Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/100635

A Fast Malicious User Detection Scheme Based on POMDP for Cooperative Spectrum Sensing in Cognitive Radio networks

Hiep Vu-Van and Insoo Koo[*]

The School of Electrical Engineering, University of Ulsan, 680-749 Ulsan, Republic of Korea

Received: 4 Feb. 2016, Revised: 3 Aug. 2016, Accepted: 5 Aug. 2016
Published online: 1 Nov. 2016

Abstract: Cooperative spectrum sensing (CSS) can improve spectrum sensing accuracy, but it can be injured due to potential attacks from malicious cognitive radio user who reports false sensing results to the fusion center (FC). Many researchers focus on reducing the effect of malicious users on the accuracy of spectrum sensing. A promising method to detect malicious users is to determine their abnormal spectrum sensing behavior. In this paper, we provide a novel malicious users detection scheme for cognitive radio (CR) based on the truth rate of each CU, which is defined as the correlation level between the Markov property of the CU's reported sensing information and the states of the PU signal. The truth rate may distinguish an honest user from a malicious user by giving an honest CU a high trust rate and giving a malicious user a low one. In the malicious user detection process, a partially observable Markov decision process (POMDP) is applied to consider the effect of the current action (that action is to classify a CU as an honest or a malicious user) on the reward in future time slot (that reward is achieved by classifying a CU as an honest or a malicious user). By taking advantage of POMDP, the proposed scheme may detect the presence of malicious users in a shorter required time.

Keywords: cognitive radio, Makov property, POMDP, fast malicious user detection, robust cooperative spectrum sensing

1 Introduction

Cognitive radio (CR) [1,2] is a promising technique to improve spectrum utilization. In a CR network, cognitive radio users (CUs) can exploit the unused spectrum that is assigned to the license user (called the primary user (PU)). To avoid interference with the PU, the CU is allowed to access to the frequency only when it is free, and when the presence of the PU is detected, the CU must vacate the occupied frequency. Reliably sensing the PU's signal is a requirement of CR network implementation.

Improved sensing performance can be obtained by allowing some CUs to perform cooperative spectrum sensing (CSS) [3,4,5]. However, CSS is sensitive to attacks by malicious users who send false sensing data to the fusion center (FC) [6,7]. The research presented in [6,7] determined that the presence of a few malicious users can severely reduce the performance of a CSS scheme. Algorithms used to identify the malicious users have been proposed in the studies of [6,7]. In [6,7], a simple technique (i.e., outlier-detection) is used to detect malicious CUs, and so it only considers for low damage type malicious CUs such as *Attack* or *Selfish* CU. In addition, the technique is unable to protect the CSS in the event of a large number of malicious users in the network. Studies in [8,9,10] apply an event detection technique to detect malicious users and protect CSS. A hidden Markov model (HMM) is utilized to defend malicious users in [11]. Almost of those malicious user detection schemes do not consider the Markov property of spectrum states for an improved robust CSS. In addition, when the number of malicious users is much higher than the number of honest CUs, it is difficult to maintain high reliable cooperative spectrum sensing in those schemes.

In general, the spectrum states are correlated and are often modeled as Markov states. In this paper, we proposed a novel robust CSS that takes advantage of the Markov property of spectrum states to detect abnormal behavior of the CU. In the proposed scheme, a truth rate is defined as the correlation level between the Markov property of the CU's reported sensing information and the states of the PU signal. The malicious users report false sensing information to the FC, so that the correlation

[*] Corresponding author e-mail: iskoo@ulsan.ac.kr

between the Markov property of their reported sensing information and of the PU state is low. In contrast, the honest CUs send the correct sensing information to the FC, and so the correlation of the honest CU will be at a high level. Subsequently, the trust rate can distinguish an honest CU from a malicious user by giving a honest CU a high trust rate and giving the malicious user a low trust rate. In the malicious users detection process, a partially observable Markov decision process (POMDP) is applied to consider the effect of the current action (that action is to classify a CU as an honest or a malicious user) on the reward in future time slots (that reward is achieved by classifying a CU to be an honest or a malicious user). By applying POMDP, the proposed scheme may detect the presence of malicious users in a shorter required time. Malicious users are classified into three types: "*selfish*", "*attack*" and "*adversary*" users. How harmful each malicious user is depends on its "*malicious rate*", which is defined as the probability that the CU acts like a malicious user. Simulation results show the effectiveness of the proposed scheme.

2 System Model

In this paper, we consider a CR network including N CUs that cooperate to sense the PU signal by using their energy detectors. Sensing results of the CUs are reported to the fusion center (FC) in order to make a global decision about the PU status. To quantify the sensing performance of the CUs, the probability of detection P_d and the probability of false alarm P_f are utilized. The CUs can be classified as an honest CU or a malicious user according to their report behavior.

2.1 Honest users

An honest CU works under the control of the fusion center (FC) for the common benefit of the CR network. It always reports real sensing information to the FC. Let denote B and R^h as the sensing information of the honest user and the information that the honest user reports to the FC, respectively. Subsequently, we have $R^h = B$, where $R^h, B \in \{0,1\}$, $R^h = 1$ and $B = 1$ indicate that the honest user has detected the presence of the PU signal, otherwise, $R^h = 0$ and $B = 0$ indicate that the honest user has not detected any signal from the PU. The FC can determine the sensing performance of the honest CU as $P_d^h = P_d$ and $P_f^h = P_f$, where P_d^h and P_f^h are the estimated probability of detection and the estimated probability of false alarm according to the reported sensing information that the honest CU reports to the FC, respectively, while P_d and P_f are its real sensing performance.

