Secure Cooperative Sensing in IEEE 802.22 WRANs Using Shadow Fading Correlation

Alexander W. Min, *Student Member, IEEE,* Kang G. Shin, *Fellow, IEEE,* and Xin Hu

Abstract—Cooperative (or distributed) sensing has been recognized as a viable means to enhance the incumbent signal detection by exploiting the diversity of sensors. However, it is challenging to secure such distributed sensing due mainly to the unique features of dynamic spectrum access networks—openness of low-layer protocol stacks in software-defined radio devices and the absence of interactions/coordination between primary and secondary devices. To meet this challenge, we propose an *attack-tolerant distributed sensing protocol* (ADSP) for DTV signal detection in IEEE 802.22 WRANs, under which sensors in close proximity are grouped as a cluster, and sensors within a cluster cooperate to safeguard the integrity of sensing. The heart of ADSP is a novel filter based on shadow-fading correlation, by which the fusion center cross-validates reports from the sensors to identify and penalize abnormal sensing reports. By realizing this correlation filter, ADSP significantly reduces the impact of an attack on the performance of distributed sensing, while incurring minimal processing and communication overheads. ADSP also guarantees the detectability requirements of 802.22 to be met even with the presence of sensing report manipulation attacks by scheduling sensing within the framework of sequential hypothesis testing. The efficacy of ADSP is validated on a realistic two-dimensional shadow-fading field. Our extensive simulation-based study shows that ADSP reduces the false-alarm rate by $99.2\,\%$ while achieving $97.4\,\%$ of maximum achievable detection rate, and meets the detection requirements of IEEE 802.22 in various attack scenarios.

Index Terms—Cognitive radio, cooperative sensing, shadowing correlation, attack tolerance, IEEE 802.22, sensing scheduling.

1 Introduction

Accurate sensing of spectrum condition is key to the opportunistic use of licensed spectrum bands in dynamic spectrum access (DSA) networks, thus mitigating the anticipated spectrum-scarcity problem. The goal of spectrum sensing is to accurately and reliably detect, in real time, the presence or absence of primary signals on a spectrum band. To achieve this goal, numerous sensing techniques and algorithms have been proposed, including physical-layer signal detection [2], [3], MAC-layer sensing scheduling and sensor selection [4], sensor mobility [5], and associated performance tradeoffs [6], to name a few.[1]

In particular, cooperative sensing [7], [8] has recently received considerable attention as a viable means to enhance the detection performance by exploiting spatial diversity in received signal strengths (RSSs) at spectrum sensors. However, reports from the sensors can be manipulated by attackers in various ways, such as primary signal emulation [9], [10] and sensing results falsification [11]. These sensing-targeted attacks can severely undermine the incumbent detection performance because the fusion rule for a final detection decision relies solely on the reported RSSs. Sensing-targeted attacks pose a significant threat as they can disrupt opportunistic spectrum access, the basic premise of DSA. We call these unique sensing-targeted attacks in DSA networks *sensing-disorder attacks*.

A sensing-disorder attack aims to obscure the existence/absence of a primary signal by manipulating the spectrum sensing information (e.g., measured RSSs) either by raising or lowering the signal strength. When no primary signal exists, attackers or compromised sensors can manipulate their reports (i.e., RSSs) to generate an illusion of a primary signal. For example, in the IEEE 802.22 wireless regional area networks (WRANs) [12], an attacker can report a fake sensing report to force all users in the entire cell (of radius up to $100\,$km) to immediately vacate the channel [13]. Once users in the cell vacate the channel, the attacker can freely use the channel without any interruption. When there is a primary signal, on the other hand, attackers can lower the RSSs to veil the presence of a primary signal, leading to an unacceptable level of interference to the primary users. In both cases, attackers mislead the fusion center, i.e., base station (BS), to make an incorrect decision on the presence/absence of a primary signal, wasting spectrum resources or causing unacceptable interference to the primary communications. Therefore, there is a clear incentive for attackers to launch the sensing-disorder attacks.

While the sensing-disorder attacks can be easily launched with the aid of programmable software-defined radio (SDR) devices, their detection is difficult. Unlike the ordinary Denial-of-Service (DoS) attacks that exhaust all the network resources, they can be easily mounted by using SDR devices, such as USRP [14] and

• *A preliminary version of this paper will appear in IEEE ICNP 2009 [1]. The authors are with the Department of Electrical Engineering and Computer Engineering, The University of Michigan, 2260 Hayward Street, Ann Arbor, MI 48109-2121. E-mail: {alexmin, kgshin, huxin}@eecs.umich.edu.*

Manuscript received October XX, 2009; revised XXXX XX, 200X.

1. In this paper, we use terms *secondary user* and *sensor* interchangeably as we focus on the secondary users' role as spectrum sensors.

Digital Object Indentifier 10.1109/TMC.2010.252

Sora [15]. These open-source SDR platforms can be an attractive target for attackers because of their accessibility of low-layer protocol stacks like PHY and MAC [16]. Detecting these attacks, however, is not an easy task. While secure mechanisms such as MAC-layer or crypto-based authentication work well in traditional wireless networks, lack of primary-secondary communications precludes their usage. Moreover, the detection of attacks is exacerbated by the volatile nature of wireless medium itself, which makes it hard to differentiate between legitimate and deliberately-manipulated sensing reports. We thus need to devise a mechanism that can protect cooperative sensing from the above-mentioned attacks.

In this paper, we propose an attack-tolerant cooperative sensing protocol for the IEEE 802.22 WRANs that filters out the abnormal sensing reports (caused by either adversaries or malfunctioning sensors) by exploiting shadow-fading correlation in RSSs. This RSS-based filtering is motivated by the fact that attackers cannot control the physical-layer signal propagation.

This paper makes several main contributions as follows.

- Proposal of a novel *correlation filter* for detection of abnormal sensing reports that (i) exploits *shadow-fading correlation* in RSSs without any additional communication, (ii) safeguards spectrum sensing against attacks that increase either the incumbent false-alarm (type-1) or mis-detection (type-2) rates, and (iii) minimizes processing and sensing overheads. Despite their importance, type-2 attacks have not been considered before.
- Introduction of cluster-based cooperative sensing to exploit shadowing correlation. Correlation between sensors, which is entailed by sensor clustering, is known to have a detrimental impact on incumbent detection performance [7], [8], [17]. Our evaluation study, however, shows that the proposed clustering does not incur any perceivable performance degradation even in a very low SNR environment. Therefore, the sensor clustering is an efficient and useful approach to the sensing-disorder attacks.
- Development of a new data fusion rule tailored to attack-tolerance. Specifically, we propose *weighted gain combining* (WGC) that adaptively assigns different weights to sensing reports according to their statistical significance based on the normal shadowing profile. As a result, it minimizes the influence of the unfiltered attacks (due to their small deviations) on a final decision, further improving attack-tolerance.
- Design of a sensing scheduling scheme that guarantees satisfaction of the detection requirements of 802.22 even in the presence of attacks, while minimizing the number of sensing rounds. Although ADSP significantly improves the attack-tolerance, our simulation results indicate that the detection requirements of 802.22 may not be satisfied with one-time sensing. To solve this problem, we propose an optimal stopping time for sensing scheduling using sequential hypothesis testing so as to meet the detectability requirements.
- In-depth evaluation of ADSP in a realistic two-dimensional shadow fading environments in IEEE 802.22 WRANs. Most previous work uses a simple but inaccurate one-dimensional model. Our simulation results show that the proposed filtering scheme successfully withstands the attacks by reducing the false-alarm rate up to 99.2% and achieving up to 97.4% of maximum achievable detection rate.

The remainder of this paper is organized as follows. Section 2 describes the system and attack models used in this paper. Section 3 presents our proposed approach for attack detection, and the generation of a realistic two-dimensional shadowing field. Section 4 details our approaches to filter design and data-fusion, and Section 5 proposes a sensing scheduling algorithm. Section 6 evaluates the performance of ADSP and Section 7 concludes the paper.

1.1 Related Work

The problem of ensuring the robustness in distributed sensing has been studied in [11], [18], [19]. Chen *et al.* [11] proposed a robust data-fusion scheme that dynamically adjusts the reputation of sensors based on the majority rule. Similarly, in the IEEE 802.22 standard draft, a voting rule [19] has been proposed for secure decision fusion. Kaligineedi *et al.* [18] presented a pre-filtering scheme based on a simple outlier method that filters out extremely low or high sensor reports. However, their method may not suitable for a very low SNR environment such as 802.22 WRANs where a final data-fusion decision is very sensitive to small deviations in RSSs. The defense against Primary User Emulation Attack (PUEA) has also been studied in [9], [10]. Chen *et al.* [9] proposed an RSS-based location verification scheme to detect a fake primary transmitter. This scheme, however, requires the deployment of a dense sensor network for estimating the location of a signal source, and thus, incurs a high system overhead. Anand *et al.* [10] analyzed the feasibility of PUEA and presented a lower-bound on the probability of a successful PUEA. However, they did not address the impact of PUEA on the performance of cooperative sensing.

The problem of enforcing/enticing secondary users to observe the spectrum etiquette has also been studied. Woyach *et al.* [20] studied how to entice secondary users to observe the spectrum etiquette by giving them incentives. In a similar context, Liu *et al.* [21] studied the problem of detecting unauthorized use of a licensed spectrum. They exploited the path-loss effect as a main criterion for detecting anomalous spectrum usage and presented a machine-learning approach for more general cases. In contrast, we focus on intelligent filtering of suspicious sensor reports.

In a broader context, our paper is related to work on secure data aggregation [22]–[24] and insider attack

detection [25] in wireless sensor networks. However, the problem considered in this paper differs from them in that it focuses on an important, realistic case where attackers manipulate the sensor reports to mislead the fusion center in making a final decision on detection of a primary signal.

2 System and Attack Model

We first describe the IEEE 802.22 WRANs and the signal propagation and sensing models to be used throughout the paper. We then introduce the data-fusion model, and finally, present the attack model.

