Can the Privacy of Primary Networks in Shared Spectrum be Protected?

Matthew Clark*†, Konstantinos Psounis*
*University of Southern California, Los Angeles, CA
{clarkma,kpsounis}@usc.edu
†The Aerospace Corporation, El Segundo, CA

***Abstract*—In an effort to meet growing demands on the radio frequency spectrum, regulators are exploring methods to enable band sharing among a diverse set of user devices. Proposed spectrum access systems would dynamically assign spectrum resources to users, maintaining databases of spectrum use information. While these systems are anticipated to increase the efficiency of spectrum sharing, incumbent users have raised concerns about exposing details of their operations and have questioned whether their privacy can be protected.**

In this paper, we explore whether primary users can retain a critical level of privacy when a system uses their information to enable dynamic access to the spectrum by other users. Under a variety of operational scenarios and user models, we examine adversary techniques to exploit the spectrum access system and obfuscation strategies to protect user privacy. We also develop analytical methods to quantify the performance of both the adversary and obfuscation strategies. To our knowledge, this is the first work that considers the privacy of a primary user in the setting of a highly dynamic spectrum access system. Privacy analysis of this kind will help to enable adoption of shared spectrum access systems by allowing incumbent users to quantify and mitigate risks to their privacy.

I. Introduction

In an attempt to meet the growing demands of the many radio frequency spectrum hungry applications, regulators are exploring the feasibility of sharing spectrum bands among increasingly broad types of systems. For example, in the United States, the Federal Communications Commission recently issued a ruling that the 3550-3700 MHz band will be opened up to new spectrum uses through advanced shared spectrum access systems [1]. The proposed systems are intended to protect a tier of primary users (PUs) from harmful interference while dynamically assigning spectrum resources to lower tier, or secondary user (SU), devices. Assignments are determined from databases of policy and spectrum use information. In prior database based spectrum sharing efforts, such as for television white space [2], the PU systems were well defined and anticipated to change relatively slowly. Assignment policies were determined a priori based on interference analysis of the most severe potential cases. The advanced spectrum access systems here, however, are envisioned to be more dynamic, responsive and generally capable of supporting a diverse set of operational scenarios and heterogeneous networks.

Between the databases of information held by these sharing systems and their dynamic nature, incumbent users have raised concerns about maintaining the privacy of their operations with these systems. Unlike the television white space model where the PUs were primarily public television broadcasts, many of the incumbent systems in 3550-3700 MHz are operated by government entities e.g., Department of Defense radars. The information that a spectrum access system would need to assign spectrum resources, such as locations, frequencies, time of use, and susceptibility to interference, may be considered very sensitive by the incumbents and should be protected from exposure to a potential adversary. However, typical cyber security approaches are not alone sufficient as the normal operation of the access system may allow an adversary to deduce critical aspects of a PU operation. For example, an adversary may legitimately operate a number of cell phones within a secondary network and leverage assignments from the access system to make inferences about the characteristics of a military radar.

In this paper, we explore whether PUs can retain a critical level of privacy when a system uses their information to enable dynamic access to the spectrum by SUs. We consider a variety of operational scenarios consisting of PU models and adversary techniques to exploit the spectrum access system. Our focus is on the inherent privacy exposure associated with the intended operations of the access system and we refer the reader to the literature for works on other cyber security threats [3]. We also examine PU obfuscation strategies to protect privacy without precluding use by SUs. As a key result, we offer a lower bound for the expected time PU privacy can be maintained with a given access system and threat model, where this bound can be tied directly to PU operational requirements. While other works have studied PU privacy under relatively static spectrum access systems (e.g., based on television white space models) [4], [5], [6], to our knowledge, this is the first work that provides privacy guarantees with an access system model that is generally applicable to highly dynamic use cases and heterogeneous users. We find that the dynamic system model substantially affects the analysis as well as considerations for both PU privacy and adversary inference strategies. Privacy analysis of this kind may help to enable wider adoption of shared spectrum access systems by allowing incumbent users to quantify the risk posed to their privacy and by providing specific techniques to mitigate that risk.

The organization of the rest of this paper follows. Related work is reviewed in Section II and a detailed model of the spectrum access system is provided in Section III. We discuss PU privacy in Section IV and formulate a PU privacy problem. The general adversary model is provided in Section V. As a

case study, we focus on analysis and results for the location privacy of stationary PUs in Section VI. Our conclusions and a discussion of future work are offered in Section VII.

II. Related Work

The requirements of the spectrum access system limit the applicability of other works on the privacy of general statistical databases. Differential privacy for example, introduced in [7], provides a measure of privacy and useful results for statistical database design [8], but the competing goals of the spectrum access system, i.e., protecting PUs from harmful interference and offering non-negligible utility to the SUs, prevents it from being differentially private. Other methods, including relaxed versions of differential privacy, present similar problems [9].

Of the many works specific to spectrum access systems, there are relatively few that address privacy issues. Several works consider SU location privacy loss due to participation in collaborative sensing [10], [11], [12], but the resulting problem model cannot be directly applied to PU privacy. In [13] the PU information is assumed to be known by the attacker and used to geo-locate the SU. This is the converse of the PU privacy problem and the difference in modeling precludes direct application of their analysis and mitigation strategies to the PU problem in our setting.

PU privacy with a spectrum access system is considered directly in [6] where the authors assess strategies for the database falsely denying SU resource assignments to protect PU privacy. The analysis is provided in the context of a system with binary decisions where either any given channel is available to an SU at a particular timeslot or it is not. This assumption limits the applicability of the analysis and results to PU time-of-use privacy. While more general PU privacy is beyond the scope of this work, the basic falsification/obfuscation strategy is useful more generally.

The most relevant works addressing the PU privacy problem are [4] and [5]. In both works, several PU privacy preserving obfuscation methods are considered and adversary inference methods are described. The authors assume a spectrum access system that grants assignments based on a pre-established lookup table for relating SU to PU distances and permissible transmit power assignments. This approach is consistent with the television white space model, where the PUs are well defined and conservative assumptions can be made without substantial loss of efficiency. Despite its merit in such settings, we argue this approach is too limiting for more general future systems designed to be efficient for a diverse group of PU and SU devices. In a more heterogeneous system, worst case assumptions, e.g., fixed margins for aggregate interference, will quickly degrade the utility of the spectrum for SU networks. Efficient use of the spectrum requires a more dynamic access system that accounts not just for the point-to-point distance between a single PU and SU, but for the overall topology of both the PU and SU systems, along with their operating characteristics. In this paper, we formulate a model for a more dynamic spectrum access system and find that this model significantly affects the analysis and resulting strategies in the PU privacy problem.