2.2 Malicious users

On the other hand, a malicious user may tamper with its local decision before reporting to the FC. Let's define a_{10} and a_{01} as the "*malicious rate*" of the malicious user, where a_{10} is the probability that the malicious user flips its local decision from "1" (the PU signal is present) to "0" (the PU signal is absent), and a_{01} is the probability that the malicious user flips its local decision from "0" to "1". Accordingly, at the FC, the sensing performance (i.e., P_d^m and P_f^m) of the malicious user is given by

$$P_d^m = (1 - a_{10}) P_d + a_{01} (1 - P_d) \tag{1}$$

and

$$P_f^m = (1 - a_{10}) P_f + a_{01} \left(1 - P_f\right). \tag{2}$$

In the case that $a_{10} = 0$ and $a_{01} = 0$, the malicious user works identically to the honest user, which means that $P_d^m = P_d^h = P_d$ and $P_f^m = P_f^h = P_f$. Therefore in this paper, in order to differ between honest and malicious users, the user is only considered to be a malicious user when at least one of a_{10} or a_{01} is nonzero. According to the values of a_{10} and a_{01}, we classify the malicious users as three types: a "*Selfish* " user (SeU) when $a_{10} = 0$ and $a_{01} > 0$; an "*Attack*" user (AtU) when $a_{10} > 0$ and $a_{01} = 0$; and an "*Adversary*" user (AdU) when $a_{10} > 0$ and $a_{01} > 0$. A "*Selfish* " user cheats the FC by reporting "1", even when it does not detect the PU signal that leads the FC to believe that the PU is active. Then, the FC will not allow others CUs in the network to access the channel. A "*Selfish*" user can exclusively use that channel. On the other hand, an "*Attack* " user tries to disrupt the considered channel by reporting "0" to the FC, even when it detects the PU signal. The "*Attack* " user makes the FC think that the PU signal is idle and allows the CUs to access the channel. Subsequently, the collision occurs when the PU is actually active. An "*Adversary*" user is the most harmful user because of its *flipping* behavior, in which it inverts the sensing results before reporting to the FC. "*Adversary*" users may selfishly use the channel when the PU signal is not detected, and may destroy the channel (i.e., increase the collision probability) when the active PU is detected.

2.3 Markov Property of The Channel States

In this paper, we assume that the PU works in a time slotted manner. The spectrum state is defined as $S \in \{1\,(\text{presence}), 0\,(\text{absence})\}$, following the Markov property as shown in Fig. 1, where b_{xy} is the transition probability that the PU signal changes from state $x \in \{0,1\}$ to state $y \in \{0,1\}$ in two continuous time slots; b_{xy} is given as

$$b_{xy} = \Pr\left\{S_{xy}\right\}, \tag{3}$$

where S_{xy} is the state of the PU in two continuous time slots, where the state "x" is in the first slot and "y" is in the next. That is, $S_{xy} = \{S(t), S(t+1) | S(t) = x, S(t+1) = y\}$, where t is the time index. We also define the state probability of the PU signal as

$$p_a = \Pr\{S_a\}, \tag{4}$$

where $S_a = \{S | S = a\}$ and $a \in \{0,1\}$. When $a = 0$, p_0 is defined as the absent probability of the PU signal and when $a = 1$, p_1 is defined as the present probability of the PU signal.

According to the transition probability and the state probability, we defines *behaviors* of the spectrum as follows:

$$BS = \{TP, SP\}, \tag{5}$$

where TP and SP are given as,

$$TP = \{b_{xy} | \forall x, y \in \{0,1\}\} \tag{6}$$

and

$$SP = \{p_a | \forall a \in \{0,1\}\}. \tag{7}$$

In the CR network, a group of CUs is assigned to perform spectrum sensing to detect the state of the considered channel. Then, they report the sensing results to the FC. The FC can estimate the spectrum *behavior* (eSB) by using the sensing information received from each CU:

$$\widehat{BS}^j = \left\{\widehat{TP}^j, \widehat{SP}^j\right\}, \tag{8}$$

where j is the CU index, $\widehat{TP}^j = \left\{\widehat{b}_{xy}^j | \forall x, y \in \{0,1\}\right\}$ and $\widehat{SP}^j = \left\{\widehat{p}_a^j | \forall a \in \{0,1\}\right\}$.

$$\widehat{b}_{xy}^j = \sum_{ab} \Pr\{S_{ab}\} \Pr\{R_{xy}^j | S_{ab}\}, \tag{9}$$

where $b \in \{0,1\}$ and R_{xy}^j is the report of the j^{th} CU in two continuous time slots, with the reports "x" and "y" in the first and second slots, respectively. That is, $R_{xy}^j = \left\{R^j(t), R^j(t+1) \middle| R^j(t) = x, R^j(t+1) = y\right\}$.

$$\widehat{p_a}^j = \sum_{x} \Pr\{S_x\} \Pr\{R_a^j | S_x\}, \tag{10}$$

where $R_a^j = \left\{R^j \middle| R^j = a\right\}$.

It can be seen that eSB depends on the original behavior of the spectrum and sensing performance of the CU. However, all CUs monitor the same channel (i.e., the same original behavior), and the difference between eSB is caused only by the sensing performance. Since the malicious user reports fake sensing results to the FC, eBS of the malicious user and the honest CU will be largely different.

