Location Spoofing Attack and Its Countermeasures in Database-Driven Cognitive Radio Networks

Kexiong (Curtis) Zeng, Sreeraksha Kondaji Ramesh and Yaling Yang
Department of Electrical and Computer Engineering
Virginia Polytechnic Institute and State University
Email: {kexiong6, kondaji, yyang8}@vt.edu

***Abstract*—The recent FCC ruling has enforced *database-driven* cognitive radio networks (CRNs), which include white space networks in TV bands (TV band CRNs) and newly proposed small cell networks in 3.5 GHz (3.5 GHz CRNs). In database-driven CRNs, a secondary user (SU) queries the database for available spectrum at its location. However, this creates a critical vulnerability to GPS spoofing attacks. Under this attack, an adversary compromises SUs' GPS localization system, which results in SUs querying the database with false locations and obtaining incorrect spectrum information. In this paper, we examine the impact of GPS spoofing attacks in database-driven CRNs and propose corresponding spoofing attack detection and countermeasure solutions. To the best of our knowledge, this is the first paper to study the impact and countermeasures of GPS spoofing attacks in database-driven CRNs.**

I. Introduction

Today's explosion of data communication needs is stretching the capacity limit of wireless networks. The principal limiting factor for the capacity of wireless networks is spectrum. Military communications, broadcast TV, WiFi, cellular systems and many such applications all compete for spectrum. Currently, the available spectrum is licensed to these different applications. However, some applications, like cellular systems, have grown much faster than others, such as broadcast radio and broadcast TV. This leads to overcrowding in some spectrum bands and underutilization of other spectrum bands. Cognitive radio networks (CRNs) based on dynamic spectrum access help to ease this imbalance in spectrum utilization. In CRNs, two kinds of users are defined - primary users (PUs) and secondary users (SUs). PUs always have the full access to the spectrum whenever they need it. SUs are permitted to use the spectrum only if they do not interfere with the PUs.

Two potential applications are in TV White Spaces (TVWS) and 3550-3650 MHz band (3.5 GHz Band). TVWS refers to the unused TV channels in any location. In November 2008, FCC issued a report that specifies the requirements for SUs to operate in licensed TV bands [1]. According to the requirements, a trusted geolocation database will be used to assign spectrum to SUs so that they will impose no interference to licensed PUs. In December 2012, FCC proposed a new Citizens Broadband Service in the 3.5 GHz Band [2]. The Citizens Broadband Service incorporates database-driven dynamic spectrum access and small cell technology[1] to enable more efficient use of radio spectrum.

With FCC laying stress on database-driven methods, it is imperative to examine the security threats of such methods. One critical security loophole is that this method relies on an SU to obtain its location information from GPS (Global Positioning System). GPS has been shown to be very vulnerable to spoofing attacks that make GPS receivers lock on spoofed GPS signals and compute false locations. With the development of programmable radio platforms such as USRPs (Universal Software Radio Peripheral), it has become quite easy to build GPS signal simulators to generate spoofed GPS signals with arbitrary date, time and location. Although there are some proposed countermeasures for GPS spoofing attacks, none of them have been implemented in commercial GPS receivers yet.

Since database-driven CRNs depend on GPS system and GPS system is vulnerable to attack, the focus of this paper is to understand the impact of GPS spoofing attacks on database-driven CRNs and explore several methods to defend against such attacks. Our study includes CRNs in both TV bands and 3.5 GHz band. We formulate various attack models and identify metrics for studying the impact of such attacks. Through simulation, we demonstrate that GPS spoofing attacks can not only create denial-of-service attacks, but also cause harmful interference to PUs. To detect and defend against GPS spoofing attacks, we propose three kinds of schemes: centralized detection scheme (CDS), environmental-radio-based location verification (ELV) and peer location verification (PLV). Moreover, we comprehensively analyze the effectiveness and limitations of the proposed countermeasures. In order to evaluate the proposed schemes, we practically implement the ELV and analyze the PLV by simulations. Experimental and simulation results show that the proposed countermeasures can effectively defend against GPS spoofing attacks on database-driven CRNs and mitigate each other's limitations if these three countermeasures are combined.

The contributions of this work are summarized as follows:

- To the best of our knowledge, this is the first paper to examine the impact of GPS spoofing attacks in dynamic spectrum access, which used to be two unrelated topics.
- We formulate different attack models and analyze possible damage of such attacks. Based on the knowledge that a single attacker can easily spoof a group of SUs to an arbitrary location, we define random and optimal attack model in TV band CRNs. Simulation results show that even a simple random attack can cause PU interference in extremely sparse network with only 100 SUs in a $16km \times 16km$ cell. Additionally, we discuss attack models in 3.5 GHz CRNs, which can result in serious interference between SUs and critical DoD radar system.
- We propose various solutions for defending against GPS spoofing attacks. By thoroughly discussing the

[1]Small cells are low-powered wireless base stations (also SUs) intended to cover small indoor or outdoor areas.

effectiveness and limitations of each scheme, we demonstrate that the proposed countermeasures can significantly mitigate the threat of GPS spoofing attacks if a hybrid approach is taken.

- We evaluate the countermeasures by implementation, simulation and mathematical analysis. We implement the ELV by using a commercial spectrum analyzer to collect data from real world. The experiment results show that the ELV can effectively thwart GPS spoofing attacks. Besides, we present an analytical model for the PLV and use spectral analysis to derive an upper bound for convergence time. Finally, we evaluate the PLV by extensive simulations. Simulation results show that our mechanism can achieve nearly 0 false negative and false positive in most cases.

The rest of the paper is organized as follows. We discuss related work in Section II, introduce GPS spoofing attack in Section III, formulate attack model in Section IV, propose countermeasures in Section V, present implementation and evaluation in Section VI, discuss limitations and future work in Section VII and conclude the paper in Section VIII.

II. Related Work

Our work is related to existing works in two areas: (1) GPS spoofing attacks and (2) security issues in database-driven CRNs.

