Gaussian Mixture Model based Anomaly Detection for Defense Against Byzantine Attack in Cooperative Spectrum Sensing

Ashok Parmar, *Student Member, IEEE*, Karan Shah, Kamal Captain, *Member, IEEE*, Miguel López-Benítez, *Senior Member, IEEE*, Jignesh Patel, *Member, IEEE*

***Abstract*—Cognitive radio (CR) serves as an effective solution to the spectrum scarcity issue in current wireless communication. Spectrum sensing is one of the key enabling technologies for CR. Spectrum sensing based on single user detection often suffers from wireless channel ailments such as path loss, fading, and shadowing. Cooperative spectrum sensing (CSS) is proposed to overcome these adverse channel effects. However, CSS is often an easy target for malicious users (MUs). The Byzantine attack is a major hurdle in the success of the CR. Hence, the identification of MUs in the CR network is essential to improve the detection performance of CSS. In this work, we propose a Gaussian mixture model based anomaly detection algorithm for the identification of MUs. We first show that the presence of MUs degrades the CSS performance. Theoretical analysis is carried out to understand the intuition behind the proposed algorithm. The effectiveness of the proposed algorithm in detecting attackers is demonstrated for different attack scenarios. The performance of the proposed algorithm in detecting MUs is compared with existing algorithms. Based on the MU detection algorithm, a weighted sum based CSS algorithm is proposed that can eliminate the effects of attackers on the CSS performance.**

***Index Terms*—Cognitive Radio, Cooperative Spectrum Sensing, Byzantine Attack, Gaussian Mixture Model.**

I. Introduction

An exponential rise in the number of wireless services has been caused by the growing demand for wireless communication in many aspects of human life. Due to this exponential increase, the electromagnetic spectrum is being overcrowded, and hence there is a shortage of spectrum to host the upcoming wireless services. The Federal Communications Commission (FCC) has reported that some spectrum bands are underutilized. This is due to the current static spectrum allocation policy [1]. This motivates the researchers to change the spectrum allocation policy. Dynamic spectrum allocation can help increase spectrum utilization by identifying unused frequency bands and allocating them to other applications. These unused frequency bands are called spectrum holes. The cognitive radio (CR) technology makes it possible for unlicensed users, also known as secondary users (SUs), to use licensed frequencies without interfering with authorized users [2]. The identification of spectrum holes is called spectrum sensing, which is one of the significant tasks of CR. In CR, the SUs sense the presence or absence of the primary user (PU), which has access to the spectrum, by spectrum sensing and access the spectrum accordingly.

The detection performance of spectrum sensing using a single SU suffers from wireless channel effects such as multipath fading and shadowing. Cooperative spectrum sensing (CSS) is an ideal candidate to overcome these issues. In CSS, multiple SUs utilize local spectrum sensing technology, reporting their individual spectrum data to a fusion center (FC). The FC is responsible for determining the PU status and channel access based on a fusion rule. Cooperative spectrum sensing in the open wireless environment may be tricked by malicious alliances and vulnerable to spectrum sensing attacks [3]. The users who perform this kind of practice are called malicious users (MUs) or Byzantine attackers, and the practice is called the Byzantine attack also known as the spectrum sensing data falsification attack (SSDF). The Byzantine attack is typically carried out on purpose by malicious users with the objective to mislead honest secondary users (HSUs) into learning that the PU channel is occupied when it is free and to mislead HSUs into learning that the PU channel status is free in case it is occupied [4], [5]. Although the above mentioned objectives are listed separately, this does not imply that a malicious user must only carry out attacks for one of them. Malicious users may attack for both reasons to maximize their attack utility [6]. Hence, it becomes necessary to identify and eliminate the Byzantine attackers from the CR network.

A. Related Works

Researchers have proposed techniques to detect the MUs from the CR network [7]. A Bayesian learning based defense strategy against Byzantine attacks is investigated in [8]. In [9], a robust defense framework is proposed for defense against data falsification attacks, and the effectiveness of the proposed algorithm is analyzed under different practical scenarios. An entropy-based weighted CSS scheme is proposed in [10] for performance improvement under Byzantine attack. An effective entropy based weighted algorithm is recently proposed

Ashok Parmar and Kamal Captain are with the Department of Electronics Engineering, Sardar Vallabhbhai National Institute of Technology, Surat, India (email: {ds20ec010, kamalcaptain}@eced.svnit.ac.in)

Karan Shah is with the Department of Electronics and Communications Engineering, Institute of Technology, Nirma University, India (email: 20bec052@nirmauni.ac.in)

Miguel López-Benítez is with the Department of Electrical Engineering and Electronics, University of Liverpool, UK, and also with the ARIES Research Centre, Antonio de Nebrija University, Madrid, Spain (email: M.Lopez-Benitez@liverpool.ac.uk)

Jignesh Patel is with the Indian Institute of Information Technology Vadodara-ICD, Diu, Gujrat, India (email: jignesh_patel@diu.iiitvadodara.ac.in)

for defense against Byzantine attack in [11]. A sequential $0/1$ for CSS algorithm is proposed in [12] where cooperating secondary users (CSUs) are restricted to report either 0 or 1, and it is shown that the proposed algorithm improves CSS performance under SSDF attack. A similar algorithm called sequential single voting (SSV) rule is proposed in [13], where trust value mechanism is used to evaluate the reliability of the data received from CSUs. The information received from MUs are suppressed using thresholding. A secure CSS called generalized voting-sequential and differential reporting is proposed in [14]. Authors also propose a single signaling transmission scheme similar to SSV. In [15], a robust weighted algorithm based on the reputation value called weighted differential sequential symbol is proposed. A double reputation based detection algorithm is proposed in [16] to separate out MUs from honest users. In [17], an online learning based algorithm for cooperating user selection in CSS is proposed. A joint spectrum sensing and resource allocation algorithm is proposed in [18], where spectrum resources are allocated to different CSUs based on their past behavior. Reduced resources are allocated to CSUs that are performing poorly, incentivizing them for good behavior and motivating them to stop attacking or leave the network. The techniques reported in [8]–[18] consider hard combining based CSS. The performance of CSS under the soft combining rule in the presence of Byzantine attack is analyzed in [19]. To detect malicious users, a block outlier method (BOM) is proposed in [20] which uses a block-wise approach to estimate the energy levels of the received signals and identifies the outliers within each block. In [21], the performance of CSS is analyzed under the presence of Byzantine attackers and an algorithm is proposed where the fusion center removes most likely malicious value from the data fusion improving the performance under attack. A modified outlier removal (ROM) based secure fusion strategy is proposed in [22], where the FC conditionally removes outliers from the data fusion process which are most likely to have come from MUs.

Recently, researchers have started exploring the use of machine learning for defense against Byzantine attack [23], [24]. An unsupervised machine learning based MUs detection algorithm is investigated in [23]. The authors used a one-class support vector machine (SVM) to identify the MUs. In [24], linear and non-linear SVM algorithm is proposed for PU boundary detection. The effects of MUs on PU boundary detection are also investigated. Three different versions of the K-nearest neighbor (KNN) algorithm are explored for attacker detection. An SVM based CSS algorithm against Byzantine attack is proposed in [25]. In [26], a two stage algorithm based on the combination of semisupervised SVM and fast convergence K-means algorithm is proposed for defense against SSDF attack. The algorithm in [26] requires a set of labeled dataset which is difficult to get in ever changing CR environment. A semi-supervised fuzzy c-means-based (SSFCM) detection algorithm is proposed in [27] for defense against SSDF attack. However, the paper considers the quantise soft combining based CSS whereas in this paper soft combining is considered. The SVM and c-means clustering algorithms form circular clusters, which can be ill-suited for non-circular data distributions, leading to poor model fits. Moreover, these algorithms are computationally expensive, limiting their practical use. In contrast, the Gaussian Mixture Model (GMM) is a soft clustering approach well-suited for overlapping and non-circular clusters of Gaussian distributed data. Given that data received at the FC in CSS often exhibits such characteristics, GMM emerges as the preferred choice for modeling the data distribution. Therefore, we propose a GMM-based attacker detection algorithm in this work, with further key contributions detailed in the next section.

