This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2892447, IEEE Transactions on Dependable and Secure Computing

Data-Driven Spectrum Trading with Secondary Users' Differential Privacy Preservation

Jingyi Wang, *Student Member, IEEE*, Xinyue Zhang, *Student Member, IEEE*, Qixun Zhang, *Member, IEEE*, Ming Li, *Member, IEEE*, Yuanxiong Guo, *Member, IEEE*, Zhiyong Feng, *Senior Member, IEEE* and Miao Pan, *Senior Member, IEEE*

***Abstract*—Spectrum trading benefits both secondary users (SUs) and primary users (PUs), while it poses great challenges to maximize PUs' revenue, since SUs' demands are uncertain and individual SU's traffic portfolio contains private information. In this paper, we propose a data-driven spectrum trading scheme which maximizes PUs' revenue and preserves SUs' demand differential privacy. Briefly, we introduce a novel network architecture consisting of the primary service provider (PSP), the secondary service provider (SSP) and the secondary traffic estimator and database (STED). Under the proposed architecture, PSP aggregates available spectrum from PUs, and sells the spectrum to SSP at fixed wholesale price, directly to SUs at spot price, or both. The PSP has to accurately estimate SUs' demands. To estimate SUs' demand, the STED exploits data-driven approach to choose sampled SUs to construct the reference distribution of SUs' demands, and utilizes reference distribution to estimate the demand distribution of all SUs. Moreover, the STED adds noises to preserve the demand differential privacy of sampled SUs before it answers the demand estimation queries from the PSP. With the estimated SUs' demand, we formulate the revenue maximization problem into a risk-averse optimization, develop feasible solutions, and verify its effectiveness through both theoretical proof and simulations.**

***Index Terms*—Spectrum Trading; Differential Privacy; Data-Driven Modeling; Risk-Averse Stochastic Optimization**

I. Introduction

The last decades have witnessed the proliferation of wireless smart devices, such as smartphones, touchable tablets, intelligent voice assistants (e.g., Amazon Echo or Google Home), etc., and the explosion of various wireless services, which exploit wireless accessing technologies to make people's daily life more convenient and comfortable. Correspondingly, there is a dramatic increase in demand for radio spectrum, while most licensed spectrum bands are underutilized in both temporal and spatial domains [1]–[3]. Cognitive radio (CR) is a promising technology to improve spectrum utilization, which enables secondary users (SUs) to access the licensed spectrum opportunistically [1]–[5] when primary users (PUs) are not active. Due to high economic values of spectrum resources, CR technology will potentially initiate spectrum trading, which benefits PUs with monetary gains and SUs with accessing opportunities to satisfy their service demands. Despite those benefits, there are many challenges for pushing spectrum trading in practice. For example, due to hardware limitation of either PUs' or SUs' devices, they may have too limited sensing capability to know some spectrum trading opportunities nearby [4]–[6]; aiming to maximize the revenue, the PU may feel challenging to develop optimal selling strategies due to the SUs' traffic demand uncertainty; the SU may feel difficult to preserve its spectrum trading privacy (i.e., the SU's locations, true evaluation values of certain spectrum, traffic portfolio, etc.) [7]–[9], and so on. Those concerns may make PUs or SUs reluctant to participate in spectrum trading.

To facilitate PUs' and SUs' participation and make spectrum trading practical, recent studies [4]–[6] have introduced spectrum trading architectures based on existing wireless network infrastructure. Under those architectures, primary service provider (PSP) aggregates vacant spectrum bands from PUs [6], and sells the spectrum bands to secondary service provider (SSP) at wholesale price. The SSP will evaluate the spectrum supply uncertainty [5], [6], make the spectrum purchasing decision, and further sell the purchased spectrum to SUs at retailed price. Here, the role of PSP/SSP can be played by base station in cellular networks, eNodeB in LTE networks, or mobile virtual network operator (MVNO), where the PSP/SSP has more sensing power [5], [6] than the individual SU. Although the spectrum trading architectures in [4]–[6] help to capture spectrum accessing opportunities, and the algorithms in [4], [6] mathematically characterize spectrum supply uncertainty, they ignore the SUs' traffic demand uncertainty, which may have negative impact on PSP's revenue maximization. That is, without the accurate knowledge of SUs' traffic demands, the PSP cannot choose the optimal selling strategies to maximize its revenue. Moreover, the approach of using random variables to model the traffic uncertainty in [4], [6] may be good enough to reflect the PU's traffic patterns over a relatively long-term period, but it will not be able to represent SUs' traffic demands in real-time manner.

Therefore, following the framework of spectrum trading

This work was supported in part by the U.S. National Science Foundation under grants US CNS-1343361, CNS-1350230 (CAREER), CNS-1646607, CNS-1702850, and CNS-1801925. This work of Q. Zhang, and Z. Feng was partly supported by the Beijing Natural Science Foundation (No. L172049), National Science and Technology Major Project (2017ZX03001014), National Natural Science Foundation of China (NSFC) (Grant No. 61525101, 61631003), 111 Project of China (B16006). This work of M. Li was partly supported by the US National Science Foundation under grants CNS-1566634 and ECCS1711991.

J. Wang, X. Zhang and M. Pan are with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204. Email: jwang86@uh.edu, xzhang67@uh.edu, mpan2@uh.edu. Q. Zhang and Z. Feng is with the Wireless Technology Innovation Institute, Beijing University of Posts and Telecommunications, Beijing 100876, China. Email: zhangqixun@bupt.edu.cn, fengzy@bupt.edu.cn. M. Li is with the Department of Computer Science and Engineering, the University of Texas at Arlington, TX 76019. Email: ming.li@uta.edu. Y. Guo is with the School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078. Email: richard.guo@okstate.edu.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2892447, IEEE Transactions on Dependable and Secure Computing

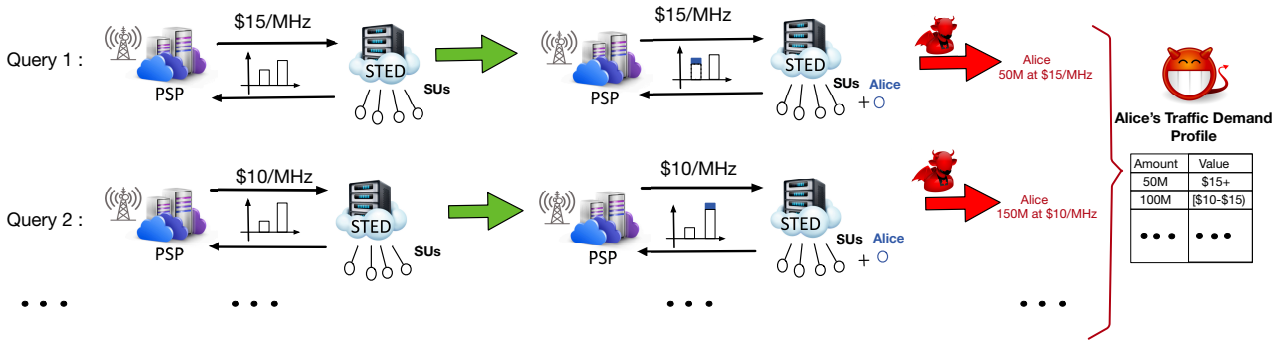


Fig. 1. Illustrative examples for the traffic demand privacy breach of SUs in spectrum trading.

architectures in [4]–[6], in this paper, we further introduce a new entity, called secondary traffic estimator and database (STED), which is responsible for estimating the SUs' traffic demands in real-time manner and answering PSP's queries about SUs' traffic demands as shown in Fig. 1. Considering the large population of SUs in the PSP's coverage boundary, it is not efficient to crowdsource SUs' traffic demands by collecting each SU's demands in terms of time consumption and communication overhead. Thus, we propose to let the STED employ data-driven approach to collect sampled SUs' demands, construct reference demand distribution from sampled demands, and leverage reference distribution to estimate the demand distribution of all SUs.