As a narrow band technology that relies on weak satellite signals with poor authentication mechanism, GPS signal can be easily jammed, delayed, or emulated in a large area with a small amount of transmit power. In 2001, the Volpe report identified jamming and spoofing as vulnerabilities of GPS receivers [3]. Starting with this report, GPS spoofing attacks are studied in several publications. In [4], a WelNavigate GS720 GPS signal simulator along with two GPS amplifiers were used to attack a GPS receiver in a truck. The authors of [5] built a software-defined GPS receiver-spoofer and used it to successfully demonstrate a spoofing attack. In June 2013, Dr. Todd Humphreys and his students from University of Texas at Austin successfully performed a "proof-of-concept" GPS spoofing attack [6]. They took over the navigation system of the luxury yacht "White Rose" with their GPS spoofing device built for about $\$3,000$. Apart from numerous demonstrations of spoofing attacks, there has also been an in-depth analysis about the effect of a GPS spoofing attack on a group of receivers. Tippenhauer et al. proved that any number of receivers can easily be spoofed to one arbitrary location by a single attacker [7]. In order to thwart GPS spoofing attacks, some prospective countermeasures are proposed. Kuhn presented an asymmetric security mechanism for GPS system [8]. In [9], the authors proposed three diverse defense mechanisms based on knowledge of location, time and Doppler shift. A comprehensive summary of anti-spoof methods has been provided by [10]. All these existing anti-spoof methods require modifications of existing GPS receivers and are not available in commercial products.

Along with GPS vulnerabilities, the other area related to our work is security issues in database-driven CRNs. The latest IETF draft discussed the potential attacks towards the querying progress between an SU and the database [11]. In such an attack, the adversary can track locations and identity of the SU and respond the SU with malicious spectrum information, which results in PU interference. It also introduced Transport Layer Security (TLS) as a solution to mitigate the threats. [12] is the latest work that focuses on location privacy in database-driven CRNs. This work pointed out that an adversary can infer an SU's location through the SU's channel usage and proposed corresponding countermeasure. To the best of our knowledge, location spoofing attacks in database-driven CRNs have not been systematically discussed in any existing work yet.

III. Overview of GPS Spoofing Attack and Countermeasures

In this section, we introduce the background of GPS spoofing attacks and countermeasures for such attacks, which are necessary for building our attack model.

A. GPS Spoofing Attack

An attacker can launch a spoofing attack by using GPS spoofing device (like a GPS signal simulator) to generate and broadcast counterfeit GPS signals to the target receiver. The attacker can also mount a replay attack by rebroadcasting signals captured at a different place or time. A spoofing attack begins by transmitting signals synchronized with the genuine GPS signals observed by the target receiver. Then, the fake signals gradually overpowers the authentic GPS signals. Finally the target receiver deviates from real GPS signals and locks on the counterfeit signals. After taking over the GPS, the attacker can create arbitrary locations for the target receiver. In this paper, we consider a single attacker with an omni-directional antenna. It has been proved in [7] that all receivers in the tranmission range of an attacker can be spoofed to any location L' using a single attacker antenna. The reason is that the computed location is determined by the time-differences of arrival of the individual satellite signals. If the spoofed signals are all sent from the same attacker antenna, all victims will obtain the same time-differences, which then result in the same spoofed location L' as shown in Figure 1.

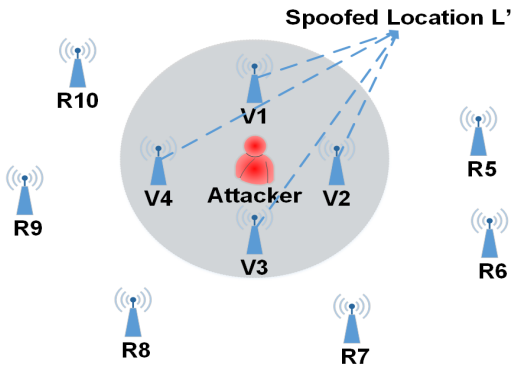


Fig. 1. Single-attacker GPS-spoofing attack

B. Countermeasures

There are a few proposed countermeasures to thwart GPS spoofing attacks. One possible solution is using self-check algorithms to find abnormality in the GPS receiver. For example, the receiver can monitor absolute or relative power of carriers, check the Doppler shift or detect abrupt changes. Another possible countermeasure is using smart antennas to detect the angle-of-arrival of each satellite signal. Since all counterfeit signals are transmitted from the exact same direction, the

attack is easy to be detected. However, all countermeasures discussed above have not been adopted and tested in civilian GPS receivers yet and they all demand some modifications on the current GPS receiver design.

IV. Attack Model

In database-driven CRNs, the system assigns available spectrum to an SU entirely based on its location report. This unique mechanism causes a serious security loophole, which can be utilized by GPS spoofing attackers. In what follows, we formulate attack model and present possible damage of GPS spoofing attacks in database-driven TV band CRNs and 3.5 GHz CRNs respectively.

A. Attack Model for Database-driven TV band CRNs

In this section, we first study the framework for simulating database-driven TV band CRNs, then introduce a random and optimal attack model, and finally present simulation results for the attack.

1) Database and Spectrum Allocation: In order to evaluate the impact of GPS spoofing attacks in database-driven TV band CRNs, we need to define the model of such a network. We use WhiteSpaceFinder [13], which is a database that uses Longley Rice model with terrain data along with TV-tower information to predict the availability of white spaces at any location. We consider a single cell coverage area of 16km-radius in Blacksburg, VA region with $100m \times 100m$ resolution.

With the objective of maximizing spectrum utilization, we use round-robin scheduling and list-coloring based greedy algorithm proposed by Wang and Liu in [14] to assign available TV channels to SUs. In each time slot, we first sort all channels in ascending order of their geographic availability. A channel's geographic availability is defined as the number of grids where the channel is available. Then, we assign these channels to SUs following the sorted sequence. For each channel, there are a bunch of candidate SUs that can use the channel at their locations. The SU that has waited the longest for a channel will be assigned the channel. When a tie exists, the SU whose location has less available channels will be picked. We do not involve link degrees in the spectrum allocation algorithm, because we assume that there is no spatial reuse of channels in a single cell.

2) Random Attack Model: A random attack is launched by an attacker that knows nothing about the spectrum database. The attacker just simply spoofs all SUs in his transmission range to a random location in the cell. For demonstration purpose, we assume the range of the attacker to be 1 km. A 1 km range is very conservative for any spoofing devices, which allows us to not impose any heavy restrictions on the attacker's capabilities.

We use this model to evaluate the impact of the attack. We first evaluate how a GPS spoofing attack may cause PU interference in TV band CRNs. We capture PU interference by the number of SUs who interfere with PUs in the simulated time. We run 50 runs of simulations under 10 different SU densities (i.e. 5 simulations for each density). Figure 2 shows the largest number of SUs interfering with PU for each SU density level. We can observe that as we increase SU density, PU interference becomes more serious. Also, even in an extremely sparse network with only 100 SUs in the 16km-radius base station coverage range, we still observe PU interference.

