Wang, J; Gong, Y.; Qian, Ljun; Jäntti, Riku; Miao, Pan; Han, Zhu

Operational Privacy Preservation via Data-Driven Optimization

Published in:
GLOBECOM 2017 - 2017 IEEE Global Communications Conference

DOI:
10.1109/GLOCOM.2017.8254108

Published: 01/12/2017

Please cite the original version:
Primary Users’ Operational Privacy Preservation via Data-Driven Optimization

Jingyi Wang*, Yanmin Gong†, Lijun Qian‡, Riku Jäntti§, Miao Pan*, and Zhu Han*

*Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204
†School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078
‡Department of Electrical and Computer Engineering, Prairie View A&M University, Prairie View, TX 77446
§Department of Communications and Networking, Aalto University, Espoo, Finland

Abstract—Recently opened spectrum within 3550-3700 MHz provides more accessing opportunities to secondary users (SUs), while it also raises concerns on the operational privacy of primary users (PUs), especially for military and government. In this paper, we propose to study the trade-off between PUs’ temporal privacy and SUs’ network performance using the data-driven approach. To preserve PUs’ temporal operational privacy, we develop an obfuscation strategy for PUs, which allows PUs to intentionally add dummy signals to change the distribution of temporal spectrum availability, and confuse the adversary. While generating the dummy signals for privacy, the PUs have to consider the utility of SUs and try their best to satisfy SUs’ uncertain traffic demands. Based on the historical data, we employ a data-driven risk-averse model to characterize the uncertainty of SUs’ demands. With joint consideration of PUs’ privacy and uncertain SUs’ demands, we formulate the data-driven risk-averse stochastic optimization, and provide corresponding solutions. Through numerical simulations, we show that the proposed scheme is effective in preserving PUs’ temporal operational privacy while offering good enough spectrum resources to satisfy SUs’ traffic demands.

I. INTRODUCTION

In recent years, the exploding increase of mobile wireless devices and the proliferation of wireless services have accelerated the growth in demand for radio spectrum. With limited unlicensed spectrum, regulators are turning to dynamic spectrum sharing and looking for advanced techniques to improve spectrum utilization. As one promising technology, cognitive radio (CR) allows secondary users (SUs) to access the idle spectrum in temporal and spatial domain opportunistically, when primary users (PUs) are not active. To further meet the ever-increasing demand for spectrum, Federal Communication Commission (FCC) and National Telecommunications and Information Administration (NTIA) have agreed to open up the 3550-3700 MHz band for unlicensed communications [1], [2]. Note that most frequencies within 3550-3700 MHz are traditionally used by government agencies, e.g., Department of Defense [2], [3], and the operational information (such as time of use, geographical locations, anti-jamming capability, and so on) of government facilities, e.g., military radars, are very sensitive or even classified. Therefore, how to maintain the PUs’ operational privacy while increasing SUs’ spectrum accessing opportunities poses great challenges.

