

Incentivizing Spectrum Sensing in Database-Driven Dynamic Spectrum Sharing

Bo Gao*, Sudeep Bhattarai†, Jung-Min “Jerry” Park†, Yaling Yang†, Min Liu*, Kexiong Zeng†, and Yanzhi Dou†

*Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

Email: {gaobo, liumin}@ict.ac.cn

†Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg VA, USA

Email: {sbhattar, jungmin, yyang8, kexiong6, yzdou}@vt.edu

Abstract—The legacy concept of exclusion zones (EZs) is inept at enabling efficient utilization of fallow spectrum by secondary users (SUs), since legacy EZs are static and overly-conservative. The notion of a static EZ implies that it has to protect incumbent users (IUs) from the union of likely interference scenarios, leading to a worst-case, conservative solution. In this paper, we propose the concept of dynamic, multi-tier EZs, which takes advantage of participatory spectrum sensing carried out by SUs to support efficient database-driven spectrum sharing while protecting IUs against SU-induced aggregate interference. Specifically, the database directly incentivizes SUs to participate in spectrum sensing, which augments geolocation database by defining smaller EZs with dynamic boundaries and creating additional spectrum access opportunities for SUs. We propose an incentive mechanism based on a two-level game-theoretic model, in which the database conducts dynamic pricing in a first-level Stackelberg game in the presence of SUs who strategically contribute to spectrum sensing in a second-level stochastic game. The existence of an equilibrium solution is proven. According to our findings, the proposed incentive mechanism for the concept of dynamic, multi-tier EZs is effective to improve spectrum utilization efficiency while guaranteeing incumbent protection.

I. INTRODUCTION

In database-driven dynamic spectrum sharing [1][2][3], each secondary user (SU) queries a geolocation database to identify fallow spectrum that is not being occupied by incumbent users (IUs). Fallow spectrum in TV bands is often referred to as “white spaces”, but we will use this term in this paper to denote fallow spectrum in arbitrary bands. A licensed channel serves as a white space channel for the requesting SU only if the SU is not located in the *exclusion zones* (EZs) for IUs operating on this channel. Hence, the accessible area outside the EZs, which is the service coverage of this channel as a white space channel, is directly determined by the regulation of EZs. The boundary of a legacy EZ is static, and it is calculated by the database based on incumbent registrations and operating characteristics, interference protection requirements, and numerical results from signal propagation models. The static regulation of an EZ is defined in such a way that the IUs are protected from the union of all possible interference scenarios, and thus can easily result in a worst-case, conservative solution for incumbent protection. However, overly-conservative EZs

cause high probability of false alarms (i.e., erroneous detection of IUs when they are absent) to keep low probability of false dismissals (i.e., missed detection of IUs when they are present), and thus secondary spectrum access opportunities have to be unnecessarily wasted. For example, the static EZs estimated in a NTIA report [4] to protect radars in 3.5 GHz band preclude approximately 60% of the U.S. population from accessing fallow spectrum, which would severely undermine the economic incentive for realizing spectrum sharing.

Recently, a number of studies have attempted to address the problems associated with the legacy concept of EZs. Generally, the boundaries of legacy EZs can be refined by adding additional details to a conventional propagation model. Due to the dynamic nature of spectrum sharing, incorporating real-time information on spectrum availability is much more effective in lowering the probability of false alarms than considering static terrain data [5] or average aggregate interference to IUs [5][6]. In fact, the information gathered from *spectrum sensing* offers what a geolocation database lacks—real-time identification of spectrum access opportunities that would have been overlooked otherwise. The FCC has accepted the use of spectrum sensing to define smaller EZs in 3.5 GHz band [7]. In particular, spectrum sensing is invaluable in understanding the interference environment when constructing an accurate propagation model is infeasible, e.g., in indoor environments [8]. Experimental results from recent studies [9][10][11] have shown that the databases alone may offer inaccurate and stale information of spectrum availability. However, none of the past studies has prescribed a specific approach that integrates spectrum sensing into a methodology for dynamically refining the boundaries of EZs to enhance spectrum utilization efficiency.

In this paper, we focus on the concept of dynamic, multi-tier EZs, which incorporates *participatory spectrum sensing* carried out by SUs into geolocation database. Particularly, we focus on the core problem of incentivizing spectrum sensing to enable efficient database-driven spectrum sharing. We propose a pragmatic strategy—incentivizing sensing-capable SUs to report spectrum sensing results to the database, in return for creating extra spectrum access opportunities for SUs in legacy EZs. Our approach integrates this incentive mechanism with the concept of dynamic, multi-tier EZs.

To achieve this goal, we tackle the following challenges. First, our incentive mechanism has to jointly consider why, where, and how spectrum sensing is incentivized. In fact, spectrum sensing is not always worthy of incentivization. Existing

This work was supported in part by the National Natural Science Foundation of China (NSFC) under grants 61120106008, 61132001, 61202410, 61272474, 61472402, 61472404, 61502457, 61572476; and by the National Science Foundation (NSF) under grants 1054697, 1228903, 1265886, 1314598, 1431244, 1547241; and by the industry affiliates of the Broadband Wireless Access & Applications Center (BWAC) and the Wireless@Virginia Tech.

proposals of single-tier EZs [7] and multi-tier EZs [5][6] do not offer a specific solution to effectively utilize spectrum sensing, especially participatory sensing. Second, our incentive mechanism has to meet the needs of dynamic zoning. To promptly and reliably capture spectrum access opportunities, the timeliness and fidelity of the sensing results to be collected need to be of sufficient quality. Existing incentive mechanisms [12] are not sufficient to meet these requirements. Third, our incentive mechanism has to ensure incumbent protection when white space capacity is used as incentives. Note that each SU serves as both a contributor and a consumer of local white space capacity. Intuitively, the SUs who have contributed to participatory sensing can be paid with spectrum access grants inside an EZ (originally prohibited), but the IUs in this EZ should be protected against the aggregate interference caused by the paid co-channel SUs. Existing incentive mechanisms do not address this issue involving incumbent protection.

In this paper, we make the contributions listed as follows.

- We propose the concept of dynamic, multi-tier EZs through a dynamic zoning framework, in which the geolocation database incentivizes participatory spectrum sensing to create secondary spectrum access opportunities in legacy EZs.
- We propose an incentive mechanism based on a two-level game-theoretic model, in which the database conducts dynamic pricing in a first-level Stackelberg game in the presence of SUs who strategically contribute to spectrum sensing in a second-level stochastic game. This defines smaller EZs with dynamic boundaries and creates spectrum access opportunities outside the revised EZs (but inside legacy EZs).
- We propose the use of white space capacity as incentives based on an aggregate interference analysis. This creates spectrum access opportunities inside the revised EZs.
- We prove the convergence of our approach to an equilibrium solution, which supports efficient database-driven spectrum sharing while guaranteeing incumbent protection.

The remainder of this paper is organized as follows. Related work is discussed in section II. Our dynamic zoning framework is introduced in section III. Our incentive mechanism is outlined in section IV. Our two-level game-theoretic model is proposed in section V. Our aggregate interference analysis is given in section VI. Simulation results are presented in section VII. Conclusions are summarized in section VIII.

II. RELATED WORK

To improve the single-tier EZs, recent studies have proposed multi-tier EZs. In [5], multi-tier EZs are defined based on differential spectrum access hierarchy to allow SUs with a higher interference tolerance threshold to operate closer to IUs. In [6], multi-tier EZs are defined based on average aggregate interference analysis to allow a number of SUs to harmlessly operate closer to IUs. However, timely and accurate capture of dynamic spectrum availability requires real-time spectrum sensing rather than average interference modelling.