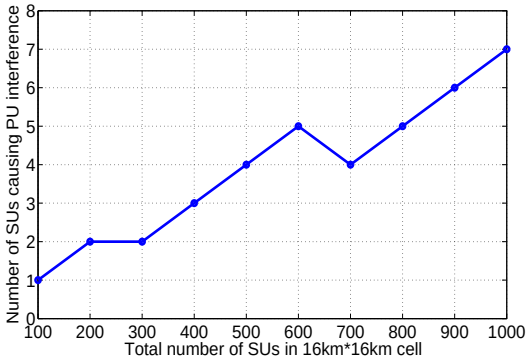


Fig. 2. Variation in severity of PU interference with SU density.

We also examine the probability distribution of PU interference among 51 TV channels. Figure 3 shows that most of the PU interference occur on TV channels 29, 33 and 21. This is due to the large variation in the availability of these 3 channels over the simulated geographic area. For a channel whose geographic availability varies a lot, it is more likely that GPS spoofing attacks can cause interference.

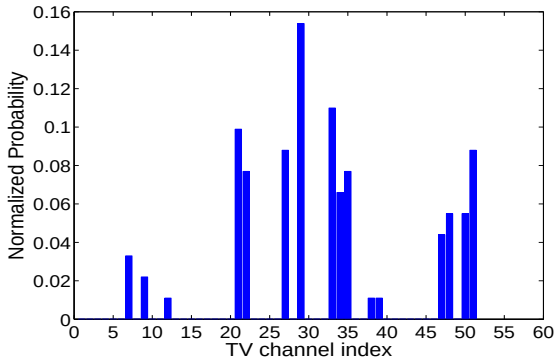


Fig. 3. There are 21 available channels in total. In a 1000-SU network, 16 of them are interfered with at least once in 50 simulations.

3) Optimal Attack Model: We define an attack as optimal if the attacker has access to the database and knows complete geolocation information of all registered SUs. Based on that knowledge, he runs the same spectrum allocation algorithm on his own computer and uses brutal force algorithm to traverse all grids in the cell with different available channels to find the optimal solution, i.e. the spoofed location that maximizes the impact of his attack. In order to analyze the damage caused by the optimal attack, we run 30 individual simulations. In each case, we launch a random and optimal attack respectively. Figure 4 compares the performance of the two kinds of attacks mentioned above. We average the number of SUs who interfere with PUs over simulated time and plot the cumulative distribution function (CDF). As we can see, the performance of the optimal attack is significantly better than the random attack. The optimal attack serves as the upper bound on the damage that GPS spoofing attack can achieve.

B. Attack Model for Database-driven 3.5 GHz CRNs

In this section, we first introduce the geographic exclusion zones in 3.5 GHz and then define two kinds of attacks targeting at database-driven 3.5 GHz CRNs.

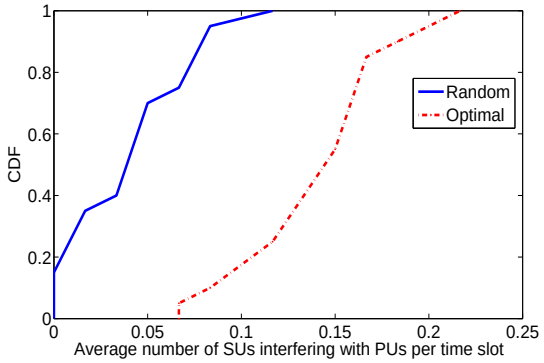


Fig. 4. The performance of random and optimal attack in a 1260-SU network.

1) Geographic Exclusion Zones: NTIA's Fast Track Report collects the information of critical DoD radars that operate from sea, land and airborne positions in or near 3.5 GHz band [15]. In order to prevent the potential interference between these radars and commercial WiMAX broadband technology, it determines "geographic exclusion zones" by calculating the separation distances between radar systems and a prospective outdoor WiMAX system. As illustrated in Figure 5, the geographic exclusion zones are imposed along the East, West, and Gulf Coasts. Exclusion zones are developed not only to protect incumbent DoD radar system but also to prevent interference from the high-powered radar operations into federal uses. Based on the exclusion zones in the Fast Track Report, FCC proposes similar geographic exclusion zones that considers low-powered small cell deployment in 3.5 GHz.

2) Attack Model: In database-driven 3.5 GHz CRNs, low-powered wireless base stations are required to submit their accurate locations to query the database before operating. By checking the location with geographic exclusion zones, the database decides to activate the base station or not. However, this spectrum allocation mechanism creates a security loophole. Because an attacker is able to spoof a group of base stations to an arbitrary location, it can easily manipulate the spectrum allocation through manipulating the locations of base stations. Generally speaking, there are two kinds of attacks. One is spoofing the base stations from places inside the exclusion zones to a location outside, which causes all these base stations to interfere with critical DoD radars. The other attack causes SUs outside of exclusion zones to be identified as inside and hence results in denial of service to these SUs.

Fig. 5. Geographic exclusion zones in 3.5 GHZ in NTIA's Fast Track Report.

V. GPS Spoofing Attack Detection and Countermeasures

In this section, we propose different solutions to detect and defend against GPS spoofing attacks, and analyze their effectiveness and limitations respectively.

A. Centralized Detection Scheme

In database-driven CRNs, all SUs have to register their locations in the database, which can be utilized by the system to detect GPS spoofing attacks. For example, the system can ask every active SU to update his location information periodically. Based on the reported location and time stamp, the system maintains trace information for all SUs. Once abnormal SU behaviors are detected, like a group of nearby SUs suddenly move to one place at tremendous speed, the system considers the network as under attack.

While the centralized detection scheme is fairly intuitive, it might result in false alarms. For example, thousands of people will crowd into a single stadium to watch a football game, which can cause a lot of SUs simultaneously report the same location to the database. A centralized detection scheme that is not intelligent enough may mistake this as a result of GPS spoofing attacks. In addition, the database-based approach can only detect GPS spoofing attacks but cannot restore the normal network operation with correct location information. Finally, the tracking and analysis of every SU's location trace may impose a heavy overhead in the database system and also create potential violation to user privacy.

B. Environmental-radio-based Location Verification

In this section, we propose an Environmental-radio-based location verification (ELV) method to detect and thwart GPS spoofing attacks. Our method leverages the fact that an SU is often also a software defined radio (SDR), which can tune to different wireless systems operating at a wide range of spectrum bands. In addition, today's wireless devices often are equipped with multiple communication interfaces (e.g. WiFi, TV, FM and cellular) that work at different frequencies. Thus, in our scheme, an SU listens to the signals emitted by existing infrastructures (i.e. WiFi access points, TV towers and FM towers). We assume that the SU prestores a Radio Environment Map (REM), which contains the fingerprints of ambient wireless signals at different locations [16]. By comparing the local signal fingerprints with the ones in the REM, an SU can estimate his location without relying on GPS. This estimation method can then be used to verify GPS location computation, detect GPS spoofing attacks and also be used as a backup localization mechanism when GPS is under attack.

