IET Research Journals

ISSN 1751-8644
doi: 0000000000
www.ietdl.org

Improved Likelihood Ratio Statistic based Cooperative Spectrum Sensing for Cognitive Radio

Dhaval K. Patel[1], *Brijesh Soni*[1*], *Miguel López-Benítez*[2,3]

[1] *Information, Communication and Technology, School of Engineering and Applied Science, Ahmedabad University, Ahmedabad, India.*
[2] *Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, United Kingdom.*
[3] *ARIES Research Centre, Antonio de Nebrija University, Madrid, Spain.*
* *E-mail: brijesh.soni@ahduni.edu.in*

Abstract: Cooperative spectrum sensing (CSS) is a technique where multiple cognitive radio users cooperate among themselves to make the binary decisions about the presence of primary user. Single cognitive user often faces the hidden terminal problem. However, CSS tackles this problem by sending their local sensing based decisions to the fusion center. Major drawback of conventional energy detection is that, at low signal to noise ratio (SNR) regime its performance is very poor. In this work, likelihood ratio statistics is considered as a test-statistic due to its highest statistical power and an improved likelihood ratio statistic based CSS scheme is proposed by considering several past sensing events. Proposed scheme mitigates the poor detection at low SNR regime and misdetections arising due to sudden drops in signal energy. We compute the lower bound for likelihood ratio statistic based sensing. We also analyse the effect of non-Gaussian noise on the sensing performance. Furthermore, the generalized Byzantine attack is taken into account considering a security aspect. The proposed scheme is also shown to outperform Anderson Darling based malicious user detection in CSS at low SNR regime. The proposed scheme is verified and validated over empirical spectrum data. The improvement in performance is at cost of computational time at lower sample size, which in practice is very low and is justified by the significant performance improvements of the proposed scheme at low SNR regime.

1 Introduction

The wireless sector has witnessed a paradigm shift in recent past in terms of number of subscribers for different wireless services and technologies. Also, the fixed spectrum management policy which has an advantage of avoiding the interference among different wireless systems suffer two disadvantages. One being its inability to roll out new radio technologies and services; and another that although the spectrum has been allocated to wireless services, it is highly underutilized [1]. As per the recent survey [2], many spectrum occupancy campaigns across the globe have proved that the utilization of spectrum is very low.

To mitigate the problem of increasing spectrum demand, the concept of Dynamic Spectrum Access (DSA) or Opportunistic Spectrum Access (OSA) or Cognitive Radio (CR) has emerged as a promising solution [3]. CR has two types of users, licensed users generally referred to as primary users (PU), and unlicensed users referred to as secondary users (SU). SU scavenges for the opportunity (referred to as *spectrum sensing*) to find some vacant space (referred as *spectrum holes or white spaces* [1]) across frequency, time or space and utilize that without interference as long as PU is inactive (i.e, spectrum hole is available). However, as soon as PU becomes active, the SU has to immediately vacate the allocated spectrum in a non-interfering manner.

1.1 Background and Motivation

Spectrum sensing is a primary function of CR and has been an important aspect of research. It can be broadly classified into parametric and non-parametric approaches. Parametric approach is the one in which the secondary CR user requires the a-priori knowledge of the PU's signal. As per the survey in [4], techniques like using information from geolocation and database using beacon signals, cyclostationary feature detection, matched filter detection and so on fall under the category of parametric sensing. On the other hand, more robust non-parametric approaches include energy detection (ED) [5], improved energy detection (IED)[6], Anderson Darling(AD) test [7], and likelihood ratio statistics (LRS-$\mathbb{G}^2$) based spectrum sensing [8], wherein no prior knowledge of the PU signal is required.

Spectrum sensing, although appearing simple, in reality is a difficult task and various challenges have been addressed in the literature. In [9] it is stated that sensing is difficult as different PUs would be employing different modulation schemes, transmission powers, different data rates, etc. Another problem encountered in implementing spectrum sensing is when singly operating SU terminal is shadowed in severe multipath fading or due to buildings with high penetration loss while the PU operates in the vicinity, usually referred to as hidden terminal problem [9]. The solution to the above mentioned problems is addressed in [10]-[11] by a technique called *Cooperative Spectrum Sensing* (CSS). In this technique different SUs in close geographical location report their sensing based decisions to the common node, referred to as fusion center (FC).

ED based CSS was discussed in [9] and [12]. Furthermore, a better approach called Improved Energy Detector based CSS was proposed in [13] wherein ED is modified by replacing the squaring operation of the received signal amplitude with an arbitrary power. However, a major drawback for ED based spectrum sensing is that its performance is very poor in low SNR regime [14]. Another limitation of ED based localized sensing at the SUs is that at low SNR it might send an erroneous decision to the FC and this may have impact specially when all the SUs send wrong decisions and the FC uses *AND fusion rule*.

Furthermore, in addition to the above mentioned limitation, users in CSS may act maliciously and send a erroneous decision at FC. There are few works in the literature which have addressed the malicious user detection problem in CSS. For instance, a robust malicious user detection scheme was studied in [15]. Authors in [16] have focused on the security aspects of CSS using AD based goodness of fit test. Trust based mechanism for malicious user detected in CSS was proposed in [17]. Similar recent works can be found in [18],

Table 1 Literature summary of related works

Reference	CSS	ED based sensing	IED and improved energy detector	AD based sensing	LRS based sensing	ILRS based sensing	Malicious user detection	Analysis under non-Gaussian noise	Use of empirical data for verification
[5]	×	✓	×	×	×	×	×	×	×
[6]	×	✓	✓	×	×	×	×	×	✓
[7]	×	×	×	✓	×	×	×	×	×
[8]	×	×	×	×	✓	×	×	×	×
[9]	✓	✓	×	×	×	×	×	×	×
[12]	✓	✓	×	×	×	×	×	×	×
[13]	✓	×	✓	×	×	×	×	×	×
[16]-[19]	✓	✓	×	✓	×	×	✓	×	×
[21]	×	×	×	×	✓	×	×	×	×
[26]	×	×	×	×	✓	×	×	✓	×
Our contribution	✓	✓	×	✓	✓	✓	✓	✓	✓

[19] and references therein. However, the previous studies have considered the attacks restricted only to specific type and not in general. An indepth study of generalized Byzantine attacks under dynamic and static CSS scenario is carried out in [20] and [21]. Moreover, all the above mentioned works have considered the Gaussian noise at the receiver. On the contrary, we have considered the effect of non-Gaussian noise in CSS.

Numerous work has been carried out in literature in which spectrum sensing has been treated as a hypothesis testing problem or goodness of fit test problem. In [7], AD sensing based on goodness of fit was proposed which was corrected in [22]. In [8], LRS-$\mathbb{G}^2$ has been proposed which has outperformed prevailing sensing schemes like ED based, AD sensing, and other order statistic based sensing. In [23] the likelihood ratio based goodness of fit using χ^2 distribution was proposed which outperformed ED at low SNR. In [24], sensing method based on statistical test which outperforms ED based sensing at low SNR has been proposed. Also in [25], Kernalized generalized likelihood ratio test based statistical method for spectrum sensing was proposed, which used nonlinear kernel to map input data onto high dimensional feature space. However, it needs historical database of PU and thus falls under parametric sensing. All the above methods are used for non CSS scenario. Table 1 provides a brief summary of related work in literature in the context of CSS.

To the best of authors' knowledge, there is very limited work carried out in literature like [26] in which statistical methods based sensing have been applied at the SUs to overcome the sensing problem at low SNR regime in CSS. In this work, we apply the statistical method based spectrum sensing as in [8] at all the SU nodes. Further, we propose an improved version of likelihood ratio statistic test (ILRS) in which abrupt or instantaneous changes in the signal leading to misdetections are taken into consideration. Moreover, the malicious user detection is taken into consideration and implemented using proposed scheme. The proposed scheme is verified with hard combining rule (AND and OR) for both (LRS and ILRS) schemes using empirical data captured through a test-bed setup. Furthermore, we consider the realistic assumption of non-Gaussian noise at the receivers, which has a more prominent effect in CSS.

1.2 Contributions

Our main contributions in this article can be summarized as below:

• Firstly, the effect of non-Gaussian noise arising due to various sources at the receiver in case of cooperative spectrum sensing by considering middleton class A noise is studied. Also, the effect of ratio of Gaussian to non-Gaussian component on proposed sensing scheme, LRS and CED at SUs is demonstrated. The proposed scheme performs better than LRS and CED even under non-Gaussian noise.
• Secondly, we derive the lower bound for probability of detection P_d for LRS based sensing in accordance with the comments from [22] and verify the result using different values of sample size.
• Thirdly, we apply the statistical method based spectrum sensing (LRS) at all the SU nodes to alleviate the degrading effect of ED based sensing at low SNR regime. We then propose an ILRS based sensing at all the SU nodes in which we apply an additional check in comparing the few previous values of test statistic with the threshold so as to avoid the misdetections arising due to abrupt signal changes. Further, we verify the proposed scheme on empirical data of various radio technologies acquired using an USRP test-bed setup. Experimental results prove that ILRS hard decision based sensing at all the nodes outperforms the LRS and ED based CSS.
• Lastly, security aspect is taken into consideration whereby users in cooperation act maliciously and report the falsified hard decisions at FC. An ILRS based malicious user detection is proposed and is shown to outperform the AD based malicious user detection in CSS [16], which considers that malicious user always report yes at the FC. In addition, we also consider the scenario wherein the SU may not always report yes at FC, known as Generalized Byzantine attack which is more realistic.