Recently, the FCC has accepted the use of spectrum sensing to define smaller EZs in 3.5 GHz band [7]. However, only limited existing studies have attempted to take advantage of spectrum sensing to augment database-driven white space discovery. In [9], an experimental testbed for identifying white

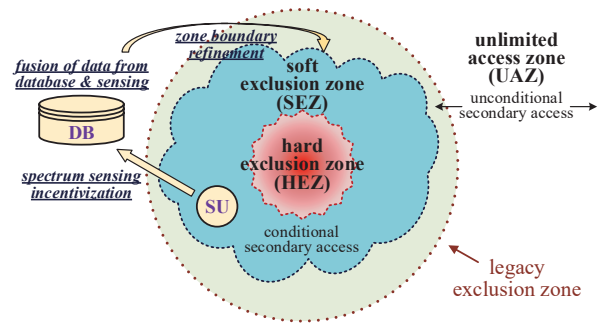


Fig. 1. An example of a dynamic, multi-tier EZ.

spaces in TV bands has been proposed. Spectrum sensing has been leveraged here to verify the available channels given by the database. However, the verification of database does not improve spectral efficiency. Besides this, spectrum sensing in our framework can find extra white space channels from the unavailable channels assumed by the database. In [10], an outdoor spectrum sensing experiment has been conducted, in which a bus-carried spectrum analyzer has been used to scan TV bands for weeks. The database that has wasted white spaces over a wide area can be corrected offline by the collected dataset of long-term sensing results. However, the database in our framework can be augmented by participatory sensing results on the fly. In [11], another outdoor experiment has been conducted. All the locations can be empirically grouped into either the locations where the database is accurate or the locations where spectrum sensing is preferred. However, sensing-required locations in our framework can be selected on the fly as directed by varying spectrum sensing demands.

To collect spectrum sensing results without relying on the deployment of dedicated sensors, the issue of incentivizing spectrum sensing needs to be addressed. Incentive mechanisms have been studied in economics and mobile crowdsourcing [12]. A typical incentive mechanism is based on a principal-agent model. In general, the negotiations between the principal and the agents are usually studied in a game [13][14][15] or an auction [15][16][17][18]. Due to the leading role of the database in database-driven spectrum sharing, the platform-centric model in [15] is most appropriate for our incentive mechanism. In this model, the principal applies a pricing strategy to lead the agents to make contributions. However, existing incentive mechanisms including the one-time pricing strategy in [15] cannot directly meet the timeliness and quality requirements of spectrum sensing for dynamic zoning.

III. DYNAMIC ZONING FRAMEWORK

In this section, we propose the concept of dynamic, multi-tier EZs through a dynamic zoning framework, which answers why, where, and how spectrum sensing is incentivized.

The regulation of EZs is the *ex-ante* spectrum enforcement technique used by the regulators to protect IUs against harmful interference from SUs. For a licensed channel, each legacy EZ is defined as a monolithic spatial separation region around the IUs operating on this channel, where co-channel and/or adjacent-channel secondary transmissions are prohibited.

To enlarge the service coverage of a certain channel as a white space channel or, equivalently, to create secondary

spectrum access opportunities in a legacy EZ of this channel, spectrum sensing can be performed inside the legacy EZ. We define a dynamic, multi-tier EZ, a part of which can potentially benefit from spectrum sensing. As an example given in Fig. 1, the local map for the legacy EZ that protects IUs located at the center is partitioned into three zones as follows.

- In *hard exclusion zone* (HEZ), secondary channel access is always prohibited, since co-channel SUs have very high chance of interfering with neighboring IUs;
- In *soft exclusion zone* (SEZ), secondary channel access is possible (originally prohibited), but SUs have to obtain explicit permission from the database prior to access by providing spectrum sensing results to the database, which discover local channel availability;
- In *unlimited access zone* (UAZ), secondary channel access is always allowed, since co-channel SUs have very low chance of interfering with faraway IUs.

In this three-zone framework, the SUs have incentives to perform spectrum sensing in SEZ only. The sensing results reported by SUs in SEZ can be used to iteratively refine the inner and outer boundaries of SEZ. Initially, the boundaries of SEZ are defined by conventional propagation modelling based on the desired rates of false dismissals and false alarms. In each iteration of dynamic zoning, three problems should be solved: a) *spectrum sensing incentivization* (SSI), b) *fusion of data from database & sensing* (FDS), and c) *zone boundary refinement* (ZBR). The database first solves the SSI problem to collect sensing results in desired amount and quality from the current SEZ. The database then solves the FDS problem to fuse the results of propagation modelling (stored in the database) and spectrum sensing to update a spectrum availability map. Based on this map, the database solves the ZBR problem to refine c-i) the inner boundary of SEZ by reallocating the locations where secondary access is always harmful to IUs from SEZ to HEZ, and c-ii) the outer boundary of SEZ by reallocating the locations where secondary access is always harmless to IUs from SEZ to UAZ. The revised SEZ is used as a baseline SEZ in the next iteration.

In this paper, we mainly focus on the SSI problem, which plays a key role in our framework. The FDS problem is similar to the construction of a radio environment map [10][19][20], which involves statistical fusion of multiple sources of incumbent characteristics. The ZBR problem is equivalent to the reallocation of sensed locations from SEZ to either HEZ or UAZ to save the cost of spectrum sensing, which is avoided at the locations that have been reallocated to HEZ or UAZ.

IV. SPECTRUM SENSING INCENTIVIZATION

In this section, we provide an overview of our incentive mechanism, which is based on a Stackelberg leadership model [15] due to the leading role of the database in dynamic zoning. The database (as the leading principal) announces spectrum sensing tasks and a *pricing strategy*. Each task is to sense a certain channel at a certain location in the SEZ of this channel. In response to the database, each SU (as the following agent) applies a *contributing strategy* to either take or ignore the tasks mapped to the locations in its vicinity. The SU who takes a task once contributes a spectrum sensing report to the database and then receives a payment for the contribution.

A. White Space Capacity as Incentives

To incentivize selfish SUs to willingly and truthfully make contributions, the payments as incentives in our model can be in the form of *white space capacity* instead of popular yet impractical monetary credits [12]. This is because each SU in SEZ serves as both a contributor and a consumer of white space capacity. In this way, any loss of transmission opportunities at the cost of performing participatory sensing can be compensated by a payment of new spectrum access opportunities. Besides supporting white space discovery, our incentive mechanism also addresses white space allocation in SEZ through making payments to contributing SUs.

Our incentive mechanism enhances spectrum utilization efficiency in two aspects. On the one hand, the collection of sensing results enables the reallocation of always-harmless locations from SEZ to UAZ, and thus allows unconditional secondary access in extended UAZ (green area in Fig. 1). On the other hand, the use of white space capacity as incentives allows conditional secondary access in SEZ. Overall, secondary spectrum access opportunities are created in the legacy EZ.

B. Pricing Strategy vs. Contributing Strategy

In our model, the database competes with the SUs in a Stackelberg game. On the one hand, the database incentivizes spectrum sensing to create spectrum access opportunities indirectly in extended UAZ and directly in SEZ. Because of the limited white space capacity as payment budgets for sensing tasks in SEZ, the database aims to minimize the total payment for each sensing task so that limited budgets are efficiently utilized to gather maximum good-quality sensing reports and accommodate maximum contributing SUs. On the other hand, each SU aims to maximize the received payment from the database for the same cost of performing spectrum sensing.

1) *Dynamic Pricing as Leading Strategy*: More advanced than the one-time pricing strategy in [15], the pricing strategy of the database in our model is dynamic to control the progress of sensing task completion. To ensure the timeliness and quality of spectrum sensing for dynamic zoning, each sensing task, which is the collection of sensing reports that can be integrated to reach a target for cumulative sensing reliability, must be completed before a deadline for sensing participation. Intuitively, the collection of sensing reports towards the target by the deadline can be compared to the sale of perishable items without inventory replenishment [21][22]. A more profitable pricing strategy for such a problem is preferred.

In our model, the database takes a dynamic pricing strategy [21][22][23][24], which balances the database's supply for sensing tasks and the SUs' demands for sensing participation better than the one-time pricing strategy. For over-demand case, on the one hand, the payments can be reduced without harming the task progress. For over-supply case, on the other hand, the payments have to be raised to attract more contributions and accelerate the task progress before the deadline. Our dynamic pricing strategy jointly considers limited white space supply, time-varying sensing participation demands, and sensing task deadline to support dynamic zoning.

