

Economic Analysis of Crowdsourced Wireless Community Networks

Qian Ma, *Student Member, IEEE*, Lin Gao, *Member, IEEE*, Ya-Feng Liu, *Member, IEEE*, and Jianwei Huang, *Fellow, IEEE*

Abstract—Crowdsourced wireless community networks can effectively alleviate the limited coverage issue of Wi-Fi access points (APs), by encouraging individuals (users) to share their private residential Wi-Fi APs with others. In this paper, we provide a comprehensive economic analysis for such a crowdsourced network, with the particular focus on the users' behavior analysis and the community network operator's pricing design. Specifically, we formulate the interactions between the network operator and users as a two-layer Stackelberg model, where the operator determining the pricing scheme in Layer I, and then users determining their Wi-Fi sharing schemes in Layer II. First, we analyze the user behavior in Layer II via a two-stage membership selection and network access game, for both small-scale networks and large-scale networks. Then, we design a partial price differentiation scheme for the operator in Layer I, which generalizes both the complete price differentiation scheme and the single pricing scheme (i.e., no price differentiation). We show that the proposed partial pricing scheme can achieve a good tradeoff between the revenue and the implementation complexity. Numerical results demonstrate that when using the partial pricing scheme with only two prices, we can increase the operator's revenue up to 124.44% comparing with the single pricing scheme, and can achieve an average of 80% of the maximum operator revenue under the complete price differentiation scheme.

Index Terms—Mobile Crowdsourcing, Wireless Community Network, Economic Analysis, Price Differentiation

1 INTRODUCTION

1.1 Background and Motivation

Global mobile data traffic grows rapidly nowadays, with an unprecedented anticipated annual growth rate of 57% from 2014 to 2019 [2]. The capacity of cellular networks, however, increases much slower than the mobile data traffic. Hence, Wi-Fi networks are playing an increasingly important role in bridging such a gap by carrying a significant amount of mobile data traffic. The fast development of Wi-Fi technology is due to several reasons such as low deploying costs of Wi-Fi access points (APs) and high Wi-Fi data transmission rates [3]. However, the large-scale deployment of a Wi-Fi network is often restricted by the limited coverage of each single Wi-Fi AP (typically tens of meters indoors and hundreds of meters outdoors [4]), which is much smaller than the coverage of a cellular tower. Hence, it is very expensive for a single operator to deploy enough Wi-Fi APs to entirely cover a large area such as a city or a nation.

- Q. Ma and J. Huang are with Department of Information Engineering, The Chinese University of Hong Kong. E-mail: {mq012, jwuang}@ie.cuhk.edu.hk. L. Gao (Corresponding Author) is with School of Electronic and Information Engineering, Harbin Institute of Technology (Shenzhen), China. E-mail: gaolin@hitsz.edu.cn. Y.-F. Liu is with State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China. E-mail: yafliu@lsec.cc.ac.cn.

Part of this work has been presented at the 13th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), Mumbai, India, May 25-29, 2015 [1]. This work is partially supported by the General Research Funds (Project Number CUHK 14202814) established under the University Grant Committee of the Hong Kong Special Administrative Region, China and partially supported by the National Natural Science Foundation of China under Grants 11301516 and 11331012.

The crowdsourced wireless community network turns out as a promising solution to expand the Wi-Fi coverage with a low cost. The key idea is to encourage individuals (users) to share their owned private Wi-Fi APs with each other, hence aggregating/crowdsourcing the coverage and capacity of private Wi-Fi APs [5], [6]. By crowdsourcing millions of Wi-Fi APs already installed by users, this new type of network can reduce or even remove new Wi-Fi installations for the network operator. Meanwhile, users can also benefit from joining such a community network, as they can use not only their own APs when staying at home, but also other users' APs when traveling to corresponding locations. Clearly, the success of such a crowdsourced network largely depends on the active participations and contributions of many individual Wi-Fi owners, and hence requires a careful design of economic incentive mechanism.

One prominent commercial example of wireless community networks is FON [7], which has more than 20 millions member Wi-Fi APs globally. In FON, the operator incentivizes Wi-Fi AP owners to share their private APs by using two different incentive schemes, corresponding to two kinds of memberships: *Linus* and *Bill* [8]. As a *Linus*, a user can get free Wi-Fi access within the community network coverage. As a *Bill*, a user can earn money from sharing his AP with other users. Moreover, a user who does not own a Wi-Fi AP can still access the FON network as an *Alien*, by purchasing Wi-Fi passes from the operator.

1.2 Model and Contributions

In this work, we consider a wireless community network launched by a FON-like network operator. There are two types of mobile users in the network: *subscribers* and *Aliens*,

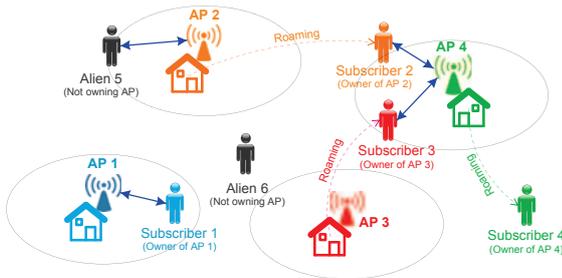


Figure 1. Wireless Community Network Model

each traveling (roaming) randomly according to certain mobility pattern. Each subscriber owns a private residential Wi-Fi AP at a fixed home location, and opens up his AP for the access of other users. An Alien does not own a Wi-Fi AP (hence does not contribute to the network), but can access subscribers' APs (when roaming to the corresponding locations) with a certain fee. Figure 1 illustrates such a wireless community network, where subscriber 1 (owner of AP 1) stays at home and connects to his own AP, subscribers 2 and 3 travel to subscriber 4's home location and connect to AP 4, and Alien 5 travels to subscriber 2's home location and connects to AP 2. Subscriber 4 and Alien 6 roam at areas without Wi-Fi coverage and have no available connections.

Similar as FON, the network operator offers two different memberships for subscribers, i.e., *Linus* or *Bill*, corresponding to two different incentive schemes:

- As a *Linus*, a subscriber contributes his AP without any monetary compensation, and as compensation, he can use other APs free of charge;
- As a *Bill*, a subscriber needs to pay for using other APs (according to a pricing scheme specified by the operator), and can obtain a portion of the revenue collected at his own AP by the network operator.

Moreover, an Alien has to pay for using any AP in the network (according to a usage-based pricing scheme specified by the operator), as he does not contribute to the network. The payments of Alien and Bill (for using other APs) are often *time usage-based* (i.e., proportional to the Wi-Fi connection time) [9].

The network operator and the users (subscribers and Aliens) interact in the following order. First, the operator announces the pricing scheme, i.e., the usage-based price (charged to Bills and Aliens) at each AP. Second, each subscriber chooses his membership type for a given *time period* (e.g., six months), considering his mobility (travel) pattern within that time period as well as his demand and evaluation for network access during travel. Third, when travelling to a particular AP's location in a particular *time slot* (e.g., five minutes), each user further decides his network access time on that AP in that time slot, taking the network congestion into consideration.

In this work, we will study the above operator pricing design and user decision problems systematically. We formulate the interactions between the operator and users as a two-layer Stackelberg model, where the operator decides the pricing scheme in Layer I and subscribers decide the membership selection and network access in Layer II, as illustrated in Figure 2. We want to emphasize that the operator's pricing decision and the subscribers' membership selection decisions are made once at the beginning

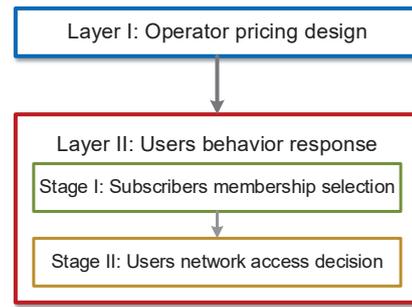


Figure 2. Stackelberg Model of the Operator and Users Interactions

of each time period (consisting of many time slots) and remain unchanged during the whole period; however, the subscribers' network access decisions are updated in each time slot, according to their membership selections and the operator's pricing scheme, as well as the real network congestions of their roaming APs. Next we explain the challenges for solving the two stage problems and our corresponding contributions.

We first consider the operator's decision problem in Layer I. To extract the maximum surplus from users, the operator can implement a *complete* price differentiation scheme, i.e., charging one price per AP, which will lead to a high implementation complexity and potential customer aversion [10]. To balance the revenue and implementation complexity, we propose a *partial* price differentiation scheme, where APs are divided into multiple groups according to their attributes such as the location popularity and the owner's network access evaluation, and the operator charges different prices to different groups. Such a partial price differentiation scheme includes both complete price differentiation scheme and single pricing scheme as special cases, but is much challenging to design than the latter two.

Next we consider the user decision problem in Layer II. As will be shown later in Section 3, to make the best decision, the computational complexity (for each user) to reach the decision increases exponentially with the number of users. Thus, in practice, users may not be able to make the fully rational decisions due to the bounded rationality [11],¹ especially in a large-scale network with many users (which is more general). To this end, we propose an *approximate* Stackelberg model for large-scale networks, and understand the system equilibrium when users make their decisions according to approximately best responses.

As far as we know, this is the first work that makes the comprehensive economic analysis in a crowdsourced wireless community network. We summarize the key contributions of this work as follows.

- *Model Novelty*: Our model is novel and captures several key practical issues, such as user mobility pattern, network access evaluation, demand response, and network congestion effect. These issues have not been fully considered before in the context of wireless community networks.
- *Small-Scale Network Analysis*: We propose a Stackelberg model to capture the interactions between the operator and users in the crowdsourced wireless

1. Bounded rationality is the concept that the rationality of decision makers is limited by the cognitive limitations of their minds [12].

community networks for small networks. We design a partial price differentiation scheme, which includes the complete price differentiation scheme and single pricing scheme as special cases, to help the operator achieve a balance between revenue and implementation complexity.

- *Large-Scale Network Analysis:* We propose an approximate Stackelberg model for the large-scale network with a large number of users, where users are bounded rational due to the limited computation capability and make the approximately best responses. We analyze the user behavior and the operator pricing design in the large-scale network systematically.

The rest of the paper is organized as below. In Section 2, we present the system model. In Section 3, we analyze users' two-stage dynamic game in Layer II. In Section 4, we analyze the operator's pricing design in Layer I. In Section 5, we analyze the user behavior and operator pricing design in the large-scale network. We show simulation results and derive engineering insights in Section 6, and conclude in Section 7. Due to space limit, we put some proofs at the appendix in the supplemental material section of the manuscript center.

1.3 Literature Review

There are several closely related studies in wireless community networks, regarding incentive issues [6], the network expansion and interactions with traditional ISP [13] [14], and the pricing mechanism design [15] [16]. Camponovo *et al.* in [6] concluded based on surveys that getting free Internet access from other members and revenue sharing are the two main incentives for users to join the FON network in Switzerland. Manshaei *et al.* in [13] modeled a user's payoff as a function of the subscription fee and network coverage, and studied the evolution dynamics of wireless community networks. Biczok *et al.* in [14] studied the competition and cooperation among users, wireless community network operator, and ISPs. Afrasiabi *et al.* in [15] proposed a low introductory price policy to promote the service adoption. Mazlounian *et al.* in [16] proposed a fair pricing, which is proportional to users' predicted network coverage.

In this work, we focus on the user behavior analysis and the operator pricing design in a crowdsourced wireless community network. Neither problem has been systematically studied in the existing literature. Our model not only captures the Internet access sharing and revenue sharing, but also incorporates the impact of users mobility and the network congestion effect. This makes our model more comprehensive and practically significant.