Now, the leftover challenge hindering spectrum trading is the traffic privacy preservation of the sampled SUs. Taking the query procedure of SUs' demands shown in Fig. 1 as an example, the SU's traffic portfolio privacy is breached as follows. For Query 1, the PSP will send a query about SUs' demand to STED, and the query is what the SUs' demand distribution is, if the price for spectrum accessing is \$15/MHz. The STED will respond to this query with a traffic demand distribution of SUs at the cost of \$15/MHz (e.g., 30% SUs would like to purchase 50M and 70% SUs would like to purchase 150M from 100 SUs in total). If a new SU, Alice, joins the group and she would like to purchase 50M at \$15/MHz, the STED will update the SUs' demand distribution to the PSP's query (i.e.,30.7% SUs would like to purchase 50M, 60.3% SUs would like to purchase 150M from 100 SUs in total). From the differences of distributions, the PSP will derive that Alice would like to purchase 50M at \$15/MHz or above. Through multiple queries, the PSP can easily learn Alice's traffic demand profile, which not only discloses Alice's true evaluation values of spectrum resources [10], but also classifies her personal traffic demands (e.g., voice, video, web browsing, social networking, online gaming, etc.) at different price levels.

In order to protect SUs' traffic demand differential privacy (DP) [9], [11], [12], in this paper, we assume the STED is trustworthy, and entitle the STED to transform the SUs' demand distribution by adding noises before it responds to the PSP's queries. Instead of brutally hammering data-driven approach and DP together, we melt SUs' traffic demand DP into data-driven based spectrum trading, and mathematically prove its effectiveness. Based on that, we propose a novel data-driven based spectrum trading scheme with secondary users' differential privacy preservation (3DPP), whose objective is maximizing the PSP's revenue. Our salient contributions are summarized as follows.

- We propose a novel spectrum trading architecture consisting of the PSP, the SSP, and the STED. Under the proposed architecture, PSP aggregates available spectrum from PUs, and sells the spectrum to the SSP at fixed wholesale price, directly to SUs at spot price, or both as shown in Fig. 2. To optimally split the spectrum sold to SSP/SUs, the PSP sends queries to the STED to estimate SUs' demands. The STED will jointly employ data-driven approach and DP preserving techniques to choose sampled SUs, collect their traffic demands, and respond to the PSP's queries.
- We propose a novel 3DPP spectrum trading scheme, which entitles the STED to construct reference distribution $\mathbb{P}_0$ from sampled SUs' demands via data-driven approach. We employ data-driven risk-averse modeling to characterize the uncertainty of SUs' traffic demands, and ensure the uncertainty distance between the reference distribution $\mathbb{P}_0$ and the real traffic demand distribution of all SUs $\mathbb{P}$ is close enough. Besides, we let the STED add noises drawn from Laplace distribution to $\mathbb{P}_0$, and further establish a SUs' traffic demand reference distribution under ϵ-DP, $\mathbb{P}_0'$.
- We mathematically prove that the 3DPP scheme is able to preserve the sampled SUs' traffic demands under ϵ-DP, the references distribution under ϵ-DP, $\mathbb{P}_0'$, and real distribution $\mathbb{P}$ satisfy the data-driven requirements, and the uncertainty distance between the two distributions is close enough, i.e., $\mathbb{P}(d_k(\mathbb{P}_0', \mathbb{P} \leq \theta)) \geq 1 - \exp(-\frac{\theta^2}{2\varnothing^2}V + V\epsilon)$ for Kantorovich metric. Similar proof is applicable for other distribution distance metrics[1].
- Based on the modeling above, we formulate the PSP's revenue maximization into a risk-averse two-stage stochastic problem (RA-SP). To resolve the problem, we utilize ζ-structure probability metric to construct confidence set, and convert the problem into a traditional two-stage robust optimization. We develop algorithms

[1]Please refer to Sec. IV for details

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2892447, IEEE Transactions on Dependable and Secure Computing

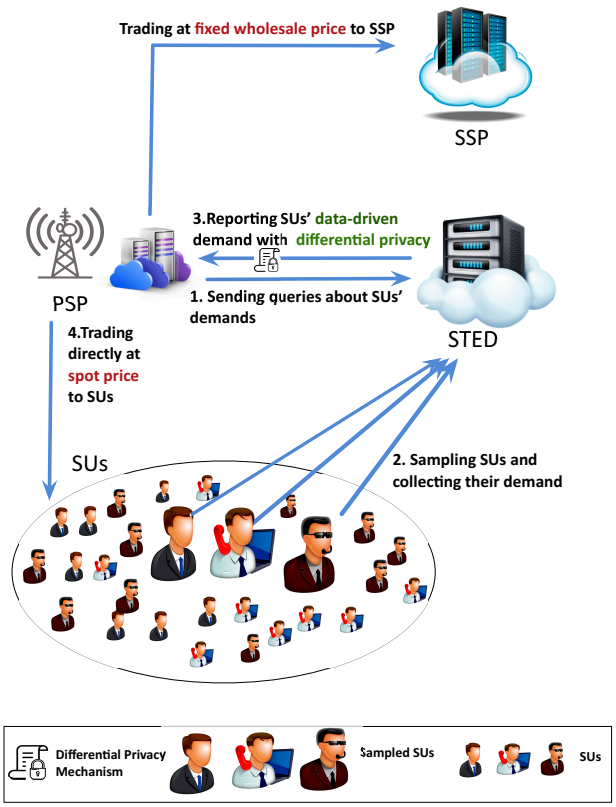


Fig. 2. The spectrum trading procedure of 3DPP.

for feasible solutions and verify the effectiveness of the proposed 3DPP by simulations.

II. System Description and 3DPP Outline

A. System Model and Adversary Model

Our proposed spectrum trading market consists of the PSP, the STED, the SSP, and $\mathcal{N} = \{1, 2, \cdots, i, \cdots, N\}$ SUs as shown in Fig. 2. As introduced in Sec. I, the PSP and the SSP are entities similar to MVNOs, and the STED is a trustworthy database server for SUs, which can collect the traffic demand information from SUs, and temporarily store it. The PSP is entitled to aggregate vacant spectrum resources from $\mathcal{M} = \{1, 2, \cdots, j, \cdots, M\}$ PUs with unequal sized bandwidth $\mathcal{W} = \{W_1, W_2, \cdots, W_j, \cdots, W_M\}$, and sell those available spectrum bands for monetary gains.

Similar to power market/cloud resource market in smart grid/cloud computing systems, the PSP has the following spectrum trading options: (i) selling available spectrum to the SSP at fixed wholesale price, i.e., c; (ii) selling available spectrum bands to the SUs directly at spot price, i.e., b; or (iii) dividing available spectrum resources and selling to both. Thus, before splitting the spectrum and deciding the selling strategy, the PSP will send queries about SUs' demands to the STED as shown in Fig. 1. Due to the large number of SUs within the PSP's coverage, the STED will sample some SUs, build up a reference traffic demand distribution of SUs, and respond to the PSP's queries.

The adversaries could be the dishonest PSP or eavesdropping attackers, who are always monitoring the information exchange between the PSP and the STED. As shown in Fig. 1, without enforcing any privacy preserving schemes, the adversaries can easily learn the sampled SUs' traffic demand profiles. That may help the adversaries make some illegal monetary gains, or even launch jamming attacks on some valuable services of chosen SUs. It also makes the SUs reluctant to participate in spectrum trading. The meaning of the notations are shown in TABLE I.

