COOPERATIVE SENSING IN COGNITIVE NETWORKS UNDER MALICIOUS ATTACK

Mai Abdelhakim Lei Zhang Jian Ren Tongtong Li
Department of Electrical & Computer Engineering, Michigan State University, East Lansing, MI 48824, USA.
Email: {abdelhak, zhangle3, renjian, tongli}@egr.msu.edu

***Abstract*—This paper considers cooperative sensing in cognitive networks under Spectrum Sensing Data Falsification attack (SSDF) in which malicious users can intentionally send false sensing information. One effective method to deal with the SSDF attack is the q-out-of-m scheme, where the sensing decision is based on q sensing reports out of m polled nodes. The major limitation with the q-out-of-m scheme is its high computational complexity due to exhaustive search. In this paper, we prove that for a fixed percentage of malicious users, the detection accuracy increases almost exponentially as the network size increases. Motivated by this observation, as well as the linear relationship between the scheme parameters and the network size, we propose a simple but accurate approach that significantly reduces the complexity of the q-out-of-m scheme. The proposed approach can easily be applied to the large scale networks, which can be much more reliable under malicious attacks.**

***Index Terms*—cognitive radio, cooperative sensing, malicious attack, data fusion.**

I. INTRODUCTION

Cognitive network is a potential wireless communication paradigm that is proposed to alleviate spectrum scarcity and underutilization of spectrum bands [1], [2]. Cognitive networks are not licensed to any particular band, however they allow radios to access the spectrum opportunistically when it is not used by the licensed users (also referred to as primary users). Cognitive radios must be cautious to their secondary access of the spectrum. Hence, spectrum sensing is an essential and critical function in cognitive radios.

A single cognitive radio may not be able to detect the presence of a primary user due to shadowing, fading, propagation losses and interference effects. Therefore, cooperative sensing was proposed for cognitive radio networks in order to improve the sensing operation [3], [4]. However, cooperation sensing itself arouses some physical layer security concerns that have recently been discussed in the literature. Cognitive network may contain some malicious users that could attack the network and corrupt the spectrum sharing operation. Two types of malicious attacks are generally considered in cooperative sensing [5], [6]: the first one is the Incumbent Emulation attack (IE), where some malicious users know the characteristics of the primary signal and transmit a signal with similar characteristics so that other secondary users would believe that a primary user is present. The other type is the Spectrum Sensing Data Falsification Attack (SSDF), where malicious users send false sensing information intentionally. In this paper we focus on the SSDF attack.

In [7], we proposed a q-out-of-m scheme for final decision making. In this scheme, the final decision is based on q sensing reports out of m polled nodes. However, due to the large computational complexity required by exhaustive search, the optimal q-out-of-m scheme would be unfeasible as the network size increases. In order to solve this problem, in this paper we propose a suboptimal q-out-of-m scheme for collaborative sensing under SSDF attack. The proposed suboptimal scheme exploits the linear relation between the scheme parameters and the network size, and hence reduces the computational complexity significantly. The significance of the proposed approach is further demonstrated as we show that for a fixed percentage of malicious users and a predefined miss detection constraint, the false alarm rate decreases almost exponentially as the network size increases. Due to the low computational complexity the proposed approach can be easily applied to large scale networks, which are much more reliable under malicious attack.

This paper is organized as follows: the problem formulation and the optimal q-out-of-m scheme are presented in section II. The proposed simplified collaborative sensing approach and the performance bounds are provided in section III. The simulation results are given in Section IV. Finally, the paper is concluded in section V.

