Defense against Spectrum Sensing Data Falsification Attacks in Cognitive Radio Networks

Chowdhury Sayeed Hyder, Brendan Grebur, and Li Xiao

Department of Computer Science and Engineering
Michigan State University
East Lansing, MI 48823, USA
{hydercho,greburbr,lxiao}@cse.msu.edu

Abstract. IEEE 802.22 is the first standard based on the concept of cognitive radio. It recommends collaborative spectrum sensing to avoid the unreliability of individual spectrum sensing while detecting primary user signals. However, it opens an opportunity for attackers to exploit the decision making process by sending false reports. In this paper, we address security issues regarding distributed node sensing in the 802.22 standard and discuss how attackers can modify or manipulate their sensing result independently or collaboratively. This problem is commonly known as spectrum sensing data falsification (SSDF) attack or Byzantine attack. To counter the different attacking strategies, we propose a reputation based clustering algorithm that does not require prior knowledge of attacker distribution or complete identification of malicious users. We compare the performance of our algorithm against existing approaches across a wide range of attacking scenarios. Our proposed algorithm displays a significantly reduced error rate in decision making compared to current methods. It also identifies a large portion of the attacking nodes and greatly minimizes the false detection rate of honest nodes.

Keywords: Cognitive Radio Network, SSDF attack, 802.22.

1 Introduction

As wireless devices are dominating the methods in which people communicate with one another, the necessary resources to support these conveniences are being ever harder to obtain. In contrast, licensed bandwidth spectrums often go underutilized as demands for those services shift temporally or spatially. Static spectrum allocation cannot efficiently support the demand of such pervasive wireless devices. To combat this salient impedance, the concept of Cognitive Radio Networks (CRN) has been proposed [9].

In order to maximize radio spectrum usage, CRNs utilize an opportunistic approach to allocate frequencies. Under the scheme, two types of users exist: primary users (PU) and secondary users (SU). Individuals who have obtained a license to broadcast in a fixed spectrum range are classified as primary users. On

M. Rajarajan et al. (Eds.): SecureComm 2011, LNICST 96, pp. 154–171, 2012.

the other hand, secondary users attempt to "fill in the gaps" by utilizing unused spectrums. The users complement each other allowing maximal utilization of a specified spectrum.

Naturally, complications arise as secondary users must release a spectrum when the primary user for that channel starts to transmit. Several research groups are working to develop standards to meet these requirements. 802.22, the first CR based network standard, defines a centralized, single hop, point to multipoint communication standard for wireless regional area network (WRAN). This standard defines the implementation of opportunistic spectrum sharing (OSS) by outlining how/when wireless devices are able to utilize temporarily idle bands in a licensed radio spectrum. The proposal also defines the cellular like communication interface between a base station (BS) and secondary users called Consumer Premise Equipments (CPE). The BS is responsible for controlling the spectrum usage and channel assignment to CPEs. All CPEs in a cell must periodically monitor primary user signals and leverage the distributed sensing power of CPEs through continual spectrum reports obtained from secondary users.

To coordinate the process, a centralized BS collects sensing information from the secondary users residing in the cell. Each user submits a hypothesis regarding whether they suspect the primary user is transmitting. As radio waves are affected by physical barriers or environmental conditions, the detection accuracy of any node within sensing range of the PU's signal varies from time to time. Malfunctions associated with the sensing equipment may also influence the node's observed measurements. From the hypotheses supplied by the secondary users, the BS must decide on the actual state of the associated spectrum. Once a decision is made, the base station can inform SUs and revoke permission for those users currently transmitting on that spectrum.

Due to its unique characteristics, CRNs face new security threats in addition to the common existing security challenges in wireless network. One typical type of attack is the Spectrum Sensing Data Falsification (SSDF) attack or Byzantine attack. During such an assault, the malicious user compromises one or more of the secondary users and may begin sending modified sensing results to the BS. In this way, an attacker tries to influence the BS into producing a wrong decision about the channel status. Compromised nodes may work independently or may collaborate to reduce spectrum utilization and degrade overall performance of the network.

Constructing a decision-making strategy that mitigates the impact of both types of attackers will prove invaluable as the reach of CRNs expands into more places. By strengthening the base station against malicious or malfunctioning users, the interference produced from CRNs will be minimized, potentially expediting the implementation of such network alternatives. Ultimately, both users and businesses can reap the benefits of efficient radio spectrum usage through CRNs.

There are very limited research works that address SSDF attack and related security problems. Existing approaches like [1], [8], [10] mainly consider independent malicious attack. However, these approaches either require prior informa-

tion of attackers (e.g. number of attackers [10], attackers' distribution, attacking strategy [8] etc.) or depend on careful threshold selection [1]. For instance, algorithm in [10] does not work in presence of multiple attackers. Similarly, performance of the algorithm proposed in [1] degrades significantly if incorrect threshold is chosen. To our best knowledge, we find only one paper [2] that handles both independent and collaborative attacks. This approach uses a reputation based method to limit the error rate in deciding channel status and in identifying attackers. Although its identification rate of attackers is high, it also misdetects a large number of honest users as attackers. Additionally, this approach fails to defend against collaborative attack and error rate (i.e. number of incorrect decision) increases almost linearly with number of attackers.

