A Business Model Analysis of Mobile Data Rewards

Haoran Yu, Ermin Wei, and Randall A. Berry
Department of Electrical Engineering and Computer Science, Northwestern University, IL, USA
Email: {haoran.yu, ermin.wei, r-berry4}@northwestern.edu

***Abstract*—Conventionally, mobile network operators charge users for data plan subscriptions. To create new revenue streams, some operators now also incentivize users to watch ads with data rewards and collect payments from advertisers. In this work, we study two such 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' existing 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 a capacity-constrained 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 provide some counter-intuitive insights for the design of data rewards. For example, the operator's optimal choice between the two schemes is sensitive to the users' data consumption utility function. 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.**

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 [1]. As suggested in [2], 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 [2]: (i) 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 [2]; (ii) 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); (iii) The users earn free mobile data to satisfy their growing data demand.

There has been an increasing number of businesses entering this space. Unlockd and Aquto 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). Unlockd has collaborated with operators, such as Tesco Mobile (in the United Kingdom), Boost Mobile (in the United States), Lebara Mobile (in Australia), and AXIS (in Indonesia) [3]. Aquto has collaborated with Verizon, Telefonica, and several other operators [4]. Furthermore, AT&T is acquiring AppNexus (an online advertising company) and making a significant investment in the advertising business [5]. Offering mobile data rewards could become a natural and effective approach to monetize its mobile service.

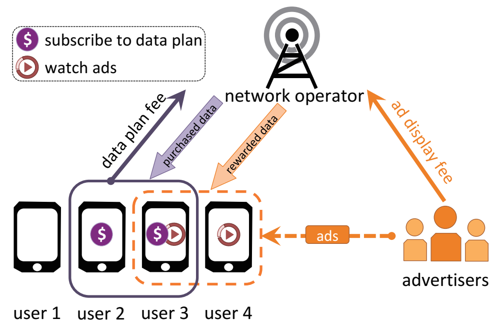


Fig. 1: Data rewarding ecosystem (user 4 is feasible under the *SUR* scheme, but is infeasible under the *SAR* scheme).

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.[1] 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.

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

This research was supported in part by the NSF grants AST-1343381, AST-1547328 and CNS-1701921.

[1] 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 [6].

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 [7], [8]).

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 total 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 its 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 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 and 2): *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, 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 advertising's 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.

Our comparison between the SAR and SUR schemes also provides insights to a more general problem, where the operator offers *multiple* data plans and decides whether to only allow the subscribers of the *expensive* data plans to earn rewards. Our analysis of the SAR and SUR schemes captures the key considerations of choosing these schemes (e.g., whether to motivate more subscriptions to the expensive data plans or incentivize more ad watching).

B. Related Work

There have been some references studying markets where providers offer both a fee-based service and an ad-based free service. For example, in [9], a Wi-Fi network provider allows users to either directly pay or watch ads to access the Wi-Fi network. In [10], an app developer offers virtual items, and each app user will either pay or watch ads to obtain them at 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. Furthermore, our work considers the operator's capacity for providing the service and the advertising's wear-out effect, which were not considered in [9] and [10].

As studied in [11]–[14], operators can create new revenue streams by offering sponsored data: content providers sponsor the data usage of their content, and users can access the content free of charge. There are several key differences between the sponsored data and data rewards as studied here: (i) 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; (ii) 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.

Other related work includes [15]–[18]. Bangera *et al.* in [15] conducted a survey, which shows that 76% of the respondents are interested in watching ads in exchange for mobile data. Sen *et al.* in [16] conducted an experiment to study the effectiveness of monetary rewards in increasing ads' viewership. Both [15] and [16] did not analyze the equilibrium strategies of the entities, such as operators, advertisers, and users. Harishankar *et al.* in [17] studied monetizing the operator's idle network capacity by providing users with supplemental discount offers, which are not related to advertising. In our prior work [18], we discussed the operators' design of the data rewards, but focused on a restricted problem setting. For example, the operators mainly reward the subscribers, and each user has a logarithmic utility function and a uniformly distributed type. In this work, we comprehensively compare the SAR and SUR schemes, and consider a general user utility function and a general user type distribution.