To clearly define the progress of sensing task completion for dynamic pricing, the quality of each contribution should be quantified. However, it is hard to judge the significance of

crowd-sensed data due to the absence of trustworthy authoritative benchmarks [18]. Similar to the quality value in [18], the contribution of each sensing report here is evaluated by a *usefulness* score, which is an overall evaluation in several aspects, including accuracy, relevance, and trustworthiness. For example, a sensing report has a higher usefulness score if its contributor has better sensing capability, is located closer to the targeted location, or has higher reputation score. The overall usefulness score should be a weighted sum of the evaluation scores for the above aspects. The weight values can dynamically change with the progress of sensing task completion. For example, a higher weight can be assigned to trustworthiness in the beginning and relevance in the end. The progress of sensing task completion is characterized by the cumulative usefulness score from the contributions heretofore.

2) *Strategic Contributing as Following Strategy*: Given the pricing strategy of the database, the contributing strategy of each SU in our model can be derived accordingly. In response to dynamic pricing, a contributing SU can be either myopic or strategic [23]. Specifically, a myopic SU arrives to submit a sensing report according to a stochastic process [24], while a strategic SU times the moment of submitting a sensing report according to a game-theoretic model [22].

In our model, we focus on dynamic pricing in the presence of strategic SUs, whose contributing strategies are not known a priori. Our model can also cover the myopic case.

V. TWO-LEVEL GAME-THEORETIC MODEL

In this section, we extend the Stackelberg game addressing the SSI problem to a two-level game-theoretic model. The first-level Stackelberg game describes the vertical competition between the database and the set of SUs, and a second-level stochastic game describes the horizontal competition within the set of SUs. It takes a number of rounds to alternately refine the pair of pricing and contributing strategies in one iteration of dynamic zoning. In the first round, we derive the pricing strategy of the database in a Markov decision process (MDP) under myopic assumption (perfect knowledge). Given the pricing strategy, we derive the contributing strategies of the SUs in the stochastic game under strategic assumption (incomplete knowledge). In return, the contributing strategies are used to refine the pricing strategy in the next round, and so forth. In the end, our incentive mechanism should make the database lead the second-level stochastic game towards an equilibrium solution to the first-level Stackelberg game.

A. Database's Pricing Strategy

1) *Problem Formulation*: From the view of the database, suppose that there are a set of licensed channels \mathcal{K} , and there are a set of location pixels \mathcal{L}_k in the SEZ of each channel $k \in \mathcal{K}$. Making a contribution to a certain sensing task (k, l) is to submit a sensing report that describes the availability of channel $k \in \mathcal{K}$ at location $l \in \mathcal{L}_k$. Among the SUs registered with the database, there are a set of sensing-capable ones, $\mathcal{N}^{(k,l)}$, who are qualified to contribute to (k, l) .

The period for dynamic pricing in one iteration of dynamic zoning is discretized into a set of time slots \mathcal{T} . The database adjusts the announced payment (action) to control the cumulative usefulness score (state) towards the target \hat{S} by the

deadline \hat{T} . Hence, the database can solve a finite-horizon MDP [25] to obtain the optimal pricing strategy for each sensing task, say (k, l) . For time slot $t = 1, 2, \dots, \hat{T}$, the MDP is defined by a 5-tuple $\langle \mathcal{S}, \mathcal{P}, \{Q_{p_t}(\cdot, \cdot)\}, \{R_{p_t}(\cdot, \cdot)\}, \{\bar{R}(\cdot)\} \rangle$:

- \mathcal{S} is a finite set of states including all possible values of cumulative usefulness score, and $s_t \in \mathcal{S}$ is the cumulative usefulness score from contributions up to time slot t ;
- \mathcal{P} is a finite set of actions for database including all possible values of announced payment, and $p_t \in \mathcal{P}$ is the announced payment for contributions in time slot t ;
- $Q_{p_t}(s_{t-1}, s_t) = Pr\{s_t | s_{t-1}, p_t\}$ is a state transition probability, i.e., the probability that the announcement of p_t leads to the state transition from s_{t-1} to s_t ;
- $R_{p_t}(s_{t-1}, s_t)$ is an immediate reward for database, if the state transits from s_{t-1} to s_t after the announcement of p_t ;
- $\bar{R}(s_{\hat{T}})$ is a terminal reward for database, if the state ends up with $s_{\hat{T}}$ in time slot \hat{T} .

In each time slot t , the database determines the payment p_t to trigger the state transition from s_{t-1} to s_t . Upon the reception of p_t , the SUs respond according to their contributing strategies, and the database accepts a set of contributing SUs, $\mathcal{N}_t \subseteq \mathcal{N}^{(k,l)}$, who submit their sensing reports. As assumed in [15], the database accepts all the contributions before the deadline, and p_t is shared by the contributing SUs. Let \mathcal{U} be a set of usefulness scores, and each SU $n \in \mathcal{N}_t$ contributes a usefulness score $u_t^n \in \mathcal{U}$. Hence, the state transits through

$$s_t = s_{t-1} + \tilde{u}_t, \quad (1)$$

where \tilde{u}_t is the sum of usefulness scores from contributions in time slot t . Namely, we define

$$\tilde{u}_t = \sum_{n \in \mathcal{N}_t} u_t^n. \quad (2)$$

Other definitions of \tilde{u}_t are possible depending on the winner selection strategy of the database.

The state transition probabilities $Q_{p_t}(s_{t-1}, s_t)$ under p_t are defined by either user arrival patterns in a stochastic process (if myopic) or user best responses in a game (if strategic). Each SU n comes to contribute with an arrival probability ω_t^n , which is a function of p_t . For a certain s_t , we have

$$\begin{aligned} Q_{p_t}(s_{t-1}, s_t) &= Pr\{\tilde{u}_t = s_t - s_{t-1} | s_{t-1}, p_t, s_t\} \\ &= \sum_{\mathfrak{N}_t \in \mathfrak{N}^{(s_t - s_{t-1})}} \prod_{n \in \mathcal{N}_t} \omega_t^n \prod_{m \in (\mathcal{N}^{(k,l)} - \mathcal{N}_t)} (1 - \omega_t^m), \end{aligned} \quad (3)$$

where $\mathfrak{N}^{(s_t - s_{t-1})}$ is a collection of all possible sets \mathcal{N}_t that satisfy $\sum_{n \in \mathcal{N}_t} u_t^n = s_t - s_{t-1}$. For myopic case, each ω_t^n can directly be defined as a non-decreasing function of p_t as the empirical demand functions in dynamic pricing [24]. For strategic case, which will be studied later, each ω_t^n can be learnt from the best response of SU n in a game.

The immediate rewards $R_{p_t}(s_{t-1}, s_t)$ under p_t and the terminal rewards $\bar{R}(s_{\hat{T}})$ are used to minimize the total payment $\tilde{p} = \sum_{t=1}^{\hat{T}} p_t$ while guaranteeing the completion of sensing task, i.e., $s_{\hat{T}} \geq \hat{S}$. Minimizing \tilde{p} for reaching \hat{S} is equivalent to maximizing the cumulative usefulness score achieved by unit payment, i.e., $\frac{\hat{S}}{\tilde{p}}$. A heuristic definition of $R_{p_t}(s_{t-1}, s_t)$ favors the pricing strategy that maximizes the

usefulness-payment ratio in each time slot. We have

$$R_{p_t}(s_{t-1}, s_t) = \frac{\tilde{u}_t}{p_t}. \quad (4)$$

Each $\bar{R}(s_{\hat{T}})$ is a non-positive reward as a punishment for the unfulfillment of sensing task, i.e., $s_{\hat{T}} < \hat{S}$. To ensure $s_{\hat{T}} \geq \hat{S}$, let $\bar{R}(s_{\hat{T}})$ decrease with the lag of $s_{\hat{T}}$ behind \hat{S} . We have

$$\bar{R}(s_{\hat{T}}) = \begin{cases} -(\hat{S} - s_{\hat{T}})^\xi & \text{if } s_{\hat{T}} < \hat{S}; \\ 0 & \text{if } s_{\hat{T}} \geq \hat{S}. \end{cases} \quad (5)$$

In this definition, if $s_{\hat{T}} < \hat{S}$, the negative $\bar{R}(s_{\hat{T}})$ drops with $\hat{S} - s_{\hat{T}}$, and ξ is a parameter to control the punishment value. Hence, the database has to make up the lag (if exists) before the deadline to avoid unnecessary negative terminal reward.

