Secure Cooperative Spectrum Sensing Based on Sybil-Resilient Clustering

Jerry T. Chiang
Advanced Digital Sciences Center
Singapore
jerry.chiang@adsc.com.sg

Yih-Chun Hu
University of Illinois at Urbana–Champaign
Urbana, IL, USA
yihchun@illinois.edu

Pulkit Yadav
Indian Institute of Technology Delhi
New Delhi, India
cs1100234@cse.iitd.ac.in

***Abstract*—The Sybil attack has devastating effect on many distributed decision protocols such as voting: By disguising with multiple identities, a Sybil attacker can amplify his impact on the final outcome. A cooperative spectrum sensing protocol, which aims to enhance the sensing performance over the individual sensing protocols, is a kind of cooperative decision protocol. If not carefully designed, cooperative spectrum sensing can also be very vulnerable to the false-reporting Sybil attack, in which an attacker seeks to degrade the sensing performance by submitting multiple incorrect measurements using multiple identities.**

In this paper, we exploit the attacker's limited radio resources and propose a Sybil-resilient clustering mechanism, and adopt it as the basis of a secure cluster-based cooperative sensing protocol. We perform extensive simulation and show that naïve soft data combination and statistics-based false-report-resilient cooperative sensing protocols are susceptible to the Sybil attack; however, our proposed protocol can still provide reasonable sensing outcome despite the presence of Sybil attackers.

I. Introduction

In a Sybil attack, a single malicious network participant masquerades as multiple participants, each with a unique identity [1]. The Sybil attack has devastating effect on many cooperative decision schemes such as voting; *ballot stuffing* is a form of Sybil attack that is known to significantly affect the decision outcome.

Many Sybil defenses exist in the literature that equalize the impact of a Sybil attacker. The general approach is to weigh the input from each participant with the amount of resources (be they computation, monetary, or even social) he is able to invest to support his inputs. Since a Sybil attacker is still one single participant, his resources are not likely to be significantly more abundant than his peers, hence the weight associated with the Sybil attacker's input would effectively limit his impact. In order to disambiguate the number of adversaries, in this paper we refer to each physical Sybil network participant as a *Sybil attacker*, and refer to each identity controlled by the attacker as a *Sybil node*. Similarly, we call each physical network participant a *network user*, and each identity in the network a *network node*.

Mitola III and Maguire Jr. conceptualized the *cognitive radio* technology, in which an unlicensed radio (known as a *secondary user*) opportunistically accesses a wireless spectrum vacant from licensed users (called *primary users*) [2]. In order to accurately determine whether a spectrum is vacant from its licensed users, the research community proposed many *cooperative spectrum sensing* techniques, in which all secondary users pool together their individual sensing results in order to derive a consensus. Generally, a decision-maker, known as the *fusion center*, makes a binary decision by aggregating, directly or otherwise, the inputs from all secondary users.

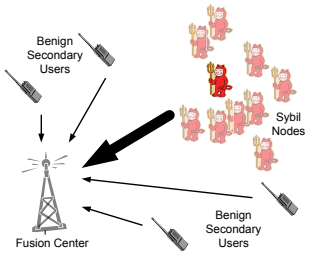


Fig. 1: Illustration of the false-reporting Sybil attack. The shaded adversaries, though with different identities, are controlled by the same Sybil attacker. The weight of the arrow indicates the weight of the input in making a decision at the fusion center.

Similar to other cooperative decision schemes, many cooperative spectrum sensing schemes are also susceptible to the Sybil attack. For example, in a K-out-of-N binary-decision scheme, if more than K out of N network nodes detect the presence of a primary user, the fusion center declares the channel occupied. A Sybil attacker that controls the majority of the N nodes can hence singlehandedly manipulate the sensing outcome: If $K \leq \lfloor \frac{N}{2} \rfloor$ (or if $K \geq \lceil \frac{N}{2} \rceil$), the Sybil nodes can all claim to (not) have detected primary presence and increase the false-alarm (missed-detection) probability to one. The attack strategy of reporting incorrect values in order to mislead the fusion center is known as the *false-reporting attack*. Fig. 1 illustrates the false-reporting Sybil attack.

In this paper, we exploit a particular Sybil defense mechanism: A Sybil attacker's radio resource stops it from digitizing a wireless band wider than $f_{\lim}$. In this paper:

1) We propose a centralized Sybil-resilient clustering algorithm, such that only a small fraction of cluster heads would be Sybil nodes; and
2) We adapt our clustering algorithm in a cluster-based cooperative sensing protocol, and show that even when a Sybil attacker is present, our proposed sensing protocol is resilient against the false-report attack.