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.[2] 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., [13], [14]).

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 pricing-based mechanism, where the operator sells the ad slots in advance at a fixed price.[3]

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 (we allow x and the advertisers' purchasing decisions to be fractional [9], [20]). 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 two concrete choices of $u(z)$ used in the literature:

- *Logarithmic function [21], [22]:* $u(z)=\ln(1+z)$;
- *Exponential function [23]:* $u(z)=1-e^{-\gamma z}, \gamma>0$.

One reason for considering these is that the logarithmic function is 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. Note that some other functions, such as the α-fair function $u(z)=\frac{z^{1-\alpha}}{1-\alpha}$ $(0<\alpha<1)$ [13], also satisfy the required properties. 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

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

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 [10].

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 $N^{\rm ad}(\omega)$ denote the mass of users who watch ads (i.e., have $x^*(\theta,\omega)>0$), and let y be a random variable denoting the value of $x^*(\theta,\omega)$ when $x^*(\theta,\omega)>0$.[4] The distribution of y gives the distribution of the number of ads watched by one of the $N^{\rm ad}(\omega)$ users. Therefore, the expected total number of created ad slots is $\mathbb{E}[y]N^{\rm ad}(\omega)$.

C. Advertisers

We consider K homogeneous advertisers. When $N^{\rm ad}(\omega)>0$, we assume that to display the ads to a user, the operator randomly draws ads from all the $\mathbb{E}[y]N^{\rm ad}(\omega)$ 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 ad slots does not exceed $\mathbb{E}[y]N^{\rm ad}(\omega)$). If a user watches y ads, on average, $\frac{my}{\mathbb{E}[y]N^{\rm ad}(\omega)}$ 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

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

where $B>0$ and $A\geq 0$ are parameters. The above equation captures 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

[2] Compared with designing data rewards, the operator has less flexibility to adjust its data plan (e.g., subscribers may sign long-term contracts with the operator). Hence, we study the operator's reward design, *given* its existing data plan. In our future work, we plan to extend our analysis by jointly optimizing the data plan and reward.

[3] As shown in [3], the operator and advertisers have large-scale collaborations (e.g., an advertiser's ads were displayed around 300,000 times per promotion activity). In this case, the pricing-based mechanism facilitates the customization and communication process [19]. The operator can also sell the slots via the real-time auction, especially when it has some user profiles and the advertisers want to target different user categories [19]. We leave the study of heterogeneous advertisers and real-time auction to future work.

[4] The randomness of y is from the uncertainty of θ. Moreover, the distribution of y depends on the operator's decision ω. For the simplicity of presentation, we omit this dependence in the notation.

impression of the product. The wear-out effect has been widely observed in the literature [7], [8]. Some studies, such as [24] and [25], explicitly considered a quadratic relation between the ad repetition and the advertising's effectiveness, which is similar to Eq. (2). Note that a larger A in (2) reflects a stronger degree of wear-out effect.

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 $\mathbb{E}_y\left[\psi\left(m, y, \omega\right)\right] N^{\mathrm{ad}}\left(\omega\right)$. Considering the advertiser's payment for purchasing m ad slots, the advertiser's payoff is

$$\Pi^{\mathrm{ad}}\left(m, \omega, p\right) = \mathbb{E}_y\left[\psi\left(m, y, \omega\right)\right] N^{\mathrm{ad}}\left(\omega\right) - mp. \quad (3)$$

When $N^{\mathrm{ad}}\left(\omega\right) = 0$, we simply define $\Pi^{\mathrm{ad}}\left(m, \omega, p\right) \triangleq -mp$, 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 [26], [27]. 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.

A. Users' Decisions in Stage II

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

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

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

In Lemma 1, we introduce several thresholds of θ, which will be used to characterize the users' decisions. Note that $\left(u'\right)^{-1}\left(\cdot\right)$ denotes the inverse function of $u'\left(\cdot\right)$.