B. Contributions

In this work, we consider soft combining based CSS in the presence of a Byzantine attack. The key contributions of the work are as follows:

- The GMM based anomalies detection algorithm is proposed for MUs detection.
- A theoretical analysis is carried out to derive the expression for the probability density function (pdf) of the received energy at the FC from the MUs, which is then used to build an intuition for the proposed algorithm.
- The effectiveness of the proposed algorithm in detecting the MUs is demonstrated by considering various attack scenarios in the result section.
- Finally, it is also shown that the proposed algorithm can also be used to decide weights that can be assigned to CSUs for weighted sum algorithm at the FC. It demonstrated that using weights obtained using proposed algorithm for different CSUs, the performance of CSS can be improved significantly under attack from MUs and the effects of MUs on the performance of CSS can be completely eliminated.

Rest of the paper is organized as follows. Section II presents the system model and attack model considered in the paper. The proposed GMM based algorithm is discussed in section III. Theoretical analysis is presented in section IV followed by the results and discussion in section V. Finally, section VI concludes the paper.

II. System and Attack Model

In this section, we first discuss the system model considered in this paper and then we discuss the Byzantine attack model.

A. System Model

In this work, we consider a single primary user and M number of CSUs collaborating to sense the presence or the absence of the PU. The received signal at the i$^{\text{th}}$ CSU can be modeled as

$$y_i(n) = \begin{cases} w_i(n); & H_0, \\ h_i \cdot s_i(n) + w_i(n); & H_1, \end{cases} \tag{1}$$

where $s_i(n)$ and $w_i(n)$ are the n^{th} sample of the primary user signal and the additive white Gaussian noise (AWGN), respectively, with $n = 1, 2, \cdots, N$ and $i = 1, 2, \cdots, M$, and h_i represents the channel gain between PU and the i^{th} CSU. The main objective of this paper is to explore the application

of GMM in detecting Byzantine attackers. Therefore, our discussion is limited to the AWGN channel, and we assume a constant channel gain of $h_i = 1$. Similar channel model is assumed in [20], [22], [28]–[30]. Nonetheless, it is important to note that the analysis presented in this paper can readily be extended to incorporate fading effects by considering h to follow a specific fading distribution. The performance of spectrum sensing over fading channels has been investigated thoroughly showing that the effect of fading with respect to AWGN is a reduction of the detection performance that affects different sensing approaches to similar extents [31]–[34]. Therefore, if a given method performs better than another one in an AWGN channel model it can also be expected to perform better under channel models with fading and shadowing. However, the performance would be slightly lower under fading than in AWGN, but would still improve those methods that it is able to outperform under AWGN. Additionally, we make the assumption of a homogeneous cognitive radio network, where all CSUs operate in similar environments, resulting in comparable signal-to-noise ratios across the network. The signal and the noise samples are independent and identically distributed (i.i.d) with $s_i(n) \sim \mathcal{N}(s; 0, \sigma^2_{s_i})$ and $w_i(n) \sim \mathcal{N}(w; 0, \sigma^2_{w_i})$. The notation $\mathcal{N}\left(x; m, \sigma^2\right)$ corresponds to the standard Gaussian distribution with argument x and it is parameterized by mean m and variance σ^2. The free and occupied primary channels are represented by hypotheses H_0 and H_1, respectively. In this work, we consider centralised CSS where each CSU performs energy detection [35], computes the received signal's energy and reports it to the FC which takes the final decision on the occupancy of the PU channel. The energy, i.e., decision statistic, at the i^{th} CSU is computed as

$$E_i = \frac{1}{N} \sum_{n=1}^{N} |y_i(n)|^2. \tag{2}$$

After receiving the decision statistic from all M CSUs, the FC computes final decision statistic T_{FC} as

$$T_{FC} = \frac{1}{M} \sum_{n=1}^{M} E_i. \tag{3}$$

The FC compares T_{FC} with the threshold τ and if $T_{FC} < \tau$ PU channel is declared as free otherwise it is declared as busy.

B. Attack Model

In this work, we consider a Centralized Independent Probabilistic Small-Scale (CIPS) attack with M users out of which M_L are malicious users [7]. In our analysis, we are examining a soft Byzantine attack scenario. In this scenario, the MUs monitor the energy levels of the channel. When they detect that the energy exceeds a threshold value denoted as η, they occasionally decrease the energy by a certain amount Δ, with a probability of α_1. Conversely, if the sensed energy falls below η, they sometimes increase the energy by Δ, with a probability of α_0. If the MUs attack with $\alpha_0 = \alpha_1 = 1$, the attack scenario is considered to be a hard Byzantine attack. These modified energies are sent to FC. The attack model can be mathematically formulated as follows

$$\begin{cases} P(E_m^{'} = E_m + \Delta | E_m < \eta) = \alpha_0; \\ P(E_m^{'} = E_m - \Delta | E_m > \eta) = \alpha_1; \end{cases} \tag{4}$$

where $m = 1, 2, \ldots, M_L$, E_m is the malicious user's observed energy, $E_m^{'}$ is falsified sensing results from m$^{\text{th}}$ MU, α_0 is the attack probability of false alarm, α_1 is the attack probability of miss detection, η represents the attack threshold, and Δ is a positive value called the attack strength. The attack strength Δ represents the degree to which a malicious user diverges from their actual measurement when reporting it to the fusion center. The greater the value of Δ, the larger the deviation. If malicious users attack with a higher Δ, the likelihood of the fusion center making an incorrect decision based on receiver information increases, thus raising the probability of a successful attack. However, malicious users employing a high Δ are more likely to be detected by the fusion center, as their reports deviate significantly from other reports. Hence, the MUs generally do not attack with abnormally high values.

C. Dataset

In this work we make use of simulated dataset. We consider M number of CSUs out of which there are M_L number of CSUs which are malicious. The MUs attack with false alarm attack probability α_0 and miss detection attack probability α_1. We assume that the number of MUs in the CR network is less than the number of honest users, i.e., $M_L < \frac{M}{2}$ [7], [36]. At each sensing instance (i.e., each sensing round), all the CSUs perform spectrum sensing on the PU channel using energy detection and compute the energy using Eq. (2). Each CSU reports its energy value to the FC. The energy vector received at the FC at the j^{th} sensing instance is given by

$$E^j = \left[E_{j,1}, E_{j,2}, \ldots, E_{j,M}\right]^T, \tag{5}$$

where $[\cdot]^T$ represents the transpose operator. The energy vector reported at the FC are recorded for every sensing instance to create the dataset. At every sensing instance, the PU is considered to be active with probability p so that we have data samples consisting of both the hypothesis H_0 and H_1. Let us say the data is recorded for P number of sensing instances, then the dataset will be of $P \times M$ dimension. Experiments are carried out for different parameter settings.

III. Proposed GMM based algorithm

Utilizing the Central Limit Theorem, we observe that the data follows a Gaussian distribution, making the application of GMM an appropriate choice for modeling the data distribution. In this scenario with only two classes, GMM can be easily applied to capture the Gaussian mixture distribution. After obtaining the model, we can straightforwardly determine outliers by assessing the likelihood of each data point. Data points residing in the lower probability regions are identified as anomalies. These anomaly counts are subsequently employed in the proposed algorithm for the detection of MUs. Next, we will discuss the GMM clustering algorithm, followed by

the GMM-based attacker detection algorithm and the weighted CSS algorithm. GMM is an unsupervised machine learning algorithm for clustering. The GMM fits mixture of Gaussian on the training data and forms the clusters by finding parameters of Gaussian distributions lying in the data. In this work, due to the one dimensional nature of the data, the univariate GMM model is used. The mixture Gaussian probability density function (pdf) with K number of Gaussian pdfs is given by

$$f\left(x;\mu,\sigma,\phi\right)=\sum_{j=1}^{K}\phi_j\mathcal{N}\left(x;\mu_j,\sigma_j\right), \tag{6}$$

where ϕ_j represents the wight of j^{th} distribution with $\sum_{j=1}^{K}\phi_j=1$ and $\mathcal{N}\left(x;\mu_j,\sigma_j\right)$ is the Gaussian pdf such that

$$\mathcal{N}\left(x;\mu_j,\sigma_j\right)=\frac{1}{\sigma_j\sqrt{(2\pi)}}\exp\left(-\frac{(x-\mu_j)^2}{2\sigma_j^2}\right). \tag{7}$$

Here, μ_j and σ_j^2 represent the mean and variance of j^{th} distribution, respectively, and $j=1,2,\ldots,K$. The GMM estimates the unknown parameters μ_j,σ_j and ϕ_j for each cluster using expectation maximization (EM) algorithm [37]. Let us consider a training vector $\bar{E}=\{E_1,...,E_L\}$ with L training samples. The EM algorithm is used to maximize the log-likelihood function given by

$$U\left(\bar{E};\mu,\sigma,\phi\right)=\sum_{l=1}^{L}\ln\left(\sum_{j=1}^{K}\phi_j.\mathcal{N}\left(E_l;\mu_j,\sigma_j\right)\right). \tag{8}$$