II. PROBLEM FORMULATION

We assume a centralized setup, where each node in the network performs spectrum sensing and reports its one bit hard decision result to a central node (fusion center) through a control channel. The control channel is assumed to be error free. The sensing result is either '1' which means primary user is present or '0' which means that the band is not used by the primary. The fusion center is then responsible for making the final decision based on the received sensing reports from all nodes. For simplicity, we adopt the assumptions in [8] that all users experience independent and identically distributed (i.i.d) fading with the same average signal to noise ratio (SNR), such that each user has a probability of false alarm P_f, and a probability of detection P_d. When energy detection is used, these probabilities are related as follows [8]:

$$P_d = \int_{\gamma} Q_l(\sqrt{2lx}, \sqrt{G_l^{-1}(P_f)}) f_\gamma(x) dx, \tag{1}$$

where l is time-bandwidth product assumed to be an integer, γ is the SNR, $G_l^{-1}(P_f)$ is the energy detection threshold expressed as the function of false alarm probability P_f, which is defined in [8], and $Q_l(\cdot,\cdot)$ is the generalized Marcum Q-function [9]. In Rayleigh fading channels, γ is a random variable and $f_\gamma(x)$ is the probability density function of γ. The presence of malicious users that send false sensing data would severely disrupt the cognitive network. In our system model, we assume that the network consists of *n* active users including *k* malicious users. We also refer to *n* as the network size. We assume that malicious users can detect the primary signal with no error. The percentage of malicious users *k/n* is denoted by α and it is assumed to be known at the fusion center.

A. Optimal q-out-of-m Cooperative Sensing Scheme

The main task of the fusion center is to decide whether the band is idle or busy based on the sensing reports. In a generalized sensing strategy, the fusion center randomly polls m out of n users and relies on q-out-of-m rule for final decision making (the fusion center decides that a primary is present if q out of the m polled users report '1') [7]. The main objective is to minimize the overall false alarm rate (Q_f) while keeping the overall miss detection (Q_m) below

certain predefined value β. Hence, it is desired to get the optimum parameters m and q that can achieve the objectives. The problem can be formulated as follows:

$$\begin{aligned} &\min_{m,q} Q_f(m,q); \\ &s.t.\ \ Q_m(m,q) \leq \beta; \\ &s.t.\ \ 1 \leq q \leq m \leq n; \quad q,m \in \mathbb{N}. \end{aligned} \tag{2}$$

In order to obtain a closed form expressions for Q_f and Q_m, we define $P_{k,n-k}^{d,m-d}$ as the probability of polling $m-d$ out of $n-k$ benign users and d out of k malicious users such that:

$$P_{k,n-k}^{d,m-d} = \frac{\binom{k}{d}\binom{n-k}{m-d}}{\binom{n}{m}} \tag{3}$$

According to our system model, the false alarm rate is expressed in (4), where three different scenarios are implied [7]:

Scenario 1: $k \leq q$. When the number of malicious users is smaller than or equal to q (consequently $k \leq m$), there would be some benign users involved, among the m polled users, in the final decision making. If the m polled users contain d out of the k malicious users, then the false alarm occurs when there are at least $q-d$ benign users sending false alarms.

Scenario 2: $m+k-n > q$. When the number of malicious users is large enough to make $m+k-n > q$ (Since $m-n \leq 0$, this implies $k > q$), there are so few benign users such that among m polled users, there are at least $m-(n-k) > q$ malicious users. In this case, false alarm happens with probability 1, leading to spectrum waste. This is because that the secondary user will not use the channel even if there is a white space.

Scenario 3: $k > q$ *but* $m+k-n \leq q$. When the number of malicious users is moderate, the false alarm probability depends on how many malicious users are polled. If among m polled users, there are at least q malicious users included, then the secondary system is jammed regardless of the behavior of other benign users. Otherwise, if there are $d < q$ malicious users are polled, then false alarm occurs when there are at least $q-d$ benign users reporting erroneous results. The miss detection probability Q_m can be expressed in terms of the detection probability Q_d, such that $Q_m = 1 - Q_d$. In q-out-of-m scheme, Q_d is given as follows:

$$Q_d = \begin{cases} 0, & \text{if } n-k < q; \\ \displaystyle\sum_{d=\max(0,\ m+k-n)}^{\min(k,m-q)} P_{k,n-k}^{d,m-d} \sum_{i=q}^{m-d} P_d^i (1-P_d)^{m-d-i}, & \\ \quad \text{if } n-k \geq q. & \end{cases} \tag{5}$$