2.1 IEEE 802.22 WRANs

We consider an IEEE 802.22 WRAN, an infrastructure-based cellular system where each cell consists of a BS and the associated end-users called *consumer premise equipments* (CPEs). The CPEs represent households in a rural area, and are thus stationary. The typical coverage of each 802.22 cell is 33 km (up to 100 km). The main goal of IEEE 802.22 WRANs is to provide broadband wireless access in rural areas by allowing opportunistic access of TV white spaces recently opened up by the FCC [26]. The BS, which we assume adversaries cannot compromise, schedules the sensing of channels and decides on the presence/absence of a primary signal in each channel based on the sensing reports from a set $\mathcal{C}$ of collaborating sensors. Among different types of primary users in TV bands, we focus on detecting DTV signals with 6 MHz channel bandwidth in the US. We consider an 802.22 cell located at the edge of the keep-out-radius (i.e., 150.3 km) of a TV transmitter, and the entire secondary network (or cell) lies within the detection range of the DTV signal.

2.2 Signal Propagation and Sensing Models

The received primary (DTV) signal strength at sensor (CPE) i can be expressed as the propagation model [27]:

$$P_i = P_o\Big(\frac{d_o}{d_i}\Big)^{\alpha} e^{X_i}, \quad \text{(Watt)} \tag{1}$$

where P_o is the signal strength at the primary transmitter, α the path-loss exponent, d_o the reference distance, and d_i the distance from the primary transmitter to the sensor i. Shadow fading is accounted for in e^{X_i} where $X_i \sim \mathcal{N}(0,\sigma^2)\ \forall i$. The log-normal shadow fading is often characterized by its dB-spread, σ_{dB}, which has the relationship $\sigma = 0.1\log_e(10)\sigma_{dB}$. We assume the energy detector for PHY-layer sensing which measures the power level over the wide 6 MHz-wide DTV channel, the effect of multi-path fading can be ignored [2], [3] as is commonly assumed in the literature [4], [21].[2]

The energy detector is widely used for its simple design and efficiency [2], [29]. Although the feature detector is more reliable, it takes much longer (e.g., 24 ms for the field-sync detector for ATSC) [3] because it looks for a specific signature of the primary signal that appears infrequently. The test statistic of the energy detector is an estimate of average RSS (including the noise power), and can be approximated as a Gaussian using the Central Limit Theorem (CLT) as [12]:

$$T_i \sim \begin{cases} \mathcal{N}\big(N_o, \frac{N_o^2}{M}\big) & \mathcal{H}_0 \text{ (no primary signal)} \\ \mathcal{N}\big(P_i+N_o, \frac{(P_i+N_o)^2}{M}\big) & \mathcal{H}_1 \text{ (primary signal exists),} \end{cases} \tag{2}$$

where P_i is the received power of a primary signal, N_o the noise power, and M the number of signal samples. We assume that sensors measure the entire 6 MHz DTV channel at the Nyquist rate for 1 ms, i.e., $M = 6\times 10^3$.

2.3 Data-Fusion Model

We consider data fusion as the rule for incumbent detection. While the decision fusion reduces the overhead in reporting the sensing results, it is difficult to thwart the sensing-disorder attacks since it only provides a binary value based on a local decision.

In fading channels, equal gain combining (EGC) is known to have near-optimal performance without requiring estimation of the channel gains. EGC has the following decision statistic:

$$T_\Sigma \triangleq \sum_{i=1}^{N_s} w_i\, T_i, \tag{3}$$

where T_i is the test statistic of the energy detector at sensor i, N_s is the number of collaborating sensors, and the sensors have an identical weight, i.e., $w_i = 1\,\forall i$. The decision threshold η to achieve the desired level of false-alarm probability Q_{FA}^* can be derived as [29]:

$$\eta = Q^{-1}(Q_{FA}^*)\,\frac{\sqrt{N_s}N_o}{\sqrt{M}} + N_s N_o, \tag{4}$$

where $Q(\cdot)$ is the well-known Q-function. The performance of EGC will be used as a baseline in evaluating the efficacy of the proposed scheme.

In order to achieve better attack-tolerance, we propose *weighted gain combining* (WGC) in ADSP that adjusts the weights $\{w_i\}_{i\in\mathcal{C}}$ so as to minimize the impact of attack mis-detection on the final decision.

2.4 Attack Model

2.4.1 Attack Scenarios and Types

Sensing can be disrupted as follows.

- A sensor is compromised, and then manipulates its sensing reports, i.e., raises or lowers RSSs.
- A sensor is malfunctioning or faulty, yielding readings that differ from the actual RSS.

A common consequence of the above two cases is that the sensing reports to the fusion center are distorted, thus increasing the probability for the fusion center to make a wrong decision. To solve this problem efficiently, we focus on the detection of any abnormal sensing report instead of pinpointing the actual cause of abnormality.

2. For signal-specific sensing techniques, e.g., FFT-based pilot sensing [28], the effect of multipath fading may not be ignored.

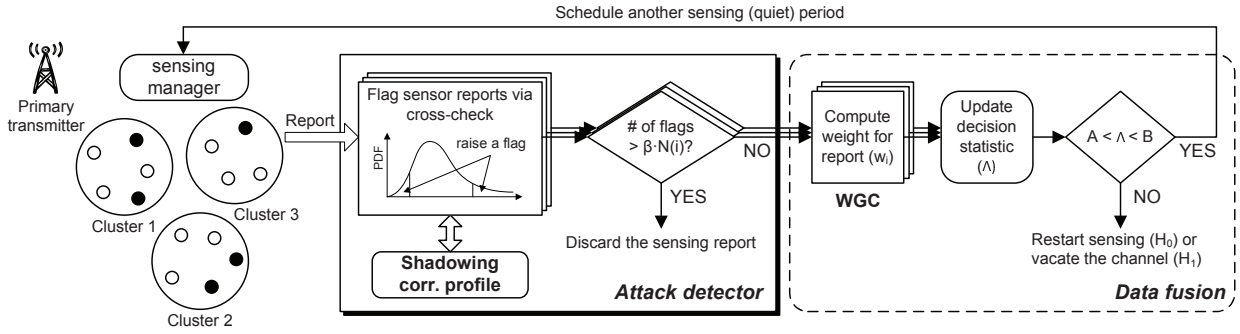


Fig. 1. *The ADSP framework*: Compromised (or malfunctioning) sensors might contaminate their sensing reports $\{R_i\}$. The attack detector filters out these contaminated sensing reports based on the shadowing correlation profile and then feeds the remaining ones to the fusion center. This process is repeated until the decision statistic at the fusion-center reaches one of the predefined thresholds, i.e., A and B, in order to guarantee satisfaction of the detection requirements of 802.22.

Note that another possible attack scenario is a primary user emulation attack (PUEA), as studied in [9], [10], [21]. However, PUEA is relatively easy to detect mainly because the attacker has only a coarse-grained control of RSSs at sensors since signals are broadcast. In the above two scenarios, however, the attacker has a fine-grained control of RSSs at individual sensors, making their detection harder. Therefore, we will focus on the above two attack scenarios.

We consider two types of attacks that can be mounted (caused) by attackers (faulty nodes):

- **Type-1 Attacks** increase the *false-positive* rate (classifying a non-primary signal or no signal as a primary signal) by raising RSSs, and
- **Type-2 Attacks** increase the *false-negative* rate (causing failure to detect a primary signal) by lowering RSSs.

We assume that the attackers know the presence/absence of a primary signal regardless of the decision made by the fusion center, and launch type-1 (type-2) attacks under $\mathcal{H}_0$ ($\mathcal{H}_1$); otherwise, attacks only serve to improve the incumbent detection performance.

2.4.2 Sensing Reports in the Presence of Attacks

Under the above model, a final sensing report to the fusion center can be expressed (in Watt) as:

$$R_i = \underbrace{P_i \cdot \mathbf{1}_{\{\mathcal{H}_1\}} + N_o + E_i}_{\text{energy detector output } (T_i)} + D_i \quad \forall i \in \mathcal{C}, \tag{5}$$

where $\mathbf{1}_{\{.\}}$ is an indicator function, T_i is the test statistic of the energy detector (in Eq. (2)) including the measurement error E_i, and $D_i \in \mathbb{R}$ is the deviation or *attack strength*, tampered with by a compromised (or faulty) sensor; $D_i = 0$ for normal sensors. Note that no loss of reporting packets is assumed, so we can focus on the detection of abnormal sensing reports.

3 The Proposed Approach

We now present the design rationale behind ADSP, its framework, and the methodology to generate a spatially-correlated shadow fading field.

3.1 Design Rationale

To maximize attack-tolerance and preserve the detection accuracy of data fusion, ADSP employs anomaly detection based on statistics. Specifically, ADSP exploits physical-layer signal propagation characteristics, or the spatial correlation in RSSs among neighboring sensors. The key insight behind ADSP is that, in shadow fading environments, RSSs at nearby sensors are likely to be highly correlated, which can be used to identify the manipulated sensing reports. The adversaries must be aggressive in raising or lowering the RSSs reported to the fusion center in order to influence the outcome of the final decision. However, any sensing report that significantly deviates from what is expected is deemed suspicious of being compromised or erroneous, and will hence be discarded or penalized by the fusion center in making a final decision. Adversaries must, therefore, lower their attack strength, reducing the chance for the fusion center to make a wrong decision; otherwise, they must risk getting caught by the detector. This way, the fusion center can achieve a high level of attack-tolerance, provided the majority of its neighbors are well-behaving.

3.2 ADSP Framework

ADSP resides at the fusion center (i.e., BS) and consists of the following three building blocks:

- **sensing manager** that manages sensor clusters and directs the sensors to report their readings at the end of each scheduled sensing period,
- **attack detector** that detects and discards (or penalizes) the abnormal sensing reports based on the pre-established shadowing correlation profile, and
- **decision maker** that determines the presence or absence of a primary signal based on the filtered sensing results using sequential hypothesis testing.

These three components closely interact with each other and form a robust distributed sensing system. Fig. 1 depicts the ADSP framework, which can be implemented at the 802.22 BS without requiring any modification to sensors (i.e., CPEs).

One important and unique feature of the attack detector is the ability to tolerate *both* type-1 and 2 attacks. This feature is attributed to the fact that the detector *cross-checks* the sensing reports and the assumption that majority of the sensors are well-behaving. As a result, under type-1(2) attacks, the sensing reports with relatively high (low) values are likely to be flagged by more of its neighboring sensors, thus making our scheme applicable regardless of the existence of a primary signal. This makes the system design simple and efficient, while achieving high attack-tolerance.