An EM is an iterative algorithm which iteratively optimizes the parameters of GMM. The steps involved in EM are as follows

1) Firstly, the parameters μ_j,σ_j and ϕ_j are initialized randomly to calculate the initial value of log-likelihood using Eq. (8).
2) **E step:** Using the current parameter values, quantities $v(z_{lj})$ called responsibilities are evaluated as

$$v(z_{lj})=\frac{\phi_j\mathcal{N}\left(E_l;\mu_j,\sigma_j\right)}{\sum_{p=1}^{K}\phi_p\mathcal{N}\left(E_l;\mu_p,\sigma_p\right)}. \tag{9}$$

3) **M step:** The parameters μ_j,σ_j and ϕ_j are re-estimated using the current responsibilities as

$$\mu_j^{new}=\frac{1}{L_j}\sum_{l=1}^{L}v\left(z_{lj}\right)E_l, \tag{10}$$

$$\sigma_j^{new}=\sqrt{\frac{1}{L_j}\sum_{l=1}^{L}v\left(z_{lj}\right)\left(E_l-\mu_j^{new}\right)^2} \tag{11}$$

$$L_j=\sum_{l=1}^{L}v\left(z_{lj}\right)\ \text{ and }\ \phi_j^{new}=\frac{L_j}{L}. \tag{12}$$

4) The log-likelihood is again evaluated using new parameters and checked if the log-likelihood is converged. Convergence can be assessed by examining either the parameter values or the log-likelihood function. If these values exhibit negligible changes, typically denoted by a small value such as ϵ, it indicates that convergence has been achieved. If the convergence criteria is not satisfied then return to E step (step 2) and perform another EM iteration.

Next, we discuss the GMM based anomaly detection algorithm for MUs detection. Since we know that the received energy in CSS corresponds to H_0 or H_1, we fit GMM on the data using $K=2$. As the likelihood function in Eq. (8) exhibits multiple local maxima, the GMM algorithm may not always converge to the global maximum; instead, it may converge to a local maximum based on the initial conditions. To address this challenge, the EM algorithm is executed iteratively with various initializations, thereby increasing the probability of discovering the global maximum. The parameters that yield the maximum log-likelihood are selected among the multiple runs of the EM algorithm. In our experimentation, we have executed EM for GMM with ten different initializations. Once the GMM has converged, it can be used to find the anomalies, also known as the outliers. Note that the anomaly and outlier words are used interchangeably in this paper. The anomalies are obtained using the GMM based algorithm, and the anomaly counts are derived for all the CSUs. The number of outliers in a particular CSU is used to decide whether the CSU is malicious or not. The anomaly counts for different users are compared with the threshold, and if the anomaly count of a particular CSU is less than the threshold, then it is declared as a malicious user; otherwise, it is declared as a genuine CSU. The steps involved in the proposed approach are mentioned in Algorithm 1. We use the OTSU algorithm to derive the threshold given in Algorithm 2. The algorithm in Algorithm 1 considers only one PU. However, it can be readily extended to accommodate multiple PUs. In this paper, we have considered the case where $K=2$, meaning we fitted a mixture of two Gaussian pdfs to the data received at the FC using the GMM. In the scenario with more PUs, the number of Gaussian pdfs used for modeling the mixture of Gaussian pdfs will depend on the total number of PUs considered. If there are, for example, P PUs, we would need to consider $K=P+1$ pdfs to fit the data. Consequently, we can apply the GMM to model the data by fitting a mixture of Gaussian pdfs to the information received at the FC. Once the model is successfully fitted, we can apply the proposed algorithm to detect MUs, as detailed in Algorithm 1. Once attackers are detected, the FC has the option to exclude the compromised CSUs from cooperation. While removing the MUs from the CRN may seem like a logical step if the MUs are always attacking, it is important to note that MUs often refrain from constant attacks to avoid easy detection. In such scenarios, we can employ a weighted sum-based algorithm. This approach does not involve simply removing the MUs from the network but instead employs a weighted algorithm that assigns less significance to the data received from them. The proposed algorithm for MUs detection can also be utilized to determine the appropriate weights for the weighted sum-based CSS algorithm. The detailed description of the weighted sum-based algorithm is provided in Algorithm 3.

Algorithm 1 Anomaly Detection Algorithms

- **Step 1:** Reshape the dataset of size $P \times M$ in the shape $L \times 1$ to form $\bar{E}$ training vector, where $L = M \cdot P$.
- **Step 2:** Set $K = 2$ and fit GMM on the data using EM algorithm to find the parameters μ_j, σ_j and ϕ_j for $j = 1, 2$.
- **Step 3:** With the trained GMM, compute density for each element in vector $\bar{E}$.
- **Step 4:** Set density threshold by flagging ω % of lower density values as anomalies or outliers.
- **Step 5:** Declare elements in vector $\bar{E}$, which are having densities lower than the density threshold as anomalies.
- **Step 6:** From the total anomalies detected in $L \times 1$ vector, calculate the number of anomalies coming from each CSU in the original $P \times M$ dataset.
- **Step 7:** Declare CSU as MU if the number of anomalies in that CSU is less than threshold $\lambda = \min[\lambda_{otsu}, \Omega\% \text{ of } (\omega\% \text{ of } P)]$, where Ω defines the expected percentage of minimum anomalies from GUs and λ_{otsu} is computed using an OTSU Algorithm [38] given in Algorithm 2.

Algorithm 2 OTSU Thresholding to compute λ_{otsu}

Input: Let a vector $A = [A_1, A_2, ..., A_M]$ contains the number of anomalies in each CSU. Initially take i = 1.

- **Step 1:** Choose A_i as threshold and split the vector A in lower (A_l) and upper (A_u) sets as

$$A_l = \{A_j : \forall A_j < A_i\}$$

$$A_u = \{A_j : \forall A_j \geq A_i\},$$

  for $j = 1, 2, \ldots, M$.
- **Step 2:** Find: N_{all} number of elements in A, N_l Number of elements in A_l and N_u number of elements in A_u.
- **Step 3:** Compute weights W_l and W_u as: $W_l = \frac{N_l}{N_{all}}$, $W_u = \frac{N_u}{N_{all}}$
- **Step 4:** Compute variances σ_l^2 and σ_u^2 of A_l and A_u respectively.
- **Step 5:** Find the variance for the threshold A_i using: $\sigma_{A_i}^2 = W_l \sigma_l^2 + W_u \sigma_u^2$.
- **Step 6:** Increase i by 1 and repeat steps 1 to 5 for each element in A.
- **Step 7:** The threshold with least variance $(\sigma_{A_i}^2)$ is selected as λ_{otsu}.

IV. THEORETICAL ANALYSIS

In this section, we first give the theoretical analysis for the CSS and derive the probability of false alarm (Q_f) and detection (Q_d).

A. Detection Probabilities for CSS

In CSS, each CSUs compute energy of the received signal using Eq. (2) and report it to the FC. For relatively large N, using central limit theorem, the energy computed at the CSUs can be modeled using a Gaussian distribution [21], [22], [39]. Using this, the pdf of energy at the i^{th} CSU is given by

$$f_{T_i}(x) = \begin{cases} \mathcal{N}\left(x; \mu_0, \sigma_0^2\right); & H_0, \\ \mathcal{N}\left(x; \mu_1, \sigma_1^2\right); & H_1, \end{cases} \tag{15}$$

where $\mu_0 = \sigma_w^2$, $\sigma_0^2 = \frac{2}{N}\sigma_w^4$, $\mu_1 = (1+\gamma)\sigma_w^2$, $\sigma_1^2 = \frac{2}{N}(1+\gamma)^2 \sigma_w^4$ and $\gamma = \frac{\sigma_s^2}{\sigma_w^2}$ represents the signal to noise ratio. The FC computes the decision statistic as given in Eq. (3) which is the sum of energies received from all the CSUs scaled by the number of CSUs M. Since the pdf of energy at the CSU is modeled using Gaussian pdf, its sum will also follow Gaussian pdf. Using this property, the pdf of decision statistic at the FC can be modeled as

$$f_{FC}(x) = \begin{cases} \mathcal{N}\left(x; \mu_{fc,0}, \sigma_{fc,0}^2\right); & H_0, \\ \mathcal{N}\left(x; \mu_{fc,1}, \sigma_{fc,1}^2\right); & H_1, \end{cases} \tag{16}$$

where $\mu_{fc,0} = \mu_0$, $\sigma_{fc,0}^2 = \frac{\sigma_0^2}{M}$, $\mu_{fc,1} = \mu_1$ and $\sigma_{fc,1}^2 = \frac{\sigma_1^2}{M}$. Using the pdf derived in Eq. (16), the probabilities Q_f and Q_d can be obtained as

$$Q_f = Q\left(\frac{\tau - \mu_{fc,0}}{\sigma_{fc,0}}\right), \text{ and } Q_d = Q\left(\frac{\tau - \mu_{fc,1}}{\sigma_{fc,1}}\right), \tag{17}$$

where τ represents the detection threshold.