2 SYSTEM MODEL

2.1 The Network Model

As illustrated in Figure 1, we consider a crowdsourced wireless community network consisting of a set $\mathcal{K}_S = \{1, \dots, K\}$ of *subscriber* (users), each owning a private residential Wi-Fi AP, and a set $\mathcal{K}_A = \{K + 1, \dots, K + K_A\}$ of *Alien* (users) without owning any Wi-Fi AP. Each subscriber is associated with a "home", corresponding to the location of his Wi-Fi AP. Hence, \mathcal{K}_S also represents the set of AP locations. For convenience, we refer to the set of all subscribers and Aliens

Table 1
A Summary of Three User Types

User Type	Pay for using other APs	Paid by sharing his AP
Linus	No	No
Bills	Yes	Yes
Aliens	Yes	Not Applicable

as *user set*, denoted by $\mathcal{K}_U = \mathcal{K}_S \cup \mathcal{K}_A$. Subscribers open their private Wi-Fi APs for the access of other users, hence constitute a community network.

Each subscriber can choose two different memberships: Linus and Bill, corresponding to different incentive schemes. Specifically, as a *Linus*, a subscriber contributes his AP without any monetary compensation, and can use other APs free of charge. As a *Bill*, a subscriber needs to pay for using other APs, and can obtain a portion of the revenue collected at his own AP by the network operator. Moreover, each Alien has to pay for using any AP in the network, as he does not contribute to the network. For clarity, we summarize the properties of these user types in Table 1.

We consider the operation in a long time period (e.g., six months), which is divided into T time slots (e.g., five minutes per time slot). For notational convenience, we normalize the length of each time slot to be one. We consider a quasi-static mobility model, where each user moves randomly across time slots, and remains at the same location within one time slot. Let $\eta_{i,j}$, $i \in \mathcal{K}_U$, $j \in \mathcal{K}_S$ denote the stationary probability that a user $i \in \mathcal{K}_U$ appears at the location of AP $j \in \mathcal{K}_S$ in any time slot, and $\eta_{i,0}$ denote the probability that user i appears at a location that is not covered by any Wi-Fi AP in the community network. We further define $\boldsymbol{\eta}_i = [\eta_{i,0}, \eta_{i,1}, \dots, \eta_{i,K}]$ as user i 's *mobility pattern*. Obviously, $\sum_{j=0}^K \eta_{i,j} = 1$, $\forall i \in \mathcal{K}_U$.

Furthermore, to ensure a subscriber's Quality of Service (QoS) at his home location, we assume that each subscriber splits the bandwidth of his AP into two separate channels (similar as the current practice of FON [17]): a *private channel* for serving himself, and a *public channel* for serving other users roaming at this location. Hence, roaming users' communications will not interfere with a subscriber's own communication, and the network congestion only occurs among roaming users on the public channel.

2.2 Multi-Stage Interactions

The operator and users interact in the following order.

First, the operator determines and announces the *pricing scheme* at the beginning of the time period, which specifies the Wi-Fi access price on each AP paid by Aliens and Bills (except the AP owner), denoted by $\mathbf{p} = \{p_i, \forall i \in \mathcal{K}_S\}$. The operator's goal is to choose a proper set of prices to maximize the total revenue collected at all member APs.

Second, given the operator's pricing scheme $\mathbf{p} = \{p_i, \forall i \in \mathcal{K}_S\}$, each subscriber $i \in \mathcal{K}_S$ chooses his membership $x_i \in \{0, 1\}$ for the entire period of T time slots, where 0 and 1 correspond to "Linus" and "Bill", respectively. The goal of each subscriber is to choose the best membership that maximizes his expected payoff during the whole period of T time slots, taking his mobility pattern and demand (or evaluation) for network access as well as other users' membership selections into consideration.

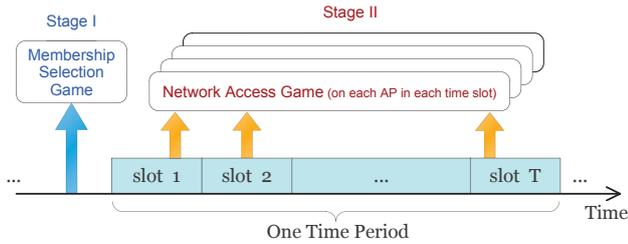


Figure 3. Two-Stage Dynamic Game in Layer II (In each time slot, there is a set of parallel network access games, each associated with an AP.)

Third, given the operator's pricing scheme $\mathbf{p} = \{p_i, \forall i \in \mathcal{K}_S\}$ and the subscribers' membership selections $\mathbf{x} = \{x_i, \forall i \in \mathcal{K}_S\}$, each user (subscriber or Alien) further decides the network usage in each time slot, i.e., the *network access time* on the AP of his current location. When staying at home, a subscriber uses his private channel exclusively, and his network access decision is independent of other users' decisions. When accessing the Internet through another subscriber's AP, a user (subscriber or Alien) needs to compete for the public channel with other users connecting to the same AP (except the owner of that AP), hence his best network access decision depends on other users' decisions.

In this work, we will study the operator pricing design problem and the user joint membership selection and network access decision problem comprehensively.

2.3 Problem Formulation

We formulate the interactions between the operator and the users as a two-layer Stackelberg model, as illustrated in Figure 2. In Layer I, the operator acts as the *leader* and optimizes the pricing scheme, based on his anticipation of users' responses (i.e., membership selection and network access) to the pricing scheme. The operator announces the pricing scheme to users at the beginning of each time period, and the pricing scheme will not change throughout the whole time period. In Layer II, the users act as the *followers* and decide their membership selections and network access decisions, given the operator's pricing scheme.

Moreover, we formulate users' joint membership selection and network access problem (in Layer II) as a two-stage dynamic game, as illustrated in Figure 3. In Stage I, subscribers participate in a *membership selection game* at the beginning of each time period, where each subscriber chooses his membership for the whole time period. In Stage II, at each time slot, users travelling to the same AP participate in a *network access game*, where each user decides his network access time on that AP. Namely, each AP is associated with a network access game at each time slot.

In this work, we consider both the small-scale and large-scale networks. In practice, each user has the limited computation capability, which may be enough for computing his best decision in a small-scale network, while not enough in a large-scale network. Hence, we assume that users will make the approximately best responses in large-scale networks, and hence are *bounded rational*. Namely, we propose an approximate Stackelberg model for large-scale networks, where users act in a bounded rational manner.

In what follows, we will study the Stackelberg game by backward induction, starting from the users' two-stage

dynamic game in Layer II (Section 3), and then moving to the operator's pricing design in Layer I (Section 4). We will further study the approximate Stackelberg model with bounded rational users in Section 5.

3 LAYER II: USER BEHAVIOR ANALYSIS

In this section, we study the users' two-stage dynamic game (for small-scale networks) in Layer II, given the operator's pricing scheme. We analyze the game by backward induction. We will first analyze the network access game in Stage II for each time slot and then analyze the membership selection game in Stage I for the whole time period.

3.1 Stage II: Network Access Game on Each AP

We first study the network access game in Stage II (on each AP in each time slot), given the subscribers' membership selections $\mathbf{x} = \{x_i, \forall i \in \mathcal{K}_S\}$ in Stage I and the operator's pricing scheme $\mathbf{p} = \{p_i, \forall i \in \mathcal{K}_S\}$ in Layer I. In this game, each user decides the network access time on the AP at his current location, aiming at maximizing his payoff in the current time slot.

3.1.1 Network Access Game Formulation

Without loss of generality, we consider the network access game on a particular AP k in a particular time slot t . Recall that the length of each time slot is normalized to be one for notational convenience.

The *players* of the game are all users traveling to AP k (except the owner of AP k) in time slot t , denoted by $\mathcal{K}(k, t) = \mathcal{K}_S(k, t) \cup \mathcal{K}_A(k, t)$, where $\mathcal{K}_S(k, t)$ and $\mathcal{K}_A(k, t)$ are the sets of subscribers and Aliens (at AP k and time slot t), respectively. For notational convenience, we will ignore the time index t , and hence write the player set as $\mathcal{K}(k) = \mathcal{K}_S(k) \cup \mathcal{K}_A(k)$ in the rest of this section, since we stick on the operations in time slot t .

The *strategy* of each player $i \in \mathcal{K}(k)$ is to decide the network access time $\sigma_{i,k} \in [0, 1]$ on AP k in time slot t . We denote the strategies of players in $\mathcal{K}(k)$ except i as $\sigma_{-i,k} = \{\sigma_{j,k}, j \neq i, j \in \mathcal{K}(k)\}$. The *payoff* of player i is a function of both his own strategy $\sigma_{i,k}$ and other players' strategies $\sigma_{-i,k}$, denoted by $v_{i,k}(\sigma_{i,k}, \sigma_{-i,k})$ (to be defined later).

More formally, the network access game on a particular AP k (in time slot t) and the corresponding Nash equilibrium are defined as follows.

Definition 1 (Network Access Game on AP k).

- *Players*: the set $\mathcal{K}(k)$ of users traveling to AP k ;
- *Strategies*: the network access time $\sigma_{i,k} \in [0, 1]$ of each user $i \in \mathcal{K}(k)$ on AP k ;
- *Payoffs*: $v_{i,k}(\sigma_{i,k}, \sigma_{-i,k}), \forall i \in \mathcal{K}(k)$.

Definition 2 (Nash Equilibrium). A *Nash equilibrium* of the Network Access Game on AP k (in time slot t) is a profile $\sigma_k^* = \{\sigma_{i,k}, \forall i \in \mathcal{K}(k)\}$ such that for each user $i \in \mathcal{K}(k)$,

$$v_{i,k}(\sigma_{i,k}^*, \sigma_{-i,k}^*) \geq v_{i,k}(\sigma_{i,k}, \sigma_{-i,k}^*), \quad \forall \sigma_{i,k} \in [0, 1].$$

Note that the Nash equilibrium σ_k^* depends on the player set $\mathcal{K}(k)$, hence can be written as $\sigma_k^*(\mathcal{K}(k))$.

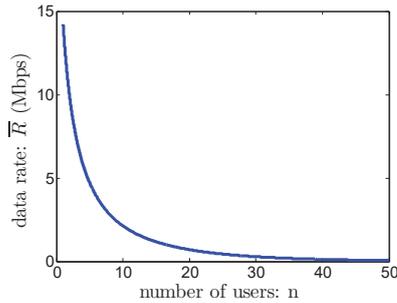


Figure 4. Average Data Rate per Wi-Fi User [20]

3.1.2 Utility and Payoff Definition

Before analyzing the Nash equilibrium, we first define users' utility and payoff functions.

Utility: The *utility* captures a user's satisfaction for accessing the Internet for a certain amount of time. Due to the diminishing marginal returns principle [18], [19], we assume the utility function is increasing and concave. As a concrete example, we define the utility of user $i \in \mathcal{K}(k)$ on AP k as

$$u_i(\sigma_{i,k}, \sigma_{-i,k}) = \rho_i \log(1 + \bar{r}_{i,k}(\sigma_{-i,k}) \cdot \sigma_{i,k}), \quad (1)$$

where ρ_i is user i 's network access evaluation, characterizing user i 's valuation of data consumption. Here, $\bar{r}_{i,k}(\sigma_{-i,k})$ is the expected data rate that user i can achieve on AP k , which is a decreasing function of other users' network access times $\sigma_{-i,k}$ on AP k . Intuitively, with more users accessing AP k 's public channel simultaneously, user i 's achieved data rate will decrease due to the increased congestion. Obviously, $\bar{r}_{i,k}(\sigma_{-i,k}) \cdot \sigma_{i,k}$ denotes the total expected data amount that user i consumes on AP k (in time slot t).