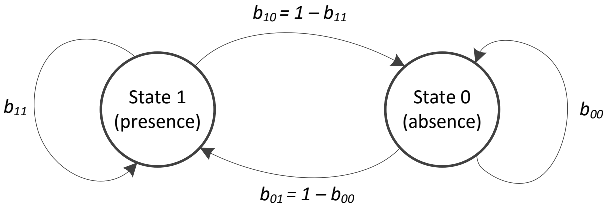

Fig. 1: Markov chain states of the PU

3 The Proposed Fast-Robust CSS based on POMDP

In this section, we proposed a fast-robust CSS scheme for a CR network. In the proposed scheme, the FC monitors the sensing results received from the CUs to determine their abnormal behavior. A CU that has abnormal behavior will be considered as a malicious user and its sensing information will not be used for making a global decision. The POMDP will be applied to consider the effect of the current action (that action is to classify a CU as an honest or a malicious user) on the reward in a future time slot (that reward is achieved by classifying a CU to be an honest or a malicious user). The problem of finding out which CU is a malicious user will be formulated within the framework of POMDP. The definition of POMDP spaces are described as follows:

3.1 State space

Because of the difference between the eBS of a malicious user and eBS of a honest CU, we use the eBS as the information to detect malicious users. Subsequently, we define the state space of POMDP of the j^{th} CU as

$$\Gamma^j = \{\varepsilon_{xy}^j, \eta_a^j, \varepsilon_{xy}^{-j}, \eta_a^{-j} | \forall x, y, a \in \{0,1\}\}, \tag{11}$$

where

$$\varepsilon_{xy}^j = \frac{\hat{b}_{xy}^j}{\hat{b}_{xy}^j + b_{xy}}, \quad \eta_a^j = \frac{\hat{p}_a^j}{\hat{p}_a^j + p_a} \tag{12}$$

and

$$\varepsilon_{xy}^{-j} = \frac{b_{xy}}{\hat{b}_{xy}^j + b_{xy}}, \quad \eta_a^{-j} = \frac{p_a}{\hat{p}_a^j + p_a}, \tag{13}$$

and $\widehat{b}_{xy}^j$ and $\widehat{p}_a^j$ are determined by using the sensing information collected from the j^{th} CU for the considered window size D. In addition, b_{xy} and p_a are the real statistic of the PU states which can be determined as in Eqn. (3) and (4), respectively.

Using the definition of the state space, it can be seen that Γ^j includes $n_{\Gamma^j} = 12$ elements. In the case of a honest CU, its performance has to satisfy reliability requirements in terms of $Pf_j \leq Pf^{thr}$ and $Pd_j \geq Pd^{thr}$, where Pf^{thr} and Pd^{thr} are requirements of false alarm and detection probability, respectively. Further, their $\widehat{b}_{xy}^{j}$ and $\widehat{p}_{a}^{j}$ are closed to b_{xy} and p_a. Subsequently, all of the elements will be close together and converge to the same value (i.e., $\frac{1}{2}$). On the contrary, a malicious user makes all of the elements different and converging to various values.

3.2 Action space

For robust cooperative spectrum sensing, the FC needs to classify which CU is a malicious user and which CU is an honest CU. Subsequently, the FC considers two actions for each CU $A_j \in \{0(\text{Reject}), 1(\text{Accept})\}$, where $A_j = 0(\text{Reject})$ indicates that the j^{th} CU is a malicious user and its sensing information will not be used for cooperative spectrum sensing, and $A_j = 1(\text{Accept})$ indicates that the j^{th} CU is an honest CU and will be polled for sensing information.

3.3 Value function

We define the *trust rate* of the j^{th} CU at time slot t (i.e., the current time slot) as

$$TR^{j}(t) = \frac{\left(\sum_{i=1}^{n_{\Gamma^j}} \Gamma_i^{j}(t)\right)^2}{n_{\Gamma^j} \sum_{i=1}^{n_{\Gamma^j}} \left(\Gamma_i^{j}(t),\right)^2} \tag{14}$$

where $\Gamma_i^{j}(t)$ is the i^{th} element of state space of the j^{th} CU at time slot t.

Since all of the elements of state space of an honest CU converged to the same value, the *trust rate* of an honest CU will converge to nearly 1, which is the maximum value of the *trust rate*. On the other hand, a malicious user has various elements of the state space then its *trust rate* will be very small compared to that of an honest CU. We define the *malicious threshold* as M^{th}, which can be selected in the range $\{0,1\}$ based on the experiment of the network. If the CU has a *trust rate* that is smaller than the threshold, it may be considered as a malicious user. Subsequently, we define the *reward* for each CU as,

$$RW^{j}\left(\Gamma^{j}, A^{j}\right) = \begin{cases} 0, & \text{if } A^{j} = 0 \\ TR^{j} - M^{th}, & \text{otherwise} \end{cases}. \tag{15}$$

The value function $\Phi(\Gamma^{j}(t))$ is defined as the maximum total discounted reward from the current time slot when the current state of the CU is $\Gamma^{j}(t)$. The value function is given as:

$$\Phi\left(\Gamma^{j}(t)\right) = \max_{A^{j}(t)\in\{0,1\}} \mathrm{E}\left\{\sum_{k=t}^{\infty} \rho^{k-t} RW^{j}\left(\Gamma^{j}(k), A^{j}(k)\right) \left|\Gamma^{j}(t)\right.\right\}, \tag{16}$$

where $0 \le \rho < 1$ is the discount factor, and $\Gamma^{j}(t)$ and $A^{j}(t)$ are the state and action of the current time slot (i.e., the t^{th} time slot), respectively. The reward value $RW^{j}\left(\Gamma^{j}(t), A^{j}(t)\right)$ depends on the current state and the chosen action mode.

3.3.1 Reject mode ($A^{j} = 0$)

For this action mode, the *reward* will be

$$RW^{j}\left(\Gamma^{j}(t), 0\right) = 0,\ \forall \Gamma^{j}(t). \tag{17}$$

We define $four$ observations for this action mode according to report of the CU in the previous and current time slots as:

Observation 1 (Θ_1): Reports from the j^{th} CU are "absence" in both the t (i.e., the current time slot) and $(t-1)$ time slots ($R^{j}(t-1) = 0$ and $R^{j}(t) = 0$).