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

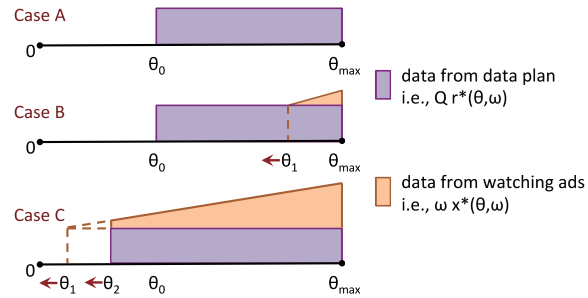


Fig. 2: Illustration of data obtained under the SAR scheme (based on Proposition 1). For $u\left(z\right) = \ln\left(1+z\right)$, the amount of data obtained via watching ads (i.e., $\omega x^*\left(\theta, \omega\right)$) linearly increases with θ when $x^*\left(\theta, \omega\right) > 0$.

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

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

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

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

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

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

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

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 $x^*\left(\theta, \omega\right)$ increases with θ ($\left(u'\right)^{-1}\left(\cdot\right)$ is decreasing because of the strict concavity of $u\left(\cdot\right)$). In particular, the marginal increase of $x^*\left(\theta, \omega\right)$ with respect to θ is affected by the utility function $u\left(z\right)$:

- If $u\left(z\right) = \ln\left(1+z\right)$, we can show that $x^*\left(\theta, \omega\right)$ linearly increases with θ (as illustrated in Fig. 2);
- If $u\left(z\right) = 1 - e^{-\gamma z}, \gamma > 0$, then $x^*\left(\theta, \omega\right)$ 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. We can prove that θ_2 decreases (i.e., more users subscribe) as ω increases.

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

Moreover, each subscriber watches a positive number of ads, i.e., $x^*(\theta,\omega)>0$ for $\theta\ge\theta_2$.

Based on these results, we can see one key advantage of the SAR scheme: it leads to more 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 $p\ge B$, 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. Given the concrete utility function $u(\cdot)$ and the distribution of θ, we can derive the expression of $x^*(\theta,\omega)$ based on Proposition 1, and further compute $\mathbb{E}[y]$, $\mathbb{E}\left[y^2\right]$, and $N^{\rm ad}(\omega)$.

In (6), $m^*(\omega,p)$ decreases with the degree of wear-out effect A. Moreover, since $\mathbb{E}\left[y^2\right]=(\mathbb{E}[y])^2+{\rm Var}[y]$, 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.

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 in Stage I as follows:

$$\max_{\omega\ge0,p>0} R^{\rm total}(\omega,p)\triangleq R^{\rm data}(\omega)+R^{\rm ad}(\omega,p) \tag{10}$$

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

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

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 [13], [14]. To ensure that choosing $\omega=0$ (i.e., no data reward) is feasible to the problem, we assume that $C\ge D(0)$. 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 to guarantee the fairness among the users choosing to watch ads [9].

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 $\omega\in\left(\frac{\Phi}{u'(Q)\theta_{\max}},\infty\right)$,*

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

Note that the random variable y is the value of $x^*(\theta,\omega)$ when $x^*(\theta,\omega)>0$. Hence, both $\mathbb{E}\left[y^2\right]$ 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 $Km^*(\omega,p^*(\omega))=\mathbb{E}[y]\,N^{\rm ad}(\omega)$). When $A\ge\frac{BK\mathbb{E}[y]}{4\mathbb{E}[y^2]}$, 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 $R^{\rm total}(\omega,p^*(\omega))$, subject to $D(\omega)\le C$. We first introduce Proposition 3.

Proposition 3. *Given $C\ge D(0)$, 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 $D(\omega)\le C$ as $\omega\le D^{-1}(C)$. From numerical experiments, $R^{\rm total}(\omega,p^*(\omega))$ 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 $R^{\rm total}(\omega,p^*(\omega))$ is unimodal, we can apply the Golden Section method [28]). 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* $\frac{(\mathbb{E}[y])^2}{\mathbb{E}[y^2]} N^{\mathrm{ad}}(\omega)$ *and* $\mathbb{E}[y]\, N^{\mathrm{ad}}(\omega)$ *increase with* ω*, the operator's optimal unit data reward is given by* $\omega^* = D^{-1}(C)$.