The rest of the paper is organised as follows. Section 2 describes the system model under Gaussian noise and non-Gaussian noise. Section 3 describes the CSS which further provides the insights of ED based localized sensing in section 3.1, LRS based sensing in section 3.2 and decision based sensing at the FC in section 3.3. Lower bound for detection probability for LRS is computed in Section 4. The proposed ILRS based cooperative sensing scheme is described in Section 5. Proposed scheme based malicious user detection in CSS is studied in Section 6. Section 7 provides the comprehensive study of empirical test bed setup for spectrum data acquisition. Simulation results are discussed in Section 8. Section 9 focuses on the comparative study of computational time and complexity of various spectrum sensing schemes. Section 10 draws conclusion from this work.

2 System Model

2.1 Cooperative spectrum sensing under Gaussian noise

We consider the total number of CR users to be N under cooperation, one FC and one PU transmitter (T_x) as shown in Fig. 1. The channel is sensed by the N CR users to determine whether it is idle or busy. Thus depending on the channel occupancy state, we form two hypotheses. One, the channel being idle i.e, PU is inactive (absent) denoted as hypothesis H_0 and other that the channel is busy i.e, PU is active (present) denoted as hypothesis H_1.

Mathematically, it can be represented as

$$y_i(t) = \begin{cases} w_i(t), & H_0 \\ h_i m(t) + w_i(t), & H_1 \end{cases} \tag{1}$$

where $i = 1, 2, ..., N$, $y_i(t)$ is the received signal at i^{th} SU, $m(t)$ is the transmitted signal from PU, h_i represents the complex channel

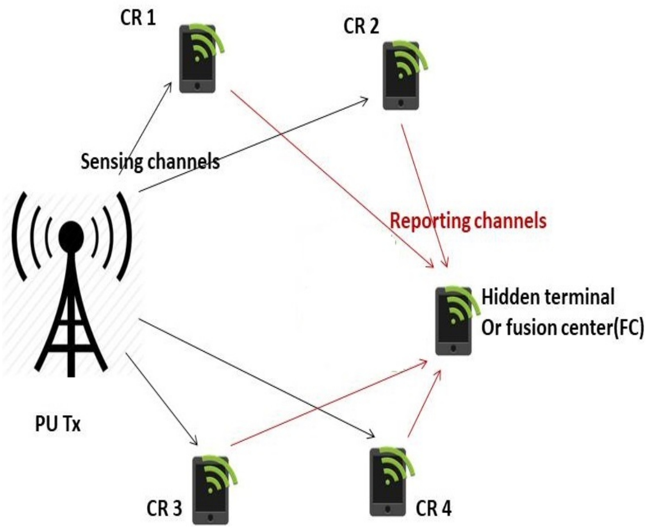


Fig. 1: Considered system model

coefficient of sensing channel between PU and i^{th} SU and $w_i(t)$ denotes white and circularly symmetric complex Gaussian (CSCG) noise with zero mean and variance σ_w^2, represented as $\mathcal{CN}(0, \sigma_w^2)$. PU and SU are assumed to be equipped with single antenna.

However there are few assumptions considered in this article as in [27]:

- Channel is assumed to be time invariant during during each sensing event, but can change across different sensing events.
- Each SU remains silent during each sensing event.
- Each SU remains more or less static at its position. This ensures that the decisions from SU does not change due to mobility of SU.
- Decision is made by the FC only after the decisions arrive from all the SUs, regardless of the fusion rule.
- We assume that the SUs are in the sensing range of FC and PU.

2.2 *Cooperative spectrum sensing under non-Gaussian noise*

There are instances where the non-Gaussian behaviour of noise needs to be considered. As in [28], broadly there are three sources of non-Gaussian noise. Firstly, this may be due to natural phenomena such as lightning and radiation from the sun, cosmic, and extraterrestrial solar, etc. This type of interference is commonly called atmospheric noise, due to inherent source (inherent noise within electronic equipment) or may be due to man-made source [29]. Although non-Gaussian behaviour can be modelled as Class-A, B, C and symmetric alpha stable $(s\alpha s)$ distributions, for the sake of brevity we restrict ourselves to class-A model. As there are as many receivers in cooperative sensing as SUs, it is important to consider the effect of non-Gaussian noise.

We consider the non-Gaussian behaviour of noise in cooperative sensing and reformulate the hypothesis testing problem as,

$$\begin{aligned} H_0 &: F_Y(y) = F_0(y), \\ H_1 &: F_Y(y) \neq F_0(y), \end{aligned} \tag{2}$$

where $F_0(y)$ will be the cumulative distribution function (CDF) of Middleton class-A noise and $F_Y(y)$ is the empirical distribution of the received signal, represented as

$$F_Y(y) = \frac{|\{s : y_s \leq y, 1 \leq s \leq n\}|}{n}, \tag{3}$$

where $s = 1, 2, .., n$, $|.|$ denotes the cardinality and n is the sample size. $F_0(y)$ is the CDF of independent and identically distributed $(i.i.d)$ noise samples.

The Class-A probability density function (PDF) can be expressed as [30]:

$$f_X(x) = \sum_{d=0}^{\infty} e^{-d} \frac{A^d}{d!} \left[\frac{1}{\sqrt{2\pi\sigma_d^2}} \right] e^{\frac{-x^2}{2\sigma_d^2}} \tag{4}$$

where $\sigma_d^2 = \frac{d/A + \mathbf{\Gamma}}{1 + \mathbf{\Gamma}}$. The class-A model can be fully described by parameters A and $\mathbf{\Gamma}$. A is the overlap or impulse index, which is the product of the average number of emission events impinging on the receiver per second. Mean duration of a typical interfering source emission is $A \in [10^{-2}, 1]$ time units. The smaller the value of A, the more “structured" (in time) the interference is. Conversely, the larger the value of A, the more Gaussian and less structured the noise is. When A = 1, the noise is Gaussian [30]. Also $\mathbf{\Gamma}$ is called the Gaussian factor and it is the ratio of intensity of the independent Gaussian component to the intensity of the impulsive non-Gaussian component. $\mathbf{\Gamma}$ lies in the range of $[10^{-6}, 1]$.

3 Cooperative spectrum sensing

3.1 *Energy detection based localized sensing*

Energy detection is an non-parametric approach for spectrum sensing. In this method we compute the energy of the received signal from PU as,

$$Y = \frac{2}{N_0} \int_0^T |y_i(t)|^2 dt \tag{5}$$

where N_0 is the one sided noise power spectral density. When hypothesis H_0 is accepted (i.e, when PU is not present) we have $y_i(t) = w_i(t)$. Y follows central chi-square (χ^2) distribution with 2u degrees of freedom. On the other hand, when hypothesis H_1 is accepted (i.e, when PU is present) we have $y_i(t) = h_i m(t) + w_i(t)$, under such case Y follows non central chi-square (χ^2) distribution with 2u degrees of freedom [31]. The parameter u denotes the time-bandwidth product.

One important parameter in deciding the above formulated hypothesis is decision threshold (λ_i). For the sake of brevity we directly state the closed form expression for probability of detection $P_{d,i}$, probability of false alarm $P_{f,i}$ and probability of missed detection $P_{m,i}$ at i^{th} SU respectively from [31].

$$P_{d,i} = Q_m(\sqrt{2\gamma_i}, \sqrt{\lambda_i}) = 1 - P_{m,i} \tag{6}$$

$$P_{f,i} = \frac{\Gamma(u, \frac{\lambda_i}{2})}{\Gamma(u)} \tag{7}$$

where Q_m denotes Marcum Q function, $\Gamma(.,.)$ is the incomplete gamma function and $\Gamma(.)$ is the Gamma function. Energy threshold is represented by λ_i and γ_i represents instantaneous SNR at i^{th} SU.

3.2 *Likelihood Ratio statistic (LRS) based localized sensing*

In this technique, we consider the sensing as a goodness of fit test problem. When sufficiently large samples (n) of the received signal is taken and if PU is absent, Y can be regarded as a samples drawn from noise distribution $F_0(y)$ and vice-versa as considered in (2). However, as H_0 corresponds to absence of PU, i.e. noise only samples, the distribution $F_0(y)$ will have Gaussian distribution [8].

This scheme is applied at N SUs and the localized binary decisions about the channel occupancy is forwarded to the FC. Expression in (8) is used to compute the measure between $F_Y(y)$ and $F_0(y)$. The detailed summary is explained in detail in Algorithm-1.

