Monetizing Mobile Data via Data Rewards

Haoran Yu, *Member, IEEE,* Ermin Wei, *Member, IEEE,* and Randall A. Berry, *Fellow, IEEE*

***Abstract*—Most mobile network operators generate revenues by directly charging users for data plan subscriptions. Some operators now also offer users data rewards to incentivize them to watch mobile ads, which enables the operators to collect payments from advertisers and create new revenue streams. In this work, we analyze and compare two data rewarding schemes: a *Subscription-Aware Rewarding (SAR) scheme* and a *Subscription-Unaware Rewarding (SUR) scheme*. Under the SAR scheme, only the subscribers of the operators' data plans are eligible for the rewards; under the SUR scheme, all users are eligible for the rewards (e.g., the users who do not subscribe to the data plans can still get SIM cards and receive data rewards by watching ads). We model the interactions among an operator, users, and advertisers by a two-stage Stackelberg game, and characterize their equilibrium strategies under both the SAR and SUR schemes. We show that the SAR scheme can lead to more subscriptions and a higher operator revenue from the data market, while the SUR scheme can lead to better ad viewership and a higher operator revenue from the ad market. We further show that the operator's optimal choice between the two schemes is sensitive to the users' data consumption utility function and the operator's network capacity. We provide some counter-intuitive insights. For example, when each user has a logarithmic utility function, the operator should apply the SUR scheme (i.e., reward both subscribers and non-subscribers) if and only if it has a small network capacity.**

***Index Terms*—Stackelberg game, network economics, mobile data rewards, business model.**

I. Introduction

Despite the rapid growth of global mobile traffic, several leading analyst firms estimate that global mobile service revenue has nearly reached a saturation point. For example, Strategy Analytics forecasts that the global mobile service revenue will only increase by 3% between 2018 and 2021 [2]. As suggested in [3], one promising approach for the mobile network operators to create new revenue streams is to offer *mobile data rewards*: the network operators reward users with free mobile data every time the users watch mobile ads delivered by the operators, and the operators are paid by the corresponding advertisers.

The data rewarding paradigm leads to a "win-win-win" outcome [3]. First, the operators monetize their services based on the mobile advertising, the global revenue of which was estimated to reach \$80 billion at the end of 2017 [3]. Second, the advertisers gain *incentivized advertising*, where the rewards incentivize the users to better engage with ads and the advertisers allow the users to have more control over their experiences (e.g., whether and when to watch ads). According to surveys conducted by Forrester Consulting, IPG Media Lab, and Kiip, most mobile app users prefer to watch ads with rewards than to watch targeted ads [4]. Third, the users earn free mobile data to satisfy their growing data demand.

There has been an increasing number of businesses entering this space. Aquto and Unlockd are two leading companies that provide technical support for data rewarding (e.g., they develop mobile apps that display ads and track the amount of rewarded data). Aquto has collaborated with operators, such as Verizon and Telefonica [5]. Unlockd has collaborated with Tesco Mobile (in the United Kingdom), Boost Mobile (in the United States), Lebara Mobile (in Australia), and AXIS (in Indonesia) [6]. Other examples of operators that have offered data rewards include DOCOMO, Optus, and ChungHwa Telecom [7], [8]. Furthermore, AT&T recently acquired AppNexus (a leading online advertising company) and will make a significant investment in the advertising business [9]. Offering mobile data rewards could become a natural and effective approach to further monetize an operator's mobile service.

We use an example in Table I to show that offering data rewards might lead to a significant revenue improvement for an operator. Suppose that an operator rewards 0.5MB of data per image ad.[1] If a user watches 40 image ads every day, it can get 600MB of data after 30 days. When the CPM (cost per thousand impressions, also called cost per mille) is \$8.2 [11], the operator's corresponding ad revenue is \$9.84. In other words, the operator gets \$9.84 by rewarding 600MB of data to the user. As a comparison, the conventional data pricing is less profitable to the operator. As shown in [12], operators only charge a user an extra \$4 when the user switches from a 1GB data plan to a 2GB data plan.

Based on the eligibility of receiving rewards, there are two basic types of data rewarding schemes. In the *Subscription-Aware Rewarding (SAR) scheme*, the operators only allow the users who subscribe to the operators' existing data plans (with monthly fees) to watch ads for rewards. In the *Subscription-Unaware Rewarding (SUR) scheme*, the operators reward all users for watching ads, regardless of whether the users subscribe to the data plans.[2] Intuitively, the SAR scheme leads to more subscriptions and the SUR scheme incentivizes more users to watch ads. The optimal design and comparison of the two schemes are crucial for realizing the full potential of the mobile data rewards, which motivates our work.

Manuscript received September 1, 2019; accepted November 9, 2019. This research was supported by the NSF grants AST-1343381, AST-1547328 and CNS-1701921. Some results in the paper were presented at IEEE INFOCOM, Paris, France, April 2019 [1]. (*Corresponding author: Haoran Yu.*)

The authors are with the Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208, USA (email: yhrhawk@gmail.com; {ermin.wei, rberry}@northwestern.edu).

[1] To ensure that users carefully watch the ads, the operator can ask ad-related questions before giving the rewards [10].

[2] The operators can offer free specialized SIM cards to the users who do not subscribe to the data plans. These users can top up the cards by watching ads, as shown in [7].

TABLE I: Example of Data Rewards

Rewarding Plan	A User's Views and Reward (Per Month)		Calculation of Operator's Ad Revenue	
	Views	Reward	CPM	Views/1000×CPM=Ad Revenue
0.5MB per image ad	1200 image ads	600MB	\$8.2	1200/1000×\$8.2=\$9.84

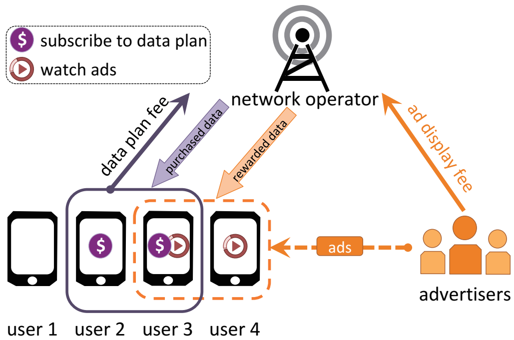


Fig. 1: Data rewarding ecosystem (user 4 is feasible under the *SUR* scheme, but is infeasible under the *SAR* scheme).

A. Our Contributions

We illustrate the data rewarding ecosystem in Fig. 1. The purple arrows indicate that an operator charges the users for data plan subscriptions. The orange arrows indicate that the operator rewards the users for watching ads and gets payments from the advertisers.

We model the interactions among the operator, users, and advertisers by a two-stage Stackelberg game. In Stage I, the operator decides the unit data reward (i.e., the amount of data rewarded for watching one ad) for the users, and the ad price (i.e., the payment for purchasing one ad slot) for the advertisers. In Stage II, the users with different valuations for the mobile service make their data plan subscription and ad watching decisions. We consider a general data consumption utility function and a general distribution of user valuation. Meanwhile, the advertisers decide the number of ad slots to purchase, considering the advertising's *wear-out effect* (i.e., an ad's effectiveness can decrease if it reaches a user who has watched the same ad for several times [13], [14]).

We analyze the two-stage game for both the SAR and SUR schemes. In particular, we characterize the operator's optimal strategy that maximizes the total revenue from the data market and ad market. Our key findings in this work are as follows.

I. Design of Unit Data Reward (Theorems 2 and 3): *Under both the SAR and SUR schemes, the operator should not always use up the available network capacity for data rewards.* Under the SAR scheme, increasing the unit data reward can lead to more data plan subscriptions and motivate more users to watch ads. However, it also allows a user to obtain a larger amount of data after watching a few ads. Hence, a user may watch fewer ads under a larger unit data reward. As a result, increasing the unit data reward may decrease the operator's revenue. Under the SUR scheme, (besides the above negative impact) increasing the unit data reward may lead to a loss in data plan subscriptions, and even generate a revenue that is lower than the revenue when the operator does not offer any data reward. In our work, we derive two sufficient conditions, under which the operator does and does not use up the capacity for data rewards, respectively.

II. Design of Ad Price (Theorems 1 and 4): *Given the unit data reward, the operator's optimal ad price is affected by the wear-out effect if and only if the wear-out effect is small.* If the wear-out effect is small, the operator should sell all ad slots and its optimal ad price should decrease with the wear-out effect; otherwise, the operator should not sell all ad slots and its optimal ad price will be independent of the wear-out effect. Moreover, *under the SUR scheme, the operator can differentiate the ad slots generated by the subscribers and non-subscribers when selling the ad slots to the advertisers and displaying the ads to the users. We numerically show that this can improve the operator's total revenue by up to* 20.3%. Under the SUR scheme, both the subscribers and non-subscribers watch ads. Since the subscribers also obtain data from the data plan, the subscribers and non-subscribers may watch different numbers of ads. Because of the advertising's wear-out effect, each advertiser has a different willingness to purchase the ad slots generated by the subscribers and non-subscribers, and it is beneficial for the operator to differentiate these ad slots.

III. Choice of Rewarding Scheme (Theorem 5; Observations 1, 2, and 3): *The operator's choice between the SAR and SUR schemes is heavily affected by the users' data consumption utility function and network capacity.* When each user has a logarithmic utility function or each user has a generalized α-fair utility function [15], if the network capacity is limited, the operator should apply the SUR scheme (i.e., reward both subscribers and non-subscribers); if the capacity is large, it should apply the SAR scheme (i.e., only reward the subscribers). When each user has an exponential utility: (i) under a large wear-out effect, the choice between the two schemes is similar to the logarithmic utility case; (ii) under a small wear-out effect, the operator should always apply the SUR scheme, regardless of the capacity.

B. Related Work

1) Provision of Fee-Based and Ad-Based Services: There has been some work studying markets where providers offer both a fee-based service and an ad-based free service. In [16], a Wi-Fi network provider allows users to either directly pay or watch ads to access the Wi-Fi network. In [17], an app developer offers virtual items, and each app user will either pay or watch ads to obtain them in the equilibrium. In these studies, the fee-based and ad-based services are always *substitutes*, and each user chooses between these two options. In our work, their relation is more complicated, since a user may subscribe to the data plan and meanwhile watch ads for more data. Under the SAR scheme, increasing the reward for watching ads can increase the number of subscribers, which shows the *complementary* relation between the subscription and data rewards. Therefore, our work studies a novel structure, and derives new insights for the joint provision of fee-based and ad-based services.