1) WiFi-based Location Verification: As WiFi access points become increasingly prevalent, using WiFi signals for SUs to detect and countermeasure GPS spoofing attacks will be fairly effective, especially in urban areas. Since a WiFi access points' coverage radius is often less than 100 m, an SU can verify if its GPS location is within the coverage of a specific WiFi access point and also use the WiFi access points' position as a backup location when under attack. Therefore, an SU only needs to narrow down its location to the coverage of a specific WiFi access point, whose information is stored in the REM. The accuracy of this WiFi localization scheme is more than enough to satisfy the location requirement for database-driven networks.

Methodology: In order to determine the surrounding WiFi access points, an SU first decodes the received IEEE 802.11 packets and records the received information like SSID and MAC address. Then, he looks up the local information in the REM to locate the corresponding WiFi access point. Finally, he checks if his GPS location is within the coverage of the WiFi access point. If not, he considers himself as under attack and uses the backup location, i.e. the WiFi access point's position.

As for the situations where an SU hears no WiFi signals, he checks the REM to see if there should be WiFi access points at his location. If so, he also regards himself as a victim under spoofing attack.

Limitations: The limitation of the WiFi-based location verification scheme is obvious. It is of no use when an SU is at the blind spot of the REM, such as rural areas, where the fingerprints have not been collected yet.

2) Television-based Location Verification: Recently, Rosum company has developed a chip called "Alloy" that can use TV broadcast signals to accurately localize people and objects. Rosum TV-positioning technology utilizes the time of arrival of TV broadcast signals and the locations of corresponding terrestrial broadcast television infrastructures to calculate location. With support for a variety of types of TV signals, Alloy is able to provide less than 150-meter accuracy even in the worst case [17]. Terrestrial broadcast TV signals are high-power, low-frequency signals that easily penetrate buildings and urban areas. Furthermore, TV signals have a quite large bandwidth (i.e. from 54 MHz to 890 MHz). Thus, TV signals are robust to jamming or spoofing attacks.

Methodology: TV localization is a great auxiliary positioning method for small cell base stations in database-driven networks, especially in indoors or urban areas where GPS signals are unavailable. In addition, TV localization can detect and work as a backup plan against GPS spoofing attack. For example, the SUs equipped with the Alloy chips can use the location calculated by TV-positioning technology to verify the GPS location. We conservatively assume that the accuracy of GPS and TV localization are 10 m and 150 m respectively. If the error distance between these two estimated locations is 160 m or more, the SU consider himself as under attack and switch to the countermeasure, i.e. TV localization. Additionally, if the SU operates in TV bands, it neither needs extra radios or antennas for TV localization nor does it consume more energy.

Limitations: The "Alloy" chip is still under development and not commercially available in the market. Even if it is available, not every SU can afford the cost of the chip.

3) FM-based Location Verification: FM signals are ubiquitous and widely-available across all environments - outdoor, indoor and urban. Therefore, we exploit FM radio localization to detect and thwart GPS spoofing attacks.

Methodology: To evaluate the effectiveness of FM-based location verification, we develop a two-phase localization system using FM radio received signal strength indicator (RSSI). In the first phase (offline phase), we collect realistic FM RSSI fingerprints by a commercial spectrum analyzer and store them in the REM. In the second phase (online phase), the SU estimates his location by comparing the measured RSSI fingerprints and the pre-constructed model. Then, the SU checks if the error distance between the GPS location and the FM estimated location falls into a reasonable region. If not, a spoofing attack is detected and the SU activates the backup plan, i.e. FM radio localization. Our experiment results show that we can achieve 50-meter accuracy by 8 strongest FM channels and effectively thwart GPS spoofing attacks.

Limitations: The limitations of FM radio localization are as the following: (1) The SU will need an extra FM antenna for signal reception. (2) Its performance highly depends on the number of local FM channels and the number of reference locations collected by the REM.

C. Peer Location Verification

Since environmental localization imposes requirements of hardware or supporting infrastructure, only limited number of SUs can use it to detect and thwart spoofing attacks. Therefore, we propose a distributed peer location verification (PLV) scheme, which propagates from a number of initial SUs to the whole network. We also describe an analytical model and study the efficiency of our mechanism by spectral analysis.

1) Location Verification Scheme: At the very beginning of the PLV process, we assume that there are a certain number of SUs who can use the ELV to verify their GPS locations. Hence, they become initial anchor nodes. Then, each anchor node transmits a r-radius beacon signal containing his position to surrounding SUs with probability β. Anchors transmit beacons in this way to avoid collisions among neighboring anchors. When an unverified SU hears a beacon signal from any anchor, he checks if his GPS location is within radius r of the anchor's location. If so, that SU trusts his GPS location and becomes a new anchor. Otherwise, the SU will infer that he is under GPS spoofing attack and remain silent. As this location verification propagates through the whole cell, an equilibrium is achieved wherein no further nodes can be verified. The verification process ends at this point. If any SU fails the location verification, he indicates that he is the victim of a spoofing attack. The victim SUs can then counter the attack by estimating their locations through the average of the coordinates of surrounding anchor nodes.

The PLV scheme also has its limitations. On the one hand, when SU distribution is sparse, it is possible that some victim SUs fail to hear any anchor node i.e., they are isolated. In such a case, they cannot positively identify whether they are under attack. This can lead to missed detection (i.e. false negative) of GPS spoofing attack. On the other hand, it is possible that some malicious SUs claim themselves as anchor nodes but transmit beacon messages containing false location information to confuse other SUs. These malicious SUs can cause false alarms (i.e. false positive) of GPS spoofing attacks. We discuss these two cases further in the evaluation section.

2) Analytical Model: We propose an analytical model to study the convergence speed of PLV under no-attack condition, which is similar to virus spreading model in random geometric networks [18]. We assume that n SUs, $S_n = \{s_1, s_2, ..., s_n\}$, are located at random positions denoted as $L_n = \{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$, where (x_i, y_i) are uniformly distributed in a 16×16 km^2 area. Each anchor has a beacon transmission range of r. We call two SUs, $s_i, s_j \in S_n$, correlated if and only if $\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \leq r$. We denote this random geometric graph by $G(L_n; r)$.

Our model divides time into discrete slots. During each time slot, an anchor node tries to verify its neighbors by transmitting a beacon signal with probability β. We denote the probability that an SU i is verified at time t as $p_{i,t}$. We assume that an SU i is verified at time t if

- i was already verified before t.
- i was not verified before t, but receives beacon signals from neighboring anchors and gets verified at t.