We explain this sufficient condition by discussing the unit data reward ω's influence on $R^{\mathrm{data}}(\omega)$ and $R^{\mathrm{ad}}(\omega, p^*(\omega))$. First, increasing ω can increase $R^{\mathrm{data}}(\omega)$, because more users subscribe. Second, increasing ω has the following impacts on $R^{\mathrm{ad}}(\omega, p^*(\omega))$: (i) (*positive impact*) It increases $N^{\mathrm{ad}}(\omega)$, i.e., more users watch ads; (ii) (*negative impact*) It may decrease $\mathbb{E}[y]$. Under a larger ω, a user can obtain a larger amount of data after watching a few ads. Then, the user's willingness to watch more ads may decrease because of the concavity of the utility function; (iii) (*negative impact*) It may increase $\mathrm{Var}[y]$. Under a larger ω, more users with different valuations θ watch ads, which can increase the variance of y. As discussed in Section III-B, increasing $\mathrm{Var}[y]$ decreases the advertisers' willingness to purchase ad slots. Under a general utility function $u(\cdot)$ and a general distribution of θ, it is challenging to analyze the net effect of the above impacts. Theorem 2 implies that when both $\frac{(\mathbb{E}[y])^2}{\mathbb{E}[y^2]} N^{\mathrm{ad}}(\omega)$ and $\mathbb{E}[y]\, N^{\mathrm{ad}}(\omega)$ increase with ω, the positive impact dominates the negative impacts. In this case, the operator should set ω as large as possible without violating the capacity constraint (11) under the SAR scheme.

A widely considered setting is that each user has a logarithmic utility function (e.g., [21], [22]) and a uniformly distributed type (e.g., [9], [27]). 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., $u(z) = 1 - e^{-\gamma z}$), we can find a numerical example where $\mathbb{E}[y]\, N^{\mathrm{ad}}(\omega)$ may decrease with ω and $\omega^* < D^{-1}(C)$ (i.e., the operator does not use up the capacity for rewards).

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 (we will use symbol ˆ to indicate that the results are obtained under the SUR scheme).

Lemma 2. *Define* $\theta_3 \triangleq \frac{\Phi}{\omega u'(0)}$. *When* $\omega \in \left(\frac{\Phi u(Q)}{F u'(0)}, \frac{\Phi Q}{F}\right)$*, 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.

Based on the thresholds introduced in Lemmas 1 and 2, we characterize the users' decisions in the following proposition.

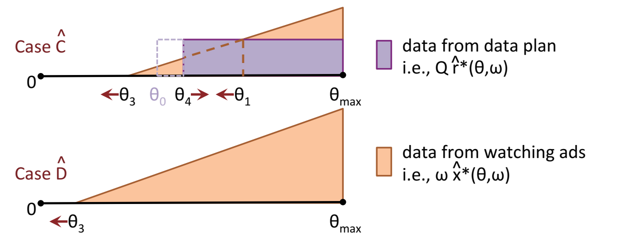


Fig. 3: Illustration of data obtained under the SUR scheme (based on Proposition 5). For $u(z) = \ln(1+z)$, the amount of data obtained via watching ads (i.e., $\omega \hat{x}^*(\theta, \omega)$) linearly increases with θ when $\hat{x}^*(\theta, \omega) > 0$.

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

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

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

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

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

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

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

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

Note that 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 different 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* $\omega \in \left(\frac{\Phi u(Q)}{F u'(0)}, \frac{\Phi Q}{F}\right)$ *(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 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 $\mathbb{E}\left[y^2\right]$ in Proposition 2 by $\hat{N}^{\mathrm{ad}}(\omega)$, $\mathbb{E}[\hat{y}]$, and $\mathbb{E}\left[\hat{y}^2\right]$.