978-1-4799-1353-4/13/$31.00 ©2013 IEEE

II. Attacker and System Models

A. Attacker Model

In this paper, we consider a legitimate (i.e. insider) Sybil attacker that seeks to degrade the spectrum sensing performance by incorrectly reporting each of his node's observed signal energy. The attacker pretends to be as many attacking nodes as possible. We assume the attacker has only one radio (RF) front end, and there exists a finite real number ρ, such that the attacker is not more resourceful than the equivalence of ρ participants.

In this paper, we focus on the *ability to digitize wireless spectrum* as the limited resource: We assume that an attacker cannot digitize more than $f_{\lim}$ of bandwidth. With today's technology, a radio can easily digitize 20 MHz of bandwidth (e.g. any commercial off-the-shelf 802.11 receiver); however, digitizing more than a gigahertz of bandwidth is uncommon even for ultra-wide band (UWB) communications.

We do not make any other assumptions on the radio resources of an attacker. Specifically, we do not restrict the attacker to using omni-directional antenna. We do not restrict the antenna type because antennas, in contrast to the entire RF front-end, are easy to replace. We assume each Sybil attacker knows the topology of the network. Moreover, if multiple Sybil attackers are present in the network, we assume the attackers could collude, thereby deteriorating the network performance by at least as much as without collusion.

B. System Model

We assume the network has access to M different channels, any pair of which are more than $f_{\lim}$ apart in frequency. For example, if $f_{\lim} = 1$ GHz, then the three ISM bands at 900 MHz, 2.4 GHz, and 5.9 GHz satisfy the above.

We then assume that at the time of cluster formation, the fusion center controls M radios, so that exactly one radio is tuned to each of the M channels. For a static network, cluster membership does not change significantly over time, and the fusion center can simply rent the radios on demand. For a mobile network, in which the secondary users should be re-clustered periodically, we can establish $M-1$ secondary users as "trust anchors" so that together with the fusion center there are M trustworthy radios. We assume all network participants are loosely synchronized in time. That is, there exists a $\tau \ll \infty$ such that the difference in local time between any pair of network participants is less than τ. The exact method for the fusion center to acquire M radios and the time synchronization technique are outside the scope of this paper.

We assume that the fusion center shares a secret key with each network node. This can be done by registering new nodes when they join the network. We do not assume the registration process to be Sybil-resilient: One Sybil attacker can register many Sybil nodes and obtain many keys. We let each secret key be associated to a unique spread-spectrum code, so that it also enables the fusion center to communicate with many network nodes simultaneously. Finally, we make the necessary assumption that the fusion center is benign.

III. Related Work

A. Sybil Attack and Defenses

Douceur proposed the Sybil attack and noted that a single Sybil attacker, by disguising as many nodes and controlling a substantial fraction of any distributed systems, can bypass the security redundancy built into the systems [1]. Since a Sybil attacker misbehaves by pretending to have multiple identities, a natural defense methodology is to attest whether each identity is associated with some *minimum resource level*. Douceur noted that, by simultaneously issuing resource challenges to all identities, a faulty entity that controls ρ times as much resources as a minimally capable entity can present itself as $\lfloor \rho \rfloor$ entities to another entity [1].

Newsome et al. examined the Sybil attack in a sensor network setting and noted that several network operations, from distributed storage to data aggregation to misbehavior detection, are susceptible to the Sybil attack [3]. These authors also presented several possible defenses particular to a wireless network: 1) radio testing; 2) registration and key predistribution; 3) position verification; and 4) remote code attestation. In our paper, we exploit radio testing to make sure that almost all of the cluster heads are benign; these cluster heads can then assist the fusion center and ensure that most clusters are free from the Sybil attack. We then use the Sybil-resilient clustering algorithm to construct a cooperative sensing protocol.

We do not consider using position verification since secure position verification protocols often require special hardware and are hence unfit for the purpose of cognitive radio networks. Simple position verification protocols, such as using the received signal strength indicators of neighboring network nodes to detect a Sybil attacker [4], [5], are susceptible to signal manipulation, especially by use of directional antennas.

B. Cooperative Spectrum Sensing

In a cognitive radio network, the network first determines whether a set of particular wireless bands are occupied by their respective primary users. If a wireless band is vacant, the cognitive radio network can allocate it to the secondary users until the primary user returns. While each secondary user may have spectrum sensing capability, due to fading and shadowing, the sensing outcome is significantly more reliable if several secondary users sense the spectrum and combine their observations [6], a process known as *fusion*.