On the contrary, we propose an adaptive reputation based clustering algorithm to defend against both independent and collaborative SSDF attack that does not require any prior information about number of attackers or attacking strategies. The whole process goes through a sequence of steps in each time step. To start with, the algorithm clusters the nodes based on the sensing history and initial reputation of nodes. Each cluster takes its decision about the channel status according to the relative closeness of nodes from the median of that cluster. Finally, channel status is decided on majority of clusters' decision. At the end of the time step, the final decision is propagated back to the clusters and then to the individual nodes. Each node is assigned a share (positive or negative) of the final decision and the reputation of each node is adjusted based on its participation in the decision making process. The adjusted reputation of nodes is used to adjust the number of clusters for the next step. In this way, the algorithm works through several steps in forward and backward direction in each time step and recursively updates the clusters and the reputation of nodes. We compare performance of our algorithm with that of the algorithm proposed in [2] under different attacking scenarios. Our algorithm handles SSDF attack significantly better than the one in [2] and minimizes error in deciding channel status. Our algorithm also identifies a significant number of attackers while keeping the misdetection rate to a minimum level.

The next section explores various approaches currently proposed and specifically identifies their limitations in the problem domain. Section 3 formally defines the problem area including the setup used to measure each method. Section 4 describes a high-level overview of our proposed method mainly focusing on design choices. Section 5 covers a detailed description of the algorithm. Section 6 compares the results with current methods and Section 7 concludes with contributions and future work.

2 Related Work

Until recently, security issues in CRN have not been addressed well in research works. However, in this section, we present existing solutions to combat against SSDF attack into three categories - reputation-based, neighborhood distance based, and artificial intelligence approaches.

2.1 Reputation Based Approaches

Wang et al. [8] propose an onion peeling approach based on bayesian statistics to assign suspicion levels for all nodes in the network. If the suspicion level of any node exceeds a certain threshold, it is marked as malicious and removed from decision making. They tested their heuristic based approach for false alarm attacks, miss detection attacks, and combinations thereof. However, they assume that base station has prior knowledge about the activities of attackers which is not very common. Without such information, the thresholds are approximated, resulting in significant false detections of attackers.

Chen et al. [3] propose a hybrid method named weighted sequential probability ratio test (WSPRT) that combines reputation and a sequential probability ratio test to identify malicious or faulty units. This method outperforms standard fusion center decision making strategies, including OR, AND, and SPRT during simulations in both minimizing missed detections and maximizing the correct sensing ratio. However, WSPRT was only tested against attackers utilizing an always-false or always-free response. Such methods represent an unsophisticated attack strategy that is not likely to reflect encountered attackers. The method also requires an additional number of secondary user sensing reports to generate the final fusion center hypothesis, which can impede the overall performance of the system and potentially cause primary user interference.

Recently, Rawat et. al. in [2] explores independent and collaborative SSDF attacks. They determined optimal attacking strategies for collaborating attackers where the fusion center cannot possibly discriminate between honest and attacking CRs. A mathematically rigorous analysis of detection performance is carried out using the Kullback-Leibler divergence (KLD). According to their result, in presence of 50% independent attackers, fusion center cannot differentiate the difference between the honest users and the attackers. However, for collaborative attack, this ratio reduces to 35%. Furthermore, they proposed a simple reputation-based method to identify attackers. A major weakness of the method stems from its massive misdetection of attackers during the identification stage. The proposed method uses a relatively small sensing window for analyzing reporting patterns to identify attackers. Under such limited time spans, temporary sensing errors of honest users cause their sensing signatures to deviate from the consensus. As more honest users are removed from the voting process, the method leaves the responsibility of final decision making up to only a few users. In such scenarios, the system is left in an extremely fragile state. Any attack on the remaining users causes the entire cell to be compromised. In addition, the method's probability of error increases dramatically when as little as 35% of the nodes are collaborating in attacks.

2.2 Data Mining Approaches

In [1], a new approach based on K-neighborhood distance algorithm is presented to detect independent malicious users. The approach does not need any prior knowledge of attacker distribution and exposes attackers across multiple sensing

rounds. However, when attackers collaborate and have secondary user data, they can successfully evade detection.

Further work has been done by [6] in establishing a more robust fusion center decision algorithm. Specifically, particular pieces of sensing information are used to validate the primary user hypothesis presented by each secondary user. Information regarding PU positioning and path loss to the secondary user can corroborate the hypothesis. The compiled set of sensory reports are analyzed using a biweight estimate and median absolute deviation to calculate magnitudes, which are then compared against thresholds to identify the attackers.

The proposed method dramatically increases misdetections when using incorrect static thresholds. Inaccurately identified secondary users could be excluded from the decision making process, resulting in a PU signal being ignored. Ultimately, the correct setting of the detection thresholds can only be achieved with prior knowledge of attacker distribution. Again, the information is unlikely to be available.

2.3 Artificial Intelligence Approaches

Clancy et al. [4] take a practical look into devising security for the physical transport layer of CRNs, focusing on CRs with artificial intelligence. When implementing such schemes, the CRs are highly susceptible to short-term and long-term manipulations caused by corrupted sensory data, altered node statistics, and inaccurate beliefs regarding the current environment. The paper addresses a series of steps to combat these sensitive areas by assuming a noisy environment, implementing levels of common sense, and programmatically resetting learned values to avoid extended corruption from attackers. They offer up the use of swarm behavior in determining a global decision on whether a sensed signal was actually generated by a primary user, along with a trust-based scheme. The proposals on how these CRs should operate in the field are presented without details for verification. They also did not address how to incorporate this new information into the current 802.22 system.

The current state of research holds very few proposals that work on realistic knowledge of the operating environment. Approximating these values fundamentally skews the proposed approaches' effectiveness. Furthermore, misidentification of attackers could also severely impact the effectiveness of strategies. Such considerations must be respected to develop a truly robust scheme. Ultimately, the approaches will need to face real attacks while producing acceptable error rates. In this paper, we explore strategies that exhibit these characteristics without being hindered by any assumptions of the operating environment or attacker strategy.

