Goodness-of-Fit-based Malicious User Detection in Cooperative Spectrum Sensing

Gosan Noh†, Sungmook Lim, Seokwon Lee, and Daesik Hong‡

Information and Telecommunications Lab., Dept. of Electrical and Electronic Eng., Yonsei Univ.
50 Yonsei-ro, Seodaemun-gu, Seoul, Korea, 120-749
Phone: +82-2-2123-3558, Fax: +82-2-312-4887
Homepage: http://mirinae.yonsei.ac.kr
E-mail: gsnoh@yonsei.ac.kr†, daesikh@yonsei.ac.kr‡

***Abstract*—Cooperative spectrum sensing improves sensing accuracy in primary user detection, but can be threatened by malicious users. Malicious users may try to falsify the sensing result to indicate that the primary user exists even when there is no primary user in order to monopolize the spectrum usage, thereby depriving other users of their spectrum opportunities. To address this, we propose a malicious user detection scheme where the malicious users are identified and cut off from the cooperative sensing process. The proposed scheme exploits the Anderson-Darling (AD) goodness-of-fit technique which tests whether the empirical distribution of the sensing data from each secondary user fits the expected distribution for a malicious user. In addition, we derive false alarm and detection probabilities for when malicious users are cut off by the malicious user detection scheme. Simulation results show that the proposed goodness-of-fit-based malicious user detection significantly improves sensing performance in comparison with conventional outlier detection-based schemes.**

***Index Terms*—Cognitive radio, spectrum sensing, malicious user detection.**

I. Introduction

Accurate spectrum sensing for the detection of a primary user (i.e., licensed user) is crucial for cognitive radio systems [1]. If a secondary user (i.e., unlicensed user) is not aware of the existence of the primary user due to sensing failure, it may transmit while the primary user is using the spectrum, thereby causing a collision between the primary and secondary users. However, accurate spectrum sensing is particularly difficult in a wireless environment due to the existence of shadowing and multipath fading.

Cooperative diversity is an approach employed for spectrum sensing as a way of overcoming these channel uncertainties [2], [3], [4]. In cooperative spectrum sensing, the local sensing data from geographically distributed secondary users are combined at the fusion center and used to decide on the existence of the primary user. Even though some of the secondary users may not be able to detect the primary user signal due to deep fading, the final decision on the existence of the primary user can be made using combined sensing data from other secondary users that are not in deep fades.

Despite the merits of cooperative sensing, malicious users have the potential to degrade sensing performance [5]. Specifically, some malicious users might intentionally send false sensing data to the fusion center in order to disrupt the final sensing result. If a malicious user falsely reports that a primary user exists even though there is no primary user, the false alarm probability will be increased, thus reducing the opportunity for other users to utilize the spectrum. To counter this possibility, several malicious user detection techniques have been proposed [6], [7]. Kaligineedi *et al.* proposed a malicious user detection scheme based on outlier detection which identifies any observation that is far away from the rest of the others [6]. Zeng *et al.* introduced a reputation-based mechanism in addition to the outlier detection in order to enhance detection reliability as the sensing interval elapses [7]. However, the sensing performance of the conventional outlier detection-based schemes that exploit the statistical peculiarity of the observation is degraded when the signal-to-noise ratio (SNR) of the primary user is low. Spectrum sensing should be able to be done at a very low SNR, e.g., -15 dB [8]. In this situation, it becomes difficult to discriminate between the falsified signal of the malicious user and the noise.

Therefore, in this paper, we propose a *goodness-of-fit*-based malicious user detection scheme that aims to minimize sensing performance degradation due to the existence of malicious users by utilizing behavioral differences between malicious and non-malicious users. Assuming energy detection for local spectrum sensing, malicious users always report high energy values (i.e., signal-plus-noise energy) to the fusion center while non-malicious users report either low (i.e., noise-only energy) or high energy values according to the state of the primary user (i.e., idle or busy state) [6], [7]. This behavioral difference is reflected in the reporting energy distributions. The energy distribution of a malicious user will be concentrated around the point of the signal-plus-noise energy value, while that of a non-malicious user will be divided into two points, i.e., noise-only and signal-plus-noise energy values.

In the proposed goodness-of-fit-based malicious user detection procedure, we first obtain an Anderson-Darling (AD)

This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2010-0018938), and in part by the KCC (Korea Communications Commission), Korea, under the R&D program supervised by the KCA (Korea Communications Agency) (KCA-2011-(11913-04004)).