Algorithm 3 Weighted Sum based CSS Algorithm

Input: Let a vector $C = [c_1, c_2, ..., c_M]$ contains the number of anomalies in each CSU.

- **Step 1:** Compute weights for each CSU as

$$w_i = \frac{c_i}{\sum_{j=1}^{M} c_j} \tag{13}$$

  where $i = 1, 2, \ldots, M$.
- **Step 2:** Using the weights computed in step 1, derive the decision statistic at the FC as

$$T_{FC}^{weighted} = \frac{1}{M} \sum_{n=1}^{M} w_i \cdot E_i. \tag{14}$$

- **Step 3:** The FC compares $T_{FC}^{weighted}$ with the threshold τ and if $T_{FC}^{weighted} < \tau$, the PU channel is declared as free otherwise it is declared as occupied.

B. Intuition behind the Proposed Algorithm

In this section, we give theoretical basis for the intuition behind the proposed MUs detection algorithm. We first derive the probability density function of the energy received at the FC from the MUs. Let us consider the first scenario when the PU is absent. In this case, the pdf of reported energy under false alarm attack from the MUs can be modeled using truncated Gaussian distribution as

$$f_{H_0}^{fa}(x) = \begin{cases} \frac{\mathcal{N}(x-\Delta;\ \mu_0,\ \sigma_0^2)}{\frac{1}{2}\left[1+erf\left(\frac{\eta-\mu_0}{\sigma_0\sqrt{2}}\right)\right]}, & -\infty < x \leq \eta + \Delta \\ 0, & \text{elsewhere} \end{cases} \tag{18}$$

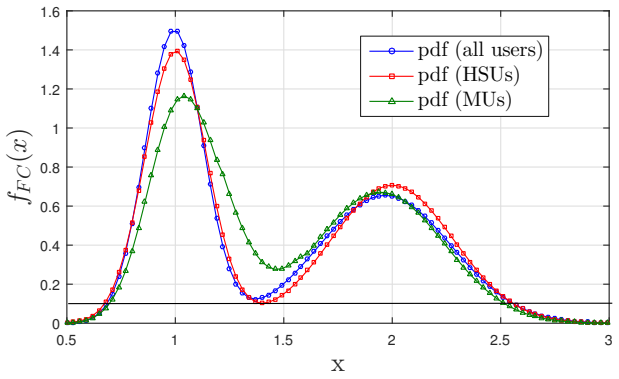

Fig. 1. Probability density functions considering different scenarios.

where $erf(x)$ is the error function. The detailed derivation of $f_{H_0}^{fa}(x)$ in Eq. (18) is given in Appendix A. Similarly, the pdf of reported energy under miss detection attack from the MUs can be modeled as

$$f_{H_0}^{md}(x) = \begin{cases} \frac{\mathcal{N}(x+\Delta;\ \mu_0,\ \sigma_0^2)}{1-\frac{1}{2}\left[1+erf\left(\frac{\eta-\mu_0}{\sigma_0\sqrt{2}}\right)\right]}, & \eta-\Delta \le x < \infty \\ 0, & \text{elsewhere} \end{cases} \tag{19}$$

The detailed derivation of Eq. (19) is given in Appendix B. The MUs make false alarm attack with probability α_0 and they attack when their computed energy is greater than attack threshold η. Also, the MUs carry out miss detection attack with probability α_1 and attacks when their measured energy is below η. Using this, the pdf of received energy at the FC under H_0 from the MUs can be obtained as

$$f_{FC,H_0}^{mu}(x) = \alpha_0 P_{H_0}^0 f_{H_0}^{fa}(x) + \alpha_1 P_{H_0}^1 f_{H_0}^{md}(x) + \left(1-\alpha_0 P_{H_0}^0 - \alpha_1 P_{H_0}^1\right) f_{H0}(x),\ -\infty < x < \infty. \tag{20}$$

where

$$P_{H_0}^0 = \int_{\eta}^{\infty} f_{H_0}(x)dx \text{ and } P_{H_0}^1 = \int_{-\infty}^{\eta} f_{H_0}(x)dx. \tag{21}$$

Following the similar procedure, the pdf of energy received at FC under hypotheses H_1 can be obtained as

$$f_{FC,H_1}^{mu}(x) = \alpha_0 P_{H_1}^0 f_{H_1}^{fa}(x) + \alpha_1 P_{H_1}^1 f_{H_1}^{md}(x) + \left(1-\alpha_0 P_{H_1}^0 - \alpha_1 P_{H_1}^1\right) f_{H1}(x),\ -\infty < x < \infty. \tag{22}$$

where

$$f_{H_1}^{fa}(x) = \begin{cases} \frac{\mathcal{N}(x-\Delta;\ \mu_1,\ \sigma_1^2)}{\frac{1}{2}\left[1+erf\left(\frac{\eta-\mu_1}{\sigma_1\sqrt{2}}\right)\right]}, & -\infty < x \le \eta+\Delta \\ 0, & \text{elsewhere} \end{cases} \tag{23}$$

$$f_{H_1}^{md}(x) = \begin{cases} \frac{\mathcal{N}(x+\Delta;\ \mu_1,\ \sigma_1^2)}{1-\frac{1}{2}\left[1+erf\left(\frac{\eta-\mu_1}{\sigma_1\sqrt{2}}\right)\right]}, & \eta-\Delta \le x < \infty \\ 0, & \text{elsewhere} \end{cases} \tag{24}$$

$$P_{H_1}^0 = \int_{\eta}^{\infty} f_{H_1}(x)dx, \text{ and } P_{H_1}^1 = \int_{-\infty}^{\eta} f_{H_1}(x)dx \tag{25}$$

In Fig. 1 we plot the pdfs considering different scenarios considering $M = 10$, $M_L = 3$, $\alpha_0 = \alpha_1 = 0.3$, $\eta = 1.4$, $\Delta = 0.2$, $N = 100$ and $\gamma = 0\ dB$. The dataset is generated using the mentioned parameters and GMM is fitted considering two classes. The pdf plot considering all users are plotted using the parameters obtained from GMM. The plot for pdf of HSUs, the weighted sum of pdfs given Eq. (16) under H_0 and H_1 is obtained considering weights obtained using GMM. Similarly, the pdf of MUs is obtained by taking the weighted sum of pdfs in Eq. (20) and Eq. (22) considering weights obtained in GMM. We can see in Fig. 1 that the pdf of HSU is closer to the pdf obtained after GMM whereas the pdf of MU is deviating from the other pdfs. The reason behind this is that the number of MUs are less than the number of HSUs and hence the parameters obtained after GMM will be biased towards the pdf of honest users (HUs). Since the attack is such that when the energy computed by MU is less than η, it attacks by adding Δ to it and when it is above η it attacks by subtracting Δ with certain probability, the pdf of MUs will shrink. The outliers are defined as the samples which lie in low probability region. In Fig. 1, the area under any curve below the horizontal line will give the probability of outliers. We can see that this probability will have small value for MU compared to the HU. For the particular position of the line in Fig. 1, the value obtained for MU is 0.0164 whereas for HU it is 0.0179. Hence, the number of outliers will be less in data received from MUs compared to those in HUs.

V. Results and Discussion

In this section, we first demonstrate the effects of the existence of MUs in the CSS network on the performance of CSS using the receiver operating characteristic (ROC) plot. It is also shown that the detection performance improves significantly once the MUs are detected and eliminated from cooperating. We then demonstrate the effectiveness of the proposed algorithm in detecting the MUs in different scenarios. First, we consider a scenario in which we have varied the activity patterns of the PU by incorporating different percentages of H_0 and H_1 cases when creating the dataset. Then, we explore various attack scenarios, including different attack probabilities and varying numbers of attackers within cognitive radio networks. Without loss of generality, we have kept first M_L users as MUs, and the remaining $M-M_L$ CSUs as honest users.