III. The Spectrum Access System

The spectrum access system in our model operates on a database of policies and real-time information provided by user devices. In the 3.5 GHz band, there are three tiers of users with the higher tiers having priority over the lower. We will only include two tiers in our model however, PUs and SUs, recognizing that loss of privacy is greatest between the first two tiers. Further, results based on two tiers may be applicable more generally since, e.g., a third tier would not be able to differentiate between the first and second tier users, and would treat them together as a single group.

An SU requests an assignment from the access system by sending information on its present location and useful parameter ranges via a connection that does not rely on spectrum from the access system. Specifically, for SU devices indexed from 1 to n_s, the ith SU will provide the range of transmission powers it can profitably use as elements of the vectors $P^{min} = [P_i^{min}] \in \mathbb{R}_+^{n_s}$ and $P^{max} = [P_i^{max}] \in \mathbb{R}_+^{n_s}$ that will be kept by the access system. We use uppercase to denote vectors and their elements, and subscripts to index the elements. The SU will also provide a set of frequency tuning ranges as well as a range of useful bandwidths, with each reflecting the capabilities of the device and the requirements of the application. To simplify the notation, we will treat the SU location as a scalar index for a cell in a discretized region, although in practice, it may be given as a vector of coordinates. The access system stores all SU locations in a set we denote $\mathcal{X}_s$. To allow the access system to take advantage of frequency dependent scheduling, an SU might also send channel state information for the links in the SU network. For a given frequency channel, the access system will store all SU reported channel gains in a vector $G^s = [G_i^s]$, where i is the index for the ith SU.

To enable the access system to protect them from harmful interference, PUs will send their location information, stored by the system in a set denoted by $\mathcal{X}_p$. The access system will assume a propagation model with uncertainty, i.e., $\mathbf{G}^p = [\mathbf{G}_{ij}^p] \in \mathbb{R}^{n_s \times n_p}$ is the matrix for the channel gains between each PU and SU with rows and columns indexed by i and j for the SUs and PUs respectively. The bold face is used to specify that the elements are random variables. The PUs must also report their interference criteria, given by vectors $I^{th} \in \mathbb{R}_+^{n_p}$ and $\Lambda^{th} \in \mathbb{R}_+^{n_p}$. I_j^{th} is the harmful received interference power threshold for the jth PU where PUs are ordered arbitrarily. $0 < \Lambda_j^{th} < 1$ is a reliability requirement for the PU, i.e., the maximum probability that the threshold given by I_j^{th} can be exceeded. This reliability parameter accounts for inherent uncertainty in the resulting interference due to the random $\mathbf{G}^p$.

The access system uses discrete time slots, where the duration of these slots are chosen as a trade between efficiency and complexity of the system. Similarly, the access system will use discrete frequency channels and discrete power levels. For each upcoming time slot, the access system will need to solve a scheduling problem to determine how to assign frequency channels and maximum transmit power levels to SUs. Consider an individual frequency channel. The access system should de-

termine the operating parameters of the SUs which maximize some utility function subject to constraints protecting the PUs from harmful interference. Deferring discussion of a specific implementation for the moment, a frequency channel power assignment function $f()$ operating on n_s SUs and n_p PUs should return a vector $P = [P_i] \in \mathbb{R}_+^{n_s}$ of maximum transmit power assignments for the SUs as follows:

$$\begin{aligned} P =& f(\mathcal{X}_p, \mathcal{X}_s, G^s, I^{th}, \Lambda^{th}, P^{min}, P^{max}) \\ \approx& \arg\max_{P'} \{U(P', G^s, P^{min}, P^{max}) \\ & ; Pr((P')^t \mathbf{G^P} \geq I^{th}) \leq \Lambda^{th}, P' \in \mathbb{R}_+^{n_s}\}, \end{aligned} \tag{1}$$

where $U()$ is some utility function to be defined for the SUs. Note that a power assignment of zero is possible and corresponds to excluding a particular SU from this frequency channel. In this way, $f()$ acts as both an admission control and power assignment function. The basic exchange of information between the PUs, SUs and the access system for a single frequency channel and timeslot is diagrammed in Figure 1.

We will offer a general methodology, but when a specific form for $f()$ is called for in the following results, we make use of the algorithm in [14]. There are many prior works in the literature that develop power assignment algorithms, see, for example, [15], [16], [17]. We opt to use [14] because it accounts for the network topology of the SUs, the aggregate interference due to multiple SUs interfering with a single PU, the need for the algorithm to be low in complexity, and the uncertainty in estimating the interference that will result for a particular scheduling decision, all of which are critical for a dynamic access system. The basic method in [14] is to determine with a function $I()$, the maximum mean equal interference power level that can be caused by each of the SUs and still satisfy the interference constraints. $I()$ is specific to the uncertainty model assumed and is negatively correlated with the number of SUs that will receive nonzero power assignments, which we denote by l, $0 \leq l \leq n_s$, and where l is selected to maximize $U()$. The power assignment P_i for the ith SU is given by

$$P_i = \frac{I(l, I^{th}, \Lambda^{th})}{\max_j \bar{G}_{ij}^p}, \tag{2}$$

where the denominator deviates from the form in [14] which was provided for the case of a single PU. The system will use the locations in $\mathcal{X}_p$ and $\mathcal{X}_s$ to compute the mean SU-to-PU channel gain, $\bar{G}_{ij}^p$, i.e., the mean of $\mathbf{G}_{ij}^p$. The maximization corresponds to application of the algorithm for each PU in a clustered fashion, i.e., for PU with location $x_j \in \mathcal{X}_p$ we select all SUs that are closer to x_j than any other element in $\mathcal{X}_p$ and apply the assignment algorithm to those. l therefore depends on the clustering, and we define a vector $L = [L_i] \in \mathbb{R}_+^{n_s}$ to denote, for each SU, the number of SUs that receive nonzero power assignments in the same PU cluster. We also include a small margin in the output of $I()$ to account for interference sources other than those in the specific PU cluster and we assume that, for practical implementation, each P_i is rounded to the nearest 1 dB increment.

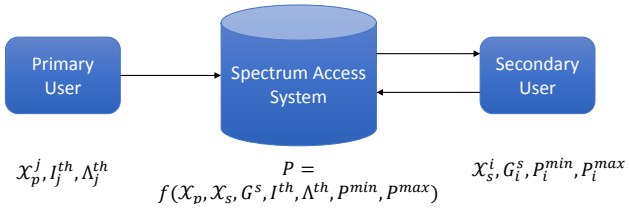


Fig. 1. Spectrum Access System data exchange model.

The utility $U()$ may be left general, but in the remainder of this work, when a specific form for $U()$ is required, we use the sum-rate of the SUs as our metric, i.e., $U(P, G^s, P^{min}, P^{max}) = \sum_i^{n_s} U_i(P_i, G_i^s, P_i^{min}, P_i^{max})$, where for $P_i \geq P_i^{min}$,

$$U_i(P_i, G_i^s, P_i^{min}, P_i^{max}) = \log_2\left(1 + \frac{\min(P_i, P_i^{max})G_i^s}{\eta}\right) \tag{3}$$

and where η is the thermal noise in the SU receiver. For $P_i < P_i^{min}$, the utility is assumed to be zero.