978-1-4673-1881-5/12/$31.00 ©2012 IEEE

statistic for measuring how well the empirical distribution of the reported energy values fits the expected distribution of a malicious user. We can then detect whether a secondary user is malicious by comparing the obtained AD statistic with the cut-off threshold. In this way, the proposed goodness-of-fit-based scheme can detect and cut off malicious users from having their sensing data combined with the rest at the fusion center. In addition, we derive false alarm and detection probabilities for when malicious users are cut off per a cut-off probability given by the malicious user detection scheme. Simulation results verify the effectiveness of the proposed scheme.

The rest of this paper is organized as follows. Section II presents the system model. In Section III, we introduce the proposed goodness-of-fit-based malicious user detection scheme and its effect on sensing performance. We conclude this paper in Section IV.

II. System Model

For our model, we begin by considering a cognitive radio network with M secondary users. Each secondary user conducts local spectrum sensing within a sensing time of N samples. We assume energy detection for the physical spectrum sensing technique due to its ability to detect primary user signals without prior information [3], [8]. Let $y_i(k)$ be the local test statistic of the i-th secondary user, which can be obtained by energy detection. The binary hypothesis test for local spectrum sensing at the k-th sensing interval is then formulated as

$$\begin{aligned}\mathcal{H}_0 : y_i(k) &= \sum_{n=T_k}^{T_k+N-1} \left|v_i(n)\right|^2, \;\; i = 1, 2, \ldots, M \\ \mathcal{H}_1 : y_i(k) &= \sum_{n=T_k}^{T_k+N-1} \left|h_i(n)s(n) + v_i(n)\right|^2, \;\; i = 1, 2, \ldots, M,\end{aligned} \tag{1}$$

where T_k is the time instant at which the k-th sensing interval starts, $s(n)$ is the transmitted signal from the primary user, and $h_i(n)$ is the channel gain between the primary user and the i-th secondary user. Let us suppose that $s(n)$ is the complex phase-shift keying (PSK) modulated signal with unit power and $h_i(n)$ is characterized by independent Rayleigh fading, i.e., $h_i(n) \sim \mathcal{CN}(0, \sigma_h^2)$. The noise $v_i(n)$ is the complex-valued AWGN with variance σ_v^2, i.e., $v_i(k) \sim \mathcal{CN}(0, \sigma_v^2)$. According to the central limit theorem, if N is large enough (i.e., $N \geq 10$ [3]), $y_i(k)$ approaches the Gaussian random variable, as follows:

$$\begin{aligned}\mathcal{H}_0 : y_i(k) &\sim \mathcal{N}\big(N\sigma_v^2, N\sigma_v^4\big) \\ \mathcal{H}_1 : y_i(k) &\sim \mathcal{N}\big(N(\sigma_h^2 + \sigma_v^2), N(\sigma_h^2 + \sigma_v^2)^2\big).\end{aligned} \tag{2}$$

After local spectrum sensing is finished, the obtained local test statistics are transmitted to the fusion center via a common control channel, which is assumed to have a time-bandwidth product high enough to achieve error-free transmission [3], [9]. In addition, we assume soft combining at the fusion center, which is known to outperform hard combining [3], [4]. The final decision on the existence of the primary user is made by comparing the sum of the local test statistics, i.e., the global test statistic $u(k)$, with a detection threshold λ, as follows:

$$u(k) = \sum_{i=1}^{M} y_i(k) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda. \tag{3}$$

The cooperative sensing model presented above is designed with the assumption that there is no malicious user among the secondary users. However, if a malicious user exists, the performance of the cooperative sensing will be degraded. Therefore, malicious users need to be detected and excluded from participating in this cooperation for the protection of the non-malicious users.

III. Proposed Malicious User Detection

In this section, we first propose a new secure malicious user detection scheme using a goodness-of-fit test, and then evaluate the sensing performance when malicious users are detected and cut off with a cut-off probability given by the cooperative sensing.

A. Malicious User Detection with Goodness-of-Fit Test

The basic idea of the malicious user detection scheme being proposed here is to identify whether or not a secondary user is malicious by comparing its empirical distribution with the distribution expected for a malicious user. If the empirical distribution of a secondary user coincides with the expected distribution for malicious users, we can conclude that the secondary user is malicious and exclude its local test statistic from the combining process at the fusion center.

