Trusted Collaborative Spectrum Sensing for Mobile Cognitive Radio Networks

Shraboni Jana, Kai Zeng*, Wei Cheng⁺ and Prasant Mohapatra University of California, Davis, CA 95616 *University of Michigan, Dearborn, MI 48128 ⁺Virginia Commonwealth University, VA 23284

Email: {sjana, pmohapatra}@ucdavis.edu, *kaizeng@umich.edu, +wcheng3@vcu.edu

Abstract—Collaborative spectrum sensing is a key technology in cognitive radio networks (CRNs). Although mobility is an inherent property of wireless networks, there has been no prior work studying the performance of collaborative spectrum sensing under attacks in mobile CRNs. Existing solutions based on user trust for secure collaborative spectrum sensing cannot be applied to mobile scenarios, since they do not consider the location diversity of the network, thus over penalize honest users who are at bad locations with severe pathloss. In this paper, we propose to use two trust parameters, Location Reliability and Malicious Intention (LRMI), to improve both malicious user detection and primary user detection in mobile CRNs under attacks. Location Reliability reflects pathloss characteristics of the wireless channel and Malicious Intention captures the true intention of secondary users, respectively. We propose a primary user detection method based on location reliability (LR) and a malicious user detection method based on LR and Dempster-Shafer (D-S) theory. Simulations show that mobility helps train location reliability and detect malicious users based on our methods. Our proposed detection mechanisms based on LRMI significantly outperforms existing solutions. In comparison to the existing solutions, we show an improvement of malicious user detection rate by 3 times and primary user detection rate by 20% at false alarm rate of 5%, respectively.

I. INTRODUCTION

With the ever-increasing wireless applications and traffic demand, spectrum shortage becomes a more severe and urgent problem. Cognitive radio technology [1] is considered as a promising solution to improve the spectrum utilization and alleviate the spectrum shortage. The basic idea of cognitive radio networks (CRNs) is that when the primary (licensed) users are absent, the secondary (unlicensed) cognitive users can opportunistically access the primary users' spectrum, but have to evacuate when the primary users emerge.

The DARPA's Next Generation Program [2] is based on spectrum sensing and dynamic spectrum utilization. FCC has mentioned the need of cognitive radio [3] for emergency situations. Thus, CRNs will play an important role in the future communications for both tactical military forces and emergency responders. A CRN deployed and operated by the military or government emergency units could be hampered or interfered by an adversary. The CRNs should have the capability to overcome any security threat.

The collaborative spectrum sensing paradigm in CRN opens a hole to the attackers who can falsify the sensing results. The motivation of an attacker can be either selfish or malicious. Being selfish, an attacker may report the presence of the primary user when there is actually none in order to deny the legitimate users' access to the spectrum (Denial of Service attack). While being malicious, an attacker may report an absence of the primary user when there is one, thus causing chaos and interference for primary and secondary users.

Existing solutions to detecting the sensing falsification attacks have focused on identifying the attackers as abnormal or outlier nodes within a small cell [4], [5], [6], [7], [8], [9], [10]. Basically, when a user's report deviates from common readings beyond a certain threshold, its trust value is degraded. A dishonest attacker can thus be identified, and its negative impact on the spectrum sensing can be weakened or eliminated. However, these solutions have two major limitations. First, they assume the whole area has the same channel propagation characteristics, which is not practical. It has been found that the pathloss are different at different sensing regions [11]. Second, they assume the users are static and cannot be directly applied to mobile scenarios. For example, existing trust-based solutions tend to over penalize an honest user who is at a bad location with large pathloss. Even when the user moves to a good location later on, its contribution to the spectrum sensing will be limited since it has been assigned a low trust value.

In this paper, we propose to use two trust parameters, Location Reliability and Malicious Intention (LRMI), to improve both malicious user detection and primary user detection in mobile CRNs under attacks. Location Reliability (LR) reflects pathloss characteristics of the wireless channel and Malicious Intention (MI) captures the true intention of secondary users, respectively. We therefore, propose to evaluate the reports at the fusion center using Dempster-Shafer(D-S) Theory based on two sources of evidence associated with each report - cell from which the report is generated (Location Reliability) and who has generated the report (Malicious Intention). Location Reliability captures trust over different positions as distributions of path-loss are not identical although they are independent. Malicious Intention defines how much trustworthy a user is. To the best of our knowledge, no prior work has studied the impact of mobility on the collaborative spectrum sensing under attacks or provided applicable solutions.

Our basic idea is as follows. For an honest user, when it moves from a good location to a bad location, we should give low weight to its report or even ignore it, since the report will not be reliable because it is now in a bad location. While at the same time, we should not decrease its credibility because it happened to be in a bad location. Considering an attacker, when it is in a good location, if it just blindly lies on the sensing result, it will be detected with a better chance since it is supposed to report a high quality result. When the attacker is in a bad location, its damage will be limited if we do not count much on the report from a bad location. Therefore, we should separate location reliability from user trust in mobile CRNs.

The major contributions of this paper are summarized as follows:

- We propose to evaluate collaborative spectrum sensing in mobile CRN based on location-trust and user-trust.
 Our novel solutions for trusted collaborative spectrum sensing takes into consideration both location-diversity and mobility of secondary users.
- We study the performance of our solutions in terms of receiver operating characteristics for both malicious users detection and primary user detection.
- We conduct extensive simulations to evaluate our proposed mechanisms and compare their performance with existing solutions. We find that with increase in number of users, mobility and system observation time, performance of our proposed scheme improves. Our proposed detection mechanisms based on LRMI significantly outperform existing solutions in terms of improving the malicious user detection rate by 3 times and primary user detection rate by 20 % at false alarm rate of 5%, respectively.

The rest of this paper is organized as follows. Section II discusses the related work. The system model is introduced in Section III followed by the problem formulation and our proposed solutions in Section IV. We evaluate the impact of mobility and our solutions and conclude this paper in Section VII and Section VIII, respectively.

II. RELATED WORK

The performance gains, achieved by collaborative spectrum sensing in CRNs is well established in literature. The centralized collaborative spectrum sensing has been included in the IEEE 802.22 standard draft [12]. The secondary users report sensing results to a base station (fusion center) on a periodic or on-demand basis about the presence and absence of primary user using spectrum sensing. The secondary user trust is critical for such a cooperative systems to operate reliably. Trust-based mechanisms have been widely suggested for collaborative spectrum sensing under report falsifying attacks, where dishonest attackers lie on their sensing results.