There are several pioneering works about PUs’ privacy preservation in existing literature. For example, Clark et al. in [3] discussed several attack models and PUs’ obfuscation strategies, based on the assumption that all the information of PUs and SUs are stored in a database. The adversary might hack into the database or compromise SUs’ devices to infer PUs’ location information. Robertson et al. in [4] proposed to add false spectrum allocation entries into the database to prevent the adversary from learning the operational privacy of PUs. Baharak et al. in [5] employed obfuscation methods to preserve PUs’ privacy by constructing forbidden contours. A perturbation strategy was used to develop a pentagon-shaped contour, which envelops the PU’s actual contour to preserve PU’s location privacy. Dwork et al. [6] added noises in the sensing reports before publishing the desired sensing statistics, which achieved differential privacy of PUs. Since simply adding noise signals may degrade the performance of collaborative sensing results, Gao et al. in [7] further proposed a distributed dummy report injection protocol, which jointly prevents the pollution of the aggregation results and preserves location privacy of PUs. Based on attributed-based encryption techniques, Liu et al. in [2] developed the query policy for PUs’ spectrum usage database to protect PUs’ location privacy. In military communications, Fu et al. in [8] proposed a method that hides traffic characteristics from eavesdroppers by padding the traffic with constant/variable interarrival times, to mitigate the traffic analysis attacks. In addition, there are some previous works related to the time-based traffic model. For instance, Bonal et al. in [9]–[11] show that if the underlaying scheduler is fair, the flow-level (TCP) [12] throughput and delay admit simple time based form, which is independent of the actual inter-arrival distribution between MAC layer packets. However, most existing schemes are limited about the privacy preservation of PUs’ temporal operational privacy (i.e., the time of usage). The temporal operations of PUs might include highly confidential or even classified information (e.g., the operational time of military radars). If such information is obtained by a malicious party, it may jeopardize national security and people’s safety [13]. In addition, most existing PUs’ privacy preserving designs have very limited consideration about how to create more accessing opportunities to satisfy SUs’ traffic demands and improve spectrum utilization, which is the original purpose for opening up 3550-3700 MHz.
From the aspect of PUs’ obfuscation strategy designing, it will be good for the PUs to precisely predict SUs’ traffic demands. In this way, the PUs can intentionally add dummy signals to obfuscate the attackers while trying their best to satisfy SUs’ traffic demands. However, it is a challenging problem to characterize the uncertainty of SUs’ traffic demands. Some previous efforts tried to employ robust optimization to address this issue. For instance, Lunden et al. in [14] proposed a robust computationally nonparametric cyclic correlation estimator, which does not require the distribution information of users’ traffic. Gong et al. in [15] designed an algorithm to search optimal detection bound considering signal uncertainty. However, the robust optimization approach can be very conservative, since its objective is to minimize the worst case cost or worst case effectiveness. If PUs add too many dummy signals, according to the overly conservative analysis for privacy preservation, it would reduce the utility of SUs.

To address the issues above, we propose a novel PUs’ obfuscation strategy design, formulate the PUs’ operational privacy preservation problem into a data-driven risk-averse optimization, and provide robust solutions. Our salient contributions are summarized as follows:

- We introduce a new privacy preserving framework for PUs’ obfuscation strategy design, which jointly considers PUs’ operational privacy in the temporal domain, the obfuscation cost of PUs, the uncertainty of SUs’ demands, and SUs’ traffic demand satisfaction. Under such a framework, when PUs add dummy signals to obfuscate the adversary, they also need to consider the trade-off between preserving PUs’ temporal privacy and satisfying SUs’ traffic requirements, and thus cannot arbitrarily generate dummy signals for privacy preserving purposes.
- Under the proposed framework, with abundant historical data of SUs’ traffic demands, we propose to let the PUs employ data-driven modeling to characterize the uncertainty of SUs’ traffic demands. The PUs can build up a reference SUs’ demand distribution from the historical data, and generate the predicted SUs’ demand distribution, which is close enough to the reference distribution at a certain confidence level.
- Based on the modeling of SUs’ uncertain traffic demands and temporal operational privacy metrics, we formulate the PUs’ temporal privacy preservation problem into a risk-averse two-stage stochastic optimization, develop algorithms for robust solutions, and conduct simulations to verify our theoretical analysis.

The rest of paper is organized as follows. In Sec. II, we introduce the network model and formulate the PUs’ operational privacy preservation problem. In Sec. IV, we develop the solutions to the proposed problem. Simulation results and discussions are presented in Sec. V, and the conclusion remarks are drawn in Sec. VI.

II. SYSTEM DESCRIPTION

A. Network Configuration

As shown in Fig. 1, we consider a CR network consisting of \( N = \{1, 2, \cdots, i, \cdots, N\} \) SUs and \( M = \{1, 2, \cdots, j, \cdots, M\} \) PUs sharing 3550-3700 MHz bands.