The degree of coincidence between the empirical and expected distributions is measured by a goodness-of-fit test. Let $F_k^{(i)}(y)$ be the empirical distribution of the i-th secondary user after the k-th sensing interval is finished. $F_k^{(i)}(y)$ is then defined as

$$F_k^{(i)}(y) = \frac{1}{k}\sum_{j=1}^{k} \mathbf{1}\big(y_i(j) \leq y\big), \tag{4}$$

where $\mathbf{1}(\alpha)$ is the indicator function, which equals 1 if α is true and 0 if α is false [10].

Depending on whether the secondary user is malicious, $F_k^{(i)}(y)$ converges to either $F_{\mathrm{N}}(y)$ (defined as the non-malicious user distribution) or $F_{\mathrm{M}}(y)$ (defined as the malicious user distribution) as $k \to \infty$. The distributions $F_{\mathrm{N}}(y)$ and $F_{\mathrm{M}}(y)$ reflect the specific reporting behavior for non-malicious and malicious users respectively, as follows:

- **Non-malicious user case:** Since a non-malicious user reports the local test statistic without any manipulation of the sensing data, the reporting frequency for the local test statistics between the $\mathcal{H}_0$ and $\mathcal{H}_1$ cases is the same as the proportion between the idle and busy periods of the primary user. Therefore, the distribution $F_{\mathrm{N}}(y)$ is the weighted combination of $F(y|\mathcal{H}_0)$ and $F(y|\mathcal{H}_1)$, i.e., $F_{\mathrm{N}}(y) = F(y|\mathcal{H}_0)P(\mathcal{H}_0) + F(y|\mathcal{H}_1)P(\mathcal{H}_1)$.
- **Malicious user case:** A malicious user always reports a high energy value to the fusion center to simulate the

presence of a primary user, thus preventing the other secondary users from utilizing the spectrum [6], [7]. Therefore, the distribution $F_{\rm M}(y)$ consists of only the $\mathcal{H}_1$ case, i.e., $F_{\rm M}(y) = F(y|\mathcal{H}_1)$.

In order to decide whether a secondary user is non-malicious or malicious (that is, whether $F_k^{(i)}(y)$ approaches $F_{\rm N}(y)$ or $F_{\rm M}(y)$ as $k \to \infty$), we adopt the Anderson-Darling (AD) test from among the various goodness-of-fit tests. The AD test is known to be one of the most powerful goodness-of-fit tests due to its enhanced sensitivity and lower computational burdens [10]. The AD statistic for determining whether the empirical distribution approaches the malicious user distribution is given by [10]

$$A_{\rm M}^{(i)}(k) = -\frac{1}{k}\sum_{j=1}^{k}(2j-1)\big[\log F_{\rm M}(y_{j:k}) + \log\big(1 - F_{\rm M}(y_{k+1-j:k})\big)\big], \tag{5}$$

where $y_{j:k}$ is the j-th smallest value among k reported local test statistics, i.e., $y_{1:k} \le y_{2:k} \le \cdots \le y_{k:k}$.

Since the AD statistic $A_{\rm M}^{(i)}(k)$ measures how likely it is that the i-th secondary user is a malicious user, the fusion center compares $A_{\rm M}^{(i)}(k)$ with a cut-off threshold η. If $A_{\rm M}^{(i)}(k)$ is lower than η (i.e., $A_{\rm M}^{(i)}(k) < \eta$), the i-th secondary user is considered a malicious user and its local test statistic is cut off from the global sensing decision.

The cut-off threshold η is calculated such that the cut-off probability for a malicious user $p_{\rm M}^{\rm cut}$ (i.e., the probability that a malicious user is detected and cut off) remains higher than a target value ζ as follows:

$$p_{\rm M}^{\rm cut} = \Pr\{A_{\rm M}^{(i)}(k) < \eta \,|\, \text{User } i \text{ is malicious}\} \ge \zeta. \tag{6}$$

The calculation of (6) can be done using the distribution of the AD statistic $A_{\rm M}^{(i)}(k)$. Although the exact distribution of $A_{\rm M}^{(i)}(k)$ is not analytically tractable, the limiting distribution can be obtained [11] as follows:

$$\begin{aligned}\lim_{k\to\infty} &\Pr\{A_{\rm M}^{(i)}(k) < \eta \,|\, \text{User } i \text{ is malicious }\} \\ &= \frac{\sqrt{2\pi}}{\eta}\sum_{j=0}^{\infty}\binom{-\frac{1}{2}}{j}(4j+1)\exp\left(-\frac{(4j+1)^2\pi^2}{8\eta}\right) \\ &\times \int_0^{\infty}\exp\left(\frac{\eta}{8(w^2+1)} - \frac{(4j+1)^2\pi^2 w^2}{8\eta}\right)dw.\end{aligned} \tag{7}$$