Ma et al. derived the optimal binary fusion rule for an energy detection system in which observations are honestly reported and not correlated [7]: The test statistic is a weighted sum of the observed signal strength of the secondary users, where higher weight is given to stronger observations. However, this fusion rule is susceptible to the false-report attack, in which an attacker maliciously report an incorrect value: By always reporting a strong observation, a single attacker can increase the false alarm probability arbitrarily high.

The community has proposed several cooperative sensing schemes based on anomaly detection that penalize secondary users who make different decisions or statistically-different

observations from the majority [8], [9], [10], or from the history [11], or both [12]. These schemes are susceptible to the false-reporting Sybil attack: If a Sybil attacker controls the majority of the network nodes and false reports, the network may mistakenly penalize the benign users since the benign users may appear abnormal instead. Zeng et al. thus suggest that the network install trust anchors – secondary users that the fusion center knows to be trustworthy – in order to escape from the Sybil attack [10].

Sun et al. proposed a cluster-based cooperative sensing scheme [13] in which each secondary user in cluster i makes a binary decision, where 1 (0) indicates the presence (absence) of the primary user; the cluster uses K-out-of-N rule to fuse the decisions of its members and make a binary decision. The fusion center finally uses the "or" (1-out-of-N) rule to further fuse the cluster-level decisions. Since the proposed protocol is simply a hierarchy K-out-of-N fusion rule, it is still susceptible to the false-reporting Sybil attack. Min et al. proposed combining anomaly-detection with cluster-based cooperative sensing scheme [14]; the authors proposed first geographically partitioning secondary users into clusters, and then use the channel shadowing characteristics to eliminate false-reporting users. Although not an emphasis of their protocol, the protocol by Min et al. can be made Sybil-resilient if position-verification can be done economically.

IV. Proposed Protocol

In this section, we first present the intuition behind our proposed protocol, then detail our Sybil-resilient clustering scheme and incorporate it in a cooperative sensing scheme.

A. Intuition

The security of our proposed protocol relies heavily on the assumption that a Sybil attacker cannot simultaneously communicate on two channels separated by more than $f_{\lim}$ apart. Thus, at the cluster-formation phase, with a one-time cost, the fusion center exploits this assumption to prevent selecting Sybil nodes as cluster heads.

Once we are sure that almost all cluster heads are benign, we can bootstrap this assurance into the reporting mechanism, so that the Sybil nodes can report to at most a small fraction of the cluster heads. By constraining the number of Sybil nodes reporting in each round, the network can use a K-out-of-N fusion scheme to mitigate the attacker's impacts.

B. Sybil-Resilient Clustering

One crucial goal of a Sybil-resilient clustering scheme is to minimize the number of Sybil nodes acting as cluster heads. This is because a cluster head aggregates the results from its cluster members, makes a decision, and reports the decision to the fusion center; if the cluster head is malicious, he can easily manipulate the cluster decision, regardless of whether the network nodes in the cluster are benign.

If there exists inexpensive but secure and precise location verification protocols, such that the fusion center can quickly verify every network participant's position down to inches, then the fusion center can use geographical clustering (such as k-mean) and all Sybil nodes would likely be grouped into one cluster. As long as there are many clusters, conceding one cluster to the Sybil attacker would not significantly impact the cooperative sensing result.

Without considering location verification schemes, we resort to using radio-resource-based Sybil detection in this paper. In our proposed protocol, the fusion center first learns the network topology by probing the network. The learned topology is a corrupted version of the true topology: 1) benign-benign links are preserved; 2) the Sybil attacker can arbitrarily set up Sybil-Sybil links; and 3) the Sybil attacker can hide benign-Sybil links. The fusion center then selects a set of cluster head candidates from the nodes that have joined the network before the start of the candidate selection process. The candidates and their one-hop neighbors span the network; in particular, the fusion center selects itself to double as the cluster head of all secondary users without neighbors. The security of our proposed protocol does not depend on the selection algorithm itself; our evaluation uses the highest-connectivity cluster algorithm for concreteness.

