Attack-Proof Collaborative Spectrum Sensing in Cognitive Radio Networks

Wenkai Wang‡, Husheng Li†, Yan (Lindsay) Sun‡, and Zhu Han*
‡ Department of Electrical, Computer and Biomedical Engineering
University of Rhode Island, Kingston, RI 02881
† Department of Electrical Engineering and Computer Science
University of Tennessee, Knoxville, TN 37996
* Department of Electrical and Computer Engineering
University of Houston, Houston, TX 77004

Abstract—Collaborative sensing in cognitive radio networks can significantly improve the probability of detecting the transmission of primary users. In current collaborative sensing schemes, all collaborative secondary users are assumed to be honest. As a consequence, the system is vulnerable to attacks in which malicious secondary users report false detection results. In this paper, we investigate how to improve the security of collaborative sensing. Particularly, we develop a malicious user detection algorithm that calculates the suspicious level of secondary users based on their past reports. Then, we calculate trust values as well as consistency values that are used to eliminate the malicious users’ influence on the primary user detection results. Through simulations, we show that even a single malicious user can significantly degrade the performance of collaborative sensing. The proposed trust value indicator can effectively differentiate honest and malicious secondary users. The receiver operating characteristic (ROC) curves for the primary user detection demonstrate the improvement in the security of collaborative sensing.

Keywords: Collaborative Sensing, Cognitive Radios, Security.

I. INTRODUCTION

Cognitive radio is a revolutionary paradigm providing high spectrum efficiency for wireless communications, in which transmission or reception parameters are dynamically changed to achieve efficient communication without introducing interference to traditionally licensed users (i.e., primary users). One major technical challenge is to detect the presence of the primary users’ transmission, or in other words, the spectrum holes. Popular detection techniques include matched filter, energy detection, cyclostationary detection and wavelet detection [1], [2]. It is recently discovered that collaboration among multiple cognitive radios can significantly improve the performance of spectrum sensing. The collaborative spectrum sensing techniques can be classified into two categories. The first category involves multiple users exchanging information [3]–[7], and the second category uses relay transmission [8], [9].

Most of the current schemes assume that secondary users always tell the truth. However, it is well known that wireless devices can be compromised and under the control of malicious parities. The malicious secondary user can send false information and mislead the spectrum sensing results to cause collision or inefficient spectrum usage. For example, some secondary users can always report the existence of the primary user such that they can occupy the spectrum themselves.

To overcome these problems, in this paper, we develop a malicious user detection algorithm that calculates the suspicious level of secondary users and then utilizes the suspicious level to eliminate the malicious users’ influence on the primary user detection results. Simulations are conducted to compare the performance of the collaborative sensing scheme without security protection, the scheme using a straightforward existing defense method, and the proposed scheme under various scenarios. We show that even a single malicious user can significantly degrade the performance of collaborative sensing.

The proposed detection scheme can effectively differentiate honest and malicious secondary users and greatly improve the security of collaborative spectrum sensing. For example, when there are 10 secondary users, with the primary user detection rate being 0.99, one malicious user can make the false alarm rate ($P_f$) increases to 59%. While an existing defense scheme can reduce $P_f$ to 16%, the proposed scheme reduces $P_f$ to 5%. Furthermore, when a good user suddenly turns bad, the proposed scheme can quickly reduce the trust value of this user. If this user only behaves badly for a few times, its trust value can recover after a large number of good behaviors. If the bad behavior is consistent, the trust value becomes almost impossible to recover.

This paper is organized as follows. In Section II, the system model is given. The attack models and the proposed scheme are described in Section III, followed by simulation results in Section IV and conclusion in Section V.

II. SYSTEM MODEL

Suppose that there are $N$ collaborative secondary users reporting their detected power of primary user transmission to a decision maker as shown in Figure 1. The objective is to detect whether the primary users are transmitting or not, based on the reports from the secondary users. For node $n$, its observation about the existence of the primary user at time slot $t$ is denoted by $X_n(t)$. The collection of $X_n(t)$, i.e. all

This work is supported by NSF Award # 0910461 and # 0831315.
observations in time slot $t$, is denoted by $X(t)$. These nodes report their observations to a centralized decision maker whose decision is broadcast to all secondary nodes.

When a secondary user is malicious, by changing its reporting values, this malicious user can significantly affect the decision and therefore degrade the performance of spectrum sensing.

The following assumptions and notations are used throughout this paper:

There is no more than one malicious node. We denote by $T_n$, the type of secondary node $n$ which could be $H$ (honest) or $M$ (malicious).

We denote the channel state by $S(t)$ which takes value in \{B, I\} (B means that primary users exist and the channel is busy; I means that there are no primary users and the channel is idle); we assume that the channel states in different time slots are mutually independent and denote by $q_B(t)$ and $q_I(t)$ the a priori probabilities of states $B$ and $I$, respectively.