We proceed to illustrate the efficacy of our proposed algorithm in detecting MUs under diverse PU activity patterns. Employing a mixture model with $K = 2$, as expressed in Eq. (6), we employ GMM to fit mixture of two Gaussian pdfs, one corresponding to H_0 and the other to H_1, to our dataset. The proportion of each Gaussian distribution is determined by the PU's activity pattern. When the dataset exhibits a higher proportion of H_0 hypotheses, the respective Gaussian distribution within the mixture model will have high ϕ_1 value. We proceed to showcase the effectiveness of our proposed algorithm in MUs detection under varying proportions of H_0 and H_1 samples within the data. This demonstration highlights the algorithm's robustness in scenarios where clusters exhibit unequal sizes. We first consider the case when the samples in the dataset belong to 50% H_0 and 50% H_1 case at SNR

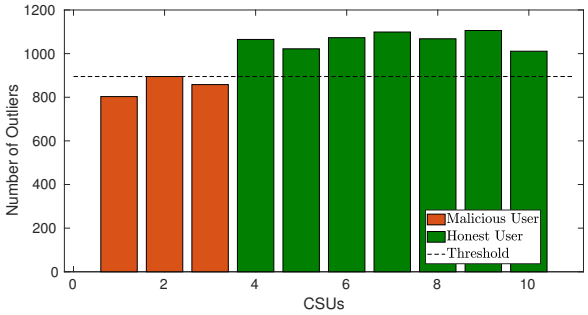

Fig. 2. The bar plot of outlier counts vs. CSU number considering $M = 10$, M_L=3, $\alpha_0 = 0.3$, $\alpha_1 = 0.3$ for H_0=H_1=50% in the dataset.

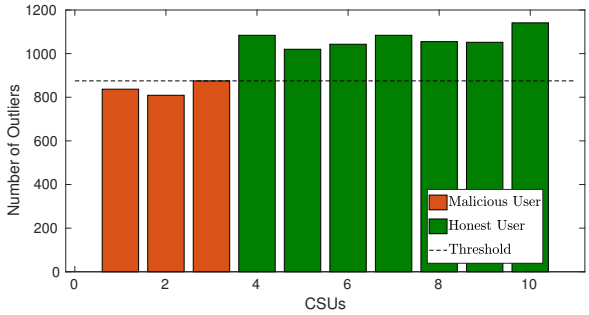

Fig. 3. The bar plot of outlier counts vs. CSU number considering $M = 10$, M_L=3, $\alpha_0 = 0.3$, $\alpha_1 = 0.3$ for $H_0 = 60\%$ and $H_1 = 40\%$ in the dataset.

$\gamma = -10$ dB. The number of sensing instances are kept as $P = 5000$ with $M = 10$ CSUs. The number of MUs is selected as $M_L = 3$ with attack threshold $\eta = 1.08$ and attack strength $\Delta = 0.2$. The size of the dataset is kept the same for all analyses in this section. In the proposed algorithm, ω is selected as 20 indicating 20% outliers, and the Ω is chosen to be 95. Using this, the threshold λ turns out to be 950. Please take note that increasing the value of ω will result in more energy values being flagged as anomalies, potentially causing non-anomalous data points to be misclassified as anomalies. Nevertheless, the key observation stands: MUs will consistently exhibit fewer anomalies compared to HSUs. As a result, the algorithm retains its effectiveness, even in scenarios where some valid reports are erroneously labeled as anomalies. Each CSU's outliers count is compared with the threshold λ, and if the count is higher than λ, the respective CSU is declared as an HU; otherwise, it is a MU. Fig. 2 shows the bar plot for outliers counted from each CSU using the proposed algorithm. The black horizontal line in the figure represents the threshold λ. The MUs are colored in red, which are the first three CSUs, which we intentionally kept during data generation. Hence, the results indicate that the proposed algorithm successfully detects the MUs.

Next, we consider the case when 60% samples belong to the H_0 case in the dataset. Hence, from the total of 5000 data samples, we have 3000 samples belonging to the H_0 case, and the remaining 2000 belongs to the H_1 case. Other parameters are kept the same as in previous section. The dataset is fed to the proposed algorithm for MUs detection, and the result outliers count values are plotted as bar plots in Fig. 3. The figure shows that the first three CSUs are declared as MUs, which are indeed the MUs. A similar analysis is carried out for the case when 80% data samples belong to the H_0 class and only 20% belong to the H_1 class indicating that the channel occupancy is very sparse. The plot in Fig. 4 indicates that the proposed algorithm efficiently detects the MUs. We see that the proposed algorithm is efficient in detecting MUs under different proportion data samples belonging to two hypothesis H_0 and H_1 in the dataset. In Fig. 5, we demonstrate the effectiveness of the proposed algorithm in detecting the MUs under different attack probabilities. The plots are obtained for CSU number vs. the number of outliers obtained using the proposed algorithm. The values of attack probabilities

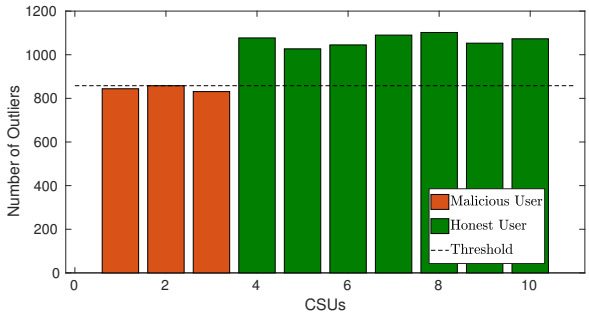

Fig. 4. The bar plot of outlier counts vs. CSU number considering $M = 10$, M_L=3, $\alpha_0 = 0.3$, $\alpha_1 = 0.3$ for $H_0 = 80\%$ and $H_1 = 20\%$ in the dataset.

are varied from 0.3 to 0.6. We can see that under all the attack probabilities, the outlier counts from the first three MUs are below the threshold, which indicates that the proposed algorithm correctly detects the MUs under all the considered scenarios. We can also note that the count values of MUs are very close to the threshold for small attack probabilities, and it may become difficult to detect them. This happens because, with low attack probabilities, the MUs attack very rarely, resulting in the pdf of MUs very similar to the honest one. However, the MUs do not attack with very low attack probability in general because the detection performance of CSS will be affected only marginally with small attack probabilities. The MUs would like to have higher degradation in the CSS performance; hence they do not attack with a very small attack probability. We can also note that the difference in the count between MUs and the HSUs increases with an increase in the attack probability, making it easier to declare them as MUs. In Fig. 6, we demonstrate the effectiveness of the proposed algorithm when there are different numbers of attackers in the network. The plots are obtained for $M_L = 0, 1, \ldots, 4$, which corresponds to 0% attackers to the 40% attackers. We can see that the proposed algorithm efficiently detects the variable number of MUs. Also, note that when there are no MUs, the count values are higher than the threshold, indicating that there are no MUs in the network. Note that the effectiveness of the proposed algorithm under different scenarios such as different SNR, different α_0 and α_1, and different attack strengths can be demonstrated using similar plots, but due to the page limitations, we have not added those plots. In Fig. 7 we show

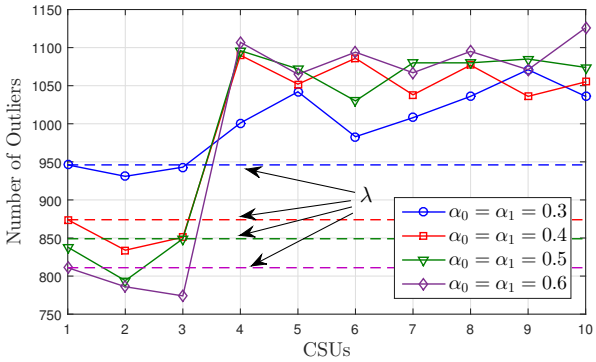

Fig. 5. Plot of outlier counts vs. CSUs for different values of α_0, α_1, $M = 10$, M_L=3 and $\Delta = 0.5$.