The calculation of the trust of secondary users has been addressed using different techniques in the literature. The trust values can be calculated from the reports received from the secondary users, comparing deviation suffered by each from average [5]. The secondary users are penalized according to the deviations calculated. In another paper by the same authors [8], outlier techniques are studied in detail and based on the knowledge of partial primary user activity, malicious user(s) identification is done. Among other techniques, the Bayesian rule can be applied to compute the a posteriori probability of

being an attacker for each secondary user. When the posteriori probability of a certain secondary user exceeds the suspicious level threshold, it is claimed to be an attacker and is removed from the collaboration [10]. For multiple attackers, the large number of combinations of attackers and honest users is removed by using an onion-peeling based approximation to reduce computational complexity.

Abnormality detection algorithm based on proximity, which is widely used in the field of data mining has been introduced in [4], to solve the problem of malicious users in the system using history reports of each secondary user. The proposed architecture in [6], needs to collect spectrum sensing data from multiple sources or equipment on consumer premises. This process is known as crowdsourcing. In [6], the area of interest is divided in to cells and the credibility of these devices are kept in check by corroboration among neighboring cells in a hierarchical structure to identify cells with significant number of malicious nodes.

In the solution proposed by authors in [9], focus is on a small region for enhancing the primary user detection by exploring the spatial diversity in user reports. In another paper by the same authors, [13], impact of mobility in spectrum sensing is analyzed. The authors show that because of mobility, the secondary user sensing results get uncorrelated faster thus giving better performance compared to spectrum sensing performed by static secondary users.

To the best of our knowledge, none of the existing work studied the impact of mobility on the malicious user detection and primary user detection under attack in CRNs. None of the existing trust-based collaborative spectrum sensing solutions are directly applicable for mobile scenarios, either. Our proposed solutions [14] are different from all the existing solutions that we separate the location reliability from the user trust, thus achieve better performance on malicious user detection which in turn improve the primary user detection under attacks in mobile scenarios.

III. SYSTEM MODEL

We divide the area of interest into a grid (Figure 1) and each cell in a grid is assumed to experience path-loss exponent and shadowing characteristic of that cell. The assumption is reasonable since some areas will have deep fade caused by buildings, trees etc. compared to others. We use the term location and cell interchangeably in the rest of the paper for cell in the grid. In Figure 1, we divide the grid into equal size cells, but our approach supports any cell shape with any size. In other words, the area of interest can be divided into any number of cells in any manner depending on the granularity required. Note that it is not necessary to have tiny size cells as the signal attenuation is generally stable in a small area. We therefore, target mainly urban areas less than 5000 square meters, where building, trees etc. will cause deep fadings and cells are approximately more than street block size i.e. approx 300 - 500 square meters.

Let there be L finite cells in the area of interest. $C=\{c_1,c_2,c_3,...,c_L\}$ be the set of cell identification numbers of L cells in the grid. Let $U=\{u_1,u_2,u_3,...,u_N\}$ be set

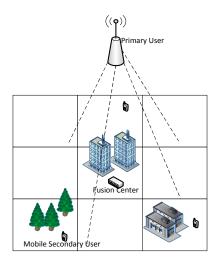


Fig. 1. System Model with static primary user and mobile secondary users.

of N users in the area. The secondary users are mobile and may be present at different cells at different times. They use energy detection for spectrum sensing and report their measurements to fusion center. Fusion center uses these spectrum sensing reports to determine primary user status. Let the system be observed for total K observation slots and any slot $k \in \{1, 2, 3, ..., K\}$.

TABLE I Main Symbols Used

Symbols	Descriptions
N	Total number of secondary users in the system.
M	Total number of malicious secondary users in the system.
L	Total number of cells.
$Y_{i,k}^{j}$	Energy detector output for user u_i at k^{th} sensing from cell c_j
$H_{0(k)}$	Primary user is idle during k^{th} sensing
$H_{1(k)}$	Primary user is active during k^{th} sensing
η	Primary user detection threshold
ξ	Malicious user detection threshold
α_j	Path-loss exponent of cell c_j
ψ_j	Shadowing in dB of cell c_j

A. Energy Detection

The secondary users sense the spectrum to identify if the primary user is active or idle in the area of interest. There has been three main spectrum sensing techniques suggested in literature - energy detection, matched filtering and cyclostationary feature for primary user detection by the secondary users [1]. Out of these three techniques, energy detection is the most promising solution due to its simplicity in implementation and its capability to detect any shape of waveforms. We choose energy detection at secondary users, as the underlying spectrum sensing scheme.

The primary user detection is modeled as a hypothesis test. The null hypothesis H_0 indicates the primary user is idle, while the alternative hypothesis H_1 indicates the primary user is active. We further assume that the time the system is in either of the states H_0 and H_1 follow exponential distributions

as commonly used in the literature [1], and that durations of successive active and inactive periods are independent of each other

We assume the secondary users sense the spectrum periodically slot by slot by using energy detection. Assume the bandwidth of the primary user signal is W, at each sensing slot, each user takes 2TW samples with the sample interval of T. Then, the output of the energy detector for secondary user u_i , present at cell c_j for the k^{th} sensing slot is [15],

$$Y_{i,k}^{j} = \begin{cases} \frac{1}{2TW} \sum_{m=1}^{2TW} |n_{i,k}[T_k + \frac{m}{2W}]|^2 & H_{0(k)} \\ \frac{1}{2TW} \sum_{m=1}^{2TW} |h_{i,k}^{j}[T_k + \frac{m}{2W}]s[T_k + \frac{m}{2W}] \\ + n_{i,k}[T_k + \frac{m}{2W}]|^2 & H_{1(k)} \end{cases}$$

where T_k is the time at the beginning of the k^{th} sensing slot. $h_{i,k}^j$ is the channel gain for secondary user u_i in cell c_j during k^{th} sensing. Assuming the noise, $n_{i,k}$ and primary signal, s to be uncorrelated, the distribution of the energy detector output is given as

$$Y_{i,k}^{j} \sim \begin{cases} \chi_{2TW}^{2} & H_{0(k)} \\ \chi_{2TW}^{2}(2\gamma_{i,k}^{j}) & H_{1(k)} \end{cases}$$
 (1)

where $\gamma_{i,k}^j = \frac{|h_{i,k}^j|^2 Pt}{N_0 W} = \frac{P r_{i,k}^j}{N_0 W}$ is referred as instantaneous signal-to-noise ratio experienced by a secondary user u_i for transmit power P_t and channel gain $h_{i,k}^j$. χ^2_{2TW} and $\chi^2_{2TW}(2\gamma_{i,k}^j)$ denote central and non-central chi-square distributions with 2TW degrees of freedom, respectively.