IV. Primary Privacy

A. Metrics

There are a variety of privacy metrics in the literature with varying levels of applicability to the PU privacy problem, e.g., [18], but the obvious approach to measuring privacy is to quantify how well an adversary can estimate the actual parameters of the PU as in [19] for the case of location privacy. We assume the adversary models the unknown PU parameters as random variables and estimates a probability distribution for each based on observations from the access system. For example, if the adversary were estimating the harmful interference threshold of a PU, I_j^{th}, with every observation it would update a probability mass function $p(\hat{I}_j^{th})$, where we use the $\hat{}$ to indicate an adversary estimated value. The privacy can be measured directly from this adversary estimated distribution. We can also compute an adversary guess from the distribution, e.g., with the maximum likelihood estimate or by applying some threshold, and measure the error between the guess and the actual value directly.

Given a large number of observations to estimate a static parameter, eventually the adversary may reduce the measure of privacy to an arbitrarily low level. As a PU then, the question is whether a threshold level of privacy can be maintained for at least as long as a particular operating parameter remains static. Let τ be a random variable for the number of time slots before the privacy is reduced below a threshold level, where adversaries get new observations at each time slot. We will seek here to measure $E\{\tau\}$, i.e., the expected duration that privacy can be maintained.

B. Obfuscation Strategies

We focus in this paper on the following two obfuscation strategies and leave other strategies, e.g. modifying the oper-

ational behavior of a PU, for future work.

Obfuscation Strategy #1: Inserting false PU entries into the database. In the context of location privacy, the jth PU with actual location x_j would generate a number of otherwise indistinguishible false locations and provide $\mathcal{X}_j$ to the access system such that $x_j \in \mathcal{X}_j$. The access system protects all PU locations identified and does not need to know which entries correspond to actual PU locations. We anticipate a tradeoff here where increasing the number of false locations should increase privacy but decrease the utility of the spectrum for the SUs. False entries should be inserted randomly to prevent an adversary from learning strategies over time, and should not be generated in a way that they could be excluded by an adversary based on a priori information. An adversary looking for a car, for example, may immediately exclude any false locations that do not coincide with roads. The false entry strategy is also applicable to time and frequency use of the PUs, where additional time slots and frequency channels would be protected by the access system.

Obfuscation Strategy #2: Parameter randomization. This strategy entails adding noise to the actual values of the PU parameters in the database. To ensure the PUs are protected from harmful interference, i.e., that the access system can guarantee $Pr\{P^t\mathbf{G}^p \geq I^{th}\} \leq \Lambda^{th}$, we would not randomize location, time, or frequency use entries. Randomization can be applied to the power assignment P and the thresholds I^{th} and Λ^{th}. We only allow the parameters to be reduced by the randomization to guarantee protection from harmful interference and recognize that uniformly distributed results will maximize adversary uncertainty.

We will specifically generate a random power assignment vector $\mathbf{\Psi} = [\mathbf{\Psi}_i] \in \mathbb{R}_+^{n_s}$ for the SUs according to $\mathbf{\Psi}_i = P_i/\mathbf{Q}_i$, where $\mathbf{Q} \in \mathbb{R}_+^{n_s}$ is a vector of random reductions with entries that satisfy $(\boldsymbol{Q}_i)_{dB} \leq b_P$, i.e., b_P bounds the randomization, and where $()_{dB}$ translates an argument to the dB domain. We assume a realization $\mathbf{\Psi_i} = \Psi_i < P_i^{min}$ would yield a power assignment of zero, equivalent to denying the SU access to the channel. To simplify the notation, let $S = [S_i] \in \mathbb{Z}_+^{n_s}$ be a vector whose elements are the number of possible realizations for the corresponding elements in $\mathbf{\Psi}$. For example, $S_i = b_P + 1$ if $(P_i)_{dB} - b_P \geq (P^{min})_{dB}$ and $S_i = 1$ if $P_i < P^{min}$. We can generate a uniformly distributed $\mathbf{\Psi}$ by drawing each $(\mathbf{Q_i})_{dB}$ from a uniform distribution on $\{0, ..., S_i - 1\}$. Similar bounds b_I and b_Λ can also be used for randomization of the interference parameters. Increasing the randomization bounds should increase the privacy, but as with inserting false entries, will come with a cost to the SU utility.

We consider that the PU privacy strategy may consist of a combination of inserting false entries and randomization. The PU will need to determine how much randomization to use and how many false entries to insert. The PU should also account for the impact on the SU utility for each strategy. We formulate this decision as the PU privacy strategy problem,

$$\begin{aligned} \underset{|\mathcal{X}|,|\mathcal{F}|,|\mathcal{T}|,b_P,b_I,b_\Lambda}{\text{maximize}} \quad & \mathbf{E}\{\textstyle\sum_{t=1}^{T} U(t)\} \\ \text{subject to} \quad & \mathbf{E}\{h(|\mathcal{X}|,|\mathcal{F}|,|\mathcal{T}|,b_P,b_I,b_\Lambda)\} \geq h_t, \end{aligned} \tag{4}$$

where $\mathcal{F}$ and $\mathcal{T}$ are the sets of frequency channel and time slot use by the PU, including actual and false entries. h_t is some threshold level of privacy. We leave $h()$ as a placeholder for the analyst's selected privacy metric, though in subsequent sections we will use the time metric discussed above, i.e., $E\{\tau\}$. A solution to problem (4) not only allows an existing PU to determine how to provide its information to the access system, but it also allows potential PUs to assess whether it is acceptable to share their spectrum in the first place.

V. The Adversary Model

The adversary model treats unknown PU parameters as random variables and creates estimates of their probability distributions based upon observations from the spectrum access system and any information known a priori. We will treat these random variables as discrete valued and assume the adversary will tolerate some quantization error. In estimating locations, for example, the adversary will divide an arbitrary region up into cells and create an estimate for the probability a PU is contained within each of the cells.

To facilitate specific analysis, for the remainder of this work we will focus on PU location privacy. With conservative estimation of PU privacy in mind, we will assume the adversary knows the PU interference criteria given by I^{th} and Λ^{th} a priori, e.g., the adversary is looking for a particular PU system with known characteristics. We now consider different scenarios for the observations available to the adversary.