TABLE I
The List of Notations

Symbol	Definition
$\mathcal{N}$	Sets of SUs
$\mathcal{M}$	Sets of PUs
$\mathcal{W}$	Sets of unequal sized bandwidth
c	Fixed wholesale price that PSP sells to SSP
b	Spot price that PSP sells to STED
$\mathbb{P}$	Real traffic demand distribution of all SUs
$\mathbb{P}_0$	Reference distribution from sampled SUs.
$\mathbb{P}_0'$	Distribution after STED adds noise
ϵ	DP parameter
Δf	l_1 sensitivity of a function f in DP
γ_j	Binary variable to indicate if W_j is assigned to STED
ξ	Random variable of SUs' traffic demands
$\mathcal{D}$	Confidence set
η	Confidence level
d_ζ	Distribution distance under ζ-structure probability metric
θ	Tolerance of the distance between two distributions
V	Number of sampled SUs
Ω	The sample space of ξ
$\varnothing$	The dimension of Ω

B. 3DPP Outline

To preserve the sampled SUs' DP, it takes four steps for the PSP to sell the available spectrum to SUs at spot price b as shown in Fig. 1. Firstly, the PSP sends queries about SUs' demands to the STED. Secondly, STED samples some SUs, and constructs a reference demand distribution $\mathbb{P}_0$ from sampled SUs' demands. The STED needs to ensure the uncertainty distance between the reference distribution $\mathbb{P}_0$ and the real traffic demand distribution of all SUs $\mathbb{P}$ is close enough. Thirdly, the STED adds noises drawn from Laplace distribution to $\mathbb{P}_0$, and establishes a SUs' traffic demand reference distribution $\mathbb{P}_0'$, which achieves ϵ-DP. Meanwhile, the STED needs to guarantee that $\mathbb{P}_0'$ is close enough to $\mathbb{P}$, so that $\mathbb{P}_0'$ satisfies both data-driven and ϵ-DP requirements. Then, the STED responds to the PSP's queries with $\mathbb{P}_0'$. Finally, based on $\mathbb{P}_0'$, the PSP decides how much spectrum needs to be sold to the SUs directly at b, and how much spectrum need to be sold to the SSP at c to maximize its revenue.

Following this spectrum trading procedure, in the next section, we formulate the PSP's revenue maximization problem under data-driven and DP constraints, i.e., 3DPP. In Sec. IV, we theoretically prove that $\mathbb{P}_0'$ is close enough to $\mathbb{P}$, which means the proposed 3DPP has data-driven and DP properties. We also develop solutions to 3DPP problem in Sec. IV.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2892447, IEEE Transactions on Dependable and Secure Computing

III. 3DPP Problem Formulation

In this section, we first present the preliminaries on DP, and then formulate the PSP's revenue maximization problem under data-driven and DP constraints.

A. Preliminaries on Differential Privacy

DP is a formal definition of data privacy, which ensures that any sequence of output from data set (e.g., responses to queries) is "essentially" equally likely to occur, no matter any individual is present or absent [9], [13], [14]. DP keeps the characteristic of the whole data set, and preserves information privacy of each individual. The definition of ϵ-DP is as follows.

Definition 1 *Differential Privacy*: Let $\mathcal{A}$ denote a randomized algorithm. We take the output as $\mathbf{r}$ and the input as x, i.e., $\mathcal{A}(x) = \mathbf{r}$. For all $x, x' \subseteq \mathcal{N}^{|\mathcal{X}|}$ satisfied $||x - x'|| \leq 1$,

$$log\frac{Pr(\mathbf{r}|x)}{Pr(\mathbf{r}|x')} \leq \epsilon. \tag{1}$$

Then we call $\mathcal{A}$ is ϵ-DP. The parameter ϵ represents the degree of DP offered, which is the upper bond of the differences between true $\mathcal{A}(x)$ and $\mathcal{A}(x')$. A smaller value of ϵ implies the stronger privacy guarantee and perturbation noise, and a larger value of ϵ implies a weaker privacy guarantee while having higher data utility.

Definition 2 *l_1-sensitivity of a function f*: The l_1-sensitivity of a function $f : \mathbb{N}^{|\mathcal{X}|\rightarrow\mathbb{R}^k}$ is

$$\Delta f = \max_{\substack{x,x'\in\mathbb{N}^{|\mathcal{X}|}.\\ ||x-x'||_1=1}} ||f(x) - f(x')||_1. \tag{2}$$

The l_1 sensitivity of a function f captures the magnitude by which a single individuals data can change the function f in the worst case, and therefore, intuitively, the uncertainty in the response that we must introduce in order to hide the participation of a single individual [13].

Definition 3 *The Laplace Distribution and the Laplace Mechanism*: The Laplace Distribution (centered at 0) with scale b is the PDF (Probability Density of Function) is :

$$Lap(x|b) = \frac{1}{2b}\exp(-\frac{|x|}{b}). \tag{3}$$

In the following paper, we will write Lap(b) simply to denote a random variable $X \sim$ Lap(b). The $Laplace\ Mechanism$ simply computes f, and perturb each coordinate with noise drawn from the Laplace distribution. The scale of the noise will be calibrated to the sensitivity of f. The Laplace Mechanism $\mathcal{A}_L$ is defined as

$$\mathcal{A}_L(x, f(\cdot), \epsilon) = f(x) + (Y_1, \cdots, Y_k),$$

where Y_i are i.i.d random variables drawn from Lap($\Delta f/\epsilon$). The proof of Laplace mechanism reserves ϵ-DP is shown in [13].

B. PSP's Revenue Maximization Formulation

Let γ_j be a binary variable indicating if W_j is directly sold to SUs, where γ_j = 1 if W_j is directly sold to SUs, and 0, otherwise. Thus, the PSP's revenue gained from selling spectrum to the SSP can be written as $\sum_{j=1}^{M} cW_j(1-\gamma_j)$, where $1-\gamma_j$ represents the spectrum sold to the SSP at fixed price c. Besides, let random variable ξ denote the uncertain demands from all SUs, and ξ follows distribution $\mathbb{P}$. Then, $b\Big(\min\big(\sum_{j=1}^{M} W_j\gamma_j, \xi\big)\Big)$ is the PSP's revenue gained from selling spectrum to SUs directly[2]. Here, due to the uncertainty of SUs' demands, if the spectrum supply from the PSP (i.e., the spectrum bands that the PSP decided to sell to SUs directly) is more than SUs' actual total traffic demand, i.e., $\sum_{j=1}^{M} W_j\gamma_j > \xi$, the revenue for the PSP is $b\xi$. Otherwise, if the spectrum supply from the PSP is less than SUs' actual traffic demand, i.e. $\sum_{j=1}^{M} W_j\gamma_j < \xi$, the revenue for the PSP is $b\sum_{j=1}^{M} W_j\gamma_j$.

Putting those two parts together, the PSP's revenue maximization can be formulated as follows.

$$\max_{\gamma} \quad -\sum_{j=1}^{M} cW_j\gamma_j + \sum_{j=1}^{M} cW_j + b\mathbb{E}_{\mathbb{P}}\Big(\min\big(\sum_{j=1}^{M} W_j\gamma_j, \xi\big)\Big), \tag{4}$$

$$\text{s.t.:} \quad \gamma_j \in \{0,1\}, j = 1, \cdots, M, \tag{5}$$

$$\xi = \sum_{i=1}^{N} d_i, i = 1, \cdots, N, \tag{6}$$

where γ_j is binary variable, and (6) represents the total traffic demand of all SUs.

C. Data-Driven Based PSP's Revenue Optimization

Given the huge number of SUs within PSP's coverage, the STED cannot collect traffic demand information from every possible SU, i.e., the STED is generally difficult to obtain the true probability distribution of all SUs' demand $\mathbb{P}$. Instead, we allow the STED to collect the traffic demands from a series of sampled SUs, and construct reference demand distribution $\mathbb{P}_0$. For a given set of sampled SU data, it is easy for us to construct a histogram to fit the SUs' traffic demand. For example, we can set N intervals to fit the total traffic demand of sampled SUs in each interval to be $L_1, \cdots, L_n, \cdots, L_N$ with $L = \sum_{n=1}^{N} L_n$. For instance, L_1 is the number of SUs who would like to access spectrum on price \$15/MHZ, L_2 is the number of SUs who would like to access spectrum on price \$20/MHZ, etc.. Based on this, we can construct an reference distribution for the uncertain total traffic demand of all consumers in particular time period of a day as $p_1^0 = L_1/L, \cdots, p_n^0 = L_n/L, \cdots,$ and $p_N^0 = L_N/L$. For simplicity, we let $\mathbb{P}_0 = p_1^0, p_2^0, \cdots, p_N^0$ represent the corresponding reference distribution. Since $\mathbb{P}_0$ may not be 100% represents the unique true SUs' demand distribution $\mathbb{P}$, we employ risk-averse stochastic optimization approaches (RA-SP) allowing distribution ambiguity [15] to reformulate the PSP's revenue maximization problem in (4). Instead of deriving a true distribution for ξ, this optimization approach derives a confidence set D, and allows the distribution ambiguity to be within set $\mathcal{D}$ with a certain confidence

[2]In this paper, we assume the aggregated spectrum resources can be perfectly split to satisfy SUs' traffic demands.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2892447, IEEE Transactions on Dependable and Secure Computing

level (e.g., 99%). The data-driven based RA-SP for the PSP's revenue maximization is formulated as follows.

$$\max_{\gamma} \quad -\sum_{j=1}^{M} cW_j\gamma_j + \sum_{j=1}^{M} cW_j + \min_{\mathbb{P}\in\mathcal{D}} b\mathbb{E}_{\mathbb{P}}\Big(\min\big(\sum_{j=1}^{M} W_j\gamma_j, \xi\big)\Big), \tag{7}$$

s.t.: constraints (5) and (6).