The probability that this observation happens is

$$\Pr(\Theta_1) = \frac{\hat{p}_0^{j}(t)\widehat{b}_{00}^{j}(t)}{\hat{p}_0^{j}(t)\widehat{b}_{00}^{j}(t) + \hat{p}_1^{j}(t)\widehat{b}_{10}^{j}(t)} \sum_{a\in\{0,1\}} p_a \Pr\left(R_o^{j}(t) \left| S_a\right.\right). \tag{18}$$

eSB will be updated as follows:

$$\begin{array}{l} \widehat{p}_0^{j}(t) = \widehat{p}_0^{j}(t-1)\frac{D-1}{D} + \frac{1}{D}, \\ \widehat{p}_1^{j}(t) = \widehat{p}_1^{j}(t-1)\frac{D-1}{D}, \\ \widehat{b}_{00}^{j}(t) = \widehat{b}_{00}^{j}(t-1)\frac{D-1}{D} + \frac{1}{D}, \\ \widehat{b}_{01}^{j}(t) = \widehat{b}_{01}^{j}(t-1)\frac{D-1}{D}, \\ \widehat{b}_{1x}^{j}(t) = \widehat{b}_{1x}^{j}(t-1), \forall x \in \{0,1\}. \end{array} \tag{19}$$

Observation 2 (Θ_2): Reports from the j^{th} CU are "presence" in both the t and $(t-1)$ time slots ($R^{j}(t-1) = 1$ and $R^{j}(t) = 1$).

The probability that this observation happens is

$$\Pr(\Theta_2) = \frac{\hat{p}_1^{j}(t)\widehat{b}_{11}^{j}(t)}{\hat{p}_1^{j}(t)\widehat{b}_{11}^{j}(t) + \hat{p}_0^{j}(t)\widehat{b}_{01}^{j}(t)} \sum_{a\in\{0,1\}} p_a \Pr\left(R_1^{j}(t) \left| S_a\right.\right). \tag{20}$$

eSB will be updated as follows:

$$\begin{array}{l} \hat{p}_0^{j}(t) = \hat{p}_0^{j}(t-1)\frac{D-1}{D}, \\ \hat{p}_1^{j}(t) = \hat{p}_1^{j}(t-1)\frac{D-1}{D} + \frac{1}{D}, \\ \hat{b}_{11}^{j}(t) = \hat{b}_{11}^{j}(t-1)\frac{D-1}{D} + \frac{1}{D}, \\ \hat{b}_{10}^{j}(t) = \hat{b}_{10}^{j}(t-1)\frac{D-1}{D}, \\ \hat{b}_{0x}^{j}(t) = \hat{b}_{0x}^{j}(t-1), \forall x \in \{0,1\}. \end{array} \tag{21}$$

Observation 3 (Θ_3): Reports from the j^{th} CU are "presence" in the t time slot and "absence" in the $(t-1)$ time slot ($R^{j}(t-1) = 0$ and $R^{j}(t) = 1$).

The probability that this observation happens is

$$\Pr(\Theta_3) = \frac{\hat{p}_0^{j}(t)\widehat{b}_{01}^{j}(t)}{\hat{p}_0^{j}(t)\widehat{b}_{01}^{j}(t) + \hat{p}_1^{j}(t)\widehat{b}_{11}^{j}(t)} \sum_{a\in\{0,1\}} p_a \Pr\left(R_1^{j}(t) \left| S_a\right.\right). \tag{22}$$

eSB will be updated as follows:

$$\begin{array}{l} \hat{p}_0^j(t) = \hat{p}_0^j(t-1)\frac{D-1}{D}, \\ \hat{p}_1^j(t) = \hat{p}_1^j(t-1)\frac{D-1}{D} + \frac{1}{D}, \\ \hat{b}_{01}^j(t) = \hat{b}_{01}^j(t-1)\frac{D-1}{D} + \frac{1}{D}, \\ \hat{b}_{00}^j(t) = \hat{b}_{00}^j(t-1)\frac{D-1}{D}, \\ \hat{b}_{1x}^j(t) = \hat{b}_{1x}^j(t-1), \forall x \in \{0,1\}. \end{array} \tag{23}$$

Observation 4 (Θ_4): Reports from the j^{th} CU are "absence" in the t time slot and "presence" in the $(t-1)$ time slot ($R^j(t-1) = 1$ and $R^j(t) = 0$).

The probability that this observation happens is

$$\Pr(\Theta_4) = \frac{\hat{p}_1^j(t)\widehat{b}_{10}^j(t)}{\hat{p}_1^j(t)\widehat{b}_{10}^j(t) + \hat{p}_0^j(t)\widehat{b}_{00}^j(t)} \sum_{a\in\{0,1\}} p_a \Pr\left(R_0^j(t)\,|S_a\right). \tag{24}$$

eSB will be updated as follows:

$$\begin{array}{l} \hat{p}_0^j(t) = \hat{p}_0^j(t-1)\frac{D-1}{D} + \frac{1}{D}, \\ \hat{p}_1^j(t) = \hat{p}_1^j(t-1)\frac{D-1}{D}, \\ \hat{b}_{10}^j(t) = \hat{b}_{10}^j(t-1)\frac{D-1}{D} + \frac{1}{D}, \\ \hat{b}_{11}^j(t) = \hat{b}_{11}^j(t-1)\frac{D-1}{D}, \\ \hat{b}_{0x}^j(t) = \hat{b}_{0x}^j(t-1), \forall x \in \{0,1\}. \end{array} \tag{25}$$

3.3.2 Accept mode ($A^j = 1$)

For this action mode, the *reward* will be

$$RW^j\left(\Gamma^j(t), 1\right) = TR^j(t) - M^{th}. \tag{26}$$

In accept mode, we also consider $four$ observations which are the same as the $four$ observations in reject mode. In each observation, eSB is also updated in the same way.

According to the updated eSB, the state space will be updated as in Eqn.(12) and (13).

Based on those observations, the value function in Eqn.(16) will be rewritten as in Eqn.(27). In order to find an optimal mode policy (which CUs are malicious users), the optimization problem in Eqn.(27) will be solved by using the value iterations method [12].