The limiting values of (7) are valid for $k \ge 5$ due to its fast convergence property [12]. Based on (7), we can calculate the cut-off threshold η with respect to the target cut-off probability for a malicious user ζ, as in Table I. For example, if we want to detect and cut off malicious users with a probability of 0.9, we can set the cut-off threshold to 1.933 and reject any local test statistic higher than the threshold.

The performance of this approach to malicious user detection is evaluated by simulation in Figs. 1 and 2. Since the spectrum occupied by licensed users in an urban area is between 10% and 30% [2], we assume the probability of a primary user being in a busy state to be $P(\mathcal{H}_1) = 0.2$. In addition, we set the sensing time for each secondary user at $N = 50$ and the number of secondary users to $M = 30$. Of these, $M_{\rm M} = 5$ users are assumed to be "Always Yes" malicious users [6], [7]. Although the results are based on the given simulation parameters, our approach is also useful for other simulation parameters.

Fig. 1 depicts the cut-off probability for a non-malicious user ($p_{\rm N}^{\rm cut}$) as a function of the cut-off probability for a malicious user ($p_{\rm M}^{\rm cut}$) for different primary user SNR values. "OD" denotes the conventional outlier detection-based scheme where the outlier factor measured by the sample mean and standard deviation is used to identify malicious users [6]. "GF" denotes the proposed goodness-of-fit-based scheme. To ensure that the comparison is fair, no weighting mechanism is employed in regard to the malicious user detection result for either the proposed or the conventional schemes. Note that in Fig. 1, the malicious user detection performance is better when $p_{\rm N}^{\rm cut}$ and higher $p_{\rm M}^{\rm cut}$ are achieved. It is clear that the cut-off probability for a non-malicious user under the proposed scheme is always lower than under the conventional scheme for given primary user SNRs. Note also that a lower $p_{\rm N}^{\rm cut}$ is achieved with higher primary user SNR, due to the reduced noise effect.

Fig. 2 shows the cut-off probabilities for non-malicious and malicious users for different sensing interval k. The figure shows that malicious user detection performance under the proposed goodness-of-fit-based scheme improves with the sensing interval k, while that of the conventional outlier detection-based scheme does not. This is because the empirical distribution used for the goodness-of-fit method becomes more accurate as more samples are obtained as the sensing interval elapses.

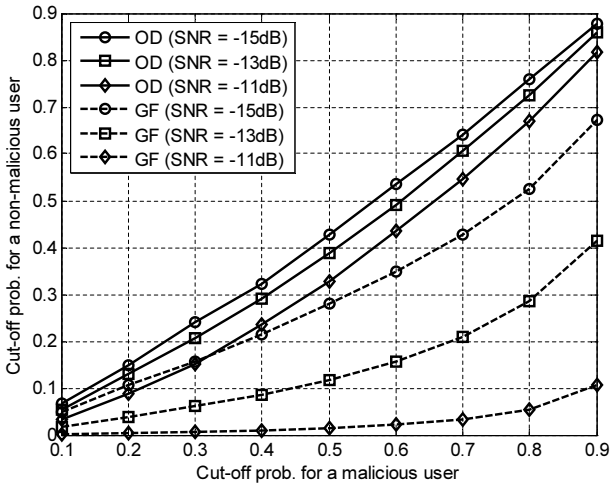

Fig. 1. Performance of malicious user detection schemes: Cut-off probability for a non-malicious user ($p_{\rm N}^{\rm cut}$) vs. cut-off probability for a malicious user ($p_{\rm M}^{\rm cut}$) for different primary user SNRs. The sensing interval is set to $k = 50$. The cut-off probability for a non-malicious user for the proposed scheme is always lower than the conventional scheme at a given cut-off probability for a malicious user and primary user SNR.

Table I
CUT-OFF THRESHOLDS WITH RESPECT TO THE TARGET CUT-OFF PROBABILITY FOR A MALICIOUS USER.