Next, we derive the concrete form of the user i 's expected data rate $\bar{r}_{i,k}(\sigma_{-i,k})$ on AP k . Let $\bar{R}(n)$ denote the average data rate of a Wi-Fi user when n users are connecting to the Wi-Fi AP simultaneously. Let $P_{i,k}(n)$ denote the probability that n other users (except i) connect to AP k . Then, user i 's expected data rate $\bar{r}_{i,k}(\sigma_{-i,k})$ can be calculated as follows:

$$\bar{r}_{i,k}(\sigma_{-i,k}) = \sum_{n=0}^{|\mathcal{K}(k)|-1} P_{i,k}(n) \cdot \bar{R}(n+1). \quad (2)$$

According to IEEE 802.11g standard [20], we have:

$$\bar{R}(n) = \frac{\tau \bar{\tau}^{n-1} L}{\bar{\tau}^n T_b + [(1 - \bar{\tau}) - n\tau \bar{\tau}^{n-1}] T_c + n\tau \bar{\tau}^{n-1} T_s}, \quad (3)$$

where τ is the average successful probability of contention (and $\bar{\tau} = 1 - \tau$), L is the average payload length, T_b is the length of a backoff slot, T_c is the length of a collision slot, and T_s is the length of a successful slot. Figure 4 illustrates an example of $\bar{R}(\cdot)$ under IEEE 802.11g standard (reproduced from [20], with parameters $\tau = 0.0765$, $L = 8192$, $T_b = 28\mu s$, and $T_c = T_s = 85.7 + L/54\mu s$). The decreasing data rate per user is due to both the reduced resource per user and the waste of resources caused by congestion among users.

For simplicity, we assume that if a user i decides to access the channel with a certain time $\sigma_{i,k}$, he will spread this access time *randomly and uniformly* across the entire time slot. Recall that the length of a time slot is normalized to 1. Hence, the probability that user i connects to AP k in an infinitely small time interval within the time slot is $\sigma_{i,k}$.

Therefore, the probabilities $P_{i,k}(n)$, $n = 0, 1, \dots, |\mathcal{K}(k)| - 1$, follow the binomial distribution (with a total of $|\mathcal{K}(k)|$ trials and a success probability $\sigma_{j,k}$ for each trial $j \in \mathcal{K}(k) \setminus \{i\}$). Formally,

$$P_{i,k}(n) = \sum_{\mathcal{K}_n \in \mathbf{K}_n(k)} \left(\prod_{j \in \mathcal{K}_n} \sigma_{j,k} \cdot \prod_{j \in \mathcal{K}(k) \setminus \{i\} \setminus \mathcal{K}_n} (1 - \sigma_{j,k}) \right),$$

where \mathcal{K}_n denotes an arbitrary subset of $\mathcal{K}(k)$ with n users (except i), and $\mathbf{K}_n(k)$ denotes the set of all possible \mathcal{K}_n . Obviously, $\prod_{j \in \mathcal{K}_n} \sigma_{j,k}$ denotes the probability that all users in \mathcal{K}_n are connecting to AP k , and $\prod_{j \in \mathcal{K}(k) \setminus \{i\} \setminus \mathcal{K}_n} (1 - \sigma_{j,k})$ denotes the probability that all other users (except user i and those in \mathcal{K}_n) are *not* connecting to AP k .

Payoff: The *payoff* of each user $i \in \mathcal{K}(k)$ is defined as the difference between the utility and the payment. Specifically, if user i is a Linus (i.e., $i \in \mathcal{K}_S(k)$ and $x_i = 0$), he does not need to pay for his network usage on AP k . Hence, the payoff of a Linus-type user i on AP k , denoted by $v_{i,k}^L$, is the same as his utility defined in (1), i.e.,

$$v_{i,k}^L(\sigma_{i,k}, \sigma_{-i,k}) = u_i(\sigma_{i,k}, \sigma_{-i,k}). \quad (4)$$

If user i is a Bill (i.e., $i \in \mathcal{K}_S(k)$ and $x_i = 1$) or Alien (i.e., $i \in \mathcal{K}_A(k)$), he needs to pay for his network usage on AP k , and the payment is proportional to his network access time $\sigma_{i,k}$. Hence, the payoff of a Bill-type or Alien user i , denoted by $v_{i,k}^B$, is the difference between utility and payment, i.e.,

$$v_{i,k}^B(\sigma_{i,k}, \sigma_{-i,k}) = u_i(\sigma_{i,k}, \sigma_{-i,k}) - p_k \sigma_{i,k}. \quad (5)$$

Based on the above, we can summarize the payoff of user $i \in \mathcal{K}(k)$ in the Network Access Game (on AP k) as follows:

$$v_{i,k}(\sigma_{i,k}, \sigma_{-i,k}) = \begin{cases} v_{i,k}^L(\sigma_{i,k}, \sigma_{-i,k}), & \text{if } i \in \mathcal{K}_S(k) \text{ and } x_i = 0; \\ v_{i,k}^B(\sigma_{i,k}, \sigma_{-i,k}), & \text{if } i \in \mathcal{K}_S(k) \text{ and } x_i = 1; \\ v_{i,k}^B(\sigma_{i,k}, \sigma_{-i,k}), & \text{if } i \in \mathcal{K}_A(k). \end{cases} \quad (6)$$

3.1.3 Nash Equilibrium Analysis

Now we study the Nash equilibrium of the above Network Access Game (on AP k).

Given all other users' strategies, a user's *best response* is the strategy that maximizes his payoff. The Nash equilibrium is a strategy profile where each user's strategy is the best response to other users' strategies.

Lemma 1. *If user i is a Linus, his best response in the Network Access Game on AP k is*

$$\sigma_{i,k}^* = 1, \quad (7)$$

regardless of other users' strategies.

Lemma 2. *If user i is a Bill or an Alien, his best response in the Network Access Game on AP k is*

$$\sigma_{i,k}^* = \min \left\{ 1, \max \left\{ \frac{\rho_i}{p_k} - \frac{1}{\bar{r}_{i,k}(\sigma_{-i,k})}, 0 \right\} \right\}, \quad (8)$$

which is a function of other users' strategies $\sigma_{-i,k}$.

We next illustrate the existence of the Nash equilibrium in the Network Access Game.

Theorem 1. *There exists at least one Nash equilibrium in the Network Access Game on AP k .*

Now we discuss the uniqueness of the Nash equilibrium in the Network Access Game on AP k .

Proposition 1. *In a Network Access Game with two players, the Nash equilibrium is unique if $\frac{\bar{R}(1)-\bar{R}(2)}{(\bar{R}(2))^2} < 1$.*

Note that the condition in Proposition 1 is always satisfied for practical WiFi systems [20]. For the cases with more than two players, however, the uniqueness of the Nash equilibrium depends on the system parameters in a more complicated fashion. Please refer to the appendix in the supplemental material section for more detailed discussions. We further propose a best response update algorithm in the appendix, which is guaranteed to linearly converge to the Nash equilibrium under the same condition for the uniqueness of the Nash equilibrium.

3.2 Stage I: Membership Selection Game

Now we study the subscribers' membership selection game in Stage I, given the operator's pricing scheme $\mathbf{p} = \{p_i, \forall i \in \mathcal{K}_S\}$. In this stage, each subscriber $i \in \mathcal{K}_S$ decides his membership type $x_i \in \{0, 1\}$ (i.e., Linus or Bill) at the beginning of the period, aiming at maximizing the overall expected payoff that he can achieve in all T time slots. Note that an Alien $i \in \mathcal{K}_A$ cannot choose his type, as he has no Wi-Fi AP and does not contribute to the network.

3.2.1 Membership Selection Game Formulation

In the Membership Selection Game, *players* are subscribers in the set \mathcal{K}_S . The *strategy* of each player $i \in \mathcal{K}_S$ is to decide his membership $x_i \in \{0, 1\}$, with $x_i = 0$ and 1 denoting Linus and Bill, respectively. We denote the strategies of all players except i by $\mathbf{x}_{-i} = \{x_j, j \neq i, j \in \mathcal{K}_S\}$. The *overall payoff* of a player i is sum of the *total expected payoff* on all APs that he may travel to and the *total expected revenue* that he may collect at his own AP (if choosing to be a Bill) during T slots. It is a function of his own strategy x_i and other players' strategies \mathbf{x}_{-i} , denoted by $V_i(x_i, \mathbf{x}_{-i})$.

Formally, the Membership Selection Game and the corresponding Nash equilibrium are defined as follows.

Definition 3 (Membership Selection Game).

- *Players:* the set \mathcal{K}_S of subscribers.
- *Strategies:* $x_i \in \{0, 1\}, \forall i \in \mathcal{K}_S$.
- *Payoffs:* $V_i(x_i, \mathbf{x}_{-i}), \forall i \in \mathcal{K}_S$.

Definition 4 (Nash Equilibrium). *A Nash equilibrium of the Membership Selection Game is a profile $\mathbf{x}^* = \{x_i^*, i \in \mathcal{K}_S\}$ such that for each subscriber $i \in \mathcal{K}_S$,*

$$V_i(x_i^*, \mathbf{x}_{-i}^*) \geq V_i(x_i, \mathbf{x}_{-i}^*), \quad \forall x_i \in \{0, 1\}.$$

We note that the Nash equilibria in Stage II (Definition 2) and Stage I (Definition 4) together form a Subgame Perfect Equilibrium (SPE) of the whole game.

3.2.2 Payoff Definition

Before analyzing the Nash equilibrium, we first calculate each subscriber's overall expected payoff during the whole period, which includes (i) the total expected payoff on all APs that he may travel to, and (ii) the total expected revenue that he may collect on his own AP (if choosing to be a Bill).

Total expected payoff: We first calculate the total expected payoff of each subscriber (on all APs that he may travel to), which depends on his mobility pattern. Recall that the mobility of a subscriber i is characterized by the probabilities of travelling to different APs, i.e., $\boldsymbol{\eta}_i = [\eta_{i,0}, \eta_{i,1}, \dots, \eta_{i,K}]$, where $\eta_{i,k}$ is the probability of subscriber i travelling to AP k , and $\eta_{i,0}$ is the probability of subscriber i travelling to an area that is not covered by any AP in the network. We calculate subscriber i 's expected payoffs (per time slot) when staying at home and when roaming outside, respectively.

(a) *When staying at home (with a probability $\eta_{i,i}$),* subscriber i communicates over the private channel of AP i and does not interfere with other users. Hence his expected payoff, denoted by $V_{i,i}(x_i, \mathbf{x}_{-i})$, is

$$V_{i,i}(x_i, \mathbf{x}_{-i}) = \rho_i \cdot \log(1 + \bar{r}_{i,i} \cdot 1),$$

where constant $\bar{r}_{i,i}$ corresponds to the average achieved data rate. The product term $\bar{r}_{i,i} \cdot 1$ implies that user i will access the Internet during the entire time slot.

(b) *When traveling to AP $k \neq i$ (with a probability $\eta_{i,k}$),* subscriber i needs to compete over the public channel with other users (except k) travelling to AP k at the same time (in the Network Access Game).