3 System Model

In this section, we briefly describe the topology of the CR network. We explain how the BS operates and takes decision regarding channel status from collective

sensing reports. We also formulate different attacking models and analyze how they exploit the decision mechanism of BS.

The BS is the central authority to coordinate and control the operation of all secondary nodes in its cell. BS instructs SUs to sense a channel according to the standard. Each node uses the same spectrum sensing technique for PU detection. Spectrum sensing itself is an ongoing research topic and is out of the scope of this paper. For simplicity, we assume that secondary users use the threshold based energy detection technique for spectrum sensing and all nodes use the same threshold provided by the BS. All nodes prepare their reports based on sensing and send their sensing results. However, different sensing techniques offer different levels of detection accuracy and may affect the sensing decision. Later in the results section, we perform simulation with varying sensing accuracy. BS then decides the channel status considering the sensing results from all the nodes. We also assume that users have no knowledge about the actual channel status.

We consider two types of users in the network - honest users and dishonest users. In each time slot, honest SUs sense the channel, compare the sensed energy with the threshold, and decide independently about the channel status. Finally, they report their sensed status to the BS without any alteration.

On the other hand, the dishonest users alter their sensed results and send it to BS. They can be selfish or malicious based on their intention. We commonly term them as 'attackers'. A selfish attacker has a different perspective from a malicious one. From a selfish attacker's point of view, the goal is to make the base station take a wrong decision about the idle channel so that it may utilize the spectrum opportunity. As a result, spectrum utilization will be significantly reduced. On the contrary, a malicious attacker's goal is not only to minimize the spectrum utilization, but also degrade the network performance. The latter one is more harmful than the former since it will also increase the interference with primary users.

Base stations usually take decisions based on an OR rule (if any of the nodes sense channel busy, BS decides a busy channel). This approach is very conservative in the sense that one single attacker or even a malfunctioning node can reduce the spectrum utilization. Another common approach is to decide according to majority voting. This resolves the spectrum underutilization problem but significantly increases the misdetection rate. Also, it becomes vulnerable when attackers collaboratively decide their attacking strategy.

3.1 Honest User Model

We assume that even an honest user cannot detect PU presence 100% accurately. We define false alarm as the probability of sensing presence of PU when it is actually not transmitting and we define misdetection as the probability of not sensing PU when it is operating. Let us assume that the probability of false alarm and misdetection rate of a user are P_{fa} and P_{md} respectively.

$$P_{fa} = P(u_i = 1|H_0), P_{md} = P(u_i = 0|H_1)$$

where H_0 and H_1 denote the channel idle and busy status and u_i represents the sensed result by user i.

As explained, honest users do not change their sensing results. Let us assume that v_i represents the report sent to BS by user i.

$$P(v_i = 1|u_i = 1) = 1, P(v_i = 0|u_i = 1) = 0$$
$$P(v_i = 0|u_i = 0) = 1, P(v_i = 1|u_i = 0) = 0$$

Accordingly, we can calculate the detection probability of an honest user using Equation 1.

$$\begin{aligned} P_d &= P(v_i = 0|H_0)P(H_0) + P(v_i = 1|H_1)P(H_1) \\ &= (1 - P_{fa})P_I + (1 - P_{md})P_B \end{aligned} \tag{1}$$

Here, P_d denotes the probability of accurate detection of channel status by any honest user and P_I and P_B denote the idle and busy rate of the channel respectively.

3.2 Attack Model

We assume that there exist at most M ($\alpha = M/N \leq 50\%$) attackers and the remaining users are honest, completely unaware of the presence of attackers. We do not consider the number of attackers more than 50% because it is not productive to study a network where a majority of nodes are attackers. We consider attackers devise their plan independently or collaboratively.

Independent Attack. Each attacker node changes its sensing result with probability P_{mal}. As a result, the detection probability of an attacker changes.

$$\begin{aligned} P_d^m = &[(1 - P_{mal})(1 - P_{fa}) + P_{mal}P_{fa}]P_I \\ &+ [(1 - P_{mal})(1 - P_{md}) + P_{mal}P_{md}]P_B \end{aligned} \tag{2}$$

Here, P_d^m denotes the detection probability of an attacker while working independently. Similarly, we can find the false alarm probability of an attacker (P_{fa}^m).

Collaborative Attack. In case of a collaborative SSDF attack, attackers exchange their sensing information and decide their response collaboratively. We study different collaboration strategies to see their impacts on decision making of BS. Let us assume that Q_d^m and Q_{fa}^m denote the detection probability and false alarm probability of attackers. To start with, we follow the same collaboration strategy used in [2]. Attackers follow 'L out of M' rule to decide their final decision where 'L' is determined according to [2]. In this case, Q_d^m and Q_{fa}^m will be

$$Q_d^m = \sum_{i=L}^{M} \binom{M}{i} (P_d^m)^i (1 - P_d^m)^{M-i}$$
$$Q_{fa}^m = \sum_{i=L}^{M} \binom{M}{i} (P_{fa}^m)^i (1 - P_{fa}^m)^{M-i} \quad (3)$$

Here, L is defined in [2]

$$L = min(M, \left\lceil \frac{M}{1+\beta} \right\rceil) \text{ where } \beta = \frac{\ln \frac{P_{fa}}{P_d}}{\ln \frac{1-P_d}{1-P_{fa}}}$$