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:

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

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

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

To solve (14)-(15), we first compute $\hat{p}^*(\omega)$ by replacing $\mathbb{E}[y]$ and $\mathbb{E}\left[y^2\right]$ in Theorem 1 under the SAR scheme by $\mathbb{E}[\hat{y}]$ and $\mathbb{E}\left[\hat{y}^2\right]$. 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) $\hat{R}^{\mathrm{total}}(\omega,\hat{p}^*(\omega))$ is discontinuous at $\omega=\frac{\Phi Q}{F}$, and (ii) $\hat{D}(\omega)$ can be decreasing in $\omega\in\left(\frac{\Phi u(Q)}{Fu'(0)},\frac{\Phi Q}{F}\right)$. Based on $\hat{D}(\omega)$, we first search for ω's feasible region, which can be numerically shown to consist of at most three intervals. We can further show that $\hat{R}^{\mathrm{total}}(\omega,\hat{p}^*(\omega))$ is either monotone or unimodal in each interval. Hence, we can determine $\hat{\omega}^*$ by comparing the local optimal unit data rewards searched 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* $A>\frac{B^2K}{8F\int_{\theta_0}^{\theta_{\max}}g(\theta)d\theta}$, *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 largely 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.

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,[6] 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 $\hat{N}_{\mathrm{I}}^{\mathrm{ad}}(\omega)$ and $\hat{N}_{\mathrm{II}}^{\mathrm{ad}}(\omega)$ to denote the number of the subscribers that watch ads and the number of the non-subscribers that watch ads, respectively. Let random variables $\hat{y}_{\mathrm{I}}$ and $\hat{y}_{\mathrm{II}}$ 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 $p_{\mathrm{I}}>0$. If $\hat{N}_{\mathrm{I}}^{\mathrm{ad}}(\omega)>0$, the number of these slots purchased by each advertiser is $\hat{m}_{\mathrm{I}}^*(\omega,p_{\mathrm{I}})=\frac{\max\{B-p_{\mathrm{I}},0\}}{2A}\frac{(\mathbb{E}[\hat{y}_{\mathrm{I}}])^2}{\mathbb{E}\left[\hat{y}_{\mathrm{I}}^2\right]}\hat{N}_{\mathrm{I}}^{\mathrm{ad}}(\omega)$; otherwise, $\hat{m}_{\mathrm{I}}^*(\omega,p_{\mathrm{I}})=0$;
- For the slots generated by the non-subscribers, the operator can set $p_{\mathrm{II}}>0$. If $\hat{N}_{\mathrm{II}}^{\mathrm{ad}}(\omega)>0$, the number of these slots purchased by each advertiser is $\hat{m}_{\mathrm{II}}^*(\omega,p_{\mathrm{II}})=\frac{\max\{B-p_{\mathrm{II}},0\}}{2A}\frac{(\mathbb{E}[\hat{y}_{\mathrm{II}}])^2}{\mathbb{E}\left[\hat{y}_{\mathrm{II}}^2\right]}\hat{N}_{\mathrm{II}}^{\mathrm{ad}}(\omega)$; otherwise, $\hat{m}_{\mathrm{II}}^*(\omega,p_{\mathrm{II}})=0$.

The operator's problem under the differentiation is given by:

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

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

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

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

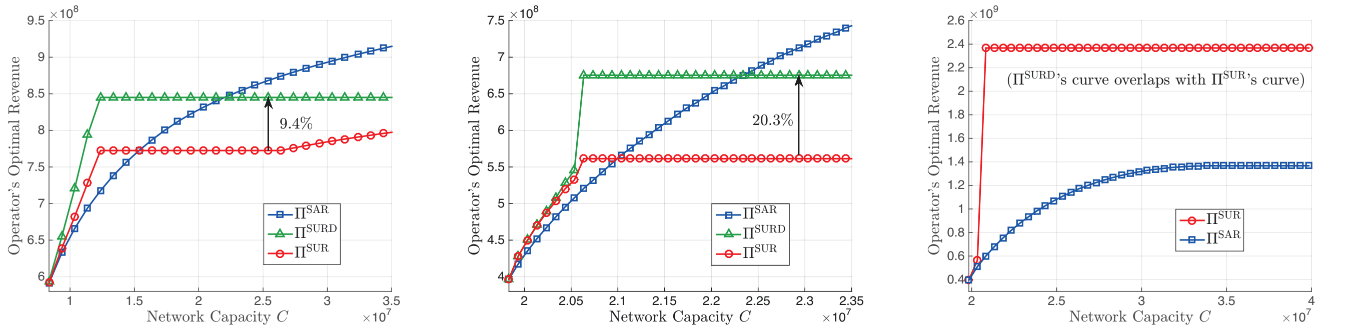
Constraint (18) means that the total number of sold ad slots that correspond to the subscribers should not exceed the number of ad slots generated by the subscribers. Constraint (19) can be explained similarly for the non-subscribers. In fact, only when ω satisfies Case $\hat{C}$ in Proposition 5, both the subscribers and non-subscribers watch ads (i.e., $\hat{N}_{\mathrm{I}}^{\mathrm{ad}}(\omega),\hat{N}_{\mathrm{II}}^{\mathrm{ad}}(\omega)>0$), 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 $\Pi^{\mathrm{SUR}}\triangleq\hat{R}^{\mathrm{total}}(\hat{\omega}^*,\hat{p}^*(\hat{\omega}^*))$, which is the optimal objective value of problem (14)-(15). Let Π^{SURD} denote the optimal objective value of problem (16)-(19). Π^{SURD} is the operator's optimal total revenue under the SUR scheme, considering the differentiation. We compare Π^{SUR} and Π^{SURD} in the following theorem.