It is shown in (5) that if q is greater than the number of benign users, then the primary users will never be detected (i.e. $Q_d = 0$). This case will result in severe interference to the primary system. Thus, q should not be too large. In the case that the number of benign users is greater than q, then at least there should be q users reporting the presence of the primary in order to be able to detect it. The number of malicious users $d = max(0, m+k-n)$ indicates that when the number of users being polled, m, is greater than that of the benign users, then there are at least $m-(n-k)$ copies of malicious reports received by the base station.

B. Complexity of Optimal q-out-of-m Scheme

With the above expressions for false alarm and miss detection, it is noted that finding the optimum m and q in (2) is a nonlinear integer optimization problem that is hard to solve theoretically. The optimal approach is to perform exhaustive search over all possible *m* and *q* values and choosing the (m_o, q_o) pair that result in the lowest false alarm rate while satisfying the QoS constraints. The complexity of the optimal algorithm is $O(n^2)$. As a result, it will not be feasible at larger network sizes. It is noted that the sensing operation is done periodically and the network has to adapt to the environmental variations and react in a quick manner. Hence, a simplified approach is required for cooperative sensing under the SSDF attack.

III. Simplified Collaborative Sensing Algorithm

In this section we first highlight our motivation to the proposed approach and describe the procedure we follow in the simplification process. In addition, we provide performance analysis and derive performance bounds on the overall false alarm rate.

A. Motivation

We aim at investigating the asymptotic network performance as n increases. However, it is not possible to get an upper bound on Q_f directly from equation (4), since the optimal m and q are function of the network size (n) and the percentage of malicious users (α). Moreover, it is not feasible to get the optimal m and q for large network sizes using the optimal exhaustive approach due to the huge computational complexity. However, we search for the optimal parameters at relatively small network sizes using $\beta = 0.01$, and two main observations are made:

Observation 1: The optimal *m* is almost independent of the percentage of malicious users and in all cases it is equal, or very close, to the value of *n* as shown in Figure 1. One possible interpretation for this is that since the polling is random, it is better to know the decision of all the nodes. This is not the case when a malicious node detection scheme is employed, as in this case the reports of malicious users would be discarded.

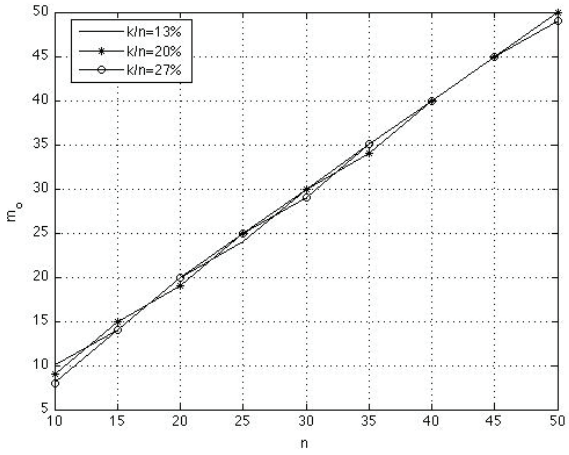

Fig. 1. Optimal m vs. n

Observation 2: q also follows an approximate linear function of n with different slopes depending on the percentage of malicious users as shown in Figure 2.

$$Q_f = \begin{cases} \sum\limits_{d=\max(0,\ m+k-n)}^{k} P_{k,n-k}^{d,m-d} \sum\limits_{i=q-d}^{m-d} \binom{m-d}{i} P_f^i (1-P_f)^{(m-d-i)}, & \text{if } k \le q; \\ 1, & \text{if } m+k-n > q; \\ \sum\limits_{d=q}^{\min(k,m)} P_{k,n-k}^{d,m-d} + \sum\limits_{d=\max(0,\ m+k-n)}^{q-1} P_{k,n-k}^{d,m-d} \sum\limits_{i=q-d}^{m-d} \binom{m-d}{i} P_f^i (1-P_f)^{m-d-i}, & \text{if } \substack{k>q \\ m+k-n \le q} \text{ and}. \end{cases} \quad (4)$$