2) Sponsored Mobile Data: As studied in [15], [18]–[20], sponsored data provides another way for operators to create new revenue streams: content providers sponsor the data usage of their content, and users can access the content free of charge. There are several key differences between sponsored data and data rewards as studied here. First, the users can consume sponsored data only for the content specified by the content providers, while they can use reward data to access any online content. Second, with sponsored data, the content providers benefit from the users' data consumption on the corresponding content. With data rewards, the advertisers aim to deliver ads effectively, and do not benefit from the users' data consumption.

3) Other Related References: Other related work includes [21]–[23]. Bangera *et al.* in [21] conducted a survey, which shows that 76% of the respondents are interested in watching ads in exchange for mobile data. Sen *et al.* in [22] conducted an experiment to study the effectiveness of monetary rewards in increasing ads' viewership. Both [21] and [22] did not analyze the equilibrium strategies of the entities, such as operators, advertisers, and users. Harishankar *et al.* in [23] studied monetizing the operator's idle network capacity by providing users with supplemental discount offers, which are not related to advertising.

II. MODEL

In this section, we model the strategies of the operator, users, and advertisers, and introduce the two-stage game. We use capital letters to denote parameters, and lower-case letters to denote decision variables or random variables.

A. Network Operator

We consider a monopolistic operator, who offers a predetermined (monthly) flat-rate data plan (F,Q) to users. Parameter $F>0$ denotes the subscription fee, and $Q>0$ denotes the data amount associated with a subscription. To derive insights into the data reward design, we focus on a single-operator, single-data plan scenario, which has been widely considered in literature (e.g., [15], [20]).

The operator decides two variables: (i) a unit data reward $\omega\in[0,\infty)$, which is the amount of data that a user receives for watching one ad; (ii) an ad price $p\in(0,\infty)$, which is the price that the operator charges the advertisers for buying one ad slot. Here, we consider a price-based mechanism, where the operator sells the ad slots in advance at a fixed price.

B. Users

We consider a continuum of users, and denote the mass of users by N. Let θ denote a user's type, which parameterizes its valuation for mobile service. We assume that θ is a continuous random variable drawn from $[0,\theta_{\max}]$, and its probability density function $g(\theta)$ satisfies $g(\theta)>0$ for all $\theta\in[0,\theta_{\max}]$.

Let $r\in\{0,1\}$ denote a user's data plan subscription decision, and $x\in[0,\infty)$ denote the number of ads that a user chooses to watch (during one month). We allow x and the advertisers' purchasing decisions to be fractional [16], [24]. The amount of data that a user obtains from its subscription and ad watching is $Qr+\omega x$. We use $\theta u(Qr+\omega x)$ to capture a type-θ user's utility of using the mobile service. Here, $u(z), z\geq 0$, is the same for all users, and can be any strictly increasing, strictly concave, and twice differentiable function that satisfies $u(0)=0$ and $\lim_{z\to\infty}u'(z)=0$. The concavity of $u(z)$ captures the diminishing marginal return with respect to the data amount. Unless otherwise specified, our results are derived under a general $u(z)$ that satisfies these properties. To study the impact of $u(z)$'s shape, we will also consider three concrete choices of $u(z)$ used in the literature:

- *Logarithmic function [25], [26]:* $u(z)=\ln(1+z)$;
- *Generalized α-fair function [15]:* $u(z)=\frac{(z+\mu)^{1-\alpha}}{1-\alpha}-\frac{\mu^{1-\alpha}}{1-\alpha}, 0<\alpha<1, \mu\geq 0$;
- *Exponential function [27]:* $u(z)=1-e^{-\gamma z}, \gamma>0$.

One reason for considering these is that the logarithmic function and generalized α-fair function are not upper bounded for $z\geq 0$, while the exponential function is upper bounded. This difference will affect the optimal choice between the SAR and SUR schemes. For ease of exposition, we call $u(\cdot)$ a user's utility function (although the actual utility is $\theta u(\cdot)$).

A type-θ user's payoff is

$$\Pi^{\text{user}}(\theta,r,x,\omega)=\theta u(Qr+\omega x)-Fr-\Phi x, \qquad (1)$$

where F is the subscription fee, and $\Phi>0$ denotes a user's average disutility (e.g., inconvenience) of watching one ad. We assume that the total disutility of watching ads linearly increases with the number of watched ads [17], [28].

In Sections III-A and IV-A, we will analyze the users' optimal decisions $r^*(\theta,\omega)$ and $x^*(\theta,\omega)$. Next, we introduce two notations to capture the total number of ad slots created by users. Let $N^{\text{ad}}(\omega)$ denote the mass of users with $x^*(\theta,\omega)>0$ (i.e., who watch ads), and let y be the value of $x^*(\theta,\omega)$ chosen by one of these $N^{\text{ad}}(\omega)$ users. Because these $N^{\text{ad}}(\omega)$ users may have different types θ, they may have different values of $x^*(\theta,\omega)$, i.e., watch different numbers of ads. Therefore, y is a *random variable*. The distribution of y gives the distribution of the number of ads watched by a user *given that the user watches ads*.[3] The expected total number of created ad slots is simply the expected total number of ads watched by the users, given by $\mathbb{E}[y]N^{\text{ad}}(\omega)$.

C. Advertisers

We consider K homogeneous advertisers. When $N^{\text{ad}}(\omega)>0$, we assume that to display the ads to a user, the operator randomly draws ads from all the $\mathbb{E}[y]N^{\text{ad}}(\omega)$ ad slots without replacement.

Suppose an advertiser purchases $m\in[0,\infty)$ ad slots from the operator (in Sections III-C and IV-C, the operator will choose its ad price p to ensure that the total number of sold

[3]The distribution of y depends on the operator's decision ω. For the simplicity of presentation, we omit this dependence in the notation.

ad slots does not exceed $\mathbb{E}[y] N^{\text{ad}}(\omega)$). If a user watches y ads, on average, $\frac{my}{\mathbb{E}[y]N^{\text{ad}}(\omega)}$ ads among the y watched ads belong to this advertiser. We let $\psi(m,y,\omega)$ denote the overall effectiveness of the advertiser's advertising on the user (e.g., a large $\psi(m,y,\omega)$ implies that the user has a good impression of the advertiser's product). We model $\psi(m,y,\omega)$ by

$$\psi(m,y,\omega) = B\frac{my}{\mathbb{E}[y] N^{\text{ad}}(\omega)} - A\left(\frac{my}{\mathbb{E}[y] N^{\text{ad}}(\omega)}\right)^2, \quad (2)$$

where $B > 0$ and $A \geq 0$ are parameters. In (2), the first term linearly increases with $\frac{my}{\mathbb{E}[y]N^{\text{ad}}(\omega)}$, and the second term quadratically decreases with $\frac{my}{\mathbb{E}[y]N^{\text{ad}}(\omega)}$. This reflects the advertising's *wear-out effect*: the advertising's effectiveness may first increase and then decrease with the number of ads delivered by this advertiser to the user. This is because too much repetition may lead the user to have a bad impression of the product. The wear-out effect has been widely observed in the literature [13], [14]. Some studies, such as [29] and [30], explicitly considered a quadratic relation between the ad repetition and the advertising's effectiveness, which is similar to (2). A larger A reflects a stronger degree of wear-out effect.[4]

We define an advertiser's utility as the expected total value of its advertising's effectiveness on all users. If a user does not see the advertiser's ads, the advertising's effectiveness on the user is zero. Therefore, an advertiser's utility is simply $\mathbb{E}_y[\psi(m,y,\omega)] N^{\text{ad}}(\omega)$. Considering the advertiser's payment for purchasing m ad slots, the advertiser's payoff is

$$\begin{aligned}\Pi^{\text{ad}}(m,\omega,p) &= \mathbb{E}_y[\psi(m,y,\omega)] N^{\text{ad}}(\omega) - mp \\ &= (B-p)m - \frac{A\mathbb{E}[y^2]}{(\mathbb{E}[y])^2 N^{\text{ad}}(\omega)} m^2. \quad (3)\end{aligned}$$

When $N^{\text{ad}}(\omega) = 0$, we simply define $\Pi^{\text{ad}}(m,\omega,p) \triangleq -mp$, and it is easy to see that the advertiser will not purchase any ad slot in this case.

D. Two-Stage Stackelberg Game

We model the interactions among the operator, users, and advertisers by a two-stage Stackelberg game. In Stage I, the operator decides the unit data reward ω and ad price p. In Stage II, each type-θ user chooses the subscription decision r and the number of watched ads x, and each advertiser decides the number of purchased ad slots m.

We assume that the users' maximum valuation $\theta_{\max}$ satisfies $\theta_{\max} > \frac{u'(0)F}{u'(Q)u(Q)}$. Similar assumptions about the range of users' attributes have been made in [31]–[33]. As shown in Sections III and IV, this assumption implies that the high-valuation users may both subscribe to the data plan and watch ads under a small reward ω. In fact, we can easily see that the user equilibrium under $\theta_{\max} \leq \frac{u'(0)F}{u'(Q)u(Q)}$ will be a special case of that under $\theta_{\max} > \frac{u'(0)F}{u'(Q)u(Q)}$.

III. Subscription-Aware Rewarding

In this section, we analyze the two-stage game under the SAR scheme, i.e., the operator only allows the subscribers of the data plan to watch ads for rewards. Note that we do not study the scheme which only rewards the non-subscribers for watching ads. This scheme is less reasonable in practice, i.e., the subscribers should not have a lower priority of using the service than the non-subscribers.

A. Users' Decisions in Stage II

Given ω, a type-θ user solves the following problem:

$$\max_{r\in\{0,1\},\ x\in[0,\infty)} \Pi^{\text{user}}(\theta,r,x,\omega), \qquad \text{s.t.} \quad x = xr, \quad (4)$$

where $\Pi^{\text{user}}(\theta,r,x,\omega)$ is given in (1), and $x = xr$ implies that a user can watch ads ($x > 0$) only if it subscribes ($r = 1$).