After selecting the cluster head candidates, the fusion center initiates several rounds of Sybil detection and elimination. In each round, for each cluster head candidate, the fusion center selects one of the M channels on which to communicate with that candidate. The fusion center informs the candidates their respective channels to use; subsequently, the fusion center simultaneously sends a nonce to each candidate on the channel associated with that candidate. Each cluster head candidate then responds to the fusion center by sending back his received nonce. If any candidate cannot respond with the correct nonce, he is eliminated from the network and the fusion center selects a new set of cluster head candidates. For some predetermined threshold P, the fusion center repeats the process until no candidates are eliminated for P consecutive rounds. We refer to P as the *paranoia level*. The cluster head selection process can be summarized in Fig. 2.

C. Cluster-Based Cooperative Sensing

After the clustering phase, the network operates similar to existing cluster-based cooperative sensing scheme [13]: All cluster heads agree on a reporting time, and each cluster head selects a random channel on which its members report. Each member of a particular cluster sends his data to the cluster head on the predetermined channel at the predetermined time; the cluster head then makes a decision based on the received data and sends its decision to the fusion center; the fusion center finally broadcasts a unified binary decision, based on the inputs from all clusters, to all secondary users.

The decision of the secondary users, as well as the decision of the cluster heads, can be either soft or hard. A hard decision is a binary decision; and a soft decision is a quantized observation that carries more information than the hard decision. It is obvious that using soft decisions at each level would result in better sensing performance (i.e. higher detection and lower false-alarm probabilities) in benign environments; however,

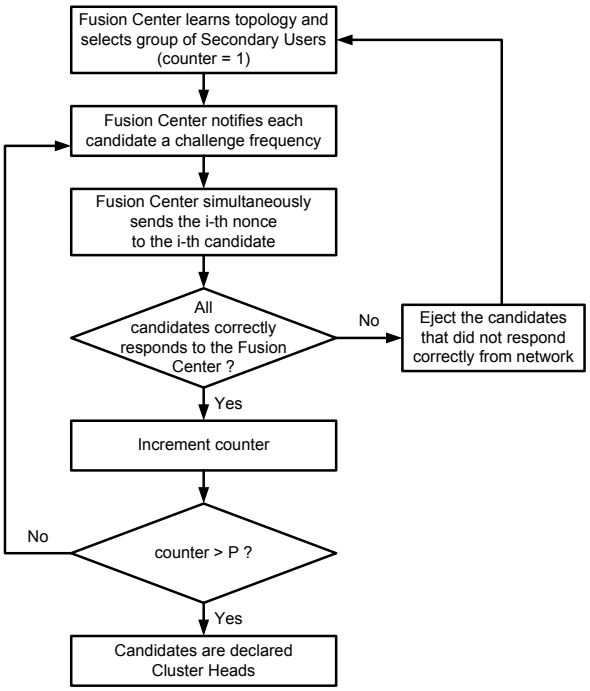


Fig. 2: Flowchart of the Sybil-resilient cluster head selection process

using hard decisions takes away an attacker's ability to amplify his impact by falsely reporting.

In this paper, we seek to reach a compromise between soft and hard decisions and propose that each secondary user reports to its cluster head its observed energy level without further quantizing the ADC output; The cluster head then uses an equal-weight aggregation decision rule to make a binary decision, and report the binary decision to the fusion center. The fusion center finally uses the majority rule to make a unified binary decision. Since all nodes are loosely time synchronized to within τ, the fusion center selects time t_r, and requires each node to report at its local time t_r, and pad its response so the entire transmission is τ in duration.

As a hierarchy design with Sybil-resilient clustering, the above design is natural: If a cluster is free of Sybil attackers, that cluster head can make an optimal decision; If a cluster has Sybil members and is suffering from false-reports, then that cluster's decision is not weighed more favorably against other clusters.

Additionally, the proposed design reduces the dependency on the channel condition between individual secondary user and the fusion center. Zhang and Letaief showed that when the channel between secondary users and the fusion center is subjected to data corruptions, sensing performance is limited by the probability of reporting errors [15]. By learning the network topology during the clustering process, the fusion center is able to connect most secondary users to cluster heads that are closer in proximity than the fusion center. The decrease in physical distance generally corresponds to better channel condition. Although the communication channel

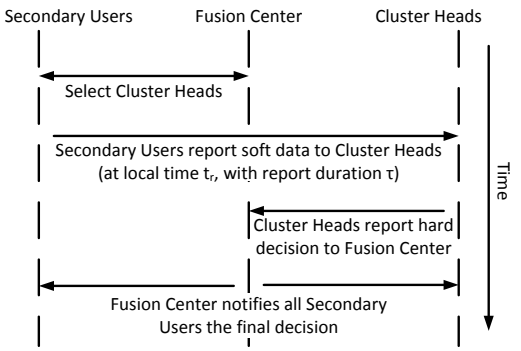


Fig. 3: Illustration of the interactions between different users in our proposed protocol

between each cluster head and the fusion center still has the same distance; however, the cluster head only sends one bit of hard decision over the channel, and more resources (such as error correction coding) can be deployed to protect that bit. We illustrate the order of interactions between different entities in our proposed protocol in Fig. 3.