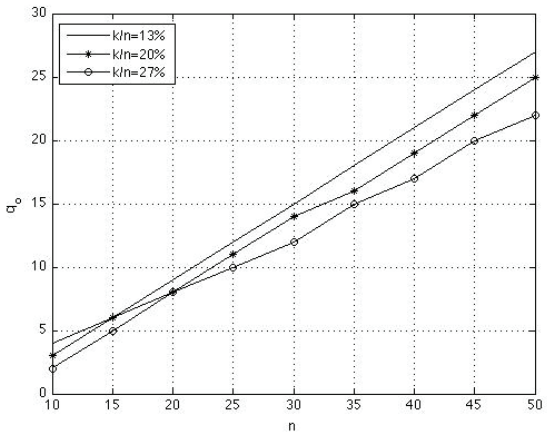

Fig. 2. Optimal q vs. n

From these observations, the solution of (2) can be simplified as will be illustrated in the following subsection.

B. Proposed Algorithm

We showed that the optimal m and q can be approximated as linear function of n, with m almost equals to n. We exploit this observation to obtain a suboptimal m and q for large network sizes without the need to perform exhaustive search that would be impractical as n increases. In our simplified approach we set $m = n$, and hence, our problem is reduced to finding the optimal q that achieves the network objectives. Therefore, the problem can be reformulated as:

$$\begin{aligned} &\min_q Q_f(n,q); \quad (6) \\ &s.t.\ \ Q_m(n,q) \le \beta; \\ &s.t.\ \ 1 \le q \le n; \quad q \in \mathbb{N}. \end{aligned}$$

Solving the above problem using exhaustive search would reduce the complexity of the algorithm to $O(n)$ instead of $O(n^2)$ as compared to the scheme in [7]. We can further simplify the algorithm by using the following linear function:

$$q_{n_1,\alpha} = q_{n_0,\alpha} + S_o(\alpha)(n_1 - n_0) \quad (7)$$

where $S_o(\alpha)$ is the slope of the q_o versus n curve at a specific percentage of malicious users (α), $q_{n_1,\alpha}$ is the suboptimal q value at a network size n_1, and $q_{n_o,\alpha}$ is the optimal q value at a network size n_0. Both $q_{n_1,\alpha}$ and $q_{n_o,\alpha}$ are at α percent of malicious users. $q_{n_o,\alpha}$ is considered as a known reference point and is obtained using the optimal q-out-of-m scheme at a relatively small network size n_o. It is noted that q is an integer value, hence we let $q = \lceil q_{n_1,\alpha} \rceil$, which is the smallest integer larger than or equal to $q_{n_1,\alpha}$. Thus, we can get the optimal scheme results at relatively small value for n. These (m, q) pairs for different network size n and k/n ratio, can be obtained once and stored in a look-up table, and can be used to get the suboptimal scheme parameters for larger network sizes.

C. System Performance Versus The Network Size

One interesting result we observed in [7] is that if we increase the network size, the overall network performance is improved even if the percentage of malicious users is kept the same. In this section, we provide a theoretical explanation for this result.

Proposition: For a fixed malicious users ratio (α) and network size n, $Q_f \le exp\left(-\frac{(Bn+C)^2}{An}\right)$, where A, B and C are constants.