Following the principles of CR communications, SUs can opportunistically use the band when the PU owning that band is not active, and SUs must evacuate if the PU comes back. Here, we assume each PU is licensed to use a dedicated band, and each SU can only opportunistically access one band at a time. For an available spectrum band, different SUs are allowed to dynamically share the band in time-division multiple access (TDMA) manners. To preserve temporal operational privacy, PUs will send obfuscating dummy signals periodically, where the fixed period is denoted by \( \mathcal{T} \). Let \( T_j \) represent the actual temporal spectrum availability for band \( j \) (i.e., available time for SUs’ opportunistic spectrum accessing before PU \( j \) adds dummy signals), and \( y_j \) (\( y_j \leq T_j \)) be the transformed temporal spectrum availability for band \( j \) (i.e., the available time for SUs’ opportunistic spectrum accessing after PU \( j \) adds dummy signals). Given the transmission rate, let a random variable \( \xi_i \) denote the required time to deliver the uncertain traffic demands of SU \( i \) within \( \mathcal{T} \). For simplicity, we call \( \xi_i \) the demand of SU \( i \) in the rest of this paper, and let \( \mathbb{P}_i \) be the distribution of \( \xi_i \). For instance, \( \mathcal{T} = 60 \) mins, and PU \( j \) is actively using band \( j \) for 20 mins, so that \( T_j \) is equal to 40 mins. After PU \( j \) executes obfuscation strategy, \( y_j = 30 \) mins, and the demand of SU \( i \) is \( \xi_i = 25 \) mins.

B. Attack Model

In this work, we consider passive adversaries, who may learn the operational time of PUs from either spectrum database or collective spectrum sensing results by compromised SUs. The compromised SUs do not intercept or modify the messages sent by PUs. Specifically, the adversaries can either eavesdrop the communication between the spectrum...
For the PU’s perspective, to preserve the temporal operational privacy from passive attackers, the PU executes obfuscation strategy by generating dummy signals for a certain time period when it actually has no traffic. In that way, the adversary cannot distinguish them from true signals, when the adversary detects those dummy signals, either through database or collective spectrum sensing. Thereafter, the adversary cannot obtain transformed temporal spectrum occupation of the PUs based on detected signals, which is a combination of the dummy and true signals. As long as the dummy signals are sent out frequently enough, the PUs’ true operations can be hidden in those signals and the operational privacy of PUs can be preserved. Thus, the utility function of PUs’ operational privacy preservation can be written as

\[ U_{PU,j}(y_j) = c(T_j - y_j), \]  

where \( c \) is a temporal privacy coefficient, \( T_j \) is the actual spectrum availability, and \( y_j \) is the transformed spectrum availability after the PU \( j \)’s obfuscation strategy is executed. We can see that if \( T_j - y_j \) is sufficiently large, PUs’ temporal operational privacy is preserved effectively.

From the SUs’ perspective, they attempt to transmit on available spectrum to satisfy their own demand. Since SUs can only observe the transformed spectrum availability of PUs, i.e., the spectrum availability after PUs execute obfuscation strategy, we denote the transformed spectrum availability for SU \( i \) over spectrum band \( j \) as \( x_{i,j} \). Assuming the SU’s traffic can be perfectly split, we let \( \sum_{j=1}^{M} x_{i,j} \) denote the total available time that SU \( i \) can transmit over all spectrum bands. We define \( d(\xi_i) \) as the actual time needed to satisfy the traffic demand of SU \( i \). Then, \( \min(\sum_{j=1}^{M} x_{i,j}, d(\xi_i)) \) represents the traffic delivery time of SU \( i \). Specifically, when \( d(\xi_i) < \sum_{j=1}^{M} x_{i,j} \), which indicates that the traffic demand is less than the transformed available spectrum supply, SU \( i \) will only transmit \( d(\xi_i) \) to meet its service demands. On the other hand, if transformed available spectrum supply for SU \( i \) is less than its real demand, i.e., \( \sum_{j=1}^{M} x_{i,j} < d(\xi_i) \), then SU \( i \) will deliver \( \sum_{j=1}^{M} x_{i,j} \).

So, the utility function of SU \( i \) is

\[ U_{SU,i}(\xi_i) = bE_{\xi_i}(\min_{j=1}^{M} x_{i,j}, d(\xi_i)), \]  

where \( b \) is SUs’ traffic delivery coefficient, and traffic demand \( \xi_i \) follows the distribution \( F_i \).