We use a distribution distance measurement proposed in [16], [17] to quantify the distance of distributions. Specifically, a predefined distance measure $d(\mathbb{P}_0, \mathbb{P})$ is constructed on confidence set $\mathcal{D}$, where $\mathbb{P}$ is the true distribution and $\mathbb{P}_0$ is the ambiguous distribution conducted from sampled SUs. The distance d_ζ and confidence set $\mathcal{D}$ can be defined as follows,

$$\mathcal{D} = \{\mathbb{P} : d_\zeta(\mathbb{P}_0, \mathbb{P}) \le \theta\}, \tag{8}$$

$$d_\zeta(\mathbb{P}_0, \mathbb{P}) = \sup_{h\in\mathcal{H}} \left| \int_\Omega h d\mathbb{P}_0 - \int_\Omega h d\mathbb{P} \right|. \tag{9}$$

Here, $d_\zeta(\cdot,\cdot)$ represents the distance under ζ structure probability metric, θ denotes the tolerance, and $\mathcal{H}$ is a family of real-valued bounded measurable functions on Ω (the sample space on ξ). Tolerance θ is correlated to data size, i.e., the number of SUs' demand samples. It can be easily inferred that the more demand samples that the STED can collect, the tighter $\mathcal{D}$ would be, and the closer ambiguous distribution $\mathbb{P}_0$ would be to $\mathbb{P}$. More details of ζ-structure probability metric is introduced in the next Section.

D. 3DPP: Data-Driven Based PSP's Revenue Optimization under ϵ-DP

To protect the sampled SUs' traffic demand profiles, the STED will employ Laplace mechanism to add noises into $\mathbb{P}_0$. Here, we denote $\mathbb{P}_0'$ as the distribution after employing Laplace mechanism, and p_0' as its density of probability function accordingly. According to the definition of ϵ-DP, we have $p_0' \le p_0 e^\epsilon$. Thus, the data-driven based PSP's revenue maximization under ϵ-DP, i.e., 3DPP problem, can be reformulated as follows.

$$\max_{\gamma} \quad -\sum_{j=1}^{M} cW_j\gamma_j + \sum_{j=1}^{M} cW_j + \min_{\mathbb{P}\in\mathcal{D}'} b\mathbb{E}_{\mathbb{P}}\Big(\min\big(\sum_{j=1}^{M} W_j\gamma_j, \xi\big)\Big), \tag{10}$$

s.t.: (5), (6)

$$\mathcal{D}' = \{\mathbb{P} : d_\zeta(\mathbb{P}_0', \mathbb{P}) \le \theta\}, \tag{11}$$

$$d_\zeta(\mathbb{P}_0', \mathbb{P}) = \sup_{h\in\mathcal{H}} \left| \int_\Omega h d\mathbb{P}_0' - \int_\Omega h d\mathbb{P} \right|. \tag{12}$$

IV. 3DPP Proof and Solutions

This section is organized as follows. First, we present how to determine converge rate under ζ-structure probability structure. We show the relation between DP parameter ϵ and distribution tolerance θ in ζ-structure probability structure, and prove our DP mechanism satisfies the requirement of data-driven, which is $d_\zeta(\mathbb{P}_0', \mathbb{P}) \le \theta$. Second, we reformulate the problem under ζ-structure probability metrics, and convert it to a traditional two-stage robust optimization. We develop algorithms to solve the problem w.r.t. different probability metrics.

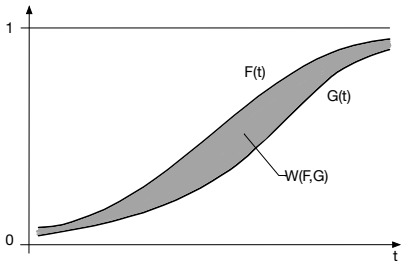


Fig. 3. Wasserstein metrics (one-dimensional case).

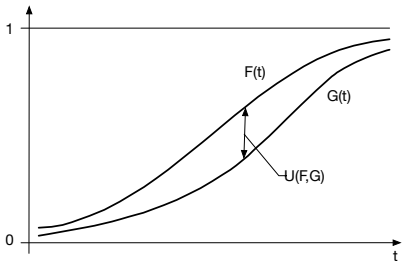


Fig. 4. Uniform metric

A. Converge Rate under ζ-structure Probability Metrics

As described in Sec.II, we employ three different ζ-structure probability metrics and solve our problem under these constraints correspondingly. We define $\rho(x, y)$ as the distance between two variables x and y, and $\varnothing$ as the dimension of Ω. $\mathbb{P} = \mathcal{L}(x)$ represents random variables x follows distribution $\mathbb{P}$. We denote V as the size of sampled SUs. The metrics are shown as follows.

- **Kantorovich metric**: denoted as $d_K(\mathbb{P}, \mathbb{Q})$, $\mathcal{H} = \{h : ||h||_L \le 1\}$, where $||h||_L := \sup\{h(x) - h(y)/\rho(x, y){:}x \ne y \text{ in } \Omega\}$. By the Kantorovich-Rubinstein theorem. the Kantorovich metric is equivalent to the Wasserstein metric. In particular, when $\Omega = R$, let d_w denote the Wasserstein metric, then

$$d_w(\mathbb{P}, \mathbb{Q}) = \int_{-\infty}^{+\infty} |F(x) - G(x)| dx, \tag{13}$$

  where F and G are the distribution function derived from $\mathbb{P}_i^0$ and $\mathbb{P}_i$ respectively. It is illustrated in Fig 3.
- **Fortet-Mourier metric**: denoted as $d_{FM}(\mathbb{P}, \mathbb{Q})$, $\mathcal{H} = \{h : ||h||_C \le 1\}$, where $||h||_C := \sup\{h(x) - h(y)/c(x, y){:}\ x \ne y \text{ in } \Omega\}$ and $c(x, y) = \rho(x, y) max\{1, \rho(x, a)^{p-1}, \rho(y, a)^{p-1}\}$ for some $p \ge 1$ and $a \in \Omega$. Note that when $p = 1$, Fortet-Mourier metric is the same as Kantorovich metric. The Fortet-mourier metric is usually utilized as a generalization of Kantorovich metric, with the application on mass transportation problems.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2892447, IEEE Transactions on Dependable and Secure Computing

- **Uniform metric**: denoted as $d_U(\mathbb{P},\mathbb{Q})$, $\mathcal{H} = \{I_{(-\infty,t]}, t \in R^n\}$. According to the definition, we have $d_U(\mathbb{P},\mathbb{Q}) = \sup_t |\mathbb{P}^0(x \le t), \mathbb{Q}(x \le t)|$. It is illustrated in Fig 4, where F and G are the distribution function derived from $\mathbb{P}$ and $\mathbb{Q}$, respectively.

From the definition of metrics above, we can derive the convergence property and convergence rate accordingly. By utilizing Dvoretzky-Kiefer-Wolfowitz inequality, the convergence rate of $d_K(\mathbb{P}_0,\mathbb{P})$ is shown as follows [18].