3.3 *Decision at the Fusion Center*

Once the locally sensed decision from all the SUs about the channel activity (either idle or busy) is made, it is reported at the FC through a reporting channel. At FC, all the decision are fused together (step 20

Algorithm 1 Cooperative Spectrum Sensing using LRS

1: Generate randomly located N SUs and define sample size(n), values of P_f, read the signal and compute SNR
2: **for** $kk \leftarrow 1$ to length($\mathrm{P_f}$) **do**
3: $\mathrm{P}_f\_AND(kk) = P_f(kk)^N$
4: $\mathrm{P}_f\_OR(kk) = 1 - (1 - P_f(kk)^N)$
5: prob_majority=0
6: **for** c = floor($N/2$)+1:N **do**
7: prob_majority=prob_majority+ nchoosek(N, c)*(Pf(kk^c)*((1-Pf(kk))$^N - c$)
8: **end for**
9: P_f_Majority(kk) = prob_majority
10: Declare LRS_detection_AND, LRS_detection_Majority and LRS_detection_OR of Montecarlo length times
11: **for** $b \leftarrow 1$ to iter i.e, length (montecarlo times) **do**
12: Define a temporary variable LRS_detection_temp
13: **for** $K \leftarrow 1$ to last frame of signal **do**
14: **for** $N \leftarrow 1$ to no. of SU's **do**
15: Model the channel, noise and compute the received signal
16: Compute the Z_c statistic as below

$$Z_c = \sum_{s=1}^{n}\left[\log\left\{\frac{F_0(y[s])^{-1} - 1}{\left(n-\frac{1}{2}\right)/\left(s-\frac{3}{4}\right)-1}\right\}\right]^2 \tag{8}$$

where n is the sample size and $F_0(y[s])^{-1}$ is the inverse
17: of the CDF of the received signal (as computed above)
18: **if** $Z\_C >= c\_value\_C(P_f == P_f(kk))$ **then**
19: LRS_detection_temp(N) $\leftarrow 1$
20: **else**
21: LRS_detection_temp(N) $\leftarrow 0$
22: **end if**
23: Repeat for all SU nodes
24: **end for**
25: **if** Product(LRS_detection_temp) == 1 **then**
26: LRS_detection_AND $\leftarrow 1$
27: **end if**
28: **if** Sum(LRS_detection_temp) > floor($N/2$) **then**
29: LRS_detection_Majority $\leftarrow 1$
30: **end if**
31: **if** 1-Product(LRS_detection_temp) $\neq 0$ **then**
32: LRS_detection_OR $\leftarrow 1$
33: **end if**
34: **end for** ▷ Increment the frame
35: **end for** ▷ Repeat for the Monte Carlo times
36: LRS_Pd_AND(kk) = sum(LRS_detection_AND) / iter
37: LRS_Pd_Majority(kk) = sum(LRS_detection_Majority) / iter
38: LRS_Pd_OR(kk) = sum(LRS_detection_OR) / iter
39: **end for**

through 24 in Algorithm-1 and step 27 through 31 in Algorithm-2) as per the following *"l out of N"* logic rule [9].

$$Y_{FC} = \begin{cases} \sum_{i=1}^{N} Decision_i < l; & H_0 \\ \sum_{i=1}^{N} Decision_i \geq l; & H_1 \end{cases} \tag{9}$$

where l represents the decision threshold at FC. When decision is made at l=1, it implies that even if any one SU has sensed the presence of PU, the FC will conclude that the PU is present. This is also called OR fusion rule. On the other hand, the FC will conclude the presence of PU only when all the SUs have sensed the presence of PU. This is called AND fusion rule. Similar to the $P_{f,i}$ and $P_{m,i}$ at SUs, false alarm and missed detection probability for cooperative sensing is given by (10)-(11). As we have a binary decision about the presence of PU, it is reasonable to assume that both probabilities follow binomial distribution.

$$Q_f = Prob[H_1|H_0] = \sum_{l=l}^{N}\binom{N}{l}P_f^l(1-P_f)^{N-l}, \tag{10}$$

$$Q_m = Prob[H_0|H_1] = 1 - \sum_{l=l}^{N}\binom{N}{l}P_d^l(1-P_d)^{N-l}, \tag{11}$$

As both Q_f and Q_m infer to the error probabilities, the sum of Q_f and Q_m, (Q_f + Q_m) implies the total error rate [13].

4 Computation of Lower Bound on Detection Probability for LRS

The signal received at the SU can be written as,

$$y_i(t) = h_i\sqrt{\gamma}m(t) + w_i(t), \tag{12}$$

where $y_i(t)$ is the received signal at the i^{th} CR, γ is the received SNR, $m(t)$ is the transmitted signal and $w_i(t)$ is the Gaussian noise with zero mean and unit variance. Considering AWGN channel, we have $h_i = 1$ in (12). Also, one SU is considered for simplicity.

In absence of PU, $y_i(t)$ is equal to $w_i(t)$ i.e, hypothesis H_0. Thus the distribution function $F_0(y)$ under such hypothesis can be written as,

$$F_0(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-x^2/2}dx$$

However when the PU signal is present, the received observations $y_i(t)$ will follow hypothesis H_1. Also the transmitted signal $m(t)$ from PU is the empirical signal captured using test bed setup, in accordance to [22] but as opposed to $m(t) = 1$ in [7]. The mean of such distribution will be $m\sqrt{\gamma}$ and will have a unit variance. Thus, the distribution $F_1(y)$, under hypothesis H_1 can be written as

$$F_1(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-(x-m\sqrt{\gamma})^2/2}dx$$

In [32], authors have proposed a new hypothesis test based on power divergence statistics for null-hypothesis testing as,

$$2nI^{\varphi} = \frac{2n}{\varphi(\varphi+1)}\left\{F_n(t)\left[\frac{F_n(t)}{F_0(t)}\right]^{\varphi} + [1-F_n(t)]\left[\frac{1-F_n(t)}{1-F_0(t)}\right]^{\varphi} - 1\right\} \tag{13}$$

where, φ represents a parameter for selection of goodness of fit test, n and $F_n(t)$ denote sample size of received observations and empirical CDF respectively.

By selecting $\varphi = 0$, (13) represents Likelihood Ratio Statistics (LRS-$\mathbb{G}^2$) as,

$$\mathbb{G}^2 = 2n\left\{F_n(t)\log\frac{F_n(t)}{F_0(t)} + [1-F_n(t)]\log\frac{1-F_n(t)}{1-F_0(t)}\right\} \tag{14}$$

In [33], authors have proposed a parametrization approach to construct a generalized omnibus GoF test for a specified distribution (F_0) under hypothesis H_0 as normal distribution using different weight functions. They have proposed general test statistics called as Z statistics using,

$$Z = \int_{-\infty}^{\infty} z_t \, \phi(t) \, dt, \tag{15}$$

where z_t indicates a type of GoF test statistics and $\phi(t)$ denotes weighting function. For likelihood ratio statistic as per (14) and (15) we have

$$Z = \int_{-\infty}^{\infty} \mathbb{G}^2 \, \phi(t) \, dt, \tag{16}$$

$$\begin{aligned} Z = \int_{-\infty}^{\infty} 2n \Bigg[\bigg(F_n(t) \log \frac{F_n(t)}{F_0(t)} \bigg) \\ + [1 - F_n(t)] \log \frac{1 - F_n(t)}{1 - F_0(t)} \Bigg] \times \phi(t) dt, \end{aligned} \tag{17}$$

On further simplification, we have

$$\begin{aligned} Z = \int_{-\infty}^{\infty} 2n \Bigg[\bigg(F_n(t) (\log F_n(t) - \log F_0(t)) \bigg) \\ + [1 - F_n(t)] (\log(1 - F_n(t)) \\ - \log(1 - F_0(t)) \Bigg] \times \phi(t) dt. \end{aligned} \tag{18}$$

We approximate the $\log(x)$ terms and $\log(1-x)$ terms in the above equation as per the Taylor series expansion [34]. Considering only the first order term, as $F_n(t)$ and $F_0(t)$ are the CDF values, it is fair to ignore the higher order squared terms. Thus approximating $\log(x) \approx (x-1)$ and $\log(1-x) \approx -x$

$$\begin{aligned} Z = \int_{-\infty}^{\infty} 2n \Bigg[F_n(t) \bigg((F_n(t) - 1) - (F_0(t) - 1) \bigg) \\ - \bigg(1 - F_n(t) \bigg) \bigg(F_n(t) - F_0(t) \bigg) \Bigg] \times \phi(t) dt \end{aligned} \tag{19}$$

On solving further, we obtain

$$Z = \int_{-\infty}^{\infty} 2n \Bigg[\bigg(F_n(t) - F_0(t) \bigg) \bigg(2F_n(t) - 1 \bigg) \Bigg] \times \phi(t) dt, \tag{20}$$

The detection probability P_d of LRS is given by,

$$P_{d,Z} = P(Z > \tau | H_1) = 1 - F_{Z,H1}(\tau), \tag{21}$$

where τ is the threshold. We upper bound the $F_{Z,H1}(\tau)$. Thus we have,

$$\begin{aligned} Z \geq \int_{-\infty}^{\infty} 2n \Bigg[\bigg(F_1(t) - F_0(t) \bigg) \bigg(2F_n(t) - 1 \bigg) \Bigg] \phi(t) dt \\ - \int_{-\infty}^{\infty} 2n \Bigg[\bigg(F_n(t) - F_0(t) \bigg) \bigg(2F_n(t) - 1 \bigg) \Bigg] \phi(t) dt, \end{aligned} \tag{22}$$

where the inequality is due to inequality of triangle. Let us assume,

$$C = \int_{-\infty}^{\infty} \Bigg[\bigg(F_1(t) - F_0(t) \bigg) \bigg(2F_n(t) - 1 \bigg) \Bigg] \times \phi(t) dt$$

and

$$B_n = \int_{-\infty}^{\infty} 2n \left[\left(F_n(t) - F_0(t) \right) \left(2F_n(t) - 1 \right) \right] \times \phi(t) dt$$

Thus we have, $Z \geq C(2n) - B_n$. Now, as per the definition of CDF and using Markov's inequality,

$$\begin{aligned} F_{Z,H1}(\tau) &= P\{Z \leq \tau | H_1\} \\ &= P\left\{ e^{-\kappa Z} \geq e^{-\kappa \tau} | H_1 \right\} \\ &\leq \frac{\boldsymbol{E}\left[e^{-\kappa Z} \right]}{e^{-\kappa \tau}} \\ &\leq \frac{e^{-\kappa C(2n)} \boldsymbol{E}\left[e^{-\kappa B_n} \right]}{e^{-\kappa \tau}} \end{aligned} \tag{23}$$

For simplicity, considering the multiplier constant, $\kappa = 1$. Thus we have,

$$F_{Z,H1}(\tau) \leq \frac{e^{-C(2n)} \boldsymbol{E}\left[e^{B_n} \right]}{e^{-\tau}} \tag{24}$$

and

$$P_d \geq 1 - \frac{e^{-C(2n)} \boldsymbol{E}\left[e^{B_n} \right]}{e^{-\tau}} \tag{25}$$

The lower bound value for probability of detection is shown in (25). As per the analysis in [35], the distribution $B_n^2 | H_1$ will converge as $n \to \infty$ and C is a constant. A similar analysis was carried out in [7] in which the PU signal considered was $m(t) = 1$. However as per [22] it is not very appropriate to consider the primary signal $m(t) = 1$. Adhering to the comments in [22], we have considered the PU signals $m(t)$ as empirical data of different radio technologies, the details of which are given in section 7.