Suppose that a set $\mathcal{M}(k)$ of other users (except i and k) are travelling to AP k at the same time. That is, the game player set in the Network Access Game on AP k is $\mathcal{K}(k) = \mathcal{M}(k) \cup \{i, k\}$. For more clarity, let us rewrite the equilibrium payoff of subscriber i on AP k , i.e., $v_{i,k}(\sigma_{i,k}, \boldsymbol{\sigma}_{-i,k})$ defined in (6), as $v_{i,k}(\sigma_{i,k}, \boldsymbol{\sigma}_{-i,k} | \mathcal{M}(k))$, when competing with a set $\mathcal{M}(k)$ of other users (in the Network Access Game on AP k). Hence, the expected payoff of subscriber i on AP k is

$$V_{i,k}(x_i, \mathbf{x}_{-i}) = \sum_{\mathcal{M}(k) \in \mathbf{K}_{-\{i,k\}}} \phi(\mathcal{M}(k)) v_{i,k}(\sigma_{i,k}^*, \boldsymbol{\sigma}_{-i,k}^* | \mathcal{M}(k)),$$

where $\phi(\mathcal{M}(k))$ is the probability that a set $\mathcal{M}(k)$ of users are travelling to AP k , $(\sigma_{i,k}^*, \boldsymbol{\sigma}_{-i,k}^*)$ is the corresponding equilibrium of the Network Access Game, and $\mathbf{K}_{-\{i,k\}}$ is the power set of $\mathcal{K}_U \setminus \{i, k\}$, i.e., the set of all subsets of $\mathcal{K}_U \setminus \{i, k\}$. The probability $\phi(\mathcal{M}(k))$ is given by

$$\phi(\mathcal{M}(k)) = \prod_{j \in \mathcal{M}(k)} \eta_{j,k} \cdot \prod_{j \in \mathcal{K}_U \setminus \{i,k\} \setminus \mathcal{M}(k)} (1 - \eta_{j,k}),$$

where $\prod_{j \in \mathcal{M}(k)} \eta_{j,k}$ denotes the probability that all users in $\mathcal{M}(k)$ are travelling to AP k , and $\prod_{j \in \mathcal{K}_U \setminus \{i,k\} \setminus \mathcal{M}(k)} (1 - \eta_{j,k})$ denotes the probability that all other users (except users i, k , and those in $\mathcal{M}(k)$) are *not* travelling to AP k .

(c) *When traveling to an area that is not covered by any AP (with a probability $\eta_{i,0}$),* the expected payoff of subscriber i , denoted by $V_{i,0}(x_i, \mathbf{x}_{-i})$, is

$$V_{i,0}(x_i, \mathbf{x}_{-i}) = 0.$$

Based on the above, the total expected payoff of subscriber i (on all APs that he may travel to during the whole period of T time slots) is

$$V_i^\dagger(x_i, \mathbf{x}_{-i}) = T \cdot \sum_{k=0}^K \eta_{i,k} \cdot V_{i,k}(x_i, \mathbf{x}_{-i}). \quad (9)$$

Total expected revenue: Next, we calculate the total expected revenue that each subscriber i may collect on his own AP. Specifically, if choosing to be a Linus, subscriber i obtains a zero revenue. If choosing to be a Bill, subscriber i obtains a fixed portion δ of the revenue collected at his AP.

Suppose that a set $\mathcal{K}(i)$ of other users (except i) are travelling to AP i . That is, the player set in the Network Access Game on AP i is $\mathcal{K}(i)$. Then, the Nash equilibrium in the Network Access Game on AP i can be written as

$$\{\sigma_{j,i}^*(\mathcal{K}(i)), \forall j \in \mathcal{K}(i)\}. \quad (10)$$

Recall that the revenue collected on each AP is the total payment of all Aliens and Bills accessing that AP. Hence, the total revenue collected on AP i is

$$\Pi_i(\mathbf{x}_{-i}, \mathcal{K}(i)) = \sum_{j \in \mathcal{K}(i) \cap \mathcal{K}_A} p_i \cdot \sigma_{j,i}^*(\mathcal{K}(i)) + \sum_{j \in \mathcal{K}(i) \cap \mathcal{K}_S} x_j \cdot p_i \cdot \sigma_{j,i}^*(\mathcal{K}(i)),$$

where the first term is the payment of Aliens, and the second term is the payment of Bills. Hence, the total expected payment of Bills and Aliens on AP i is

$$\bar{\Pi}_i(\mathbf{x}_{-i}) = \sum_{\mathcal{K}(i) \in \mathbf{K}_{-i}} \phi(\mathcal{K}(i)) \cdot \Pi_i(\mathbf{x}_{-i}, \mathcal{K}(i)),$$

where $\phi(\mathcal{K}(i))$ is the probability that a set $\mathcal{K}(i)$ of users are travelling to AP i , and \mathbf{K}_{-i} is the power set of $\mathcal{K}_U \setminus \{i\}$. The probability $\phi(\mathcal{K}(i))$ is given by

$$\phi(\mathcal{K}(i)) = \prod_{j \in \mathcal{K}(i)} \eta_{j,i} \cdot \prod_{j \in \mathcal{K}_U \setminus \{i\} \setminus \mathcal{K}(i)} (1 - \eta_{j,i}). \quad (11)$$

Based on the above, the total expected revenue that a subscriber i can achieve at his own AP (during the whole time period of T time slots) is

$$V_i^\ddagger(x_i, \mathbf{x}_{-i}) = T \cdot x_i \cdot \delta \cdot \bar{\Pi}_i(\mathbf{x}_{-i}). \quad (12)$$

Overall payoff: Combining the total expected payoff in (9) and the total expected revenue in (12), the overall payoff of each subscriber in the Membership Selection Game is

$$\begin{aligned} V_i(x_i, \mathbf{x}_{-i}) &= V_i^\ddagger(x_i, \mathbf{x}_{-i}) + V_i^\dagger(x_i, \mathbf{x}_{-i}) \\ &= T \cdot \left(x_i \cdot \delta \cdot \bar{\Pi}_i(\mathbf{x}_{-i}) + \sum_{k=0}^K \eta_{i,k} \cdot V_{i,k}(x_i, \mathbf{x}_{-i}) \right). \end{aligned} \quad (13)$$

3.2.3 Nash Equilibrium Analysis

A subscriber i will make the membership decision to maximize his overall payoff defined in (13). Specifically, he will choose to be a Linus if $V_i(0, \mathbf{x}_{-i}) > V_i(1, \mathbf{x}_{-i})$, and choose to be a Bill otherwise. For notational convenience, we denote $f_i(\mathbf{x}_{-i})$ as the gap between $V_i(1, \mathbf{x}_{-i})$ and $V_i(0, \mathbf{x}_{-i})$:

$$f_i(\mathbf{x}_{-i}) = V_i(1, \mathbf{x}_{-i}) - V_i(0, \mathbf{x}_{-i}). \quad (14)$$

Hence, subscriber i will choose to be a Linus ($x_i = 0$) if $f_i(\mathbf{x}_{-i}) < 0$, and choose to be a Bill ($x_i = 1$) if $f_i(\mathbf{x}_{-i}) \geq 0$. Mathematically, this is equivalent to choosing x_i from $\{0, 1\}$, such that the following condition holds:

$$(2x_i - 1) \cdot f_i(\mathbf{x}_{-i}) \geq 0.$$

Next, we study the Nash equilibrium of the Membership Selection Game.

Lemma 3. *A membership profile \mathbf{x}^* is an Nash equilibrium of the Membership Selection Game, if and only if*

$$(2x_i^* - 1) \cdot f_i(\mathbf{x}_{-i}^*) \geq 0, \quad \forall i \in \mathcal{K}.$$

Proposition 2. *For each subscriber i , if*

$$\eta_{i,i} > \underline{\eta}_i \triangleq 1 - \frac{\delta \cdot \bar{\Pi}_i(\mathbf{x}_{-i}^*)}{\sum_{k \in \mathcal{K}_S \setminus \{i\}} (V_{i,k}(0, \mathbf{x}_{-i}^*) - V_{i,k}(1, \mathbf{x}_{-i}^*))},$$

then his best response is to choose to be a Bill (i.e., $x_i = 1$).

Intuitively, a subscriber with a large probability of staying at home will choose to be a Bill, as his network usage on other APs is small, hence the benefit of obtaining revenue at his own AP outweighs the payment at other APs.

Unfortunately, the Membership Selection Game may not always possess an Nash equilibrium in Definition 4, which is essentially a *pure strategy* equilibrium (i.e., each subscriber chooses a particular membership). To illustrate this, we provide a simple example with 3 APs in the appendix. Hence, in what follows, we will further look at the case of *mixed-strategy* Nash equilibrium, where each subscriber may choose both membership types with probabilities.

3.2.4 Mixed-Strategy Nash Equilibrium

For each subscriber i , his *mixed strategy* can be characterized as the probability $\alpha_i \in [0, 1]$ of choosing to be a Bill (hence the probability of choosing to be a Linus is $1 - \alpha_i$). The pure strategy x_i is a special case of the mixed strategy when α_i equals 1 or 0. For writing convenience, we denote the mixed strategy profile of all subscribers except i as

$$\alpha_{-i} = \{\alpha_j, j \neq i, j \in \mathcal{K}_S\}.$$

Then, the expected payoff of subscriber i can be defined as

$$\omega_i(\alpha_i, \alpha_{-i}) = \alpha_i \cdot \bar{V}_i(1, \alpha_{-i}) + (1 - \alpha_i) \cdot \bar{V}_i(0, \alpha_{-i}), \quad (15)$$

where $\bar{V}_i(1, \alpha_{-i})$ and $\bar{V}_i(0, \alpha_{-i})$ are subscriber i 's expected payoffs when choosing to be a Bill and a Linus, respectively. Note that $\bar{V}_i(1, \alpha_{-i})$ and $\bar{V}_i(0, \alpha_{-i})$ are the expected values over all possible membership selections of all other users. Specifically, there are $K - 1$ other subscribers, hence 2^{K-1} possible membership selection combination of those subscribers, forming a set \mathcal{X}_{-i} . Each subscriber j chooses $x_j = 1$ and 0 with probabilities α_j and $1 - \alpha_j$, respectively. Then, the probability that a particular $\mathbf{x}_{-i} \in \mathcal{X}_{-i}$ is

$$\psi(\mathbf{x}_{-i}) = \prod_{j \in \mathcal{K}_S \setminus \{i\}} (\alpha_j \cdot x_j + (1 - \alpha_j) \cdot (1 - x_j)). \quad (16)$$

Then, $\bar{V}_i(1, \alpha_{-i})$ and $\bar{V}_i(0, \alpha_{-i})$ can be calculated by

$$\bar{V}_i(x_i, \alpha_{-i}) = \sum_{\mathbf{x}_{-i} \in \mathcal{X}_{-i}} \psi(\mathbf{x}_{-i}) V_i(x_i, \mathbf{x}_{-i}), \quad x_i \in \{0, 1\}, \quad (17)$$

where $V_i(x_i, \mathbf{x}_{-i})$ is the overall payoff of subscriber i under the pure strategy profile defined in (13).

Definition 5 (Mixed-Strategy Nash Equilibrium). *A mixed-strategy Nash equilibrium of the Membership Selection Game is a probability profile α^* such that for each subscriber $i \in \mathcal{K}_S$:*

$$\omega_i(\alpha_i^*, \alpha_{-i}^*) \geq \omega_i(\alpha_i, \alpha_{-i}^*), \quad \forall \alpha_i \in [0, 1].$$

We first show the existence of the mixed-strategy Nash equilibrium in the Membership Selection Game.