### III. Obfuscation Strategy and Problem Formulation

#### A. Utility Functions of PUs and SUs

From the PU’s perspective, to preserve the temporal operational privacy from passive attackers, the PU executes obfuscation strategy by generating dummy signals for a certain time period when it actually has no traffic. In that way, the adversary cannot distinguish them from true signals, when the adversary detects those dummy signals, either through database or collective spectrum sensing. Thereafter, the adversary cannot obtain transformed temporal spectrum occupation of the PUs based on detected signals, which is a combination of the dummy and true signals. As long as the dummy signals are sent out frequently enough, the PUs’ true operations can be hidden in those signals and the operational privacy of PUs can be preserved. Thus, the utility function of PUs’ operational privacy preservation can be written as

\[ U_{PU,j}(y_j) = c(T_j - y_j), \]  

where \( c \) is a temporal privacy coefficient, \( T_j \) is the actual spectrum availability, and \( y_j \) is the transformed spectrum availability after the PU \( j \)’s obfuscation strategy is executed. We can see that if \( T_j - y_j \) is sufficiently large, PUs’ temporal operational privacy is preserved effectively.

From the SUs’ perspective, they attempt to transmit on available spectrum to satisfy their own demand. Since SUs can only observe the transformed spectrum availability of PUs, i.e., the spectrum availability after PUs execute obfuscation strategy, we denote the transformed spectrum availability for SU \( i \) over spectrum band \( j \) as \( x_{i,j} \). Assuming the SU’s traffic can be perfectly split, we let \( \sum_{j=1}^{M} x_{i,j} \) denote the total available time that SU \( i \) can transmit over all spectrum bands. We define \( d(\xi_i) \) as the actual time needed to satisfy the traffic demand of SU \( i \). Then, \( \min(\sum_{j=1}^{M} x_{i,j}, d(\xi_i)) \) represents the traffic delivery time of SU \( i \). Specifically, when \( d(\xi_i) < \sum_{j=1}^{M} x_{i,j} \), which indicates that the traffic demand is less than the transformed available spectrum supply, SU \( i \) will only transmit \( d(\xi_i) \) to meet its service demands. On the other hand, if transformed available spectrum supply for SU \( i \) is less than its real demand, i.e., \( \sum_{j=1}^{M} x_{i,j} < d(\xi_i) \), then SU \( i \) will deliver \( \sum_{j=1}^{M} x_{i,j} \).

So, the utility function of SU \( i \) is

\[ U_{SU,i}(\xi_i) = bE_{\xi_i}(\min_{j=1}^{M} x_{i,j}, d(\xi_i)), \]  

where \( b \) is SUs’ traffic delivery coefficient, and traffic demand \( \xi_i \) follows the distribution \( F_i \).

#### B. PUs’ Operational Privacy Preserving Optimization

Based on the utility functions of PUs and SUs, we expect an obfuscation strategy jointly considering PUs’ operational privacy preservation and the satisfaction of SUs’ uncertain traffic demands. Obviously, regardless of the PU’s power consumption, generating more dummy signals can better protect the PU’s operational privacy while reducing the available opportunistic accessing time of the SUs and thus diminishing SUs’ traffic delivery. Taking the trade-off between PUs’ privacy and SUs’ utility into account, we can formulate the PUs’ obfuscation strategy design into an optimization, a classic two-stage stochastic programming (SP) problem, described as follows:

\[
\begin{align*}
\text{Max} & \quad \sum_{j=1}^{M} c(T_j - y_j) + b \sum_{i=1}^{N} \mathbb{E}_{\xi_i} \left( \min_{j=1}^{M} x_{i,j}, d(\xi_i) \right), \\
\text{s.t.} & \quad T_j - y_j \geq \lambda, \\
& \quad \sum_{i=1}^{N} x_{i,j} \leq y_j.
\end{align*}
\]  

The function \( \min(\cdot, \cdot) \) in (3) returns the smaller value of \( \sum_{j=1}^{M} x_{i,j} \) and \( d(\xi_i) \). The constraint (5) indicates that total transmission time for SUs over PU \( j \)’s spectrum should be less than the total transformed available spectrum supply of PU \( j \). Besides, to preserve PU \( j \)’s operational privacy, the time period of the sent dummy signals, i.e., \( T_j - y_j \), is then required to be larger than a certain predefined privacy threshold \( \lambda \), which is a constant, as shown in (4).