Adversary Case #1: Compromised spectrum access system. In the worst case, we consider a scenario where the adversary has direct access to the information provided to the access system by the PU. This could result if the adversary is able to eavesdrop on the communications between the PU and the access system. In this case, the adversary observes $\mathcal{X}_p$ and the PU privacy is only protected by any false entries that were included. Assuming that the false entries are indistinguishable from the actuals, if we suppose there is one actual location for the jth PU, then the adversary's estimate for a candidate location $\hat{x}_j \in \mathcal{X}_p$ is $p_{\boldsymbol{x_j}}(\hat{x}_j) = |\mathcal{X}_p|^{-1}$. This can be thought of as the minimum PU privacy for all scenarios.[1]

Adversary Case #2: Compromised SU networks. In this case, the adversary has either hacked the SU networks, or has found a way to eavesdrop, and can observe communication between the access system and all SUs. At every time slot, the adversary will get the current $\mathcal{X}_s$ and a vector Ψ, which denotes realizations of the randomized power assignment. The adversary estimate at time slot T for a current candidate set of PU locations $\hat{\mathcal{X}}_p$, given the sequence of power assignment observations $\{\Psi_p^t; t = 1, ..., T\}$ is

$$p_{\boldsymbol{\mathcal{X}_p}}(\hat{\mathcal{X}}_p|\Psi^1, ..., \Psi^T) = \frac{p_{\mathbf{\Psi^1},...,\mathbf{\Psi^T}}(\Psi^1, ..., \Psi^t|\hat{\mathcal{X}}_p) p_{\boldsymbol{\mathcal{X}_p}}(\hat{\mathcal{X}}_p)}{p_{\mathbf{\Psi^1},...,\mathbf{\Psi^T}}(\Psi^1, ..., \Psi^T)}, \tag{5}$$

where $p_{\mathbf{\Psi^1},...,\mathbf{\Psi^T}}(\Psi^1, ..., \Psi^T|\hat{\mathcal{X}}_p)$ is the probability of the observed power assignment given the candidate PU loca-

[1]The adversary may reduce the privacy by other mechanisms of course, e.g., with a sensing capability, but that is beyond the scope of this paper.

tions, and the equality is a direct application of Bayes' theorem. Because the non-randomized powers P are deterministic functions of $\mathcal{X}_p$, and the randomization is assumed to be i.i.d., there is no memory in the system such that $p_{\mathbf{\Psi^1},...,\mathbf{\Psi^T}}(\Psi^1,...,\Psi^T|\hat{\mathcal{X}}_p) = \prod_{t=1}^{T} p_{\mathbf{\Psi}}(\Psi^t|\hat{\mathcal{X}}_p)$. Given $\hat{\mathcal{X}}_p$ and our assumption that the adversary knows $\mathcal{X}_s$, we can directly compute the non-randomized assignment $\hat{P}$ from (1) for each time slot, and, therefore

$$p_{\mathbf{\Psi}}(\Psi^t|\hat{\mathcal{X}}_p) = p_{\mathbf{\Psi}}(\Psi^t|\hat{\mathcal{X}}_p, \hat{P}) = \begin{cases} \prod_{i=1}^{n_s} S_i^{-1}; & (\Psi^t)_{dB} \in \left\{(\hat{P})_{dB} - S + 1, ..., (\hat{P})_{dB}\right\} \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

where we have conservatively assumed that the adversary has learned the general randomization strategy, possibly gained by long-term observation of the system. If the adversary assumes all candidates are equally likely a priori, i.e., $p_{\boldsymbol{\mathcal{X}_p}}(\hat{\mathcal{X}}_p)$ is constant, then it follows from (5) that

$$p_{\boldsymbol{\mathcal{X}_p}}(\hat{\mathcal{X}}_p|\Psi^1,...,\Psi^T) = \frac{\prod_{t=1}^{T} p_{\mathbf{\Psi}}(\Psi^t|\hat{\mathcal{X}}_p)}{\sum_{\boldsymbol{\mathcal{X}_p}=\mathcal{X}} \prod_{t=1}^{T} p_{\mathbf{\Psi}}(\Psi^t|\mathcal{X})}, \tag{7}$$

where $p_{\mathbf{\Psi}}(\Psi^t|\hat{\mathcal{X}}_p)$ is given in (6). The probability estimate for whether a PU is present at any single location can be found by summing (7) over all possible candidate location sets $\hat{\mathcal{X}}_p$ that contain the location of interest. Note that if the adversary divides the region into w discrete cells, testing 2^w candidates quickly becomes intractable. The adversary could also use (7) to identify the maximum likelihood estimate, but we will not necessarily have a closed form to facilitate finding the maximum and the general optimization would be non-trivial. Despite these questions about tractability, (7) will be useful as an ideal adversary estimate.

Adversary Case #3: Compromised SU devices. Another form of adversary may only have access to the assignments for a subset of SUs in the secondary networks. This could be the case if the adversary owns and operates its own SU systems, e.g. its own cellphones registered on a cellular SU. The adversary is challenged with this case in that the power assignment P cannot be treated as a deterministic function of $\mathcal{X}_p$ if $\mathcal{X}_s$ is not entirely known. Since P, and thus the adversary observations Ψ, is dependent on the entire topology of PUs and SUs, the adversary problem in the spectrum access system setting is potentially harder than in other location privacy problem settings. We could treat $\mathcal{X}_s$ as a set of random variables and follow a similar approach to what was used to derive (7), but considering the potential intractableness of that result, we do not pursue such an approach here. Instead we will offer efficient scenario-specific heuristics.

VI. Case Study: Protection of Stationary PU Location Privacy

In this section we consider the location privacy of PU locations that remain fixed for many time slots. A PU system that dwells in one location will want to ensure the observations made by an adversary over the duration of the PU operation will be insufficient to reduce the PU privacy below some threshold.

A. Adversary Estimation Schemes

For this case study, (7) can be applied in theory if the adversary can observe all SU assignments and has a priori knowledge about the PU interference thresholds (I^{th} and Λ^{th}) and the randomization strategy. Because (7) is potentially intractable we consider heuristic approaches. At each observation, using I^{th} and Λ^{th}, the adversary can identify locations where a PU could not possibly exist and still be protected from harmful interference. The first and simplest approach then is to start with a set of all possible PU locations and eliminate those that are found to receive harmful interference at one or more time slots. Once enough locations are eliminated it becomes feasible to compute (7) directly. We will refer to this approach of eliminating locations in the first phase and then computing likelihood estimates directly in a second phase as the *composite method*.

For candidate locations that have not been eliminated, we argue that those locations which have come closest to receiving harmful interference are more likely to be the source of limitation to the assigned SU transmit powers. Thus, we suppose they are more likely to be protected PU locations. With this in mind we calculate the harmful interference constraint margin for each location and use the inverse as a weight, where the normalized weights can be used as an approximation of the likelihood. We'll refer to this approach as the *interference constraint margin (ICM)*.

We also recognize that a power assignment resulting from a set of multiple PU locations should have some similarity to an assignment that was based on a single location within that set due to the clustered approach to implementing the power assignment algorithm. We therefore consider an approach where we compute the maximum power assignment $\hat{P}$ that would result for each location as if it were the only PU protected. By comparing the squared error between $\hat{P}$ and the observed power assignment Ψ, we have a measure of the similarity and consider the normalized inverses as another basis for an estimated likelihood. We'll refer to this approach as *single PU error (SPE)*. More detail and performance assessments are provided for these approaches in VI-C.