We also denote by $p_B$ and $p_I$ the observation probabilities under busy and idle states, respectively, i.e.

\[
\begin{align*}
p_I(X_j(t)) &= P(X_j(t)|S(t) = I), \\
p_B(X_j(t)) &= P(X_j(t)|S(t) = B).
\end{align*}
\]

III. SECURE COLLABORATIVE SENSING

The performance of collaborative sensing can be severely damaged by malicious secondary users. To address this security vulnerability, we propose a defense scheme that contains two major components: (1) detection algorithm that determines suspicious level of secondary users; and (2) decision combining algorithm that considers the trustworthiness of secondary users. In this section, we describe the attack model, the detection algorithm, and the decision combining algorithm.

A. Attack Model

The behavior of an attacker (i.e. malicious secondary user) is described by the attack threshold ($\eta$), attack strength ($\Delta$), and attack probability ($P_a$). We consider two types of attacks.

- **False Alarm (FA) Attack:** With probability $P_a$, the attacker selects collaborative sensing rounds in which it aims to cause false alarm. In a selected round, when the sensed power $X_n(t)$ is higher than $\eta$, the attacker reports $X_n(t)$; otherwise, it reports $X_n(t) + \Delta$.

- **False Alarm & Miss Detection (FAMD) Attack:** With probability $P_a$, when the sensed power $X_n(t)$ is higher than $\eta$, the attacker reports $X_n(t) - \Delta$; otherwise, it reports $X_n(t) + \Delta$.

Note that the OR rule is commonly used in collaborative sensing because it has the best performance when there is no attackers. According to the OR rule, whenever one secondary user reports the existence of the primary user, the decision maker believes the existence of the primary user. With the FA attack, the malicious secondary user can easily cause high false alarm. When the malicious user is the only user who detects the primary user, it can also cause miss detection by not reporting the existence of the primary user. In the FAMD attack, the malicious user aims to cause both false alarm and miss detection.

B. Suspicious Level Calculation

In this section, we assume that there is only one malicious user. We define

\[
\pi_n(t) \triangleq P(T_n = M|\mathcal{F}_t)
\]

as the suspicious level of node $n$ at time $t$, where $T_n(= H \text{ or } M)$ is the type of node and $\mathcal{F}_t$ represents all observations from time slot 1 to time slot $t$. By applying Bayesian criterion, we have

\[
\pi_n(t) = \frac{P(\mathcal{F}_t|T_n = M)P(T_n = M)}{\sum_{j=1}^N P(\mathcal{F}_t|T_j = M)P(T_j = M)}.
\]

Suppose that $P(T_n = M) = \rho$ for all nodes. Then, we have

\[
\pi_n(t) = \frac{P(\mathcal{F}_t|T_n = M)}{\sum_{j=1}^N P(\mathcal{F}_t|T_j = M)}.
\]

It is easy to verify

\[
\begin{align*}
P(\mathcal{F}_t|T_n = M) &= \prod_{\tau=1}^t P(X(\tau)|T_n = M, \mathcal{F}_{\tau-1}) \\
&= \prod_{\tau=1}^t \left[ \prod_{j=1, j \neq n}^N P(X_j(\tau)|T_j = H) \right] P(X_n(\tau)|\mathcal{F}_{\tau-1}) \\
&= \prod_{\tau=1}^t \rho_n(\tau),
\end{align*}
\]

where

\[
\rho_n(t) = P(X_n(t)|\mathcal{F}_{\tau-1}) \prod_{j=1, j \neq n}^N P(X_j(t)|T_j = H),
\]

which means the probability of reports at time slot $t$ conditioned that node $n$ is malicious. Note that the first equation is obtained by repeatedly applying the following equation

\[
\begin{align*}
P(\mathcal{F}_t|T_n = M) &= P(X(t)|T_n = M, \mathcal{F}_{t-1})P(\mathcal{F}_{t-1}|T_n = M). (8)
\end{align*}
\]
Recall that \( q_B(t) \) and \( q_I(t) \) are the priori probability of whether or not the primary user exists. Therefore, the honest user’s report probability is given by
\[
P(X_j(t)|T_j = H) = P(X_j(t), S(t)_B|T_j = H) + P(X_j(t), S(t)_I|T_j = H)
\]
\[
= p_B(X_j(t))q_B(t) + p_I(X_j(t))q_I(t). \quad (9)
\]
Thus, the computation of \( \pi_n(t) \) is given by
\[
\pi_n(t) = \frac{\prod_{\tau=1}^{t} p_n(\tau)}{\sum_{j=1}^{N} \prod_{\tau=1}^{L} p_j(\tau)}. \quad (10)
\]

C. Primary User Detection

The suspicious level calculation provides a foundation for dealing with malicious secondary users. In this section, we demonstrate how to perform secure primary user detection.