Assuming channel bandwidth is much larger than the coherent bandwidth, the effect of multipath fading is negligible. The received primary user power at secondary user u_i at a distance $d_{i,k}$ from primary user can be expressed as [16] in dB:

$$Pr_{i,k}^{j}(dB) = P_{t}(dB) - \{PL_{0} + 10\alpha_{j}log_{10}(\frac{d_{i,k}}{d_{0}}) + \psi_{j}\}$$
 (2)

where PL_0 is a path-loss at a reference distance d_0 in dB and is close to $20log10(\frac{4\pi d_0}{\lambda})$, where λ is wavelength. Path-loss exponent α_j for cell c_j ranges from 2 to 5 [11]. Empirical measurements support the log-normal distribution where ψ_j in dB -

$$p(\psi_j) = \frac{1}{\sqrt{2\pi}\sigma_{\psi_j}} exp[-\frac{(\psi_j - \mu_{\psi_j})^2}{2\sigma_{\psi_j}^2}]$$
 (3)

At each sensing slot k, each user u_i reports $Y_{i,k}^{j}$ along with their coordinates or current cell id (each cell has a unique cell id) to the fusion center. A secondary user can find its current coordinates, based on localization techniques [17]. For detection threshold η at the fusion center, if $Y_{i,k}^{j} > \eta$, the fusion center concludes the primary user is active otherwise primary user is idle. Hence, the probability of detection (P_d) and probability of false alarm (P_f) for a sensing report, $Y_{i,k}^{j}$ [18]:

$$P_d = P\{Y_{i,k}^j > \eta | H_{1(k)}\}$$
 (4)

$$= \int_{\gamma_{i,k}^j} Q_m(\sqrt{2TWx}, \sqrt{\eta}) f_{\gamma_{i,k}^j}(x) dx \qquad (5)$$

$$P_f = P\{Y_{i,k}^j > \eta | H_{0(k)}\}$$
 (6)

$$= \frac{\Gamma(TW, \eta/2)}{\Gamma(TW)} \tag{7}$$

where Q_m is marqum-function, Γ is a gamma function and η is the detection threshold. Probability of false alarm is independent of SNR since under $H_{0(k)}$ there is no primary signal present. Since the channel gain, $h_{i,k}^j$ is varying due to shadowing or fading for each sensing report, probability of detection is conditioned on instantaneous SNR $\gamma_{i,k}^j$.

B. Collaborative Spectrum Sensing

The uncertainty in a sensing report due to fading may be mitigated at the fusion center by considering spectrum sensing results from multiple users. Such collaboration to decide on primary user occupancy has been widely studied [19], [18] and is a part of IEEE 802.22 standard draft [12]. Commonly used techniques in the literature for collaborative spectrum sensing are soft-combining and hard-combining. In soft-combining, raw sensed signal power values are sent from secondary users to the fusion center, whereas in hard-combining techniques a 0/1 decision from each secondary user is considered. We consider soft-combining in this paper because its performance is much better than hard-combining with only a slightly higher communication overhead [18].

As we consider soft-combining, fusion center will process the reports and make a decision whether there is an active primary user or not. Let \widetilde{Y}_k be the soft-combined spectrum sensing report processed by the fusion center. The probability of detection and false alarm per slot for collaborative spectrum sensing will be given as

$$Q_{d(k)} = Pr(\widetilde{Y}_k > \eta | H_{1(k)})$$
 (8)

$$Q_{f(k)} = Pr(\widetilde{Y_k} > \eta | H_{0(k)}) \tag{9}$$

We will detail our solution on how to combine the user's report and how to evaluate the performance of primary user detection and malicious user detection in Section IV.

C. Attack model

We assume a very strong attack model where malicious users lie about their sensing reports and locations (i.e. the cell in which they are present during sensing). A malicious user is aware of the primary user status. Our approach works even if the malicious users are not always aware of the primary user status and lie intermittently. The difference will be in the detection time of malicious users in such cases which is out of the scope of our work. Each malicious user thwarts the system performance by-

- Reporting an increased observation $(Y_{i,k}^j+\Delta)$ when the primary user is inactive, thus increasing the false-alarm
- Reporting a decreased observation $(Y_{i,k}^j \Delta)$ when the primary user is active, thus increasing the missed-detection rate.
- Reporting incorrect cell number.

The deviation Δ is any random value chosen by malicious user for each observation. Faulty nodes are not a part of the system. The fusion center is unaware of attack strategy by malicious users. It is also unaware of primary user location and primary user transmit power. We assume that each malicious

user acts alone and the attack strategy is independent during each sensing slot. If the reliability of a user assigned by fusion center drops below a certain threshold (ξ), the user report is not considered. The number of malicious secondary users are always less than the number of honest users in the system.

Note that a malicious user may either take all the above attacks or randomly choose some of them. In either cases, the report received at the fusion center is erroneous. In this paper, we primarily consider the strongest attack models, and the proposed approaches can be applied to the simpler attack models as well.

IV. LOCATION RELIABILITY (LR)

In this section we propose our algorithms for evaluating location reliability and malicious intention in mobile CRNs. One of the key cognitive radio application addressed in, FCC report [3], is the capability of public safety enhancements. For such mobile ad-hoc emergency networks it is therefore, more realistic that the secondary users in mobile CRNs can be turned on at any location and at any time. It is not feasible to have prior knowledge of fading and shadowing characteristics of the area of interest at the fusion center neither availability of large training-data for data-mining approach seems feasible. Moreover, such networks when set-up at urban locations can have a high variance in shadowing and fading characteristics within a few feet. To add to the uncertainty imposed by different path-loss at different locations, the spectrum sensing reports received at the fusion center can be corrupted by the malicious users, from the very onset of the system.

In this section we discuss the evaluation of location trust in mobile CRNs. The papers, [18] and [11] have studied the performance of collaborative spectrum sensing under different shadowing and fading characteristics. Large-scale propagation loss depends significantly on the users positions and in practical situations, it is almost impossible that the effects would occur with identical distribution in different geographical locations. It is therefore, very important to know the reliability of location from where the sensing result is generated. Apart from the cell characteristics, the sensing reports are affected by two other reasons - 1) primary user activity and 2) malicious secondary user activity. We need to understand the impact of these two activities before deriving LR from spectrum sensing report as there is no training period.