5 Proposed Improved likelihood ratio statistic (ILRS) based Cooperative sensing scheme

The results obtained using LRS scheme suggest that the detection performance can be improved if the misdetections caused by the sudden changes in the signal energy drops are avoided [6]. This motivates us to apply an improved version of likelihood ratio statistics in which apart from the present event, the average of past L sensing events is taken into consideration for calculating the Z_c statistic (Algorithm-2; line 16).

Consider the scenario as shown in Fig. 2 where the PU signal is present throughout, however at instant **'b'** there is a sudden drop in the energy of signal which may cause misdetection. If average of the Z_c statistic of past L events is taken into consideration, then Z_c average will be greater than decision threshold resulting in declaring the channel to be busy (i.e, hypothesis H_1). The summary of the Algorithm is explained in detail in Algorithm-2. An important parameter that needs to be addressed is the number of past sensing events L that is to be considered for computing Z_c average. This is discussed in detail in the results section.

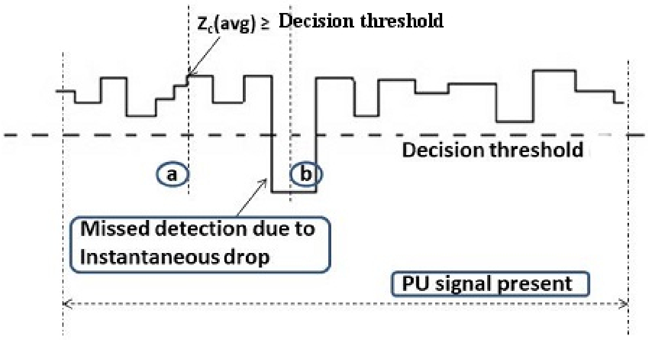


Fig. 2: Motivation for the proposed scheme

Algorithm 2 Proposed scheme for Cooperative spectrum sensing

Generate randomly located N SUs and define sample size(n), values of P_f, number of past events (L), read the signal and compute SNR
for $kk \leftarrow 1$ to length($\mathrm{P_f}$) **do**
  $\mathrm{P}_f\_AND(kk) = P_f(kk)^N$
  $\mathrm{P}_f\_OR(kk) = 1 - (1 - P_f(kk)^N)$
  prob_majority=0
  for c = floor($N/2$)+1:N **do**
    prob_majority=prob_majority+ nchoosek(N, c)*(Pf(kk^c)*((1-Pf(kk))$^N - c$)
  end for
  P_f_Majority(kk) = prob_majority
  Declare ILRS_detection_AND, OR and Majority.
  for $b \leftarrow 1$ to iter i.e, length (montecarlo times) **do**
    Define a temporary variable ILRS_detection_temp
    for $K \leftarrow 1$ to last frame of signal **do**
      for $N \leftarrow 1$ to no. of SU's **do**
        Model the channel, noise and compute the rx. signal
        Compute the Z_c statistic as below

$$Z_c = \sum_{s=1}^{n} \left[\log \left\{ \frac{F_0(y[s])^{-1} - 1}{\left(n - \frac{1}{2}\right) / \left(s - \frac{3}{4}\right) - 1} \right\} \right]^2$$

        if No. of past events $> L$ **then**
          test_mean $\leftarrow$ (Z_C(K-1)+...+Z_C(K-L)) /L
        else
          test_mean $\leftarrow$ Z_C(K)
        end if
        if $Z\_C >= c\_value\_C(P_f == P_f(kk))$ **then**
          ILRS_detection_temp(N) $\leftarrow$ 1
        else if
          $Test\_mean >= c\_value\_C(P_f == P_f(kk))$
          ILRS_detection_temp(N) $\leftarrow$ 1
        else
          ILRS_detection_temp(N) $\leftarrow$ 0
        end if
        Repeat for all SU nodes
      end for
      if Product(ILRS_detection_temp) == 1 **then**
        ILRS_detection_AND $\leftarrow$ 1
      end if
      if Sum(ILRS_detection_temp) > floor($N/2$) **then**
        ILRS_detection_Majority $\leftarrow$ 1
      end if
      if 1-Product(ILRS_detection_temp) $\neq$ 0 **then**
        ILRS_detection_OR $\leftarrow$ 1
      end if
    end for ▷ Increment the frame
  end for ▷ Repeat for the Monte Carlo times
  ILRS_Pd_AND(kk) = sum(ILRS_detection_AND) / iter
  ILRS_Pd_Majority(kk) = sum(ILRS_detection_Majority) / iter
  ILRS_Pd_OR(kk) = sum(ILRS_detection_OR) / iter
end for

6 Proposed Scheme based Malicious user detection in Cooperative Spectrum Sensing

Security is an important concern in cooperative sensing. In this section we consider the situation wherein SUs in cooperation act maliciously in a way that they try to falsify the sensing result by indicating the presence of PU (i.e., PU in busy state), even when they are actually not present in order to monopolize the spectrum and prevent other SUs from accessing the spectrum opportunities [16]. To address the above scenario, AD based GoF test for malicious user detection scheme was used in [16], which tests whether empirical distribution from SUs fit the expected distribution for malicious user.

The binary hypothesis for local sensing at SUs are formulated similar to (1). However, for simulation and to show a fair comparison of AD based malicious user detection in cooperative sensing [16] and proposed scheme (LRS and ILRS based malicious user detection), we consider that the transmitted signal $m(t)$ from PU is quadrature phase shift keying (QPSK) modulated signal with unit transmit power and the channel is characterized as Rayleigh faded, i.e, $h_i(t) = \mathcal{CN}(0, \sigma_v^2)$.

In malicious user detection, as mentioned earlier, the empirical distribution from SUs is compared with the expected distribution for malicious user. If both distributions match well, we can conclude that the SU is malicious and exclude its test statistic for consideration at the fusion centre [16]. The empirical distribution of the i^{th} secondary user after the k^{th} sensing interval is finished is given as [36],

$$F_k^{(i)}(y) = \frac{1}{k} \sum_{p=1}^{k} \mathbf{1}(y_i(p) \le y) \tag{26}$$

where $\mathbf{1}(\alpha)$ is the indicator function which equals 1 if α is true and zero otherwise.

For sufficiently long sensing period k (i.e, as $k \to \infty$), $F_k^{(i)}(y)$ converges to the malicious user distribution $F_M(y)$ or non-malicious user distribution $F_N(y)$. As the malicious user always reports binary one, i.e, presence of PU (H_1 hypothesis), the distribution $F_M(y)$ can be written as,

$$F_M(y) = F(y|H_1) \tag{27}$$

However, as non-malicious or legitimate users send the true decision to the FC about the presence of PU, the distribution $F_N(y)$ according to the total probability theorem can be written as,

$$F_N(y) = P(H_0)F(y|H_0) + P(H_1)F(y|H_1) \tag{28}$$

where $P(H_0)$ and $P(H_1)$ are the probabilities of PU being in the free or busy state, respectively.

In [16], AD statistic is used to determine the empirical distribution

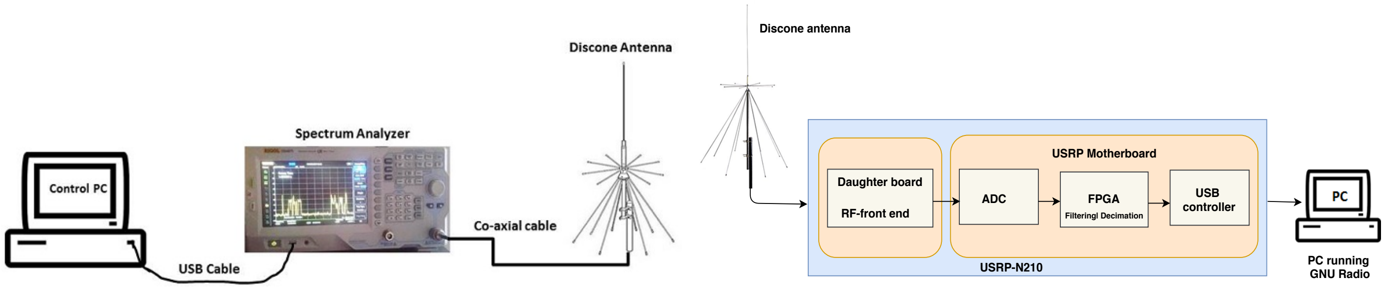
$$\begin{aligned} A_M^{(i)}(k) = -\frac{1}{k} \sum_{j=1}^{k} (2j-1)[\log F_M(y_{j:k}) \\ + \log(1 - F_M(y_{k+1-j:k}))] \end{aligned} \tag{29}$$

where $y_{j:k}$ is the j^{th} smallest value among k reported local test statistics. Once $A_M^{(i)}(k)$ is computed for each SU, it is compared with the cutoff threshold η. If $A_M^{(i)}(k) < \eta$, the i^{th} SU is considered malicious and its local test statistic is cut off from the global sensing decision [16].