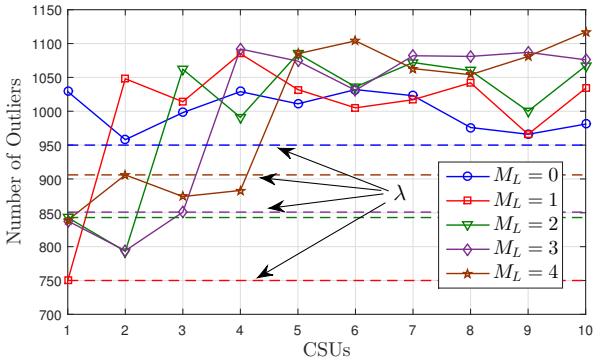

Fig. 6. Plot of outlier counts vs. CSUs for different values of M_L with $\alpha_0 = \alpha_1 = 0.3$, $M = 10$ and Δ=0.5.

the plot for detection probability vs. accuracy for the proposed algorithm. The x-axis represents the values of α_0 and α_1, and the plot is obtained at the SNR of $\gamma = 0\ dB$ with $\Delta = 0.5$, $M = 10$ and $M_L = 3$. We can see that, the accuracy of the proposed algorithm increases with the attack probability and achieves almost 100% accuracy at the attack probability of 0.5 or above that. We see that, if the attack probability is very low, the algorithm fails in identifying the attackers. However, if the MUs are attacking with very small probability, their effect on the performance of CSS is very marginal. Hence, even if the MUs are not detected in such scenario will not affect the performance of CR network. In Fig. 8, we compare the performance of proposed algorithm with the Tietjen-Moore (TM) test based algorithm proposed in [20]. Since TM test based algorithm decides about the MUs based on observations received in single sensing instance, it is sensible to compare the performance considering $\alpha_0 = \alpha_1 = 1$, i.e., always attack scenario. In addition, the TM test algorithm requires an information about the number of MUs in the network in order to detect them. To get the information about the number of MUs, a clustering based algorithm is used in [20]. In obtaining the plot for TM test, we have given the information about the number of MUs as an input to the TM test algorithm. No such information is required in the proposed algorithm and it can detect MUs blindly. We can see from the plot that the proposed algorithm detects the MUs with high accuracy even

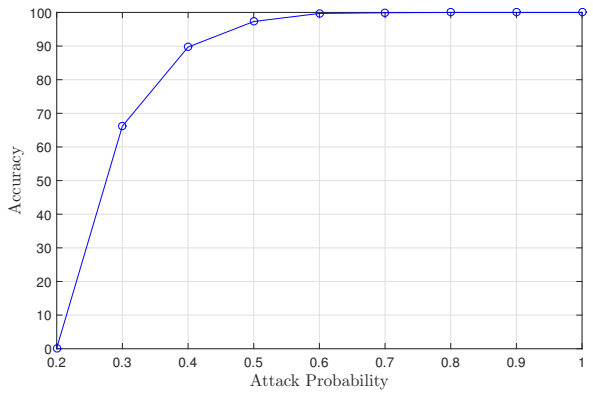

Fig. 7. Detection accuracy vs. Attack probability for proposed algorithm.

with very small values of Δ. To make the plot comparable, the values of Δ selected for TM test is very high compared to that considered for the proposed algorithm. We can observe that, even with high values of Δ, the TM test algorithm performs poorly compared to the proposed algorithm. For example, the proposed algorithm achieves 100% accuracy at $\Delta = 0.3$ whereas it requires $\Delta = 45$ to achieve 100% accuracy with TM test algorithm. One class SVM based algorithm is proposed in [24] for MUs detection. The plot for accuracy of one class SVM is not included here since its performance in detecting the MUs is very poor with the parameters selected. The one class SVM can detect MUs only if the MUs are using very high value of Δ. Furthermore, the algorithm also requires the percentage of MUs as an input which is also not available to the FC. In addition, the proposed algorithm has lower complexity than the single class SVM based algorithm proposed in [24] and the c-means clustering based algorithm proposed in [27]. The GMM based anomaly detection algorithm has complexity of order of $O(ILK)$, whereas the OTSU thresholding and the weighted sum based CSS algorithms' complexity is of order of $O(M)$, hence the overall complexity of the proposed algorithm is $O(2M + ILK)$, where, I is the number of iterations. The single class SVM based algorithm exhibits a significantly higher complexity of $O\left(2M^3 I_1 + L^2 + M^2 + MkI_2\right)$, where, I_1 and I_2 denote the number of iterations in their respective algorithms, k signifies the number of nearest neighbors. Additionally, the c-means clustering-based algorithm also exhibits a higher complexity of $O\left(\sqrt{J_c} L_{nz} + J_c L + MQL_c + 2MQK^2 I\right)$, where, L_{nz} denotes the number of non-zero elements in a matrix obtained from Micali's blossom algorithm, J_c is a parameter specific to the c-means clustering algorithm, L_c represents the number of intervals in which FC collects falsified data and the maximum quantization level is denoted as Q.

Finally, the effect of three MUs on the performance of CSS is demonstrated in Fig. 9. We see that the presence of MUs significantly degrades the detection performance. For example, for $Q_f = 0.1$, we get $Q_d = 0.8$ with CSS having all HUs whereas we get $Q_d = 0.5$ with CSS having three MUs. This indicates that the detection probability has been reduced by 36.7% due to the attack from MUs. The plots

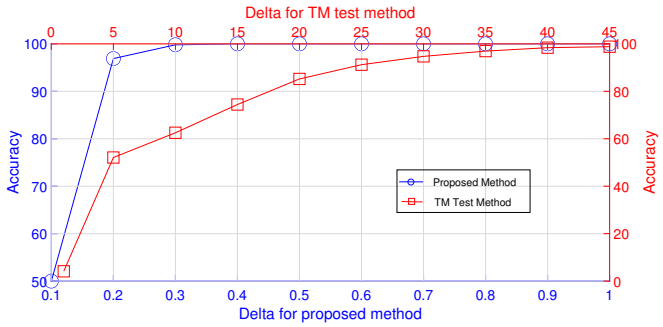

Fig. 8. Comparison of proposed algorithm with existing TM test based algorithm [20].

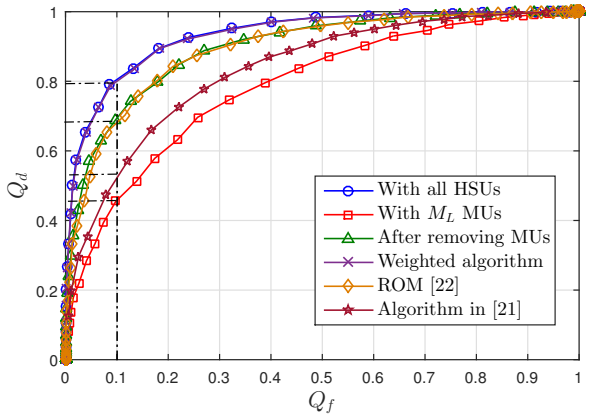

Fig. 9. Q_f vs. Q_d plot considering $\gamma = -10\ dB$, $M = 10$, $M_L = 3$, $\alpha_0 = \alpha_1 = 0.6$ comaparing the performance of proposed algorithm with existing algorithms..

are shown for $\alpha_0 = \alpha_1 = 0.6$. The degradation in the performance will be even higher if the MUs attack with higher attack probabilities. We also show that the detection and elimination of MUs from cooperating in decision making would result in improved performance. We can see from the plot in Fig. 9 that the detection probability after removing the MUs increases to 0.7, indicating an improvement in the performance. We can utilize the proposed algorithm in further improvement of the CSS performance by assigning weights to different CSUs based on the outlier counts and the weighted sum can be taken at the FC to compute the decision statistic as given in Algorithm 3. Since the outlier counts are less in case of MUs and higher for HSUs, the weights obtained using Eq. (13) are smaller for MUs and higher for HSUs. Hence, in computing the decision statistic for weighted sum based algorithm using Eq. (14) more importance is given to data received from HSUs and less to the data received from suspected MUs. It can be observed from the Fig. 9 that the Q_f vs. Q_d plot obtained using the weighted sum algorithm coincides with the plot when all HSUs are considered. This shows that using the weights obtained from the proposed algorithm in the weighted sum algorithm eliminates the effects of MUs completely. Fig. 9 also compares the performance of the proposed weighted algorithm with the algorithms proposed in [21] and a ROM based algorithm in [22]. We can see that the proposed algorithm significantly outperforms the algorithm proposed in [21]. For example, for $Q_f = 0.1$, Q_d with the proposed algorithm is 71.6% higher than Q_d achieved with an algorithm in [21]. The algorithm in [21] performs poorly because it only eliminates the effect of one of the malicious values, which is abnormally maximum from obtaining the decision statistic. The algorithm can perform well when only one MU exists, which attacks by sending a higher energy value than it has observed. However, in practice, there can be multiple MUs. In such cases, the performance of the algorithm degrades. The proposed algorithm also performs better than the ROM algorithm [22]. We can see that, for $P_f = 0.1$, the detection probability for the proposed algorithm is 15.4% higher than that with the ROM algorithm. The performance of the ROM algorithm is similar to the one where we eliminate MUs from cooperation. The ROM algorithm eliminates only one malicious value, abnormally high or small. The algorithm does not perform well when there are more MUs in the network.