To accommodate an open network membership, in our protocol, when a node joins the network after the beginning of the cluster head selection phase, it joins the cluster around it with the least number of nodes. If the new node cannot establish a reliable communication link with any other cluster heads, the new node joins the fusion center's cluster.

V. Evaluation

We evaluate the effectiveness of our protocol in this section. In particular, we are interested in understanding:

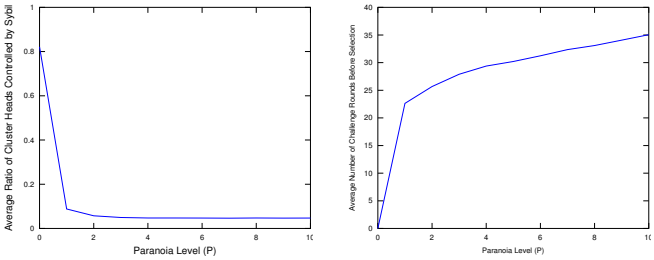
1) The relationship between the number of Sybil nodes as cluster heads, the number of challenges necessary during the selection process, and the level of paranoia; and
2) The ultimate accuracy and precision of the sensing result as compared to prior methods that are susceptible to the Sybil attack.

A. Cluster Head Selection Process

a) Methodology

We simulate the cluster head selection process using GNU Octave. We create a topology in which there are 100 benign secondary users and one Sybil attacker disguising as 500 Sybil nodes. These 600 network nodes are scattered in a 1-by-1 kilometer square area uniformly randomly. The attacking nodes pretend to have different locations and are not concentrated artificially into any particular region.

We assume that two nodes within 100 meters away from each other are connected with probability 1; and two nodes that are farther than 500 meters apart are connected with probability 0 (i.e. disconnected). If the distance between two nodes is larger than 100 meters but smaller than 500 meters, the two nodes are connected with probability linearly correlated to the received signal strength with path loss exponent $\alpha = 4$. That



(a) Ratio of cluster heads controlled by Sybil versus paranoia level.

(b) Number of rounds until termination of cluster head selection versus paranoia level.

Fig. 4: Results of Monte Carlo simulation on the behavior of proposed cluster head selection algorithm

is, if two nodes are 2×100 meters apart, the two nodes are connected with probability $\frac{1}{2^\alpha}$.

We implemented a greedy cluster head selection scheme. A secondary user is called "uncommitted" if it is not associated with any cluster heads. The fusion center repeatedly selects as cluster head candidate a secondary user that has the most positive (> 0) number of uncommitted neighbors. The fusion center itself serves as the cluster head of all remaining secondary users. Unlike the previous section, the greedy cluster head selection scheme does not always cluster the secondary users into the same number of clusters.

b) Simulation Results

Fig. 4 shows our simulation results. In Fig. 4(a), we observe that as the paranoia level increases, the number of Sybil cluster heads quickly falls from 80% of all selected cluster heads to a single cluster head.

The number of challenge rounds the fusion center takes to eliminate the Sybil nodes increases almost linearly with high paranoia level (see Fig. 4(b)). Specifically, when the paranoia level is low ($P = 1$), an increase in paranoia level makes our clustering algorithm run longer and eliminates more Sybil cluster heads; however, when the paranoia level passes a threshold ($P = 4$ in our simulation), further increase in paranoia level does not bring more benefits. This demonstrates that most Sybil cluster heads are eliminated at the very beginning, and for our topology, where the Sybil nodes outnumber benign users five-to-one, only around 30 rounds of challenges are needed to select a group of cluster heads that consists of very few (usually only one) Sybil nodes.