The second attacking strategy we consider here is termed as 'Going Against MAjority (GAMA)'. Each attacker shares its true sensing result and in collaboration with other attackers decides against the majority sensing result with a certain probability. For example, if 2 attackers sense the channel idle and 1 user senses the channel busy, all 3 attackers report the busy status of the channel to the BS. The idea behind this attacking strategy is that sensing results of majority nodes may reflect the actual channel status. So, when the attackers collaborate, they change the sensing result of the majority and go against that. It may help them manipulate BS taking a wrong decision. In this case, $L = M/2 + 1$ and the collaborative detection probability will be

$$Q_d^m = \sum_{i=L}^{M} \binom{M}{i} (1 - P_d)^i (P_d)^{M-i}$$
$$Q_{fa}^m = \sum_{i=L}^{M} \binom{M}{i} (1 - P_{fa})^i (P_{fa})^{M-i} \quad (4)$$

Third, we also investigate the impact of collaboration among subgroups. In this approach, we assume that attackers exist in small groups, and each group changes their sensing result according to the first approach. Finally, one group is chosen randomly and all the attackers in that group report the same sensing result. This approach tries to attack in small groups without exposing all collaborators at a time.

4 Algorithm Design - Attackers vs BS

In this section, we discuss the viewpoints of attackers and BS and explain the defense mechanism taken by BS to defend against different attacking strategies. As stated in Section 3, attackers' detection rate varies with their strategy and is different from that of honest users. So, if the attackers can successfully manipulate the decision making process, detection rate will be significantly low, error rate in decision making will be high and spectrum utilization will be degraded.

From the attackers' point of view, the more error they make in decision making, the more successful they are. So, the most common attacking strategy is to

falsify about channel status in every time step and send it to BS. In collaborative attack, since attackers share their information, they may have better idea about the actual channel status and devise their attacking plan in a more effective way. The collaboration makes it easier to manipulate the BS decision mechanism than independent attack and increases their success rate. However, if the malicious users try to strengthen their attacks and continuously send false channel status, the pattern of their sensing report will be almost the same. In this way, their sensing history will be significantly different from honest users and will be easily identifiable. So, the best attacking strategy is to attack occasionally or try to behave like an honest user otherwise. In summary, attackers' success depends on attacking frequency (i.e. when to attack) and how long they can attack without being identified. Together, all attackers can follow the same plan and can make the decision making process more complicated.

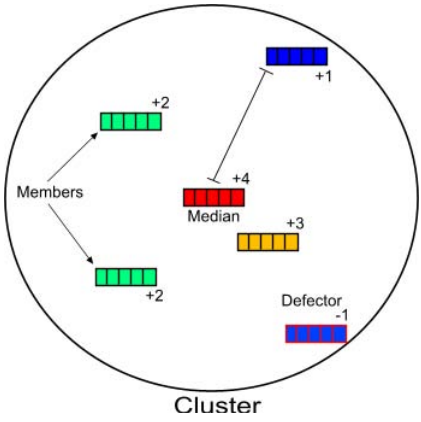

Fig. 1. Reputation Distribution

Now, from BS's point of view, its decision mechanism should be robust and capable of defending against any attacking strategy adopted by any number of malicious users. However, BS does not have any exact information about the attacking strategies or number of attackers. The only information available to BS is the sensing reports sent by users. So, the defense mechanism should be able to nullify (or at least reduce) the impact of collaboration of attackers, identify them and quarantine them from the decision process.

Accordingly, we design an adaptive reputation based clustering (ARC) algorithm to defend against both types of SSDF attack. The algorithm works against the intention and motivation of malicious users and tries to nullify their influence on the final decision. To reduce the impact of attackers, we create clusters so that nodes with similar sensing history will be in the same cluster. Then, each cluster has only one vote to cast and channel status is decided based on majority voting of clusters. The idea behind this defense mechanism is that if the attackers attack frequently, attackers and honest nodes will be in separate

clusters due to their different sensing reports. Also, collaboration of attackers will not help to increase the error rate since each cluster has only one vote.

The key to attackers' success is to avoid being in the same cluster and take control of the majority of the clusters. To handle these issues, we introduce distance weighted voting in a cluster and a feedback component in each node's reputation. Voting power of each node in the cluster is inversely proportional to its distance from the median of that cluster. Similarly, each node gets reputation inversely proportional to its distance from the median of that cluster. By distributing the reputation based on distance from the median, nodes are only impacted relative to their 'confidence' of that group (see Figure 1). Furthermore, from the next round, nodes' modified reputation is also used to cluster nodes in addition to sensing history. In this way, even if an attacker and an honest user incorrectly fall in the same cluster, attackers cannot establish their decision. Furthermore, as time goes, the distance between an honest user and an attacker will be amplified due to the joint consideration of reputation and sensing history.

5 Adaptive Reputation Based Clustering (ARC) Algorithm

In this section, we explain our adaptive reputation based clustering (ARC) algorithm in detail. The algorithm goes through a sequence of phases to reach the final decision. The phase sequences are illustrated in Figure 2. In the first phase, the BS collects the sensing result from all the nodes. BS maintains sensing history of all nodes for last d time steps. In the next phase, partitioning around medoids (PAM) algorithm is applied on the sensing reports to create k equal sized virtual clusters. In the third phase, each cluster makes its decision based on the response of each individual node and their relative distance from the median of that cluster. Then the final decision is made based on majority voting of clusters. The final result is then used to adjust the number of clusters and to update the reputation of all nodes.