According to [8], LRS outperforms AD sensing. Moreover, LRS has highest statistical power in all GoF tests [8]. To the best of our knowledge, no work in the literature has considered the LRS-based detection of malicious user in cooperative spectrum sensing. Thus, instead of calculating the empirical distribution of SUs as per (29), we calculate the empirical distribution at each SU, as mentioned in [8] as follows:

$$Z_c^{(i)} = \sum_{s=1}^{n} \left[\log \left\{ \frac{F_0(y[s])^{-1} - 1}{\left(n - \frac{1}{2}\right) / \left(s - \frac{3}{4}\right) - 1} \right\} \right]^2 \tag{30}$$

where $Z_c^{(i)}$ is the Zhang statistic computed at the i^{th} SU and other terms are similar as per the equation of Z_c in Algorithm-1 and Algorithm-2.



(a) Spectrum analysis.

(b) Spectrum data acquisition.

Fig. 3: Measurement platform deployed in this study.

Table 2 Channels measured in this study and USRP configuration

Radio Technology	Channel Number	F_{start}(MHz)	F_{center}(MHz)	F_{stop}(MHz)	Signal Bandwidth (MHz)	Gain (dB)	Decimation Rate (M)	Sampled Bandwidth (MHz)
FM Broadcasting		96.500	96.700	96.900	0.2	45	64	1
UHF Television Band-IV	U - 29 U - 33	534 566	538 570	542 574	8	45	8	8
E-GSM(900) DL	27 77 120	940.2 950.2 958.8	940.4 950.4 959	940.6 950.6 959.200	0.2	45	64	1
DCS (1800) DL	690	1839.6	1840.8	1841	0.2	45	64	1

The cut-off probability for a malicious user p_M^{cut} (i.e., the probability that a malicious user is detected and cut off) is calculated in a way such that it remains higher than a target value ζ as follows [16]:

$$p_M^{cut} = Pr(Z_c^{(i)} < \eta \mid \text{user i is malicious}) \geq \zeta \tag{31}$$

Further, simulation shows that ILRS and LRS performs similar for the detection of malicious user and thus we have restricted our analysis to LRS based malicious user detection.

7 Measurement and Experimental Setup

7.1 Empirical Setup

We deployed a measurement setup for data acquisition on the rooftop of School of Engineering and Applied Science (SEAS), GICT building of Ahmedabad University. The measurement setup is as shown in Fig. 3. It includes both hardware and software. Hardware consists of digital spectrum analyzer (Rigol DSA-875), a universal software radio peripheral (USRP-N210) with a WBX daughter board surmounted, two D3000N super discone antenna and a computer system to interface the hardware and software. The software part includes MATLAB and GNURadio. The measurement setup is similar to the one deployed in [37] and [38].

7.2 Spectrum Data Acquisition

Spectrum data is crucial in validating any proposed algorithm/scheme. Table 2 shows the channels measured in this study and the USRP configuration. We captured the spectrum data across four radio technologies namely FM broadcasting, UHF TV Band, E-GSM 900 downlink and DCS 1800 downlink. The prime reason for capturing the data across varied bands is to check the generalization of the proposed scheme. To make sure that the signal is actually present in the channel, we tune the spectrum analyzer to the desired channel/frequency band and look at the power spectral density on the analyzer's screen. We can clearly discriminate whether signal is present or absent (i.e, channel is idle or busy). This ensures that we are not receiving any noise only samples. This exercise is more useful in case of discontinuous power transmitters (e.g, GSM-900 and GSM-1800 downlink) to determine the presence of signal in the desired channel. Table 3 shows the tuning parameters of spectrum analyzer.

Table 3 Parameters of spectrum analyzer

Parameter	Value
Frequency range	75-2000 MHz
Frequency span	45-600 MHz
Frequency bin	Depends on band selected
Resolution Bandwidth-RBW	10 kHz
Video Bandwidth-VBW	10 kHz
Measurement period	5-15 mins
Sweep time	1 second
Scale	10 dB/division
Input attenuation	0 dB
Detection type	RMS detector

Two D3000N super discone antennas were connected, one with USRP and other with spectrum analyzer. The D3000N super discone antenna is an ultra-wideband antenna covering amateur radio, commercial 2-way, cellular, air traffic control and various utility frequency bands. It covers a frequency range of 25-3000 MHz. It is vertically polarized and has a nominal gain of 2 dBi. As the sensitivity is a key issue in spectrum measurements, we used the inbuilt pre-amplifier which amplifies all the incoming signals. As we have restricted our measurements maximum up-to 3000 MHz, there was no need of external amplifier as such. Also the length of co-axial cable was limited to the required range so as to avoid the loss of signal strength of incoming signals. Rigol DSA-875 spectrum analyzer supports 601 frequency points. Also as per the discussions in [39] a resolution bandwidth (RBW) of 10 kHz was selected and a sweep period of 1 second was kept. The frequency bins selected in the spectrum analyzer were kept slightly wider than those selected in USRP. This ensures that we are not missing the extreme points of the selected frequency bins.

Data acquisition and spectrum measurement was conducted on monday morning from 10:00 AM to 11:00 AM. These timings only

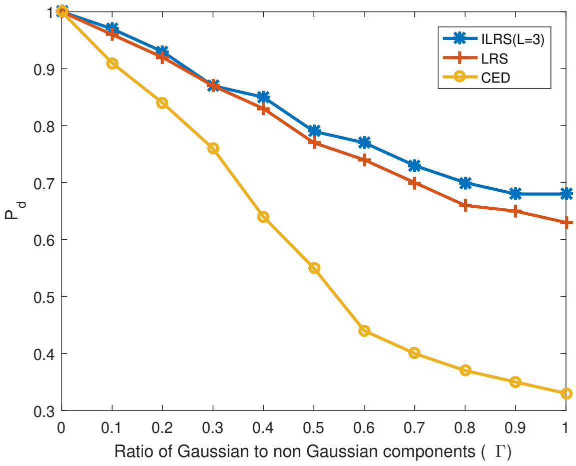


Fig. 4: P_d vs $\mathbf{\Gamma}$ at SNR=-5 dB and impulse index A=0.1

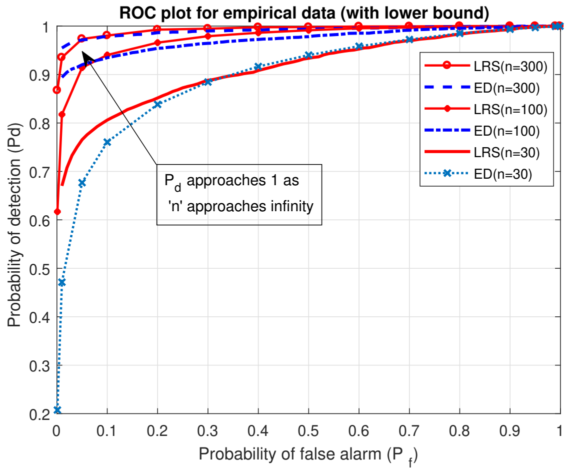
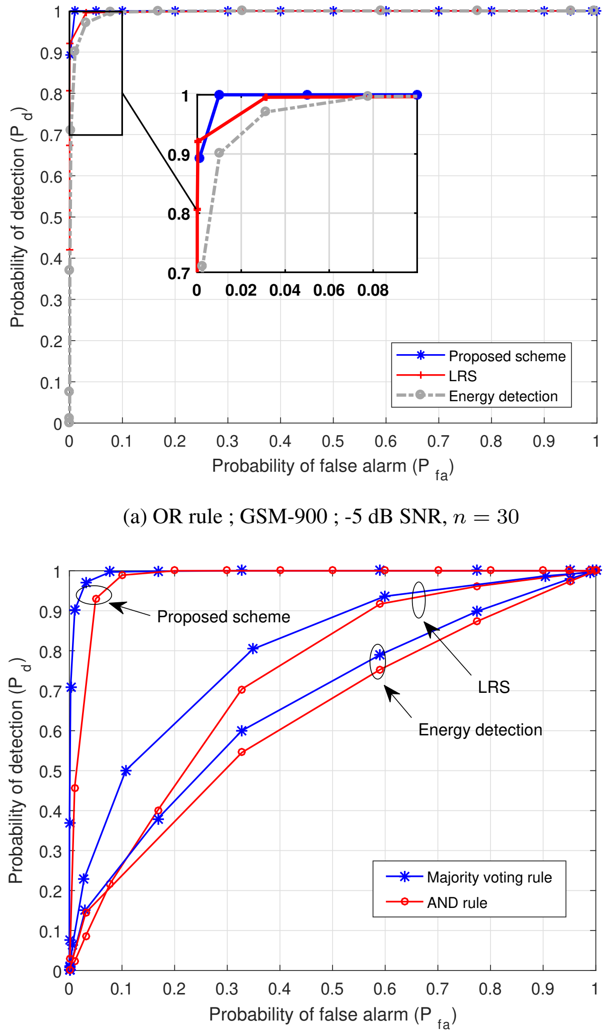


Fig. 5: Lower bound of LRS



(a) OR rule ; GSM-900 ; -5 dB SNR, $n = 30$

(b) Majority voting rule and AND rule; GSM-900 ; -5 dB SNR, $n = 30$

Fig. 6: ROC comparison of proposed scheme, LRS and ED

affect the discontinuous transmitters. We used a Dell-i7 embedded processor personal computer (PC) system and connected the spectrum analyzer with PC using instrument control toolbox of MATLAB for quick access. Various environmental factors like temperature, humidity, moisture, dew, rain, etc., need to be considered as elaborated in [38]. However, as our measurement was of short duration we did not consider such factors into measurement account.