Theorem 2. *There exists at least one mixed-strategy Nash equilibrium in the Membership Selection Game.*

To compute the Nash equilibrium effectively, we design a *smoothed* best response updated algorithm, where each player updates his mixed strategy in a smoothed best response manner according to the other players' mixed strategies in the previous iteration. The basic idea is as follows. First, given the mixed strategy profile α^n at the n -th round, each player computes the corresponding expected payoff when choosing to be Bill (i.e., $\bar{V}_i(1, \alpha_{-i}^n)$) or to be Linus (i.e.,

$\tilde{V}_i(0, \alpha_{-i}^n)$). Then, each player updates his mixed strategy at the $(n + 1)$ -th round according to the following smoothed best response method [23]:

$$\alpha_i^{n+1} = \frac{e^{\tilde{V}_i(1, \alpha_{-i}^n)/\gamma}}{e^{\tilde{V}_i(1, \alpha_{-i}^n)/\gamma} + e^{\tilde{V}_i(0, \alpha_{-i}^n)/\gamma}},$$

where γ is a parameter that determines the degree to which the function deviates from the true best response (a larger γ implies that the player is more likely to act randomly). Using the result in [23], we can show that such a smoothed best response with some learning rules (as in fictitious play) converges to the mixed strategy Nash equilibria.

Algorithm 1 Smoothed Best Response Update Algorithm

Input: $\alpha^0, \gamma, \varepsilon$.

Output: α^* .

- 1: Set $n = 0$ and $Flag = 0$.
 - 2: **while** $Flag = 0$ **do**
 - 3: **for** $i = 1 : K$ **do**
 - 4: Calculate $\tilde{V}_i(1, \alpha_{-i}^n)$ and $\tilde{V}_i(0, \alpha_{-i}^n)$.
 - 5: Update $\alpha_i^{n+1} = \frac{e^{\tilde{V}_i(1, \alpha_{-i}^n)/\gamma}}{e^{\tilde{V}_i(1, \alpha_{-i}^n)/\gamma} + e^{\tilde{V}_i(0, \alpha_{-i}^n)/\gamma}}$.
 - 6: **end for**
 - 7: **if** $|\alpha^{n+1} - \alpha^n| \leq \varepsilon$ **then**
 - 8: Set $Flag = 1$.
 - 9: **end if**
 - 10: Set $n = n + 1$.
 - 11: **end while**
 - 12: Set $\alpha^* = \alpha^n$.
-

4 LAYER I: OPERATOR PRICING DESIGN

In this section, we study the operator's pricing design (for small-scale networks) in Layer I. The operator designs the pricing scheme based on his anticipation of users' membership selections and network access decisions in Layer II.

Note that the operator can choose different pricing schemes. Two typical examples are the complete price differentiation scheme, where the operator charges different prices on different APs, and the single-pricing scheme, where the operator charges the same price on all APs. The former scheme can achieve a high performance (operator revenue) with the cost of a high computational complexity, while the latter scheme reduces the complexity with the cost of revenue loss. To balance the complexity and performance, we will propose a partial price differentiation scheme, which generalizes both the complete price differentiation scheme and the single-pricing scheme.

4.1 Complete Price Differentiation

We first study the complete price differentiation scenario, where the operator charges different prices on different APs. The operator optimizes his pricing scheme to maximize his total expected revenue in one time period, which is proportional to the network usage of Bills and Aliens. For convenience, we denote the network usage of Bills and Aliens as the *charged* network usage, which brings revenue for the operator directly.

Charged network usage: We first derive the charged network usage on a particular AP $i \in \mathcal{K}_s$ in a particular time

slot, which depends on the set of users that are traveling to AP i . Suppose that a set $\mathcal{K}(i)$ of users (except i) are traveling to AP i in that time slot. That is, the player set in the corresponding network access game is $\mathcal{K}(i)$. The Nash equilibrium of the game is given in (10), from which we can further compute the charged network usage on AP i . Hence, given the user set $\mathcal{K}(i)$, the charged network usage on AP i in one time slot is:

$$\sum_{j \in \mathcal{K}(i) \cap \mathcal{K}_A} \sigma_{j,i}^*(\mathcal{K}(i)) + \sum_{j \in \mathcal{K}(i) \cap \mathcal{K}_S} x_j \cdot \sigma_{j,i}^*(\mathcal{K}(i)),$$

where $\sigma_{j,i}^*$ is user j 's equilibrium network usage on AP i , which is given in Lemmas 1 and 2. The first term denotes the total network usage from Aliens in $\mathcal{K}(i)$, and the second term denotes the total network usage from Bills in $\mathcal{K}(i)$.

Next we derive the charged network usage on AP i in the whole period of T time slots, which is the expected value over all possible membership selection combinations of the other $K - 1$ subscribers (except i).² For notational convenience, we use $\mathbf{x}_{-i} = \{x_j : j \neq i, j \in \mathcal{K}_S\}$ to represent a membership selection combination of other $K - 1$ subscribers, and use \mathcal{X}_{-i} to represent the set of all 2^{K-1} possible membership selection combinations. Then, the expected charged network usage on AP i in the whole time period, denoted as $\bar{\sigma}_i^*(\mathbf{p})$, can be calculated by Eq. (18) on the top of next page, where \mathbf{K}_{-i} is the power set of $\mathcal{K}_U \setminus \{i\}$. The term $\phi(\mathcal{K}(i))$ denotes the probability that a set $\mathcal{K}(i)$ of users are travelling to AP i in a time slot, and is calculated by Eq. (11). The term $\psi(\mathbf{x}_{-i})$ denotes the probability that a particular membership selection combination $\mathbf{x}_{-i} \in \mathcal{X}_{-i}$ occurs, and is calculated by Eq. (16).

Operator's expected revenue: The total expected revenue generated on AP i is proportional to the charged network usage, i.e., $p_i \cdot \bar{\sigma}_i^*(\mathbf{p})$, where p_i is the price per unit connection time charged to Bills and Aliens on AP i . We use α_i^* to represent subscriber i 's probability of choosing to be a Bill at equilibrium. Recall that if subscriber i chooses to be a Linus (with probability $1 - \alpha_i^*$), the operator can obtain all the revenue generated on AP i . If subscriber i chooses to be a Bill (with probability α_i^*), the operator can only obtain $1 - \delta$ of the revenue generated on AP i . Hence, the operator's expected revenue collected on AP i is:

$$h_i(\mathbf{p}) = [1 \cdot (1 - \alpha_i^*) + (1 - \delta) \cdot \alpha_i^*] \cdot p_i \cdot \bar{\sigma}_i^*(\mathbf{p}). \quad (19)$$

Thus, the operator's total expected revenue collected on all APs in the whole time period is

$$H^C(\mathbf{p}) = \sum_{i=1}^K h_i(\mathbf{p}).$$

Based on the above analysis, the operator's complete price differentiation problem can be defined as follows:

Problem 1: Complete Price Differentiation

$$\begin{aligned} \max \quad & H^C(\mathbf{p}) = \sum_{i=1}^K h_i(\mathbf{p}) \\ \text{var:} \quad & p_i \geq 0, \forall i \in \mathcal{K}_S \end{aligned}$$

² User i 's own membership selection does not affect the charged network usage on AP i , as each AP owner uses his private channel exclusively and does not account for the charged network usage.

$$\bar{\sigma}_i^*(\mathbf{p}) = T \cdot \sum_{\mathbf{x}_{-i} \in \mathcal{X}_{-i}} \psi(\mathbf{x}_{-i}) \left[\sum_{\mathcal{K}(i) \in \mathbf{K}_{-i}} \phi(\mathcal{K}(i)) \left(\sum_{j \in \mathcal{K}(i) \cap \mathcal{K}_A} \sigma_{j,i}^*(\mathcal{K}(i)) + \sum_{j \in \mathcal{K}(i) \cap \mathcal{K}_S} x_j \sigma_{j,i}^*(\mathcal{K}(i)) \right) \right]. \quad (18)$$

We can easily compute an effective upper bound for each price variable p_i in the above Problem 1:

$$p_i \leq \bar{p} \triangleq \max_{i \in \mathcal{K}_U} \rho_i \cdot \bar{R}(1), \quad (20)$$

where ρ_i is the network access evaluation of user i (characterizing user i 's valuation of data consumption), and $\bar{R}(1)$ is the average data rate that a user can achieve when accessing a Wi-Fi AP exclusively, and can be calculated by Eq. (3). By Lemma 2, we can easily find that any price p_i larger than \bar{p} will lead to a zero network usage for all Bills and Aliens on AP i , hence a zero revenue for the operator. Thus, we can focus on finding the optimal prices within their respective upper bounds, without affecting the optimality. Moreover, such upper bounds are also very useful for implementing the DYCORS algorithm to solve Problem 1 numerically, which will be discussed soon later.

Unfortunately, it is very challenging to solve Problem 1, as it is a mixed multi-level problem, where continuous and binary variables are coupled with each other in a highly nonlinear manner. Nevertheless, we can explore some useful characteristics of Problem 1, which will help us find the solution numerically.

First, the derivative of the objective function is not readily computable, since we do not have the explicit function relationship of the membership profiles with respect to the prices. Second, given any price vector \mathbf{p} , we can compute the objective value, since the optimal total charged network usage on each AP $\bar{\sigma}_i^*(\mathbf{p})$ is given by Eq. (18), and the equilibrium $\alpha^*(\mathbf{p})$ of Membership Selection Game can be computed by the smoothed best response update algorithm. The unavailability of the derivative information of the objective function makes the use of gradient-based methods impossible, and availability of the objective value makes the use of derivative-free algorithms possible. These factors motivates us to use the derivative-free algorithm [24].

We propose to use a recently developed DYCORS (Dynamically COordinate search using Response Surface models) algorithm [25], which is one of the derivative-free algorithms, to solve Problem 1. Such an algorithm is designed to solve the box-constrained optimization problem. The main idea of the DYCORS algorithm is to build and maintain a surrogate model [26] of the objective function at each iteration, and generate trial solutions by using a dynamic coordinate search strategy. It selects the iterate from a set of random trial solutions obtained by perturbing only a subset of the coordinates of the current best solution, which is helpful in finding the global minimum. Moreover, the probability of perturbing a coordinate decreases as the algorithm reaches the computational budget. If the objective function of Problem 1 is continuous, then the DYCORS algorithm converges to its global optimal solution with probability one. For details of the DYCORS algorithm, please refer to [25].

4.2 Partial Price Differentiation

The complete price differentiation is of high implementation complexity and user aversion. However, if the operator implements the single pricing scheme and ignores the difference among different APs, it may suffer from a high revenue loss. To this end, we propose a partial price differentiation scheme, where the operator charges the same price on APs with similar attributes, to achieve a tradeoff between the revenue and implementation complexity. Note that the complete price differentiation scheme and the single pricing scheme are special cases of the partial price differentiation scheme.

In the partial price differentiation scheme, the operator charges the same price on the APs in the same group (i.e., those with similar attributes), while different prices on different groups. This is analogous to the third-degree price discrimination in economics [27], where prices are set according to user segmentation (based on user attributes such as ages, occupations, and genders).

In our partial price differentiation, we first segment the APs in the network into different groups, based on the attributes, such as location hotness (i.e., the summation of other users' probabilities of roaming to that location, also called location popularity) and AP owners' network access evaluation. We then set one price on the APs in the same group.