4 Implementation of the Proposed Robust CSS based on POMDP

In this paper, we propose a robust CSS that detects and rejects harmful effects from malicious users. POMDP is applied to make the scheme detect malicious users in a shorter required time. The implementation of the proposed scheme can be described by the flow chart in Fig. 2. First, the FC collects and stores the sensing information reported from all of the CUs. For the POMDP, the state space will be determined by using the information from the D past time slots (i.e., $(t-D+1)$ to t, where t is the current time slot). For implementation, this state space can also be updated at each time slot according to the observations as maintained above in the subsection *Value function*. Second, POMDP will be run in order to to determine whether or not the considered CU is a malicious user. If the CU is concluded to be a malicious user, its reported sensing information will be not used to make a global decision in the current time slot, and it must wait for the next time slot. If the result is an honest user, its reported sensing information will be used to a make a global decision.

Sensing information of the honest CUs will be combined to make a global decision by using log-likelihood combination rule, given by

$$\begin{cases} Gb(t) = 1, \text{if} \Delta(t) \geq 0 \\ Gb(t) = 1, \text{otherwise} \end{cases}, \tag{28}$$

where

$$\Delta(t) = \sum_{k\in\Psi_1} \log\frac{p_d^k}{p_d^k} + \sum_{l\in\Psi_0} \log\frac{1-p_d^l}{1-p_d^l}. \tag{29}$$

where Ψ_1 and Ψ_0 are the sets of honest CUs who report the local decisions "1" and "0" to the FC, respectively.

In order to determine the state space of a CU, we also need information about the PU signal, which are the transition probability $b_{xy}, \forall x,y \in \{0,1\}$ and the state probability $p_a, \forall a \in \{0,1\}$. According to the availability of the PU signal information, the state space is determined in different ways.

4.1 Information of the PU signal is available

In practice, the FC has difficulty knowing the exact statistics of the PU signal, meaning that b_{xy} and p_a are often not available at the FC. However, the FC can reliably estimate this information by long time statistic of sensing process. We assume the FC can perfectly know $b_{xy}, \forall x,y \in \{0,1\}$ and $p_a, \forall a \in \{0,1\}$ of the PU signal. This information will be used to update the state space of the CU as in Eqn.(12) and (13), in each time slot.

4.2 Information of the PU signal is unavailable

In this case, we do not have enough information to update the state space of the CU. Therefore, we must estimate b_{xy} and p_a to run the malicious user detection scheme. We utilize two methods to estimate this information, "channel feedback information" and "trust node assistant". These methods will estimate the status of the PU signal ($ePS(t) \in \{0,1\}$) for each time slot and from which $\hat{b}_{xy}$ and $\hat{p}_a$ can be determined by using $ePS(t)$.

- *Estimation method based on channel feedback information*

$$\Phi\left(\Gamma^{j}(t)\right)=\max_{A^{j}(t)\in\{0,1\}}\left\{\sum_{k=t}^{\infty}\rho^{k-t}\sum_{\Theta_i\in A^{j}(k)}\Pr(\Theta_i)RW^{j}\left(\Gamma^{j}(k),A^{j}(k)\right)\Big|\Gamma^{j}(t)\right\} \tag{27}$$

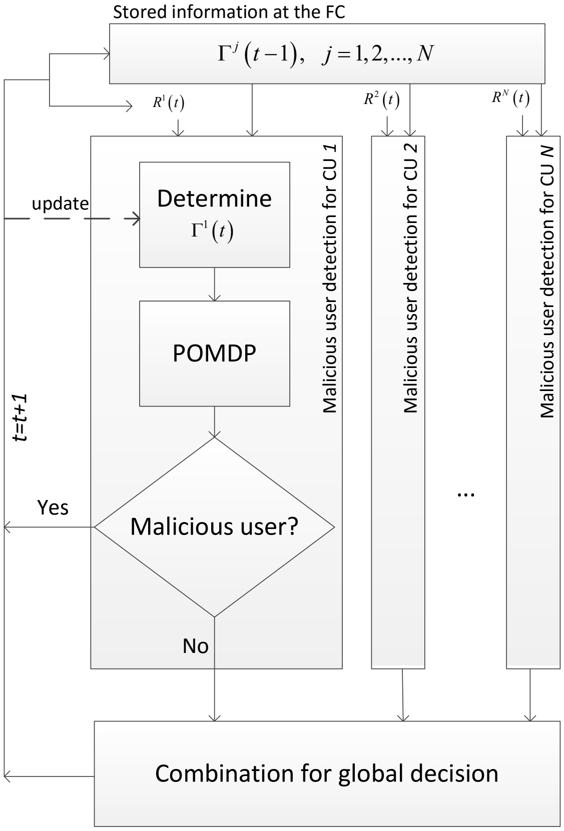
Stored information at the FC
$\Gamma^{j}(t-1),\quad j=1,2,...,N$
$R^{1}(t)$
$R^{2}(t)$
$R^{N}(t)$
update
Determine
$\Gamma^{1}(t)$
POMDP
Malicious user?
Yes
No
t=t+1
Malicious user detection for CU 1
Malicious user detection for CU 2
Malicious user detection for CU N
...
Combination for global decision

Fig. 2: Flow-chart of the proposed robust CSS based on POMDP

When FC recognizes that the PU signal is absent, it is allowed to use the free channel. Based on the feedback information indicating whether the transmission is a success or failure, the FC may know the real status of the PU signal. If the transmission is a success, the sensing process is correct and the PU signal is actually absent (i.e., $ePS(t)=S_0(t)$). Otherwise, the transmission is a failure, the sensing process has a miss detection event and the PU signal is present (i.e., $ePS(t)=S_1(t)$). Since the feedback information in the data channel is highly reliable, this method can provide high reliable statistic information of the PU signal. However, this method can be applied only when the CR network transmits in the considered channel (i.e., when FC recognizes that the PU signal is absent). Therefore, we use the trust node assistant method to estimate statistic information of the PU signal when the CR network is not allowed to use the channel.