We use $(u')^{-1}(\cdot)$ to denote the inverse function of $u'(\cdot)$. In Lemma 1, we introduce several thresholds of θ, which will be used to characterize the users' decisions (due to space limits, we leave all proofs in our online report [34]).

Lemma 1. *Define* $\theta_0 \triangleq \frac{F}{u(Q)}$ *and* $\theta_1 \triangleq \frac{\Phi}{\omega u'(Q)}$. *When* $\omega \in \left(\frac{\Phi u(Q)}{Fu'(Q)}, \infty\right)$, *there is a unique* $\theta \in (\theta_1, \theta_0)$ *that satisfies* $\theta u\left((u')^{-1}\left(\frac{\Phi}{\omega\theta}\right)\right) - F - \frac{\Phi}{\omega}\left((u')^{-1}\left(\frac{\Phi}{\omega\theta}\right) - Q\right) = 0$, *and we denote it by* θ_2.

Although θ_1, θ_2 in Lemma 1 (and θ_3, θ_4 in Lemma 2) are functions of ω, we omit this dependence in the notation to simplify the presentation. Based on these thresholds, we characterize the users' decisions in the following proposition.

Proposition 1. *Under the SAR scheme, the optimal decisions of a type-θ user ($\theta \in [0, \theta_{\max}]$) are as follows:*[5]

Case A: When $\omega \in \left[0, \frac{\Phi}{u'(Q)\theta_{\max}}\right]$,

$$r^*(\theta,\omega) = \mathbb{1}_{\{\theta\geq\theta_0\}}, \quad x^*(\theta,\omega) = 0;$$

Case B: When $\omega \in \left(\frac{\Phi}{u'(Q)\theta_{\max}}, \frac{\Phi u(Q)}{Fu'(Q)}\right]$,

$$r^*(\theta,\omega) = \mathbb{1}_{\{\theta\geq\theta_0\}}, x^*(\theta,\omega) = \frac{1}{\omega}\left((u')^{-1}\left(\frac{\Phi}{\omega\theta}\right) - Q\right)\mathbb{1}_{\{\theta\geq\theta_1\}};$$

Case C: When $\omega \in \left(\frac{\Phi u(Q)}{Fu'(Q)}, \infty\right)$,

$$r^*(\theta,\omega) = \mathbb{1}_{\{\theta\geq\theta_2\}}, x^*(\theta,\omega) = \frac{1}{\omega}\left((u')^{-1}\left(\frac{\Phi}{\omega\theta}\right) - Q\right)\mathbb{1}_{\{\theta\geq\theta_2\}}.$$

In Fig. 2, we illustrate the data that users with different valuations θ obtain from data plan subscriptions (i.e., $Qr^*(\theta,\omega)$) and watching ads (i.e., $\omega x^*(\theta,\omega)$).

In Case A, only the users with $\theta \geq \theta_0$ subscribe, and no user watches ads because of the small unit data reward ω.

In Case B, the users who subscribe are the same as those in Case A. Users with $\theta \geq \theta_1$ watch ads, and the threshold θ_1 decreases (i.e., more users watch ads) as ω increases. Next, we focus on the users with $\theta \geq \theta_1$. We can show that the number of watched ads $x^*(\theta,\omega)$ increases with θ (note that $(u')^{-1}(\cdot)$ is decreasing because of the strict concavity of $u(\cdot)$). In particular, the marginal increase of $x^*(\theta,\omega)$ with respect to θ is affected by the utility function $u(z)$:

[4]When advertising its product, the advertiser can make several different versions of ads, and fill the m purchased ad slots with them. This can reduce A, as it mitigates the feeling of repetition from the perspective of the users.

[5]Here, $\mathbb{1}_{\{\cdot\}}$ denotes the indicator function. It equals 1 if the event in braces is true, and equals 0 otherwise.

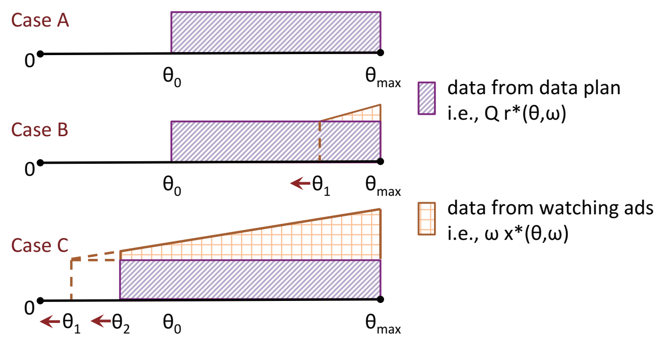


Fig. 2: Illustration of data obtained under the SAR scheme (based on Proposition 1). For $u(z) = \ln(1+z)$, the amount of data obtained via watching ads (i.e., $\omega x^*(\theta,\omega)$) linearly increases with θ when $x^*(\theta,\omega) > 0$. The red arrows indicate the change of θ_1 and θ_2 as ω increases.

- If $u(z) = \ln(1+z)$, we can show that $x^*(\theta,\omega)$ linearly increases with θ (as illustrated in Fig. 2);
- If $u(z) = \frac{(z+\mu)^{1-\alpha}}{1-\alpha} - \frac{\mu^{1-\alpha}}{1-\alpha}, 0 < \alpha < 1, \mu \geq 0$, then $x^*(\theta,\omega)$ convexly increases with θ;
- If $u(z) = 1 - e^{-\gamma z}, \gamma > 0$, then $x^*(\theta,\omega)$ concavely increases with θ.

In Case C, more users subscribe compared with Cases A and B, i.e., the subscription threshold θ_2 is smaller than θ_0. This is because the unit reward ω is large and users with $\theta \in [\theta_2, \theta_0)$ subscribe to be eligible for the data rewards. In Appendix D in our report [34], we prove that θ_2 decreases (i.e., more users subscribe) as ω increases. Moreover, each subscriber watches a positive number of ads, i.e., $x^*(\theta,\omega) > 0$ for $\theta \geq \theta_2$.

Based on these results, we can see one key advantage of the SAR scheme: it leads to a large number of data plan subscriptions.

B. Advertisers' Decisions in Stage II

Given p and ω, each advertiser solves the following problem:

$$\max_{m\in[0,\infty)} \Pi^{\rm ad}(m,\omega,p), \tag{5}$$

where the payoff $\Pi^{\rm ad}(m,\omega,p)$ is given in (3). We characterize the optimal number of purchased ad slots in Proposition 2.

Proposition 2. *If $N^{\rm ad}(\omega) = 0$ or $p \geq B$, then $m^*(\omega,p) = 0$; otherwise,*

$$m^*(\omega,p) = \frac{B-p}{2A}\frac{(\mathbb{E}[y])^2}{\mathbb{E}[y^2]}N^{\rm ad}(\omega). \tag{6}$$

Recall that the random variable y denotes the value of $x^*(\theta,\omega)$ when $x^*(\theta,\omega) > 0$, and $N^{\rm ad}(\omega)$ is the mass of users watching ads. In (6), $m^*(\omega,p)$ decreases with the degree of wear-out effect A. Moreover, since $\mathbb{E}[y^2] = (\mathbb{E}[y])^2 + {\rm Var}[y]$, we can see that $m^*(\omega,p)$ decreases with ${\rm Var}[y]$ (i.e., the variance of y). This implies that the advertisers prefer a low variation in the number of ads watched by each of the $N^{\rm ad}(\omega)$ users. The reason is that the advertising's effectiveness is concave in y given $\mathbb{E}[y]$ (as shown in (2)).

Given the concrete utility function $u(\cdot)$ and the distribution of θ, we can derive $x^*(\theta,\omega)$ based on Proposition 1, and further compute $\mathbb{E}[y]$, $\mathbb{E}[y^2]$, and $N^{\rm ad}(\omega)$ in (6).

C. Operator's Decisions in Stage I

The operator obtains revenue from both the mobile data market and ad market. In the mobile data market, each user who subscribes to the data plan should pay F to the operator. The operator's corresponding revenue is

$$R^{\rm data}(\omega) = NF\int_0^{\theta_{\max}} r^*(\theta,\omega)\, g(\theta)\, d\theta. \tag{7}$$

In the ad market, each advertiser pays p for each purchased ad slot. The operator's corresponding revenue is

$$R^{\rm ad}(\omega,p) = Km^*(\omega,p)\,p. \tag{8}$$

Let $D(\omega)$ denote the total data demand, i.e., the total amount of mobile data that users request (by subscription and watching ads) under reward ω. We can compute $D(\omega)$ as

$$D(\omega) = N\int_0^{\theta_{\max}} \left(Qr^*(\theta,\omega) + \omega x^*(\theta,\omega)\right) g(\theta)\, d\theta, \tag{9}$$

where $Qr^*(\theta,\omega)$ and $\omega x^*(\theta,\omega)$ are illustrated in Fig. 2.

Based on $R^{\rm data}(\omega)$, $R^{\rm ad}(\omega,p)$, and $D(\omega)$, we formulate the operator's problem as follows:

$$\max_{\omega\geq 0, p>0} R^{\rm total}(\omega,p) \triangleq R^{\rm data}(\omega) + R^{\rm ad}(\omega,p) \tag{10}$$

$$\text{s.t.}\quad D(\omega) \leq C, \tag{11}$$

$$Km^*(\omega,p) \leq \mathbb{E}[y]\, N^{\rm ad}(\omega). \tag{12}$$

Here, $R^{\rm total}(\omega,p)$ is the operator's total revenue. Constraint (11) implies that the total data demand $D(\omega)$ cannot exceed a capacity C [15], [20]. To ensure that choosing $\omega = 0$ (i.e., no data reward) is feasible to the problem, we assume that $C \geq D(0)$. Here, $D(0)$ is the data demand when the operator only offers the data plan without any data reward. Constraint (12) implies that the total number of sold ad slots (i.e., $Km^*(\omega,p)$) should not exceed the number of available ad slots. When the operator does not sell all ad slots, it can fill the unsold slots with content like public news and pictures to guarantee the fairness among the users choosing to watch ads (e.g., Optus displayed wallpapers to users when there were unsold ad slots [35]).