4.2.1 AP Segmentation

We assume that the operator segments the set $\mathcal{K}_S = \{1, 2, \dots, K\}$ of APs into G ($G \leq K$) groups. We denote the set of APs in group g ($g = 1, 2, \dots, G$) as set \mathcal{S}_g , and we denote the AP segmentation result, i.e., the set of all G AP sets, as $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_G\}$. The AP segmentation is based on AP attributes. The APs in the same set \mathcal{S}_g have similar attributes.

The AP segmentation problem has several characteristics. First, the number of APs involved in the AP segmentation problem can be very large, e.g., several hundred or several thousand. Second, the AP segmentation is based on AP attributes, such as location hotness and AP owner's network access evaluation. So each AP is described by a multiple dimensional data profile in the AP segmentation problem. Third, different AP attributes can be of different importance in the segmentation problem. The above three characteristics motivate us to use the weighted k-means clustering [28], [29], [30] to deal with the AP segmentation problem.

The k-means clustering is popular for many real world applications such as user segmentation in marketing research [28]. The weighted k-means clustering [29] assigns a weight to each attribute, where the weight measures the importance of the weight in the clustering. Details of the weighted k-means clustering can be found in the appendix.

By applying the weighted k-means clustering, we can obtain the AP segmentation result $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_G\}$.

In Section 6, we provide numerical results to illustrate how the AP attributes and corresponding weights affect the AP segmentation result.

4.2.2 Price Optimization

In the partial price differentiation scheme, the operator sets one price to APs in the same group.

Given the AP segmentation $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_G\}$, the operator's partial price differentiation problem is as follows:

Problem 2: Partial Price Differentiation

$$\begin{aligned} \max \quad & H^P(\mathbf{p}^P) = \sum_{i=1}^K h_i(\mathbf{p}^P) \\ \text{var:} \quad & p_g^P \geq 0, \forall g = 1, 2, \dots, G \end{aligned}$$

Here, $\mathbf{p}^P = \{p_1^P, p_2^P, \dots, p_G^P\}$. The objective function $H^P(\mathbf{p}^P)$ is similar as $H^C(\mathbf{p})$, by replacing p_i in $H^C(\mathbf{p})$ with p_g^P if $i \in \mathcal{S}_g$. Obviously, Problem 2 has the similar structure as Problem 1, hence we can use the DYCORDS algorithm to solve it.

5 APPROXIMATE MODEL AND ANALYSIS FOR LARGE-SCALE SYSTEMS

It is important to note that the complexity of the Stackelberg model (in Sections 3 and 4) increases exponentially with the number of APs. Hence, in large-scale networks with a large number of APs, users may not be able to make the fully rational decisions due to the limited computation abilities. In this section, we will propose and analyze an approximate Stackelberg model for such a large-scale system.

The basic framework of the approximate Stackelberg model is similar as that of the Stackelberg model in Figure 2. Namely, the operator optimizes the pricing scheme in Layer I, and users decide their probabilities of choosing to be Bills/Linues and their network access in Layer II. The key difference is summarized as follows. In a large-scale network, each user may *not* be able to enumerate all possible behaviors of other users and then calculate his payoff by Eqs. (15)-(17) accordingly, due to the limited computation capability of each user. Hence, in the approximate Stackelberg model, each user will behave based on the estimated *expected* behaviors of other users. In the previous Stackelber model, however, each user will behave based on the complete enumeration of all possible behaviors of all other users.

Next, we will analyze the approximate Stackelberg model by backward induction.

5.1 Layer II: User Behavior Analysis

Users' approximate two-stage dynamic game is the same as the one illustrated in Figure 3. Subscribers participate in an approximate membership selection game in Stage I and decide their probabilities of choosing to be Bills/Linus at the beginning of one time period. Users participate in an approximate network access game in Stage II and decide their network access time on the APs in each time slot. We analyze the approximate two-stage dynamic game by backward induction.

5.1.1 Stage II: Approximate Network Access Game on Each AP

The approximate network access game in the approximate Stackelberg model is different from the network access game in the Stackelberg model. In the large-scale system, when a user travels to an AP location, he can neither derive the set of other users who travel to the same AP location, nor those users' precise network access decisions. Hence, we assume that each user only knows other users' *expected* network access decisions.

We define the approximate network access game on AP k in the large-scale system as follows.

Game (Approximate Network Access Game on AP k).

- *Players:* the set $\mathcal{K}_U = \mathcal{K}_S \cup \mathcal{K}_A$ of users in the network;
- *Strategies:* $\tilde{\sigma}_{i,k} \in [0, 1], \forall i \in \mathcal{K}_U$;
- *Payoffs:* $\tilde{v}_{i,k}(\tilde{\sigma}_{i,k}, \tilde{\sigma}_{-i,k}), \forall i \in \mathcal{K}_U$.

The payoff is the difference between the utility and the payment (charged to Bills and Aliens).

The utility of user $i \in \mathcal{K}_U$ on AP k is

$$\tilde{u}_i(\tilde{\sigma}_{i,k}, \tilde{\sigma}_{-i,k}) = \rho_i \log(1 + \tilde{r}_k(\tilde{\sigma}_k) \cdot \tilde{\sigma}_{i,k}). \quad (21)$$

Here $\tilde{r}_k(\tilde{\sigma}_k)$ is the data rate that user i can achieve on AP k . When the number of user in the network is large, we assume that each user will achieve the same data rate $\tilde{r}_k(\tilde{\sigma}_k)$ on AP k , which is a function of all users' network access decisions $\tilde{\sigma}_k = \{\tilde{\sigma}_{i,k}, \forall i \in \mathcal{K}_U\}$:

$$\tilde{r}_k(\tilde{\sigma}_k) = \bar{R} \left(\sum_{i \neq k} \eta_{i,k} \tilde{\sigma}_{i,k} \right). \quad (22)$$

Here $\sum_{i \neq k} \eta_{i,k} \tilde{\sigma}_{i,k}$ represents the expected number of users who access AP k simultaneously in a single time slot, and $\bar{R}(\sum_{i \neq k} \eta_{i,k} \tilde{\sigma}_{i,k})$ calculates (by Eq. (3)) the average data rate of a user when $\sum_{i \neq k} \eta_{i,k} \tilde{\sigma}_{i,k}$ users connect to AP k simultaneously.

Regarding the payoff, if user i is a Linus, his payoff is his utility, since he does not need to pay for his network usage on AP k :

$$\tilde{v}_{i,k}^L(\tilde{\sigma}_{i,k}, \tilde{\sigma}_{-i,k}) = \tilde{u}_i(\tilde{\sigma}_{i,k}, \tilde{\sigma}_{-i,k}). \quad (23)$$

If user i is a Bill or Alien, his payoff is the difference between his utility and payment:

$$\tilde{v}_{i,k}^B(\tilde{\sigma}_{i,k}, \tilde{\sigma}_{-i,k}) = \tilde{u}_i(\tilde{\sigma}_{i,k}, \tilde{\sigma}_{-i,k}) - \tilde{p}_k \tilde{\sigma}_{i,k}. \quad (24)$$

Here \tilde{p}_k is the unit price charged to Bills and Aliens on AP k in the approximate Stackelberg model.

We summarize the payoff of user $i \in \mathcal{K}_U$ in the Approximate Network Access Game (on AP k) in the large-scale system as follows:

$$\tilde{v}_{i,k}(\tilde{\sigma}_{i,k}, \tilde{\sigma}_{-i,k}) = \begin{cases} \tilde{v}_{i,k}^L(\tilde{\sigma}_{i,k}, \tilde{\sigma}_{-i,k}), & \text{if } i \in \mathcal{K}_S \text{ and } x_i = 0; \\ \tilde{v}_{i,k}^B(\tilde{\sigma}_{i,k}, \tilde{\sigma}_{-i,k}), & \text{if } i \in \mathcal{K}_S \text{ and } x_i = 1; \\ \tilde{v}_{i,k}^B(\tilde{\sigma}_{i,k}, \tilde{\sigma}_{-i,k}), & \text{if } i \in \mathcal{K}_A. \end{cases} \quad (25)$$

We derive the best response of user $i \in \mathcal{K}_U$ as follows.

Lemma 4. *If user i is a Linus, his best response in the Approximate Network Access Game on AP k in the large-scale system is*

$$\tilde{\sigma}_{i,k}^{L*} = 1, \quad (26)$$

regardless of other users' strategies.

Lemma 5. *If user i is a Bill or an Alien, his best response in the Approximate Network Access Game on AP k in the large-scale system is*

$$\tilde{\sigma}_{i,k}^{B*} = \min \left\{ 1, \max \left\{ \frac{\rho_i}{\tilde{p}_k} - \frac{1}{\tilde{r}_k(\tilde{\sigma}_k)}, 0 \right\} \right\}, \quad (27)$$

which is a function of users' strategy profile $\tilde{\sigma}_k$.

Regarding the existence of the Nash equilibrium, we have the same conclusion as that in the network access game of the Stackelberg model, i.e., Theorem 1 in Section 3.1. We can use the same best response update algorithm in Section 3.1 to derive the Nash equilibrium for the approximate network access game.

5.1.2 Stage I: Approximate Membership Selection Game

Now we study the approximate membership selection game in the large-scale system.

The players are the set \mathcal{K}_s of subscribers in the network. Each subscriber decides his probability of choosing to be a Bill. Because of the limited computation capability, users are bounded rational. Each user aims to maximize his *perceived* payoff which is based on the *expected* behaviors of other users.

We define the approximate membership selection game in the large-scale system as follows.

Game (Approximate Membership Selection Game).

- *Players:* the set \mathcal{K}_s of subscribers.
- *Strategies:* $\tilde{\alpha}_i \in [0, 1], \forall i \in \mathcal{K}_s$.
- *Payoffs:* $\tilde{\omega}_i(\tilde{\alpha}_i, \tilde{\alpha}_{-i}), \forall i \in \mathcal{K}_s$.

The perceived expected payoff of subscriber i choosing to be a Bill with probability $\tilde{\alpha}_i$ is:

$$\tilde{\omega}_i(\tilde{\alpha}_i, \tilde{\alpha}_{-i}) = (1 - \tilde{\alpha}_i)\tilde{V}_i^L(0, \tilde{\alpha}_{-i}) + \tilde{\alpha}_i\tilde{V}_i^B(1, \tilde{\alpha}_{-i}), \quad (28)$$

where $\tilde{V}_i^L(0, \tilde{\alpha}_{-i})$ is the perceived payoff when subscriber i chooses to be a Linus, and $\tilde{V}_i^B(1, \tilde{\alpha}_{-i})$ is the perceived payoff when subscriber i chooses to be a Bill. Due to the limited computation capability of subscriber i , subscriber i can not calculate $\tilde{V}_i^L(0, \tilde{\alpha}_{-i})$ and $\tilde{V}_i^B(1, \tilde{\alpha}_{-i})$ based on the enumeration of all possible membership selections of other users, as in Eq. (17). The calculation of $\tilde{V}_i^L(0, \tilde{\alpha}_{-i})$ and $\tilde{V}_i^B(1, \tilde{\alpha}_{-i})$ is based on the *expected* behaviors of other users perceived by subscriber i .

Specifically, if subscriber i chooses to be a Linus, his perceived payoff $\tilde{V}_i^L(0, \tilde{\alpha}_{-i})$ includes his perceived utility on all possible APs that he may travel to during the whole time period.