A. Primary User Activity

We are interested in the sensing measurement from each cell in order to determine its reliability. Let $S(c_j) = \{Y_1^j, Y_2^j,, Y_n^j\}$ be a sequence of n reports from the cell c_j , the fusion center receives in K slots where $n \leq K$ as secondary users may not be present in each cell during a sensing slot k. From Equation 1, each element of $S(c_j)$ is an independent and identically distributed random variable generated at different times from different secondary users. The elements follow central or non-central chi-square distributions depending on the status of primary user. Let p_B and p_I be the probability that primary user is active or idle respectively

 $(p_B + p_I = 1)$. The sum of such n random reports from cell c_i

$$Y^{j} = \frac{\sum_{m=1}^{n} Y_{m}^{j}}{n} \tag{10}$$

Since the sum of such random variables(> 250 [19], [9]) follow a central limit theorem, the distribution of spectrum sensing measurements from cell c_i .

$$Y^{j} \sim \mathcal{N}((\mu_{c_{j}} + \mu_{o})p_{B} + \mu_{o}p_{I}, \frac{(\sigma_{c_{j}}^{2} + \sigma_{o}^{2})p_{B}^{2} + \sigma_{o}^{2}p_{I}^{2}}{n})$$
 (11)

where μ_{c_j} and $\sigma_{c_j}^2$ is the mean and variance of power received from cell c_j due to active primary user. μ_o and σ_o^2 is the mean and variance of the noise. Thus μ_{c_j} and p_B are determining factors for the trust level of cell c_j . Since p_B is same for all cells, μ_{c_j} determines location reliability of a cell. We take 600 consecutive samples (sensing results) for different pathloss and shadowing, with primary user on-off model. From Figure 2, it is evident that Y^j follows gaussian.

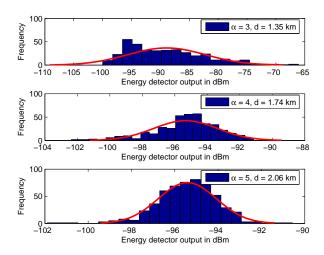


Fig. 2. Distribution of average sensing measurements $E[Y^j]$ from cells with different path-loss exponent, $\psi=4dB,\,d$ is the distance from primary user.

B. Malicious User Activity

The spectrum sensing measurements from a cell, will be further infected with malicious data. There is no closed form relation between velocity and spectrum sensing for variable large-scale propagation losses. We find that as the user speed varies for a fixed area of interest, the type of users (honest or malicious) visiting a cell keeps varying. In other words, the user diversity in a cell increases over the time with increase in speed. It shows that for different speed of the mobile users the average cell change rate increases with increase in speed [20]. We further, simulate the behavior of mobile nodes in Figure 3 to validate if the number of distinct users visiting a cell during a given period depend upon speed. For N = 10, 12, 16 and 20 at V = 0m/s, 10m/s, 20m/s and 30m/s, the average distinct users per cell increases with N and V. Moreover, we formally analyze the impact of malicious data on sensing results in mobile CRNs as follows.

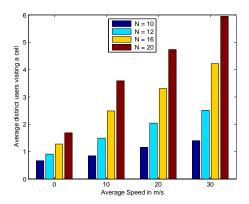


Fig. 3. Average number of distinct users visiting a cell.

In static setting, malicious users would have dominated the location reliability for the infected cell. However, in mobile CRNs, their attack gets distributed across all the cells over time. On the other hand, mobile honest users help to train location reliability of each cell. As a result, the reports generated from a cell at different times are from different users, and the reliability of the report should be able to converge to a stable value even with the presence of malicious nodes in the system.

Theorem 1. Given the mobility models of users in a mobile CRN, the malicious users' impacts on the sensing reports converge to a stable value.

Proof: Assuming that user u_i is traveling with velocity V_{u_i} in the network, the distance S_{u_i} travelled by the user after Δ_t time can be represented by

$$S_{u_i} = \Delta_t \times V_{u_i}$$

The number of cells visited by u_i in Δ_t time is $\frac{S_{u_i}}{K_{u_i} \times r}$, where K_{u_i} is the parameter that is related to u_i 's mobility model, and r is the cell size. Consequently, given a cell c_j , its expected number of time of being visited by u_i is $\frac{S_{u_i}}{K_{u_i} \times r \times L}$. Then, expected number of time, denoted as Υ_{c_j} , which c_j is visited by all the N users in Δ_t time, is:

$$\Upsilon_{c_j} = \sum_{i=1}^{N} \frac{\Delta_t \times V_{u_i}}{K_{u_i} \times r \times L}$$
$$= \frac{\Delta_t}{r \times L} \times \sum_{i=1}^{N} \frac{V_{u_i}}{K_{u_i}}$$

Therefore, the number of reports generated by u_i is a function of the user's mobility model (V_{u_i} and K_{u_i}). We define u_i 's impact (τ_i) on sensing result as the number of reports generated by him over the number of all the reports in the cell.

$$\tau_{i} = \frac{\frac{V_{u_{i}}}{K_{u_{i}}}}{\sum_{i=1}^{N} \frac{V_{u_{i}}}{K_{u_{i}}}}$$

Then, the impacts of all the malicious users on cell c_j 's report in Δ_t time can be represented by (12), where i' is the index of malicious users.

$$\tau = \frac{\sum_{i'=1}^{M} \frac{V_{u_{i'}}}{K_{u_{i'}}}}{\sum_{i=1}^{N} \frac{V_{u_i}}{K_{u_{i'}}}}$$
(12)

Note that τ is only determined by the user's mobility model. Therefore, τ is a stable value when the mobility models are given.

Assuming that all the users have similar mobility models, we can have (13) according to (12).

$$\tau \approx \frac{M}{N} \tag{13}$$

It means that the sensing results are still trustable if the malicious users are minority and the time Δ_t is long enough.

From Equation 11, the average sensing result from a cell c_j will be higher for location with less attenuation and fading and can be evaluated as,

$$E[Y^{j}] = \frac{\sum_{k=1}^{K} \sum_{i=1}^{N} Y_{i,k}^{j} \delta_{i,k}^{j}}{\sum_{k=1}^{K} \sum_{i=1}^{N} \delta_{i,k}^{j}}$$
(14)

$$tot_{rep}(c_j) = \sum_{k=1}^{K} \sum_{i=1}^{N} \delta_{i,k}^j$$
 (15)

 $\delta_{i,k}^{j} = 1$ for user u_i in cell c_j at time slot k otherwise is 0. $tot_{rep}(c_j)$ represents number of reports received from the cell c_j .