Hence, we define the probability of an SU i getting verified at time t to be

$$p_{i,t} = p_{i,t-1} + (1 - p_{i,t-1})(1 - \prod_{j \in Neighbors\ of\ i} (1 - \beta p_{j,t-1})) = 1 - (1 - p_{i,t-1}) \prod_{j \in Neighbors\ of\ i} (1 - \beta p_{j,t-1}), \quad (1)$$

where $i = 1, 2, ..., n$. Given a network topology and specific β value, we can numerically solve equation (1) and obtain time evolution of total number of verified SUs $\mathrm{N}_t = \sum_{i=1}^{n} p_{i,t}$.

3) Spectral Analysis of Convergence Time: The efficiency of the PLV mechanism highly depends on the convergence time of location verification process. Therefore, we use spectral analysis on our model to find an upper bound of convergence time.

Assuming β is much smaller than 1, we can approximate (1) as:

$$p_{i,t} = p_{i,t-1} + \beta \sum_{j} p_{j,t-1}. \quad (2)$$

This uses the approximation $(1 - \varepsilon)(1 - \mu) \approx 1 - \varepsilon - \mu$, where $\varepsilon \ll 1, \mu \ll 1$.

Using a column vector $\mathbf{P_t} = (p_{1,t}, p_{2,t}, ..., p_{n,t})^T$ to convert equation (2) to matrix form, we have

$$\mathbf{P_t} = (\mathbf{I} + \beta \mathbf{A})\mathbf{P_{t-1}} = \mathbf{B}\mathbf{P_{t-1}} = \mathbf{B}^t\mathbf{P}_0, \quad (3)$$

where $\mathbf{B} = (\mathbf{I} + \beta \mathbf{A})$ and $\mathbf{A}$ is the adjacency matrix of G.

$\mathbf{V}_{i,A}$ is the eigenvector of $\mathbf{A}$ corresponding to eigenvalue $\lambda_{i,A}$. By definition, we have $\mathbf{A}\mathbf{V}_{i,A} = \lambda_{i,A}\mathbf{V}_{i,A}$. So,

$$\mathbf{B}\mathbf{V}_{i,A} = (\mathbf{I} + \beta\mathbf{A})\mathbf{V}_{i,A} = \mathbf{V}_{i,A} + \beta\lambda_{i,A}\mathbf{V}_{i,A} = (1 + \beta\lambda_{i,A})\mathbf{V}_{i,A}.$$

Hence, $\mathbf{V}_{i,A}$ is also the eigenvector of $\mathbf{B}$ but corresponding to eigenvalue $\lambda_{i,B} = 1 + \beta\lambda_{i,A}$. Using spectral decomposition on real symmetric matrix $\mathbf{B}$, we have

$$\mathbf{B}^t = \sum_{i} {\lambda^t}_{i,B} \mathbf{V}_{i,B} {\mathbf{V}_{i,B}}^T. \quad (4)$$

Sorting the eigenvalues in non-increasing order such that $\lambda_{1,A} \geq \lambda_{2,A} \geq ... \geq \lambda_{n,A}$ and $\lambda_{1,B} \geq \lambda_{2,B} \geq ... \geq \lambda_{n,B}$. Substituting equation (4) into equation (3), we have

$$\mathbf{P_t} = \sum_{i} {\lambda^t}_{i,B} \mathbf{V}_{i,B} {\mathbf{V}_{i,B}}^T \mathbf{P}_0 = \sum_{i} {\lambda^t}_{i,B} \mathbf{C_i}, \quad (5)$$

where $\mathbf{C}_i$ are constant column vectors. Thus, time evolution of total number of verified SUs is

$$\mathrm{N}_t = \sum_{i=1}^{n} p_{i,t} = \sum_{i=1}^{n} ({\lambda^t}_{i,B} \sum_{j=1}^{n} c_{ij}),$$

where c_{ij} is the jth element of the constant matrix $\mathbf{C}_i$.

Furthermore, we can say that

$$\mathrm{N}_t > {\lambda^t}_{1,B} \sum_{j=1}^{n} c_{1j} \Rightarrow t < \log_{(1+\beta\lambda_{1,A})} \frac{N_t}{\sum_{j=1}^{n} c_{1j}}. \quad (6)$$

In equation (6), we see that the upper bound of convergence time is a log function with base $1 + \beta\lambda_{1,A}$ indicating that the convergence time scales very well with N_t and the system can converge fairly fast even with a large N_t.

In order to use simulation to evaluate the correctness of our theoretical analysis, we set up an initial anchor ratio $\gamma = \frac{Number\ of\ initial\ anchors}{Total\ SU\ number}$ and begin each simulation case with $\gamma \cdot n$ randomly chosen initial anchors. To calculate theoretical bound ${\lambda^t}_{1,B} \sum_{j=1}^{n} c_{1j}$, we obtain ${\lambda^t}_{1,B}$ and $\mathbf{V}_{i,B}$ from B and set $p_{i,0} = \gamma$ to match the initial anchor ratio with the simulation settings. Figure 6 plots time evolution of verified SUs in a simulation of 5000-SU network. We can see that the convergence time is indeed upper bounded by what indicates in equation (6).

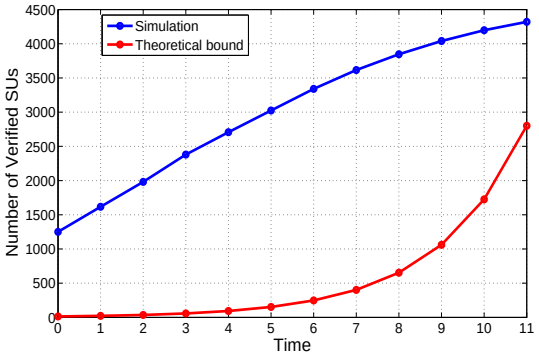


Fig. 6. Simulation results are averaged over 30 individual runs in a 5000-SU network with $\gamma = 25\%$, $\beta = 0.03$ and $r = 500m$.

Since PLV converges very fast, we can reasonably assume that all SUs maintain a static topology in such a short time. Hence, we do not consider the model of SUs with mobility in this paper.

VI. Implementation and Evaluation

In this section, we implement and evaluate the performance of the ELV. We also evaluate the effectiveness of the PLV by simulations.

A. Implementation of ELV

In order to evaluate the proposed ELV, we implement FM-based location verification (We skip the implementation of WiFi-based and TV-location verification. The reason is that for WiFi-based location verification, it is universally known that software defined radios can decode 802.11 packets. As for TV-based location verification, the "Alloy" chip is not commercially available yet.). To implement the FM radio localization, we drive around the Blacksburg, Virginia region to collect real RSSI fingerprints with a commercial spectrum analyzer and evaluate the performance by extensive experiments.