To solve (10)-(12), we first analyze $p^*(\omega)$, which is the optimal ad price under a given ω. Then, we substitute $p = p^*(\omega)$ into $R^{\rm total}(\omega,p)$, and analyze the optimal unit data reward ω^*. We characterize $p^*(\omega)$ in the following theorem.

Theorem 1. *If $\omega \in \left[0, \frac{\Phi}{u'(Q)\theta_{\max}}\right]$, any positive price is optimal; if $\omega \in \left(\frac{\Phi}{u'(Q)\theta_{\max}}, \infty\right)$,*

$$p^*(\omega) = \max\left\{\frac{B}{2}, B - \frac{2A\mathbb{E}[y^2]}{K\mathbb{E}[y]}\right\}. \tag{13}$$

Note that the random variable y is the value of $x^*(\theta,\omega)$ when $x^*(\theta,\omega) > 0$. Hence, both $\mathbb{E}[y^2]$ and $\mathbb{E}[y]$ depend on ω.

If $\omega \in \left[0, \frac{\Phi}{u'(Q)\theta_{\max}}\right]$, no user watches ads (based on Proposition 1). In this case, the advertisers do not purchase ad slots, regardless of the ad price p. If $\omega \in \left(\frac{\Phi}{u'(Q)\theta_{\max}}, \infty\right)$, Eq. (13) implies that $p^*(\omega)$ decreases with A (the degree of wear-out effect) when A is small, but does not change with A when A is large. When $A < \frac{BK\mathbb{E}[y]}{4\mathbb{E}[y^2]}$, the wear-out effect is small, and the advertisers have high willingness to purchase ad slots. Hence, the operator chooses $p^*(\omega) = B - \frac{2A\mathbb{E}[y^2]}{K\mathbb{E}[y]}$ to sell all the ad slots (which leads to $Km^*(\omega, p^*(\omega)) = \mathbb{E}[y] N^{\text{ad}}(\omega)$). When $A \geq \frac{BK\mathbb{E}[y]}{4\mathbb{E}[y^2]}$, the large wear-out effect decreases the advertisers' willingness to purchase slots. The operator will not sell all slots, and will choose $p^*(\omega) = \frac{B}{2}$, which is independent of A.

Next, we analyze ω^*, which maximizes $R^{\text{total}}(\omega, p^*(\omega))$, subject to $D(\omega) \leq C$. We first introduce Proposition 3.

Proposition 3. *Given $C \geq D(0)$, there is a unique $\omega \in \left[\frac{\Phi}{u'(Q)\theta_{\max}}, \infty\right)$ such that $D(\omega) = C$. We denote this ω by $D^{-1}(C)$. Moreover, $D^{-1}(C)$ strictly increases with C.*

Based on Proposition 3, we can rewrite $D(\omega) \leq C$ as $\omega \leq D^{-1}(C)$. From numerical experiments, $R^{\text{total}}(\omega, p^*(\omega))$ is either always increasing or unimodal in $\omega \in [0, \infty)$. Hence, we can easily search for ω^* in the interval $\left[0, D^{-1}(C)\right]$ (e.g., when $R^{\text{total}}(\omega, p^*(\omega))$ is unimodal, we can apply the Golden Section method [36]). Next, we study when the operator will choose ω to be $D^{-1}(C)$, i.e., use up the network capacity for data rewards. In Theorem 2, we show a sufficient condition under which $\omega^* = D^{-1}(C)$.

Theorem 2. *Under the SAR scheme, if both $\frac{(\mathbb{E}[y])^2}{\mathbb{E}[y^2]} N^{\text{ad}}(\omega)$ and $\mathbb{E}[y] N^{\text{ad}}(\omega)$ increase with ω for $\omega \in \left(\frac{\Phi}{u'(Q)\theta_{\max}}, \infty\right)$, the operator's optimal unit data reward is given by $\omega^* = D^{-1}(C)$.*

A widely considered setting is that each user has a logarithmic utility function (e.g., [25], [26]) and a uniformly distributed type (e.g., [16], [32]). We can verify that this setting satisfies the sufficient condition in Theorem 2, and hence we have the following proposition.

Proposition 4. *When $u(z) = \ln(1+z)$ and $\theta \sim \mathcal{U}[0, \theta_{\max}]$, the operator's optimal unit data reward is given by $\omega^* = D^{-1}(C)$.*

When each user has an exponential utility function (i.e., $u(z) = 1 - e^{-\gamma z}$), $\mathbb{E}[y] N^{\text{ad}}(\omega)$ may decrease with ω and ω^* can be smaller than $D^{-1}(C)$ (i.e., the operator does not use up the capacity for rewards). We show an example in Appendix K in [34].

IV. Subscription-Unaware Rewarding

In this section, we consider the SUR scheme, i.e., both the subscribers and non-subscribers can watch ads for rewards.

A. Users' Decisions in Stage II

Since the users can watch ads without subscription, each type-θ user simply chooses r and x to maximize its payoff without the constraint $x = xr$, as in (4) in Section III-A.

In Lemma 2, we introduce two new thresholds of θ.

Lemma 2. *Define $\theta_3 \triangleq \frac{\Phi}{\omega u'(0)}$. When $\omega \in \left(\frac{\Phi u(Q)}{Fu'(0)}, \frac{\Phi Q}{F}\right)$, there is a unique $\theta \in (\theta_3, \theta_1)$ that satisfies $\theta u\left((u')^{-1}\left(\frac{\Phi}{\omega\theta}\right)\right) - \frac{\Phi}{\omega}(u')^{-1}\left(\frac{\Phi}{\omega\theta}\right) = \theta u(Q) - F$, and we denote it by θ_4.*

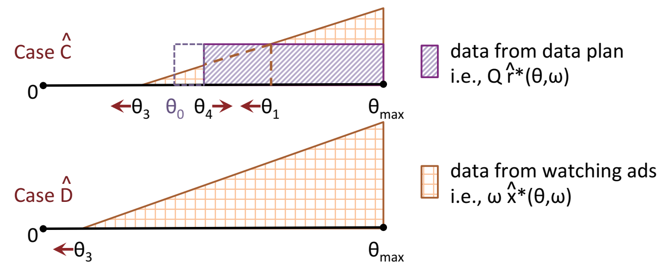


Fig. 3: Illustration of data obtained under the SUR scheme (based on Proposition 5). For $u(z) = \ln(1+z)$, the amount of data obtained via watching ads (i.e., $\omega\hat{x}^*(\theta, \omega)$) linearly increases with θ when $\hat{x}^*(\theta, \omega) > 0$. The red arrows indicate the change of θ_1, θ_3, and θ_4 as ω increases.

Recall that $(u')^{-1}(\cdot)$ denotes the inverse function of $u'(\cdot)$. Based on the thresholds introduced in Lemmas 1 and 2, we characterize the users' decisions in the following proposition (we use symbol ˆ to indicate that the results are obtained under the SUR scheme).

Proposition 5. *Under the SUR scheme, the optimal decisions of a type-θ user ($\theta \in [0, \theta_{\max}]$) are as follows:*

Case $\hat{A}$: When $\omega \in \left[0, \frac{\Phi}{u'(Q)\theta_{\max}}\right]$,

$$\hat{r}^*(\theta, \omega) = \mathbb{1}_{\{\theta \geq \theta_0\}}, \quad \hat{x}^*(\theta, \omega) = 0;$$

Case $\hat{B}$: When $\omega \in \left(\frac{\Phi}{u'(Q)\theta_{\max}}, \frac{\Phi u(Q)}{Fu'(0)}\right]$,

$$\hat{r}^*(\theta,\omega) = \mathbb{1}_{\{\theta \geq \theta_0\}}, \hat{x}^*(\theta,\omega) = \frac{1}{\omega}\left((u')^{-1}\left(\frac{\Phi}{\omega\theta}\right) - Q\right)\mathbb{1}_{\{\theta \geq \theta_1\}};$$

Case $\hat{C}$: When $\omega \in \left(\frac{\Phi u(Q)}{Fu'(0)}, \frac{\Phi Q}{F}\right)$,

$$\hat{r}^*(\theta,\omega) = \mathbb{1}_{\{\theta \geq \theta_4\}},$$
$$\hat{x}^*(\theta,\omega) = \frac{1}{\omega}(u')^{-1}\left(\frac{\Phi}{\omega\theta}\right)\mathbb{1}_{\{\theta_3 \leq \theta < \theta_4\}} + \frac{1}{\omega}\left((u')^{-1}\left(\frac{\Phi}{\omega\theta}\right) - Q\right)\mathbb{1}_{\{\theta \geq \theta_1\}};$$

Case $\hat{D}$: When $\omega \in \left[\frac{\Phi Q}{F}, \infty\right)$,

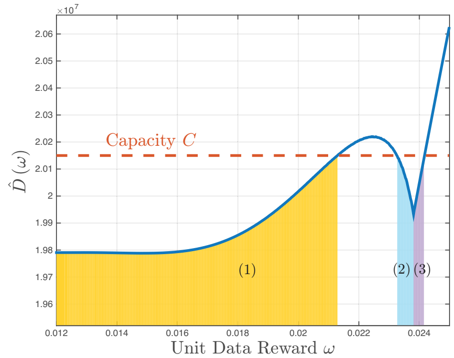
$$\hat{r}^*(\theta, \omega) = 0, \quad \hat{x}^*(\theta, \omega) = \frac{1}{\omega}(u')^{-1}\left(\frac{\Phi}{\omega\theta}\right)\mathbb{1}_{\{\theta \geq \theta_3\}}.$$

The users' optimal decisions in Cases $\hat{A}$ and $\hat{B}$ are the same as those in Cases A and B (under the SAR scheme), respectively. Hence, in Fig. 3, we only illustrate the data obtained by users via subscription (i.e., $Q\hat{r}^*(\theta, \omega)$) and watching ads (i.e., $\omega\hat{x}^*(\theta, \omega)$) in Cases $\hat{C}$ and $\hat{D}$.