Target cut-off prob. ζ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Cut-off threshold η	0.346	0.448	0.546	0.652	0.774	0.923	1.120	1.408	1.933

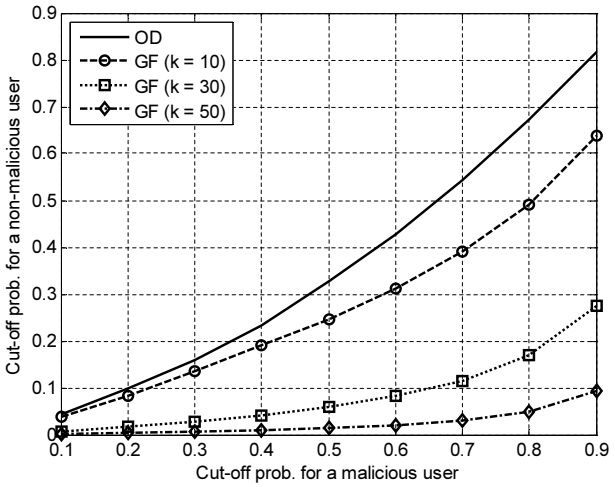

Fig. 2. Performance of malicious user detection schemes: Cut-off probability for a non-malicious user ($p_{\mathrm{N}}^{\mathrm{cut}}$) vs. cut-off probability for a malicious user ($p_{\mathrm{M}}^{\mathrm{cut}}$) for different sensing intervals k. The primary user SNR is -11 dB. The cut-off probability for a non-malicious user for the proposed goodness-of-fit-based scheme decreases with the sensing interval.

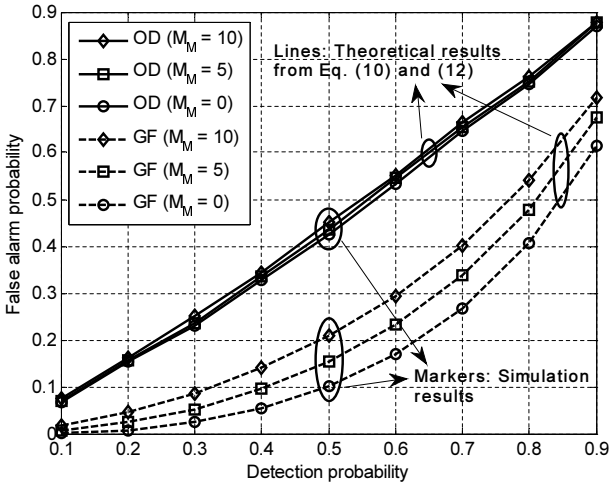

Fig. 3. Sensing performance with malicious user detection schemes: False alarm probability (P_{FA}) vs. detection probability (P_D) for different numbers of malicious users (M_{M}). Theoretical results from Eq. (10) and (12) (solid and dashed lines) and simulation results (markers) are plotted. The sensing interval is set to $k = 50$. The primary user SNR is -11 dB. False alarm probability of the proposed scheme is always lower than the conventional scheme.

B. Sensing Performance Improvement with the Proposed Malicious User Detection

We now proceed to investigate the degree to which the sensing performance (e.g., false alarm and detection probabilities) is improved by the proposed malicious user detection in conjunction with the malicious user detection performance discussed in Subsection III-A.

Let us suppose that the total number of secondary users of M consists of M_{N} non-malicious users and M_{M} malicious users, i.e., $M = M_{\mathrm{N}} + M_{\mathrm{M}}$. Without any malicious user detection scheme, all the local test statistics reported from M_{N} and M_{M} secondary users will be combined. In contrast, if the proposed malicious user detection is used, some of them will be detected as malicious users and cut off from the combining procedure. Since each non-malicious and malicious user can be cut off with a probability of $p_{\mathrm{N}}^{\mathrm{cut}}$ and $p_{\mathrm{M}}^{\mathrm{cut}}$ respectively, the numbers of cut-off non-malicious and malicious users, $m_{\mathrm{N}}^{\mathrm{cut}}$ and $m_{\mathrm{M}}^{\mathrm{cut}}$, are binomial random variables with probability mass functions (PMFs):