- Estimation method based on trust node assistant

In this method, we base on reported sensing information of some trust devices in order to estimate the PU signal information. In the CR network, some devices cannot be malicious users, for example, the FC or base station (BS), who are often equipped with the full ability of spectrum sensing. In this paper, we assume that the FC also performs spectrum sensing as an honest CU, and sensing results of the FC are called $B^{FC}\in\{0,1\}$. In the time slot, when the feedback information in the data channel is not available (i.e., the CU does not transmit), the sensing result from the FC will be used as an estimated status of the PU signal, that is, $ePS(t)=B^{FC}(t)$.

Estimated values of $\hat{b}_{xy}$ and $\hat{p}_a$ can be updated by using ePS(t) as

$$\hat{b}_{xy}(t)=\hat{b}_{xy}(t-1)\tfrac{t-1}{t}+\tfrac{1}{t}(ePS(t-1)=x)(ePS(t)=y)$$
$$\forall x,y\in\{0,1\} \tag{30}$$

and

$$\widehat{p}_a(t)=\widehat{p}_a(t-1)\tfrac{t-1}{t}+\tfrac{1}{t}(ePS(t-1)=a)$$
$$\forall a\in\{0,1\}, \tag{31}$$

where t is the index of the time slot, $(A=z)$ is a logic function given by $(A=z)=1$ if $A=z$ and $(A=z)=0$ if $A\neq z$.

Estimated values of $\hat{b}_{xy}$ and $\hat{p}_a$ will be used to update the state space of the CU as in Eqn.(12) and (13), at each time slot.

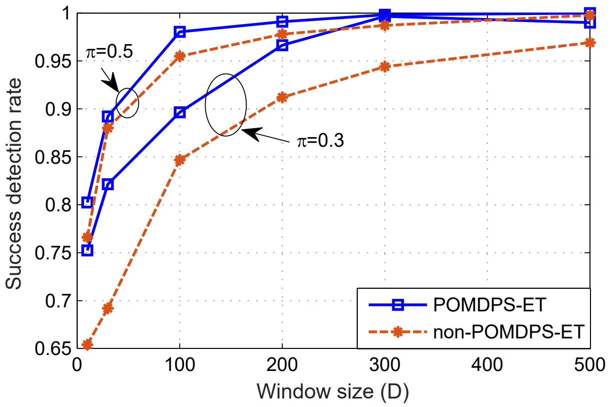

Fig. 3: Success detection rates of the proposed schemes versus window size D when $\pi=0.3$ and $\pi=0.5$.

5 Simulation Results

In order to show the effectiveness of the proposed scheme, we provide the simulation results of some schemes as follows:

–The proposed scheme based on POMDP; information of the PU signal is available at the FC (called POMDPS-IA).
–The proposed scheme based on POMDP; information of the PU signal is not available at the FC (called POMDPS-ET).
–The proposed scheme does not apply POMDP; information of the PU signal is available at the FC (called non-POMDPS-IA).
–The proposed scheme does not apply POMDP; information of the PU signal is not available at the FC (called non-POMDPS-ET),
–The perfect malicious detection scheme (called PDS).
–The conventional scheme, which does not have any malicious user detection and combines all of the received sensing information (including information from malicious users) to make a global decision (called non-MDS).

Here "the proposed scheme does not apply POMDP" (non-POMDPS) detects malicious users by using only the trust rate information, which is defined in Eqn. (14) as

$$\begin{cases} \text{The } j^{th} \text{ CU is a malicious user, if } TR^j(t) < M^{th} \\ \text{The } j^{th} \text{ CU is not a malicious user, } otherwise \end{cases}. \quad (32)$$

For simulation, a success detection is defined when the scheme correctly classify a CU to be an honest or a malicious user. Then "success detection rate" is the average success detection rate of the four types of CUs, an honest CU, a "*selfish*" user, an"*attack*" user and an "*adversary*" user. We set the same malicious rate for all malicious users as $a_{01} = \pi$ and/or $a_{10} = \pi$, with the SNR of the sensing channel in the range of -13dB to -10dB.

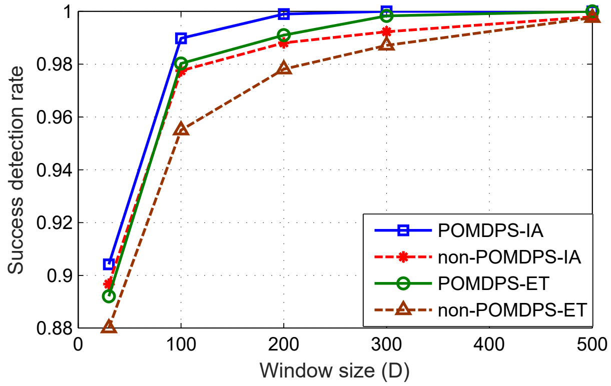

Fig. 4: Comparison of success detection rates of the proposed detection schemes with and without information of the PU signal when $\pi = 0.5$.

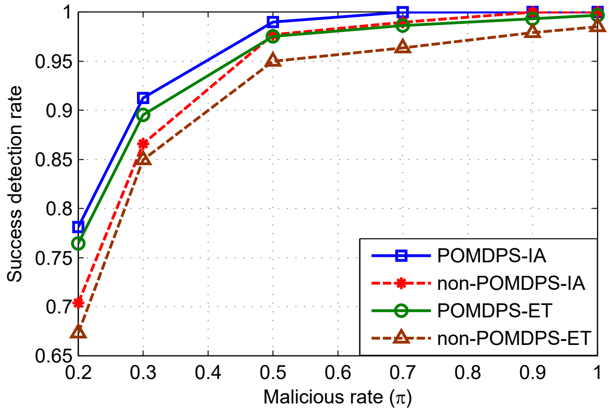

Fig. 5: Success detection rates of the proposed schemes with and without information of the PU signal versus malicious rates when $D = 100$.