One of the key features in our algorithm is how we reach the final decision and use that decision recursively to update the clustering. The information flow of our algorithm from one step to another in each time step is depicted in Figure 3. The BS considers the most recent d sensing reports of each node in addition to their reputation during cluster formation. To enable this recursive approach, we

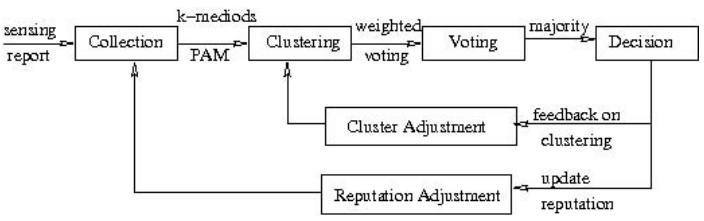

Fig. 2. Different Phases of the Algorithm

add an extra dimension to the sensing report of all nodes. This extra dimension represents the current reputation of that node (see Figure 3). So, each node provides a d+1 dimensional vector ($X_1 = [r_{1,1}, x_{1,1}, ... x_{d,1}]$) for cluster formation where the first dimension represents reputation and the remaining ones represent sensing report of last d time steps. Initially, all nodes are assigned the same reputation value.

Each cluster then finalizes its decision about channel status in a unique way. Only last round sensing report of each node in the cluster is considered. However, each response is weighted with an impact factor that is inversely proportional to the distance between the node and the median of that cluster. The impact factor of a node j at time t denoted by $I_j(t)$ is defined as

$$I_j(t) = \frac{1}{d_t(j, m_i)}$$

where m_i is the median of the cluster i and $d_t(j, m_i)$ denotes the distance between node j and median m_i of the same cluster at time t. Nodes closer to median have higher influence in decision making than the far ones. Accordingly, the cluster voting $v_i(t)$ at time t is determined by Equation 5.

$$v_i(t) = \frac{\sum_{j=1}^{N/k} I_j(t) * y_j(t)}{\sum_{j=1}^{N/k} I_j(t)} \tag{5}$$

Here, $y_j(t)$ is the sensing report of node j at time t which takes value from {0,1}.

After each cluster finalizes its decision, the BS makes the final decision $v(t)$ on the basis of majority voting among the valid clusters. If the reputation score of a cluster goes below a threshold, they cannot vote and all the nodes are marked as attackers. Therefore, $v(t) = \lceil 2 * \sum_{i=1}^{k} v_i(t)/k \rceil$.

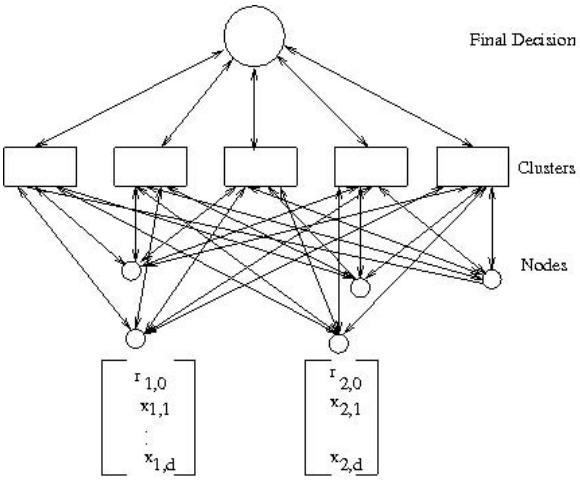

Fig. 3. Cluster Voting and Reputation Propagation

At the end of every time step, the base station updates number of clusters and reputation of all nodes according to the algorithm. The final result is propagated back to the clusters, and then to the individual nodes. If the final decision matches with the cluster decision, the cluster gets a positive feedback, and it gets negative feedback otherwise. Similarly, if a node's decision matches with its cluster decision, it gets positive feedback while it receives negative feedback for a mismatch. Each node's reputation is then adjusted according to Equation 6.

$$r_j = r_j + \Pi(v_i(t), v(t)) * \frac{\sum_{j=1}^{N/k} \Pi(v_i(t), y_j(t)) * I_j(t)}{\sum_{j=1}^{N/k} I_j(t)} \tag{6}$$

where r_j denotes the reputation of node $jand\Pi(a, b)$ is an indicator function that returns 1 if a equals b, it returns -1 otherwise.

The final result is also used to adjust the number of clusters. Initially, we start with 5 clusters with 5 random medoids. After each step, if all clusters pass the validation (i.e. reputation score exceeds threshold ϵ), we increment the number of clusters and continue the same process. Otherwise, we remove all the nodes in the cluster that fails the test.

6 Results

In this section we discuss results from the implementation of our proposed method, specifically comparing its effectiveness against a previously proposed method in [2]. We compare the two across both independent and collaborative attacks, as well as various probabilities of attack under a range of sensing conditions.

For each test, the methods are run over the same number of time steps, in this case 80 frames. For each time frame, the methods must produce a final

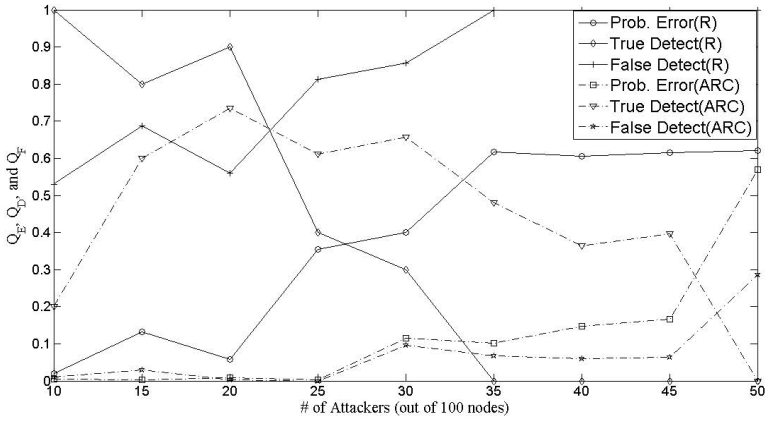

Fig. 4. Q_E, Q_D, Q_F with varying number of attackers (Collaborative SSDF Attack)

hypothesis, which is compared against actual transmission state of the primary user to determine the method's probability of error (Q_E). Rates for the correct detection of attacking nodes (Q_D) and the incorrect detection of honest users as attackers (Q_F) are also reported at the end of the test. Each test is then repeated 10 times with an average of the values displayed in the graphs. A test consists of randomly generated reports for each secondary user, adhering to labeled probability distributions. For validation test, we consider $\epsilon = 0.5$.