2) *Optimal Solution:* In the above MDP, a pricing strategy $\pi : \mathcal{S} \times \mathcal{T} \rightarrow \mathcal{D}(\mathcal{P})$ is a mapping from the progress of sensing task completion $\mathcal{S} \times \mathcal{T}$ to a probability distribution $\mathcal{D}(\mathcal{P})$ over pricing actions \mathcal{P} . Then, a mixed strategy π includes $\varphi_{p_t} \in \mathcal{D}(\mathcal{P})$, which is the probability of taking each pure strategy $p_t \in \mathcal{P}$ at any state $s_{t-1} \in \mathcal{S}$, $t \in \mathcal{T}$. We have

$$\varphi_{p_t} = \pi(s_{t-1}, p_t). \quad (6)$$

In this finite-horizon MDP, the expected total reward is computed to evaluate the pricing strategy of the database. Under a certain π , if the initial state is at s_0 , the expected total reward from $t = 1$ to \hat{T} can be defined by

$$V^\pi = \mathbb{E}^\pi \left[\sum_{t=1}^{\hat{T}} R_{p_t}(s_{t-1}, s_t) + \bar{R}(s_{\hat{T}}) \right]. \quad (7)$$

A pricing strategy π^* is optimal with respect to the expected total reward criterion if it is true that

$$V^{\pi^*} = \sup_{\pi} V^\pi. \quad (8)$$

Several algorithms have been proposed in the literature that solve this standard finite-horizon MDP offline [25].

However, the user arrival probabilities ω_t^n in (3) are initially assumed a priori to derive the pricing strategy of the database. For dynamic pricing in the presence of strategic SUs, we need to relax this assumption and specifically determine the values of ω_t^n as the contributing strategies of the SUs, which further refine the pricing strategy in the next round.

B. User's Contributing Strategy

1) *Problem Formulation:* We turn our focus to the second-level game played by strategic SUs, in which we aim to find an equilibrium solution of user best responses such that it can further be utilized by the database to refine its pricing strategy.

Given a pricing strategy π of the database, each SU times the moment of submitting a sensing report to the sensing task (k, l) in order to receive the maximum expected payment. Due to the multi-stage nature of dynamic pricing, the SUs in $\mathcal{N}^{(k, l)}$ can play a finite-horizon stochastic game [22][26][27]. For time slot $t = 1, 2, \dots, \hat{T}$, the game is defined by a 5-tuple $\langle \mathcal{N}^{(k, l)}, \mathcal{S} \times \mathcal{P}, \mathcal{A}, \{q(\cdot, \cdot)\}, \{r(\cdot)\} \rangle$:

- $\mathcal{N}^{(k, l)}$ is a finite set of players, and each SU $n \in \mathcal{N}^{(k, l)}$ strategically behaves in response to the given π ;

- $\mathcal{S} \times \mathcal{P}$ is a finite set of states, and $(s_{t-1}, p_t) \in \mathcal{S} \times \mathcal{P}$ includes s_{t-1} before the actions of SUs in time slot t and p_t after the action of database in time slot t ;
- $\mathcal{A} = \{0, 1\}$ is a finite set of actions for each SU n , and $a_t^n = 1(0) \in \mathcal{A}$ represents SU n submits (does not submit) a sensing report in time slot t , and $\mathbf{a}_t = \{a_t^n \mid n \in \mathcal{N}^{(k, l)}\}$;
- $q_{\mathbf{a}_t}((s_{t-1}, p_t), (s_t, p_{t+1})) = Pr\{(s_t, p_{t+1}) \mid (s_{t-1}, p_t), \mathbf{a}_t\}$ is a state transition probability, i.e., the probability that the response of \mathbf{a}_t to database leads to the state transition from (s_{t-1}, p_t) to (s_t, p_{t+1}) ;
- $r_{\mathbf{a}_t}^n(s_{t-1}, p_t)$ is an immediate payoff for each SU n , after the response of \mathbf{a}_t to database at the state (s_{t-1}, p_t) .

In each time slot t , each SU n observes the state (s_{t-1}, p_t) and takes an action $a_t^n = 0/1$ (wait/take). The SUs take actions simultaneously and independently. After the database receives the response of \mathbf{a}_t to the announcement of p_t , the state transits from (s_{t-1}, p_t) to (s_t, p_{t+1}) . The resulting s_t is uniquely determined by (1) and the rewritten (2) that

$$\tilde{u}_t = \sum_{n \in \mathcal{N}^{(k, l)}} a_t^n u_t^n, \quad (9)$$

and the resulting p_{t+1} is probabilistically determined by (6).

The state transition probabilities $q_{\mathbf{a}_t}((s_{t-1}, p_t), (s_t, p_{t+1}))$ under \mathbf{a}_t are defined by π . For any (s_t, p_{t+1}) , we have

$$\begin{aligned} q_{\mathbf{a}_t}((s_{t-1}, p_t), (s_t, p_{t+1})) \\ = Pr\{\tilde{u}_t = s_t - s_{t-1} \mid s_{t-1}, \mathbf{a}_t, s_t\} \varphi_{p_{t+1}} \\ = \mathbf{1}_{\{\tilde{u}_t = s_t - s_{t-1}\}} \pi(s_t, p_{t+1}), \end{aligned} \quad (10)$$

where $\mathbf{1}_{\{\cdot\}}$ is an indicator function that is equal to 1 (0) if the condition inside $\{\cdot\}$ is true (false).

The immediate payoffs $r_{\mathbf{a}_t}^n(s_{t-1}, p_t)$ under \mathbf{a}_t are defined by the sharing of announced payment p_t among the contributing SUs in $\mathcal{N}_t = \{n \mid a_t^n = 1\}$. Each SU $n \in \mathcal{N}_t$ only gets a fraction of p_t as its received payment, denoted by p_t^n , which is further deducted by its cost c_t^n . As assumed in [15], for example, each contributing SU obtains a received payment proportional to its usefulness contribution. Then, we have

$$r_{\mathbf{a}_t}^n(s_{t-1}, p_t) = p_t^n - c_t^n = \frac{a_t^n u_t^n}{\tilde{u}_t} p_t - c_t^n. \quad (11)$$

In this definition, SU n will not receive any positive payoff after the deadline. The definition of c_t^n will be discussed later with regard to system convergence.