With our heuristics and direct likelihood estimates, we will also apply a thresholding approach to identify an adversary guess for the actual PU locations. Specifically we define an x-σ approach where we take the median likelihood of all locations and include in the guess all locations that have likelihoods at least x standard deviations above the median. With the resulting adversary guess, we will directly compute the sum distance error between the guess and the actual PU locations as a measure of adversary performance and PU privacy.

In addition to the distance error, we will observe in typical cases that many locations can be eliminated quickly relative to the convergence rate of optimal inference attacks on the remaining locations. We use the number of candidate locations

which have not been eliminated yet as another simple measure of progression.

B. PU Privacy Time and Strategy

We now focus on the PU strategy problem and seek an estimate for the privacy time, measuring privacy here in terms of an optimal adversary's ability to eliminate locations. Suppose $\boldsymbol{\mathcal{X}}_s$ and $\boldsymbol{\mathcal{X}}_p$ are treated as random variables with arbitrary distributions, and assume a perfectly symmetric region, i.e., a torus. Randomly pick a location in the region and call it x_k. An adversary observing the random power assignments will have some probability of determining that $x_k \notin \mathcal{X}_p$, i.e., that x_k is not a PU location. Denote this probability of eliminating x_k as p_k. If we consider multiple trials with i.i.d. $\mathcal{X}_s$ and $\mathcal{X}_p$, then eliminating any location at a particular trial can be treated as a geometric random variable with parameter p_k and cumulative distribution function $F_k(t) = 1-(1-p_k)^t$. By the symmetry of the region, we can consider eliminating multiple locations as i.i.d. and if there are v locations in the region that do not contain PUs, then the expected time to eliminate w of the v locations can be shown to be

$$\begin{aligned} E\{\boldsymbol{\tau_w}\} &= \sum_{t=1}^{\infty} t \binom{v}{w} [(F_k(t))^w - (F_k(t-1))^w] \\ &= \binom{v}{w} \sum_{i=1}^{w} \frac{\binom{w}{i}(-1)^{i+1}}{1-(1-p_k)^i}, \end{aligned} \tag{8}$$

where the second equality follows from some algebraic manipulations and use of a series identity. To compute (8), we must first compute p_k, which, in turn requires the computation of $p_k|\mathcal{X}_s, \mathcal{X}_p$, since $p_k = \sum p_k|\mathcal{X}_s, \mathcal{X}_p \cdot p_{\mathcal{X}_s,\mathcal{X}_p}$, that is, we simply treat p_k as the marginal for the probability conditioned on the realizations of $\boldsymbol{\mathcal{X}}_s$ and $\boldsymbol{\mathcal{X}}_p$. For a single realization, we can eliminate x_k from consideration if the power assignment provided to the ith SU, Ψ_i, is greater than the maximum that could result if the candidate were actually a PU location. Note that if $\hat{\mathcal{X}}_p$ is a particular candidate set of PU locations with $x_k \in \hat{\mathcal{X}}_p$ and $\hat{P}_i$ is the maximum power assignment for the ith SU under $\hat{\mathcal{X}}_p$, for x_k to be eliminated we need $\Psi_i > \max_{\{\hat{\mathcal{X}}_p : x_k \in \hat{\mathcal{X}}_p\}} \hat{P}_i$.

Let $B = [B_i] \in \mathbb{Z}_+^{n_s}$ be the vector whose elements are the number of realizations of Ψ_i that will allow the adversary to eliminate x_k. The probability of eliminating x_k is then straightforward, given the uniform distribution assumed for Ψ_i. However, computing each B_i may be complicated depending on the power assignment algorithm. Intuitively, B_i should get larger as an SU moves closer to x_k. With the equal interference power allocation algorithm that we utilize, B_i depends on the specific topology, i.e., $\mathcal{X}_p$ and $\mathcal{X}_s$, as this determines the clustering of the SUs with each PU and the resulting power assignments. Let $D = [D_i] \in \mathbb{R}_+^{n_s}$ denote a vector whose elements are the euclidean distances between the SUs and the nearest PUs. Then, with this power assignment method (see (2)), assuming that the mean channel gain is inversely proportional to the distance raised to a path loss exponent n, and recalling the condition $\Psi_i > \max_{\{\hat{\mathcal{X}}_p : x_k \in \hat{\mathcal{X}}_p\}} \hat{P}_i$, the condition on whether a location can be eliminated based on a particular power assignment realization for the ith SU can be shown to be

$$\left(Q_i \max_{\{\hat{\mathcal{X}}_p : x_k \in \hat{\mathcal{X}}_p\}} \gamma \right)^{1/n} \hat{D}_i < D_i \tag{9}$$

where Q_i is the realization of the random power assignment reduction described in Section IV-B and γ is defined as the ratio of the equal interference levels returned by $I()$ for $\hat{\mathcal{X}}_p$ and actual PU locations $\mathcal{X}_p$, i.e.,

$$\gamma = \frac{I(\hat{L}_i, I^{th}, \Lambda^{th})}{I(L_i, I^{th}, \Lambda^{th})}. \tag{10}$$

Determining the probability of eliminating a location x_k is complicated in the dynamic access system setting because it requires eliminating all candidate PU configurations for which $x_k \in \hat{\mathcal{X}}_p$, which may be a very large set. However, we can lower bound γ over all candidates such that we lower bound the PU privacy estimate. Specifically, since the locations in $\mathcal{X}_p$ will always be viable candidate locations for the adversary, eliminating each $\hat{\mathcal{X}}_p$ will require, at a minimum, eliminating all candidates that are subsets from $\{x_k\} \cup \mathcal{X}_p$. With the assumption that the power allocation algorithm will group individual SUs with the nearest PU, the adversary can only hope to eliminate a location x_k with those SUs that are closer to x_k than any location in $\mathcal{X}_p$. Let $\mathcal{V}_k$ denote the set of SUs that satisfy this condition. If we set $\hat{L}_i = |\mathcal{V}_k|$ in (10), we can compute a bounding γ for (9), and can then compute each B_i by counting the number of realizations Q_i that satisfy (9).

By estimating p_k, we have a method, with (8), to compute a lower bound for the expected time until an adversary can eliminate w locations from consideration. This is a useful metric for the PU optimization Problem (4). Treating $\boldsymbol{\mathcal{X}}_p$ as i.i.d. is applicable to the case of a PU interested in the suitability of an access system to protect its typical operations. We can use the same approach with a fixed $\mathcal{X}_p$ corresponding to a specific PU operation, though the notation must be modified to account for the loss of symmetry in the topology. In either scenario, the estimation of p_k and the numerical computation of the expected sum-utility still remain as complications for solving the PU problem explicitly. We consider a simple greedy approach in Algorithm 1 to find a reasonable PU privacy strategy for a given $\mathcal{X}_p$. In words, this algorithm starts with no obfuscation and at each iteration increases either the randomization parameterized by b or adds a false location, depending on which provides the greatest privacy gain using (8). These iterations continue until the expected privacy time satisfies an input privacy time threshold h_t.