According to the analysis from Equation 13, we group the current reports with past reports based on the cell location informed by the secondary users. The steps of evaluating LR have been detailed in Algorithm 1. At the beginning of the Algorithm 1, all cells are given same trust values. $\mathcal{R} = \{Y_{i,k}^j | i \in \{1,...,N\}, k \in \{1,...,K\}, j \in C\}$ is the set of evidences received in K sensing slots. $\beta_k(c_j)$ is the LR of the cell c_j .

Algorithm 1 Location Reliability (LR)

```
Initialize S
For \forall c_j, S(c_j) = \{\phi\}, E[Y^j] = 0, \beta_0(c_j) = \frac{1}{L} and tot_{rep}(c_j) = 0
For each k
for c_j = 1 \to L do
  if \mathcal{R} has c_j then
    \mathbf{r} = \text{number of reports from cell } c_j.
    tot_{rep}(c_j) = tot_{rep}(c_j) + r
S(c_j) = S(c_j) \bigcup \{Y_{i,k}^j\}
E[Y^j] = \frac{\sum S(c_j)}{tot_{rep}(c_j)}
end if
end for
\beta_k(c_j) = \frac{E[Y^j]}{\sum_j E[Y^j]}
```

V. MALICIOUS INTENTION (MI)

Locations do not lie but users may. At the same time, a user's performance may be hampered because of its instantaneous position during spectrum sensing. The true intention of a secondary user cannot be captured entirely by their respective sensing reports as honest users can be in bad locations experiencing deep path-loss and shadowing or malicious users can be in good locations and vice versa.

We use Dempster-Shafer (D-S) theory to evaluate trust-worthiness in collaborative spectrum sensing in mobile CRN. D-S theory is a mathematical theory of evidence. It can be viewed as a method for reasoning under uncertainty (epistemic uncertainty) to logically arrive at decisions based on available knowledge [21]. In dynamic mobile cognitive radio networks, the D-S theory is well suited for two reasons - 1) it reflects uncertainty and the D-S theory rule of combination, 2) combines evidences from two or more sources to form inferences. In a finite discrete space, D-S theory can be interpreted as a generalization of probability theory where probabilities are assigned to sets as opposed to mutually exclusive singletons.

The frame of discernment $\Theta_{u_i} = \{T, -T\}$ denotes a set of mutually exclusive and exhaustive hypotheses about the problem domain - if user u_i is trustworthy or malicious. The power set $2^{\Theta_{u_i}}$ is $\{\phi, T, -T, \{T, -T\}\}$. The Belief Mass Assignment (bma), represented by m, defines a mapping of the power set to the interval between 0 and 1. For each k,

$$m: 2^{\Theta_{u_i}} \to [0\ 1], m(\phi) = 0, \sum_{A_k \in 2^{\Theta_{u_i}}} m(A_k) = 1$$
 (16)

Intuitively, trust bma should be lesser when the deviation increases. However, since this deviation is not limited, for normalization we choose trust bma to be negative exponential. This limits the trust bma in $[0\ 1]$. Based on Equation 16, the bma function for k^{th} sensing,

$$m_{u_i}(A_k = T) = e^{-|D|}$$
 (17)

$$m_{u_i}(A_k = -T) = 0 (18)$$

$$m_{u_i}(A_k = \{T, -T\}) = 1 - m_{u_i}(A_k = T)$$
 (19)

where D, is the deviation in the user report depending upon its location. As the deviation D decreases, our belief in u_i increases and vice versa. The uncertainty due to the noise level experienced by the user is incorporated into $m_{u_i}(A_k = \{T, -T\})$.

The trust penalty or deviation, of evidence from average be $\zeta_k(u_i)$ for user u_i for any location c_j .

$$\zeta_k(u_i) = \frac{Y_{i,k}^j - avg\{Y_{i,k}^j\}_{i=1}^N}{std\{Y_{i,k}^j\}_{i=1}^N}$$
(20)

avg stands for average and std stands for standard deviation. Some users are more vulnerable to misreading due to their instantaneous location. The evidence received from a user in a cell is discounted based on location reliability $\beta_k(c_j)$, computed in Algorithm 1, to reflect the user's credibility. Hence, the deviation in the user trust for sensing slot k is discounted based on the location-

$$D = (1 - \beta_k(c_i))\zeta_k(u_i) \tag{21}$$

where $(1 - \beta_k(c_i))$ reflects the unreliability of the reported location. In the trust computation, the deviation in users reports will reduce more if location unreliability is higher and vice versa. This discounting of deviation, thus, incorporates unreliability of the locations. Since the user trust is evaluated and updated based on current and past reports, we use D-S rule of combination for updating subsequent bma in each sensing slot.

$$m_{u_i}(\{A=T\}) = \frac{\sum_{\bigcap A_r = A} \prod_{r=1}^k m_{u_i}(A_r)}{1 - \sum_{\bigcap A_r \neq A} \prod_{r=1}^k m_{u_i}(A_r)}$$
(22)

where $A_r \in 2^{\Theta_{u_i}}$ and $m_{u_i}(A_r)$ is obtained from Equation 17 - 19.

Elaborating Equation 22, the bmas for sensing slot k will be updated as in Equation 23 where it incorporates past and present evidences. At slot k, $m_{u_i}(\{A=T\}) = m_{u_i}(A_k=T)$ where $m_{u_i}(A_k = T)$ is evaluated as in Equation 23.

The formalized function of user trust evaluation at each slot k, will be

$$T_k(u_i) = m_{u_i}(\{A = T\})$$
 (24)

VI. LRMI - RECEIVER OPERATING CHARACTERISTICS

We evaluate Equations 7 - 12 proposed in [5] against our algorithm. We address this approach as Malicious Detection (MD) in our paper. MD assigns trust factors such that they are exponentially decreasing according to their distance from the median. The approach is similar to LRMI (algorithm 2) in terms of using exponential function for trust evaluation but does not take into consideration location reliability.

Algorithm 2 LRMI

```
Initialize S
T_k(u_i) = 1 \forall u_i
For each k
Evaluate LR based on Algorithm 1
for u_i = 1 \rightarrow N do
   Evaluate Equation 20 - 24
   if T_k(u_i) < \xi then
     Remove Y_{i,k}^j from \mathcal{R}
   end if
for Y_{i,k}^j in \mathcal{R} do
   Apply Equation 26
end for
if Y_k^{LRMI} > \eta then
else
   H_0
end if
```

We analyze the performance of our solution in terms of Receiver Operating Characteristics (ROC) for both primary user and malicious users detection. ROC is the plot of probability of detection vs. probability of false alarm rate.