$$\begin{aligned} P(m_{\mathrm{N}}^{\mathrm{cut}}) &= \binom{M_{\mathrm{N}}}{m_{\mathrm{N}}^{\mathrm{cut}}} \left(p_{\mathrm{N}}^{\mathrm{cut}}\right)^{m_{\mathrm{N}}^{\mathrm{cut}}} \left(1-p_{\mathrm{N}}^{\mathrm{cut}}\right)^{M_{\mathrm{N}}-m_{\mathrm{N}}^{\mathrm{cut}}}, \\ P(m_{\mathrm{M}}^{\mathrm{cut}}) &= \binom{M_{\mathrm{M}}}{m_{\mathrm{M}}^{\mathrm{cut}}} \left(p_{\mathrm{M}}^{\mathrm{cut}}\right)^{m_{\mathrm{M}}^{\mathrm{cut}}} \left(1-p_{\mathrm{M}}^{\mathrm{cut}}\right)^{M_{\mathrm{M}}-m_{\mathrm{M}}^{\mathrm{cut}}}. \end{aligned} \tag{8}$$

The false alarm probability P_{FA} and detection probability P_D depend on the number of non-malicious and malicious users that are not cut off. P_{FA} is the probability that the combined energy from $M_{\mathrm{N}} - m_{\mathrm{N}}^{\mathrm{cut}}$ non-malicious users (which send noise-only energy when $\mathcal{H}_0$) and $M_{\mathrm{M}} - m_{\mathrm{M}}^{\mathrm{cut}}$ malicious users (which send primary signal-plus-noise energy when $\mathcal{H}_0$) exceeds the detection threshold λ when there is no primary user ($\mathcal{H}_0$). The false alarm probability with fixed $m_{\mathrm{N}}^{\mathrm{cut}}$ and $m_{\mathrm{M}}^{\mathrm{cut}}$ can be then expressed as

$$\begin{aligned} &p_{FA}(m_{\mathrm{N}}^{\mathrm{cut}}, m_{\mathrm{M}}^{\mathrm{cut}}) \\ &= Q\left(\frac{\lambda - N\left[(M_{\mathrm{N}} - m_{\mathrm{N}}^{\mathrm{cut}})\sigma_v^2 + (M_{\mathrm{M}} - m_{\mathrm{M}}^{\mathrm{cut}})(\sigma_h^2+\sigma_v^2)\right]}{\sqrt{N\left[(M_{\mathrm{N}} - m_{\mathrm{N}}^{\mathrm{cut}})\sigma_v^4 + (M_{\mathrm{M}} - m_{\mathrm{M}}^{\mathrm{cut}})(\sigma_h^2+\sigma_v^2)^2\right]}}\right), \end{aligned} \tag{9}$$

where $Q(x) = 1/\sqrt{2\pi}\sum_x^{\infty} e^{-t^2/2}dt$ is the Gaussian Q-function. Since $m_{\mathrm{N}}^{\mathrm{cut}}$ and $m_{\mathrm{M}}^{\mathrm{cut}}$ are random variables, P_{FA} can be obtained by averaging $p_{FA}(m_{\mathrm{N}}^{\mathrm{cut}}, m_{\mathrm{M}}^{\mathrm{cut}})$ over $m_{\mathrm{N}}^{\mathrm{cut}}$ and $m_{\mathrm{M}}^{\mathrm{cut}}$, as follows:

$$P_{FA} = \sum_{m_{\mathrm{N}}^{\mathrm{cut}}=0}^{M_{\mathrm{N}}} \sum_{m_{\mathrm{M}}^{\mathrm{cut}}=0}^{M_{\mathrm{M}}} p_{FA}(m_{\mathrm{N}}^{\mathrm{cut}}, m_{\mathrm{M}}^{\mathrm{cut}}) P(m_{\mathrm{N}}^{\mathrm{cut}}) P(m_{\mathrm{M}}^{\mathrm{cut}}). \tag{10}$$

P_D is the probability that the combined energy from the non-malicious and malicious users will exceed λ when a primary user exists. In this case, however, both the non-malicious and malicious users send high energy which consists of primary user signal and noise. The detection probability

with fixed $m_{\mathrm{N}}^{\mathrm{cut}}$ and $m_{\mathrm{M}}^{\mathrm{cut}}$ can be written as

$$p_D(m_{\mathrm{N}}^{\mathrm{cut}}, m_{\mathrm{M}}^{\mathrm{cut}}) = Q\left(\frac{\lambda - N\left[(M - m_{\mathrm{N}}^{\mathrm{cut}} - m_{\mathrm{M}}^{\mathrm{cut}})(\sigma_h^2 + \sigma_v^2)\right]}{\sqrt{N\left[(M - m_{\mathrm{N}}^{\mathrm{cut}} - m_{\mathrm{M}}^{\mathrm{cut}})(\sigma_h^2 + \sigma_v^2)^2\right]}}\right). \quad (11)$$