VI. Conclusion

In this paper, we study the performance of centralized CSS under attack from MUs. A centralized independent probabilistic small scale attack is considered, and it is demonstrated that the performance of CSS degrades under attack from MUs. We then propose an algorithm based anomaly detection using the Gaussian mixture model for malicious user detection in CSS. A theoretical analysis is carried out, and the pdf of energies received at the FC from the MUs is derived. The derived pdf is then used to discuss the intuition behind the proposed algorithm. The proposed algorithm effectively detects the MUs in the cognitive radio network. Once the MUs are detected, one can eliminate them from cooperating in spectrum sensing leading to improved sensing performance. The outlier counts obtained using GMM based algorithm can be utilized to obtain weights for weighed sum based CSS algorithm. The weights obtained are such that, the algorithm gives less importance to the information received from the MUs, and higher importance to the data received from honest users leading to improved CSS performance.

Appendix A
Derivation of $f_{H_0}^{fa}(x)$ given in Eq. (18)

In Eq. (18), $f_{H_0}^{fa}(x)$ represents the pdf of energy received at the FC from the MUs when they carry out false alarm attack under hypothesis H_0. The MUs carry out attack by reporting $E^{'} = E_m + \Delta$ whenever the measured energy $E_m < \eta$. Since E_m follows a Gaussian distribution given by Eq. (15) under hypothesis H_0, the pdf $f_{H_0}^{fa}(x)$ follows same shape as he Gaussian pdf in Eq. (15) in the interval $-\infty < x \leq \eta$. Hence, to derive the pdf, we can truncate the Gaussian pdf in the given interval as

$$f_1^{trunc}(x) = \mathcal{N}\left(x; \mu_0, \sigma_0^2\right); -\infty < x \leq \eta \qquad (26)$$

Since Eq. (26) is truncated version of Gaussian distribution function, it is not a proper pdf and does not satisfy the property of the pdf, i.e., area under the curve is not 1. In order to

make the function a proper pdf, we can divide the Eq. (26) by its area. The area under the curve given in Eq. (26) is obtained by computing the cumulative distribution function (cdf) of Gaussian pdf which is given as

$$F_{H_0}(x) = \int_{-\infty}^{\eta} \mathcal{N}\left(x; \mu_0, \sigma_0^2\right) dx = \frac{1}{2}\left[1 + erf\left(\frac{\eta - \mu_0}{\sigma_0\sqrt{2}}\right)\right]. \tag{27}$$

Using this, the pdf of truncated Gaussian distribution can be obtained by dividing Eq. (26) by $F_{H_0}(x)$ as

$$f_{H_0}^{trunc1}(x) = \frac{\mathcal{N}\left(x; \mu_0, \sigma_0^2\right)}{\frac{1}{2}\left[1 + erf\left(\frac{\eta - \mu_0}{\sigma_0\sqrt{2}}\right)\right]};\ -\infty < x \leq \eta \tag{28}$$

The Eq. (28) represents the truncated Gaussian distribution where the Gaussian pdf is truncated in the interval $-\infty < x \leq \eta$. However, in case of false alarm attack, the MUs attack by adding Δ to the measured energy and reporting this increased energy to the FC. Hence, the pdf of under false alarm attack will exist in the interval $-\infty < x \leq \eta + \Delta$. Hence, we have to modify the truncated Gaussian pdf given in Eq. (28) to exist in interval $-\infty < x \leq \eta + \Delta$ which will give us the pdf $f_{H_0}^{fa}(x)$ in Eq. (18).

APPENDIX B
DERIVATION OF $f_{H_0}^{md}(x)$ GIVEN IN EQ. (19)

The $f_{H_0}^{md}(x)$ represents the pdf of energy received at the FC under H_0 when a MU carry out miss detection attack. The MU attack by reporting $E_m^{'} = E_m - \Delta$ whenever $E_m > \eta$. Because of Gaussian nature of E_m, the pdf $f_{H_0}^{md}(x)$ will have the same shape as that of Gaussian pdf in the interval $\eta \leq x < \infty$, and can be modeled using a truncated Gaussian pdf. To derive the intended pdf, we can truncate the Gaussian pdf as

$$f_2^{trunc}(x) = \mathcal{N}\left(x; \mu_0, \sigma_0^2\right); \eta \leq x < \infty \tag{29}$$

To make the function given in Eq. (29) a proper pdf, we can divide it by its area. The area under the truncated Gaussian function given in Eq. (29) can be obtained by computing the complementary cdf of Gaussian pdf as

$$\bar{F}_{H_0}(x) = \int_{\eta}^{\infty} \mathcal{N}\left(x; \mu_0, \sigma_0^2\right) dx = 1 - \frac{1}{2}\left[1 + erf\left(\frac{\eta - \mu_0}{\sigma_0\sqrt{2}}\right)\right]. \tag{30}$$

Using this, the pdf of truncated Gaussian distribution can be obtained by dividing the function in Eq. (29) by $\bar{F}_{H_0}(x)$ as

$$f_{H_0}^{trunc2}(x) = \frac{\mathcal{N}\left(x; \mu_0, \sigma_0^2\right)}{1 - \frac{1}{2}\left[1 + erf\left(\frac{\eta - \mu_0}{\sigma_0\sqrt{2}}\right)\right]};\ \eta \leq x < \infty \tag{31}$$

The pdf of truncated Gaussian distribution is given in Eq. (31) which exists in the interval $\eta \leq x < \infty$. When the MUs carry out miss detection attack, they subtract Δ from the measured energy and report this reduced energy to the FC. Hence, the pdf of under miss detection attack will exist in the interval $\eta - \Delta \leq x < \infty$. Hence, we have to modify the truncated Gaussian pdf given in Eq. (31) to exist in interval this interval which will give us the pdf $f_{H_0}^{md}(x)$ in Eq. (19).

REFERENCES

[1] F. C. Commission, "Spectrum policy task force report, FCC 02-155," 2002.

[2] K. Arshad, M. A. Imran, and K. Moessner, "Collaborative spectrum sensing optimisation algorithms for cognitive radio networks," *International Journal of Digital Multimedia Broadcasting*, vol. 2010, 2010.

[3] T. C. Clancy and N. Goergen, "Security in cognitive radio networks: Threats and mitigation," in *2008 3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2008)*. IEEE, 2008, pp. 1–8.

[4] O. Fatemieh, R. Chandra, and C. A. Gunter, "Secure collaborative sensing for crowd sourcing spectrum data in white space networks," in *2010 IEEE Symposium on New Frontiers in Dynamic Spectrum (DySPAN)*. IEEE, 2010, pp. 1–12.

[5] O. Fatemieh, A. Farhadi, R. Chandra, and C. A. Gunter, "Using classification to protect the integrity of spectrum measurements in white space networks." in *NDSS*, 2011.

[6] G. Ding, J. Wang, Q. Wu, L. Zhang, Y. Zou, Y.-D. Yao, and Y. Chen, "Robust spectrum sensing with crowd sensors," *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3129–3143, 2014.

[7] L. Zhang, G. Ding, Q. Wu, Y. Zou, Z. Han, and J. Wang, "Byzantine attack and defense in cognitive radio networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1342–1363, 2015.

[8] G. Nie, G. Ding, L. Zhang, and Q. Wu, "Byzantine defense in collaborative spectrum sensing via bayesian learning," *IEEE Access*, vol. 5, pp. 20 089–20 098, 2017.

[9] L. Zhang, G. Nie, G. Ding, Q. Wu, Z. Zhang, and Z. Han, "Byzantine attacker identification in collaborative spectrum sensing: A robust defense framework," *IEEE Transactions on Mobile Computing*, vol. 18, no. 9, pp. 1992–2004, 2019.

[10] Y. Fu and Z. He, "Entropy-based weighted decision combining for collaborative spectrum sensing over byzantine attack," *IEEE Wireless Communications Letters*, vol. 8, no. 6, pp. 1528–1532, 2019.

[11] A. Chouhan, K. Captain, A. Parmar, and J. Patel, "Defending cooperative spectrum sensing from byzantine attacks: An effective entropy-based weighted algorithm," *IEEE Wireless Communications Letters*, 2023.