6.1 Collaborative Attack

First, we tested each method against a collaborative byzantine attack (see Figure 4), where the number of malicious users range from 10 to 50 out of 100 total secondary users. The byzantine attackers utilize the decision-making algorithm defined in [2]. Malicious users attack with $P_{mal} = 1$. Sensing probabilities for correctly detecting a signal and falsely detecting a signal were set to $P_d = 0.9$ and $P_{fa} = 0.1$ respectively.

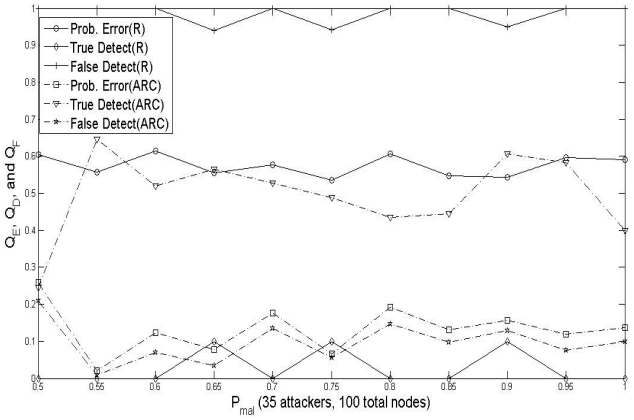

Fig. 5. Q_E, Q_D, Q_F with varying attacking probability (Collaborative SSDF Attack)

Our proposed method outperforms consistently with respect to (Q_E) showing a markedly decreased error rate until roughly 50% of the population becomes attackers. Once the population contains a majority of malicious users, it is impossible for any sensing strategy to sustain an error rate under 50%. The base stations are incapable of distinguishing between honest users and attackers. They can only resort to a blind guess for each sensing round. The Rawat method shows a high Q_D initially but quickly diminishes after 20% of nodes are attackers. At approximately the same attacker concentration, our method exceeds and maintains a marked increase in identifying attackers. Conversely, the Rawat method begins with a significant false detection rate (Q_F) while our method minimizes this rate across the entire range of attackers. Maintaining a low misdetection rate

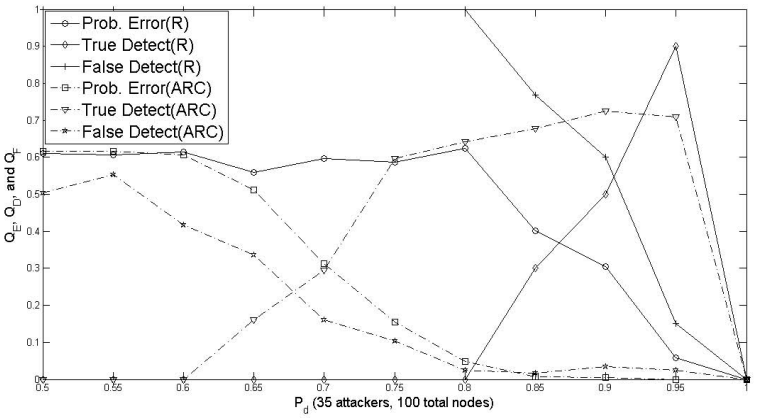

Fig. 6. Q_E, Q_D, Q_F with varying detection probability (Collaborative SSDF Attack)

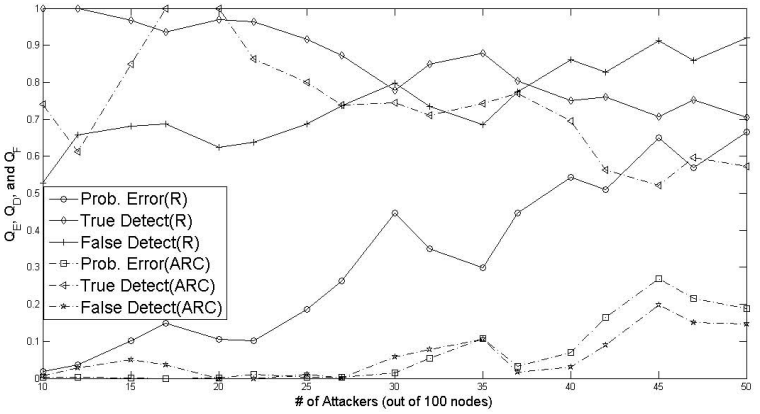

Fig. 7. Q_E, Q_D, Q_F with varying detection probability (Subgroup SSDF Attack)

allows our method to maximize honest user reports and mitigate the impact of attackers even under heavy attacks. A second set of measurements observed the impact of collaborating malicious users when varying their probability of attack. Malicious users can utilize this technique to escape detection from high dimensional clustering methods. In Figure 5, attackers produce on average less than 20% error rates while the Rawat method sustains significant errors. Regardless of attacking rate, our method consistently identifies 50% of the attackers. The Rawat method exhibits an unusually high attacker misdetection rate, which likely leads to the high error rate.