3.3 Generation of Spatially-Correlated Shadow Fading

To incorporate the spatially-correlated shadow fading in our analysis and simulation, we need a shadowing correlation model in which the statistics accurately reflect the real-world wireless shadowing environment. Note that one must rely on a model-based approach since measurement data for shadow fading is very scarce, and conducting a field test is too expensive to do. Gudmundson's model [30] is one of the most widely-used models in accounting for the shadowing correlation. However, it cannot capture spatial shadowing correlation, and hence, analyses based on this model might yield results that are significantly different from those in real-world wireless environments, as evidenced in both the theoretical study in [31] and empirical measurements in [32]. Recently, the authors of [33] proposed a statistical modeling approach to characterization of the spatial spectrum behavior of primary signals in the context of DSA networks.

Along the same line as in [33], we generate spatially-correlated shadow fading in a two-dimensional area by applying the convolution method proposed in [34]. We refer to the thus-generated data set as a *shadowing random field* $\mathbf{p}$ where $\mathbf{p}(x, y)$ represents the shadowing gain at a unit grid area, i.e., $\Delta\,\mathrm{m}\times\Delta\,\mathrm{m}$, centered at the coordinate $(x, y) \in \mathbb{R}^2$.

The shadowing random field $\mathbf{p}(\cdot, \cdot)$ is assumed to be an isotropic,[3] wide-sense stationary, and log-normally distributed random field with zero mean and exponentially-decaying spatial correlation. Then, the covariance between the two points $\boldsymbol{\theta}_i = (x_i, y_i)$ and $\boldsymbol{\theta}_j = (x_j, y_j)$ in $\mathbf{p}$ is given as:

$$\mathbb{E}\big[\mathbf{p}(\boldsymbol{\theta}_i), \mathbf{p}(\boldsymbol{\theta}_j)\big] = R_{\mathbf{p}}(d_{ij}) = \sigma^2 \cdot e^{-d_{ij}/D_{corr}}, \qquad (6)$$

where $d_{ij} = \|\mathbf{p}(\boldsymbol{\theta}_i) - \mathbf{p}(\boldsymbol{\theta}_j)\|$ is the Euclidean distance between the locations $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_j$, σ is the standard deviation of shadow fading, and D_{corr} is the decorrelation distance, which depends on local wireless environments (e.g., urban or suburban).[4]

Fig. 2(a) shows an example *shadowing random field* in a $2\,\mathrm{km}\times 2\,\mathrm{km}$ region, which clearly exhibits a strong spatial correlation in shadow fading. This is clearly shown in Fig. 2(b), which depicts the two-dimensional auto-correlation of the shadow fading. To demonstrate the accuracy of this method, Fig. 2(c) compares the one-dimensional auto-correlation function (ρ) of the random field against the Gudmundson's empirical model with with $\sigma_{dB} = 4.5\,\mathrm{dB}$ and $D_{corr} = 150\,\mathrm{m}$. The figure indicates that the synthetic data in the shadowing random field accurately emulates the real-world shadowing correlations. Note that our attack detection scheme in ADSP only requires the one-dimensional auto-correlation function of the shadowing field, which can be estimated by the service provider at the time of system deployment.

4 Detection of Abnormal Sensor Reports via Correlation Analysis

In this section, we formulate the anomaly-detection problem as a hypothesis testing, and present the design of a correlation-based filter. To further improve the attack-tolerance of ADSP, we propose a new data-fusion rule, called the *weighted gain combining* (WGC).

For cooperative sensing, the designated sensors (grouped in clusters) report their energy-detector's output along with their location information to the fusion center, at the end of each sensing period.[5] The location information is required to exploit the shadowing correlation in RSSs; it may be available at the fusion center since the sensors (i.e., CPEs) in 802.22 are stationary and 802.22 standard draft mandates the BS to have sensors' location information. Sensors can employ existing secure localization protocols (e.g., [36], [37]) to obtain accurate sensor location information.

4.1 Characterization of the Correlation in Sensing Reports

We first study the correlation structure of the sensing reports. A key observation is that the correlation structure of shadowing components $\{e^{X_i}\}$ is preserved in the sensing reports $\{R_i\}$ when there is no attack (or misbehavior), i.e., $D_i = 0$. To simplify the analysis, we further assume that the variance of the measurement error can be approximated as $\sigma_E^2 \approx \frac{N_o^2}{M}$ regardless of the presence/absence of a primary signal.[6]

Under the above conditions, and treating all the other terms in Eq. (1) (except e^{X_i} and E_i) as constants, we can express sensor i's report in Eq. (5) as:

$$R_i = C_1\, e^{X_i} + C_2 + E_i \quad \text{(Watt)}, \qquad (7)$$

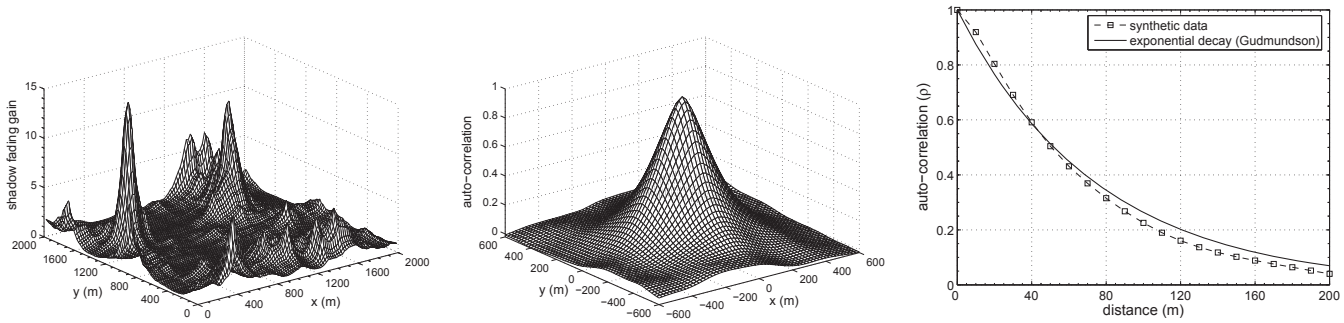
where $C_1 = P_o\big(d_o/d_i\big)^{\alpha}$, $C_2 = N_o$, and $E_i \sim \mathcal{N}(0, \frac{N_o^2}{M})$ is the measurement error of the energy detector. The correlation in shadowing component e^{X_i} does not change

3. Note that we do not consider the angular dependency in shadowing correlation for analytical tractability.

4. The measurement study in [35] indicates that a typical decorrelation distance is in the range of $120 - 200\,\mathrm{m}$ in suburban areas.

5. We consider two-dimensional sensor coordinates for simplicity, while the actual terrain profile is three-dimensional.

6. This assumption is reasonable in a very low SNR environment, e.g., $-20\,\mathrm{dB}$, where the average primary signal power is only about $1\,\%$ of the noise power, i.e., $\mathbb{E}[P_i] = 0.01\times\mathbb{E}[N_o]$.



(a) Shadowing random field (b) Two-dimensional auto-correlation function (c) Theoretical vs. synthetic data

Fig. 2. *Spatially-correlated shadowing random field* $\mathbf{p}(\cdot,\cdot)$: (a) An example of $\mathbf{p}(\cdot,\cdot)$ with exponentially-decaying spatial correlation, where the dB-spread and decorrelation distance are assumed to be $\sigma_{dB} = 4.5\,\text{dB}$ and $D_{corr} = 150\,\text{m}$, respectively, (b) Illustration of two-dimensional auto-correlation function of shadow fading, and (c) Comparison of auto-correlation function: Theoretical model (solid line) vs. synthetic data from a random field $\mathbf{p}(\cdot,\cdot)$ (dotted line).

when we add/multiply the same number to all of the shadowing components.

Moreover, the variance of measurement error is much smaller than that of a shadowing component, i.e., $\sigma_E^2 < \sigma_X^2$, since the number of samples M is sufficiently large even with a short sensing time, e.g., $M = 6\times10^3$ for the duration of 1 ms. So, the correlation in the received sensing reports $\{R_i\}$ almost preserves the correlation of the shadow fading e^{X_i}, i.e., $Corr(R_i, R_j) \approx Corr(e^{X_i}, e^{X_j})$.

4.2 Cluster-based Hypothesis Testing

While we exploit shadowing correlation for attack detection, the degree of correlation decreases exponentially with the distance between sensors. Therefore, we form *sensor clusters* among the sensors in close proximity such that sensors within the same cluster are highly correlated. A measurement study in [38] indicates that households in rural areas tend to be clustered, and thus, it is reasonable to assume that a BS can identify several sensor (i.e., CPE) clusters within its own cell of typical radius of 33 km. If such a sensor cluster exists, the BS can easily identify them based on their location information. If such sensor clusters do not exist, additional sensors can be deployed to form such sensor clusters.

Therefore, for each collaborative sensor $i \in \mathcal{C}$, the correlation-filter checks if the sensor exhibits a proper correlation behavior based on the following hypothesis testing for each of its neighbors within its cluster:

$$\mathcal{H}_0^a : Corr(R_i, R_j) = \rho(d_{ij}) \quad \forall j \in N(i), \tag{8}$$

where the neighbor set $N(i)$ is defined as the sensors belong to the same cluster of sensor i. As a result of this cross-checking, the number of flags raised by the neighboring sensors will be used as a filtering criterion (see Section 5.3 for details). We will henceforth focus on the analysis of shadowing correlation in the sensing reports.

4.3 Correlation Analysis for Filter Design

Although the shadowing correlation coefficient (ρ) is an obvious metric for the above hypothesis testing (i.e., Eq. (8)), it is not suitable for direct use in our problem because estimation of the correlation coefficient would require a sequence of samples; this can incur significant time and energy overheads for sensing, and can also deter the detection of returning primary users. Therefore, we detect a per-sample abnormal behavior by examining their *similarity* using the conditional probability distributions of the sensing reports. This is an alternative, but efficient approach since higher correlation entails greater similarity, which can be measured via a conditional distribution of sensor reports, as we will describe next.

In order to capture the similarity between sensing reports, we first derive the probability distribution of R_i, which is the sum of non-zero mean normal (i.e., E_i) and log-normal (i.e., e^{X_i}) random variables, as indicated in Eq. (7). To the best of our knowledge, there is no closed-form expression for such a distribution. However, a close examination of Eq. (7) implies that R_i can be approximated as a *shifted log-normal random variable*, i.e., the sum of a log-normal random variable and a constant.