Primary user detection: The existing method used for soft combining is Equal Gain Combining (EGC) [18], which gives equal emphasis to all the individual measurements. For N users in collaborative spectrum sensing, with EGC rule at the fusion center

$$Y_k^{EGC} = \frac{1}{N} \sum_{i=1}^{N} Y_{i,k}^{j} . w(j), w(j) = 1 \forall i.$$
 (25)

For $Y_k^{EGC} > \eta$, the primary user status is $H_{1(k)}$ otherwise $H_{0(k)}$. Since we know the weight of each cell, we apply

$$Y_k^{LRMI} = \sum_{i=1}^{N} Y_{i,k}.w(j), w(j) = \frac{\beta_k(c_j)}{\sum \beta_k(c_j)} \forall j.$$
 (26)

We need to normalize the location weights at each sensing slot as there may not be any report originated from a cell(s). The probability of detection and false alarm per slot for collaborative spectrum sensing under LRMI is given as

$$Q_{d(k)} = Pr(Y_k^{LRMI} > \eta | H_{1(k)})$$
 (27)

$$= \int_{\gamma_{(k)}^j} Q_m(\sqrt{2TWx}, \sqrt{\eta}) f_{\gamma_{(k)}^j}(x) dx \quad (28)$$

$$Q_{f(k)} = \frac{\Gamma(NTW, \eta/2)}{\Gamma(NTW)}$$
 (29)

where $\gamma_{(k)}^j = \sum_{i=1}^N \gamma_{i,k}^j$. $\gamma_{(k)}^j$ is log-normally distributed. The probability of detection $\overline{Q_d}$ and false alarm $\overline{Q_f}$ for the system based on LRMI overall k is

$$\overline{Q_d} = \frac{\sum_{k=1}^K Q_{d(k)} \delta_{H_{1(k)}}}{\sum_{k=1}^K \delta_{H_{1(k)}}}$$
(30)

$$\overline{Q_f} = \frac{\sum_{k=1}^{K} Q_{f(k)} \delta_{H_{0(k)}}}{\sum_{k=1}^{K} \delta_{H_{0(k)}}}$$
(31)

 $\delta_{H_{1(k)}}=1$ if the primary user is active at slot k and $\delta_{H_{0(k)}}=1$ if primary user is inactive at slot k.

Malicious user detection: ROC for malicious user detection for user u_i ,

$$P_d(\xi) = \frac{1}{|\mathcal{M}|} \sum_{u_i \in \mathcal{M}} \sum_{k=1}^K \frac{\delta(T_k(u_i) < \xi | u_i \in \mathcal{M})}{K}$$
(32)

$$P_d(\xi) = \frac{1}{|\mathcal{M}|} \sum_{u_i \in \mathcal{M}} \sum_{k=1}^K \frac{\delta(T_k(u_i) < \xi | u_i \in \mathcal{M})}{K}$$
(32)
$$P_f(\xi) = \frac{1}{N-|\mathcal{M}|} \sum_{u_i \notin \mathcal{M}} \sum_{k=1}^K \frac{\delta(T_k(u_i) < \xi | u_i \notin \mathcal{M})}{K}$$
(33)

 \mathcal{M} is the set of malicious users. where $P_d(\xi)$ is the probability primary user is detected by the fusion center applying LRMI and $P_f(\xi)$ is the corresponding false alarm rate. There is no closed form solution for \bar{Q}_d for log-normal fading [18] and therefore, we evaluate the system numerically.

VII. PERFORMANCE EVALUATION

In this section, the performance of LRMI is compared with MD both in terms of ROC for malicious user detection and ROC for primary user detection. Results demonstrate that LRMI consistently outperforms MD in mobile CRN. The impact of collaborative secondary users, malicious users, secondary users mobility and cell-size is investigated. We use MATLAB to simulate the system.

$$m_{u_i}(A_k = T) = \frac{m_{u_i}(A_{k-1} = T)m_{u_i}(A_k = T) + m_{u_i}(A_{k-1} = T)m_{u_i}(A_k = \{T, -T\}) + m_{u_i}(A_{k-1} = \{T, -T\})m_{u_i}(A_k = T)}{1 - [m_{u_i}(A_{k-1} = T)m_{u_i}(A_k = -T) + m_{u_i}(A_{k-1} = -T)m_{u_i}(A_k = T)]}$$
(23)

A. Simulation Settings

Cognitive Radio Network Settings: We consider the region of interest to be 1000 m away from primary user. The region is $1000 \text{ m} \times 1000 \text{ m}$ and is divided into grid with L cells of equal area. We take average velocity V = 20m/s, cells L = 9 and sensing time K = 120s for all simulation results unless otherwise mentioned. The secondary users send their location coordinates along with the sensing report during each sensing slot. The noise power is -110dBm and primary user transmit power is 200mW. We assume the users never pause. The sensing duration of all secondary users is 1ms [12] and the users sense after every 1s. We choose users to sense after every 1s, as FCC requires secondary users to evacuate the spectrum in 2s when primary user becomes active. The time-bandwidth product for our simulation is 5. The path-loss exponent is selected randomly from 3 to 6 for each cell and shadowing between 2 to 20 dB. For simulation purpose, we assume the attack strength is $\Delta \sim \mathcal{N}(-10dBm, -5dBm)$ which fusion center is oblivious of. M is used to denote the number of malicious users in the system. We evaluate the system numerically. The malicious users detection threshold, $\xi = 0.5$. We take $\eta = -120dBm$ to -20dBm with a step size of 0.5dBm for primary user ROC. $p_B = p_I = 0.5$

Mobility Settings: Since the sensing duration ($\sim 1 \mathrm{ms}$ - $10 \mathrm{ms}$) is so small, we assume the users locations remain unchanged during each sensing. The IEEE 802.22 needs spectrum evacuation to be in 2 seconds by the secondary users when primary user become active. Hence, the time between two sensing is small (<=2 seconds), it will be unrealistic to consider the mobility models with sudden and uncorrelated change in speed and direction during each sensing slot. Therefore, we consider Smooth Random Mobility Model [22], which considers the two stochastic processes, speed (Vm/s) and direction (θ) to have their values correlated to the previous one in order to avoid unrealistic patterns. The speed/direction changes occur according to a poisson process over the time. The acceleration of all secondary users are $\pm 4m/s^2$. The speed changes on an average every 25 seconds.