Depending on environmental conditions, the achievable sensing rates of primary user signals can vary dramatically. The next test looks at consequences of

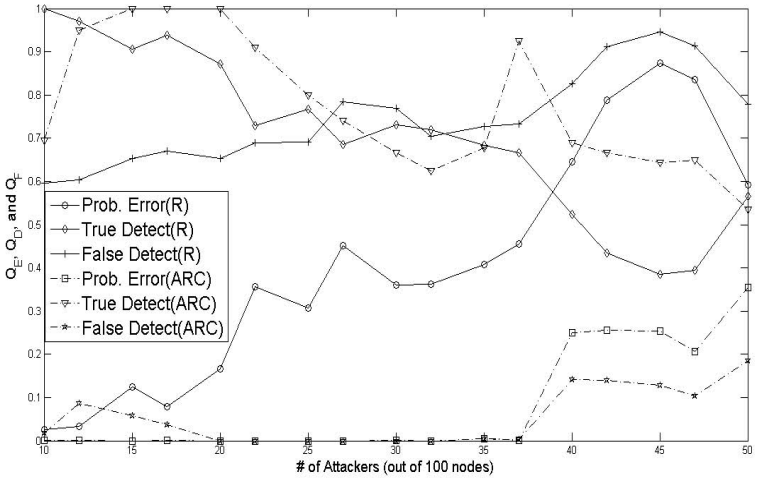

Fig. 8. Q_E, Q_D, Q_F with varying detection probability (GAMA SSDF Attack)

variable sensor accuracy (see Figure 6). Here, 35 collaborating malicious users attack during each sensing frame, and we can see the impact these sensing conditions have on the overall effectiveness of a byzantine attack. Both methods begin with relatively high error rates, as the sensing reports of honest users resemble that of attackers due to the inaccurate sensor readings. Once sensing errors fall below 65%, our proposed method shows a linear decrease in the Hypothesis error rate. The Rawat method takes significantly longer, approximately 80% detection rates, before error rates begin to decline.

We also test our algorithm in case of subgroup collaborative attack (see Figure 7). As the number of attacker increases, Q_E increases slightly in our algorithm while Q_E reaches almost 40% in the reputation method. As expected, both their true detection and false detection rate is high. On the other hand, Q_D is about 65% and Q_F is almost negligible in our algorithm.

We find interesting results for attackers with GAMA strategy. In case of our algorithm, Q_E is 0 and only increases when the number of attackers exceeds 37. On the other hand, Q_E increases almost linearly with the number of attackers in reputation based method. We get similar results in true and false detection rate. The results are plotted in Figure 8.

6.2 Independent Attack

In the next step, we compare the performance of our algorithm with reputation based scheme in [2] for independent SSDF attacks. In this attack, attackers do not collaborate to exchange their reports. Each attacker works independently to maximize its goal. Figure 9 shows the error rate of two algorithms with varying number of attackers. We keep the attacking probability $P_{mal} = 1$. Also, probabilities for true and false detection of a signal are set to $P_d = 0.9$ and $P_{fa} = 0.1$.

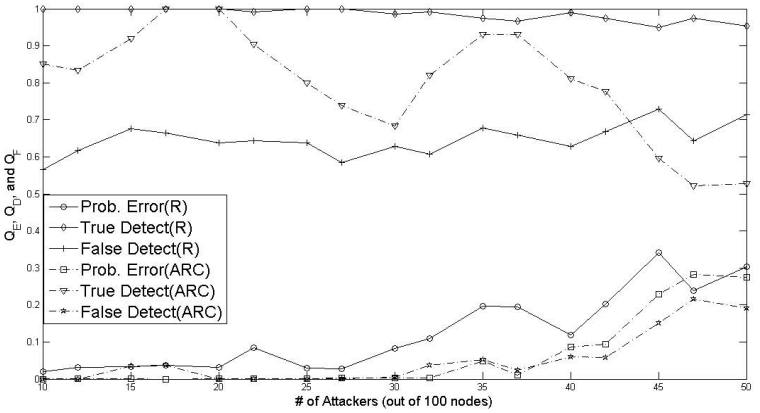

Fig. 9. Q_E, Q_D, Q_F with varying number of attackers (Independent SSDF Attack)

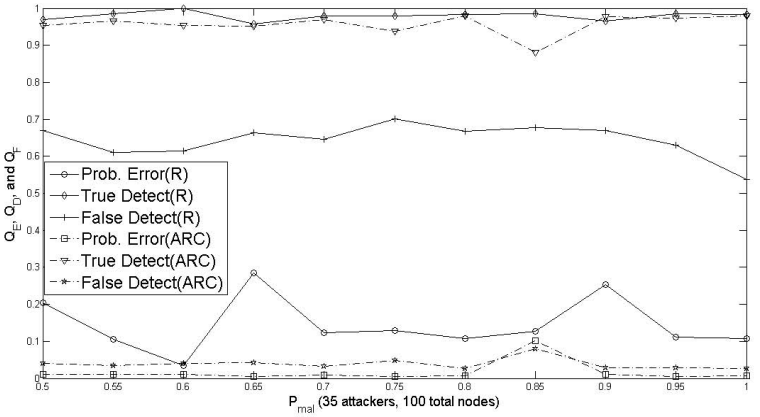

Fig. 10. Q_E, Q_D, Q_F with varying attacking probability (Independent SSDF Attack)

Our algorithm performs better up to 45 attackers and then slightly degrades its performance over their algorithm. On the other hand, our algorithm performs moderately to detect malicious attackers while their algorithm consistently identifies attackers with high precision. However, their algorithm eliminates a large number of honest users incorrectly. Figure 9 shows that about 40% honest users are miss identified as attacker. On the other hand, false detection rate of our algorithm is almost negligible. Although the reputation based algorithm performs better in detecting attacker than our algorithm, they misidentified a large number of honest users as attackers making their algorithm less effective.