Let us denote the sensing reports by a shifted log-normal random variable, i.e., $R_i = e^{Z_i} + N_o + C$ where $Z_i \sim \mathcal{N}(\mu_Z, \sigma_Z^2)$. From Eq. (7), we have the following approximation after simple manipulation:

$$e^{Z_i} + N_o + C \approx e^{X_i + \ln C_1} + N_o + E_i, \tag{9}$$

where $Z_i \sim \mathcal{N}(\mu_Z, \sigma_Z^2)$ and $X_i \sim \mathcal{N}(0, \sigma_X^2)$ with $\sigma_X = \sigma$. We set the constant $C = 4\,\sigma_E$ where $\sigma_E = \frac{N_o}{\sqrt{M}}$ so that the probability of the right-hand side of Eq. (9) become less than C is close to zero (i.e., $\approx 3 \times 10^{-5}$). This is important to preserve the non-negativeness of the log-normal random variable e^{Z_i}.

Then, we estimate the mean and variance of e^{Z_i} using a moment-matching method. By matching the mean and variance of both sides of Eq. (9), we have:

$$\hat{\sigma}_Z^2 = \log\left[\frac{C_1^2\,(e^{\sigma_X^2} - 1)\,e^{2\mu_X + \sigma_X^2} + \sigma_E^2}{(C_1\,e^{\mu_X + \sigma_X^2/2} + \mu_E + C)^2} + 1\right], \tag{10}$$

$$\hat{\mu}_Z = \log\left[\frac{C_1\,e^{\mu_X + \sigma_X^2/2} + \mu_E + C}{e^{\hat{\sigma}_Z^2/2}}\right]. \tag{11}$$

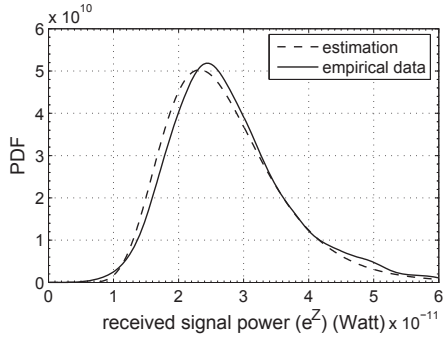


Fig. 3. *Estimation of the distribution of sensing reports as a shifted log-normal distribution*: The empirical data for sensing reports (solid line) obtained from the shadowing field can be accurately approximated as a log-normal distribution (dashed line).

The derivations of Eqs. (10) and (11) are straightforward, and thus omitted due to space limitation.

Fig. 3 shows an example of such approximation. While the figure indicates that the sensing reports can be accurately estimated by such a distribution, it becomes less accurate as the sensing duration T_S increases. Note, however, that we want to capture the correlation among sensors in a tractable form, not necessarily accurate approximation that only complicates the analysis without yielding a noticeable improvement in detection performance. The impact of the approximation error will be discussed in Section 6.

Based on Eqs. (9), (10), and (11), the p.d.f. of a sensor report can be expressed as:

$$f_R(r) = \frac{1}{(r-C)\,\sigma_Z\sqrt{2\pi}} \exp\left[-\frac{(\ln(r-C)-\mu_Z)^2}{2\,\sigma_Z^2}\right], \quad z \geq 0. \tag{12}$$

Recall that we are interested in studying the similarity of the sensing reports measured at nearby (thus spatially-correlated) sensors. To measure the similarity between sensing reports, we derive the conditional p.d.f. of sensor i's report R_i given the neighboring sensor j's report $R_j = r_j$ using Eq. (12) as:

$$f_{R_i|R_j}(r_i|r_j) = \frac{1}{(r_i-C)\,\sigma_{R_i|R_j}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln(r_i-C)-\mu_{Z_i|Z_j}}{\sigma_{Z_i|Z_j}}\right)^{1/2}\right], \tag{13}$$

where

$$\mu_{Z_i|Z_j} = \mu_{Z_i} + \rho_{ij}\frac{\sigma_{Z_i}}{\sigma_{Z_j}}\big[\ln(r_j-C)-\mu_{Z_j}\big], \tag{14}$$

$$\sigma_{Z_i|Z_j} = \sigma_{Z_i}\sqrt{1-\rho_{ij}^2(d_{ij})}. \tag{15}$$

Eq. (15) indicates that standard deviation $\sigma_{Z_i|Z_j}$ decreases as the correlation ρ_{ij} increases, and thus greater similarity between sensing reports.

Eqs. (13), (14), and (15) indicate that the conditional distribution of the sensing reports is also log-normally distributed. We thus set the lower and upper thresholds on the sensing reports based on conditional p.d.f. in Eq. (13), and then mark any outlier that resides outside of the thresholds. To set the thresholds, we first derive the cumulative distribution function (c.d.f.) of sensor i's report r_i, given sensor j's report r_j as:

$$F_{R_i|R_j}(x) = Pr(R_i \leq x \,|\, R_j = r_j) = \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\left[\frac{\ln(x-C)-\mu_{Z_i|Z_j}}{\sigma_{Z_i|Z_j}\sqrt{2}}\right], \quad x \geq 0, \tag{16}$$

where $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,\mathrm{d}t$.

Using Eq. (16), the thresholds $TH_{\{L,U\}}$ with a $100\times(1-\epsilon)\,\%$ confidence interval can be derived as:

$$TH_{\{L,U\}}(\epsilon) = \exp\left[\sqrt{2}\cdot\mathrm{erf}^{-1}\big(g(\epsilon)\big)\cdot\sigma_{Z_i|Z_j}+\mu_{Z_i|Z_j}\right]+C, \tag{17}$$

where

$$g(\epsilon) = \begin{cases} \epsilon - 1 & \text{for } TH_L \\ 1-\epsilon & \text{for } TH_U \end{cases} \quad 0 \leq \epsilon \leq 0.5, \tag{18}$$

where $\mu_{Z_i|Z_j}$ and $\sigma_{Z_i|Z_j}$ are the conditional mean and standard deviation in Eqs. (14) and (15), respectively.

Therefore, the null hypothesis $\mathcal{H}_0^a$, i.e., $Corr(R_i, R_j) = \rho(d_{ij})$, cannot be rejected if $r_i \in [TH_L, TH_U]$, as depicted in Fig. 4, whereas the attack false-alarm probability can be calculated as $P_{FA}^a = Pr(r_i < TH_L) + Pr(r_i > TH_U)$. Note that the thresholds are set differently for neighboring sensors, depending on their relative distance and measured RSSs.

Clearly, there is a tradeoff in determining the threshold parameter ϵ, i.e., the higher the threshold, the higher (lower) the false-alarm (mis-detection) rate for attack detection. The impact of the thresholds on the incumbent detection performance will be studied in Section 6.

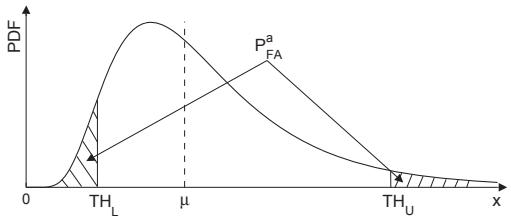


Fig. 4. *The correlation filter for anomaly detection*: Sensor i's report r_i will be flagged if it resides outside of the lower and upper thresholds, i.e., TH_L and TH_U.

4.4 The Proposed Data-Fusion Rule

While the correlation filter accurately detects RSS deviations in sensing reports, we observed that it often mis-detects small deviations (e.g., ≤ 0.3 dB). These small deviations can still influence the data-fusion results in a very low SNR environment due to the high sensitivity of the fusion decision to RSSs. Therefore, as a second line of defense, we propose a new data-fusion rule, namely *weighted gain combining* (WGC), to provide a better attack-tolerance to such small deviations. The idea is to assign different weights to the sensing reports according to their significance level based on the conditional c.d.f. in Eq. (16). This way, the mis-detected (unfiltered) attacks are highly likely to be assigned relatively

small weights compared to the legitimate sensing reports because of their lack of significance. Thus, the weights in WGC are defined as:

$$w_i \triangleq \frac{\sum_{j\in N_v(i)} w_{ij}}{|N_v(i)|} \text{ where } w_{ij} = 1 - 2\left|F_{R_i|R_j}(r_i \,|\, r_j) - 0.5\right|, \tag{19}$$

where $N_v(i)$ is the set of valid neighbors of sensor i whose reports passed the filter. The thus-obtained weights are used in calculating the decision statistic.

The simulation results (in Section 6) show that the WGC for data-fusion significantly reduces the attack false-alarm and mis-detection probabilities. However, the results also indicate that the detectability requirement of 802.22, i.e., $Q_{FA}, Q_{MD} \leq 0.1$, might not be met under weak attack strengths (e.g., $\leq$0.3 dB) as they cannot be easily differentiated from the normal sensing reports. To remedy this and meet the detectability requirements of 802.22 regardless of attack strengths, next we present sequential hypothesis testing framework for sensing scheduling.

5 The Proposed Data-Fusion Rule via Sequential Hypothesis Testing

In this section, we first formulate the incumbent detection problem as a sequential hypothesis testing subject to the detection requirements of 802.22, followed by the description of ADSP.

5.1 Attack-Tolerant Sensing Scheduling via SPRT

In ADSP, the BS schedules the sensing periods (stages) until it obtains a sufficient amount of information for making a final decision. Thus, the BS receives a sequence of measured test statistics from the sensors. This makes sequential detection suitable for our problem. In particular, among various sequential detection techniques, we adopt Wald's *Sequential Probability Ratio Test* (SPRT) [39] since it is optimal in the sense of minimizing the average number of observations, given bounded false-alarm probability Q_{FA} and mis-detection probability Q_{MD}. Therefore, by adopting the SPRT along with WGC, the BS can meet the detection requirement of 802.22 under the existence of malicious sensors by carefully designing the decision statistic as we discuss next.

5.1.1 Design of Decision Statistic

For SPRT, the distributions of the weighted test statistics of the sensors that passed the filter should be available to the BS under the both hypotheses. In practice, however, it is not feasible to derive a closed-form expression for such distributions. Therefore, instead of relying on the exact distributions of T_Σ, we exploit the threshold property of T_Σ as our main decision criterion.

Let ϑ_n denote a Bernoulli random variable defined as:

$$\vartheta_n \triangleq \begin{cases} 0 & \text{if } T_{\Sigma,n} \leq \eta_n \\ 1 & \text{if } T_{\Sigma,n} > \eta_n, \end{cases} \tag{20}$$

where $T_{\Sigma,n}$ is the sum of test statistics from the valid sensors, i.e., those who passed the filter, in sensing stage n, and η_n is the decision threshold, which depends on the number of valid sensing reports and the desired false-alarm probability Q^*_{FA} (see Eq. (4) in Section 2).