Proposition 1 For a general dimension case (i.e., n=1),

$$\mathbb{P}(d_K(\mathbb{P}_0,\mathbb{P}) \le \theta) \ge 1 - \exp\left(-\frac{\theta^2 V}{2}\right). \tag{14}$$

For a high dimension case (i.e., n > 1), and any $\alpha > 0$, there is a constant C_α that could satisfies

$$\mathbb{P}(d_K(\mathbb{P}_0,\mathbb{P}) \le \theta) \ge 1 - C_\alpha \exp\left(-\frac{\theta^2 V}{2}\right). \tag{15}$$

Then we prove the converge rate between distribution with Laplace mechanism $\mathbb{P}_0'$ and real distribution $\mathbb{P}$ under Kantorovich metric as follows.

Proposition 2 For a general dimension case (i.e., n≥1),

$$\mathbb{P}(d_K(\mathbb{P}_0',\mathbb{P}) \le \theta) \ge 1 - \exp\left(-\frac{\theta^2 V}{2\varnothing^2} - \epsilon\right). \tag{16}$$

Proof: Let us define a set

$$\mathcal{B} := \{\mu \in \mathcal{P}(\Omega) : d_k(\mu,\mathbb{P}) \ge \theta\}, \tag{17}$$

where $\mathcal{P}(\Omega)$ is the set of all probability measures defined on Ω. Let $\mathcal{C}(\Omega)$ be the set of bounded continuous function $\phi \to R$. Therefore, following the definitions, for each $\phi \in \mathcal{C}(\Omega)$, we have

$$\mathbb{P}\big(d_K(\mathbb{P}_0',\mathbb{P}) \ge \theta\big) = Pr(\mathbb{P}_0' \in \mathcal{B}) \tag{18}$$

$$\le Pr\left(\int_\Omega \phi d\mathbb{P}_0' \ge \inf_{\mu\in\mathcal{B}} \int_\Omega \phi d\mu\right) \tag{19}$$

$$\le \exp\left(-V \inf_{\mu\in\mathcal{B}} \int_\Omega \phi d\mu\right) E\left(e^{V\int_\Omega \phi d\mathbb{P}_0 e^\epsilon}\right) \tag{20}$$

$$= \exp\left(-V \inf_{\mu\in\mathcal{B}} \left\{\int_\Omega \phi d\mu - \frac{1}{V}\log E\left(e^{V\int_\Omega \phi d\mathbb{P}_0 e^\epsilon}\right)\right\}\right)$$

$$= \exp\left(-V \inf_{\mu\in\mathcal{B}} \left\{\int_\Omega \phi d\mu - \frac{1}{V}\log E\left(e^{\sum_{i=1}^V e^\epsilon \phi(\xi^i)}\right)\right\}\right) \tag{21}$$

$$= \exp\left(-V \inf_{\mu\in\mathcal{B}} \left\{\int_\Omega \phi d\mu - \log \int_\Omega e^\epsilon e^\phi d\mathbb{P}\right\}\right) \tag{22}$$

$$= \exp\left(-V \inf_{\mu\in\mathcal{B}} \left\{\int_\Omega \phi d\mu - \log \int_\Omega e^\phi d\mathbb{P} - \epsilon\right\}\right), \tag{23}$$

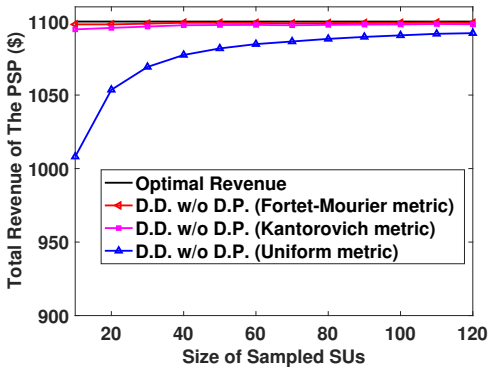
where (18) follows the definition of $\mathcal{B}$, inequality (19) is from the fact that $\mathbb{P}_0 \in \mathcal{B}$, and μ is the one distribution in $\mathcal{B}$ that satisfies the minimum of $\int_\Omega \phi d\mu$,(20) follows from the Chebyshev's exponential inequality [19], and (21) follows from the definition of $\mathbb{P}_0$.

Now we define $\Delta(\mu) := \sup_{\phi\in\mathcal{C}(\Omega)} \int_\Omega \phi d\mu - \log \int_\Omega e^\phi d\mathbb{P}$. Thus, following the definition of $\mathcal{C}(\Omega)$, there exists a series ϕ_n such that $\lim_{n\to\infty} \int_\Omega \phi d\mu - \log \int_\Omega e^\phi d\mathbb{P} = \Delta(\mu)$. For any small positive number $\theta' > 0$, there exists a constant number n_0 such that $\Delta(\mu) - (\int_\Omega \phi_n d\mu - \log \int_\Omega e_n^\phi d\mathbb{P}) \le \theta'$ for any $n \ge n_0$. Therefore, according to (23), we use substitute ϕ_n for ϕ, then we have

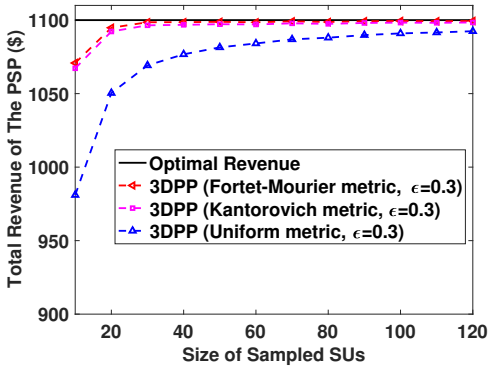
$$Pr(\mathbb{P}_0' \in \mathcal{B})$$

$$\le \exp\left(-V \inf_{\mu\in\mathcal{B}} \left\{\int_\Omega \phi d\mu - \log \int_\Omega e^\phi d\mathbb{P} - \epsilon\right\}\right) \tag{24}$$

$$\le \exp\left(-V \inf_{\mu\in\mathcal{B}} \{\Delta(\mu) - \epsilon - \theta'\}\right) \tag{25}$$



(a) Data-Driven spectrum trading without ϵ-DP.



(b) 3DPP.

Fig. 5. Total revenue of PSP under different probability distance metrics.

Algorithm 1 Algorithm1: Procedure of Solving $3DPP$

1: **Input:** Historical data ξ_1,ξ_2,$\cdots$,ξ_N from sample SUs. Set ϵ as the privacy parameter. Set η as the confidence level of D.
2: **Out:** Objective value of the η.
3: STED receives the number of sampled SUs under different traffic demand, i.e.,ξ_1, $\cdots$, ξ_N.
4: STED adds Laplace noise to the original data set of sample SUs. ξ_n'=ξ_n+$(Y_1,\cdots,Y_k)$, where Y_i are i.i.d random variables drawn from Lap($\Delta f/\epsilon$).
5: STED reports the processed data ξ_n' to PSP.
6: Obtain the reference distribution $\mathbb{P}_0'(\xi)$ and tolerance θ based on the data received from STED.
7: STED uses the reformulation (SP-M) or (SP-U) to solve the problem.
8: Output the solution.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2892447, IEEE Transactions on Dependable and Secure Computing

According to Lemma 6.2.13 in [20], we have

$$\Delta(\mu) = d_{KL}\left(\mu, \mathbb{P}\right) \tag{26}$$

where $d_{KL}\left(\mu, \mathbb{P}\right)$ is the discrete case KL-divergence defined as $\sum_i ln(p_i/\mu_i)p_i$. For the case $\mu \in \mathcal{B}$, with (17), we have $d_K\left(\mu, \mathbb{P}\right) \geq \theta$. Moreover, in "Particular case 5" in [21], we have

$$d_K\left(\mu, \mathbb{P}\right) \leq \varnothing\sqrt{2d_{KL}\left(\mu, \mathbb{P}\right)} \tag{27}$$

hold for $\forall \mu \in \mathcal{P}(\Omega)$. Consequently, following (27), we have

$$d_{KL}\left(\mu, \mathbb{P}\right) \geq \theta^2/\left(2\varnothing^2\right). \tag{28}$$