As we know, it is difficult to know the actual probability distribution of SUs’ demands in practice. In this paper, we employ a data-driven approach, i.e., the risk-averse stochastic optimization approach (RA-SP) allowing distribution ambiguity [16], to characterize the uncertainty of SUs’ demands. Instead of deriving a true distribution for the unknown parameter \( \xi \), this optimization approach constructs a confident set \( D \), which allows the distribution ambiguity to be within \( D \) under a certain confidence level (e.g., 99%). With RA-SP, considering the worst-case distribution, we can reformulate the problem as follows.

\[
\begin{align*}
\text{Max} & \quad \sum_{j=1}^{M} c(T_j - y_j) + \min_{\xi \in D} \left[ b \sum_{i=1}^{N} \mathbb{E}_{\xi_i} \left( \min_{j=1}^{M} x_{i,j}, d(\xi_i) \right) \right], \\
\text{s.t.} & \quad T_j - y_j \geq \lambda, \\
& \quad \sum_{i=1}^{N} x_{i,j} \leq y_j.
\end{align*}
\]  

2Here, we assume SUs use energy detection for spectrum sensing.
We use a distance measurement proposed in [17], [18] to quantify the distance between two distributions. Specifically, a predefined distance measure \(d(P_i^0, P_i)\) is constructed on confident set \(D\), where \(P_i^0\) is the reference distribution estimated from historical data, and \(P_i\) is the ambiguous distribution of SU \(i\). The distance \(d\) and confident set \(D\) can be defined as follows:
\[
D = \{P_i : d(P_i^0, P_i) \leq \theta\},
\]
\[
d_c(P_i^0, P_i) = \sup_{h \in \mathcal{H}} \left| \int \Omega h dP_i^0 - \int \Omega h dP_i \right|,
\]
where the distance under \(\zeta\)-structure probability metric is denoted by \(d_c(\cdot, \cdot)\), the tolerance is denoted by \(\theta\), and \(\mathcal{H}\) is a family of real-valued bounded measurable functions on \(\Omega\) (the sample space on \(\xi\)). Tolerance \(\theta\) is correlated to historical data size. It can be easily inferred that the more historical data that the PU can observe, the tighter \(D\) would be, and the closer the ambiguous distribution \(P_i\) would be to \(P_i^0\). More details of \(\zeta\)-structure probability metric will be illustrated in the following section.

IV. RISK-AVERSE STOCHASTIC PROGRAMMING FOR PRESERVING TEMPORAL OPERATIONAL PRIVACY

This section is organized as follows. First, we describe how to determine the reference distribution \(P_i^0\) for SU \(i\). Second, we represent how to determine tolerance \(\theta\) on the amount of historical data under \(\zeta\)-structure. Lastly but not least, we develop algorithms to solve the problem with respect to different probability distance metrics.

**A. Reference Distribution**

\[
P_i^0(x \leq X) = \frac{1}{Q} \sum_{q=1}^{Q} \delta_{\xi_q^0}(x)
\]

First, we introduce how we determine the reference distribution \(P_0\). Suppose we use a set of historical data \(\{\xi_1^0, \xi_2^0, \xi_3^0, \cdots, \xi_q^0\}\) to estimate the reference distribution \(P_0\). We utilize the empirical distribution of the historical data samples to construct \(P_0\). To be specific, the distribution of empirical data is defined as (9), the indicator variable \(\delta_{\xi_q^0}(x)\) is equal to 1 when \(\xi_q^0 \leq x\) and 0 otherwise. After that, the reference distribution data can be represented by its mass probability \(p_i^0\) which equals to the ratio of the number of historical data samples matching \(\xi_q^0\) and \(Q\) since the supporting space is discrete.

**B. Converge Rate under \(\zeta\)-Probability Metrics**

As described in Sec. IV, we employ three metrics and solve our problem under these constraints correspondingly. We define \(\rho(x, y)\) as the distance between two variables \(x\) and \(y\), and \(n\) is the dimension of \(\Omega\). \(\mathbb{P} = \mathcal{L}(x)\) represents random variables \(x\) follows distribution \(\mathbb{P}\). The metrics are derived as follows.