Our detection problem is thus a binary Gaussian classification problem where the observed test statistic $\vartheta_n \; \forall n$ belongs to one of two classes, $\mathcal{H}_0$ or $\mathcal{H}_1$, where:

$$\begin{aligned} \mathcal{H}_0 &: \vartheta \sim Bernoulli(\phi_0) \quad \text{(no primary signal)} \\ \mathcal{H}_1 &: \vartheta \sim Bernoulli(\phi_1) \quad \text{(primary signal exists)}, \end{aligned}$$

When there is no attack, the random variables ϕ_0 and ϕ_1 can be defined as:

$$\phi_0 \triangleq Pr(\vartheta_n = 1 \,|\, \mathcal{H}_0) = Q^*_{FA}, \tag{21}$$

$$\phi_1 \triangleq Pr(\vartheta_n = 1 \,|\, \mathcal{H}_1) = Q^*_D = 1 - Q^*_{MD}. \tag{22}$$

In this case, there should be a significant difference between ϕ_0 and ϕ_1, i.e., $\phi_1 \gg \phi_0$.[7] However, the actual achievable Q_{FA} and Q_D under attack scenarios can be higher and lower than the desired values, respectively, due to the performance deficiency of the filter. For example, Fig. 8 in Section 6 indicates that $\phi_1 - \phi_0$ can be as low as 0.08 under weak attacks, thus rendering it difficult for the BS to make a correct decision.

Therefore, ϕ_0 and ϕ_1 are the key parameters in our design of SPRT, which must be carefully set so as to meet the detection requirements of 802.22 under various attack scenarios. Thus, we set:

$$\phi_0^{'} = Q^*_{FA} + \varepsilon_0 \quad \text{and} \quad \phi_1^{'} = Q^*_D - \varepsilon_1, \tag{23}$$

where $\varepsilon_0, \varepsilon_1 \in \mathbb{R}$ with the constraint $\phi_1^{'} > \phi_0^{'}$.

We set the values of ε_0 and ε_1 empirically, based on the observations made in our simulation study. Note that an inaccurate values of $\phi_0^{'}$ and $\phi_1^{'}$ might introduce additional detection delay. However, as long as $\phi_0^{'}$ used by the BS is close to the true distribution under $\mathcal{H}_0$ than $\phi_1^{'}$, or vice versa, the SPRT will terminate with the desired level of detection probabilities.

5.1.2 Optimal Stopping Rule for Sensing Scheduling

In SPRT, a decision is made based on the observed sequence of test statistics, $\{\vartheta_n\}_{n=1}^N$, using the following rule:

$$\begin{aligned} \Lambda_N \geq B &\;\Rightarrow\; \text{accept } \mathcal{H}_1 \text{ (primary signal exists)} \\ \Lambda_N < A &\;\Rightarrow\; \text{accept } \mathcal{H}_0 \text{ (no primary signal)} \\ A \leq \Lambda_N < B &\;\Rightarrow\; \text{take another observation,} \end{aligned}$$

where A and B ($0 < A < B < \infty$) are the detection thresholds that depend on the desired values of Q_{FA} and Q_{MD}. The decision statistic Λ_N is the log-likelihood ratio based on N sequential observations (i.e., test statistics) $\vartheta_1, \ldots, \vartheta_N$ as:

$$\Lambda_N \triangleq \lambda(\vartheta_1, \ldots, \vartheta_N) = \ln \frac{Pr(\vartheta_1, \ldots, \vartheta_N \,|\, \mathcal{H}_1)}{Pr(\vartheta_1, \ldots, \vartheta_N \,|\, \mathcal{H}_0)}. \tag{24}$$

7. For example, the detection requirements of 802.22 is $\phi_1 - \phi_0 = Q^*_D - Q^*_{FA} = 0.9 - 0.1 = 0.8$.

Assuming that $\{\vartheta_n\}_{n=1}^N$ are i.i.d., Eq. (24) becomes:

$$\Lambda_N = \sum_{n=1}^{N} \lambda_n = \sum_{n=1}^{N} \ln \frac{Pr(\vartheta_n \,|\, \mathcal{H}_1)}{Pr(\vartheta_n \,|\, \mathcal{H}_0)} \tag{25}$$

Eq. (25) can be rewritten as:

$$\Lambda_N = s_N \ln \frac{\phi_1^{'}}{\phi_0^{'}} + (N - s_N) \ln \frac{1-\phi_1^{'}}{1-\phi_0^{'}}, \tag{26}$$

where $s_N = \sum_{n=1}^{N} \mathbf{1}_{\{\vartheta_n=1\}}$ denotes the number of sensing stages n where $\vartheta_n = 1$.

5.2 Performance Analysis

We now quantify the performance of our SPRT-based sensing scheduling in terms of (i) detection performance, i.e., Q_{FA} and Q_{MD}, and (ii) average number of sensing rounds needed to meet the detectability requirements.

In SPRT, the desired detection performance can be guaranteed by setting the decision thresholds A and B as follows. Let a^* and b^* denote the desired values of Q_{FA} and Q_{MD}, respectively. Then, the decision boundaries A and B are given by [39]:

$$A = \ln \frac{b^*}{1-a^*} \quad \text{and} \quad B = \ln \frac{1-b^*}{a^*}, \tag{27}$$

and the actual achievable error probabilities, denoted as a and b can only be slightly larger than the desired values a^* and b^*.

Recall that our objective is to meet the detection requirements of 802.22 even in the presence of malicious/mal-functioning sensors. Thus, we aim to minimize the number of times the spectrum needs to be sensed, with the decision thresholds derived from the target detection probabilities as shown in Eq. (27). Therefore, we are interested in analyzing the number of sensing rounds until a decision is made (i.e., either the boundary A or B is reached).

The average number of sensing rounds (also called quiet periods in 802.22) required to make a decision, denoted by $\mathbb{E}[N]$, can be computed as:

$$\mathbb{E}[N] = \mathbb{E}[\Lambda_N]^{-1} \times \mathbb{E}[\lambda \,|\, \mathcal{H}_k]. \tag{28}$$

First, using Eq. (26), the average value of λ under both hypotheses can be derived as:

$$\mathbb{E}[\lambda \,|\, \mathcal{H}_0] = \mathbb{E}\Big[\ln \frac{1-\phi_1^{'}}{1-\phi_0^{'}}\Big] \quad \text{and} \quad \mathbb{E}[\lambda \,|\, \mathcal{H}_1] = \mathbb{E}\Big[\ln \frac{\phi_1^{'}}{\phi_0^{'}}\Big] \tag{29}$$

Next, the average of Λ_N can be found as follows. Suppose $\mathcal{H}_0$ holds, then Λ_N will reach B (i.e., false alarm) with the desired false-alarm probability a^*; otherwise, it will reach A. Thus, using Eq. (27), we get:

$$\mathbb{E}[\Lambda_N \,|\, \mathcal{H}_0] = a^* \ln \frac{1-b^*}{a^*} + (1-a^*) \ln \frac{b^*}{1-a^*}. \tag{30}$$

Based on Eqs. (28), (29) and (30), we can derive the average required sensing rounds for decision-making as:

$$\mathbb{E}[N \,|\, \mathcal{H}_0] = \frac{a^* \ln \frac{1-b^*}{a^*} + (1-a^*) \ln \frac{b^*}{1-a^*}}{\mathbb{E}\Big[\ln \frac{1-\phi_1^{'}}{1-\phi_0^{'}}\Big]}. \tag{31}$$

Similarly, we can derive $\mathbb{E}[N \,|\, \mathcal{H}_1]$.

Algorithm 1 ATTACK-TOLERANT DISTRIBUTED SENSING WITH WEIGHTED GAIN COMBINING

Procedure **ADSP_WGC**($\{R_i\}, Q_{FA}, \beta$)

1: **while** each sensing round n **do**
2: $T_{\Sigma,n} \leftarrow 0$ /* Decision statistic */
3: $N_{normal} \leftarrow 0$ /* Number of normal sensing reports */
// Step 1. Check (ab)normality of sensing reports
4: **for** each sensor cluster $\mathcal{S}_k$ $k = 1, \ldots, N_c$ **do**
5: **for** each sensor $i \in \mathcal{S}_k$ **do**
6: $(\texttt{Isnormal}(i), \mathbf{w}_i) \leftarrow$ **CorrFilter**$(i, \{R_j\}_{j \in N(i)}, \beta)$
7: **end for**
8: **end for**
// Step 2. Update decision statistic
9: **for** each sensor cluster $\mathcal{S}_k$ $k = 1, \ldots, N_c$ **do**
10: **for** each sensor $i \in \mathcal{S}_k$ **do**
11: **if** $\texttt{Isnormal}(i) == 1$ **then**
12: Update w_i using Eq. (19)
13: $T_{\Sigma,n} \leftarrow T_{\Sigma,n} + w_i R_i$
14: $N_{normal} \leftarrow N_{normal} + 1$
15: **end if**
16: **end for**
17: **end for**
18: $T_{\Sigma,n} \leftarrow T_{\Sigma,n} \times N_{normal} / \sum w_i$ /* Normalization */
19: Calculate the decision threshold η_n using Eq. (4)
20: **if** $T_{\Sigma,n} > \eta_n$ **then**
21: $\Lambda_n \leftarrow \Lambda_{n-1} + \ln \frac{\phi_1^{'}}{\phi_0^{'}}$
22: **else**
23: $\Lambda_n \leftarrow \Lambda_{n-1} + \ln \frac{1-\phi_1^{'}}{1-\phi_0^{'}}$
24: **end if**
// Step 3. Make a final decision
25: **if** $\Lambda_n \geq B$ **then**
26: **return** 1 /* Primary exists */
27: **else if** $\Lambda_n < A$ **then**
28: **return** 0 /* Primary does not exists */
29: **else**
30: Schedule another sensing round and wait for the observation
31: **end if**
32: **end while**

5.3 Protocol Description

We now present the attack-tolerant distributed sensing protocol (ADSP) with the proposed WGC for final fusion. **Algorithm 1** describes the overall data-fusion procedure in ADSP. At the end of each sensing period, the fusion center collects sensing reports $\{R_i\}$ from the collaborating sensors, which are co-located in clusters. Then, the fusion center activates the correlation filter to selectively discard the abnormal sensing reports and updates the decision statistic Λ_n based on the remaining sensing reports with their weights. Note that the weights are assigned after the filtering process (line 11) so that the filtered abnormal sensing reports would have no influence on them. The fusion center repeats this process until the decision statistic reaches one of the predefined thresholds, i.e., A and B.