1) Data Collection: We perform the experiment in Blacksburg, Virginia region with an area of around 86.25 km^2. We select 26 locations shown as the red balloons in Figure 7. To measure real RSSI of FM channels, we use a Tektronix MDO4104-6 Mixed Domain Oscilloscope and a v-shaped "rabbit ear" FM antenna, as shown in Figure 8. We take measurements of 17 FM channels at each reference position and divide the measurements into two groups. The first group is called training group and the second group is called testing group. We use Gaussian regression to model the data in

training group as a Gaussian distribution for each reference location. Then, we input the data in testing group to evaluate the performance of the FM-based location verification scheme.

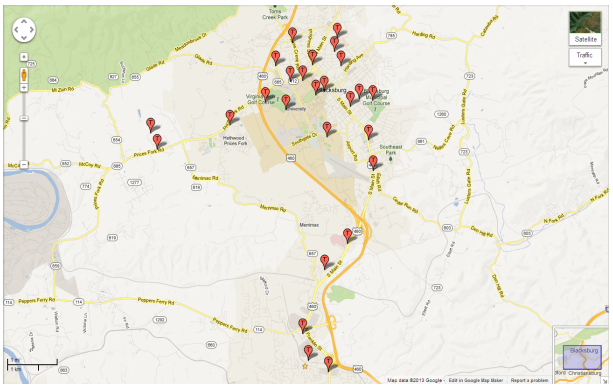


Fig. 7. This picture is obtained from the Google Map and the red balloons indicate the locations of test points.

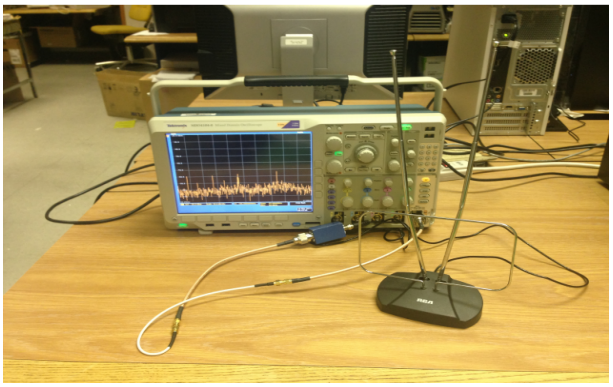


Fig. 8. Tektronix MDO4104-6 Mixed Domain Oscilloscope and the FM antenna.

2) Positioning Algorithm: The positioning algorithm we use is proposed by Fang et al. [19]. Our task is to estimate the locations from the pre-stored reference points. So we use likelihood of each reference position to calculate estimated location represented by numerical latitude and longitude. In this way, the localization problem can be formulated as follows:

$$\bar{L} = \sum_{k=1}^{n} L_k \cdot P_{norm}(\mathbf{C_k}|\mathbf{X}), \tag{7}$$

where

$$P_{norm}(\mathbf{C_k}|\mathbf{X}) = P(\mathbf{C_k}|\mathbf{X}) / \sum_{i=1}^{n} P(\mathbf{C}_i|\mathbf{X}). \tag{8}$$

$\overline{L}$ is the localization output represented by numerical latitude and longitude. L_k is the coordinate of kth reference position and n is the number of reference positions. $P_{norm}(\mathbf{C_k}|\mathbf{X})$ is the normalized likelihood that the observed RSSI vector $\mathbf{X}$ is measured in reference location $\mathbf{C_k}$.

Applying the Bayes' rule, we have

$$P(\mathbf{C}_i|\mathbf{X}) = P(\mathbf{X}|\mathbf{C}_i) \cdot P(\mathbf{C}_i)/P(\mathbf{X}), \tag{9}$$

where $i = 1, 2, ..., n$. In Equation (9), the $P(\mathbf{C}_i)$ is the a priori probability of being at the reference position i. Since we have no prori knowledge of where the SU is, we assume a uniform distribution, i.e. $P(\mathbf{C}_i) = 1/n$.

Applying Equation (9) to Equation (8), we have

$$P_{norm}(\mathbf{C_k}|\mathbf{X}) = P(\mathbf{X}|\mathbf{C_k}) / \sum_{i=1}^{n} P(\mathbf{X}|\mathbf{C}_i). \tag{10}$$

Since we already have the Gaussian distribution model for each reference point $\mathbf{C_k}$ and the observed RSSI vector $\mathbf{X}$, the likelihood $P(\mathbf{X}|\mathbf{C_k})$ is numerically computable.

B. Evaluation of ELV

ELV can only detect spoofing attacks where the spoofed location is relatively far away from an SU's true location. If the distance between the spoofed location and the true location is smaller than the confidence range of ELV's localization scheme, the spoofing attack cannot be detected. To evaluate the performance of ELV, it is important to understand if these undetected attacks can impose significant threat to CRNs.

We will focus on database-driven TV band CRNs rather than 3.5 GHz CRNs because the TV band CRNs have many channels with different availability distribution while in 3.5 GHz CRNs, there is only one channel. Thus, TV band CRNs are more likely to have variations in channel availability in a small area than 3.5 GHz CRNs. Thus, in TV band CRNs, undetected spoofing attacks have a higher chance to cause damage.

The undetected spoofing attack can be modeled as a location spoofing attack with a distance constraint so that the attacker cannot spoof an SU to a false location that is further than the constraint. Different ELV localization accuracy will impose different distance constraint. Our simulation settings are the same as described in Section IV-A. Figure 9 shows the probability that SU interferes with PU when an optimal spoofing attack is launched under the distance constraint. The interference probability is calculated as number of simulation runs that have PU interference over total number of simulation runs. It can be seen that the probability of interference caused by optimal attacks goes up as the distance constraint increases. Particularly, the probability is $\leq 10\%$ when constraint distance is $\leq 100m$. In comparison to the optimal attacks without distance constraint which causes PU interference with 100% probability, the PU interference threat is greatly mitigated.

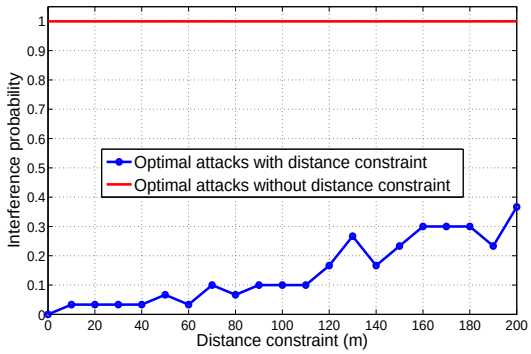


Fig. 9. The probability is calculated by 30 individual simulations in a 84-SU network.