- **Kantorovich metric**: denoted as \(d_K(P_i^0, P_i)\), \(\mathcal{H} = \{h : ||h||_L \leq 1\}\), where \(||h||_L := \sup\{h(x) - h(y) / \rho(x, y) : x \neq y \in \Omega\}\).

- **Fortet-Mourier metric**: denoted as \(d_{FM}(P_i^0, P_i)\), \(\mathcal{H} = \{h : ||h||_{C} \leq 1\}\), where \(||h||_{C} := \sup\{h(x) - h(y) / c(x, y) : x \neq y \in \Omega\}\) and \(c(x, y) = \rho(x, y) max\{1, \rho(x, a)^{p-1}, \rho(y, a)^{p-1}\}\) for some \(p \geq 1\) and \(a \in \Omega\).

- **Uniform metric**: denoted as \(d_U(P_i^0, P_i)\), \(\mathcal{H} = \{I_{(-\infty, t)} : t \in R^n\}\).

The relationships among metrics are represented as follows:
\[
d_{FM}(P_i^0, P_i) \leq \Lambda \cdot d_K(P_i^0, P_i),
\]
where \(\Lambda = \max\{1, C(p^{-1})\}\), \(C\) is the diameter of \(\Omega\). From the definition of metrics and relationships between metrics under \(\zeta\)-structure, we can derive the convergence property and convergence rate accordingly.

For the Uniform metric, the convergence rate is obtained by utilizing the Dvoretzky-Kiefer-Wolfowitz inequality [19]:

**Proposition 1** For a single dimension case (i.e., \(n = 1\)),
\[
\mathbb{P}(d_U(P_i^0, P_i) \leq \theta) \geq 1 - \exp\left(-\frac{\theta^2 Q}{2}\right).
\]

In [16], the converge rate of Kantorovich metric is shown below:

**Proposition 2** For a general dimension case (i.e., \(n \geq 1\)),
\[
\mathbb{P}(d_K(P_i^0, P_i) \leq \theta) \geq 1 - \exp\left(-\frac{\theta^2 Q}{2C^2}\right).
\]

From the relation between the Fortet-Mourier metric and Kantorovich metric (10), with Proposition 2, we can easily derive the convergence rate of other metrics.

**Corollary 1** For a general dimension (i.e., \(n \geq 1\)), we have
\[
\mathbb{P}(d_{FM}(P_i^0, P_i) \leq \theta) \geq 1 - \exp\left(-\frac{\theta^2 Q}{2C^2\Lambda^2}\right).
\]

With the convergence rate in (11)-(13), we can calculate the tolerance \(\theta\) accordingly. For instance, in the Kantorovich metric, we assume the confidence level is \(\eta\). Therefore, \(\mathbb{P}(d_K(P_i^0, P_i) \leq \theta) \geq 1 - \exp(-\frac{\theta^2 Q}{2C^2}) = \eta\) according to (11), and \(\theta = C\sqrt{2\log(1/(1-\eta))}/Q\). After that, we explore how to solve the problem in (6). The sample space is \(\Omega = \{\xi_1, \xi_2, \cdots, \xi_V\}\). Then the formulation can be simplified as:
\[
\text{Max} \sum_{j=1}^{M} c(T_j - y_j) + \text{Min}_{p_i} \sum_{i=1}^{N} \sum_{k=1}^{V} b_i k \min\left(\sum_{j=1}^{M} x_{i,j}, d(x_i)\right),
\]
s.t.
\[
T_j - y_j \geq \lambda_i \sum_{i=1}^{N} \sum_{j=1}^{M} x_{i,j} \leq y_j,
\]
\[
\sum_{k=1}^{V} p_i k = 1, \forall i = 1, \cdots, N,
\]
\[
\max_{k=1}^{V} \sum_{j=1}^{M} h_{k} s_{k,j} \leq \theta, \forall h \leq 1, j \leq 1.
\]
where $|h|_C$ defined according to different metrics. In the Kantorovich metric and Bounded-Lipschitz metric, $|h_x - h_y| \leq \rho(\zeta^x, \zeta^y)$. In the Fortet-Mourier, $|h_x - h_y| \leq \rho(\zeta^x, \zeta^y)\max\{1, \rho(\zeta^x, a)^{p-1}, \rho(\zeta^y, a)^{p-1}\}$.