In order to prove this, we consider our simplified approach where we set $m = n$, then it follows that $d = k$ and the term $P_{k,n-k}^{d,m-d}$ is equal to one. Assuming the case where $q \ge k$, the false alarm probability Q_f can be expressed as:

$$Q_f = \sum_{i=q-k}^{m-k} \binom{m-k}{i} P_f^i (1-P_f)^{m-k-i}. \quad (8)$$

It is clear that Q_f is the summation over a binomial probability density function with parameters P_f and $m-k$, where the random variable is the number of benign users having false alarm. Thus, Q_f can be upper-bounded using the Chernoff bound as follows [10]:

$$Q_f \le \exp\left(-\left(\frac{q-d}{P_f(m-d)} - 1\right)^2 (m-d)\frac{P_f}{3}\right) \quad (9)$$

Substituting with $m = n$ and using (7) while assuming the percentage of the malicious users is fixed, Q_f can be bounded as follows:

$$\begin{aligned} Q_f &\le \exp\left(-\left(\frac{q_o + (n_1 - n_o)S(\alpha) - k}{P_f n_1(1-k/n_1)} - 1\right)^2 n_1(1-k/n_1)\frac{P_f}{3}\right) \\ &\le \exp\left(-\left(\frac{q_o + (n_1 - n_o)S(\alpha) - k}{P_f n_1(1-\alpha)} - 1\right)^2 n_1(1-\alpha)\frac{P_f}{3}\right) \quad (10) \end{aligned}$$

It is noted that at a fixed percentage of malicious users, q_o, n_o, α and $S(\alpha)$ are constants and the bound can be simplified as follows:

$$Q_f \le \exp\left(-\frac{(q_o - n_o S(\alpha) + n1(S(\alpha) - P_f(1-\alpha)) - k)^2}{3P_f n_1(1-\alpha)}\right) \quad (11)$$

$$Q_f \le \exp\left(-\frac{(C + n_1 B)^2}{An_1}\right) \quad (12)$$

where $A = 3P_f(1-\alpha)$, $B = (S(\alpha) - P_f(1-\alpha))(1-\alpha/(S(\alpha) - P_f(1-\alpha)))$ and $C = q_o - n_o S(\alpha) + P_f\alpha$. A, B and C are constants.

If we take the limit as $n \to \infty$ then $Q_f \to 0$. Therefore, for a fixed α, we can improve the performance significantly by increasing the network size. An intuitive explanation for this result is that, as the network size (n) increases, the detection decision is based on the reports from a larger pool of users, and hence decreases Q_f. This observation highlights the impact of the proposed algorithm, which can easily be applied to large sized network with very low computational complexity, and comparable result with that of the optimal scheme.

IV. Simulation Results

For each benign user, the individual sensing probabilities are $P_f = 0.1$ and $P_d = 0.775$, where we have used the settings in [8] with average SNR=5dB and $l = 5$. The slope $S_o(\alpha)$ is obtained from the curves shown in Figure (2) such that it is equal to (0.56, 0.52, 0.52) for the corresponding α (0.13, 0.20, 0.27), respectively. Applying the simplified cooperative sensing scheme, the false alarm rate is obtained in Figure 3 at different percentages of malicious users, where we use the reference point at $n_o = 40$. The figure shows that even at a fixed percentage of malicious users, the performance is improved as the network size increases. Hence, networks with larger sizes can tolerate higher percentage of malicious users. The optimal method and the simplified approach are compared in Figure 4 at n=50. It can be shown that the linear approximation used in the simplified approach is almost achieving the same performance as the optimal algorithm that uses exhaustive search for m and q. The bound on Q_f is also plotted for both schemes. For the optimal scheme, Chernoff bound is used to get the upper bound on Q_f given in equation 4. The bound for the proposed scheme is directly obtained using equation (12). It is also noted from Figure 4 that using the simplified linear approach, the overall miss detection (Q_m) is still within the acceptable range when compared to our predefined limit ($\beta = 0.01$).