B. Accuracy and Precision of Our Cluster-Based Cooperative Sensing

c) Methodology

We evaluate our cooperative sensing protocol by running Monte Carlo simulation; for each scenario, we run 1000 trials with different topology, and plot the mean and standard variation of the detection probability versus the false-alarm probability. We reuse the parameters from last section and scatter 100 benign secondary users and 500 Sybil nodes across a 1-by-1 kilometer field. We place a primary user 20 kilometers

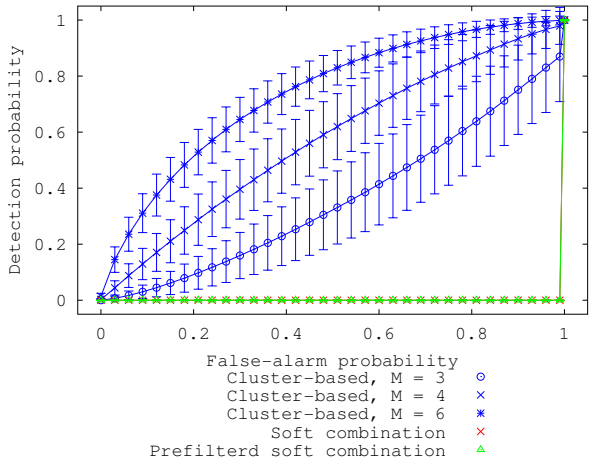


Fig. 5: ROC curves of cluster-based sensing with different number of available channels versus both naïve and prefiltered equal-weight soft decision aggregation. The error bar represents the standard deviation in the detection probability.

away from the center of the field. We model the channel between the primary user and the network nodes as a Rayleigh fading channel with a pathloss exponent of 4. We define the *primary signal-to-noise ratio (SNR$_p$)* to be the signal-to-noise ratio (SNR) of the primary signal at the center of the 1-by-1 kilometer field without fading. We vary the strength of the primary signal so that the primary signal-to-noise ratio varies between -13 dB to -7 dB. We also vary the number of channels from 3 to 10, but keep the paranoia level at 4.

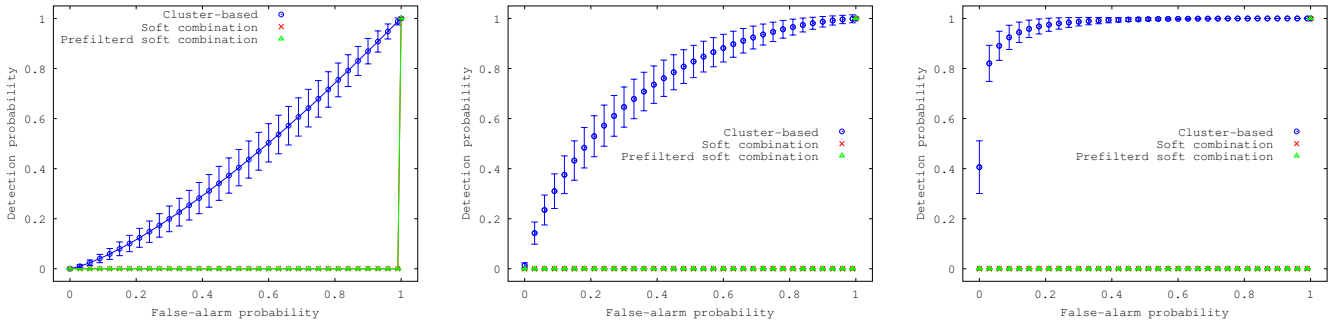
Each network node observes his received energy across 10 samples, sums the total received energy and reports it to his cluster head. The Sybil false-reporting attacker uses the most favorable channel (the channel that is used by the plurality of clusters), but only the Sybil nodes that are in clusters using the selected channel can report to their cluster heads.

We compare the performance of our proposed protocol with the performance of naïvely summing all observations and the performance of prefiltering away observations that are three inter-quartile-ranges higher than the third quartile and lower than the first quartile observations within each cluster [8].

d) Simulation Results

In Fig. 5, we fix the attack multiplier ($a = 2$) and the primary SNR (SNR$_p = -10\ dB$). we see that when we increase the number of channels the performance of our proposed protocol also becomes better. In particular, with only three channels our proposed protocol does not eliminate enough Sybil nodes and the false-alarm rate is high compared to the detection rate. When we increase the number of available channels to 6, we can enhance our protocol's performance to significantly above the performance of guessing. On the other hand, our results show that the soft combination schemes are very vulnerable to the Sybil false-reporting attack.

In Fig. 6, we fix the number of channels ($M = 6$) and the attack multiplier ($a = 2$). As the primary signal



(a) $\text{SNR}_p = -13$ dB (b) $\text{SNR}_p = -10$ dB (c) $\text{SNR}_p = -7$ dB

Fig. 6: ROC curves of cluster-based sensing versus both naïve and prefiltered equal-weight soft decision aggregation with different SNR_p. The error bar represents the standard deviation in the detection probability.

strength increases, the performance of our proposed protocol also becomes better. In particular, the average primary signal detection probability exceeds 0.8 while the average false-alarm probability stays below 0.05 when the primary SNR increases to -7 dB. On the other hand, even with the increase in signal strength, both naïve soft combination and prefiltered soft combination result in high false-alarm rates because of the Sybil false-reporting attack.