2) *Equilibrium Solution:* In the above stochastic game, a contributing strategy $\theta^n : \mathcal{S} \times \mathcal{P} \times \mathcal{T} \rightarrow \mathcal{D}(\mathcal{A})$ of each SU n is a mapping from the progress of sensing task completion along with the pricing strategy $\mathcal{S} \times \mathcal{P} \times \mathcal{T}$ to a probability distribution $\mathcal{D}(\mathcal{A})$ over contributing actions \mathcal{A} . Then, a Markov strategy θ^n depends on the current state and time [26], and includes $\omega_t^n \in \mathcal{D}(\mathcal{A})$, which is the arrival probability of SU n at any state $(s_{t-1}, p_t) \in \mathcal{S} \times \mathcal{P}$, $t \in \mathcal{T}$. We have

$$\omega_t^n = \theta^n(s_{t-1}, p_t). \quad (12)$$

In this finite-horizon stochastic game, the database finds a solution of user best responses, i.e., $\theta = \{\theta^n \mid n \in \mathcal{N}^{(k, l)}\}$, to refine the values of ω_t^n in (3). The expected total discounted payoff is computed to evaluate the contributing strategy of each SU. Under a certain $\theta = \{\theta^n, \theta^{-n}\}$, if the initial state is

at (s_0, p_1) , the expected total discounted payoff for each SU n from $t = 1$ to T can be defined by

$$v^{\{\theta^n, \theta^{-n}\}} = \mathbb{E}\{\theta^n, \theta^{-n}\} \left[\sum_{t=1}^{\hat{T}} \delta^{t-1} r_{\mathbf{a}_t}^n(s_{t-1}, p_t) \right], \quad (13)$$

where $\delta \in (0, 1)$ is a discount factor. A contributing strategy θ^{n*} of SU n is greedy with respect to the expected total discounted payoff criterion if it is true that

$$v^{\{\theta^{n*}, \theta^{-n}\}} = \sup_{\theta^n} v^{\{\theta^n, \theta^{-n}\}}. \quad (14)$$

In the following, we prove the existence of an equilibrium solution $\theta^* = \{\theta^{n*} \mid n \in \mathcal{N}^{(k,l)}\}$ in the stochastic game. Before that, we prove the existence of an equilibrium solution $\omega_t^* = \{\omega_t^{n*} = \theta^{n*}(s_{t-1}, p_t) \mid n \in \mathcal{N}^{(k,l)}, (s_{t-1}, p_t) \in \mathcal{S} \times \mathcal{P}, t \in \mathcal{T}\}$ in each stage-game of the stochastic game.

Lemma 1 There exists a Nash equilibrium ω_t^* in each stage-game of the second-level stochastic game (with observable actions) for a given state $(s_{t-1}, p_t) \in \mathcal{S} \times \mathcal{P}$, $t \in \mathcal{T}$, if $c_t = \sum_{n \in \mathcal{N}^{(k,l)}} c_t^n$ is increasing as the stage-game evolves.

Proof: Each stage-game is a normal-form game, in which each SU $n \in \mathcal{N}^{(k,l)}$ facing the state (s_{t-1}, p_t) chooses its arrival probability ω_t^n to maximize its immediate payoff $r_{\mathbf{a}_t}^n$ (and future discounted payoff). We assume that each SU n can contribute to the task (k, l) for multiple times, so its actions before t do not directly affect its current decision. What SU n only cares is to maximize its received payoff. The convergence of the stage-game to a Nash equilibrium ω_t^* depends on the definition of the cost c_t^n in (11). Besides the cost of performing spectrum sensing, the database as the central controller can deduct an amount of payment as a cost or punishment from the received payment p_t^n if SU n deviates from a Nash equilibrium. We define

$$r_{\mathbf{a}_t}(s_{t-1}, p_t) = \sum_{n \in \mathcal{N}^{(k,l)}} (p_t^n - c_t^n) = p_t - c_t. \quad (15)$$

The convergence of the stage-game can be proven through the Lyapunov's direct stability theorem as utilized in [28][29]. If c_t is increasing as the stage-game evolves, it is straightforward to show that $r_{\mathbf{a}_t} - \check{r}_{\mathbf{a}_t}$ is a Lyapunov function, where $\check{r}_{\mathbf{a}_t}$ is the lower bound of $r_{\mathbf{a}_t}$, and the point where $r_{\mathbf{a}_t} - \check{r}_{\mathbf{a}_t} = 0$ is a critical point. The proof for the convergence to a Nash equilibrium follows the same logic as that in [28]. ■

Now we prove the existence of an equilibrium solution in this finite-horizon stochastic game.

Theorem 1 There exists a Markov-perfect equilibrium θ^* in the second-level stochastic game (with observable actions) for a given pricing strategy π of the database.

Proof: Lemma 1 has ensured a finite sequence of Nash equilibria $\{\omega_1^*, \omega_2^*, \dots, \omega_{\hat{T}}^*\}$ in this finite-horizon stochastic game. Furthermore, it has been concluded that there always exists a subgame-perfect equilibrium in a finite-horizon multi-stage game if each stage-game has a Nash equilibrium [30]. Therefore, in this multi-stage stochastic game, there exists a Markov-perfect equilibrium θ^* , which is a subgame-perfect equilibrium in Markov strategies [26]. This is consistent with the finding that any N -player, general-sum, discounted-payoff stochastic game has a Markov-perfect equilibrium [31]. ■

Algorithm 1 pricing vs. contributing in a Stackelberg game

- 1: **repeat**
 - 2: **if** round $i = 1$ **then**
 - 3: **set** each ω_t^n in (3) according to an empirical demand function in economics
 - 4: **else if** round $i > 1$ **then**
 - 5: **set** each ω_t^n in (3) according to θ^* (in round $i - 1$)
 - 6: **end if**
 - 7: **solve** a finite-horizon MDP and obtain π^* (in round i) through a policy iteration algorithm
 - 8: **solve** a finite-horizon stochastic game and obtain θ^* (in round i) through a backward induction algorithm
 - 9: **until** $\{\pi^*, \theta^*\}$ is same for rounds $i - 1$ and i
-

Given a pricing strategy π , a Markov-perfect equilibrium θ^* can be derived by using a backward induction algorithm [32], which finds Nash equilibria in stage-games starting from the one in \hat{T} and moving successively towards the one in $t = 1$.

C. Stackelberg Equilibrium

It takes a number of rounds in the first-level Stackelberg game to alternately refine the pair of pricing and contributing strategies. In each round, the database solves the MDP to adapt its pricing strategy to the changes of contributing strategies in the previous round. Following the current pricing strategy, the SUs play the second-level stochastic game to agree on a new equilibrium solution of their contributing strategies. In the end, the database is expected to lead the second-level game towards an equilibrium solution to the first-level game.

Theorem 2 There exists a Stackelberg equilibrium, say $\{\pi^*, \theta^*\}$, in the first-level Stackelberg game, where the leading database derives its pricing strategy π^* in a finite-horizon MDP and the following SUs derive their contributing strategies θ^* in a finite-horizon stochastic game.

Proof: Theorem 1 has ensured an equilibrium solution of following strategies θ^* for any leading strategy π . We can actually view the set of SUs as one follower who responds to the database on behalf of the SUs, and thus this Stackelberg game is reduced to the basic one with single leader and single follower. It has been concluded that there always exists a Stackelberg equilibrium $\{\pi^*, \theta^*\}$ in such a game [33]. ■

The database can apply an algorithm that incorporates a policy iteration algorithm for the MDP [25] and a backward induction algorithm for the stochastic game [32], as summarized in Algorithm 1, to compute a Stackelberg equilibrium $\{\pi^*, \theta^*\}$. The database runs Algorithm 1 and computes π^* offline under the assumption that all the local SUs do not move out of their current location before the deadline and are willing to follow θ^* . Note that this algorithm is heuristic based on the database's non-unique estimation of user best responses in the stochastic game, and thus the uniqueness of the resulting equilibrium solution is not guaranteed.

VI. PAYMENT BUDGET

In our model, the payment for a contribution is assumed to be in the form of white space capacity, more specifically, the number of frequency-time white space blocks. However,

there are always limited white space blocks that set payment budgets for sensing tasks. The objective of dynamic pricing is to minimize the total payment for each sensing task, so that limited budgets are efficiently utilized to accommodate maximum contributing SUs and gather maximum good-quality sensing reports. However, it is still unclear how to support multiple sensing tasks under their payment budgets.