Figure 10 shows the localization error distribution (i.e. mean plus standard deviation) under different number of FM channels. In order to obtain a better localization performance, we always select the strongest FM channels at each test position to calculate estimated location. It can be seen that the FM-based localization can achieve 50-meter accuracy with

8 strongest FM channels. In addition, the transmission range of WiFi access points is generally less than 100m. Thus, comparing with Figure 9, we can conclude that both WiFi and FM-based ELV can effectively limit the impact of spoofing attacks even when these attacks evade their detection. Hence, we can say our ELV mechanism can effectively countermeasure GPS spoofing attacks.

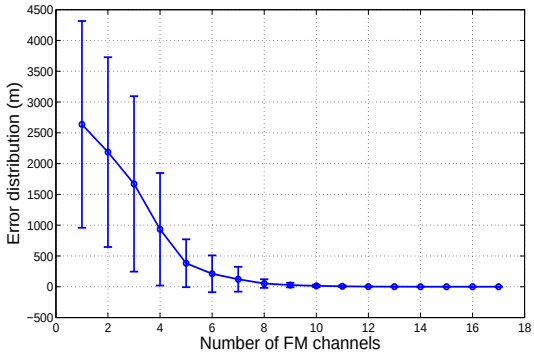


Fig. 10. Localization performance using the strongest FM channels.

C. Evaluation of PLV

PLV scheme may experience missed detection (flase negative) because of isolated SUs and false alarm (false positive) caused by malicious anchor nodes. In this section, we evaluate such situations by simulations. Each simulation plot is averaged over 100 individual runs.

1) Missed Detection: In Figure 11, we vary beacon transmission range r , SU density $n/(16 \times 16km^2)$ and initial anchor ratio γ to exam the performance under different scenarios. We define false negative SU ratio as the number of non-detected victim SUs over the total number of SUs who are not anchors.

As seen from the figure, the false negative SU ratio goes down as one of the three parameters increases. As we state in section V-C, missed detection are resulted from the isolated victim SUs, which cannot be positively identified by the system. Hence, with more initial anchors and larger beacon transmission range, the PLV can propagate more widely. It reduces the number of isolated SUs so that it is more likely to detect spoofing attacks. Additionally, as the density of SUs become larger, the graph is more connected, which also results in better detection performance. In Figure 11(b), we can achieve $\leq 0.5\%$ false negative SU ratio with a beacon transmission range $r = 500m$ if SU density $\geq 600/(16 \times 16km^2)$. In Figure 11(c), the initial anchor ratio increases to 50%. In order to achieve the same false negative SU ratio, we only need a $400m$ beacon transmission range for a 500-SU network. As the initial anchor ratio keeps increasing, the same false negative SU ratio can be achieved with even lower beacon transmission range for a sparser network. Therefore, we can adjust beacon transmission range according to initial anchor ratio and SU density. For example, it is more efficient to use a low beacon transmission range for urban area (high γ and SU density) and a higher beacon transmission range for suburb or rural area (low γ and SU density).

2) False Alarm: Malicious SUs can cause false alarms in PLV process. These malicious SUs claim themselves as anchor nodes and transmit false location information to fool other SUs. To deal with such a situation, we adopt a majority voting mechanism. An unverified SU will always select the result (verified or spoofed) indicated by most of the surrounding anchor nodes. If there is a tie, the default assumption is that the SU's GPS location is verified.

We first fix the benign anchor ratio $\gamma = 30\%$. Then we vary the ratio of the number of malicious SUs over the number of benign anchor nodes and SU density to examine the performance of majority voting mechanism under different situations. We also define false positive SU ratio as the number of SUs who report false alarms over the total number of SUs who are not anchors. The malicious SUs are uniformly distributed in the $16 \times 16km^2$ area.

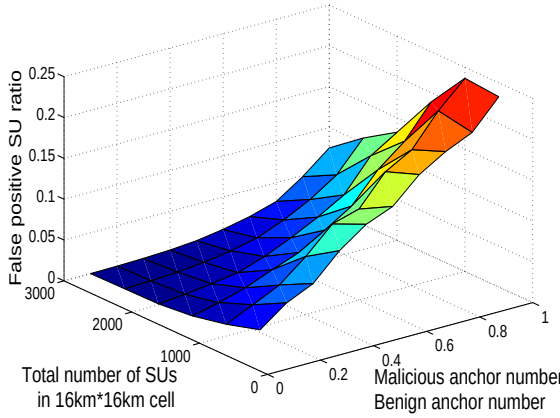


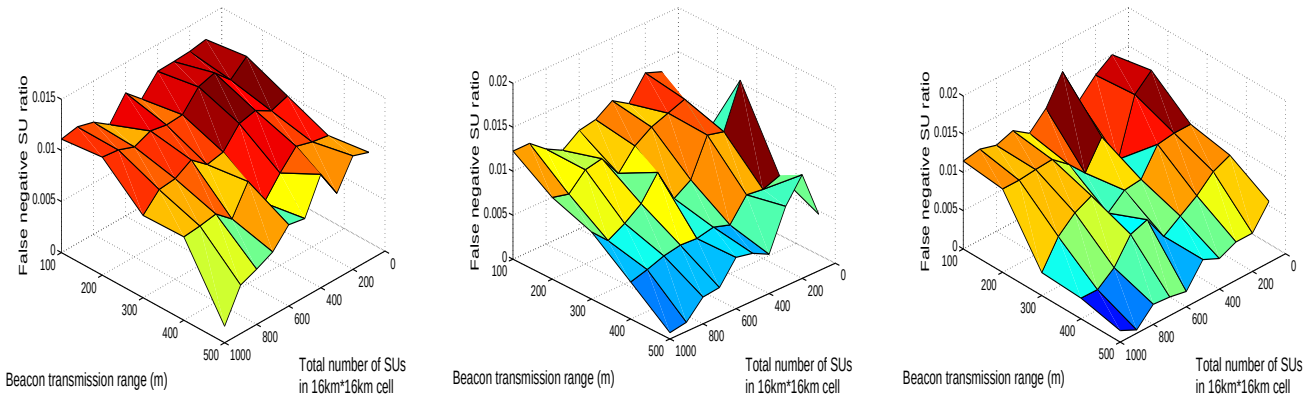
Fig. 12. False positive SU ratio with $\gamma = 30\%$ and beacon transmission range $= 500m$.