C. Simulation Results

To examine the contributions in this paper via simulation, we deploy SUs in a 20 km by 20 km rectangular region and consider a 10 MHz channel with center frequency 3645 MHz. We lay down a grid of SU access points (e.g., cellular base stations) with a 2.5 km inter-site spacing and randomly place SU devices and connect them with the nearest access point. Since we are concerned with adversary estimates based on

Algorithm 1 Greedy Algorithm for the PU Privacy Strategy

Given $\mathcal{X}_p, n_s, n_a, w, h_t$
$\mathcal{X} \leftarrow \mathcal{X}_p$
$b \leftarrow 0$
$b' \leftarrow b + 1$
$\mathcal{X}' \leftarrow \boldsymbol{x_k} \notin \mathcal{X}_p$
$E_b \leftarrow E\{\tau_w(\mathcal{X}, b', n_s, n_a)\}$
$E_x \leftarrow E\{\tau_w(\mathcal{X} \cup \mathcal{X}', b, n_s, n_a)\}$
if $E_x \geq h_t$ **then**
  $\mathcal{X} \leftarrow \mathcal{X} \cup \mathcal{X}'$
  return $\mathcal{X}, b$
else if $E_b \geq h_t$ **then**
  $b \leftarrow b'$
  return $\mathcal{X}, b$
else if $E_x \geq E_b$ **then**
  $\mathcal{X} \leftarrow \mathcal{X} \cup \mathcal{X}'$
else
  $b \leftarrow b'$
end if
go to 4

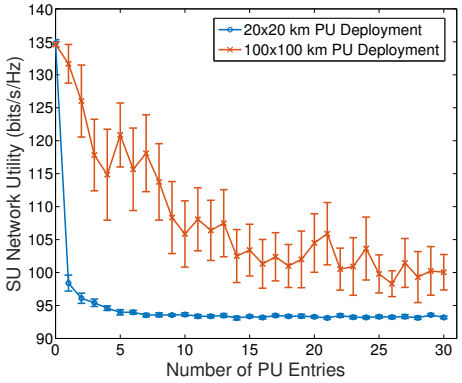


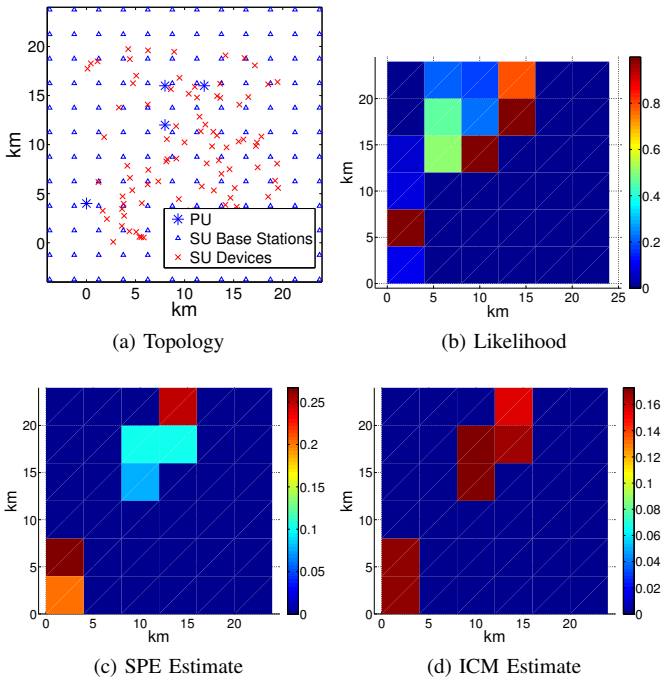
Fig. 2. Sample average SU sum-rate utility versus number of protect primary locations

power assignments granted to mobile devices, we assume the channel use is limited to SU device uplink transmissions to the access points. Lacking specific parameters for the systems that operate in 3.5 GHz, we use the parameters in [20] for both the PU and SU systems given that this is at least a real-world case where a spectrum access system could be applied. With this approach, the SU devices are representative of LTE handsets with transmission powers in the range $P^{max} = 23$ dBm and $P^{min} = -40$ dBm. The access points act as common LTE base stations and we run a proportional fair scheduler with parameters $P_0 = -90$ dBm and path loss correction coefficient $\alpha = 0.8$ in estimating the sum-rate of the SUs. The access system grants SU assignments for 30 second time slots. The PU locations are protected with $I^{th} = -121$ dBm and $\Lambda^{th} = 0.5\%$. All systems are assumed to have omnidirectional antenna gains for simplicity. For the propagation model between all transmitters and receivers, we use a two-ray model ($n = 4$ after a breakpoint) with PU heights of 15 m and SU heights of 4 m. For the distribution of $\boldsymbol{G}^p$, the channel gain of the SU to PU paths, we use a log-normal distribution with standard deviation $\sigma = 10$ dB. Finally, for the adversary's a priori information, we assume all candidate solutions are equally likely, resulting in a 50% probability of a PU in any specific location.

Adversary Case #1: We first consider the minimum PU privacy level, important for any case where the adversary knows $\mathcal{X}_p$ exactly. We consider two deployments for the PUs. In the first, the PUs are uniformly placed in the same 20 km by 20 km region as the SUs. In the second, the PUs are uniformly placed in a 100 km by 100 km region and the SUs are uniformly placed in a 20 km by 20 km corner of the region. The latter deployment is obviously more favorable for SU utility, and it is also more representative of real-world sharing scenarios. For example, in 3.5 GHz, we would expect to find large concentrations of SUs in population centers, whereas many military operations would be more likely to take place in somewhat remote areas. Since the PU privacy is only protected by the number of locations in $\mathcal{X}_p$ we provide in Figure 2, for both PU deployments, results for the sample average utility of 1,000 SU devices versus the size of $\mathcal{X}_p$ where the average is over 40 randomly generated topologies of SU and PU locations. We see from this figure a direct way to trade SU utility for PU minimum privacy, though the SU utility is, not suprisingly, much more limited when the PUs and SUs are concentrated in the same region.