Suppose that the map is partitioned into a set of location pixels \mathcal{L} . Each location $l \in \mathcal{L}$ may fall within the SEZs of multiple channels, forming a set $\mathcal{K}_l = \{k \mid l \in \mathcal{L}_k\}$, and thus the SUs co-located at l , forming a set \mathcal{N}_l , may contribute to sensing tasks (k, l) for $k \in \mathcal{K}_l$. The contributing SUs in \mathcal{N}_l have to share the payment budget, i.e., the set of local frequency-time white space blocks $\mathcal{B}_l \subseteq \mathcal{K}_l \times \mathcal{T}$. Note that the channels in $\mathcal{K} - \mathcal{K}_l$ are either completely unavailable (if l is in HEZ) or freely available (if l is in UAZ), so \mathcal{B}_l as incentives should come from \mathcal{K}_l . We assume that each white space block $(k, t) \in \mathcal{B}_l$ is only assigned to one SU in \mathcal{N}_l to avoid mutual interference among the co-located SUs. Hence, each SU $n \in \mathcal{N}_l$ who has contributed to a sensing task needs to be paid with a separate subset of blocks from \mathcal{B}_l , and the number of paid blocks is proportional to $\sum_{t=1}^{\tilde{T}} p_t^n$. However, the block (k, t) can still be shared by the contributing SUs at other locations $l' \in \mathcal{L}$, $l' \neq l$. Hence, a payment budget \mathcal{B}_l should be set based on global view of incumbent protection.

For each block $(k, t) \in \mathcal{K} \times \mathcal{T}$, it is key to derive the maximum number of contributing SUs, $N_{(k,t)}$, that can be paid with the same (k, t) in SEZ \mathcal{L}_k without causing noticeable interference with co-channel IUs. The aggregate interference from SUs to IUs occupying (k, t) is defined by

$$\tilde{I}_{(k,t)} = \sum_{n=1}^{N_{(k,t)}} I_{(k,t)}^n, \quad (16)$$

where $I_{(k,t)}^n$ is the interference caused by n^{th} contributing SU paid with (k, t) . We assume that the negligible interference from the SUs taking (k, t) in UAZ does not contribute to $\tilde{I}_{(k,t)}$. In general, we need to ensure that $\tilde{I}_{(k,t)}$ is below a threshold \hat{I} for at least $1 - \epsilon$ fraction of the time, namely

$$\Pr \left\{ \tilde{I}_{(k,t)} \leq \hat{I} \right\} \geq 1 - \epsilon. \quad (17)$$

Now we first study each $I_{(k,t)}^n$ and further derive a closed-form expression of $\tilde{I}_{(k,t)}$ under path loss and shadowing. Beyond a reference distance d_0 , the path loss in dB for an interference link with distance d from the n^{th} contributing SU to the IUs at the center of HEZ is given by

$$Y_d = Y_{d_0} + 10\gamma \log_{10} \frac{d}{d_0} + X, \quad (18)$$

where Y_{d_0} is the reference path loss in dB, γ is the path loss exponent, and X is the normal shadowing in dB with zero mean and variance σ^2 . The transmit power of n^{th} contributing SU, Y^n , is governed by the following power control law [34].

$$Y^n = \left(\frac{d}{\rho_2} \right)^\alpha \hat{Y} \quad \text{for } \rho_1 < d \leq \rho_2, \quad (19)$$

where ρ_1 and ρ_2 are the approximated inner and outer radii

Algorithm 2 payment budget assignment

- 1: **set** w_l for $l \in \cup_{k \in \mathcal{K}} \mathcal{L}_k$ according to spec. sensing demands
 - 2: **for each** $(k, t) \in \mathcal{K} \times \mathcal{T}$ **do**
 - 3: **solve** an incumbent protection prob. and obtain $N_{(k,t)}$
 - 4: **select** $\tilde{\mathcal{L}}_k \subset \mathcal{L}_k$ according to w_l for $l \in \mathcal{L}_k$
 - 5: **update** \mathcal{B}_l for $l \in \tilde{\mathcal{L}}_k$ by adding (k, t)
 - 6: **end for**
-

of SEZ, respectively, α is the power control exponent, and \hat{Y} is the transmit power cap of each SU. Such a power control law helps in abstracting away the spatial distribution of contributing SUs inside the SEZ. Letting $\alpha = \gamma$ for simplicity and taking Y^n in dB, $I_{(k,t)}^n$ in dB can be written as

$$I_{(k,t)}^n = Y^n - Y_d \sim N(\mu, \sigma^2), \quad (20)$$

where μ is a constant. Thus, $I_{(k,t)}^n$ follows a log-normal (normal in dB) distribution with mean μ and variance σ^2 .

The sum of i.i.d. log-normal random variables, i.e., $I_{(k,t)}^n$, can be approximated as another log-normal random variable [35]. Hence, $\tilde{I}_{(k,t)}$ roughly follows a log-normal distribution. Based on the approximation in [35] that is most accurate in the tail portion of the resulting CDF, as demanded by (17), the mean and variance of $\tilde{I}_{(k,t)}$ are respectively given by

$$\tilde{\mu} = \ln \left(\sum_{n=1}^{N_{(k,t)}} \left(e^{\mu + \frac{\sigma^2}{2}} \right) \right) - \frac{\tilde{\sigma}^2}{2}, \quad (21)$$

$$\tilde{\sigma}^2 = \ln \left(\frac{\sum_{n=1}^{N_{(k,t)}} \left(e^{2\mu + \sigma^2} (e^{\sigma^2} - 1) \right)}{\sum_{n=1}^{N_{(k,t)}} \left(e^{\mu + \frac{\sigma^2}{2}} \right)} + 1 \right). \quad (22)$$

Given the distribution of $\tilde{I}_{(k,t)}$, the database is able to solve an incumbent protection problem, which maximizes $N_{(k,t)}$ subject to (17). After that, the database can globally select a set of at most $\lfloor N_{(k,t)} \rfloor$ locations, denoted by $\tilde{\mathcal{L}}_k \subset \mathcal{L}_k$, to include (k, t) in their payment budgets \mathcal{B}_l . To distribute limited budgets, more blocks should be assigned to the locations where spectrum sensing is more beneficial for dynamic zoning. The priority of each location $l \in \cup_{k \in \mathcal{K}} \mathcal{L}_k$ to be selected in a $\tilde{\mathcal{L}}_k$, denoted by w_l , can be high if l is e.g., indoor, near a boundary, with low cumulative usefulness score, or within the SEZs of many channels. Before running Algorithm 1 at each location with positive payment budget, the database conducts payment budget assignment as summarized in Algorithm 2. Note that as long as $\lfloor N_{(k,t)} \rfloor$ is positive, there will be sensing tasks created. If \mathcal{B}_l for a certain location l is non-empty after running Algorithm 2, the sensing tasks (k, l) for $k \in \mathcal{K}_l$ should be created with controlled targets and deadlines to ensure that the total number of paid blocks for all the tasks, which is proportional to $\sum_{k \in \mathcal{K}_l} \tilde{p}$, does not exceed $|\mathcal{B}_l|$.

VII. SIMULATION RESULTS

In this section, we evaluate our algorithms. For Algorithm 1 based on our two-level game-theoretic model, we focus on one sensing task to evaluate the convergence of the Stackelberg game and the optimality of the equilibrium solution in the

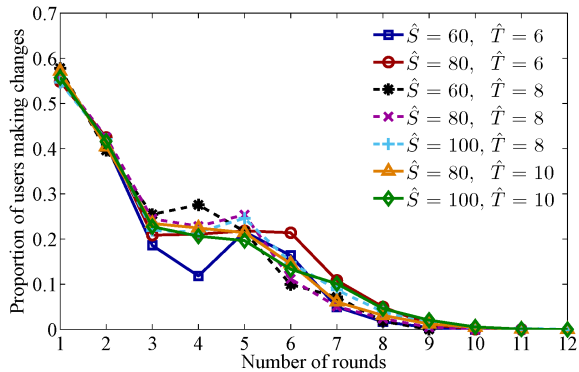


Fig. 2. Convergence of Stackelberg game.