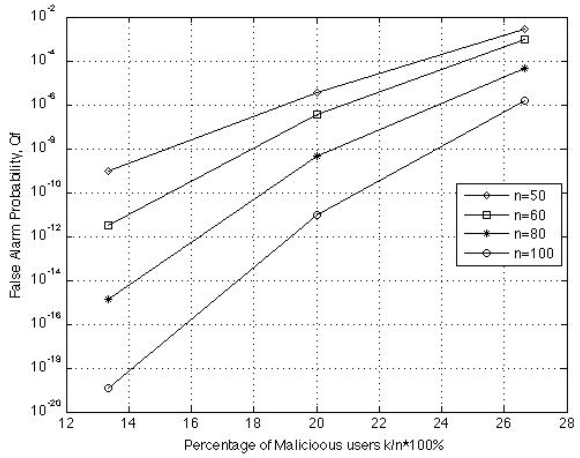

Fig. 3. False alarm rate vs. the percentage of malicious users at different values of n

V. Conclusion

In this paper, we proved that for a fixed percentage of malicious users and a predefined miss detection constraint, the false alarm rate decreases almost exponentially as the network size increases. We also proposed a simplified q-out-of-m approach for cooperative spectrum sensing under SSDF attack that can be easily applied to large scale networks. The proposed scheme exploits the linear approximation of the scheme parameters with the network size. Hence, it significantly reduces the computational complexity of the optimal q-out-of-m scheme. The simulations showed that the simplified approach is accurate and achieves comparable results with the optimal q-out-of-m method.

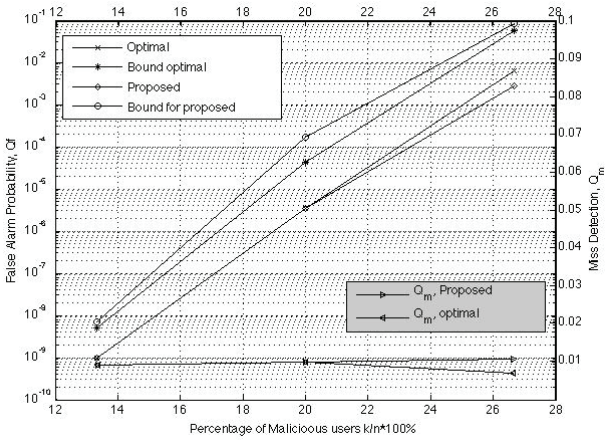

Fig. 4. Exact and bound for the optimal scheme and the proposed simplified scheme

References

[1] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201 – 220, feb. 2005.

[2] I. Mitola, J. and J. Maguire, G.Q., "Cognitive radio: making software radios more personal," *Personal Communications, IEEE*, vol. 6, no. 4, pp. 13 –18, aug. 1999.

[3] D. Cabric, S. Mishra, and R. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 772 – 776 Vol.1, nov. 2004.

[4] S. Mishra, A. Sahai, and R. Brodersen, "Cooperative sensing among cognitive radios," *IEEE International Conference on Communications, 2006. ICC '06*, vol. 4, pp. 1658 –1663, jun. 2006.

[5] R. Chen, J.-M. Park, Y. Hou, and J. Reed, "Toward secure distributed spectrum sensing in cognitive radio networks," *Communications Magazine, IEEE*, vol. 46, no. 4, pp. 50–55, apr. 2008.

[6] J. Wei and X. Zhang, "Two-tier optimal-cooperation based secure distributed spectrum sensing for wireless cognitive radio networks," *INFOCOM IEEE Conference on Computer Communications Workshops*, pp. 1–6, mar. 2010.

[7] H. Wang, L. Lightfoot, and T. Li, "On phy-layer security of cognitive radio: Collaborative sensing under malicious attacks," *Information Sciences and Systems (CISS), 2010 44th Annual Conference on*, pp. 1 –6, mar. 2010.

[8] A. Ghasemi and E. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2005.*, 2005, pp. 131–136.

[9] A. Nuttall, "Some integrals involving the Q_M function," *IEEE Transactioins on Information Theory*, vol. 21, no. 1, pp. 95–96, 1975.

[10] M. Mitzenmacher and E. Upfal, *Probability and computing randomized algorithms and probabilistic analysis*. Cambridge Univ. Press, 2009.