Adversary Case #2: The adversary no longer has direct access to the protected PU locations, but does have access to all SU assignments. We use a random direction mobility model for movement of 100 SU devices with a constant speed of 30 km per hour. The adversary divides the region into 4 km by 4 km cells, resulting in a small enough number of candidate solutions that likelihood estimates may be computed directly from (7). There are four PU locations ($n_p = 4$) and randomization is set with $b_P = 30$ dB. Figure 3 provides a visualization of the adversary probability estimate in the small scale scenario for each location with the given topology and based on observation of a single time slot. The actual likelihood estimate and the heuristics introduced in Section VI-A, SPE, and ICM each identify the actual PU locations. SPE actually assigns more relative weight to two locations that do not contain PUs in reality than it assigns to three of the actual PU locations. ICM identifies the same number of candidate locations as SPE but assigns similar weights to the actual PU locations as it assigns to other locations. The direct likelihood computation pinpoints three of the PU locations with nearly 100% likelihood but is not able to distinguish between the fourth and a few other locations.

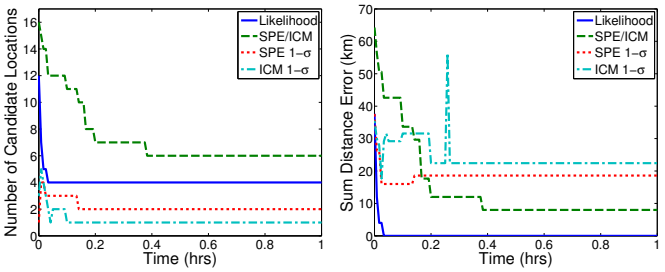
Figure 4a shows the progression of the number of remaining candidate locations over time for the three approaches along with SPE $1 - \sigma$ and ICM $1 - \sigma$, i.e., the heuristics after σ-thresholding is applied. Note that SPE and ICM without any thresholding are equivalent under this metric. While the thresholding approaches do reduce the number of candidates, we cannot guarantee that they do not eliminate an actual PU location. This clearly occurs in this example where the candidate locations fall below the number of actual PU locations. In Figure 4b, the progression of the sum-error measure described in Section VI-A is provided for the approaches. We see that SPE and ICM are outperformed by direct computation of the



(a) Topology (b) Likelihood

(c) SPE Estimate (d) ICM Estimate

Fig. 3. Topology and Estimates for Fixed PU - Low Resolution Adversary.

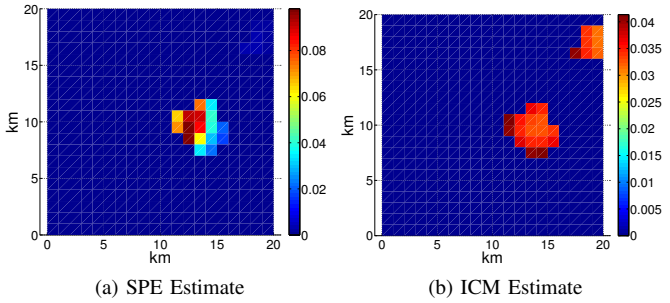


(a) Candidate Locations (b) Sum Distance Error

Fig. 4. Adversary Estimates Over Time - Low Resolution.

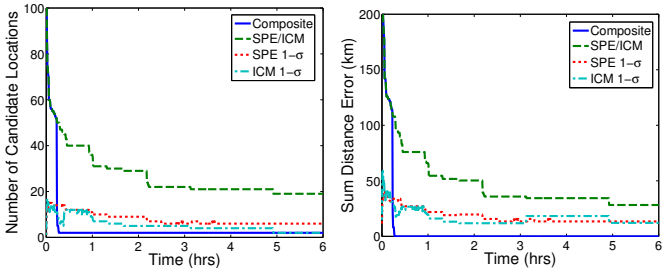
likelihood, where the direct computation converged on the correct PU location in a matter of minutes. We also see further evidence of the potential for the thresholding to introduce errors by eliminating actual PU locations.

We also consider a scenario with an adversary resolution of 1 km by 1 km cells, $n_p = 2$, $b_P = 10$ dB, and all other parameters the same as with the prior lower resolution scenario. The coordinates for the PUs are (13 km, 10 km) and (18 km, 17 km). With the finer adversary resolution, direct computation of the likelihood is impractical and we instead provide results for the composite approach of eliminating locations until few enough remain to use direct computation. In this case we set the threshold to 50 locations or fewer. Figures 5 and 6 provide analogous plots as those provided for the low resolution scenario. SPE identifies the location at (13,10) well, but almost misses the location at (18,17) entirely. ICM identifies the general area of both locations but does not do well at distinguishing within this area. The composite estimate plot is omitted for brevity. Once the number of locations



(a) SPE Estimate (b) ICM Estimate

Fig. 5. Topology and Estimates for Fixed PU - Higher Resolution Adversary.



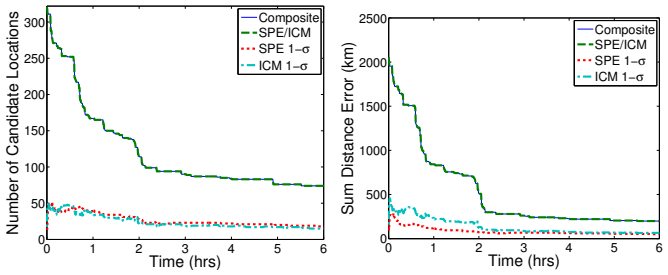
(a) Candidate Locations (b) Sum Distance Error

Fig. 6. Adversary Estimates Over Time - Higher Resolution.

are reduced to 50, the direct computation of the likelihood identifies the PU locations within a couple time slots, with the total process completing in minutes. The ICM and SPE methods are much slower, failing to identify the exact PU locations within the simulated 6 hours.

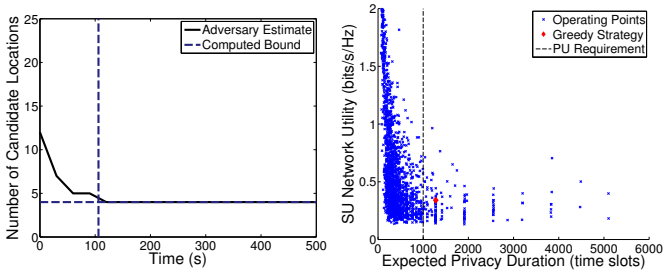
Adversary Case #3: In Figure 7 we provide the progression for the adversary estimate with the topology and higher resolution scenario considered in the prior example, but here assume the adversary only has access to five of the 100 SUs. We see a substantial degradation in the quality of the adversary estimate (note the scale of the y-axis). The total number of candidate locations is not reduced below 50 in the composite method within the simulated 6 hours, and direct computation of the likelihood is not conducted. We refrain from applying direct computation to the locations identified by the $1 - \sigma$ threshold techniques, since these are prone to eliminating actual PU locations. Clearly, limiting adversary access to SU information produces a harder adversary problem, stressing the importance of security measures for the access system and connected devices.