Algorithm 2 details the filtering procedure. For each sensing report, the filter counts the number of flags raised by its neighbors in the cluster. Then, the filter will return $\texttt{Isnormal} = 0$ if more than $\beta \in [0, 1]$ fraction of its

Algorithm 2 FILTERING ALGORITHM BASED ON CORRELATION ANALYSIS

Procedure **CorrFilter**$(i, \{R_j\}_{j\in N(i)}, \beta)$

1: `blacklist_counter`$(i) \leftarrow 0$ /* Initialize the counter */
2: $\mathbf{w}_i \leftarrow [0, \ldots, 0]^T$ /* Initialize the weight vector */
3: `Isnormal` $\leftarrow 1$ /* Initialize the indicator */
4: **for** each neighbor $j \in N(i)$ **do**
5: Update w_{ij} using Eq. (19)
6: **if** $Corr(R_i, R_j) \neq \rho(d_{ij})$ using Eq. (17) **then**
7: $++$ `blacklist_counter`(i)
8: **end if**
9: **end for**
10: **if** `blacklist_counter`$(i) > \beta \cdot N(i)$ **then**
11: `Isnormal` $\leftarrow 0$ /* Mark it as abnormal */
12: **end if**
13: **return** (`Isnormal`, $\mathbf{w}_i$)

TABLE 1
System parameters used in simulations

Parameter	Value	Comments
N_s	30	Number of collaborating sensors
N_c	5	Number of clusters
T_S	1 ms	Sensing duration
M	6×10^3	# of signal samples per sensing
σ_{dB}	4.5 dB	Shadow fading dB-spread
D_{corr}	150 m	Decorrelation distance
Δ	20 m	Dimension of a grid
N_o	-95.2 dBm	Noise power
γ	-20 dB	Signal-to-noise ratio (SNR)
Q^*_{FA}	0.01	Target false-alarm probability
β	0.34	Attack detection threshold

neighboring sensors mark it as abnormal, where β is a design parameter; otherwise, it will return `Isnormal`$=1$. The filter also returns the weight vector ($\mathbf{w}_i$) for future use in the final data-fusion process (i.e., WGC). The computational complexity of the algorithm is bounded by $\mathcal{O}(m^2)$ where m is the number of sensors in a cluster.

Remark: Although the key assumptions we have made, i.e., negligible multipath fading and presence of sensor clusters, are valid for the DTV signal detection in IEEE 802.22 WRANs, they might not always hold, depending on a given DSA environment, thus limiting the practicality of ADSP. For example, multipath fading in sensing reports may be negligible when sensors are mobile, or a primary signal is sensed with narrow channel bandwidth. However, relaxation of such assumptions may require a major modification to ADSP, and thus, extension of ADSP to such a challenging environment is left as our future work.

6 PERFORMANCE EVALUATION

The performance of ADSP is evaluated via MATLAB-based simulations. We first describe the simulation setup and then present the simulation results for both types of attacks under various attack scenarios.

6.1 Simulation Setup

To demonstrate the effectiveness of ADSP, we consider an IEEE 802.22 WRAN environment with a single DTV transmitter with 6 MHz bandwidth and multiple sensors (i.e., CPEs) located at the edge of the *keep-out radius* of 150.3 km from the DTV transmitter [2]. An 802.22 cell of radius 30 km is considered for our evaluation, and we generate a two-dimensional shadowing field (using the method discussed in Section 3.3) with a unit grid of $20 \times 20\,\mathrm{m}^2$ to emulate a realistic shadow fading environment in a cell. Throughout the simulation, we assume 5 sensor clusters located randomly within the cell, with 6 sensors in each cluster; the sensors are located in different grids, and the distances between sensors within a cluster range from $d_{min} = 20\,\mathrm{m}$ to $d_{max} = 20\sqrt{5}\,\mathrm{m}$, as shown in Fig. 5. We consider the attack scenario where a one-third of the sensors are malicious in each cluster. Table 1 lists the system parameters used in our simulation. Each simulation is run on 5×10^4 randomly-generated shadowing fields and their average values are taken as the performance measures.

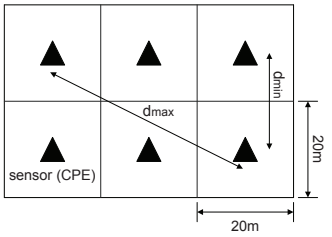


Fig. 5. *Sensor cluster*: An illustration of sensor cluster with 6 sensors in an 802.22 WRAN cell.

6.2 Impact of Sensor Clustering

While ADSP exploits shadowing correlation via sensor clustering, correlated sensor readings are, in general, known to degrade the detection performance as it limits diversity gain [7], [8], [17]. Therefore, we first study the effect of sensor clustering on detection performance to understand the efficiency vs. robustness tradeoff in ADSP. Fig. 6 compares the achieved incumbent detection probabilities (Q_D) with and without sensor clustering (i.e., sensors are randomly selected by the BS). As expected, cooperative sensing without clustering yields higher detection probability than with sensor clustering with -20 dB SNR. However, the performance gap decreases as more sensors are involved in cooperative sensing, e.g., sensing with 5 clusters achieves 94 % of that without clustering. Note that this performance with clustering gets even closer to that of random selection as the SNR increases. Therefore, we can conclude that sensor clustering is not critical to incumbent detection, while it provides an efficient means of attack detection.

6.3 Attack Detection Performance

As a first line of defense, the attack detector in ADSP must be able to correctly identify any abnormal sensors

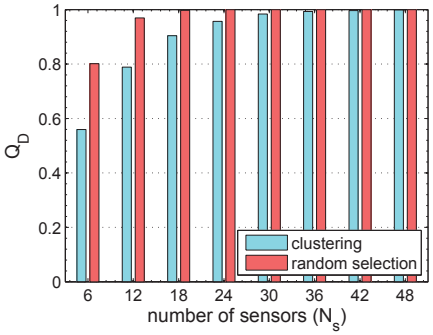
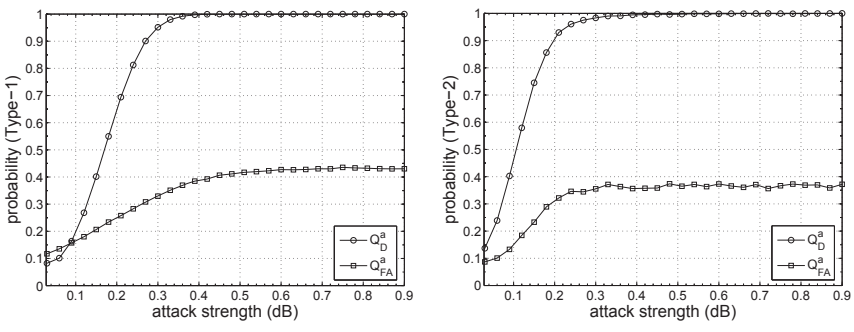


Fig. 6. *Impact of sensor clustering*: Sensor clustering with $N_c = 5$ achieves $94\,\%$ of the detection performance without clustering.



(a) Type-1 attack (b) Type-2 attack

Fig. 7. *Attack detection performance of the correlation filter*: The detection and false-alarm probabilities of our correlation filter increase with attack strengths under both types of attacks.

within each cluster and discard their reports before making a final decision. Fig. 7 shows the performance of our correlation-based filter under both types of attacks. The lower and upper thresholds (i.e., $TH_{\{L,U\}}$) for correlation filter is set using Eq. (17) with $99\,\%$ confidence interval, i.e., $\epsilon = 0.01$. The figures indicate that the attack detection rate, i.e., probability that a manipulated sensing report will be correctly filtered, increases with attack strength under both attack types. This is because the larger the deviation from the normal profile, the easier to identify them. However, the attack false-alarm rate also increases with the attack strength because the normal sensing reports will be mistakenly flagged more frequently by the manipulated sensing reports, and as a result, the normal sensing reports will be classified as attacks more frequently. The figures show that ADSP performs well against both types of attacks.

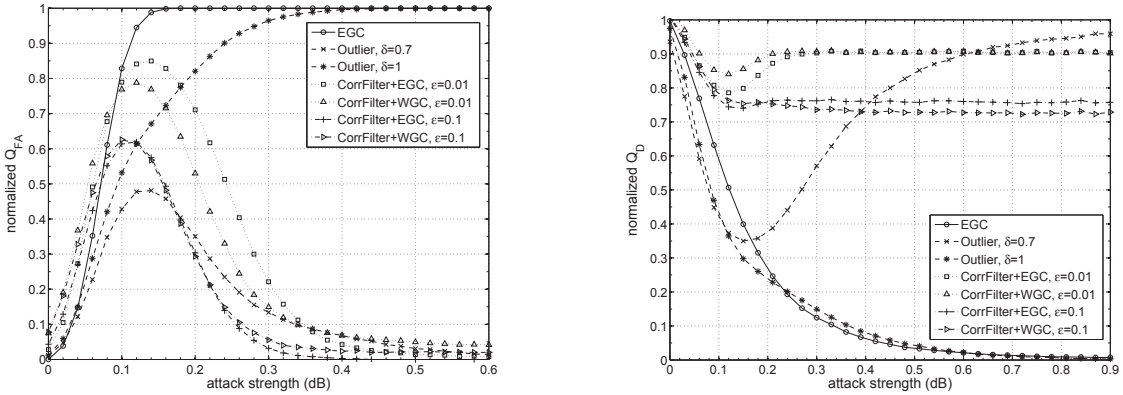
6.4 Attack-Tolerance for One-Time Sensing

We now demonstrate the robustness of ADSP to both type-1 and type-2 attacks for one-time sensing. Fig. 8 plots the incumbent false-alarm (Q_{FA}) and detection (Q_D) probabilities under type-1 and type-2 attacks, respectively. Note that Q_{FA} and Q_D are *normalized* with respect to the maximum achievable values in the absence of attacks. The figure shows that the correlation filter is efficient in mitigating the effect of attacks on incumbent detection performance, e.g., $99.2\,\%$ for type-1 and $97.4\,\%$ for type-2 attacks, thanks to its ability to accurately filter out manipulated sensing reports. By contrast, without ADSP (denoted by EGC in Fig. 8), Q_{FA} and Q_D rapidly converge to 1 and 0, respectively, as the attack strength increases, i.e., attacks have maximal influence on the data-fusion results.