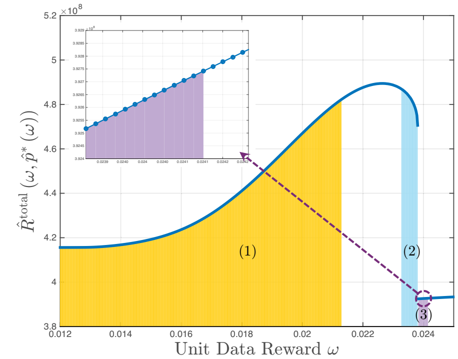
In Case $\hat{C}$, two segments of users watch ads: users with valuations $\theta \geq \theta_1$ watch ads and subscribe; users with valuations $\theta_3 \leq \theta < \theta_4$ watch ads without subscription. We characterize the properties of θ_4 in the following lemma.

Lemma 3. *When $\omega \in \left(\frac{\Phi u(Q)}{Fu'(0)}, \frac{\Phi Q}{F}\right)$ (i.e., Case $\hat{C}$), (i) θ_4 is greater than θ_0, and (ii) θ_4 increases with ω.*

In Case $\hat{B}$, the subscription threshold is θ_0. Hence, result (i) of Lemma 3 implies that some low-valuation users who subscribe in Case $\hat{B}$ become non-subscribers in Case $\hat{C}$. This is because these low-valuation users' marginal benefit of consuming data decreases after earning the data rewards, and it is no longer beneficial for them to subscribe to the data plan



(a) Function $\hat{D}(\omega)$.



(b) Function $\hat{R}^{\mathrm{total}}(\omega, \hat{p}^*(\omega))$.

Fig. 4: Examples of $\hat{D}(\omega)$ and $\hat{R}^{\mathrm{total}}(\omega, \hat{p}^*(\omega))$: We assume that $u(z) = 1 - e^{-0.7z}$ and obtain the distribution of θ by truncating the normal distribution $\mathcal{N}(125, 30)$ to interval $[0, 250]$. We choose $N = 10^7$, $F = 42$, $Q = 2$, $\Phi = 0.5$, $K = 13$, $A = 1$, $B = 5$, and $C = 2.015 \times 10^7$.

in Case $\hat{C}$. Result (ii) of Lemma 3 shows that more subscribers become non-subscribers as the unit reward increases.

In Case $\hat{D}$, since ω is large, all users simply watch ads to earn the rewards, without paying for the subscription.

B. Advertisers' Decisions in Stage II

Compared with the SAR scheme, the SUR scheme changes each advertiser's optimal decision through changing the mass of users watching ads and the distribution of the number of ads watched by each of these users.

Given $\hat{r}^*(\theta, \omega)$ and $\hat{x}^*(\theta, \omega)$ in Proposition 5, we can compute $\hat{N}^{\mathrm{ad}}(\omega)$ (i.e., the mass of users watching ads) and the distribution of $\hat{y}$ (i.e., $\hat{x}^*(\theta, \omega)$'s value when $\hat{x}^*(\theta, \omega) > 0$). To compute $\hat{m}^*(\omega, p)$, we can simply replace $N^{\mathrm{ad}}(\omega)$, $\mathbb{E}[y]$, and $\mathbb{E}[y^2]$ in Proposition 2 by $\hat{N}^{\mathrm{ad}}(\omega)$, $\mathbb{E}[\hat{y}]$, and $\mathbb{E}[\hat{y}^2]$.

C. Operator's Decisions in Stage I

Based on $\hat{r}^*(\theta, \omega)$, $\hat{x}^*(\theta, \omega)$, and $\hat{m}^*(\omega, p)$, we can compute $\hat{R}^{\mathrm{data}}(\omega)$, $\hat{R}^{\mathrm{ad}}(\omega, p)$, and $\hat{D}(\omega)$ in a similar manner as in (7)-(9). The operator's problem in Stage I is then given by:

$$\max_{\omega \geq 0, p > 0} \hat{R}^{\mathrm{total}}(\omega, p) \triangleq \hat{R}^{\mathrm{data}}(\omega) + \hat{R}^{\mathrm{ad}}(\omega, p) \tag{14}$$

$$\text{s.t.} \quad \hat{D}(\omega) \leq C, \quad K\hat{m}^*(\omega, p) \leq \hat{N}^{\mathrm{ad}}(\omega)\, \mathbb{E}[\hat{y}], \tag{15}$$

which is similar to problem (10)-(12).

To solve (14)-(15), we first compute $\hat{p}^*(\omega)$, i.e., the optimal ad price under a given ω. The analysis of $\hat{p}^*(\omega)$ is similar to that of $p^*(\omega)$ in Theorem 1 under the SAR scheme. We can prove that if $\omega \in \left[0, \frac{\Phi}{u'(Q)\theta_{\max}}\right]$, no user watches ads and hence any positive ad price is optimal; otherwise, we have $\hat{p}^*(\omega) = \max\left\{\frac{B}{2}, B - \frac{2A\mathbb{E}[\hat{y}^2]}{K\mathbb{E}[\hat{y}]}\right\}$.

Then, we compute $\hat{\omega}^*$ by maximizing $\hat{R}^{\mathrm{total}}(\omega, \hat{p}^*(\omega))$, subject to $\hat{D}(\omega) \leq C$. The computation of $\hat{\omega}^*$ is different from that of ω^* under the SAR scheme, because (i) $\hat{D}(\omega)$ can be decreasing in $\omega \in \left(\frac{\Phi u(Q)}{F u'(0)}, \frac{\Phi Q}{F}\right)$, and (ii) $\hat{R}^{\mathrm{total}}(\omega, \hat{p}^*(\omega))$ is discontinuous at $\omega = \frac{\Phi Q}{F}$. Specifically, when $\omega \in \left(\frac{\Phi u(Q)}{F u'(0)}, \frac{\Phi Q}{F}\right)$, increasing ω reduces the number of data plan subscribers, which may decrease $\hat{D}(\omega)$. Moreover, when ω increases to $\frac{\Phi Q}{F}$, all data plan subscribers quit their subscriptions and the distribution of users' ad watching times also changes. This leads to the discontinuity of $\hat{R}^{\mathrm{total}}(\omega, \hat{p}^*(\omega))$ at $\omega = \frac{\Phi Q}{F}$. We illustrate examples of $\hat{D}(\omega)$ and $\hat{R}^{\mathrm{total}}(\omega, \hat{p}^*(\omega))$ in Fig. 4(a) and Fig. 4(b), respectively.

We can compute $\hat{\omega}^*$ as follows. First, we search for ω's feasible region, where $\hat{D}(\omega) \leq C$. We can numerically show that ω's feasible region consists of at most three intervals. Then, we can show that $\hat{R}^{\mathrm{total}}(\omega, \hat{p}^*(\omega))$ is either monotone or unimodal in each interval.[6] Hence, we can determine $\hat{\omega}^*$ by comparing the local optimal unit data rewards found in these intervals.

Under the SAR scheme, the operator always uses up the capacity for rewards if $u(z) = \ln(1+z)$ and $\theta \sim \mathcal{U}[0, \theta_{\max}]$. Under the SUR scheme, this does not hold, and a large ω may even generate a total revenue that is lower than the revenue when the operator does not offer any reward. This is because a large ω may reduce the number of subscribers (as shown in Case $\hat{C}$) and hence decrease $\hat{R}^{\mathrm{data}}(\omega)$. Next, we characterize a sufficient condition under which the network capacity is not used up for rewards (given general $u(z)$ and $g(\theta)$).

Theorem 3. *Under the SUR scheme, when the network capacity $C > N(u')^{-1}\left(\frac{F}{\theta_{\max} Q}\right)$ and the degree of wear-out effect $A > \frac{B^2 K}{8F\int_{\theta_0}^{\theta_{\max}} g(\theta)d\theta}$, we have $\hat{D}(\hat{\omega}^*) < C$.*

When the operator has a large capacity and the wear-out effect is large, using up the capacity for rewards will significantly decrease $\hat{R}^{\mathrm{data}}(\omega)$ and will not significantly increase $\hat{R}^{\mathrm{ad}}(\omega, \hat{p}^*(\omega))$. Hence, we have $\hat{D}(\hat{\omega}^*) < C$ in this situation. We can show that both thresholds $N(u')^{-1}\left(\frac{F}{\theta_{\max} Q}\right)$ and $\frac{B^2 K}{8F\int_{\theta_0}^{\theta_{\max}} g(\theta)d\theta}$ decrease with F (i.e., the subscription fee). Intuitively, if the data plan is expensive, the operator should not use up the capacity for rewards under the SUR scheme.

[6] For example, in Fig. 4(a), ω's feasible region consists of the yellow, blue, and purple intervals (denoted by intervals (1), (2), and (3)). In Fig. 4(b), $\hat{R}^{\mathrm{total}}(\omega, \hat{p}^*(\omega))$ is increasing when ω is in the yellow or purple intervals, and is decreasing when ω is in the blue interval.

D. Extension: Differentiation of Ad Slots

In the above analysis, we assume that the operator does not differentiate the ad slots generated by the users. It sells all ad slots to the advertisers at the same price, and randomly draws ads from all ad slots when a user watches ads. Under the SUR scheme, the ad slots can be generated by both the subscribers and non-subscribers. In this section, we consider the differentiation of these two types of ad slots, which affects both the pricing and ad display rule. The operator can sell these two types of ad slots at different prices. When a subscriber or non-subscriber watches ads, the operator draws ads only from the corresponding type of ad slots (e.g., if an advertiser only purchases the ad slots generated by the subscribers, its ads will only be seen by the subscribers).