We first convert suspicious level \( \pi_n(t) \) into trust value \( \phi_n(t) \) as
\[
\phi_n(t) = 1 - \pi_n(t). \quad (11)
\]

Trust value alone is not sufficient for determining whether a certain user’s report is reliable or not. In fact, we find that trust values become unstable when (1) there is no enough observation or (2) there is no malicious user. The first case is easy to understand. The reason for the second case is that the derivation in Section III-B assumes one and only one malicious user. When there is no attacker, the trust values of honest users become unstable. To solve this problem, we define consistency value of user \( n \) (i.e. \( \psi_n(t) \)) as
\[
\mu_n(t) = \begin{cases} 
\sum_{t=L}^{L+t} \frac{\phi_n(t)}{t}, & t < L \\
\sum_{\tau=1}^{L} \phi_n(t), & t \geq L,
\end{cases} \quad (12)
\]
\[
\psi_n(t) = \begin{cases} 
\sum_{t=L}^{L+t} (\phi_n(t) - \mu_n(t))^2, & t < L \\
\sum_{\tau=1}^{L} (\phi_n(t) - \mu_n(t))^2, & t \geq L,
\end{cases} \quad (13)
\]

where \( L \) is the size of the window in which the variation of recent trust values is compared with overall trust value variation.

Procedure 1 Primary user detection

1: receive reports from \( N \) secondary users.
2: calculate trust values and consistency values for all users.
3: for each user \( n \) do
4: \hspace{.5cm} if \( \phi_n(t) < \text{threshold}_1 \) and \( \psi_n(t) < \text{threshold}_2 \) then
5: \hspace{1cm} the report from user \( n \) is removed
6: \hspace{.5cm} end if
7: end for
8: perform primary user detection algorithm based on the remaining reports.

Next, the trust value \( \phi_n(t) \) and the consistency value \( \psi_n(t) \) are used in the primary user detection algorithm, as shown in Procedure 1. The basic idea is to eliminate the reports from users who have consistent low trust values. In this procedure, \( \text{threshold}_1 \) and \( \text{threshold}_2 \) can be chosen dynamically. For example, \( \text{threshold}_1 \) can be set as the a fraction of average trust value of all \( N \) users, and \( \text{threshold}_2 \) can be a fraction of average consistency values of all \( N \) users. This procedure can be used together with most existing primary user detection algorithms. In this paper, we will use the hard decision combining algorithm in [3], [11] to demonstrate the performance.

IV. Simulation Results

The simulations are conducted in a simple network topology with \( N = \) 10 cognitive secondary users randomly located around the primary user. The minimum distance between the secondary users and the primary user is 1000m and the maximum distance is 2000m. For the local spectrum sensing, the bandwidth-time product \( [3], [11] \) is \( m = 5 \). The transmission power of the primary user is 200mW. The noise level \( \sigma^2 \) is \(-110\text{dBm}\). The signal-to-noise ratio(SNR) of individual secondary user depends on its location and Rayleigh fading is assumed. The propagation loss factor is 3. In the attack models, the attack threshold \( \eta = 10 \), the attack strength \( \Delta = 10 \), and the attack probability \( P_a \) is 1 or 0.5. In the proposed scheme, the trust value threshold \( \text{threshold}_1 = 0.01 \), the consistency value threshold is \( \text{threshold}_2 = 0.1 \), and the window size for calculating consistency value is \( L = 10 \).

We compare three schemes.

- **OR Rule**: the primary user is detected when at least one secondary user’s reporting values are above a certain threshold.
- **K2 Rule**: the primary user is detected when at least two secondary user’s reporting value is above a certain threshold. This scheme, presented in [12], is a straightforward defense against one malicious user.
- **Proposed Scheme**: applying the OR rule after removing reports according to Procedure 1.

Figure 2 - 5 show the ROC curves for primary user detection in 6 cases. Case 1 is for OR rule with \( N \) honest users. Case 2 is for OR rule with \( N-1 \) honest users. In Case 3 - 6, there are \( N-1 \) honest users and 1 malicious user. Case 3 is for OR rule. Case 4 is for K2 rule. Case 5 is for the proposed scheme with \( t = 250 \), where \( t \) is the index of detection rounds. Case 6 is for the proposed scheme with \( t = 500 \).

When the attack strategy is the FA Attack, Figure 2 and Figure 3 show the ROC curves when the attack probability is 1 and 0.5, respectively. The following observations are made.