B. Impact of secondary users

In this scenario, we use the settings described in the previous section. We study the performance LRMI and MD for malicious user detection in Figure 4 in mobile CRNs with average velocity of secondary users being V=20m/s. We vary number of secondary users in collaborative spectrum sensing. From Figure 4, it is obvious that as the total number of secondary users increases (N=5,10,20) keeping percentage of malicious users constant (20% malicious nodes), the system performance improves. For N=10, with decrease in number of malicious nodes in the system (M=4,M=3,M=2), LRMI performance improves. MD gives a very high false

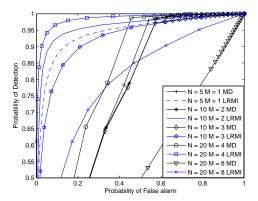


Fig. 4. Impact of secondary users - ROC for malicious user detection at V=20m/s

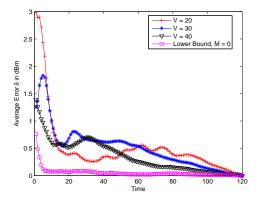


Fig. 5. Maximum average error in LR measurements for M=2, N=10.

alarm rate. It ignores the information that honest users can be at poor locations at times due to their mobility.

C. Impact of mobility

Due to mobility, the number of cells changes per unit time for mobile users increases with the speed for a fixed cell-area and cell-size [20], increasing user-diversity in a cell. We see the impact of mobility in evaluation of location trust in Figure 5. The maximum average error $\delta = \max_{c_j} |E_k(c_j) - E(c_j)| \forall k$ is the maximum error incurred in calculating sensing measurements for location reliability across all cells where $E_k(c_j)$ is the average sensing measurement evaluated at slot k. We find that as the speed is increased, the average error decreases faster with increase in sensing slots with 20% malicious nodes. The lowerbound δ is evaluated with no malicious nodes (M=0) in the system.

To see the effect of user diversity in a cell with respect to ROC, we further evaluate the performance of LRMI with malicious and non-malicious data for calculating LR in Figure 6. Note for MI, the data contains reports from malicious users.

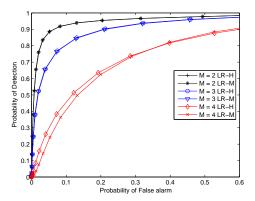


Fig. 6. Comparison of ROC curves for malicious user detection with LR evaluated both with honest data (LR-H) and malicious data (LR-M) for N=10

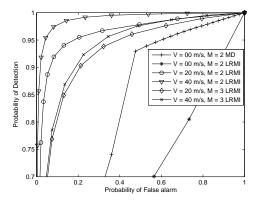


Fig. 7. Impact of velocity - ROC for malicious user detection with secondary users ${\cal N}=10.$

LR-H is for evaluation of LR with honest users in the system and LR-M is for the evaluation of LR with malicious data in the system. We find that the performance in both LR-H and LR-M cases differ only when the number of malicious nodes in the system is as high as 40%.

We study the performance of LRMI with different average velocity of secondary users. Figure 7 evaluates system performance for V=0m/s, V=20m/s and V=40m/s. As the average velocity of users is increased, performance of LRMI improves. MD performs better than LRMI at V=0m/s but for mobile secondary users, LRMI outperforms MD (Fig.7). Performance of LRMI further increases when the average speed of the mobile users is increased from 20m/s to 40m/s. Thus mobility aids in malicious user detection in collaborative spectrum sensing.

D. Impact of number of LR sensings

We find from Figure 8, that for a fixed setting of N,L,M and V, as the K is increased, the performance of malicious detection using LRMI increases. Implicitly, with increased number of sensings, the $\beta_k(c_j)$ $\forall j$ converges to the actual weight of each cell. For N=10,M=2,V=20m/s,L=9,~K=180sec performs better than K=120sec and K=60sec. Similarly for the same settings for M=3, the

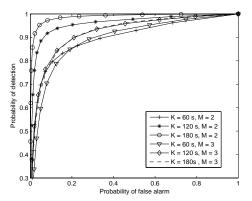


Fig. 8. Impact of LR sensings - ROC for malicious user detection. N = 10, L = 9, V = 20 m/s

system with K=180sec performs better than K=60sec. In addition, it is interesting to observe that there exists a performance upper bound for each M. In other words, increasing the number of sensing will no longer help to increase the performance after the convergence time, which should be between K=60sec and K=120sec for M=3. Note that the convergence time for M=2 may be larger than K=180sec, and this may be the reason why we cannot observe the upper bound for M=2. Moreover, the performance of M=2 is generally better than the one of M=3, which is reasonable as the increasing number of attackers will make them become less minority.

E. Impact on primary user detection

We evaluate complementary ROC for primary user detection for LRMI and MD approach with different number of malicious users in the system. For 10 secondary users at average speed of V=20m/s in the system, we increase the number of malicious nodes (M=1,2,3) and simulate the performance of both LRMI and MD. We take threshold value, $\xi=0.5$ and filter out user reports whose $T_k(u_i)$ is below ξ . The Figure 9 shows LRMI performs better than MD for all cases.

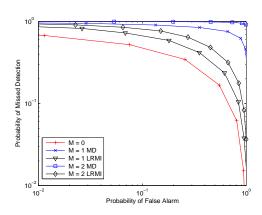


Fig. 9. Complementary ROC for primary user detection with N=10, V=20m/s.

VIII. CONCLUSIONS & FUTURE WORK

We studied the performance of spectrum sensing under different path-loss and fading conditions and came up with a solution fitting for mobile CRNs. The numerically simulated results showed that our approach (LRMI) greatly improves malicious detection in mobile CRNs and hence, performance of collaborative-spectrum sensing for primary user detection. Thus mobile CRNs, need to be evaluated considering both the location from where the report was generated and who has generated the report. Mobility is also found to be an aiding factor in malicious users detection. The simulation results also show that as the average velocity of the secondary users in the system increases, the ROC curves for the system improves.

An interesting extension of the work will be to evaluate how malicious users can exploit mobility to their advantage and avoid getting detected. The primary user is static in our current model. Another future work will be to study the malicious detection in a scenario when primary user is mobile.