Figure 3 shows the effect of the window size D on "success detection rate" of POMDPS-ET and non-POMDPS-ET. It can be seen that the bigger window size D can give a better "success detection rate". On the other hand, the POMDP may offer the proposed scheme better performance with a higher "success detection rate". When the window size $D \geq 300$ and the malicious rate $\pi \geq 0.3$, the proposed scheme based on POMDP (POMDPS-ET) may successfully distinguish between honest and malicious user nearly 100% of the time.

The success detection rate comparison of the proposed schemes with and without information of the PU signal is presented in Figure 4. The figure shows that the proposed scheme without information of the PU signal (i.e., POMDPS-ET and non-POMDPS-ET) may provide a similar performance in comparison to the case of available information of the PU signal (i.e., POMDPS-IA and non-POMDPS-IA). In both cases, POMDP may provide an advantage for the proposed scheme. When the value of D is small ($D < 100$), the success detection rate is strongly affected by D.

Relation between the success detection rate and the malicious rate is investigated in Figure 5. This figure shows that the proposed scheme is more successful at detecting malicious users when the malicious user is more harmful (i.e., a higher malicious rate user). A malicious user with a lower malicious rate is only slightly harmful to the sensing process, but it is more difficult to detect. At a certain value of π, the available information of the PU signal can provide to the proposed scheme a higher success detection rate.

In order to evaluate the performance of the whole sensing process, we define the "probability of error" as

$$Q_e = p_0 Q_f + p_1(1 - Q_d), \quad (33)$$

where Q_f and Q_d are the false alarm probability and the detection probability of the global decision, respectively.

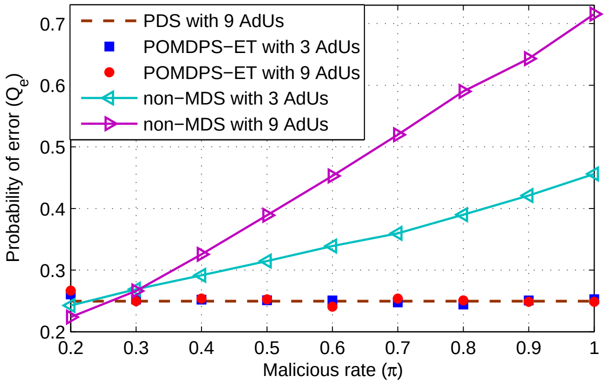

Fig. 6: Sensing performance of the considered schemes versus the malicious rates when *Adversary* users exist and $D = 100$.

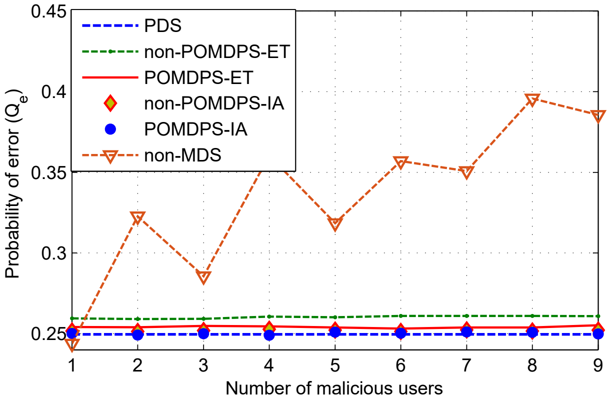

Fig. 8: Sensing performance comparison of the proposed schemes with and without information of the PU signal when various *Attack* users exist and $D = 100$, $\pi = 0.5$.

For showing the sensing performances of the proposed schemes, we consider the scenarios, for which the CR network includes only *one* honest CU and a various number of malicious users.

Figure 6 illustrates sensing performance of the proposed schemes. In this figure, we consider the presence of the most harmful type of malicious user, the "*adversary*" user (AdU). PDS is a perfect detection scheme that perfectly detects the presence of malicious users and rejects their sensing information out of the combination process. Therefore, performance of PSD-9 AdUs is not affected by the malicious rate. With the proposed scheme, it may be more difficult to detect the malicious users with a low malicious rate. However, a low malicious rate user only slightly harms the sensing process. Therefore, the proposed scheme may provide an almost similar performance to that of the PSD-9 AdUs, as shown in figure 6. On the other hand, non-MDS combines the received sensing information from both honest CUs and malicious users, and thus its sensing performance rapidly decreases when π increases. When the malicious rate is low (i.e., $\pi = 0.2$), non-MDS-9 AdUs seems to provide better performance for the proposed scheme with a lower probability of error. This is due to the malicious user with low malicious rates may behave as an honest CU for a longer time than as a malicious user; then, its reported sensing information may help improve the sensing performance of non-MDS-9 AdUs.

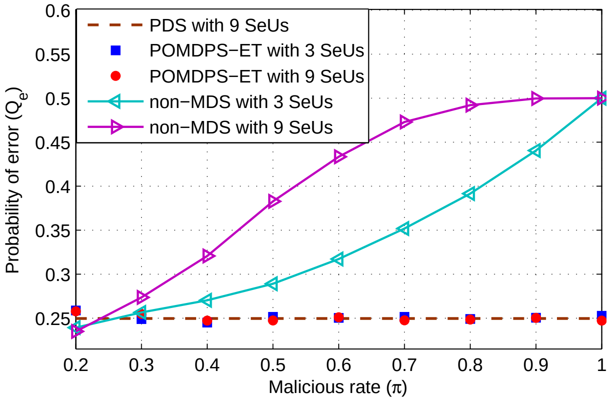

Fig. 7: Sensing performance of the considered schemes versus malicious rates when *Selfish* users exist and $D = 100$.