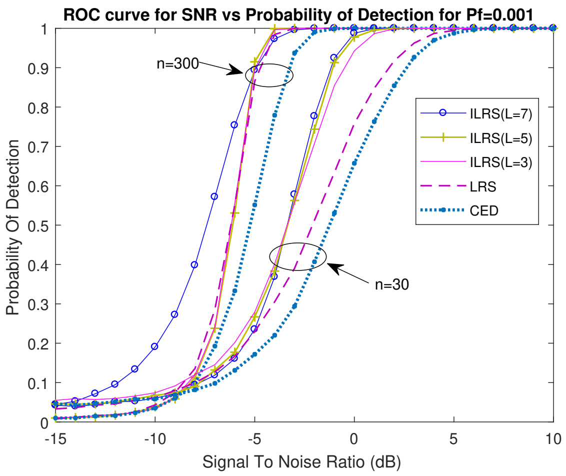
For data acquisition, the controlling computer runs the GNU radio's script to collect the digital samples (I/Q data) from USRP. Once the data are captured, off-line processing is performed in MATLAB and then the proposed scheme is applied on the stored data to check its validity.

8 Simulation Results and Discussion

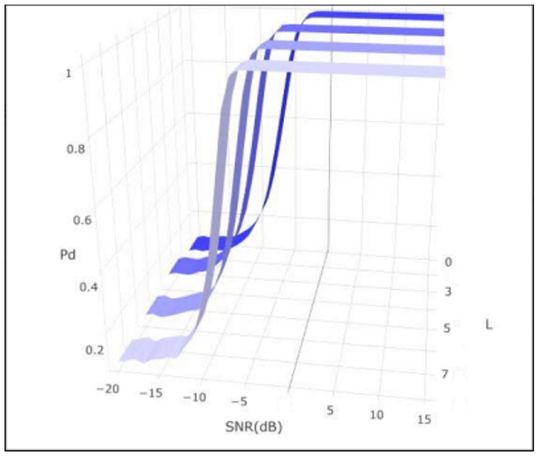
Several experimental and simulation results are provided to validate the proposed scheme for CSS. Note that in all of the figures, n is the sample size, N is the number of SUs and L is the number of past sensing events considered. For the results in Fig.7-9, the PU signal captured using the empirical test bed is forwarded to N SUs, considering the assumptions in Section-II. Further, randomly located SUs are generated and as SNR levels at SUs is unknown, AWGN is added to obtain appropriate SNR for simulations.

Fig.4 shows the effect of non-Gaussian noise components on detection probability at the SU in accordance with [28]. The plot of detection probability P_d vs ratio of Gaussian to non-Gaussian component $\mathbf{\Gamma}$ at SNR=-5dB and impulse index $A = 0.1$ for CSS is shown. Five secondary users are considered. The plot compares the detection probability under non-Gaussian noise for the proposed scheme, LRS and CED. As conveyed from the plot, the proposed scheme outperforms LRS and CED for non-Gaussian noise considered at SU. Also we can observe from plot that at low $\mathbf{\Gamma}$, the non-Gaussian component will be dominating and the proposed ILRS scheme performs better than CED and LRS in non-Gaussian environment as well. This can be ascribed to the smart combination of the test statistics across sensing events that the proposed ILRS scheme performs, which helps prevent missed detections (regardless of the type of noise) caused by sudden energy drops below the decision threshold, which are particularly likely to occur at low SNR conditions (as considered in Fig. 4). This operating principle leads to an improved detection performance of the proposed method in several operation conditions as it will be shown in the subsequent results presented in this section.

Fig.5 shows the receiver operating characteristic (ROC) plot for lower bound of LRS. It is depicted that detection probability $P_d \to 1$ with the $\mathcal{O}(e^{-C(2n)})$ as $n \to \infty$. It is interesting to note that the P_d lower bound of LRS is better than the P_d of ED for all the values of n, which suggests a better performance as demonstrated experimentally in this section. This result indicates that the LRS statistic tends to lead to a better performance than the conventional ED method and this is, in part, where the improved performance of the proposed ILRS method comes from (in addition to the smart combination of the LRS statistics from several sensing events).

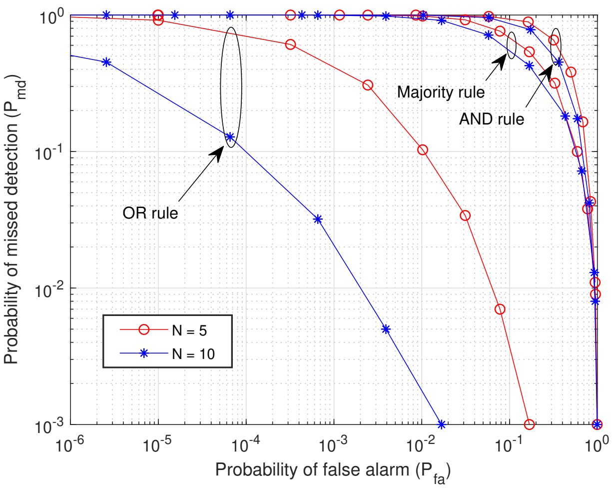


(a) P_d v/s SNR

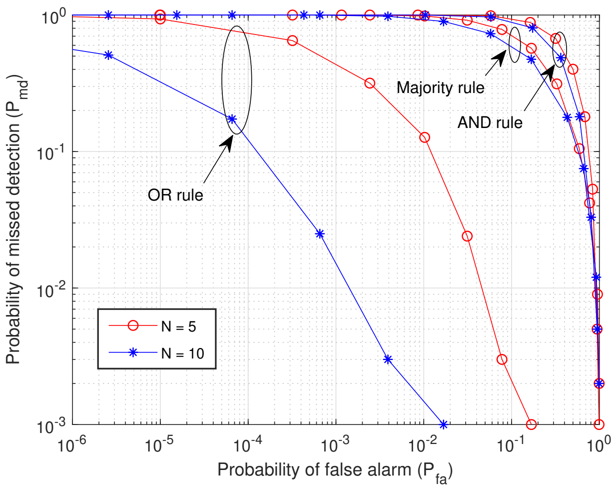


(b) 3D view for different values of L

Fig. 7: P_d v/s SNR comparison at $P_f = 0.001$ and for $n = 30$ and 300.



(a) GSM



(b) UHF TV

Fig. 8: ILRS fusion rule comparison using complementary ROC plot for different number of secondary users (N)

Fig. 6 shows the ROC plots for comparison of proposed scheme, LRS and ED sensing across GSM at -5dB SNR and for sample size $n = 30$. Also, the ROC plot compares the (a) OR rule, (b) Majority voting rule and AND fusion rules at FC. It is obvious that detection probability will be better in case of OR fusion rule. It is clear from the plot that ILRS OR rule outperforms ILRS AND fusion rule at the FC. ROC plot demonstrates that the performance of proposed scheme outperforms both, LRS and ED based CSS at low SNR. However, the improvement in performance is at the cost of nominal computational time as would be discussed in Section 9. Another interesting observation from Fig. 6 is that the performance of the proposed ILRS scheme is significantly less sensitive to the fusion rule used at the FC, showing a consistent high performance in all cases, while the other considered schemes can sometimes lead to a lower performance, in particular when the majority voting and the AND rules are employed. This is somehow natural since the OR rule is more permissive with missed detections (i.e., it is enough to have one or few busy channel states correctly detected in order to correctly report the channel as busy), while on the other hand the majority voting and AND rules require most or all of the busy states of the channel to be correctly detected, and here missed detections have a more degrading effect. By overcoming the impact of missed detections caused by sudden energy drops, the proposed ILRS scheme leads to a better performance, also with the majority voting and AND rules. As a result the proposed ILRS scheme consistently outperforms the other methods regardless of the fusion rule used at the FC.