REFERENCES

- [1] I. F. Akyildiz, W. Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks*, vol. 50, pp. 2127–2159, Sep. 2006.
- [2] "DARPA Next Generation Communications Program," http://www.sharedspectrum.com/resources/darpa-next-generation-communications-program/.
- [3] "Topic 8: Cognitive Radio for Public Safety," http://transition.fcc.gov/pshs/techtopics/techtopic8.html.
- [4] H. Li and Z. Han, "Catching attacker(s) for collaborative spectrum sensing in cognitive radio systems: An abnormility detection approach," in *IEEE Dynamic Spectrum Access Networks*, DySPAN, 2010.
- [5] P. Kaligineedi, M. Khabbazian, and V. Bhargava, "Secure cooperative sensing techniques for cognitive radio systems," in *IEEE International Conference on Communications, ICC*, 2008.
- [6] O. Fatemieh, R. Chandra, and C. A. Gunter, "Secure collaborative sensing for crowdsourcing spectrum data in white space networks," in IEEE Dynamic Spectrum Access Networks, DySPAN, 2010.
- [7] R. Chen, J.-M. Park, and K. Bian, "Robust distributed spectrum sensing in cognitive radio networks," in *IEEE International Conference on Computer Communications*, INFOCOM, 2008.
- [8] P. Kaligineedi, M. Khabbazian, and V. K. Bhargava, "Malicious user detection in a cognitive radio cooperative sensing system," *IEEE Trans*actions on Wireless Communications, vol. 9, no. 8, pp. 2488–2497, 2010.
- [9] A. W. Min, K. G. Shin, and X. Hu, "Attack-tolerant distributed sensing for dynamic spectrum access networks," in *IEEE International Confer*ence on Network Protocols, ICNP, 2009.
- [10] W. Wang, H. Li, Y. Sun, and Z. Han, "Catch it: Detect malicious nodes in collaborative spectrum sensing," in *IEEE Global Communications Conference*, Globecom, 2009.
- [11] S. M. Mishra, R. Tandra, and A. Sahai, "Coexistence with primary users of different scales," in *IEEE Dynamic Spectrum Access Networks*, DySPAN, 2007.
- [12] "IEEE 802.22 WRAN WG on broadband wireless access standards." www.ieee802.org/22.
- [13] A. W. Min and K. G. Shin, "Impact of mobility on spectrum sensing in cognitive radio networks," in ACM workshop on Cognitive radio networks, CoRoNet, 2009.
- [14] S. Jana, K. Zeng, and P. Mohapatra, "Trusted collaborative spectrum sensing for mobile cognitive radio networks," in *IEEE International Conference on Computer Communications, INFOCOM*, 2012.
- [15] H. Urkowitz, "Energy detection of unknown deterministic signals," Proceedings of the IEEE, vol. 55, no. 4, pp. 523–531, 1967.
- [16] A. Goldsmith, Wireless Communication. Cambridge University Press, 2006
- [17] R. L. Moses, O. L. Moses, D. Krishnamurthy, and R. Patterson, "A self-localization method for wireless sensor networks," *EURASIP Journal on Applied Signal Processing*, vol. 4, pp. 348–358, 2002.
- [18] A. Ghasemi and E. S. Sousa, "Opportunistic spectrum access in fading channels through collaborative sensing," *Journal of Communications*, vol. 2, pp. 71–82, 2007.

- [19] D. Cabric, A. Tkachenko, and R. W. Brodersen, "Experimental study of spectrum sensing based on energy detection and network cooperation," in ACM 1st International Workshop on Technology and Policy for Accessing Spectrum, TAPAS, 2006.
- [20] C. Bettsetetter, H. Hartenstein, and X. Perez-Costa, "Stochastic properties of the random waypoint model," Wireless Networks, vol. 10, no. 5, pp. 555–567, Sep. 2004.
- [21] K. Sentz, "Combination of evidence in dempster-shafer theory, Ph.d. Thesis," 2002.
- [22] C. Bettstetter, "Smooth is better than sharp: A random mobility model for simulation of wireless networks," in 4th ACM International Workshop on Modeling, Analysis, and Simulation of Wireless and Mobile Systems, MSWiM, 2001.

APPENDIX

The basic concepts of D-S theory used in formulation are

- Frame of discernment (Θ) denotes a set of mutually exclusive and exhaustive hypotheses about problem domains. 2^Θ is the power set of Θ.
- 2) Belief mass assignment (bma), represented by m, defines a mapping of the power set to the interval between 0 and 1.

$$m: 2^{\Theta} \to [0\ 1], m(\phi) = 0, \sum_{A \in 2^{\Theta}} m(A) = 1$$

3) Rule of bma combination for elements A, B, C in the power set, the D-S rule of combination is given as

$$m(C) = \frac{\sum_{A \cap B = C, C \neq \phi} m(A)m(B)}{1 - \left[\sum_{A \cap B = \phi} m(A)m(B)\right]}$$



Shraboni Jana is currently a Ph.D. student in the Department of Electrical and Computer Engineering at University of California, Davis. She received her M.S. from University of California, Irvine and B.E. from Regional Engineering College, Durgapur (now National Institute of Technology, Durgapur), India. Her research interests are in wireless and multimedia networking.



Kai Zeng received his Ph.D. degree in Electrical and Computer Engineering at Worcester Polytechnic Institute (WPI) in 2008. He obtained MS in Communication and Information Systems and BS in Communication Engineering both from Huazhong University of Science and Technology, China in 2004 and 2001, respectively. He was a postdoctoral scholar in the Department of Computer Science at University of California, Davis (UCD) from 2008 to 2011. He joined the Department of Computer and Information Science at University of Michigan

Dearborn as an assistant professor in 2011. He is a recipient of the U.S. National Science Foundation Faculty Early Career Development (CAREER) award in 2012. He won Excellence in Postdoctoral Research Award at UCD in 2011 and Sigma Xi Outstanding Ph.D. Dissertation Award at WPI in 2008. His current research interests are in wireless network security, physical layer security, cognitive radio networks, energy efficiency, and cyber-physical systems.



Wei Cheng received his Ph.D. degree in Computer Science from the George Washington University in 2010, and B.S. degree in Applied Mathematics, M.S. degree in Computer Science both from National University of Defense Technology, China, in 2002 and 2004. Currently, he is an Assistant Professor at Virginia Commonwealth University. He has worked as a Postdoc Scholar at University of California Davis. His research interests span the areas of wireless network, cyber-physical networking system, and algorithm design and analysis. In particular, he is

interested in localization, security, routing, and RFID system on roads.



Dr. Prasant Mohapatra is currently the Tim Bucher Family Endowed Chair Professor and the Chairman of the Department of Computer Science at the University of California, Davis. He was/is on the editorial board of the IEEE Transactions on Computers, IEEE Transactions on Mobile Computing, IEEE Transaction on Parallel and Distributed Systems, ACM WINET, and Ad Hoc Networks. He has been on the program/organizational committees of several international conferences. Dr. Mohapatra received his doctoral degree from Penn State University in

1993, and received an Outstanding Engineering Alumni Award in 2008. He also received an Outstanding Research Faculty Award from the College of Engineering at the University of California, Davis. He is a Fellow of the IEEE and AAAS. Dr. Mohapatra's research interests are in the areas of wireless networks, mobile communications, sensor networks, Internet protocols, and QoS.