Then, P_D can be calculated by averaging both $m_{\mathrm{N}}^{\mathrm{cut}}$ and $m_{\mathrm{M}}^{\mathrm{cut}}$, as follows:

$$P_D = \sum_{m_{\mathrm{N}}^{\mathrm{cut}}=0}^{M_{\mathrm{N}}} \sum_{m_{\mathrm{M}}^{\mathrm{cut}}=0}^{M_{\mathrm{M}}} p_D(m_{\mathrm{N}}^{\mathrm{cut}}, m_{\mathrm{M}}^{\mathrm{cut}}) P(m_{\mathrm{N}}^{\mathrm{cut}}) P(m_{\mathrm{M}}^{\mathrm{cut}}). \quad (12)$$

The effect of the malicious user detection schemes on the sensing performance (i.e., false alarm and detection probabilities) can thus be evaluated in this way. Fig. 3 shows the false alarm probability (P_{FA}) as a function of the detection probability (P_D) for different numbers of malicious users M_{M}. We set the target cut-off probability for a malicious user at $\zeta = 0.9$. The corresponding cut-off probabilities for non-malicious users are obtained as $p_{\mathrm{N}}^{\mathrm{cut}} = 0.8166$ for the conventional scheme and $p_{\mathrm{N}}^{\mathrm{cut}} = 0.1019$ for the proposed scheme. This difference in malicious user detection performance directly affects the sensing performance. We can see that the false alarm probability of the proposed scheme is always lower than with the conventional scheme. Lower false alarm probability is known to guarantee greater spectrum usage opportunities for secondary users [8]. Note also that the false alarm probability increases as M_{M} increases. This is because the amount of reliable sensing data from non-malicious users decreases after the malicious users have been detected and cut off.

IV. Conclusion

This paper proposed a goodness-of-fit-based malicious user detection scheme for cooperative sensing. The proposed scheme, which is based on AD statistics, tests whether the empirical distribution of each secondary user fits the expected distribution of a malicious user. By cutting off the detected malicious users from having their sensing data combined with the rest at the fusion center, sensing performance degradation caused by malicious users can be greatly reduced. In this regard, we derived the false alarm and detection probability as functions of the cut-off probabilities for malicious and non-malicious users. Simulation results show that the proposed malicious user detection outperforms the conventional outlier-based approach.

References

[1] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access: Signal processing, networking, and regulation policy," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.

[2] K. B. Letaief and W. Zhang, "Cooperative communications for cognitive radio networks," *Proc. IEEE*, vol. 97, no. 5, pp. 878–893, 2009.

[3] Z. Quan, S. Cui, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 28–40, Feb. 2008.

[4] J. Ma, G. Zhao, and Y. Li, "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4502–4507, Nov. 2008.

[5] S. M. Mishra, A. Sahai, and R. W. Brodersen, "Cooperative sensing among CRs," in *Proc. IEEE Int. Conf. Commun. (ICC 2006)*, Jun. 2006, pp. 1658–1663.

[6] P. Kaligineedi, M. Khabbazian, and V. K. Bhargava, "Malicious user detection in a cognitive radio cooperative sensing system," *IEEE Trans. Wireless Commun.*, vol. 9, no. 8, pp. 2488–2497, Aug. 2010.

[7] K. Zeng, P. Pawelczak, and D. Cabric, "Reputation-based cooperative spectrum sensing with trusted nodes assistance," *IEEE Commun. Lett.*, vol. 14, no. 3, pp. 226–228, Mar. 2010.

[8] Y.-C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[9] H. Li, H. Dai, and C. Li, "Collaborative quickest spectrum sensing via random broadcast in cognitive radio systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 7, pp. 2338–2348, Jul. 2010.

[10] P. H. Kvam and B. Vidakovic, *Nonparametric Statistics with Applications to Science and Engineering*. Hoboken, NJ: John Wiley & Sons, 2007.

[11] T. W. Anderson and D. A. Darling, "Asymptotic theory of certain 'Goodness of Fit' criteria based on stochastic process," *Ann. Math. Stat.*, vol. 23, no. 2, pp. 193–212, 1952.

[12] M. A. Stephens, "EDF statistics for goodness of fit and some comparisons," *J. Amer. Statistical Assoc.*, vol. 69, no. 367, pp. 730–737, Sep. 1974.