PU Privacy Bound: Figure 8a re-plots the adversary estimate progression using the composite method for the low resolution scenario and also plots $E\{\tau_w\}$ per (8) as the expected time to eliminate all locations but those containing actual PUs, shown with the vertical dashed line. $n_p = 4$ is also plotted as a horizontal dashed line for easy visualization of when all possible locations are eliminated. We see that the bound provides a reasonable approximation in this case. Here we note that τ_w actually corresponds to the number of i.i.d. observations before w locations are eliminated. Relating τ_w



(a) Candidate Locations (b) Sum Distance Error

Fig. 7. Adversary Estimates Over Time - Adversary Case #3.



(a) Privacy bound accuracy (b) Privacy strategies

Fig. 8. Application of the PU privacy bound.

to an actual time thus requires some interpretation. In this example, we assumed each observation at 30 second intervals would be sufficiently independent and therefore the plotted bound equals $t_s E\{\tau_w\}$ where $t_s = 30$. In other scenarios, this may not hold. By examining the stochastic mixing [21] of the SU observations, we could potentially derive a unit of time for independent observations suitable for the bounding computation. Such analysis is beyond the scope of this paper.

Returning to the overall PU privacy problem, we consider a scenario of 100 SUs of which the adversary can observe 10, and in Figure 8b we provide a scatter plot of our computed bound for all possible PU strategies involving $b_P \in \{0, ..., 64\}$ and $n_p \in \{1, ..., 30\}$. The plot shows the computed performance of each PU strategy in terms of both expected SU utility and expected privacy time. We clearly observe the wide range of options in trading SU utility for PU privacy. We also run algorithm 1 with the constraint that the privacy should be preserved for 1,000 time slots, plotting the requirement as a vertical dashed line and highlighting the resulting strategy in the figure. This demonstrates that the greedy algorithm, and thus a PU, can at least identify a reasonable, though not necessarily optimal obfuscation strategy.

VII. Conclusions

Dynamic spectrum access systems offer potential gains in efficient spectrum use but also pose risks for user privacy. We have examined general strategies for PU privacy and studied PU location privacy in detail, observing that the potential dependency of an adversary's observation on the entire user topology potentially presents a harder problem than has been considered in other location privacy settings. We anticipate the approaches we presented in this paper are also applicable in assessing the privacy of other PU parameters under a more general framework, which would also include other interesting variations of the problem such as PUs which alter their operational behavior to improve privacy and adversaries with sensing capabilities.

References

[1] *Report and Order and Second Further Notice of Proposed Rulemaking*, Federal Communications Commission 15-47 GN Docket No. 12-354, April 2015.

[2] S. Deb, V. Srinivasan, and R. Maheshwari, *Dynamic Spectrum Access in DTV Whitespaces: Design Rules, Architecture and Algorithms*, in Proceedings of the 15th Annual International Conference on Mobile Computing and Networking, New York, NY, USA, 2009, pp. 1-12.

[3] J.-M. Park, J. H. Reed, A. A. Beex, T. C. Clancy, V. Kumar, and B. Bahrak, *Security and Enforcement in Spectrum Sharing*, Proceedings of the IEEE, vol. 102, no. 3, pp. 270-281, Mar. 2014.

[4] B. Bahrak, S. Bhattarai, A. Ullah, J.-M. Park, J. Reed, and D. Gurney, *Protecting the primary users operational privacy in spectrum sharing*, IEEE International Symposium on Dynamic Spectrum Access Networks, 2014, pp. 236-247

[5] B. Bahrak, *Ex Ante Approaches for Security, Privacy, and Enforcement in Spectrum Sharing (Doctoral Thesis)*, Virginia Technical Institute and State University, Dec. 2013.

[6] A. Robertson, J. Molnar, and J. Boksiner, *Spectrum Database Poisoning for Operational Security in Policy-Based Spectrum Operations*, IEEE Military Communications Conference, 2013, pp. 382-387.

[7] C. Dwork, *Differential privacy*, In Proceedings of the International Colloquium on Automata, Languages and Programming, pp. 1-12. 2006.

[8] C. Dwork and A. Smith, *Differential Privacy for Statistics: What we Know and What we Want to Learn*, Journal of Privacy and Confidentiality, Number 2, pp. 135-154, 2009.

[9] S. Kasiviswanathan and A. Smith, *On the Semantics of Differential Privacy: A Bayesian Formulation*, Journal of Privacy and Confidentiality, vol. 6, no. 1, Aug. 2014.

[10] S. Li, H. Zhu, Z. Gao, X. Guan, K. Xing, and X. Shen, *Location privacy preservation in collaborative spectrum sensing*, Proceedings IEEE INFOCOM, 2012, pp. 729-737.

[11] W. Wang and Q. Zhang, *Privacy-Preserving Collaborative Spectrum Sensing With Multiple Service Providers*, IEEE Transactions on Wireless Communications, vol. 14, no. 2, pp. 1011-1019, Feb. 2015.

[12] Z. Gao, H. Zhu, S. Li, S. Du, and X. Li, *Security and privacy of collaborative spectrum sensing in cognitive radio networks*, IEEE Wireless Communications, vol. 19, no. 6, pp. 106-112, Dec. 2012.

[13] Z. Gao, H. Zhu, Y. Liu, M. Li, and Z. Cao, *Location privacy in database-driven Cognitive Radio Networks: Attacks and countermeasures*, Proceedings IEEE INFOCOM, 2013, pp. 2751-2759.

[14] M. Clark and K. Psounis, *Efficient resource scheduling for a secondary network in shared spectrum*, in IEEE Conference on Computer Communications (INFOCOM), 2015, pp. 12571265.

[15] X. Gong, A. Ispas, and G. Ascheid. *Outage-constrained power control in spectrum sharing systems with partial primary CSI*, IEEE Global Communications Conference (GLOBECOM), 2012, pp. 3879-3885.

[16] L. Zheng, and C. W. Tan. *Maximizing Sum Rates in Cognitive Radio Networks: Convex Relaxation and Global Optimization Algorithms*, IEEE Journal on Selected Areas in Communications 32, no. 3 (March 2014): 667-680.

[17] M. H. Islam and Z. Dziong. *Joint Link Scheduling, Beamforming and Power Control for Maximizing the Sum-Rate of Cognitive Wireless Mesh Networks*, 73rd IEEE Vehicular Technology Conference (VTC Spring), 2011, pp. 1-5.

[18] L. Sweeney, *k-ANONYMITY: A MODEL FOR PROTECTING PRIVACY*, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems., vol. 10, no. 05, pp. 557-570, Oct. 2002.

[19] R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux, *Quantifying Location Privacy*, in IEEE Symposium on Security and Privacy, 2011, pp. 247262.

[20] *Commerce Spectrum Management Advisory Committee Final Report Working Group 1 – 1695-1710 MHz Meteorological-Satellite Rev. 1*, National Telecommunications and Information Administration, July 2013.

[21] M. B. Rajarshi, *Statistical Inference for Discrete Time Stochastic Processes*. Springer Science & Business Media, 2014.