- By comparing the ROC for case 1 and case 3, we see that the performance of primary user detection degrades greatly even when there is only one malicious user. This demonstrates the vulnerability of collaborative sensing, which leads inefficient usage of available spectrum resource.
- The proposed scheme demonstrates significant performance gain over the scheme without defense (i.e. OR rule) and the straightforward defense scheme (i.e. K2 rule). For example, the following table shows the false alarm rate \( (P_a) \) for two given detection rate \( (P_d) \), when attack probability \( (P_a) \) is 1.

<table>
<thead>
<tr>
<th>When FA</th>
<th>OR Rule</th>
<th>K2 Rule</th>
<th>Proposed ((t = 250))</th>
<th>Proposed ((t = 500))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_a = 1 )</td>
<td>( 0.99 )</td>
<td>( 0.53 )</td>
<td>( 0.10 )</td>
<td>( 0.06 )</td>
</tr>
<tr>
<td>( P_d = 0.995 )</td>
<td>( 0.62 )</td>
<td>( 0.18 )</td>
<td>( 0.07 )</td>
<td>( 0.04 )</td>
</tr>
</tbody>
</table>
When the attack probability is 0.5, the performance advantage is smaller but still large.

When FA
\[ P_a = 0.5 \]
\[
\begin{array}{c|c|c|c|c}
\text{OR Rule} & \text{K2 Rule} & \text{Proposed (t = 250)} & \text{Proposed (t = 500)} \\
\hline
0.27 & 0.06 & 0.04 & 0.03 \\
0.32 & 0.11 & 0.07 & 0.05 \\
\end{array}
\]

- In addition, as \( t \) increases, the performance of the proposed scheme gets close to the performance of case 2, which represents perfect detection of the malicious nodes.

Figure 4 and Figure 5 shows the ROC performance when the malicious user adopts the FAMD attack. We observe that the FAMD attack is stronger than FA. In other words, the OR rule and K2 rule have worse performance when facing the FAMD attack. However, the performance of the proposed scheme is almost the same under both attacks. That is, the proposed scheme is highly effective under both attacks, and much better than the traditional OR rule and the simple defense K2 rule. The example false alarm rates are listed as follows.

When FAMD
\[ P_a = 1 \]
\[
\begin{array}{c|c|c|c|c}
\text{OR Rule} & \text{K2 Rule} & \text{Proposed (t = 250)} & \text{Proposed (t = 500)} \\
\hline
0.39 & 0.16 & 0.07 & 0.05 \\
0.69 & 0.29 & 0.09 & 0.06 \\
\end{array}
\]

When FAMD
\[ P_a = 0.5 \]
\[
\begin{array}{c|c|c|c|c}
\text{OR Rule} & \text{K2 Rule} & \text{Proposed (t = 250)} & \text{Proposed (t = 500)} \\
\hline
0.29 & 0.09 & 0.05 & 0.04 \\
0.34 & 0.14 & 0.07 & 0.05 \\
\end{array}
\]

Finally, trust values of honest users and malicious users are
Examined. The malicious user adopts the FAMD attack and dynamically adjusts the attack probability $P_a$. In Figure 6, the malicious user changes the attack probability from 0 to 1 at $t = 50$ and from 1 to 0 at time $t = 90$. We can see that this user’s trust value quickly drops when it turns from good to bad. However, its trust value does not recover after it turns good unless this user consistently behaves well for more than at least 5000 rounds (the recovery is not shown). In Figure 7, one user behaves badly in only 5 rounds starting at $t = 50$. In this case, its trust value drops quickly and then recovers very slowly. This means that the proposed scheme allows slow recovery of trust value after a few bad behaviors, which may due to channel variation and unintentional errors.

V. CONCLUSIONS

To overcome the malicious behavior in collaborative sensing, in this paper, we propose a defense scheme that computes suspicious level, trust values and consistency values. These values are used to guide the decision combining, in which the reports from untrustworthy users are eliminated. In its current form, the proposed scheme can effectively defense against one malicious user. It can be extended to multiple malicious user case in the future. Through simulations, the ROC curves and trust value dynamics are studied for different attack models, attack probabilities and different collaborative sensing schemes. The proposed schemes demonstrate significant performance advantage. For example, when there are 10 secondary users, with the primary user detection rate equals to 0.99, one malicious user can make the false alarm rate ($P_f$) increases to 59%. While a simple defense scheme can reduce $P_f$ to 16%, the proposed scheme reduces $P_f$ to 5%. Furthermore, when a good user suddenly turns bad, the proposed scheme can quickly reduce the trust value of this user. If this user only behaves badly for a few times, its trust value can recover after a large number of good behaviors. If the bad behavior is consistent for sometime, the trust value is almost impossible to recover unless the malicious user behave well consistently for extremely long time. The suspiciousness indicator developed in this paper is novel and effectively captures misbehavior dynamically.

REFERENCES