In figure 7, we present the sensing performance of the considered schemes with the presence of "*selfish*" users (SeUs). The proposed scheme can also effectively defend against "*selfish*" users by reject their harmful out of combination process. The proposed scheme separately considers each CU for a malicious test. Therefore, the proposed scheme can well run in the network will many malicious users that is proved by the performance of POMDPS-ET-9 SeUs in Figure 7, where only 1 honest and 9 SeUs are considered in the network.

Sensing performance comparison of the proposed schemes with and without sensing information of the PU signal is shown in Figure 8. In this simulation, we consider "*Attack*" users with malicious rate $\pi = 0.5$, i.e., the malicious users will randomly act as an Attack user in 50% of time and it randomly acts as an honest user in remaining 50% of time. This explains the reason why the higher number of malicious users (i.e., with malicious rate $\pi = 0.5$) does not ensure to make a stronger attack to the conventional scheme (non-MDS). On the other hand, Figure 8 shows that there is little difference among the considered schemes (i.e., POMDPS-IA, POMDPS-ET, non-POMDPS-IA and non-POMDPS-ET), in which non-POMDPS-ET gives the lowest performance, while POMDPS-IA and POMDPS-ET obtain a performance similar to that of PDS.

6 Conclusion

In this paper, we proposed a robust CSS scheme that can effectively defend against malicious users even when the malicious rate is low. Three types of malicious users, "*selfish*" users, "*attack*" users and "*adversary*" users are considered in this paper. By applying POMDP, the proposed scheme can detect malicious users faster (i.e., it requires smaller time slots (window size) to maintain the same success detection rate). Since the proposed scheme separately detects whether or not the CU is a malicious user, it can be robust in the cases where the number of malicious users in the network is much larger than the number of honest users. The simulation results show the effectiveness of the proposed scheme, which can reject almost all harmful effects from malicious users in order to protect CSS.

References

[1] Y. Hur, J. Park, W. Woo, K. Lim, C.-H. Lee, H. S. Kim and J. Laskar, "A wideband analog multi-resolution spectrum sensing (MRSS) technique for cognitive radio (CR) systems", in Proc. IEEE Int. Symp. Circuit and System, Greece, 2006, pp.4090-4093.

[2] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio", in Proc. Allerton Conf. on Communications, control, and computing, Monticello, 2004.

[3] G. Ganesan and Y. G. Li, "Cooperative spectrum sensing in cognitive radio networks", in Proc. IEEE Symp. New Frontiers in Dynamic Spectrum Access Networks (DySPAN05), Baltimore, USA, 2005, pp.137-143.

[4] S. M. Mishra, A. Sahai and R. W. Brodersen, "Cooperative sensing among CRs," IEEE International Conf. Commun., vol. 4, pp.1658-1663, 2006.

[5] R. Deng, J. Chen, C. Yuen, P. Cheng and Y. Sun, "Energy-Efficient Cooperative Spectrum Sensing by Optimal Scheduling in Sensor-Aided Cognitive Radio Networks," IEEE Transactions on Vehicular Technology, vol.61, no.2, pp.716-725, Feb. 2012

[6] P. Kaligineedi, M. Khabbazian and V. K. Bhargava, "Secure cooperative sensing techniques for cognitive radio systems," IEEE International Conf. Commun. (ICC08), pp. 3406-3410, 2008.

[7] P. Kaligineedi, M. Khabbazian and V. Bhargava, "Malicious User Detection in a Cognitive Radio Cooperative Sensing System," IEEE Transactions on Wireless Communications, vol.9, no.8, pp.2488-2497, 2010.

[8] J. Li, J. Liu and K. Long, "Reliable Cooperative Spectrum Sensing Algorithm Based on Dempster-Shafer Theory," Global Telecommunications Conference (GLOBECOM 2010), pp.1-5, 2010

[9] X. Zheng, J. Wang, Q. Wu and J. Chen, "Cooperative spectrum sensing algorithm based on Dempster-Shafer theory," 11th IEEE Singapore International Conference on Communication Systems, ICCS 2008 , pp.218-221, 2008.

[10] N. Nguyen-Thanh and I. Koo, "Empirical Distribution-Based Event Detection in Wireless Sensor Networks: An Approach Based on Evidence Theory," Sensors Journal, IEEE , vol.12, no.6, pp.2222-2228, 2012.

[11] H. Xiaofan H, D. Huaiyu and N. Peng, "HMM-Based Malicious User Detection for Robust Collaborative Spectrum Sensing," IEEE Journal on in Selected Areas in Communications, vol.31, no.11, pp.2196-2208, 2013.

[12] D. P. Bertsekas, Dynamic Programming and Optimal Control. Athena Scientic, 2nd edition, vol. 1 and 2, 2001.

Acknowledgement

This work was supported by the National Research Foundation (NRF) of Korea funded by the MEST (Nos. NRF-2015R1A2A1A15053452 and NRF-2014R1A1A2005378)

Hiep Vu-Van received the B.E. degree in Electronics & Telecommunications Engineering from Ton Duc Thang University, Vietnam in 2005 and the B. degree in Business Administration from University of Economy Ho Chi Minh city, Vietnam in 2007. He received Ph.D. degree in Electrical Engineering from the University of Ulsan (UOU), Korea, in 2013, where he is now research fellow. His current research interests include cognitive radio and next generation wireless communication systems.

Insoo Koo received the B.E. degree from the Kon-Kuk University, Seoul, Korea, in 1996, and received the M.S. and Ph.D. degrees from the Gwangju Institute of Science and Technology (GIST), Gwangju, Korea, in 1998 and 2002, respectively. From 2002 to 2004, he was with Ultrafast Fiber-Optic Networks (UFON) research center in GIST, as a research professor. For one year from September 2003, he was a visiting scholar at Royal Institute of Science and Technology, Sweden. In 2005, he joined University of Ulsan where he is now professor. His research interests include next generation wireless communication systems and wireless sensor networks.