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

Hence, the differentiation does not decrease the operator's optimal total revenue (given general $u(z)$ and $g(\theta)$). In

[6] Besides the subscription decision r, a user decides x, e.g., the number of ads to watch within a month. Different from r, the operator does not precisely know the user's decision of x until the end of the month. If the operator can estimate x's range based on the user's historical behavior, it can classify users into different categories and differentiate the corresponding ad slots similarly.



(a) Logarithmic Utility. (b) Exponential Utility (Large Wear-Out Effect A). (c) Exponential Utility (Small Wear-Out Effect A).

Fig. 4: Π^{SAR}, Π^{SUR}, and Π^{SURD} Under Different Network Capacity.

general, it is easy to show that allowing a seller to sell items at different prices does not decrease its revenue. However, the differentiation here affects the ad display rule as well as the pricing, so it is non-trivial to prove Theorem 4. For example, one conjecture is that given any (ω, p) which is feasible to (14)-(15), the operator can choose the same ω and set $p_{\mathrm{I}} = p_{\mathrm{II}} = p$ in (16)-(19) to ensure that the value of objective (16) is no smaller than that of (14). In fact, the conjecture does not hold, because this $(\omega, p_{\mathrm{I}}, p_{\mathrm{II}})$ *may be infeasible to (16)-(19).*

Intuitively, if the optimal unit data reward satisfies Case $\hat{C}$ and the distributions of $\hat{y}_{\mathrm{I}}$ and $\hat{y}_{\mathrm{II}}$ are significantly different, the gap between Π^{SURD} and Π^{SUR} will be large. In the next section, we will show this gap numerically.

V. Comparison Between Rewarding Schemes

We define $\Pi^{\mathrm{SAR}} \triangleq R^{\mathrm{total}}(\omega^*, p^*(\omega^*))$, which is the operator's optimal total revenue under the SAR scheme. In this section, we compare Π^{SAR}, Π^{SUR}, and Π^{SURD}. Since the comparison is challenging under a general utility function $u(\cdot)$, we focus on the logarithmic utility function and exponential utility function in Sections V-A and V-B, respectively.

A. Logarithmic Utility Function

We assume that $u(z) = \ln(1+z)$ and each user's θ follows a uniform distribution. Theorem 5 characterizes the analytical comparison between different schemes under $C \to \infty$.

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

Theorem 5 implies that if the operator has an infinitely 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 Π^{SAR}, Π^{SUR}, or Π^{SURD} has a closed-form expression, and their analytical comparison is challenging. Next, we show the comparison numerically. 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 $B = 5$.