We make the following four main observations. First, the performance of ADSP suffers in case of low attack strengths (e.g., $< 0.4\,$dB for type-1 attack). This is because they do not exhibit deviations significant enough to be detected (thus causing *under-filtering*), yet they affect data-fusion decisions. The proposed weighted gain combining (WGC) mitigates this performance deficiency for both types of attacks by adaptively adjusting sensing reports' weights based on their statistical significance. However, WGC performs as well as, or even worse than, EGC when the attack strength is either (i) extremely low so that most of attacks will not be filtered out or (ii) large enough so that most (or all) of attacks are filtered out, as can be seen in Fig. 8 with $\epsilon = 0.01$. This is because, in the first case, the unfiltered attacks will decrease the weights of the legitimate sensing reports, while sharing large weights among themselves. On the other hand, in the second case, the legitimate sensing reports with extreme values are likely to be assigned small weights despite their critical role in accurate detection of incumbents.

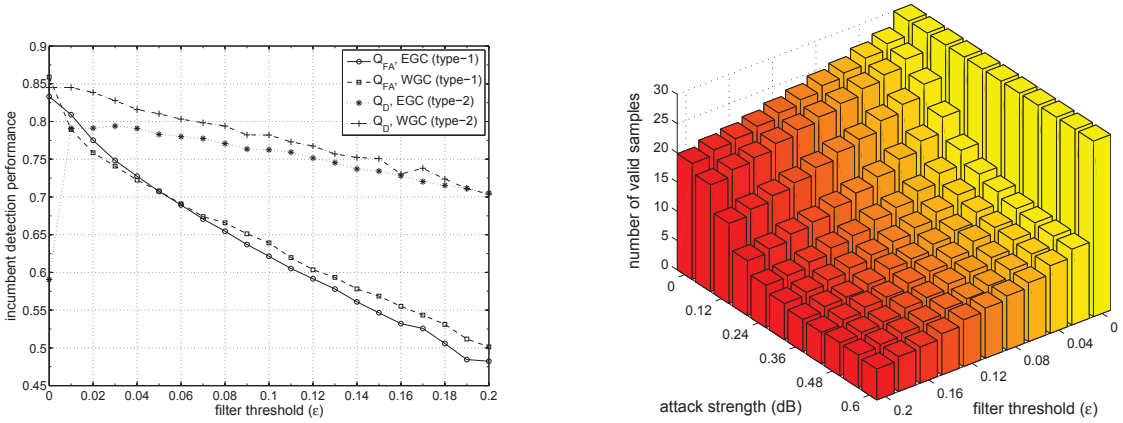
Second, ADSP outperforms the statistics-based filtering method proposed in [18] (denoted by Outlier in Fig. 8). The fusion center filters out the sensing reports outside the range $[\,e_1 - \delta \cdot e_{iqr}, e_3 + \delta \cdot e_{iqr}\,]$ where e_1 and e_3 represent the first and third quartile of the samples, respectively, and $e_{iqr} = e_3 - e_1$ is the interquartile range (see Eq. (4) in [18]). This method does not require sensor clustering, and thus, one might think that it performs well when the attack strength is strong enough to be easily detected as an outlier. However, the performance depends strongly on the filtering range, i.e., the choice of δ, the result of which varies with attack scenarios. For example, when $\delta = 0.7$, the performance suffers from *over-filtering* with a high attack mis-detection rate. On the other hand, when $\delta = 1$, the performance suffers from *under-filtering*, and as a result, Q_{FA} and Q_D converge to 1 and 0, respectively, even in the case of high attack strengths. In contrast, ADSP accurately detects the manipulated sensing reports by considering shadowing correlation.

Third, even in case of high attack strengths, ADSP does not completely eliminate the effects of attacks for the following reasons. First, the fixed threshold parameter ϵ does not work optimally for all attack strengths, thus causing either over- or under-filtering, both of which degrade the detection performance. The over-filtering caused by a large threshold value (e.g., $\epsilon = 0.1$) turned out to lower both Q_{FA} and Q_D, as shown in Fig. 8. Second, as a result of filtering, the fusion center



(a) Type-1 attack (b) Type-2 attack

Fig. 8. *Attack-tolerance of ADSP*: ADSP (a) minimizes the false-alarm probability by up to $99.2\,\%$ for type-1 attacks, and (b) achieves $97.4\,\%$ of maximum achievable detection probability (i.e., with 20 normal sensing reports in 5 clusters) for type-2 attacks.



(a) Incumbent detection performance (b) Number of valid sensing reports

Fig. 9. *Impact of threshold parameter (ϵ)*: (a) Q_{FA} and Q_D exhibit different behaviors under various ϵ values, and (b) the number of valid sensing reports for data fusion depends on both filter threshold and attack strength.

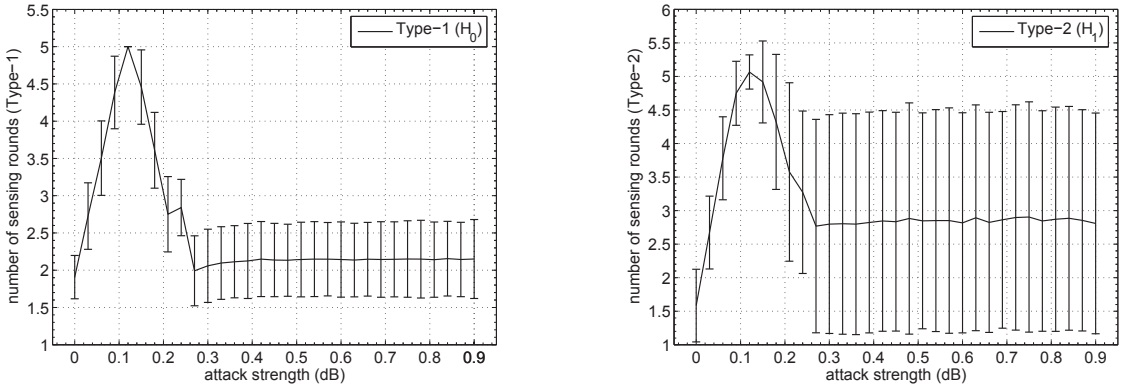
will have less samples to be used for data fusion. Since the data fusion is sensitive to the number of samples used, especially in very low SNR environments (as shown in Fig. 6), the incumbent detection performance degrades. For example, with 20 sensing reports remaining after filtering out all the 10 manipulated sensing reports, the average achievable Q_D is 0.88, which corresponds to the normalized Q_D of 0.93 in Fig. 8.

Fourth, in the absence of attacks, the correlation filter incurs a small increase in both Q_{FA} and Q_D. This is caused by the inaccuracy in the log-normal approximation of sensing reports, which causes over-filtering even in case of no attacks. We observed that this performance anomaly can be mitigated by reducing the sensing duration T_S (e.g., $<1\,$ms), which makes the approximation more accurate because the distribution of the sensing reports becomes closer to a normal distribution.

6.5 Tradeoff in Setting the Detection Threshold

We now study the impact of filtering threshold on attack detection performance. Fig. 9(a) plots the impact of the filtering threshold ϵ on incumbent detection performance. In this simulation, we fixed the attack strength at $0.1\,$dB for both types of attacks. The figure shows that Q_{FA} monotonically decreases as ϵ increases for both fusion rules, implying that filtering out more sensing reports always helps lower the false-alarm rate of incumbents. For the same reason, however, a large ϵ degrades the detection probability Q_D. This can be explained by the heavy-tail of a log-normal distribution of shadow fading; filtering out high RSSs at the tail lowers the decision statistics significantly, thus reducing the chance of generating false-alarms (or detecting incumbents). Another observation is that WGC outperforms EGC for type-2 attacks, thanks to its ability to adjust the weights for sensing reports based on their significance. However, the performance gain decreases as ϵ increases. For type-1 attacks, WGC also outperforms EGC in case of under-filtering, e.g., $\epsilon \in [0.01, 0.06]$, as discussed in Section 6.4.

Fig. 9(b) shows the average number of valid sensing reports (i.e., those that passed the filter). It clearly indicates that the filter becomes more aggressive in rejecting the sensing reports as ϵ increases, thus reducing the



(a) Type-1 attack (b) Type-2 attack

Fig. 10. *Average number of sensing rounds under various attack strengths*: the number of sensing rounds needed to meet the detectability requirement, i.e., $Q_{FA}, Q_{MD} \leq 0.01$, both under the filter threshold $\epsilon = 0.1$ and -20 dB SNR.

number of sensing reports to be used for making a final fusion decision. Therefore, the filter must be carefully designed to make the tradeoff between false-alarm and detection probabilities, while considering their dependency on attack strengths.

6.6 Meeting the IEEE 802.22 Detection Requirements via Sensing Scheduling

Here we evaluate the performance of the sensing scheduling algorithm in ADSP in terms of the number of sensing rounds (i.e., detection delay). Fig. 10 shows the number of sensing rounds needed to meet the detectability requirement of $Q_{FA}, Q_{MD} \leq 0.01$, which is below the requirements of IEEE 802.22, i.e., $Q_{FA}, Q_{MD} \leq 0.1$. Figs. 10(a) and 10(b) plot the mean and standard deviation of the number of sensing rounds. The figures indicate that the average number of sensing rounds is maximized when the attack strength is relatively small, i.e., 0.12 dB, thus confirming the observation made in Fig. 8. In 802.22, sensing rounds can be scheduled as frequent as once every 10 ms, i.e., one MAC frame size in 802.22. Therefore, Fig. 10 implies that ADSP can meet the incumbent detection timing requirement of 802.22, i.e., the returning primary signal must be detected within 2 seconds, since the maximum required number of sensing rounds remains below 5.