Combining (25), (26), (28), we have

$$Pr(\mathbb{P}_0' \in \mathcal{B}) \leq \exp\left(-V\left(\frac{\theta^2}{2\varnothing^2} - \epsilon - \theta'\right)\right). \tag{29}$$

Let $\theta' = \lambda/V$ for any arbitrary small positive λ. Then, we have

$$\begin{aligned} &P\left(d_k\left(\mathbb{P}_0', \mathbb{P}\right) \geq \theta\right) \\ &= Pr\left(\mathbb{P}_0' \in \mathcal{B}\right) \leq \exp\left(-V\left(\frac{\theta^2}{2\varnothing^2} - \epsilon\right) + \lambda\right). \end{aligned} \tag{30}$$

Since λ can be arbitrarily small, we have $\mathbb{P}(d_k(\mathbb{P}_0', \mathbb{P} \leq \theta)) \geq 1 - \exp(-\frac{\theta^2}{2\varnothing^2}V + V\epsilon)$.

With convergence rate (30), we can calculate the tolerance θ accordingly. For instance, in Kantorovich metric, we assume the confidence level is η. Therefore $\mathbb{P}(d_u(\mathbb{P}_0, \mathbb{P} \leq \theta)) \geq 1 - \exp(-\frac{\theta^2}{2\varnothing^2}V + V\epsilon) = \eta$ according to (30), and $\theta = \varnothing\sqrt{2log(e^{\epsilon V}/(1-\eta))/V}$. ■

Similar proof is applicable for other metrics. For example, following the proof procedure of **Proposition 2** in our work and using **Corollary 1** in [15], it is easy to prove that under Fortet-Mourier metric, we have

$$\mathbb{P}(d_{FM}(\mathbb{P}_0', \mathbb{P}) \leq \theta) \geq 1 - \exp\Big(-\frac{\theta^2 V}{2\varnothing^2\Lambda^2} + \epsilon V\Big), \tag{31}$$

where $\Lambda = \max\{1, \varnothing^{p-1}\}$. Due to the page limits, we omit the detailed proof procedure.

B. Problem Reformulation under ζ-Probability Metrics, and Solutions

We denote $x = \sum_{j=1}^{M} W_j\gamma_j$, $\alpha = \sum_{j=1}^{M} W_j$ where α is a constant. The sample space is $\Omega = \{\xi_1, \xi_2, \cdots, \xi_N\}$. Then the formulation can be simplified as

$$\max_{x} \quad -cx + \min_{p_i} b\sum_{i=1}^{N} p_i\Big(\min\big(x, \xi_i\big)\Big) + c\alpha \tag{32}$$

$$\text{s.t.} \quad x \in [0, \alpha], \tag{33}$$

$$\sum_i p_i = 1, \tag{34}$$

$$\max \sum_{i=1}^{N} h_i p_{0_i}' - \sum_{i=1}^{N} h_i p_i \leq \theta, \forall h_i : ||h||_\zeta \leq 1, \tag{35}$$

where the $|h||_\zeta$ is defined according to different metric. In Kantorovich metric, $|h_x - h_y| \leq \rho(\zeta^x, \zeta^y)$. The constraint (34), (35) can be summarized as $\sum_i a_{il}h_i \leq b_{il}, l = 1, \cdots, L$. To reformulate the constraint, we consider the problem

$$\min_{h_i} \quad \sum_{i=1}^{N} h_i p_{0_i}' - \sum_{i=1}^{N} h_i p_i, \tag{36}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} a_{il}h_i \leq b_{il}, l = 1, \cdots, L. \tag{37}$$

Its dual problem is represented as

$$\min \quad \sum_{l=1}^{L} b_l u_l, \tag{38}$$

$$\text{s.t.} \quad \sum_{l=1}^{l} a_{il}u_l \geq p_{0_i}' - p_i, \forall i = 1, \cdots, N, \tag{39}$$

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

$$\max_{x} \quad -cx + \min_{p_i} b\sum_{i=1}^{N} p_i\Big(\min\big(x, \xi_i\big)\Big) + c\alpha, \tag{40}$$

$$\text{(SP-M)} \quad \text{s.t.} \quad x \in [0, \alpha], \tag{41}$$

$$\sum_{i=1}^{N} p_i = 1, \sum_{l=1}^{l} b_l u_l \leq \theta, \tag{42}$$

$$\sum_{l=1}^{L} a_{il}u_l \geq p_{0_i}' - p_i, \forall i = 1, \cdots, N. \tag{43}$$

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

$$\max_{x} \quad -cx + \min_{p_i} b\sum_{i=1}^{N} p_i\Big(\min\big(x, \xi_i\big)\Big) + c\alpha \tag{44}$$

$$\text{(SP-U)} \quad \text{s.t.} \quad x \in [0, \alpha], \tag{45}$$

$$\sum_{i=1}^{N} p_i = 1, \tag{46}$$

$$\left|\sum_{i=1}^{l}\left(p_{0_i}' - p_i\right)\right| \leq \theta, \forall l = 1, \cdots, L. \tag{47}$$

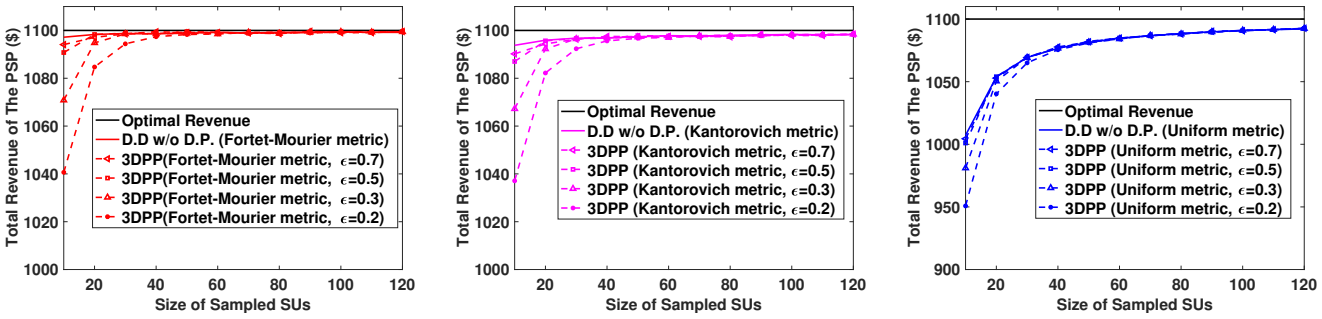
The formulation SP-M and SP-U can be solved by L-shape algorithm which is described in [22]. We summarize the procedure of solving the 3DPP problem in Alg. 1.

V. Performance Evaluation

A. Simulation Setup

For illustrative purposes, we consider a spectrum trading market with 500 SUs. We assume the true traffic demand of all SUs follows a discrete distribution: 100M with probability 0.4 and 200M with probability 0.6, respectively. Total available spectrum resources aggregated by the PSP is 300M. In addition, we set the fixed wholesale price for the spectrum sold to the SSP to be \$ 3/MHz, and the spot price for the spectrum sold directly to SUs to be \$ 5/MHz.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2892447, IEEE Transactions on Dependable and Secure Computing



(a) Fortet-Mourier metric (b) Kantorovich metric (c) Uniform metric

Fig. 6. Total revenue of the PSP with 3DPP under different ϵ-DP parameters.

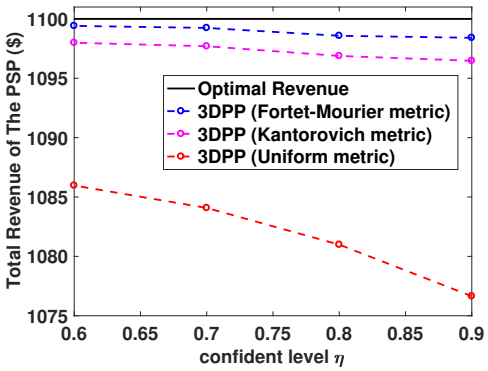


Fig. 7. Total revenue of the PSP with 3DPP under different confidence levels.