In Fig. 4(a), we plot Π^{SAR}, Π^{SUR}, and Π^{SURD} against C. We can see that only Π^{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 Π^{SAR} and Π^{SUR}. When C is close to $D(0)$, $\Pi^{\mathrm{SAR}} = \Pi^{\mathrm{SUR}}$. 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, $\Pi^{\mathrm{SAR}} < \Pi^{\mathrm{SUR}}$. This is because the SUR scheme can motivate two segments of users to watch ads (by setting $\omega \in \left(\frac{\Phi u(Q)}{F u'(0)}, \frac{\Phi Q}{F}\right)$, 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, $\Pi^{\mathrm{SAR}} > \Pi^{\mathrm{SUR}}$. 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 $\Pi^{\mathrm{SAR}} > \Pi^{\mathrm{SUR}}$). We summarize the results in Observation 1, which has been numerically verified under a truncated normal distribution of θ and other different parameter settings (the comparison between Π^{SAR} and Π^{SURD} is similar to the comparison between Π^{SAR} and Π^{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 Π^{SUR} and Π^{SURD}. We can see that the ad slots' differentiation can improve the operator's revenue under the SUR scheme by at most 9.4%.

B. Exponential Utility Function

We assume that $u(z) = 1 - e^{-\gamma z}$ and θ follows a truncated normal distribution. We 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 $B = 5$. In Fig. 4(b) and Fig. 4(c), we show the comparison between Π^{SAR}, Π^{SUR}, and Π^{SURD} under different degrees of the wear-out effect.

In Fig. 4(b), we consider a large wear-out effect ($A = 0.9$). The comparison between Π^{SAR} and Π^{SUR} (or Π^{SURD}) is similar to that in Fig. 4(a). The SAR scheme achieves a higher revenue than the SUR scheme when C is large. Comparing Π^{SUR} and Π^{SURD} in Fig. 4(b), we observe that the differentiation improves the operator's revenue under the SUR scheme by at most 20.3%. This improvement is larger than that in Fig. 4(a). Intuitively, the benefit of the differentiation is obvious when the subscribers' and non-subscribers' ad watching behaviors (i.e., the distributions of $\hat{y}_{\mathrm{I}}$ and $\hat{y}_{\mathrm{II}}$) are significantly different. Their difference is larger when each user has an exponential utility rather than a logarithmic utility. The truncated normal distribution also increases the difference between the distributions of $\hat{y}_{\mathrm{I}}$ and $\hat{y}_{\mathrm{II}}$, compared with the uniform distribution.

In Fig. 4(c), we consider a small wear-out effect ($A = 0.2$), and have three observations. First, Π^{SAR} may not change with C, which is different from the logarithmic utility situation shown in Fig. 4(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, Π^{SAR} is always no greater than Π^{SUR} (even under a large C), which is different from the logarithmic 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, Π^{SURD}'s curve overlaps with Π^{SUR}'s 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 the 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 observation below (we have numerically verified it under other parameter settings).

Observation 2. *When $u(z) = 1 - e^{-\gamma z}$, (i) under a large A, the SUR scheme achieves a higher operator's 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's revenue than the SAR scheme.*

VI. Conclusion

Mobile data rewarding is an emerging approach to monetize mobile services. We modeled the data rewarding ecosystem and analyzed an operator's optimal 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 study the operator's joint optimization of the data plan and the data rewards.

References

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

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

[3] https://unlockd.com/case-studies/.

[4] http://www.aquto.com/.

[5] https://www.theverge.com/2018/6/25/17500548/att-acquiring-appnexus-ad-tech-tv-time-warner-advertising.

[6] https://www.nttdocomo.co.jp/english/info/media_center/pr/2017/1225_00.html.

[7] 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.

[8] 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.

[9] 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.

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

[11] 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.

[12] 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.

[13] C. Joe-Wong, S. Ha, and M. Chiang, "Sponsoring mobile data: An economic analysis of the impact on users and content providers," in *Proc. of IEEE INFOCOM*, Hong Kong, 2015, pp. 1499–1507.

[14] 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.

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

[16] 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.

[17] 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.

[18] H. Yu, E. Wei, and R. Berry, "Watch ads, earn data: Economics of mobile data rewards," in *Proc. of ACM NetEcon*, Irvine, CA, USA, 2018.

[19] H. Choi, C. Mela, S. Balseiro, and A. Leary, "Online display advertising markets: A literature review and future directions," *Working Paper*, 2017.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

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

[25] 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.

[26] H. Yu, G. Iosifidis, B. Shou, and J. Huang, "Market your venue with mobile applications: Collaboration of online and offline businesses," in *Proc. of IEEE INFOCOM*, Honolulu, HI, USA, 2018.

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

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