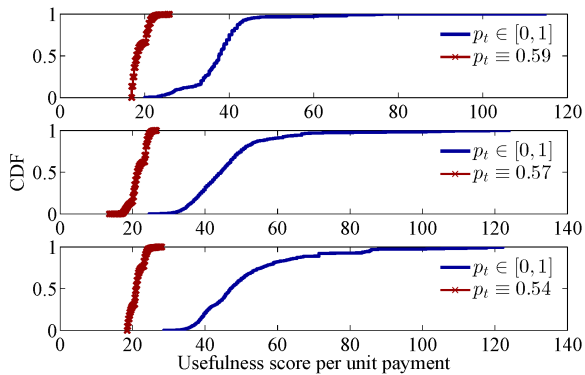


Fig. 3. Minimization of total payment by pricing strategy: (a) $\hat{S} = 60$, $\hat{T} = 6$ (upper); (b) $\hat{S} = 80$, $\hat{T} = 8$ (middle); (c) $\hat{S} = 100$, $\hat{T} = 10$ (lower).

game. For Algorithm 2 based on our aggregate interference analysis, we consider multiple sensing tasks to evaluate their sharing of limited payment budgets.

A. Convergence to Equilibrium Solution

We run Algorithm 1 to observe its convergence. Let $\mathcal{U} = \{1, 2, \dots, 5\}$, and assume that the usefulness score of a sensing report mainly depends on the sensing capability of the contributor. Then, we consider a sensing task (k, l) to be taken by five different SUs, and each SU $n \in \mathcal{N}^{(k,l)} = \{1, 2, \dots, 5\}$ (if takes the task) contributes a user-dependent usefulness score $u_t^n \equiv n$. For different settings of \hat{S} and \hat{T} , in Fig. 2, the convergence of our Stackelberg game to an equilibrium solution is guaranteed as the proportion of SUs changing their contributing strategies reaches zero in finite rounds of the algorithm. The duration of each round is finite as the stochastic game converges to an equilibrium solution.

B. Optimality of Equilibrium Solution

1) *Pricing Strategy*: We first compare our dynamic pricing strategy with the one-time pricing strategy in [15] facing strategic SUs with the same contributing strategy. Before comparison, we notice a trade-off between total payment minimization and sensing task completion. Intuitively, lower payments as incentives usually attract less contributions to a sensing task, and thus it is less likely to complete the task. For the same settings as above, we evaluate the equilibrium

TABLE I
PROBABILITY OF SENSING TASK FAILURE.

	$\hat{S} = 60$	$\hat{S} = 80$	$\hat{S} = 100$
$\hat{T} = 6$	0.0358	0.1442	1
$\hat{T} = 8$	0.0277	0.0436	0.2453
$\hat{T} = 10$	0	0.0303	0.1338

TABLE II
MAXIMUM NUMBER OF CO-CHANNEL USERS.

$\hat{Y} = 23, \sigma =$	1	2	3	4	5	6	7
$\epsilon = 0.1$	382	346	296	237	176	119	70
$\epsilon = 0.01$	377	337	282	218	151	88	35

$\sigma = 7, \hat{Y} =$	5	10	15	20	25	30
$\epsilon = 0.1$	6496	1966	575	158	40	10
$\epsilon = 0.01$	6191	1795	479	105	15	2

solution of dynamic pricing strategy outputted by Algorithm 1 in terms of the probability of sensing task failure. As shown in Tab. I, Algorithm 1 achieves low failure probability as long as the required value of $\frac{\hat{S}}{\hat{T}}$ is not too large compared with the maximum possible value of \tilde{u}_t , i.e., 15. Note that our strategy is aggressive in minimizing total payment, resulting in non-zero failure probability. Now we compare our dynamic pricing strategy with the one-time pricing strategy in terms of the usefulness score per unit payment $\frac{s_{\hat{T}}}{\hat{p}}$. As shown in Fig. 3, our dynamic pricing strategy (with varying $p_t \in [0, 1]$) outperforms the best one-time pricing strategy (with the lowest fixed p_t) that achieves comparable failure probability.

2) *Contributing Strategy*: Under the same dynamic pricing strategy, we compare our equilibrium solution of contributing strategies (in the stochastic game) with the optimal counterpart (in an optimization problem where all the SUs are centrally controlled). The minimum received payment of a contributing SU $\min_{n \in \mathcal{N}^{(k,t)}} \sum_{t=1}^{\hat{T}} p_t^n$ is maximized by the optimal solution. To achieve the same max-min objective in our stochastic game for a fair comparison, the database thwarts any SU n whose newly changed θ^n leads to a decrease in the minimum received payment, and a punishment proportional to the decrease is deducted from the total payoff $v^{\{\theta^n, \theta^{-n}\}}$. As shown in Fig. 4, the gap between our heuristic equilibrium solution and its optimal counterpart is not large.

C. Spectrum Utilization by Multiple Tasks

We move our focus from a local view of one sensing task to a global view of multiple sensing tasks sharing limited payment budgets, i.e., white space blocks. To protect IUs on channel k in time slot t against the aggregate interference from contributing SUs in $\tilde{\mathcal{L}}_k$ who are paid with the same block (k, t) , the maximum value of $N_{(k,t)}$ required in Algorithm 2 is evaluated, as listed in Tab. II. Here we set $\rho_1 = 70$ km, $\rho_2 = 126$ km, $\gamma = 2$, $Y_{d_0} = 37$ dB, $\hat{I} = -90$ dBm, and \hat{Y} is in dBm. We can see that each (k, t) can usually be shared at considerable locations $l \in \mathcal{L}_k$. Each \mathcal{B}_l for $l \in \mathcal{L}_k$ can include at most $|\mathcal{T}|$ k -channel white space blocks. Note that $|\mathcal{T}|$ for block partition can be different from \hat{T} for sensing participation. Let paying every $p_t = 0.01$ be equal to assigning one block. Setting $|\mathcal{T}| = 600$, $N_{(k,t)} = 218$ for $t = 1, 2, \dots, |\mathcal{T}|$, as an example, we evaluate the number

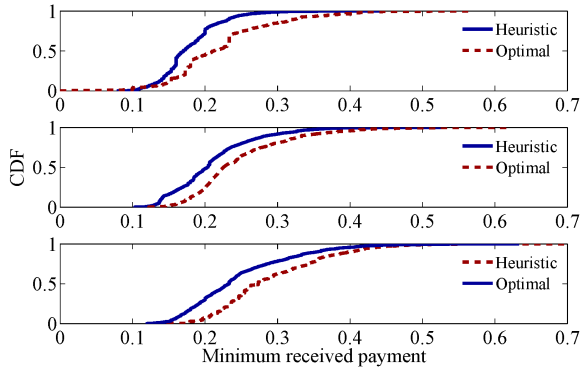


Fig. 4. Maximization of received payment by contributing strategy: (a) $\hat{S} = 60$, $\hat{T} = 6$ (upp.); (b) $\hat{S} = 80$, $\hat{T} = 8$ (mid.); (c) $\hat{S} = 100$, $\hat{T} = 10$ (low.).

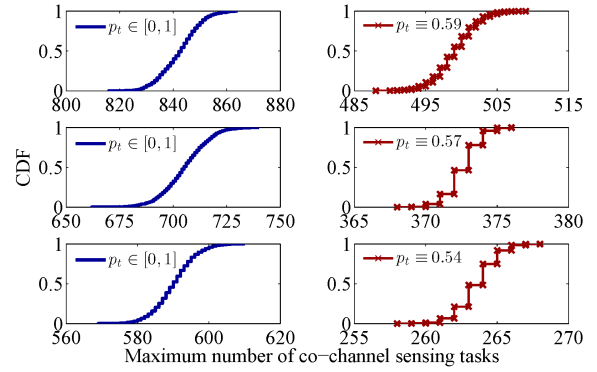


Fig. 5. Sharing of white space blocks as payment budgets: (a) $\hat{S} = 60$, $\hat{T} = 6$ (upper); (b) $\hat{S} = 80$, $\hat{T} = 8$ (middle); (c) $\hat{S} = 100$, $\hat{T} = 10$ (lower).

of sensing tasks that can be accommodated by the k -channel blocks. As shown in Fig. 5, our dynamic pricing strategy outperforms the best one-time pricing strategy in terms of the utilization efficiency of white space capacity. Our approach gathers more good-quality sensing reports to fully extend UAZ and accommodates more contributing SUs in SEZ. In other words, our approach creates spectrum access opportunities indirectly in extended UAZ and directly in SEZ.