The constraints in (16)-(17) can be summarized as $\sum_k a_{kl} h_k \leq b_{kl}, l = 1, \cdots, L$. To reformulate the constraints, we consider the problem:

$$\min_{h_k} \sum_{k=1}^V h_k \rho_{ik}^p - \sum_{k=1}^V h_k \rho_{ik}^p,$$

s.t. $\sum_{k=1}^V a_{kl} h_k \leq b_{kl}, l = 1, \cdots, L.$ (19)

Its dual problem is represented as:

$$\max \sum_{l=1}^L b_l u_l,$$

s.t. $\sum_{l=1}^L a_{kl} u_l \geq \rho_{ik}^p - \rho_{ik}^p, \forall k = 1, \cdots, V,$$ (21)

where $u$ is the dual variable. Accordingly, the formulation can be reformulated as follows:

$$\max \sum_{j=1}^M c(T_j - y_j) + \min_{p_i} \sum_{i=1}^N \sum_{k=1}^V b_p^k \min \left\{ \sum_{j=1}^M x_{i,j}, d(\xi_i) \right\},$$

(22)

(SP-M) s.t. $T_j - y_j \geq \lambda$, (23)

$$\sum_{i=1}^N x_{i,j} \leq y_j,$$ (24)

$$\sum_{k=1}^V p_i^k = 1, \sum_{l=1}^L b_l u_l \leq \theta,$$ (25)

$$\sum_{l=1}^L a_{il} u_l \geq \rho_{ik}^p - \rho_{ik}^p, \forall i = 1, \cdots, N.$$ (26)

For the Uniform metric, we can have the reformulation from the Uniform metric definition:

$$\max \sum_{j=1}^M c(T_j - y_j) + \min_{p_i} \sum_{i=1}^N \sum_{k=1}^V b_p^k \min \left\{ \sum_{j=1}^M x_{i,j}, d(\xi_i) \right\},$$

(27)

(SP-U) s.t. $T_j - y_j \geq \lambda$, (28)

$$\sum_{i=1}^N x_{i,j} \leq y_j;$$ (29)

$$\sum_{i=1}^N \sum_{k=1}^V p_i^k = 1, \forall i = 1, \cdots, N,$$ (30)

$$\sum_{k=1}^V (p_i^k - \rho_{ik}^p) \leq \theta, \forall l = 1, \cdots, L.$$ (31)

The formulation SP-M and SP-U can be solved by CPLEX, etc. We also summarize the algorithm for the problem in Algorithm 1.

### Algorithm 1 Algorithm for Obfuscation Strategy

1: **Input:** Historical data $\xi_i^1, \xi_i^2, \cdots, \xi_i^L$ from true distribution.

2: **Output:** Objective value of the added time period of dummy signals.

3: Obtain the reference distribution $P_0(x)$ and tolerance $\theta$ based on the historical data.

4: Use the reformulation (SP-M) or (SP-U) to solve the problem.

5: Output the solution.

### V. PERFORMANCE EVALUATION

For illustrative purposes, in the simulations, we consider a CR network of 1 PU and 5 SUs. The actual available time of the PU’s spectrum is $T = 30$ mins in a particular period $T = 60$ mins. Moreover, we set the utility parameter for measuring operational privacy level $c$ to be 3, and the utility parameter for SUs’ traffic delivery $b$ to be 5. We assume the SUs’ traffic demand follows a discrete distribution with two scenarios: 10
The proposed scheme can provide a design guide-
current system. However, for a more PU’s privacy oriented
than the deductions of the denied SUs’ traffic demands to
utility coefficient
signal time period increases, the overall system utility under all
the results are shown in Fig. 3. We find that as the dummy
iss e tf r o m1 0m i n st o2 0m i ns ,a n d
behind the results is that the value \( \theta \) decreases as the size of
historical data increases. Therefore, the optimized problem (6)
becomes less conservative. We can also see that when the
number of samples exceeds 200, the gaps between results
under the Fortet-Mourier metric and optimal value is very
small. Furthermore, when sample size is 300, the gaps between
system utility values are small under all metrics. Moreover, we
study the performance under preserving privacy scheme. We
set \( \lambda = 15 \) mins, which indicates that there is at least 15-
minute gap between the transformed PUs’ spectrum available
time and the actual unoccupied period of the PU’s spectrum.
It can be observed that in Fig. 2, the total utility decreases
after employing preservation privacy strategy since the PU’s
operational privacy preservation is at the cost of reducing
accessing opportunities for SUs.