Given ω, we use $\hat{N}_{\rm I}^{\rm ad}(\omega)$ and $\hat{N}_{\rm II}^{\rm ad}(\omega)$ to denote the number of the subscribers that watch ads and the number of the non-subscribers that watch ads, respectively. Let random variables $\hat{y}_{\rm I}$ and $\hat{y}_{\rm II}$ denote the numbers of ads watched by one of these subscribers and one of these non-subscribers, respectively. Similar to Proposition 2, we have the following results:

- For the ad slots generated by the subscribers, the operator can set a price $p_{\rm I}>0$. If $\hat{N}_{\rm I}^{\rm ad}(\omega)>0$, the number of these slots purchased by each advertiser is $\hat{m}_{\rm I}^*(\omega,p_{\rm I})=\frac{\max\{B-p_{\rm I},0\}}{2A}\frac{(\mathbb{E}[\hat{y}_{\rm I}])^2}{\mathbb{E}[\hat{y}_{\rm I}^2]}\hat{N}_{\rm I}^{\rm ad}(\omega)$; otherwise, $\hat{m}_{\rm I}^*(\omega,p_{\rm I})=0$;
- For the slots generated by the non-subscribers, the operator can set $p_{\rm II}>0$. If $\hat{N}_{\rm II}^{\rm ad}(\omega)>0$, the number of these slots purchased by each advertiser is $\hat{m}_{\rm II}^*(\omega,p_{\rm II})=\frac{\max\{B-p_{\rm II},0\}}{2A}\frac{(\mathbb{E}[\hat{y}_{\rm II}])^2}{\mathbb{E}[\hat{y}_{\rm II}^2]}\hat{N}_{\rm II}^{\rm ad}(\omega)$; otherwise, $\hat{m}_{\rm II}^*(\omega,p_{\rm II})=0$.

The operator's problem with differentiation is given by:

$$\max_{\omega\geq 0,p_{\rm I},p_{\rm II}>0}\hat{R}^{\rm data}(\omega)+K\hat{m}_{\rm I}^*(\omega,p_{\rm I})\,p_{\rm I}+K\hat{m}_{\rm II}^*(\omega,p_{\rm II})\,p_{\rm II} \quad (16)$$

$$\text{s.t.}\quad \hat{D}(\omega)\leq C, \quad (17)$$

$$K\hat{m}_{\rm I}^*(\omega,p_{\rm I})\leq\mathbb{E}[\hat{y}_{\rm I}]\,\hat{N}_{\rm I}^{\rm ad}(\omega), \quad (18)$$

$$K\hat{m}_{\rm II}^*(\omega,p_{\rm II})\leq\mathbb{E}[\hat{y}_{\rm II}]\,\hat{N}_{\rm II}^{\rm ad}(\omega). \quad (19)$$

Constraint (18) means that the total number of sold ad slots that correspond to the subscribers should not exceed the number of ad slots generated by the subscribers. Constraint (19) can be explained similarly for the non-subscribers. In fact, only when ω satisfies Case $\hat{C}$ in Proposition 5, both the subscribers and non-subscribers watch ads (i.e., $\hat{N}_{\rm I}^{\rm ad}(\omega),\hat{N}_{\rm II}^{\rm ad}(\omega)>0$), and problem (16)-(19) is different from problem (14)-(15) (i.e., the problem without differentiation). In the remaining cases, problem (16)-(19) reduces to problem (14)-(15).

We define $\Pi^{\rm SUR}\triangleq\hat{R}^{\rm total}(\hat{\omega}^*,\hat{p}^*(\hat{\omega}^*))$, which is the optimal objective value of problem (14)-(15). Let $\Pi^{\rm SURD}$ denote the optimal objective value of problem (16)-(19), i.e., $\Pi^{\rm SURD}$ is the operator's optimal total revenue under the SUR scheme with differentiation. We compare $\Pi^{\rm SUR}$ and $\Pi^{\rm SURD}$ in the following theorem.

Theorem 4. *We always have* $\Pi^{\rm SURD}\geq\Pi^{\rm SUR}$.

Hence, differentiation does not decrease the operator's optimal total revenue (given general $u(z)$ and $g(\theta)$). Intuitively, if the optimal unit data reward satisfies Case $\hat{C}$ and the distributions of $\hat{y}_{\rm I}$ and $\hat{y}_{\rm II}$ are significantly different, the gap between $\Pi^{\rm SURD}$ and $\Pi^{\rm SUR}$ will be large. In the next section, we will show this gap numerically.

V. Comparison Between Rewarding Schemes

We define $\Pi^{\rm SAR}\triangleq R^{\rm total}(\omega^*,p^*(\omega^*))$, which is the operator's optimal total revenue under the SAR scheme. In this section, we compare $\Pi^{\rm SAR}$, $\Pi^{\rm SUR}$, and $\Pi^{\rm SURD}$. Since the comparison is challenging under a general user type distribution and a general utility function, we focus on specific user type distributions and utility functions. In Sections V-A and V-B, we consider uniformly distributed user types and truncated normally distributed user types, respectively.

A. Uniformly Distributed User Types

In this section, we assume that each user's type θ follows a uniform distribution. We will consider logarithmic utility, generalized α-fair utility, and exponential utility.

1) Logarithmic Utility Function: We assume that $u(z)=\ln(1+z)$. Theorem 5 characterizes the analytical comparison between different schemes as $C\rightarrow\infty$.

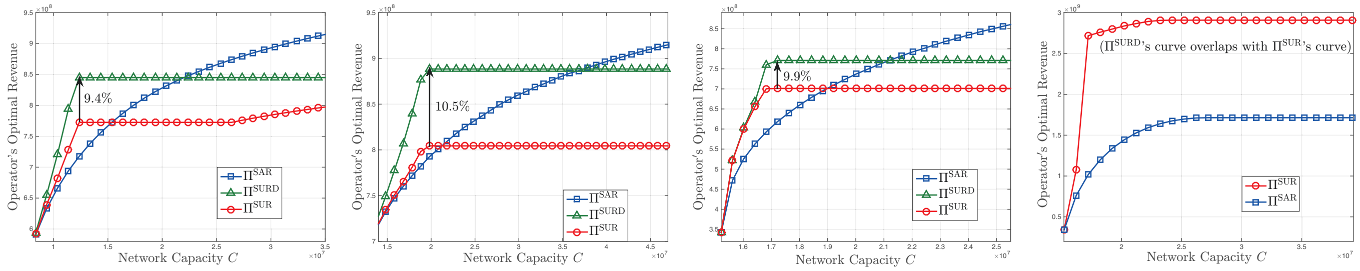
Theorem 5. *When* $\theta\sim\mathcal{U}[0,\theta_{\max}]$ *and* $u(z)=\ln(1+z)$*, if network capacity* $C\rightarrow\infty$*, then* $\Pi^{\rm SAR}>\Pi^{\rm SURD}\geq\Pi^{\rm SUR}$.

Theorem 5 implies that if the operator has sufficiently large network capacity, it should only reward the subscribers for watching ads. Intuitively, this allows the operator to motivate all users to subscribe and watch ads via high data rewards. It maximizes the operator's revenue from both the data market and the ad market.

Under a finite network capacity C, none of $\Pi^{\rm SAR}$, $\Pi^{\rm SUR}$, or $\Pi^{\rm SURD}$ has a closed-form expression, and their analytical comparison is challenging. Next, we compare them numerically. In the numerical experiment, we choose $N=10^7$, $F=30$, $Q=0.8$, $\theta\sim\mathcal{U}[0,155]$, $\Phi=0.3$, $K=23$, $A=0.6$, and $B=5$. Here, we consider an area with 10 million users. In Appendix R in [34], we consider different parameter settings (e.g., different values of N), and the key observations summarized in this section still hold under those settings.

In Fig. 5(a), we plot $\Pi^{\rm SAR}$, $\Pi^{\rm SUR}$, and $\Pi^{\rm SURD}$ against C. We can see that only $\Pi^{\rm SAR}$ strictly increases with C. As shown in Proposition 4, when each user has a logarithmic utility and a uniformly distributed type, the operator always uses up the capacity for rewards under the SAR scheme. Hence, the operator can always benefit from C's increase in this situation.

First, we compare $\Pi^{\rm SAR}$ and $\Pi^{\rm SUR}$. When C is close to $D(0)$, $\Pi^{\rm SAR}$ and $\Pi^{\rm SUR}$ are equal. In this situation, the operator can only choose a very small unit reward ω. As shown in Case B in Proposition 1 and Case $\hat{B}$ in Proposition 5, the users' optimal decisions under the two schemes are the same, which leads to the same operator's revenue. When C is from 0.84×10^7 to 1.54×10^7, $\Pi^{\rm SAR}$ is smaller than $\Pi^{\rm SUR}$. This is because the SUR scheme can motivate two segments of users to watch ads (by setting $\omega\in\left(\frac{\Phi u(Q)}{Fu'(0)},\frac{\Phi Q}{F}\right)$, as shown in Case $\hat{C}$ in Proposition 5), which generates a higher ad revenue than the SAR scheme. When C is greater than 1.54×10^7,



(a) Logarithmic Utility. (b) Generalized α-Fair Utility. (c) Exponential Utility (Large A). (d) Exponential Utility (Small A).

Fig. 5: $\Pi^{\rm SAR}$, $\Pi^{\rm SUR}$, and $\Pi^{\rm SURD}$ Under Different Network Capacity (Uniformly Distributed θ).

$\Pi^{\rm SAR}$ is greater than $\Pi^{\rm SUR}$. The operator will fully utilize the large network capacity under the SAR scheme, and set a large ω to motivate more users to both subscribe and watch ads. This is consistent with Theorem 5 (i.e., if $C \to \infty$, then $\Pi^{\rm SAR} > \Pi^{\rm SUR}$). We summarize the results in Observation 1 (the comparison between $\Pi^{\rm SAR}$ and $\Pi^{\rm SURD}$ is similar to the comparison between $\Pi^{\rm SAR}$ and $\Pi^{\rm SUR}$).

Observation 1. *When $u(z) = \ln(1+z)$, if C is small, the SUR scheme achieves a higher operator's revenue; otherwise, the SAR scheme achieves a higher operator's revenue.*

Second, we compare $\Pi^{\rm SUR}$ and $\Pi^{\rm SURD}$. When $C = 1.24 \times 10^7$, Fig. 5(a) shows that the ad slots' differentiation can improve the operator's revenue under the SUR scheme by 9.4%. This is because the value of $\hat{\omega}^*$ under the SUR scheme satisfies Case $\hat{C}$ in Proposition 5, which implies that both subscribers and non-subscribers watch ads. Moreover, the subscribers and non-subscribers have quite different ad watching behaviors. In Fig. 6(a), we illustrate the distributions of $\hat{y}_{\rm I}$ (i.e., the number of ads watched by a subscriber) and $\hat{y}_{\rm II}$ (i.e., the number of ads watched by a non-subscriber) when $C = 1.24 \times 10^7$ and the operator uses the SUR scheme. We can see that both $\hat{y}_{\rm I}$ and $\hat{y}_{\rm II}$ follow uniform distributions, but their mean values are significantly different.