[12] J. Wu, Y. Yu, T. Song, and J. Hu, "Sequential 0/1 for cooperative spectrum sensing in the presence of strategic byzantine attack," *IEEE Wireless Communications Letters*, vol. 8, no. 2, pp. 500–503, 2019.

[13] J. Wu, Z. Chen, H. Liang, Z. Chen, J. Zhang, J. Gan, and J. He, "Sequential single voting for cooperative spectrum sensing against byzantine attack," in *Proceedings of the 2023 10th International Conference on Wireless Communication and Sensor Networks*, 2023, pp. 1–8.

[14] Z. C. Jun Wu, "Secure and efficient cooperative spectrum sensing under byzantine attack and imperfect reporting channel," *Wireless Networks*, vol. 28, p. 367–380, 2022.

[15] Z. Chen, J. Wu, J. Gan, Z. Chen, J. Zhang, and J. He, "Robust and efficient cooperative spectrum sensing against probabilistic hard byzantine attack," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 4, p. e4414, 2022.

[16] Z. Xu, Z. Sun, and L. Guo, "Throughput maximization of collaborative spectrum sensing under ssdf attacks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 8378–8383, 2021.

[17] Y. Fu and Z. He, "An online learning approach for cooperator selection in css under ssdf attack," *IEEE Communications Letters*, vol. 26, no. 7, pp. 1479–1483, 2022.

[18] J. Gan, J. Wu, P. Li, Z. Chen, J. Zhang, Z. Chen, J. He, and S. Fan, "Joint spectrum sensing and resource allocation against byzantine attack in overlay cognitive radio networks," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 7, p. e4500, 2022.

[19] J. Wu, P. Li, Y. Chen, J. Tang, C. Wei, L. Xia, and T. Song, "Analysis of byzantine attack strategy for cooperative spectrum sensing," *IEEE Communications Letters*, vol. 24, no. 8, pp. 1631–1635, 2020.

[20] S. S. Kalamkar, P. K. Singh, and A. Banerjee, "Block outlier methods for malicious user detection in cooperative spectrum sensing," in *2014 IEEE 79th Vehicular Technology Conference (VTC Spring)*. IEEE, 2014, pp. 1–5.

[21] R. Gao, Z. Zhang, M. Zhang, J. Yang, and P. Qi, "A cooperative spectrum sensing scheme in malicious cognitive radio networks," in *2019 IEEE Globecom Workshops (GC Wkshps)*, 2019, pp. 1–5.

[22] L. Guo, W. Chen, Y. Cong, and X. Yan, "A robust mor-based secure fusion strategy against byzantine attack in cooperative spectrum sensing," in *International Congress on Communications, Networking, and Information Systems*. Springer, 2023, pp. 81–94.

[23] S. Rajasegarar, C. Leckie, and M. Palaniswami, "Pattern based anomalous user detection in cognitive radio networks," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, pp. 5605–5609.

[24] H. Wang and Y.-D. Yao, "Primary user boundary detection in cognitive radio networks: Estimated secondary user locations and impact of malicious secondary users," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4577–4588, 2018.

[25] H. Zhu, T. Song, J. Wu, X. Li, and J. Hu, "Cooperative spectrum sensing algorithm based on support vector machine against ssdf attack," in *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2018, pp. 1–6.

[26] Z. Chen, J. Wu, and J. Bao, "Semi-supervised learning-enabled two-stage framework for cooperative spectrum sensing against ssdf attack," in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2022, pp. 554–559.

[27] Z. Cheng, J. Zhang, T. Song, J. Hu, and X. Bao, "Detection strategy against restricted ssdf attack with potential interaction assistance," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 2, pp. 553–566, 2021.

[28] S.-Q. Liu and B.-J. Hu, "Analysis of sensing efficiency for cooperative spectrum sensing with malicious users in cognitive radio networks," *IEEE Communications Letters*, vol. 18, no. 9, pp. 1645–1648, 2014.

[29] D. Sun, T. Song, B. Gu, X. Li, J. Hu, and M. Liu, "Spectrum sensing and the utilization of spectrum opportunity tradeoff in cognitive radio network," *IEEE Communications Letters*, vol. 20, no. 12, pp. 2442–2445, 2016.

[30] A. Vosoughi, J. R. Cavallaro, and A. Marshall, "Trust-aware consensus-inspired distributed cooperative spectrum sensing for cognitive radio ad hoc networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 2, no. 1, pp. 24–37, 2016.

[31] M. Golvaei and M. Fakharzadeh, "A fast soft decision algorithm for cooperative spectrum sensing," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 68, no. 1, pp. 241–245, 2021.

[32] A. Kumar, S. Saha, and K. Tiwari, "A double threshold-based cooperative spectrum sensing with novel hard-soft combining over fading channels," *IEEE Wireless Communications Letters*, vol. 8, no. 4, pp. 1154–1158, 2019.

[33] G. Chandrasekaran and S. Kalyani, "Performance analysis of cooperative spectrum sensing over $\kappa-\mu$ shadowed fading," *IEEE Wireless Communications Letters*, vol. 4, no. 5, pp. 553–556, 2015.

[34] K. M. Captain and M. V. Joshi, "Energy detection based spectrum sensing over η-λ-μ fading channel," in *2016 8th International Conference on Communication Systems and Networks (COMSNETS)*. IEEE, 2016, pp. 1–6.

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

[36] Z. Luo, S. Zhao, Z. Lu, J. Xu, and Y. E. Sagduyu, "When attackers meet ai: Learning-empowered attacks in cooperative spectrum sensing," *IEEE Transactions on Mobile Computing*, vol. 21, no. 05, pp. 1892–1908, may 2022.

[37] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*. Springer, 2006, vol. 4, no. 4.

[38] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[39] K. Captain and M. Joshi, "SNR wall for generalized energy detector in the presence of noise uncertainty and fading," *Physical Communication*, vol. 32, pp. 172–184, 2019.

Ashok Parmar received the B.E. degree in electronics and communication engineering and the M.E. degree in electronics and telecommunication engineering from Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, India, in 2014 and 2016, respectively. He is currently pursuing the Ph.D. degree with Sardar Vallabhbhai National Institute of Technology, Surat, India. His research interests include cognitive radio and deep learning-based radio signal recognition.

Karan Shah is currently pursuing the BTECH electronics and communication degree from the Institute of Technology Nirma University and is set to graduate in 2024. His research interests include cognitive radio and machine learning.

Dr. Kamal M. Captain received his Ph.D. in the area of Spectrum Sensing for Cognitive Radio from Dhirubhai Ambani Institute of Information and Communication Technology (DAIICT), Gandhinagar, India. Before this, he completed his M.Tech in Communication Systems from Sardar Vallabhbhai National Institute of Technology (SVNIT) and B.E degree in Electronics and Communication Engineering from Veer Narmad South Gujarat University (VNSGU), Surat, Gujarat, India. He is currently serving as an assistant professor at SVNIT, Surat, Gujarat, India. Prior to joining SVNIT, he served as a senior engineer (signal processing) at eInfochips, Ahmedabad, India. He has been involved in active research in the areas of cognitive radio, wireless communication, signal processing, and machine learning and has several journals and international conference papers, including IEEE Transactions. He has also served as a reviewer for IEEE Transactions, letters, and top tier conferences.

Miguel López-Benítez (S'08, M'12, SM'17) received the BSc and MSc degrees (both with Distinction) from Miguel Hernández University, Elche, Spain in 2003 and 2006, respectively, and the PhD degree (*summa cum laude*) from the Technical University of Catalonia, Barcelona, Spain in 2011. From 2011 to 2013, he was a Research Fellow with the Centre for Communication Systems Research, University of Surrey, Guildford, UK. In 2013, he joined as a Lecturer (Assistant Professor) the Department of Electrical Engineering and Electronics, University of Liverpool, UK, where he has been a Senior Lecturer (Associate Professor) since 2018. His research interests include wireless communications and networking, including mobile communications and dynamic spectrum access in cognitive radio systems.

Jignesh R. Patel earned a B.E. degree in Electronics and Communications Engineering from North Gujarat University, Patan, India, and an M.E. degree in Electronics and Communications Engineering from Gujarat Technological University, Ahmedabad, India. He obtained his Ph.D. degree in Information and Communication Technology at DA-IICT, Gandhinagar, India. Currently, he is serving as an Assistant Professor at the Indian Institute of Information Technology, Vadodara (IIITV)- International Campus Diu (ICD). His research interests include Deep Learning, Computer Vision, and Remote Sensing. His current work focuses on Communication and Hyperspectral Imaging, utilizing Machine Learning.