMA-based resource allocation for cluster-based cognitive industrial internet of things. *IEEE Transactions on Industrial Informatics*, 16(8), 5379–5388. <https://doi.org/10.1109/TII.2019.2947435>
- Sengupta, S. K. (1993). *Fundamentals of statistical signal processing: Detection theory* (pp. 465–466). Prentice Hall PTR.
- Lin, F., Qiu, R. C., Hu, Z., Hou, S., Browning, J. P., & Wicks, M. C. (2012). Generalized FMD detection for spectrum sensing under low signal-to-noise ratio. *IEEE Communications Letters*, 16(5), 604–607.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

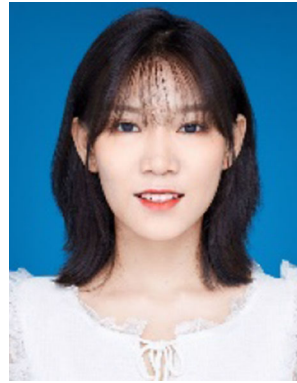


Mingchuan Yang (1977-) is currently an associate professor of Communication Research Center Harbin Institute of Technology, China. He received a B.Sc. degree in Mechanical Engineering from Harbin Institute of Technology, China, in 2001, and a M.Sc. degree and Ph.D degree in Information and Communication Engineering from Harbin Institute of Technology, China, in 2004 and 2010 respectively His research interests are in the fields of

space information network and mobile propagation channel modeling, especially mobile satellite and High Altitude Platforms (HAPs). He also has done some research on the quick reconstruction of the wireless networks under complicated electromagnetic environments. (1977-) is currently an associate professor of Communication Research Center Harbin Institute of Technology, China. He received a B.Sc. degree in Mechanical Engineering from Harbin Institute of Technology, China, in 2001, and a M.Sc. degree and Ph.D degree in Information and Communication Engineering from Harbin Institute of Technology, China, in 2004 and 2010 respectively His research interests are in the fields of space information network and mobile propagation channel modeling, especially mobile satellite and High Altitude Platforms (HAPs). He also has done some research on the quick reconstruction of the wireless networks under complicated electromagnetic environments.



Guanchang Xue (1998-) is currently a Ph.D student of Harbin Institute of Technology. Main research direction: satellite communication, satellite and terrestrial integrated network.



Bingyu Xie (1998-) is currently a M.Sc student of Harbin Institute of Technology. Main research direction: satellite and terrestrial integrated network, dynamic spectrum sharing technology, co-frequency interference cancellation.



Xinye Shao (1992-) received the B.Sc. degree and M.Sc. degree in Information and Communication Engineering from Harbin Institute of Technology, China, in 2013 and 2015 respectively. He is currently pursuing his Ph.D degree at Harbin Institute of Technology. His research interests include random access, the Internet of Things, and satellite networks.