In addition, we explore the impacts of dummy signals time
period on the system utility. The total number of historical
sample is 300, and \( \lambda \) is set from 10 mins to 20 mins, and
the results are shown in Fig. 3. We find that as the dummy
signal time period increases, the overall system utility under all
metrics decreases for chosen PU’s privacy coefficient \( c \), SUs’
utility coefficient \( b \), and confidence level. The reason is that
the contributions of PU’s privacy preservation is less important
than the deductions of the denied SUs’ traffic demands to
current system. However, for a more PU’s privacy oriented
system (e.g., \( c \gg b \)), the system utility may increase while
adding more dummy signals. For given PUs’ and SUs’ utility
parameters, the proposed scheme can provide a design guide-
line for such a CR network considering the tradeoff between
PUs’ temporal operational privacy and SUs’ performance.

VI. CONCLUSION

In this paper, we have proposed a novel obfuscation strategy
for PUs within 3550-3700 MHz, which has a joint consid-
eration of preserving PUs’ temporal operational privacy and
satisfying SUs’ uncertain traffic demands. We have employed
the data-driven risk-averse model in our scheme to characterize
SUs’ uncertain demand based on the historical data. With
such a model, we have formulated the PUs’ temporal opera-
tional privacy preservation problem into a risk-averse two-
stage stochastic optimization, developed a robust algorithm
for solutions, and presented simulation results to show the
effectiveness of the proposed scheme in terms of preserving
PUs’ temporal operational privacy and satisfying SUs’ traffic
demands.

REFERENCES

tions Commission, Et docket No. 02-135, November 2002.
preserving scheme for primary users in database-driven cognitive radio
networks,” in Proceeding of the IEEE Global Communications Confer-
ence (GLOBECOM), Washington, DC, December 2016.
shared spectrum be protected?” in Proceeding of the IEEE International
Conference on Computer Communications (INFOCOM), San Francisco,
CA, April 2016.
for operational security in policy-based spectrum operations,” in IEEE
Military Communications Conference, San Diego, CA, November 2013.
[5] B. Baharak, S. Bhattachai, A. Ullah, J.-M. Park, J. Reed, and D. Guney,
“Protecting the primary users operational privacy in spectrum sharing;”
in IEEE International Symposium on Dynamic Spectrum Access Net-
works, Mclean, VA, April 2014.
Conference on Theory and Applications of Models of Computation.
cooperative sensing in cognitive radios,” IEEE Wireless Communications,
padding for statistical traffic analysis attacks,” in 23rd International
[9] T. Bonald, L. Massoulié, A. Proutière, and J. Virtamo, “A queueing anal-
ysis of max-min fairness, proportional fairness and balanced fairness.”
contraints,” in Information Sciences and Systems, 2006 40th Annual
mance analysis of sump on a tls/tcp base,” in The Seventh IFIP/IEEE In-
ternational Symposium on Integrated Network Management (IM 2001),
authentication mechanism for small scale iot applications,” IEEE Sensors
cyclical correlation-based spectrum sensing for cognitive radio,” IEEE
Transactions on Signal Processing, vol. 58, no. 1, pp. 38–52, January
2010.
energy detection with signal uncertainty in cognitive radio networks;” in
program with \( \zeta \)-structure probability metrics;” Available on Optimization
Online, 2015.
[17] G. C. Calafiore, “Ambiguous risk measures and optimal robust port-
folios,” SIAM Journal on Optimization, vol. 18, no. 3, pp. 853–877,
October 2007.
by means of histograms,” Production and Operations Management,
[19] A. Dvoretzky, J. Kiefer, and J. Wolfowitz, “Asymptotic minimax charac-
ter of the sample distribution function and of the classical multinomial