VI. Security Analysis and Discussion

In this section, we first consider that only one Sybil attacker is present. We separately analyze the security of our Sybil-resilient clustering algorithm and our cluster-based cooperative sensing scheme. We then consider the impact on the security of the system when multiple colluding Sybil attackers are present.

A. Number of Sybil Cluster Heads

Since the cluster head makes the decision for its cluster and report to the fusion center, if a cluster head is malicious, that cluster is hopeless in reporting a correct decision. Two important security concerns are thus: 1) the number of Sybil nodes selected as cluster heads; and 2) how quickly the fusion center can eliminate Sybil cluster head candidates.

Given a paranoia level of P, a set of two Sybil nodes can be both undetected and selected as cluster heads with probability $\left(\frac{1}{M}\right)^P$; it is even less likely that more Sybil nodes go undetected. In a simple example, if $M = 4$ and $P = 5$, the probability of retaining two Sybil cluster heads is less than 0.1%. We presented a thorough evaluation of the number of Sybil nodes selected as cluster heads in Section V-A.

After selecting a set of cluster head candidates, for each of the candidate, the fusion center can *uniformly randomly* assign it a frequency channel on which to receive the challenge nonce. If at the beginning of the round there are S cluster head candidates that are controlled by the Sybil attacker, our second security concern is related to how many misbehaving attackers will remain after one round of challenge. To minimize the attacker's loss from the attacker's perspective, after learning the channel assignments, the attacker picks the channel on which the maximum number of Sybil nodes are assigned, and answer all challenges sent over that channel.

This problem is well studied in networking and database research: There are S balls (Sybil candidates) and M bins (challenge channels), what is the expected number of balls in the fullest bin? In expectation,

$$R(S, M) = \frac{S}{M} + \Theta\left(\left(\frac{S \log(M)}{M}\right)^{1/2}\right)$$

Sybil candidates remain after one challenge round [16].

It is also possible to formulate a more systematic way of picking the challenge channels so that every pair of candidates must receive their nonces over two different channels in some rounds. This ensures that the network can eliminate all-but-one Sybil candidates.

We observe that there is another possible attack exploiting insecure clustering algorithms. In particular, if a node claims to have no neighbors, the fusion center must *not* automatically cluster that node into its own cluster, since this enables Sybil nodes to be easily promoted to cluster heads. Since the nodes can already communicate with the fusion center in order to register and join the network, we simply let the fusion center double as a cluster head in order to prevent this attack.

B. Number of Clusters that Contain Sybil Attackers

Although we can guarantee that almost all cluster heads are benign, these cluster heads are themselves subject to false-report attack from the Sybil nodes belonging to their respective clusters. Since in a Sybil-dominated population, it might not be possible to accurately assign blame, we do not try to identify or even isolate the Sybil nodes from the benign network nodes. Instead, we simply limit the attackers' ability to false-report.

In our proposed protocols, each cluster head randomly selects a different frequency channel on which to collect decisions. Since the entire network is loosely time-synchronized, the network enforces that each secondary user to report to its cluster head at roughly the same global time, so that all reports overlap in time.

Even if the Sybil nodes permeate the network, the radio resource constrains the Sybil nodes to focus their effort in only

a limited number of clusters. To minimize the attacker's loss from the attacker's perspective, after learning each cluster's channel choice, the attacker picks the channel chosen by the maximum number of clusters, and false-reports to each of the corresponding cluster heads. A simple substitution of variable shows that

$$R(C, M) = \frac{C}{M} + \Theta\left(\left(\frac{C\log(M)}{M}\right)^{1/2}\right)$$

clusters are impacted by the Sybil attack, where C is the number of clusters.

Since at the fusion center, the cluster heads effectively vote for the outcome. If M is large, then a majority-rule decision is not significantly affected by the Sybil attacker, and the unified sensing result retains its accuracy and precision.

C. Collusion Resilience

When multiple colluding Sybil attackers are present, their aggregated attack capability is also a multiple of a single attacker. During the cluster head selection phase, a set of D colluding Sybil attackers is functionally equivalent to a single Sybil attacker being able to receive and reply on D channels. Thus, the analysis is a simple variant that asks for the expected value of the sum of balls in the D fullest bins, which is readily bounded by D times the number of balls in the fullest bin.