Fig.7a shows the probability of detection versus SNR of proposed scheme (for different values of L), LRS and ED based sensing at SU for captured GSM data, $P_f = 0.001$, sample size n =30 and 300, using AND fusion rule at FC. It is clear from the plot that LRS based sensing outperforms ED sensing. However for localized sensing at SUs, as mentioned in Algorithm-2, we have also considered the past L sensing events so as to avoid any misdetection due to instantaneous drop in signal. The main challenge in considering the past events is in deciding the number of past sensing events L that should be taken into consideration. It is obvious that higher the value of L, less will be the chances of misdetection due to instantaneous drop, but at the cost of longer execution time and larger memory requirements of storing the previous values. We have considered three values of L (i.e, $L = 3, 5, 7$). We can depict from Fig.7a that the proposed scheme outperforms LRS and ED based CSS in the low SNR regime. Fig.7b shows the plot of probability of detection versus SNR, with number of past sensing events L added as a third dimension for n =300. The obtained results show, as expected, that increasing the number of sensing events L combined by the proposed

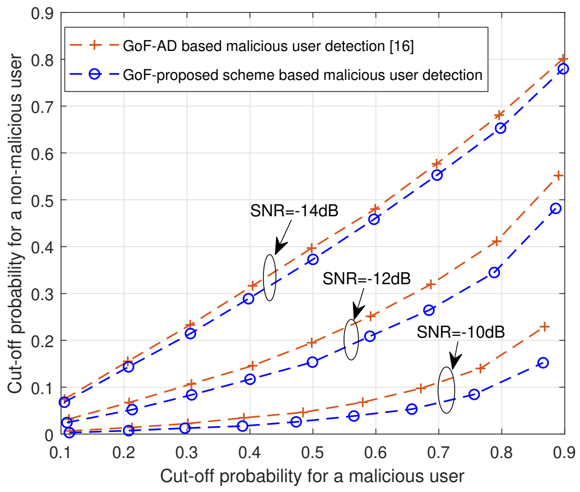


Fig. 9: Malicious user detection (p_N^{cut} vs p_M^{cut}) in CSS for $n = 50$, no. of SUs = 30 and no. of malicious users = 5.

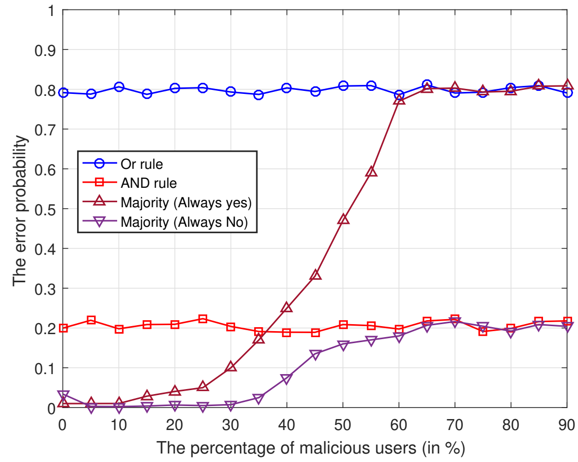


Fig. 10: Performance of generalized Byzantine attacks in CSS for $n = 50$, no. of SUs = 30.

ILRS scheme, a better probability of detection can be obtained. However, while the performance improvement from $L = 3$ to $L = 5$ is noticeable (in particular when a low sample size such as n=30 is considered), the improvement from $L = 5$ to $L = 7$ is not so significant, at least in the region of high probability of detection (0.9 and above), which is the operating region of interest. Therefore, $L = 5$ can be considered as a good tradeoff between the performance of the proposed scheme and its computation time and memory requirements (to store past sensing events). It is also worth noting in Fig. 7 that the enhanced probability of detection of the proposed ILRS scheme can lead to sensitivity improvements at $P_d = 0.9$ of up to 4 dB with respect to the conventional ED scheme and about 3 dB with respect to the standard LRS scheme.

Fig.8 shows the fusion rule comparison at SU using complementary receiver operating characteristic (CROC) for the proposed scheme for $N = 5$ and 10. The plot demonstrates three fusion rules at the FC i.e, (AND, majority & OR). In AND rule, FC makes the final decision about the channel only if the binary decisions of all the SUs are one. In majority rule, the final decision at the FC is made when majority of the SU reports one. On the other hand, in OR rule even if one SU sends its localized binary decision to be one, the FC decides the presence of PU. It is clear from the plot that OR rule outperforms AND fusion rule for both, GSM in Fig.8a and UHF TV in Fig.8b. Similar performance has been obtained for other radio technologies as well. Moreover, it is intuitive from Fig.8 that, as the number of SUs reporting to FC increases, the performance improves (i.e, lower will be the missed detection probability at particular value of false alarm). This is valid for all three, AND, Majority and OR fusion rules.

Fig.9 shows the plot of cut-off probability for non-malicious users p_N^{cut} versus cut-off probability for malicious users p_M^{cut}. As per [16] and references therein, the probability of PU being in busy state is assumed to be 20%, i.e, P(H_1) = 0.2. Total number of SUs is assumed to be 30 out of which 5 SUs are always malicious and the sample size is taken as $n = 50$ samples. The cut-off probability for a malicious user p_M^{cut} is calculated according to [19 - eq.(6)] and as per (31) for different target cut-off probabilities and similar calculation is carried out for cut-off probability for a non-malicious user p_N^{cut} at different SNR. Simulation result in Fig.9 shows that the proposed scheme clearly outperforms the AD based malicious user detection for CSS in terms of cutting off the malicious user even in the low SNR regime. For example, at -14 dB SNR, $p_M^{cut} > p_N^{cut}$ for both schemes. However, for the proposed LRS-based malicious user detection scheme, the p_N^{cut} is lower at a particular value of p_M^{cut}, which shows its applicability.

Fig. 10 indicates the plot of error probability v/s the percentage of malicious users considered for the proposed scheme. As similar to the fig. 9, we consider the no. of SUs to be 30 and $n = 50$. Additionally, we also consider that $P(H_0) = 0.8$ and $P(H_1) = 0.2$. It is evident that OR and AND saturates at $P(H_0)$ and $P(H_1)$ respectively, since they always declare the channel as either busy or idle. Thus, OR and AND rule will remain same for always yes and always no case. However, notice that the majority rule saturates at 0.8 for always yes case and at 0.2 for always no case.

9 Computational time and Complexity Analysis

Computational complexity is another metric used to quantify the computational cost of an Algorithm. In this context, computational complexity of ED, IED, LRS and proposed scheme are compared for CSS. Computational complexity for ED [5] is $\mathcal{O}(nK)$. However, for cooperative sensing one more iterative loop is included which equals the number of SUs (N) and thus it results in the complexity order of $\mathcal{O}(nNK)$. IED in [6] includes past few samples to improve the performance over ED. However that does not contribute in computational order and thus it has same computational order of $\mathcal{O}(nK)$. Also similar to cooperative ED sensing, it has the complexity order of $\mathcal{O}(nNK)$.

LRS based sensing has very high statistical power however it comes at a computational cost. LRS based sensing has the computational complexity of the order $\mathcal{O}(n^2 N \log(n))$ wherein $n \log(n)$ is contributed due to inverse CDF in line 16 of Algorithm-1, n is contributed due to summation term in the same line. Increase in the complexity order by N is due to the iterative loop which is equal to the number of SUs in line 14 of Algorithm-1 and K is due to line 13 of Algorithm-1.

The proposed scheme differs from the LRS based cooperative sensing in a way that later considers past L samples into account for calculating the test statistic as per line 17 of Algorithm-2. However it does not contribute in computational order and thus the proposed scheme has the same computational complexity of order $\mathcal{O}(n^2 N \log(n))$ as that of LRS based cooperative sensing while providing additional performance improvements. Table 4 provides the summary of computational complexities of various spectrum sensing schemes for cooperative spectrum sensing.

Table 4 Computational complexity for CSS

Spectrum sensing scheme	Computational order
Energy detection (ED)	$\mathcal{O}(nNK)$
Improved Energy detection (IED)	$\mathcal{O}(nNK)$
Likelihood Ratio Statistics (LRS)	$\mathcal{O}(n^2 N \log(n) K)$
Proposed Scheme (ILRS)	$\mathcal{O}(n^2 N \log(n) K)$

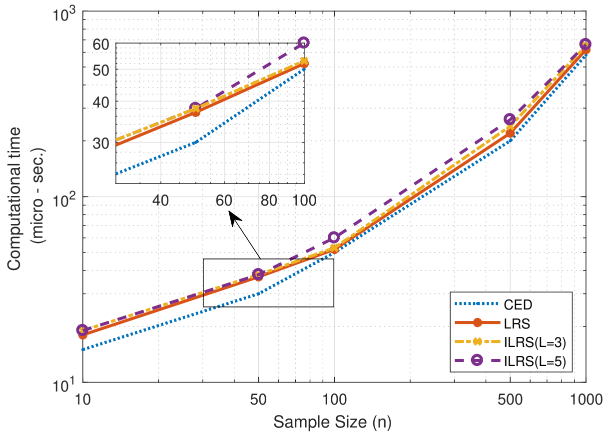


Fig. 11: Comparison of computational time

Fig.11 shows the comparison of computational time of different schemes as observed in the simulations on MATLAB platform in a PC with an Intel Core i7 processor at 3.60 GHz. It is evident from Fig.11 that at lower sample size ($n < 50$) the computational time of CED is slightly lower than for LRS and the proposed scheme. We can also note that for different values of L, the computational time is almost similar. The extra cost of computational time is in practice very low and is justified by the significant performance improvements as shown in simulation results section. Another important thing to mention here is that in order to obtain a satisfactory detection performance, the sample size n needs to be relatively large and in that region the computation time is similar. Therefore the improved performance is obtained at very low increase in the computational time.

10 Conclusions

In this work, the localized sensing at SUs in cooperation is studied. We have considered LRS as a test statistic, due to its highest statistical power, for localized sensing at SUs in cooperative spectrum sensing to mitigate the poor performance of ED at low SNR. Further, an Improved LRS for CSS is proposed in which the past L sensing events is considered so as to avoid misdetections caused due to sudden drops in signal energy. From numerical results and discussion we conclude that the proposed ILRS scheme outperforms LRS and ED based CSS at low SNR regime.