VIII. CONCLUSION

In this paper, an overly-conservative legacy EZ has been redefined as a dynamic, multi-tier EZ, which relies on incentivizing participatory spectrum sensing to create additional spectrum access opportunities for SUs. On the one hand, the collection of spectrum sensing results based on a two-level game-theoretic model contributes to extend UAZ and thus allows unconditional secondary access in extended UAZ. On the other hand, the use of white space capacity as incentives based on an aggregate interference analysis allows conditional secondary access in SEZ. Our results have shown that the proposed incentive mechanism for the concept of dynamic, multi-tier EZs can support efficient database-driven spectrum sharing while ensuring incumbent protection.

REFERENCES

- [1] D. Gurney, G. Buchwald, L. Ecklund, S. Kuffner, and J. Grosspietsch. Geo-location database techniques for incumbent protection in the tv white space. In *Proc. IEEE DySPAN'08*, Oct. 2008.
- [2] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl. Senseless: A database-driven white spaces network. *IEEE Transactions on Mobile Computing*, 11(2), Feb. 2012.
- [3] B. Gao, J.-M. Park, and Y. Yang. Supporting mobile users in database-driven opportunistic spectrum access. In *Proc. ACM MobiHoc'14*, Aug. 2014.
- [4] NTIA. An assessment of the near-term viability of accommodating wireless broadband systems in the 1675-1710 mhz, 1755-1780 mhz, 3500-3650 mhz, and 4200-4220 mhz, 4380-4400 mhz bands. Fast Track Evaluation Report, Oct. 2010.
- [5] A. Ullah, S. Bhattarai, J.-M. Park, J. H. Reed, D. Gurney, and B. Bahrak. Multi-tier exclusion zones for dynamic spectrum sharing. In *Proc. IEEE ICC'15*, Jun. 2015.
- [6] S. Bhattarai, A. Ullah, J.-M. Park, J. H. Reed, D. Gurney, and B. Gao. Defining exclusion zones on the fly: Dynamic boundaries for spectrum sharing. In *Proc. IEEE DySPAN'15*, Sept. 2015.
- [7] FCC. Report and order and second further notice of proposed rulemaking. FCC 15-47, Apr. 2015.
- [8] X. Ying, J. Zhang, L. Yan, G. Zhang, M. Chen, and R. Chandra. Exploring indoor white spaces in metropolises. In *Proc. ACM MobiCom'13*, Sept. 2013.
- [9] J. C. Ribeiro, J. Ribeiro, J. Rodriguez, R. Dionisio, H. Esteves, P. Duarte, and P. Marques. Testbed for combination of local sensing with geolocation database in real environments. *IEEE Wireless Communications*, 19(4), Aug. 2012.
- [10] T. Zhang and S. Banerjee. Inaccurate spectrum databases? public transit to its rescue! In *Proc. ACM HotNets'13*, Nov. 2013.

- [11] A. Chakraborty and S. R. Das. Measurement-augmented spectrum databases for white space spectrum. In *Proc. ACM CoNEXT'14*, Dec. 2014.
- [12] X. Zhang, Z. Yang, W. Sun, Y. Liu, S. Tang, K. Xing, and X. Mao. Incentives for mobile crowd sensing: A survey. *IEEE Communications Surveys & Tutorials*, Mar. 2015.
- [13] L. Duan, T. Kubo, K. Sugiyama, J. Huang, T. Hasegawa, and J. Walrand. Incentive mechanisms for smartphone collaboration in data acquisition and distributed computing. In *Proc. IEEE INFOCOM'12*, Mar. 2012.
- [14] Y. Zhang and M. van der Schaar. Reputation-based incentive protocols in crowdsourcing applications. In *Proc. IEEE INFOCOM'12*, Mar. 2012.
- [15] D. Yang, G. Xue, X. Fang, and J. Tang. Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing. In *Proc. ACM MobiCom'12*, Aug. 2012.
- [16] T. Luo, H.-P. Tan, and L. Xia. Profit-maximizing incentive for participatory sensing. In *Proc. IEEE INFOCOM'14*, Apr. 2014.
- [17] Z. Feng, Y. Zhu, Q. Zhang, L. M. Ni, and A. V. Vasilakos. Trac: Truthful auction for location-aware collaborative sensing in mobile crowdsourcing. In *Proc. IEEE INFOCOM'14*, Apr. 2014.
- [18] Y. Wen, J. Shi, Q. Zhang, X. Tian, Z. Huang, H. Yu, Y. Cheng, and X. Shen. Quality-driven auction based incentive mechanism for mobile crowd sensing. *IEEE Transactions on Vehicular Technology*, Oct. 2014.
- [19] Y. Zhao, B. Le, and J. H. Reed. Network support—the radio environment map. In *Cognitive Radio Technology*. Elsevier, 2006.
- [20] H. B. Yilmaz, T. Tugcu, F. Alagöz, and S. Bayhan. Radio environment map as enabler for practical cognitive radio networks. *IEEE Communications Magazine*, 51(12), Dec. 2013.
- [21] M. Cho, M. Fan, and Y.-P. Zhou. Strategic consumer response to dynamic pricing of perishable products. In *Consumer-Driven Demand and Operations Management Models*. Springer US, 2009.
- [22] Y. Levin, J. McGill, and M. Nediak. Optimal dynamic pricing of perishable items by a monopolist facing strategic consumers. *Production and Operations Management*, 19(1), Jan. 2010.
- [23] W. Elmaghraby and P. Keskinocak. Dynamic pricing in the presence of inventory considerations: Research overview, current practices and future directions. *Management Science*, 49(10), Oct. 2003.
- [24] G. Gallego and G. van Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science*, 40(8), Aug. 1994.
- [25] N. Bäuerle and U. Rieder. Theory of finite horizon markov decision processes. In *Markov Decision Processes with Applications to Finance*. Springer Berlin Heidelberg, 2011.
- [26] E. Maskin and J. Tirole. Markov perfect equilibrium: I. observable actions. *Journal of Economic Theory*, 100(2), Oct. 2001.
- [27] C. Marlats. A folk theorem for stochastic games with finite horizon. *Economic Theory*, 58(3), Feb. 2015.
- [28] B. Gao, Y. Yang, and J.-M. Park. A credit-token-based spectrum etiquette framework for coexistence of heterogeneous cognitive radio networks. In *Proc. IEEE INFOCOM'14*, Apr. 2014.
- [29] B. Gao, J.-M. Park, and Y. Yang. Uplink soft frequency reuse for self-coexistence of cognitive radio networks. *IEEE Transactions on Mobile Computing*, 13(6), Jun. 2014.
- [30] S. Tadelis. *Game Theory: An Introduction*. Princeton University Press, 2013.
- [31] K. Leyton-Brown and Y. Shoham. *Essentials of Game Theory: A Concise, Multidisciplinary Introduction*. Morgan & Claypool Publishers, 2008.
- [32] U. Rieder. Equilibrium plans for non-zero-sum markov games. In *Game Theory and Related Topics*. North-Holland Publishing Company, 1979.
- [33] B. von Stengel and S. Zamir. Leadership with commitment to mixed strategies. CDAM Research Report LSE-CDAM-2004-01, Feb. 2004.
- [34] Z. Chen, C.-X. Wang, X. Hong, J. S. Thompson, S. Vorobyov, X. Ge, H. Xiao, and F. Zhao. Aggregate interference modeling in cognitive radio networks with power and contention control. *IEEE Transactions on Communications*, 60(2), Feb. 2012.
- [35] L. F. Fenton. The sum of log-normal probability distributions in scatter transmission systems. *IRE Transactions on Communications Systems*, 8(1), Mar. 1960.