Moreover, it is possible to extend the optimal scheme to a design where every set of M candidates, in some rounds, receive their nonces over all M different channels: if $D < M$, then the D Sybil attackers can only control D cluster heads. If $D > M$, then radio-resource-based challenges cannot distinguish between D Sybil attackers and multiple benign users. In other words, the security of our proposed protocol is directly related to the number of available challenge channels.

Similarly, the number of clusters subject to the false-reporting Sybil attack also grows slightly sub-linearly with respect to the number of Sybil attackers. Since at least D of the clusters have Sybil cluster heads, and at least $\frac{CD}{M}$ of the clusters are subject to false-report attacks by the Sybil nodes, our proposed clustering scheme is not secure when $D + \frac{CD}{M} > \frac{C}{2}$, or $D > \frac{MC}{2(M+C)}$, which converges to $M/2$ from below as C approaches infinity.

VII. Conclusion

Cooperative sensing promises to offer significant performance enhancement over individual spectrum sensing. However, to the authors' knowledge, all existing secure cooperative sensing schemes that seeks to eliminate the adverse effects of false reports are susceptible to the Sybil attack, and may inadvertently penalize benign secondary users.

In this paper we propose a Sybil-resilient cluster-based cooperative spectrum sensing protocol that offers high detection probability and low false-alarm probability even in the presence of a false-reporting Sybil attacker. The intuition is to combine the reporting mechanism with radio resource testing, thereby limiting the number of clusters affected by the attacker. A simple voting scheme can then mitigate the attacker's impact on the sensing decision.

We perform extensive simulation and show that naïve soft data combination and statistics-based false-report-resilient cooperative sensing protocols are indeed susceptible to the Sybil attack; however, our proposed protocol can provide significantly better sensing outcome despite the presence of Sybil attackers.

References

[1] J. R. Douceur, "The sybil attack," in *Peer-to-Peer Systems*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2002, vol. 2429, pp. 251–260.

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

[3] J. Newsome, E. Shi, D. Song, and A. Perrig, "The sybil attack in sensor networks: analysis & defenses," in *Proceedings of the 3rd international symposium on Information processing in sensor networks (IPSN '04)*, 2004, pp. 259–268.

[4] M. Demirbas and Y. Song, "An rssi-based scheme for sybil attack detection in wireless sensor networks," in *Proceedings of the 2006 International Symposium on World of Wireless, Mobile and Multimedia Networks (WOWMOM '06)*, 2006, pp. 564–570.

[5] J. Yang, Y. Chen, and W. Trappe, "Detecting sybil attacks in wireless and sensor networks using cluster analysis," in *Proceedings of the Fifth IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS '08)*, Oct. 2008, pp. 834–839.

[6] A. Ghasemi and E. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proceedings of the First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN '05)*, Nov. 2005, pp. 131–136.

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

[8] P. Kaligineedi, M. Khabbazian, and V. Bhargava, "Secure cooperative sensing techniques for cognitive radio systems," in *IEEE International Conference on Communications (ICC), 2008*, May 2008, pp. 3406–3410.

[9] ——, "Malicious user detection in a cognitive radio cooperative sensing system," *IEEE Transactions on Wireless Communications*, vol. 9, no. 8, pp. 2488–2497, Aug. 2010.

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

[11] H. Li and Z. Han, "Catch me if you can: An abnormality detection approach for collaborative spectrum sensing in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3554–3565, Nov. 2010.

[12] T. Zhao and Y. Zhao, "A new cooperative detection technique with malicious user suppression," in *IEEE International Conference on Communications (ICC)*, Jun. 2009.

[13] C. Sun, W. Zhang, and K. Ben, "Cluster-based cooperative spectrum sensing in cognitive radio systems," in *Proceedings of the IEEE International Conference on Communications (ICC)*, Jun. 2007, pp. 2511–2515.

[14] A. Min, K. Shin, and X. Hu, "Attack-tolerant distributed sensing for dynamic spectrum access networks," in *the 17th IEEE International Conference on Network Protocols (ICNP '09)*, Oct. 2009, pp. 294–303.

[15] W. Zhang and K. Letaief, "Cooperative spectrum sensing with transmit and relay diversity in cognitive radio networks - [transaction letters]," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4761 –4766, Dec. 2008.

[16] M. Raab and A. Steger, ""balls into bins" - a simple and tight analysis," in *Proceedings of the Second International Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM '98)*, 1998, pp. 159–170.