The proposed scheme has been concretely verified with empirical spectrum data of various radio technologies. Furthermore, the proposed scheme outperforms ED and LRS based sensing under non-Gaussian noise distribution of SUs. Also, taking into consideration the security aspect in CSS wherein users in cooperation may act maliciously by reporting false hard decisions at FC, the proposed scheme has outperformed AD based malicious user detection in CSS even at low SNR of -14 dB considering that malicious user always report yes at the FC. In addition, we have also considered a more realistic Generalized Byzantine attack wherein the SU may not always report yes at FC. ROC, CROC and detection probability versus SNR plots also indicates the better performance of proposed scheme as compared to ED and LRS in CSS at low SNR regime. The computational time analysis indicates that at lower sample size, CED has lower time complexity than LRS and ILRS schemes but at higher sample size, computational time is almost similar. The extra cost of computational time at lower sample size is in practice very low and is justified by the significant performance improvements of the proposed scheme. Furthermore, the complexity order of LRS and proposed scheme is same in CSS. As LRS and ILRS perform better than CED at low SNR regime, these sensing schemes can be used as a future scope to obtain PU activity statistics with better accuracy, a problem which has been untouched in literature till date, and will be addressed in our future work.

11 Acknowledgments

This work was supported by DST-UKIERI thematic partnership under the generous grant DST/INT/UK/P-150/2016. The authors also thank School of Engineering and Applied Science, Ahmedabad University and University of Liverpool for the infrastructural support.

12 References

[1] Haykin, S.: 'Cognitive radio: brain-empowered wireless communications', *IEEE J Sel Areas Commun*, 2005, **23**, (2), pp. 201–220
[2] Chen, Y., Oh, H.S.: 'A survey of measurement-based spectrum occupancy modeling for cognitive radios', *IEEE Commun Surveys Tuts*, 2016, **18**, (1), pp. 848–859
[3] Mitola, J., Maguire, G.Q.: 'Cognitive radio: making software radios more personal', *IEEE Pers Commun*, 1999, **6**, (4), pp. 13–18
[4] Yucek, T., Arslan, H.: 'A survey of spectrum sensing algorithms for cognitive radio applications', *IEEE Commun Surveys Tuts*, 2009, **11**, (1), pp. 116–130
[5] Urkowitz, H.: 'Energy detection of unknown deterministic signals', *Proc IEEE*, 1967, **55**, (4), pp. 523–531
[6] López.Benítez, M., Casadevall, F.: 'Improved energy detection spectrum sensing for cognitive radio', *IET commun*, 2012, **6**, pp. 785 – 796
[7] Wang, H., Yang, E.H., Zhao, Z., Zhang, W.: 'Spectrum sensing in cognitive radio using goodness of fit testing', *IEEE Trans Wireless Commun*, 2009, **8**, (11), pp. 5427–5430
[8] Patel, D.K., Trivedi, Y.N.: 'LRS-G^2 Based Non-parametric Spectrum Sensing for Cognitive Radio', *Proc CROWNCOM*, 2016, pp. 330–341
[9] Letaief, K.B., Zhang, W.: 'Cooperative communications for cognitive radio networks', *Proc IEEE*, 2009, **97**, (5), pp. 878–893
[10] Sendonaris, A., Erkip, E., Aazhang, B.: 'User cooperation diversity. part i. system description', *IEEE Trans Commun*, 2003, **51**, (11), pp. 1927–1938
[11] Cabric, D., Mishra, S.M., Brodersen, R.W.: ; Ieee. 'Implementation issues in spectrum sensing for cognitive radios', *Signals, systems and computers, 2004 Conference record of the thirty-eighth Asilomar conference on*, 2004, **1**, pp. 772–776
[12] Atapattu, S., Tellambura, C., Jiang, H.: 'Energy detection based cooperative spectrum sensing in cognitive radio networks', *IEEE Trans Wireless Commun*, 2011, **10**, (4), pp. 1232–1241
[13] Singh, A., Bhatnagar, M.R., Mallik, R.K.: 'Cooperative spectrum sensing with an improved energy detector in cognitive radio network', *Proc IEEE NCC*, 2011, pp. 1–5
[14] Sahai, A., Hoven, N., Tundra, R.: 'Some fundamental limits on cognitive radio', *Proc Allerton Conf, Monticello*, Oct. 2004, pp. 131–136
[15] Chen, C., Song, M., Xin, C., Alam, M.: 'A robust malicious user detection scheme in cooperative spectrum sensing', *Proc IEEE GLOBECOM*, 2012, pp. 4856–4861
[16] Noh, G., Lim, S., Lee, S., Hong, D.: 'Goodness-of-fit-based malicious user detection in cooperative spectrum sensing', *Proc IEEE VTC-Fall*, 2012, pp. 1–5
[17] Wang, J., Chen, I.R., Tsai, J.J.P., Wang, D.C.: 'Trust-based mechanism design for cooperative spectrum sensing in cognitive radio networks', *Computer Communications*, 2018, **116**, pp. 90 – 100
[18] Gul, N., Qureshi, I.M., Naveed, A., Elahi, A., Rasool, I.: 'Secured soft combination schemes against malicious-users in cooperative spectrum sensing', *Wireless Personal Communications*, 2019, pp. 1–20
[19] Qin, Z., Gao, Y., Plumbley, M.D.: 'Malicious user detection based on low-rank matrix completion in wideband spectrum sensing', *IEEE Trans Signal Process*, 2018, **66**, (1), pp. 5–17
[20] Wu, J., Song, T., Yu, Y., Wang, C., Hu, J.: 'Generalized byzantine attack and defense in cooperative spectrum sensing for cognitive radio networks', *IEEE Access*, 2018, **6**, pp. 53272–53286
[21] Wu, J., Song, T., Yu, Y., Wang, C., Hu, J.: 'Sequential cooperative spectrum sensing in the presence of dynamic byzantine attack for mobile networks', *PloS one*, 2018, **13**, (7), pp. e0199546
[22] Nguyen.Thanh, N., Kieu.Xuan, T., Koo, I.: 'Comments and corrections comments on "spectrum sensing in cognitive radio using goodness-of-fit testing"', *IEEE Trans Wireless Commun*, 2012, **11**, (10), pp. 3409–3411
[23] Teguig, D., Nir, V.L., Scheers, B.: 'Spectrum sensing method based on likelihood ratio goodness-of-fit test', *IET Electronics Letters*, 2015, **51**, (3), pp. 253–255
[24] Arshad, K., Moessner, K.: 'Robust spectrum sensing based on statistical tests', *IET Commun*, 2013, **7**, (9), pp. 808–817
[25] Li, L., Hou, S., Anderson, A.L.: 'Kernelized generalized likelihood ratio test for spectrum sensing in cognitive radio', *IEEE Trans Veh Technol*, 2018, **67**, (8), pp. 6761–6773
[26] Yarkan, S., Toreyin, B.U., Qaraqe, K.A., Cetin, A.E.: 'An online adaptive cooperation scheme for spectrum sensing based on a second-order statistical method', *IEEE Trans Veh Technol*, 2012, **61**, (2), pp. 675–686
[27] Bera, D., Chakrabarti, I., Pathak, S.S., Karagiannidis, G.K.: 'Another look in the analysis of cooperative spectrum sensing over nakagami- m fading channels', *IEEE Trans Wireless Commun*, 2017, **16**, (2), pp. 856–871
[28] Blackard, K.L., Rappaport, T.S., Bostian, C.W.: 'Measurements and models of radio frequency impulsive noise for indoor wireless communications', *IEEE J Sel Areas commun*, 1993, **11**, (7), pp. 991–1001
[29] Patel, D.K., Trivedi, Y.N.: 'Goodness-of-fit-based non-parametric spectrum sensing under middleton noise for cognitive radio', *Electronics Letters*, 2015, **51**, (5), pp. 419–421

[30] Middleton, D.: 'Non-gaussian noise models in signal processing for telecommunications: new methods an results for class a and class b noise models', *IEEE Trans Inf Theory*, 1999, **45**, (4), pp. 1129–1149
[31] Digham, F.F., Alouini, M.S., Simon, M.K.: 'On the energy detection of unknown signals over fading channels', *IEEE Trans Commun*, 2007, **55**, (1), pp. 21–24
[32] Cressie, N., Read, T.R.: 'Multinomial goodness-of-fit tests', *Journal of the Royal Statistical Society Series B (Methodological)*, 1984, pp. 440–464
[33] Zhang, J., Wu, Y.: 'Likelihood-ratio tests for normality', *Computational statistics & data analysis*, 2005, **49**, (3), pp. 709–721
[34] Zwillinger, D.: 'Table of Integrals, Series, and Products, Eighth Edition'. (, 2014)
[35] Anderson, T.W., Darling, D.A.: 'Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes', *The Ann Math Stat*, 1952, **23**, (2), pp. 193–212
[36] Kvam, P.H., Vidakovic, B.: 'Nonparametric statistics with applications to science and engineering'. vol. 653. (John Wiley & Sons, 2007)
[37] López.Benítez, M., Casadevall, F., Martella, C.: 'Performance of spectrum sensing for cognitive radio based on field measurements of various radio technologies', *European Wireless Conference (EW)*, 2010, pp. 969–977
[38] Wellens, M., Mähönen, P.: 'Lessons learned from an extensive spectrum occupancy measurement campaign and a stochastic duty cycle model', *Mob networks Appl*, 2010, **15**, (3), pp. 461–474
[39] López.Benítez, M., Casadevall, F.: 'Methodological aspects of spectrum occupancy evaluation in the context of cognitive radio', *Transactions on Emerging Telecommunications Technologies*, 2010, **21**, (8), pp. 680–693