When Linus i stays at home and access his own AP, his perceived payoff is:

$$\tilde{V}_{i,i}^L(0, \tilde{\alpha}_{-i}) = \rho_i \log(1 + \tilde{r}_{i,i} \cdot 1).$$

Recall that $\tilde{r}_{i,i}$ corresponds to the average data rate that Linus i achieves at his private channel. When Linus i travels to AP k , his perceived payoff of connecting to AP k is:

$$\tilde{V}_{i,k}^L(0, \tilde{\alpha}_{-i}) = \rho_i \log(1 + \tilde{r}_k(\tilde{\sigma}_k) \cdot 1).$$

When Linus i travels to an area that is not covered by any of the K Wi-Fi APs, his perceived payoff is:

$$\tilde{V}_{i,0}^L(0, \tilde{\alpha}_{-i}) = 0.$$

Based on the above, the perceived payoff of Linus i is

$$\tilde{V}_i^L(0, \tilde{\alpha}_{-i}) = T \sum_{k=0}^K \eta_{i,k} \tilde{V}_{i,k}^L(0, \tilde{\alpha}_{-i}). \quad (29)$$

If subscriber i chooses to be a Bill, his perceived payoff $\tilde{V}_i^B(1, \tilde{\alpha}_{-i})$ includes his payoff on all possible APs that he may travel to during the whole time period and the expected revenue that subscriber i collects on his own AP.

Similarly, when Bill i stays at home and access his own AP, his perceived payoff is:

$$\tilde{V}_{i,i}^B(1, \tilde{\alpha}_{-i}) = \rho_i \log(1 + \tilde{r}_{i,i} \cdot 1).$$

When Bill i travels to AP k , his perceived payoff of connecting to AP k is:

$$\tilde{V}_{i,k}^B(1, \tilde{\alpha}_{-i}) = \rho_i \log(1 + \tilde{r}_k(\tilde{\sigma}_k)\tilde{\sigma}_{i,k}^{B*}) - \tilde{p}_k\tilde{\sigma}_{i,k}^{B*},$$

where $\tilde{\sigma}_{i,k}^{B*}$ is calculated by Eq. (27). When Bill i travels to an area that is not covered by any of the K Wi-Fi APs, his perceived payoff is:

$$\tilde{V}_{i,0}^B(1, \tilde{\alpha}_{-i}) = 0.$$

The expected revenue that Bill i collects on his own AP is a δ portion of the payment of all Aliens and Bills, which can be calculated as:

$$\begin{aligned} & \tilde{\Pi}_i^B(1, \tilde{\alpha}_{-i}) \\ &= T\delta\tilde{p}_i \left(\sum_{j \in \mathcal{K}_v \cap \mathcal{K}_a} \eta_{j,i} \tilde{\sigma}_{j,i}^{B*} + \sum_{j \in \mathcal{K}_v \cap \mathcal{K}_s, j \neq i} \eta_{j,i} \tilde{\alpha}_j \tilde{\sigma}_{j,i}^{B*} \right). \end{aligned}$$

Based on the above, the perceived payoff of Bill i is

$$\tilde{V}_i^B(1, \tilde{\alpha}_{-i}) = \tilde{\Pi}_i^B(1, \tilde{\alpha}_{-i}) + T \sum_{k=0}^K \eta_{i,k} \tilde{V}_{i,k}^B(1, \tilde{\alpha}_{-i}). \quad (30)$$

Regarding the existence of the Nash equilibrium, we have the same conclusions in Theorem 2 as the membership selection game in the Stackelberg model, in Section 3.2. We can use the same smoothed best response update algorithm in Section 3.2 to derive the Nash equilibrium for the approximate membership selection game.

5.2 Layer I: Operator Pricing Design

In this section, we study the operator's pricing design in Layer I of the approximate Stackelberg model. Similarly, we propose a partial price differentiation scheme.

In the partial price differentiation scheme, the operator first segment the APs in the network into different groups, based on AP attributes, as in Section 4.2.1. We denote the set of APs in group g ($g = 1, 2, \dots, G$) as set $\tilde{\mathcal{S}}_g$, and we denote the AP segmentation result as $\tilde{\mathcal{S}} = \{\tilde{\mathcal{S}}_1, \tilde{\mathcal{S}}_2, \dots, \tilde{\mathcal{S}}_G\}$.

Then the operator sets one price to APs in the same group. The operator's goal is to maximize the total revenue he collected on all APs in the network:

$$\tilde{H}(\tilde{\mathbf{p}}^P) = \sum_{i=1}^K \tilde{h}_i(\tilde{\mathbf{p}}^P).$$

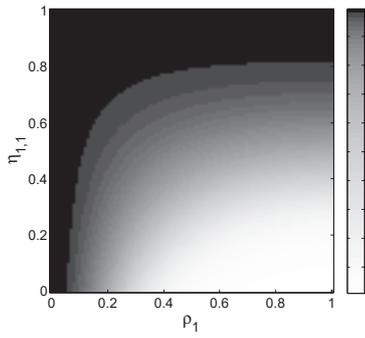


Figure 5. Membership Decision under Different Parameters (Color represents subscriber 1's probability of choosing to be a Bill.)

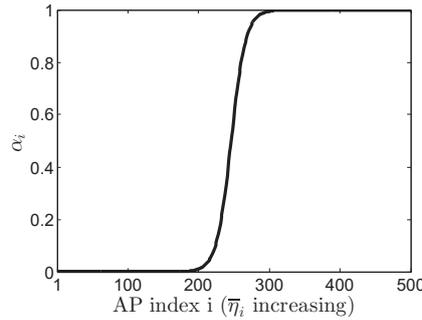


Figure 6. Membership Decisions under Different Location Popularities

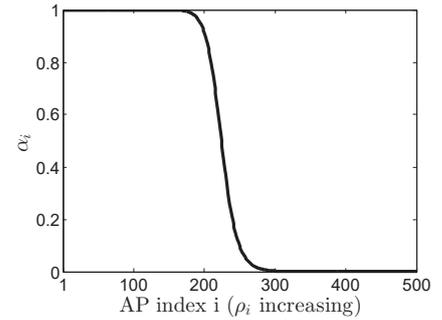


Figure 7. Membership Decisions under Different Network Access Evaluations

Here $\tilde{h}_i(\tilde{\mathbf{p}}^P)$ is the expected revenue collected by the operator on AP i , which can be written as:

$$\tilde{h}_i(\tilde{\mathbf{p}}^P) = (1 \cdot (1 - \tilde{\alpha}_i^*) + (1 - \delta) \tilde{\alpha}_i^*) \tilde{p}_g^P \tilde{\sigma}_i^*(\tilde{\mathbf{p}}^P), \quad (31)$$

Here $i \in \tilde{\mathcal{S}}_g$, and $\tilde{\sigma}_i^*(\tilde{\mathbf{p}}^P)$ is the expected *charged* network usage from all Bills (except subscriber i) and Aliens at equilibrium on AP i :

$$\tilde{\sigma}_i^*(\tilde{\mathbf{p}}^P) = T \tilde{p}_g^P \left(\sum_{j \in \mathcal{K}_v \cap \mathcal{K}_A} \eta_{j,i} \tilde{\sigma}_{j,i}^{B*} + \sum_{j \in \mathcal{K}_v \cap \mathcal{K}_s, j \neq i} \eta_{j,i} \tilde{\alpha}_j \tilde{\sigma}_{j,i}^{B*} \right)$$

Given the AP segmentation $\tilde{\mathcal{S}} = \{\tilde{\mathcal{S}}_1, \tilde{\mathcal{S}}_2, \dots, \tilde{\mathcal{S}}_G\}$, the operator's partial price differentiation problem in the approximate Stackelberg model can be written as follows:

Problem 3: Approximate Partial Price Differentiation

$$\begin{aligned} \max \quad & \tilde{H}(\tilde{\mathbf{p}}^P) = \sum_{i=1}^K \tilde{h}_i(\tilde{\mathbf{p}}^P) \\ \text{var:} \quad & \tilde{p}_g^P \geq 0, \forall g = 1, 2, \dots, G \end{aligned}$$

Problem 3 has similar structure as Problem 2. Hence, we propose to use DYCORDS algorithm to solve Problem 3.

6 SIMULATION RESULTS

We provide simulation results to illustrate the users' behaviors in Section 6.1 and evaluate the operator's revenue in Section 6.2.

6.1 User Behaviors

Regarding user behaviors, we numerically study how the network access valuation parameter ρ_i and the mobility pattern η_i affect subscriber i 's membership selection decision, given other system parameters fixed. In what follows, we will first simulate a small network with 2 APs (subscribers) and 1 Alien, to gain insights of a single user's best choice. Then, we will simulate a large network with 500 APs and 500 Aliens to understand the system-level performance.

6.1.1 A Small Network Example

We simulate a small network with 2 subscribers (each owns an AP) and 1 Alien. We study how subscriber 1's network access valuation parameter ρ_1 and his probability of staying at home $\eta_{1,1}$ affect his membership selection.

We assume that the revenue sharing ratio $\delta = 0.5$. Both APs have the same price $p = 1$. In Section 6.2, we will consider the operator's price optimization problem. The mobility patterns of subscriber 2 and the Alien are the same: $\eta_2 = \eta_a = [1/3, 1/3, 1/3]$. We assume that $\rho_1 \in [0, 1]$. Subscriber 1 stays at home with probability $\eta_{1,1}$, and travels to AP 2 and outside the Wi-Fi coverage with a same probability $\eta_{1,2} = \eta_{1,0} = (1 - \eta_{1,1})/2$.

Figure 5 shows subscriber 1's membership selection decision in the equilibrium (Definition 5 in Section 3.2.4), under different values of $\rho_i \in [0, 1]$ and $\eta_{1,1} \in [0, 1]$. The color represents the value of α_1 , which is subscriber 1's probability of choosing to be a Bill. The black region corresponds to $\alpha_1 = 1$, and the white region corresponds to $\alpha_1 = 0$. The color in between corresponds to a mixed strategy of $\alpha_1 \in (0, 1)$, as shown in the colorbar on the right.

Figure 5 shows that when $\eta_{1,1}$ is large enough (i.e., larger than 0.82), i.e., subscriber 1 stays at home most of the time, he will always choose to be a Bill (with the probability $\alpha_1 = 1$), independent of subscriber 2's membership decision. As $\eta_{1,1}$ becomes smaller and ρ_1 becomes larger, the performance and payment during roaming becomes increasingly important, so subscriber 1 starts to choose a mixed strategy with a smaller number of α_1 . When $\eta_{1,1}$ is small enough and ρ_1 is large enough, e.g., the right bottom corner of Figure 5, he will always choose to be a Linus with a probability $1 - \alpha_1 = 1$.

6.1.2 Simulation Results for Large Network

Next, we simulate a larger network with 500 APs and 500 Aliens. We assume the price $p = 1$ and the revenue sharing ratio $\delta = 0.5$. We study how the system parameters, i.e., the location popularity and the network access evaluation, affect subscribers' membership decisions.

We first study how the location popularity of an AP affects the subscriber's membership selection decision. We assume that each subscriber's network access valuation parameter $\rho_i = 1, \forall i \in \mathcal{K}_s$. Since Aliens do not have APs, we assume that Aliens value the network access more, and $\rho_i = 5, \forall i \in \mathcal{K}_A$. We assume that all users show up at a particular AP, e.g., AP i ($i \in \mathcal{K}_s$), with the same probability $\bar{\eta}_i$, i.e.,

$$\eta_{j,i} = \bar{\eta}_i, \forall j \in \mathcal{K}_U.$$

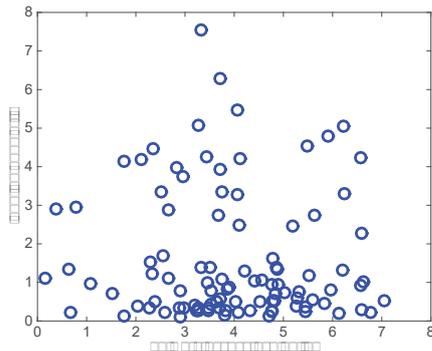


Figure 8. Subscribers' Network Access Evaluation and Location Popularity Parameters

We denote $\bar{\eta}_0$ as the probability of users showing up at areas that are not covered by any WiFi AP. Obviously, we have:

$$\sum_{i=0}^K \bar{\eta}_i = 1.$$

For simplicity, we assume that the location popularity $\bar{\eta}_i$ increases with the AP index i .