2) Generalized α-Fair Utility Function: We assume that $u(z) = \frac{(z+\mu)^{1-\alpha}}{1-\alpha} - \frac{\mu^{1-\alpha}}{1-\alpha}$. We choose $\alpha = 0.8$ and $\mu = 0.8$, and the other settings are the same as those in Fig. 5(a). In Fig. 5(b), we plot $\Pi^{\rm SAR}$, $\Pi^{\rm SUR}$, and $\Pi^{\rm SURD}$ against C. We can see that the comparison among the operator's optimal revenues under different schemes is similar to that in Fig. 5(a). We summarize the key results about the comparison between $\Pi^{\rm SAR}$ and $\Pi^{\rm SUR}$ in the following observation.

Observation 2. *When $u(z) = \frac{(z+\mu)^{1-\alpha}}{1-\alpha} - \frac{\mu^{1-\alpha}}{1-\alpha}$, if C is small, the SUR scheme achieves a higher operator's revenue; otherwise, the SAR scheme achieves a higher operator's revenue.*

3) Exponential Utility Function: We assume that $u(z) = 1 - e^{-\gamma z}$, and choose $\gamma = 0.7$, $N = 10^7$, $F = 45$, $Q = 2$, $\theta \sim \mathcal{U}[0, 250]$, $\Phi = 0.3$, $K = 23$, and $B = 5$. In Fig. 5(c) and Fig. 5(d), we show the comparison between $\Pi^{\rm SAR}$, $\Pi^{\rm SUR}$, and $\Pi^{\rm SURD}$ under different degrees of the wear-out effect.

In Fig. 5(c), we consider a large wear-out effect ($A = 0.9$). The comparison between $\Pi^{\rm SAR}$ and $\Pi^{\rm SUR}$ (or $\Pi^{\rm SURD}$) is similar to those in Fig. 5(a) and Fig. 5(b). The SAR scheme achieves a higher revenue than the SUR scheme when C is large. Comparing $\Pi^{\rm SUR}$ and $\Pi^{\rm SURD}$ in Fig. 5(c), we observe that differentiation improves the operator's revenue under the SUR scheme by at most 9.9%.

In Fig. 5(d), we consider a small wear-out effect ($A = 0.2$), and have three observations. First, $\Pi^{\rm SAR}$ may not change with C, which is different from the logarithmic utility situation shown in Fig. 5(a). When each user has an exponential utility, the operator may not benefit from the increase of C, since it may not use up the capacity for the rewards (as discussed in Section III-C). Second, $\Pi^{\rm SAR}$ is always no greater than $\Pi^{\rm SUR}$ (even under a large C), which is different from the logarithmic utility situation and the generalized α-fair utility situation. Under the SAR scheme, each user has to pay the subscription fee $F > 0$ before receiving the data rewards. The exponential utility function is upper bounded (i.e., $u(z) = 1 - e^{-\gamma z} \leq 1$), and hence the users with $\theta < F$ will never subscribe and watch ads under the SAR scheme, regardless of the unit data reward ω. When A is small, the advertisers are willing to buy more slots, and having more users watching ads significantly increases the operator's revenue. Therefore, the SUR scheme, which can motivate the users with $\theta < F$ to watch ads, achieves a higher revenue than the SAR scheme. Third, the $\Pi^{\rm SURD}$ curve overlaps the $\Pi^{\rm SUR}$ curve, because the operator chooses a large ω to incentivize the users to watch ads under a small A. In this situation, all the ad slots are generated by non-subscribers under the SUR scheme (see Case $\hat{D}$ of Proposition 5), and the differentiation cannot improve the operator's revenue.

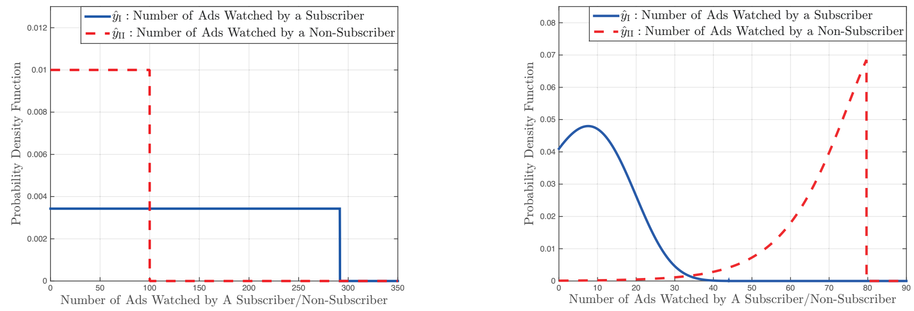
We summarize the key observations below.

Observation 3. *When $u(z) = 1 - e^{-\gamma z}$, (i) under a large A, the SUR scheme achieves a higher operator revenue than the SAR scheme if and only if C is below a finite threshold; (ii) under a small A, the SUR scheme always achieves a higher operator revenue than the SAR scheme.*

B. Truncated Normally Distributed User Types

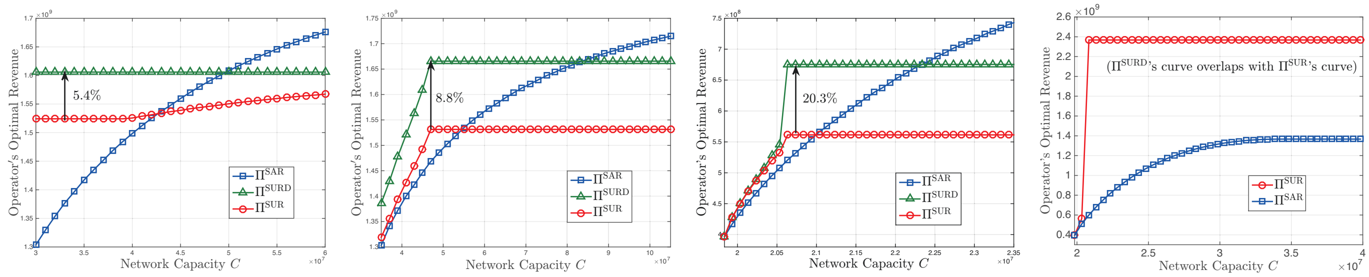
We next assume that each user's type θ follows a truncated normal distribution. We show that most observations under the uniformly distributed user types still hold.

1) Logarithmic Utility Function: We assume that $u(z) = \ln(1+z)$, and obtain the distribution of θ by truncating the normal distribution $\mathcal{N}(75, 40)$ to interval $[0, 150]$. We choose



(a) Logarithmic Utility and Uniform Distribution. (b) Exponential Utility and Truncated Normal Distribution.

Fig. 6: Probability Distribution Function of $\hat{y}_{\rm I}$ and $\hat{y}_{\rm II}$.



(a) Logarithmic Utility. (b) Generalized α-Fair Utility. (c) Exponential Utility (Large A). (d) Exponential Utility (Small A).

Fig. 7: $\Pi^{\rm SAR}$, $\Pi^{\rm SUR}$, and $\Pi^{\rm SURD}$ Under Different Network Capacity (Truncated Normally Distributed θ).

$N = 10^7$, $F = 40$, $Q = 2$, $\Phi = 0.03$, $K = 8$, $A = 0.5$, and $B = 10$. In Fig. 7(a), we plot $\Pi^{\rm SAR}$, $\Pi^{\rm SUR}$, and $\Pi^{\rm SURD}$ against C. We can see that the SUR scheme outperforms the SAR scheme if and only if C is below a threshold. This is consistent with Observation 1.

2) Generalized α-Fair Utility Function: We next assume that $u(z) = \frac{(z+\mu)^{1-\alpha}}{1-\alpha} - \frac{\mu^{1-\alpha}}{1-\alpha}$, where $\alpha = 0.8$ and $\mu = 0.8$. The other settings are the same as those in Fig. 7(a). We plot the operator's optimal revenues under different schemes in Fig. 7(b). The influence of C on the comparison is consistent with Observation 2.

3) Exponential Utility Function: We next assume that $u(z) = 1 - e^{-\gamma z}$, and obtain the distribution of θ by truncating the normal distribution $\mathcal{N}(125, 30)$ to interval $[0, 250]$. We choose $\gamma = 0.7$, $N = 10^7$, $F = 40$, $Q = 2$, $\Phi = 0.5$, $K = 16$, and $B = 5$. Fig. 7(c) and Fig. 7(d) show the comparison between $\Pi^{\rm SAR}$, $\Pi^{\rm SUR}$, and $\Pi^{\rm SURD}$ under $A = 0.9$ and $A = 0.2$, respectively.

Fig. 7(c) shows that if the wear-out effect is large, the SUR scheme outperforms the SAR scheme under a small C. Fig. 7(d) shows that if the wear-out effect is small, the SUR scheme always outperforms the SAR scheme. These results are consistent with Observation 3.

In Fig. 7(c), when $C = 2.07 \times 10^7$, the differentiation of the ad slots improves the operator's revenue under the SUR scheme by 20.3%. To explain this large improvement, we illustrate the distributions of $\hat{y}_{\rm I}$ and $\hat{y}_{\rm II}$ under $C = 2.07 \times 10^7$ and the SUR scheme in Fig. 6(b). We can observe that the difference between the two distributions is greater than that in Fig. 6(a) (where each user has a logarithmic utility function and a uniformly distributed type). For example, the value of $\frac{\mathbb{E}[\hat{y}_{\rm II}]}{\mathbb{E}[\hat{y}_{\rm I}]}$ in Fig. 6(b) is around 5.7, and the value of $\frac{\mathbb{E}[\hat{y}_{\rm I}]}{\mathbb{E}[\hat{y}_{\rm II}]}$ in Fig. 6(a) is around 2.9. Intuitively, when the difference between the subscribers' and non-subscribers' ad watching behaviors is larger, the benefit of differentiation is more obvious. Therefore, the improvement of $\Pi^{\rm SURD}$ over $\Pi^{\rm SUR}$ in Fig. 7(c) is greater than the improvement in Fig. 5(a) (which is 9.4%).