Similarly, we run the simulation for independent SSDF attacks with varying attacking probability. We vary the attacking probability from 0.5 to 1 and

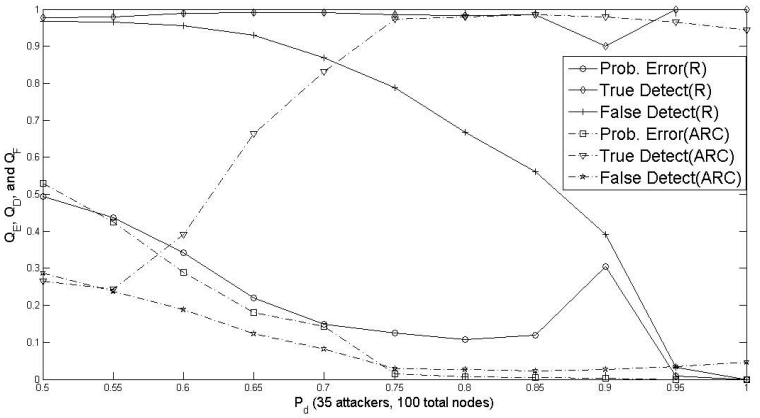

Fig. 11. Q_E, Q_D, Q_F with varying detection probability (Independent SSDF Attack)

plot Q_E, Q_E, and Q_F in 10 for our algorithm and reputation based algorithm proposed in [2]. Again, our algorithm performs better in decision making (see Figure 10). Error rate of our algorithm is almost negligible while their algorithm makes approximately 20% incorrect decisions about the channel status. The true attacker detection rate is almost the same for both algorithms. However, their algorithm constantly eliminates 60% of honest nodes as attackers for any attacking probability ranging between 0.5 and 1.0. On the other hand, our algorithm performs significantly better and keeps a false detection rate close to zero.

Next, we vary the detection probability of nodes from 0.5 to 1.0 and plot Q_E, Q_D and Q_F in Figure 11 for our algorithm and reputation based algorithm proposed in [2]. As usual, the error rate of our algorithm outperforms their algorithm. Also, our algorithm performs better in terms of misidentification of attackers. However, their algorithm identifies almost all attackers irrespective of the detection probability. On the other hand, our algorithm gradually increases the true detection rate with the increase of detection probability.

7 Conclusion

In this paper, we discussed one of the major security problems afflicting CRNs and proposed a reputation based clustering algorithm to defend against these attacks. We use reputation of nodes in addition to their sensing history to form clusters and then adjust reputation based on the cluster output. This recursive approach is tested in the presence of independent and collaborative spectrum sensing data falsification attacks. We compared the performance of our algorithm with existing approaches. With respect to current approaches, our algorithm significantly reduces the error rate in the final decision making process, thus increasing spectrum utilization. The false detection rate by our algorithm is almost negligible, while true attacker detection rate performs reasonably well.

However, the initial number of clusters plays an important role in overall performance of the algorithm. Also, it will be interesting to analyze the performance of the algorithm if attackers can overhear the honest users and decide accordingly. We will address these issues in future.

References

1. Li, H., Han, Z.: Catching Attackers for Collaborative Spectrum Sensing in Cognitive Radio Systems: An Abnormality Detection Approach. In: IEEE Symposium on New Frontiers in Dynamic Spectrum, pp. 1–12 (2010)
2. Rawat, A.S., Anand, P., Chen, H., Varshney, P.K.: Collaborative Spectrum Sensing in the Presence of Byzantine Attacks in Cognitive Radio Networks. IEEE Transactions on Signal Processing 59(2), 774–786 (2011)
3. Chen, R., Park, J.-M., Bian, K.: Robust Distributed Spectrum Sensing in Cognitive Radio Networks. In: INFOCOM: The 27th Conference on Computer Communications, pp. 1876–1884. IEEE (2008)
4. Clancy, T.C., Goergen, N.: Security in Cognitive Radio Networks: Threats and Mitigation. In: Cognitive Radio Oriented Wireless Networks and Communications (CrownCom), pp. 1–8 (2008)
5. Bian, K., Park, J.-M.J.: Security vulnerabilities in IEEE 802.22. In: Proceedings of the 4th Annual International Conference on Wireless Internet, WICON 2008, pp. 9:1–9:9 (2008)
6. Kaligineedi, P., Khabbazian, M., Bhargava, V.K.: Malicious User Detection in a Cognitive Radio Cooperative Sensing System. IEEE Transactions on Wireless Communications 9, 2488–2497 (2010)
7. Chen, R., Park, J.-M., Hou, Y.T., Reed, J.H.: Toward secure distributed spectrum sensing in cognitive radio networks. IEEE Communications Magazine 46, 50–55 (2008)
8. Wang, W., Li, H., Sun, Y., Han, Z.: CatchIt: Detect Malicious Nodes in Collaborative Spectrum Sensing. In: Global Telecommunications Conference, GLOBECOM 2009, pp. 1–6. IEEE (2009)
9. Akyildiz, I.F., Lee, W.-Y., Vuran, M.C., Mohanty, S.: A survey on spectrum management in cognitive radio networks. IEEE Communications Magazine 46, 40–48 (2008)
10. Wang, W., Li, H., Sun, Y., Han, Z.: Attack-proof collaborative spectrum sensing in cognitive radio networks. In: 43rd Annual Conference on Information Sciences and Systems (March 2009)