Figure 6 shows each of the 500 subscribers' membership selection decision. Simulation result shows that the subscriber's probability of choosing to be a Bill increases with his AP location popularity. The reason is that a subscriber whose AP is located at a more popular location can earn more revenue from other Bills and Aliens.

We then study how the network access evaluation of a subscriber affects his membership selection decision. We assume that the probability of each user showing up at each location is uniform and same. We assume that the subscribers' network access evaluation parameters ρ_i increases with the AP index i .

Figure 7 shows each of the 500 subscribers' membership selection decision. Simulation result shows that the subscriber's probability of choosing to be a Bill decreases with his network access evaluation. This is because he cares more about the network access benefit when roaming, and hence is more willing to be a Linus to enjoy free access and consume more data during roaming.

6.2 The Operator's Revenue

Regarding the operator's revenue, we numerically study the performance of the partial price differentiation scheme in a large network. We consider a network which consists of 100 APs and 100 Aliens.

We assume that subscribers' network access evaluation parameters $\rho_i, \forall i \in \mathcal{K}_s$ follow the Gaussian distribution with a mean value of 4 and a variance of 2. We obtain the APs' location popularity $\bar{\eta}_i = \sum_{j \neq i} \eta_{j,i}$ ($\forall i \in \mathcal{K}_s$) based on the realistic mobility data in [31]. Figure 8 shows the relationship between network access evaluation parameter and location popularity parameter.

We study the performance of the partial price differentiation scheme. We first use the weighted k-means clustering algorithm to segment APs into different groups considering the attributes of network access evaluation and location popularity. We assign a weight β to the attribute of network access evaluation and $1 - \beta$ to the attribute of location popularity in the weighted k-means clustering algorithm. The AP

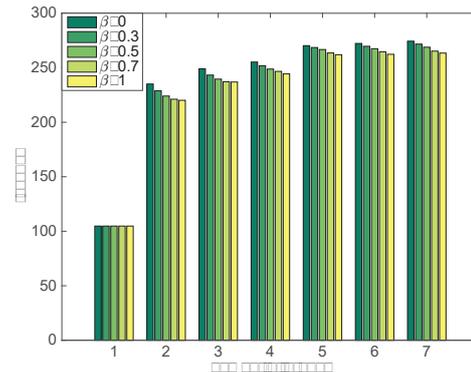


Figure 9. Operator's Revenue under the Partial Price Differentiation

segmentation results under different weight β are provided in the appendix in the supplemental material section. Then we apply the price differentiation scheme and optimize over the prices to maximize the operator's revenue.

Figure 9 shows the operator's revenue under different number of groups and different weight β . Given a particular weight β , we can see that (i) the operator's revenue increases with the number of groups; and (ii) the revenue increase rate slows down as the number of groups increase. Given a particular number of groups, the revenue difference under different β is small. Compared with the single pricing scheme, the two-price scheme can effectively increase the operator's revenue by up to 124.44%. However, the revenue increase from the 6-price scheme to the 7-price scheme is only 0.78%. Hence, the operator can achieve a good tradeoff of total revenue and implementation complexity, by applying the partial price differentiation scheme.

7 CONCLUSION

In this paper, we analyze the user behavior and the operator pricing design for the crowdsourced wireless community network. We model the interactions between the operator and users as a Stackelberg model where users are fully rational in small-scale systems. For the large-scale systems where users are bounded rational due to the limited computation capability, we propose and analyze an approximate Stackelberg model. We study the user behaviors under both full rationality and bounded rationality systematically. We show that a user with a more popular home location, a smaller probability of travelling, or a smaller network access evaluation is more likely to choose to be a Bill. Moreover, we propose a partial price differentiation scheme for the operator, based on the analysis of users response to the pricing scheme. Our numerical results show that the proposed partial price differentiation scheme with only two prices can increase the operator's revenue up to 124.44% comparing with the single pricing scheme, and can achieve an average of 80% of the maximum operator revenue under the complete price differentiation scheme.

There are several interesting and important extensions for the model in this work. For example, it is important to study the model under information asymmetry, where users have some private information, e.g., network access valuation ρ and user mobility pattern η . In that case, the operator needs to further design an incentive compatible mechanism (e.g., auction) to elicit the users' private information.

REFERENCES

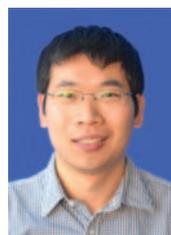
- [1] Q. Ma, L. Gao, Y.-F. Liu, and J. Huang, "A game-theoretic analysis of user behaviors in crowdsourced wireless community networks," *The 13th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, 2015.
- [2] Cisco, "Cisco visual networking index: global mobile data traffic forecast update, 2014-2019," Feb. 2015.
- [3] Wikipedia: <http://en.wikipedia.org/wiki/802.11>
- [4] Wikipedia: <http://en.wikipedia.org/wiki/Wi-Fi>
- [5] Wikipedia: http://en.wikipedia.org/wiki/List_of_wireless_community_networks_by_region
- [6] G. Camponovo and A. Picco Schwendener, "Motivations of hybrid wireless community participants: a qualitative analysis of Swiss FON members," *10th International Conference on Mobile Business*, 2011.
- [7] FON website: <https://corp.fon.com/en>
- [8] La fonera user manual [online]. Available: http://misc.nybergh.net/pub/M2100_1_01.pdf
- [9] FON passes: <https://support.fon.com/entries/22854577-Fon-Passes>
- [10] S. Li, and J. Huang, "Price differentiation for communication networks," *IEEE/ACM Transactions on Networking (TON)*, pp. 703–716, 2014.
- [11] M.E. Bratman, D.J. Israel, and M.E. Pollack, "Plans and resource-bounded practical reasoning," *Computational intelligence*, vol. 4, no. 3, pp. 349–355, 1988.
- [12] H. A. Simon, "Bounded rationality and organizational learning," *Organization science*, vol. 2, no. 1, pp. 125–134, 1991.
- [13] M.H. Manshaei, et al., "On wireless social community networks," *IEEE INFOCOM*, 2008.
- [14] G. Biczok, L. Toka, A. Vidacs, and T.A. Trinh, "On incentives in global wireless communities," *ACM workshop on User-provided networking: challenges and opportunities*, 2009.
- [15] M.H. Afrasiabi and R. Guerin, "Pricing strategies for user-provided connectivity services," *IEEE INFOCOM*, 2012.
- [16] A. Mazlounian and M.H. Manshaei, "Optimal pricing strategy for wireless social community networks," *International workshop on Economics of networked systems*, 2008.
- [17] FON website: <https://support.fon.com/hc/en-us/articles/200535292-How-do-I-set-up-the-Fonera-via-its-web-interface>
- [18] C. Courcoubetis and R. Weber, "Pricing communication networks: economics, technology and modelling," *Wiley Press*, 2003.
- [19] A. Mas-Colell, M. D. Whinston, and J. R. Green, "Microeconomic theory," *Oxford University Press*, 1995.
- [20] M. Chiang, "Networked life: 20 questions and answers," *Cambridge University Press*, 2012.
- [21] Wikipedia: http://en.wikipedia.org/wiki/Brouwer_fixed-point_theorem
- [22] R. Gibbons, "Game theory for applied economists," *Princeton University Press*, 1992.
- [23] D. Fudenberg and David K Levine, "The theory of learning in games," *MIT Press*, 1998.
- [24] L.M. Rios, and N.V. Sahinidis, "Derivative-free optimization: a review of algorithms and comparison of software implementations," *Journal of Global Optimization*, vol. 56, no. 3, 2013.
- [25] R.G. Regis, and C.A. Shoemaker, "Combining radial basis function surrogates and dynamic coordinate search in high-dimensional expensive black-box optimization," *Engineering Optimization*, vol. 45, no. 5, 2013.
- [26] N.V. Queipo, R.T. Haftka, W. Shyya, et al., "Surrogate-based analysis and optimization," *Progress in aerospace sciences*, vol. 41, no.1, 2005.
- [27] H.R. Varian, "Price discrimination," *Handbook of Industrial Organization*, pp. 597–654, 1989.
- [28] P.E. Green, K. Jonathan, and J.C. Frank, "A preliminary study of optimal variable weighting in k-means clustering," *Journal of Classification*, vol. 7, no.2, pp. 271–285, 1990.
- [29] J.Z. Huang, M.K. Ng, H. Rong, and Z. Li, "Automated variable weighting in k-means type clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.27, no.5, pp. 657–668, 2005.
- [30] J. MacQueen, "Some methods for classification and analysis of multivariate observations," *Proc. Fifth Berkeley Symp. Math. Statistics and Probability*, vol.1, no.14, pp. 281–297, 1967.
- [31] J. Ghosh, M. J. Beal, H. Q. Ngo, and C. Qiao, "On profiling mobility and predicting locations of campus-wide wireless network users," *State University of New York at Buffalo, Tech. Rep.*, 2005.



Qian Ma (S'13) is a Ph.D. student in the Department of Information Engineering at the Chinese University of Hong Kong. She received the B.S. degree from Beijing University of Posts and Telecommunications (China) in 2012. Her research interests lie in the field of wireless communications and network economics. She is the recipient of the Best Student Paper Award from the IEEE International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt) in 2015.



Lin Gao (S'08-M'10) is an Associate Professor at Harbin Institute of Technology (Shenzhen), China. He received M.S. and Ph.D. degrees in Electronic Engineering from Shanghai Jiao Tong University (China), in 2006 and 2010, respectively. He was a Postdoctoral Fellow at The Chinese University of Hong Kong from 2010 to 2015. His research interests are in the interdisciplinary area combining telecommunications and microeconomics, with particular focus on the game-theoretic and economic analysis for various communication and network scenarios, including cognitive radio networks, TV white space networks, heterogeneous networks, cooperative communications, 5G communications, mobile crowdsensing, and mobile Internet.



Ya-Feng Liu (M'12) is an Assistant Professor of the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, where he received the Ph.D degree in Computational Mathematics in 2012. His main research interests are nonlinear optimization and its applications to signal processing, wireless communications, and machine learning. He is a recipient of the Best Paper Award from the IEEE International Conference on Communications (ICC) in 2011 and the Best Student Paper Award from the International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt) in 2015.



Jianwei Huang (S'01-M'06-SM'11-F'16) is an Associate Professor and Director of the Network Communications and Economics Lab (ncl.ie.cuhk.edu.hk), in the Department of Information Engineering at the Chinese University of Hong Kong. He received the Ph.D. degree from Northwestern University in 2005. He is the co-recipient of 8 international Best Paper Awards, including IEEE Marconi Prize Paper Award in Wireless Communications in 2011. He has co-authored five books and six ESI highly cited papers. He has served as an Editor of several top IEEE Communications journals, including JSAC, TWC, and TCCN. He is an IEEE Fellow and a Distinguished Lecturer of IEEE Communications Society.