VI. Conclusion

Mobile data rewarding is an emerging approach to monetize mobile services. We modeled the data rewarding ecosystem and analyzed an operator's rewarding scheme. Our results reveal that: (i) increasing the unit data reward may decrease the number of ads watched by the users, and the operator may not use up its network capacity to reward the users; (ii) under the SUR scheme, the operator can improve its revenue by differentiating the ad slots generated by the subscribers and non-subscribers; (iii) the operator's optimal choice between the SAR and SUR schemes is sensitive to the user utility function, network capacity, and advertising's wear-out effect.

In future work, we plan to first study the operator's joint optimization of the data plan and the data rewards. Under the SAR scheme, the operator can reduce the subscription fee to motivate more users to subscribe and watch ads. Under the SUR scheme, the operator may increase the subscription fee, which (i) extracts more revenue from the users with high θ and (ii) pushes more users with low θ to become non-subscribers and watch ads. Second, we are interested in relaxing the assumptions of a monopolistic operator and homogeneous advertisers. For example, when there are multiple operators, they will compete for users as well as advertisers, which may increase the unit data rewards and reduce the ad prices. Third, we can study a general data rewarding scheme where the operator can set different unit data rewards for the subscribers and non-subscribers. The SAR and SUR schemes can be treated as two special cases of this general scheme.

References

[1] H. Yu, E. Wei, and R. A. Berry, "A business model analysis of mobile data rewards," in *Proc. of IEEE INFOCOM*, Paris, France, 2019, pp. 2098–2106.

[2] S. Analytics, "Worldwide cellular user forecasts 2018-2023," Tech. Rep., 2018.

[3] ——, "Can operator collaboration on sponsored data lead to success?" Tech. Rep., 2018.

[4] eMarketer, https://www.emarketer.com/Article/Want-App-Users-Interact-with-Your-Ads-Reward-Them/1010966, accessed on Nov. 27, 2019.

[5] Aquto, http://www.aquto.com/, accessed on Nov. 27, 2019.

[6] N. Summers, https://www.engadget.com/2016/06/09/tesco-mobile-unlockd-lock-screen-ads/, accessed on Nov. 27, 2019.

[7] DOCOMO, https://www.nttdocomo.co.jp/english/info/media_center/pr/2017/1225_00.html, accessed on Nov. 27, 2019.

[8] R. Johnston, https://www.gizmodo.com.au/2016/11/optus-wants-you-to-watch-ads-in-return-for-data/, accessed on Nov. 27, 2019.

[9] AT&T, https://about.att.com/story/2018/att_appnexus.html, accessed on Nov. 27, 2019.

[10] S. Dewing, https://tech.co/news/adwallet-future-advertising-app-2017-08, accessed on Nov. 27, 2019.

[11] AdStage, "Q3 2018 paid search and paid social benchmark report," Tech. Rep., 2018.

[12] Ultra Mobile, https://www.ultramobile.com/monthly-plans/, accessed on Nov. 27, 2019.

[13] C. Pechmann and D. W. Stewart, "Advertising repetition: A critical review of wearin and wearout," *Current issues and research in advertising*, vol. 11, no. 1-2, pp. 285–329, 1988.

[14] A. Kirmani, "Advertising repetition as a signal of quality: If it's advertised so much, something must be wrong," *Journal of advertising*, vol. 26, no. 3, pp. 77–86, 1997.

[15] C. Joe-Wong, S. Sen, and S. Ha, "Sponsoring mobile data: Analyzing the impact on internet stakeholders," *IEEE/ACM Transactions on Networking*, vol. 26, no. 3, pp. 1179–1192, 2018.

[16] H. Yu, M. H. Cheung, L. Gao, and J. Huang, "Public Wi-Fi monetization via advertising," *IEEE/ACM Transactions on Networking*, vol. 25, no. 4, pp. 2110–2121, 2017.

[17] H. Guo, X. Zhao, L. Hao, and D. Liu, "Economic analysis of reward advertising," *Working Paper*, 2017.

[18] M. Andrews, U. Özen, M. I. Reiman, and Q. Wang, "Economic models of sponsored content in wireless networks with uncertain demand," in *Proc. of IEEE INFOCOM Workshops*, Turin, Italy, 2013, pp. 345–350.

[19] M. H. Lotfi, K. Sundaresan, S. Sarkar, and M. A. Khojastepour, "Economics of quality sponsored data in non-neutral networks," *IEEE/ACM Transactions on Networking*, vol. 25, no. 4, pp. 2068–2081, 2017.

[20] L. Zhang, W. Wu, and D. Wang, "Sponsored data plan: A two-class service model in wireless data networks," in *Proc. of ACM SIGMETRICS*, Portland, OR, USA, 2015.

[21] P. Bangera, S. Hasan, and S. Gorinsky, "An advertising revenue model for access ISPs," in *Proc. of IEEE ISCC*, Heraklion, Greece, 2017.

[22] S. Sen, G. Burtch, A. Gupta, and R. Rill, "Incentive design for ad-sponsored content: Results from a randomized trial," in *Proc. of IEEE INFOCOM Workshops*, Atlanta, GA, USA, 2017, pp. 826–831.

[23] M. Harishankar, N. Srinivasan, C. Joe-Wong, and P. Tague, "To accept or not to accept: The question of supplemental discount offers in mobile data plans," in *Proc. of IEEE INFOCOM*, Honolulu, HI, USA, 2018.

[24] D. Bergemann and A. Bonatti, "Targeting in advertising markets: implications for offline versus online media," *The RAND Journal of Economics*, vol. 42, no. 3, pp. 417–443, 2011.

[25] L. Duan, J. Huang, and B. Shou, "Pricing for local and global Wi-Fi markets," *IEEE Transactions on Mobile Computing*, vol. 14, no. 5, pp. 1056–1070, 2015.

[26] D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig, and W. Utschick, "Minimum mean squared error interference alignment," in *Proc. of IEEE ACSSC*, Pacific Grove, CA, USA, 2009, pp. 1106–1110.

[27] C. Zhou, M. L. Honig, and S. Jordan, "Utility-based power control for a two-cell CDMA data network," *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 2764–2776, 2005.

[28] S. P. Anderson and B. Jullien, "The advertising-financed business model in two-sided media markets," in *Handbook of media economics*, 2015, vol. 1, pp. 41–90.

[29] B. N. Anand and R. Shachar, "Advertising, the matchmaker," *The RAND Journal of Economics*, vol. 42, no. 2, pp. 205–245, 2011.

[30] M. C. Campbell and K. L. Keller, "Brand familiarity and advertising repetition effects," *Journal of consumer research*, vol. 30, no. 2, pp. 292–304, 2003.

[31] R. Dewenter, J. Haucap, and T. Wenzel, "On file sharing with indirect network effects between concert ticket sales and music recordings," *Journal of Media Economics*, vol. 25, no. 3, pp. 168–178, 2012.

[32] A. Rasch and T. Wenzel, "Piracy in a two-sided software market," *Journal of Economic Behavior & Organization*, vol. 88, pp. 78–89, 2013.

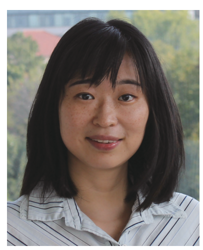
[33] H. Yu, G. Iosifidis, B. Shou, and J. Huang, "Pricing for collaboration between online apps and offline venues," *IEEE Transactions on Mobile Computing*, 2019.

[34] H. Yu, E. Wei, and R. Berry, "Monetizing mobile data via data rewards," *arXiv preprint arXiv:1911.12023*, 2019.

[35] Optus, https://www.optus.com.au/shop/mobile/apps/optusxtra, accessed on Nov. 27, 2019.

[36] D. P. Bertsekas, *Nonlinear programming*. Belmont, MA, USA: Athena Scientific, 1999.

Haoran Yu (S'14-M'17) received the Ph.D. degree from the Chinese University of Hong Kong in 2016. During 2015-2016, he was a Visiting Student in the Yale Institute for Network Science and the Department of Electrical Engineering at Yale University. During 2018-2019, he was a Post-Doctoral Fellow in the Department of Electrical and Computer Engineering at Northwestern University. His recent research interests lie in the field of mechanism design for networks. His paper in *IEEE INFOCOM 2016* was selected as a Best Paper Award finalist and one of top 5 papers from over 1600 submissions.



Ermin Wei is currently an Assistant Professor at the ECE Dept. of Northwestern University. She completed her PhD studies in Electrical Engineering and Computer Science at MIT in 2014, advised by Professor Asu Ozdaglar, where she also obtained her M.S.. She received her undergraduate triple degree in Computer Engineering, Finance and Mathematics with a minor in German, from University of Maryland, College Park. Wei has received many awards, including the Graduate Women of Excellence Award, second place prize in Ernst A. Guillemen Thesis Award and Alpha Lambda Delta National Academic Honor Society Betty Jo Budson Fellowship. Wei's research interests include distributed optimization methods, convex optimization and analysis, smart grid, communication systems and energy networks and market economic analysis.

Randall A. Berry (S'93-M'00-SM'12-F'14) received the Ph.D. degree from the Massachusetts Institute of Technology in 2000 and joined Northwestern University, where he is currently the John A. Dever Professor and Chair of Electrical and Computer Engineering. Dr. Berry's research interests span topics in network economics, wireless communication, computer networking, and information theory. He was a recipient of the 2003 CAREER Award from the National Science Foundation. He served as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2006 to 2009 and the IEEE TRANSACTIONS ON INFORMATION THEORY from 2008 to 2012, as well as a Guest Editor for special issues of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS in 2017, the IEEE JOURNAL ON SELECTED TOPICS IN SIGNAL PROCESSING in 2008, and the IEEE TRANSACTIONS ON INFORMATION THEORY in 2007. He has also served on the program and organizing committees of numerous conferences, including serving as the Co-Chair for the 2012 IEEE Communication Theory Workshop and 2010 IEEE ICC Wireless Networking Symposium and the TPC Co-Chair for the 2018 ACM MobiHoc conference.