B. Privacy and Performance Analysis

First, the confidence level η is set to be 90% and the size of sampled SUs varies from 10 to 120. We study the data-driven algorithm without DP. The results are shown in Fig. 5(a). After collecting traffic demand of sample SUs, the STED does not add Laplace noises, and submits the true reference distribution directly to the PSP. From the results in Fig. 5(a), it can be observed the total revenue of the PSP increases when the size of sample SUs increases, regardless of the distance metrics adopted. The intuition behind the result is that, as the size of sampled SUs , the value θ decreases, which stands for the distance between true distribution and reference distribution. As a result, the solutions are moving closer to the optimal one. It is also shown in Fig. 5(a) that the gap between total revenue under the Fortet-Mourier metric and the Kantorovich metric is very small, when the number of sampled SUs is over 100. When the number of sampled SU is 120, the results under all metrics are close to the optimal one. Besides, we study the 3DPP's performance in Fig. 5(b). Compared with results in Fig. 5(a), it can be observed that the total revenue of the PSP with 3DPP is less than that without ϵ-DP when the number of sampled SUs is small, but becomes close to each other, or even to the optimal revenue when the size of sampled SUs increases. That means it incurs some cost to involve ϵ-DP for the sampled SUs' traffic demands, especially when the number of sampled SUs is small. But this impact significantly diminishes when the number of samples increases. That also implies that the proposed 3DPP scheme can still successfully captures the characteristics of whole data set, i.e., the demand distribution of all SUs, while preserving individual sampled SU's traffic profile privacy. Moreover, from Fig. 5, we found the Fortet-Mourier metric is a more applicable metric, since the simulation results is more closer to the optimal revenue.

Moreover, we explore the impact of DP parameter ϵ in Fig. 6. We choose four different ϵ values, i.e., 0.7, 0.5, 0.3, 0.2, respectively, and study its impact under different metrics. We find that as the ϵ decreases, the total revenue of PSP decreases under all metrics. The reason is, ϵ stands for the upper bound of privacy loss. It means, when ϵ is smaller, the mechanism yields better privacy, and less accurate responses which leads to less revenue of the PSP. It also can be observed that, when size of sampled SUs is less than 60, the gaps of total revenue under different ϵ is large. When size of sampled SUs increases, the influence of ϵ is less, and the total revenue under 3DPP with different ϵ converges to the optimal one. Last but not the least, we study the effect of confidence level on the 3DPP in Fig. 7. We set the number of sampled SUs as 40, and test four different confidence levels, i.e., 0.6, 0.7, 0.8, 0.9, respectively. From the Fig. 7 we can observe that, as the confidence level increases, the gaps between the PSP's revenue of 3DPP and optimal one increases under all three metrics. The reason is that, as the confidence level η increases, the distance θ between reference distribution with ϵ-DP $\mathbb{P}_0'$ and true distribution $\mathbb{P}$ increases, and the true probability distribution of SUs traffic demands is more likely to be in the confidence set $\mathcal{D}$. That implies that distribution in set $\mathcal{D}$ which is not that close to $\mathbb{P}$ might be used to yield solutions. Therefore, the PSP's revenue performance degrades when confidence level increases.

VI. Related Work

There are a lot of research works focusing on preserving privacy during spectrum trading. To be specific, Errapotu et al. in [23] employ the Paillier crypto-system to preserve SUs' bidding privacy and maximize the revenue of PU simultaneously in a semi-distributed manner. Liu et al. leverage attribute-based encryption to preserve PUs' operational privacy in spectrum database. Recently, a promising mechanism, *differential privacy* (DP), proposed by Dwork [11] has been employed in dynamic spectrum allocation [9], [12], [24]. DP aims to reveal statistical information of whole dataset without compromising the privacy of each individual. Zhu et al. in [12] preserve the bidders valuation privacy with approximate revenue maximization in spectrum auction mechanism, and

theoretically proved the mechanism is differential private. In the area of internet of things and spectrum monitoring, Sun et al. in [24] propose a distributed stream monitoring system with high communication efficiency and privacy guarantee. The technique they proposed is powered by DP theory, which can ensure submitted data of every node are not substantially different with one element of the node's data stream changes. Jin et al. in [9] present a crowdsourced spectrum sensing service provider, which selects spectrum-sensing participants in a DP preserving manner. They prove the new mechanism can prevent any internal or external attackers from learning the location of mobile participants, and minimize the social cost simultaneously.

To process spectrum trading, PU service provider recruits SUs to collect their characteristic (traffic demand, location, etc.), and allocate different quantity of bandwidths to different SUs accordingly. Since the number of mobile devices increases dramatically (the mobile devices are expected to hit 12.1 billion in 2018), it is unrealistic to recruit all mobiles in a specified region. Thus, we present a new architecture with data-driven. In our work, STED samples a relatively smaller scale of SUs to collect the information of SUs' traffic demand and sends to PSP. However, since the number of sample is limited, it is difficult for PSP to learn the precise information of SUs' traffic demands. Hence, we utilize the data-driven approach to deal with uncertainty of the information. Some previous researchers have noticed the issue of distribution uncertainty and tried to employ robust optimization to address this issue. For instance, Lunden et al. [25] propose a non-parametric cyclic correlation in robust computation, which lead the algorithm doesn't require the distribution of users' traffic. Gong et al. in [26] present a model, which consider the distribution uncertainty of received primary signal in spectrum sensing, to determine the robust threshold that can guarantee the false alarm uncertainty. However, there is a lack of study to incorporate data-driven sensing and DP together in spectrum trading system. In our paper, we are trying to melt SUs' traffic demand DP into data-driven based spectrum trading. With the proposed scheme, our work effectively preserves each individual SU's traffic demand and maximizes revenue of PSP under data-driven scheme at the same time.

VII. Conclusion

In this paper, we propose a novel spectrum trading architecture consisting of the PSP, the SSP and the STED. Under this architecture, we proposed a novel 3DPP spectrum trading scheme, which jointly employs DP techniques to preserve SUs' demand, and data-driven approach to characterize the uncertainty of SUs' traffic demand. Moreover we mathematically prove that the data after employing DP mechanism satisfies the data-driven requirements under different ζ-structure probability metrics. Based on the contribution above, we formulate a RA-SP problem to maximize revenue of the PSP. We employ a confidence set by ζ-structure metric to reformulate the problem to a traditional two-stage robust optimization, and developed algorithms. Through simulations, we show the feasible solutions and verify the effectiveness of the proposed 3DPP scheme.

References

[1] I. Akyildiz, W. Lee, M. Vuran, and M. Shantidev, "Next generation/ dynamic spectrum access/ cognitive radio wireless networks: a survey," *Computer Networks (Elsevier) Journal*, vol. 50, no. 4, pp. 2127–2159, September 2006.

[2] IEEE 802.22-2011(TM) Standard for Cognitive Wireless Regional Area Networks (RAN) for Operation in TV Bands, July 2011.

[3] FCC, "Spectrum policy task force report," Report of Federal Communications Commission, Et docket No. 02-135, November 2002.

[4] M. Pan, P. Li, Y. Song, Y. Fang, P. Lin, and S. Glisic, "When spectrum meets clouds: Optimal session based spectrum trading under spectrum uncertainty," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 3, pp. 615–627, March 2014.

[5] L. Duan, J. Huang, and B. Shou, "Cognitive mobile virtual network operator: Investment and pricing with supply uncertainty," in *Proc. of IEEE Conference on Computer Communications, INFOCOM 2010*, San Diego, CA, March 2010.

[6] X. Li, H. Ding, M. Pan, Y. Sun, and Y. Fang, "Users first: Service-oriented spectrum auction with a two-tier framework support," *IEEE Journal on Selected Areas in Communications: Spectrum Sharing and Aggregation for Future Wireless Networks*, vol. 34, no. 11, pp. 2999–3013, November 2016.