As shown in Figure 12, for a fixed malicious SU number, the false positive SU ratio decreases as the network density becomes larger. The reason is that for a sparse network, some unverified SUs can only hear one malicious anchor node and then report false alarms. The majority voting mechanism is invalid in this case. With the SU distribution becomes denser, unverified SUs can hear more anchors, so that the majority voting becomes more reliable. For relatively dense SU situations, if the number of malicious anchor nodes is less than half of the number of benign anchor nodes, the false positive SU ratio is negligible. Our majority voting mechanism only cannot handle the case in which the network is extremely sparse and with a great portion of malicious SUs. However, it is very unlikely that a large portion of SUs are malicious in an extremely sparse network.

3) Discussion: The PLV mechanism requires a certain minimum initial anchor ratio and SU density to propagate, which causes an intrinsic limitation. It may have high rate of missed detection in extremely sparse SU situations with fairly low initial anchor ratio and it may create false alarms in extremely sparse SU networks with high malicious SU ratio. However, we find in our experiments that attackers are less likely to launch GPS spoofing attacks on such a sparse network. The reason is that such an attack can hardly spoof enough SUs to cause PU interference. Hence, we conclude that our detection mechanism works well for relatively dense SU situations which are very likely for such a large cell.

VII. Discussion and Future Work

We now discuss some limitations of our work, as well as future work.



(a) $\gamma = 10\%$ (b) $\gamma = 30\%$ (c) $\gamma = 50\%$

Fig. 11. False negative SU ratio varies with SU density, beacon transmission range and initial anchor ratio γ.

Hybrid Countermeasures: In this paper, we come up with three detection mechanisms and countermeasures for GPS spoofing attacks in database-driven CRNs. However, they all have their limitations. CDS cannot provide backup localization method and may create potential violation to user privacy. ELV requires additional hardware and consumes time to collect fingerprints. PLV demands a minimum initial anchor ratio and SU density to propagate and has missed detection and false alarm cases. Nevertheless, a hybrid scheme can adopt all of them and let these three mechanisms complement each other. For example, we can first use CDS to detect a suspicious spoofing attack timely, then launch ELV and PLV to verify the suspicious victims' GPS locations and provide backup locations if necessary. We plan to demonstrate the effectiveness of hybrid countermeasures by quantitative analysis in our future work.

Limited transmit power of base stations: In database-driven 3.5 GHz CRNs, the transmission range of base stations is limited by their maximum transmit power. The proposed PLV mechanism might not work well if the small cells cannot provide a required minimum beacon transmission range.

VIII. Conclusion

In this paper, we identify GPS spoofing attacks in database-driven CRNs, which can result in interference between SUs and critical PUs and false denial of service to SUs. We also propose various solutions to defend against such attacks and analyze the effectiveness, efficiency and limitations of them. Through extensive experiments and simulations, we demonstrate that the proposed schemes can significantly mitigate the threat of GPS spoofing attacks on database-driven CRNs.

Acknowledgment

We are grateful to National Science Foundation for funding this research through the following grants: CNS-1228903 and CNS-1054697.

References

[1] FCC, "Authorized ex parte contact unlicensed operation in the tv broadcast bands (et docket no. 04-186)."

[2] ——, "Amendment of the commissions rules with regard to commercial operations in the 3550-3650 mhz band (gn docket no. 12-354)."

[3] J. Volpe, "Vulnerability assessment of the transportation infrastructure relying on the global positioning system," 2001.

[4] J. S. Warner and R. G. Johnston, "A simple demonstration that the global positioning system (gps) is vulnerable to spoofing," *Journal of Security Administration*, vol. 25, no. 2, pp. 19–27, 2002.

[5] T. E. Humphreys, B. M. Ledvina, M. L. Psiaki, B. W. OHanlon, and P. M. Kintner Jr, "Assessing the spoofing threat: Development of a portable gps civilian spoofer," in *Proceedings of the ION GNSS International Technical Meeting of the Satellite Division*, 2008.

[6] J. Saarinen, "Students hijack luxury yacht with gps spoofing," Secure Business Intelligence Magazine, July 2013.

[7] N. O. Tippenhauer, C. Pöpper, K. B. Rasmussen, and S. Capkun, "On the requirements for successful gps spoofing attacks," in *Proceedings of the 18th ACM conference on Computer and communications security*. ACM, 2011, pp. 75–86.

[8] M. G. Kuhn, "An asymmetric security mechanism for navigation signals," in *Information Hiding*. Springer, 2005, pp. 239–252.

[9] P. Papadimitratos and A. Jovanovic, "Gnss-based positioning: Attacks and countermeasures," *arXiv preprint arXiv:1001.0025*, 2010.

[10] H. Wen, P. Y.-R. Huang, J. Dyer, A. Archinal, and J. Fagan, "Countermeasures for gps signal spoofing," in *ION GNSS*, 2005, pp. 13–16.

[11] Y. Cui and Y. Wu, "Protocol to access white space database: Security considerations," 2012.

[12] Z. Gao, H. Zhu, Y. Liu, M. Li, and Z. Cao, "Location privacy in database-driven cognitive radio networks: Attacks and countermeasures," in *INFOCOM, 2013 Proceedings IEEE*. IEEE, 2013, pp. 2751–2759.

[13] "http://whitespaces.msresearch.us/wswebgui/whitespaces.aspx."

[14] W. Wang and X. Liu, "List-coloring based channel allocation for open-spectrum wireless networks," in *IEEE Vehicular Technology Conference*, vol. 62, no. 1. Citeseer, 2005, p. 690.

[15] G. Locke and L. E. Strickling, "An assessment of the near-term viability of accommodating wireless broadband systems in the 1675-1710 mhz, 1755-1780 mhz, 3500-3650 mhz, and 4200-4220 mhz, 4380-4400 mhz bands," October 2010.

[16] Y. Zhao, J. Gaeddert, K. K. Bae, and J. H. Reed, "Radio environment map enabled situation-aware cognitive radio learning algorithms," in *Proc. SDR Forum Technical Conference*, 2006.

[17] C. ZIEGLER, "Rosum's alloy chip promises 'precise' location using tv signals," http://www.engadget.com/2010/03/01/rosums-alloy-chip-promises-precise-location-using-tv-signals/, MARCH 1ST.

[18] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, "Epidemic spreading in real networks: An eigenvalue viewpoint," in *Reliable Distributed Systems, 2003. Proceedings. 22nd International Symposium on*. IEEE, 2003, pp. 25–34.

[19] S.-H. Fang, J.-C. Chen, H.-R. Huang, and T.-N. Lin, "Metropolitan-scale location estimation using fm radio with analysis of measurements," in *Wireless Communications and Mobile Computing Conference, 2008. IWCMC'08. International*. IEEE, 2008, pp. 171–176.