7 Concluding Remarks

The design of reliable distributed sensing for opportunistic spectrum use is a major research challenge in DSA networks. To address this challenge, we have developed a novel attack-tolerant distributed sensing protocol (ADSP) that selectively filters out abnormal sensor reports, and thus maintains the accuracy of incumbent detection. The key idea behind this mechanism is that the measured primary signal strength at nearby sensors should be correlated due to shadow fading, which has not been considered before. To realize this idea, we proposed a sensor clustering method and designed filters and data-fusion rules based on the correlation analysis of the sensor reports. We also proposed a sensing scheduling scheme based on sequential hypothesis testing that finds an optimal stopping time for sensing, while meeting the detection requirements of 802.22. ADSP can readily be implemented in 802.22 WRANs, incurring very low processing and communication overheads. We evaluated ADSP in realistic shadowing environments of 802.22 WRANs, demonstrating its ability to tolerate both type-1 and type-2 attacks.

Acknowledgement

The work reported in this paper was supported in part by the NSF under grants CNS-0519498 and CNS-0721529.

References

[1] A. W. Min, K. G. Shin, and X. Hu, "Attack-tolerant distributed sensing for dynamic spectrum access networks," in *Proc. IEEE ICNP*, Oct 2009.

[2] S. Shellhammer, S. Shankar, R. Tandra, and J. Tomcik, "Performance of power detector sensors of DTV signals in IEEE 802.22 WRANs," in *Proc. ACM TAPAS*, Aug 2006.

[3] S. Shellhammer and R. Tandra, "An evaluation of DTV pilot power detection," IEEE 802.22-06/0188r0, Sep 2006.

[4] A. W. Min and K. G. Shin, "An optimal sensing framework based on spatial rss-profile in cognitive radio networks," in *Proc. IEEE SECON*, June 2009.

[5] ——, "Impact of mobility on spectrum sensing in cognitive radio networks," in *Proc. ACM CoRoNet*, Sep 2009.

[6] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1326–1337, April 2008.

[7] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proc. IEEE DySPAN*, Nov 2005.

[8] S. M. Mishra, A. Sahai, and R. W. Brodersen, "Cooperative Sensing among Cognitive Radios," in *Proc. IEEE ICC*, June 2006.

[9] R. Chen, J.-M. Park, and J. H. Reed, "Defense against primary user emulation attacks in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 25–37, Jan 2008.

[10] S. Anand, Z. Jin, and K. P. Subbalakshmi, "An Analytical Model for Primary User Emulation Attacks in Cognitive Radio Networks," in *Proc. IEEE DySPAN*, Oct 2008.

[11] R. Chen, J.-M. Park, and K. Bian, "Robust distributed spectrum sensing in cognitive radio networks," in *Proc. IEEE INFOCOM*, April 2008.

[12] C. Cordeiro, K. Challapali, D. Birru, and S. Shankar, "IEEE 802.22: An introduction to the first wireless standard based on cognitive radio," *J. Commun.*, vol. 1, no. 1, pp. 38–47, April 2006.
[13] W. Rose, "Enhanced protection for low power licensed devices operating in tv broadcast bands," IEEE 802.22-06/0073r2, May 2006.
[14] USRP: Universal Software Radio Peripheral. [Online]. Available: http://www.ettus.com
[15] K. Tan *et al.*, "Sora: High performance software radio using general purpose multi-core processors," in *Proc. USENIX NSDI*, April 2009.
[16] W. Xu, P. Kamat, and W. Trappe, "TRIESTE: A trusted radio infrastructure for enforcing spectrum etiquettes," in *Proc. Allerton*, Sep 2008.
[17] A. Ghasemi and E. S. Sousa, "Asymptotic performance of collaborative spectrum sensing under correlated log-normal shadowing," *IEEE Commun. Lett.*, vol. 11, no. 1, pp. 34–36, Jan 2007.
[18] P. Kaligineedi, M. Khabbazian, and V. K. Bharava, "Secure cooperative sensing techniques for cognitive radio systems," in *Proc. IEEE ICC*, May 2008.
[19] A. Mody et al., "Collaborative sensing for security," IEEE 802.22-08/0301r011, Dec 2008.
[20] K. A. Woyach, A. Sahai, G. Atia, and V. Saligrama, "Crime and punishment for cognitive radios," in *Proc. Allerton*, Sep 2008.
[21] S. Liu, Y. Chen, W. Trappe, and L. J. Greenstein, "Aldo: An anomaly detection framework for dynamic sptectrum access networks," in *Proc. IEEE INFOCOM*, April 2009.
[22] W. Zhang, S. K. Das, and Y. Liu, "A trust based framework for secure data aggregation in wireless sensor networks," in *Proc. IEEE SECON*, Sep 2006.
[23] Y. Yang, X. Wang, S. Zhu, and G. Cao, "SDAP: A secure hop-by-hop data aggregation protocol for sensor networks," in *Proc. ACM MobiHoc*, May 2006.
[24] W. Du, J. Deng, Y. S. Han, and P. K. Varshney, "A witness-based approach for data fusion assurance in wireless sensor networks," in *Proc. IEEE Globecom*, Dec 2003.
[25] F. Liu, X. Cheng, and D. Chen, "Insider attacker detection in wireless sensor networks," in *Proc. IEEE INFOCOM*, May 2007.
[26] FCC, "Notice of Proposed Rulemaking and Order," FCC 08-188, Aug 2008.
[27] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.
[28] S. J. Shellhammer, "Spectrum sensing in ieee 802.22," in *IAPR Workshop on Cognitive Information Processing*, June 2008.
[29] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the Energy Detection of Unknown Signals over Fading Channels," in *Proc. IEEE ICC*, May 2003.
[30] M. Gudmundson, "A correlation model for shadow fading in mobile radio," *Electron. Lett.*, vol. 27, no. 23, pp. 2146–2147, Nov 1991.
[31] T. Muetze, P. Stuedi, F. Kuhn, and G. Alonso, "Understanding radio irregularity in wireless networks," in *Proc. IEEE SECON*, June 2008.
[32] N. Patwari and P. Agrawal, "Effects of correlated shadowing: Connectivity, localization, and RF tomography," in *Proc. IEEE IPSN*, April 2008.
[33] J. Riihijärvi, P. Mähönen, M. Wellens, and M. Gordziel, "Characterization and modelling of spectrum for dynamic spectrum access with spatial statistics and random fields," in *Proc. IEEE PIMRC*, Sep 2008.
[34] I. Forkel, M. Schinnenburg, and M. Ang, "Generation of two-dimensional correlated shadowing for mobile radio network simulation," in *Proc. WPMC*, Sep 2004.
[35] A. Algans, K. I. Pedersen, and P. E. Mogensen, "Experimental analysis of the joint statistical properties of azimuth spread, delay spread, and shadow fading," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 523–531, April 2002.
[36] Y. Chen, W. Trappe, and R. P. Martin, "Attack detection in wireless localization," in *Proc. IEEE INFOCOM*, May 2007.
[37] T. Park and K. G. Shin, "Attack-tolerant localization via iterative verification of locations in sensor networks," *ACM T. Embed. Comp. S.*, vol. 8, no. 1, pp. 2:1–2:24, Jan 2009.
[38] M. Sawada, D. Cossette, B. Wellar, and T. Kurt, "Analysis of the urban/rural broadband divide in Canada: Using GIS in planning terrestrial wireless deployment," *Government Information Quarterly*, no. 23, pp. 454–479, Sep 2006.
[39] A. Wald, *Sequential Analysis*. Dover Publications, 2004.

PLACE PHOTO HERE

Alexander W. Min (S'08) received his B.S. degree in electrical engineering from Seoul National University, Korea, in 2005 and the M.S. degree from the University of Michigan in 2007. He is currently a Ph.D. candidate in the department of Electrical Engineering and Computer Science (EECS), the University of Michigan, Ann Arbor. Since 2006, he has been a research assistant in the Real-Time Computing Laboratory (RTCL) in EECS Department. In 2010, he was a Research Intern at Deutsche Telekom Inc., R&D Laboratory, Los Altos, USA. His research interests are in the area of cognitive radio and dynamic spectrum access networks including spectrum sensing, resource allocation, security, and secondary spectrum market. He is a student member of ACM and the IEEE Communications Society.

PLACE PHOTO HERE

Kang G. Shin (F'92) is the Kevin and Nancy O'Connor Professor of Computer Science and founding director of the Real-Time Computing Laboratory in the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor. His current research focuses on computing systems and networks as well as on embedded real-time and cyber-physical systems, all with emphasis on timeliness, security, and dependability. He has supervised the completion of 65 Ph.D.s, and authored/coauthored more than 720 technical articles. He has co-authored (with C. M. Krishna) a textbook, Real-Time Systems (McGraw Hill, 1997). He has received numerous best paper awards, including the Best Paper at the 2010 USENIX Annual Technical Conference, the IEEE Communications Society William R. Bennett Prize Paper Award in 2003, the Best Paper Award from IWQoS'03 in 2003, and an Outstanding IEEE Transactions of Automatic Control Paper Award in 1987. He has also coauthored papers with his students, which received the Best Student Paper Awards from the 1996 IEEE Real-Time Technology and Application Symposium and the 2000 UNSENIX Technical Conference. He has also received several institutional awards, including the Research Excellence Award in 1989, Outstanding Achievement Award in 1999, Service Excellence Award in 2000, Distinguished Faculty Achievement Award in 2001, and Stephen Attwood Award in 2004 from The University of Michigan (the highest honor bestowed to Michigan Engineering faculty); a Distinguished Alumni Award of the College of Engineering, Seoul National University in 2002; 2003 IEEE RTC Technical Achievement Award; and 2006 Ho-Am Prize in Engineering (the highest honor bestowed to Korean-origin engineers). He is a Fellow of ACM and has chaired numerous conferences, including the 2009 ACM MobiCom'09, IEEE SECON'08, ACM/USENIX MobiSys'05, and IEEE RTAS 2000, and IEEE RTSS'86 and '87. He also chaired the IEEE Technical Committee on Real-Time Systems, 1991-1993, and has served as an Editor of IEEE Transactions on Parallel and Distributed Computing, and an Area Editor of the International Journal of Time-Critical Computing Systems, Computer Networks and ACM Transactions on Embedded Systems.

PLACE PHOTO HERE

Xin Hu is currently a Ph.D. candidate in the department of Electrical Engineering and Computer Science, the University of Michigan, Ann Arbor. He received his B.S. degree from Zhejiang University, China in 2005 and M.S. degree from the University of Michigan in 2007, respectively. His research interests lie primarily in the area of network and system security. The goal of his research is to discover the underlying principles of real-world security problems and design effective and scalable solutions to enhance the security, availability and integrity of computer systems. His current research focuses on monitoring and detecting botnets, reverse engineering malware, and location privacy protection for mobile devices.