[7] Q. Huang, Y. Tao, and F. Wu, "Spring: A strategy-proof and privacy preserving spectrum auction mechanism," in *Proceeding of IEEE International Conference on Computer Communications (INFOCOM)*, Turian, Italy, April 2013.

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

[9] X. Jin and Y. Zhang, "Privacy-preserving crowdsourced spectrum sensing," in *Proceeding of the IEEE International Conference on Computer Communications (INFOCOM)*, 2016, pp. 1–9.

[10] M. Li, P. Li, L. Guo, and X. Huang, "PPER: Privacy-preserving economic-robust spectrum auction in wireless networks," in *IEEE Conference on Computer Communications (INFOCOM)*, Kowloon, April 2015, pp. 909–917.

[11] C. Dwork, "Differential privacy: A survey of results," in *International Conference on Theory and Applications of Models of Computation*. Springer, 2008, pp. 1–19.

[12] R. Zhu, Z. Li, F. Wu, K. Shin, and G. Chen, "Differentially private spectrum auction with approximate revenue maximization," in *Proceedings of ACM international symposium on mobile ad hoc networking and computing, ACM MobiHoc*, 2014, pp. 185–194.

[13] C. Dwork, A. Roth *et al.*, "The algorithmic foundations of differential privacy," *Foundations and Trends in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[14] A. Friedman, I. Sharfman, D. Keren, and A. Schuster, "Privacy-preserving distributed stream monitoring." in *NDSS*, San Diego, CA, Feb 2014.

[15] C. Zhao and Y. Guan, "Data-driven risk-averse two-stage stochastic program with ζ-structure probability metrics," *Available on Optimization Online*, 2015.

[16] G. C. Calafiore, "Ambiguous risk measures and optimal robust portfolios," *SIAM Journal on Optimization*, vol. 18, no. 3, pp. 853–877, October 2007.

[17] D. Klabjan, D. Simchi-Levi, and M. Song, "Robust stochastic lot-sizing by means of histograms," *Production and Operations Management*, vol. 22, no. 3, pp. 691–710, February 2013.

[18] A. Dvoretzky, J. Kiefer, and J. Wolfowitz, "Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator," *The Annals of Mathematical Statistics*, pp. 642–669, 1956.

[19] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods*. Methuen London, 1964.

[20] U. SCHMOCK, "Large deviations techniques and applications," *Journal of the American Statistical Association*, no. 452, pp. 1380–1380, 2000.

[21] F. Bolley and C. Villani, "Weighted csiszar-kullback-pinsker inequalities and applications to transportation inequalities," *Annales de la Facult e des Sciences de Toulouse*, vol. 14, pp. 331–352, 2005.

[22] C. zhao and Y. Yuan, "Data-driven stochastic unit commitment for integrating wind generation," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2587–2596, 2016.

[23] S. M. Errapotu, J. Wang, Z. Lu, W. Li, M. Pan, and Z. Han, "Bidding privacy preservation for dynamic matching based spectrum trading," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM'16)*, Washington, D.C., USA, December 2016.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2892447, IEEE Transactions on Dependable and Secure Computing

[24] J. Sun, R. Zhang, J. Zhang, and Y. Zhang, "Pristream: Privacy-preserving distributed stream monitoring of thresholded percentile statistics," in *Proceeding of the IEEE International Conference on Computer Communications (INFOCOM)*, 2016, pp. 1–9.

[25] J. Lundén, S. A. Kassam, and V. Koivunen, "Robust nonparametric cyclic correlation-based spectrum sensing for cognitive radio," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 38–52, January 2010.

[26] S. Gong, P. Wang, and W. Liu, "Spectrum sensing under distribution uncertainty in cognitive radio networks," in *IEEE International Conference on Communications (ICC)*, Ottawa, ON, USA, June 2012.

Jingyi Wang received her B.S. degree in Physics from Nankai University, China, in 2012 and M.S. degree in electrical and computer engineering from Auburn University, Auburn, AL, in 2015. She has been working towards her Ph.D. degree in the Department of Electrical and Computer Engineering at University of Houston, Houston, TX, since August 2015. Her research interests include big data analysis, cyber-physical systems and cybersecurity. She is a student member of IEEE.

Xinyue Zhang received the B.E. degree in communication engineering from Beijing Jiaotong University, China, in 2016, and the B.Sc. degree in electronic engineering from KU Leuven, Belgium, in 2016. She is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Houston. She has been a Research Assistant with the Cognitive Radio Networking, Cybersecurity, and Cyber-Physical System Laboratory since 2017. Her research interests include cognitive radio networks and wireless security. She is a student member of IEEE.

Qixun Zhang received his B.S., and Ph.D. degrees from Beijing University of Posts and Telecommunications (BUPT), China, in 2006 and 2011. He is an associate professor with the Key Laboratory of Universal Wireless Communications, Ministry of Education, China. He is a member of IEEE and active in the ITU-R WP5A/5C/5D, IEEE 1900, CCSA, and IMT-2020(5G) standards. From Mar. to Jun. 2006, he was a visiting scholar at the University of Maryland, College Park, Maryland. Currently, he is a visiting scholar in the Electrical and Computer Engineering Department at the University of Houston, Texas. His research interests include 5th generation mobile networks (5G), cognitive radio and heterogeneous networks, game theory, LAA and LTE-U system, sensing and communication integrated system design, and unmanned aerial vehicles (UAVs) communication.

Ming Li received the BE degree in electrical engineering from Sun Yat-sen University, China, in 2007, the ME degree in electrical engineering from the Beijing University of Posts and Communications, China, in 2010, and the PhD degree in electrical and computer engineering from Mississippi State University, Starkville, in 2014, respectively. She is currently an assistant professor in the Department of Computer Science and Engineering, the University of Texas at Arlington. Her research interests include usable security, privacy-preserving computing, crowd and social sensing, Internet of things, and mobile computing.

Yuanxiong Guo received the B.Eng. degree in electronics and information engineering from Huazhong University of Science and Technology, Wuhan, China, in 2009, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Florida, Gainesville, FL, USA, in 2012 and 2014, respectively. Since 2014, he has been an Assistant Professor with the School of Electrical and Computer Engineering at Oklahoma State University, Stillwater, OK, USA. His current research interests include cybersecurity, data analytics, and resource management for networked systems including Internet of things, cyber-physical systems, and cloud/edge systems. He is a recipient of the Best Paper Award in the IEEE Global Communications Conference 2011. He has been serving as an Editor of IEEE Transactions on Vehicular Technology since January 2018.

Zhiyong Feng received her B.S., M.S., and Ph.D. degrees from Beijing University of Posts and Telecommunications (BUPT), Beijing, China. She is a professor at BUPT, and the director of the Key Laboratory of the Universal Wireless Communications, Ministry of Education, P.R.China. She is a senior member of IEEE, vice chair of the Information and Communication Test Committee of the Chinese Institute of Communications (CIC). Currently, she is serving as Associate Editors-in- Chief for China Communications, and she is a technological advisor for international forum on NGMN. Her main research interests include wireless network architecture design and radio resource management in 5th generation mobile networks (5G), spectrum sensing and dynamic spectrum management in cognitive wireless networks, and universal signal detection and identication.

Miao Pan received his BSc degree in Electrical Engineering from Dalian University of Technology, China, in 2004, MASc degree in electrical and computer engineering from Beijing University of Posts and Telecommunications, China, in 2007 and Ph.D. degree in Electrical and Computer Engineering from the University of Florida in 2012, respectively. He is now an Assistant Professor in the Department of Electrical and Computer Engineering at University of Houston. He was a recipient of NSF CAREER Award in 2014. His research interests include big data privacy, cybersecurity, cyber-physical systems, and cognitive radio networking. His work won Best Paper Awards in VTC 2018, Globecom 2017 and Globecom 2015, respectively. Dr. Pan is an Associate Editor for IEEE Internet of Things (IoT) Journal from 2015 to 2018. He is a member of ACM and a senior member of IEEE.