Personalized Pricing through User Profiling in Social Networks

Qinqi Lin, Lingjie Duan, and Jianwei Huang

***Abstract*—User profiling allows product sellers to identify users' willingness to pay and enable personalized pricing. However, users' information exploited in profiling is usually private and hard to obtain accurately due to users' privacy concerns. With the increasing popularity of social networks, where users reveal their private information through social interactions, more sellers today profile users through their social data. This paper is the first to study how a seller optimizes personalized pricing through user profiling on social networks, where users proactively react by controlling their social activities and information leakage. We formulate and analyze a dynamic Bayesian game played between users and the seller. First, users decide their social activities by trading off the social network benefit against the potential risk of revealing private information. Then, the seller exploits users' profiles to determine the personalized prices for the profiled users and a uniform price for the non-profiled users. It is challenging to analyze the Perfect Bayesian Equilibrium (PBE) of this game due to i) the randomness in user profiling, and ii) the coupling among users' activity levels and that between the seller's pricing decisions and users' social activities. Despite the difficulty, we propose to alternate backward induction and forward induction to successfully solve the PBE. We show the surprising result that users' activity levels do not monotonically decrease as the profiling technology improves. Instead, when user profiling is of high accuracy, the seller strategically chooses a high uniform price to stimulate their increased social activities to profile more users.**

***Index Terms*—social networks, user profiling, personalized pricing, dynamic Bayesian game**

I. Introduction

Recent advances in information technology have enabled *user profiling*, which is a data analytic tool to outline users' preferences, demands, and other characteristics. As a consequence, a product seller can identify (to some extent) how much different users are willing to pay and offer *personalized prices* to these *profiled users*. Otherwise, a product seller offers a *uniform price* to users whom the seller fails to profile, i.e., the *non-profiled users*. For example, Orbitz was reported to differentiate users based on their computers' operating systems and charge higher hotel prices to Mac users than Windows users [1]. Uber also prices users based on their online/social behaviors, which indicates their spatial and temporal patterns of movements [2]. However, users' information exploited in profiling is usually private; hence it is challenging to obtain directly due to users' privacy concerns.

With the ever-increasing penetration of social networks and proliferation of user-generated data, a product seller can better track users' profiles through monitoring users' social network activities. For example, users reading sports news extensively or sharing photos about fitness frequently tend to value sports products and gym services more than average users. Therefore, user profiling by extracting private information from social media data has been gaining increasingly importance and attention. Indeed, recent studies shed light on user profiling in social networks. Reviews [3], [4] systematically surveyed the techniques and methodologies involved. Specifically, some studies used digital records of users' behaviors in Facebook [5], [6] and Twitter [7], [8] to successfully explore their private traits.

Though the profiling technology is ready, how the seller should exploit such technology to help their pricing decisions is unclear. So far, there is no such analytical study in existing studies between user profiling in social media and personalized pricing. Such an interaction is complicated, as personalized pricing can change users' social behaviors and user profiling efficiency. Moreover, users need to weigh between social interaction satisfaction due to *positive network externality* and the risk of being profiled. Our study aims at filling the gap in the related literature by trying to answer the following two key questions:

- *Key Questions 1:* How does the seller's user profiling technology in social networks affect users' social activities?
- *Key Questions 2:* How will the seller optimize his pricing schemes considering user profiling in social media?

To answer the above questions, we face the following challenges in game theoretic modeling and analysis:

- *Double Coupling in Decision-Making:* On one hand, users' social activities are coupled with each other in social media, as each user's satisfaction level depends on the others' activity levels. On the other hand, there is also a coupling between the seller and users. While enjoying social interactions, the users should be aware of the risk of being profiled and charged personalized prices by the seller. The seller, correspondingly, should well balance the uniform pricing to non-profiled users and the personalized pricing to profiled users, as increasing the former pricing motivates more users to be active (and can be profiled) but will also lose some potential customers.

Q. Lin and J. Huang (corresponding author) are with the Shenzhen Institute of Artificial Intelligence and Robotics for Society (AIRS) and the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Email: {qinqilin@link.cuhk.edu.cn, jianweihuang@cuhk.edu.cn}; L. Duan is with Engineering Systems and Design Pillar, Singapore University of Technology and Design, Singapore, Email: lingjie_duan@sutd.edu.sg. This work is supported by the Shenzhen Institute of Artificial Intelligence and Robotics for Society, and the Presidential Fund from the Chinese University of Hong Kong, Shenzhen.

- *Randomness in User Profiling and Pricing-driven Information Revelation:* Even if a user exposes a significant amount of personal information, the seller's probability of successfully profiling this user may still be random due to the constraints in profiling technology or privacy-related regulations. Furthermore, the seller needs to use proper pricing to incentivize (instead of directly controlling) the users' activity levels.

Based on the above, we model and analyze the coupled interaction between users and the seller under incomplete information. Specifically, we aim to shed light on guiding users' social interaction activities under user profiling and the seller's pricing decisions to maximize the sale revenue.

The contributions of this work are summarized below.

- *Novel Personalized Pricing through User Profiling in Social Networks:* To the best of our knowledge, this is the first analytical study on how to set personalized pricing by exploiting user profiling in social networks. Such an interaction is complex as personalized pricing can affect users' social activities and user profiling efficiency.
- *Dynamic Bayesian Game Formulation:* We formulate the interactions between users and the seller under information asymmetry as a dynamic Bayesian game. First, users decide their social activities by trading off the social network benefit against the potential risk of revealing private information. Then, the seller exploits users' profiles to make price offers. It is challenging to analyze it due to i) the randomness in user profiling, and ii) the coupling among users' activities as well as the coupling between the seller's pricing decisions and users' social interactions.
- *Perfect Bayesian Equilibrium Analysis:* By alternating backward induction and forward induction to ensure consistent belief updates over different stages, we fully characterize the Perfect Bayesian Equilibrium (PBE), revealing the threshold structure of users' social interactions and characterizing the seller's optimal pricing decisions. To our surprise, we show that users' social activity levels in social media do not monotonically decrease as the profiling technology improves. Instead, when user profiling is of high accuracy, users tend to increase their activity levels since the seller chooses a high uniform price to stimulate their social activities to profile more users.

The rest of the paper is organized as follows. In Section II, we review the related literature. In Section III, we present the system model about a group of users and the seller with a dynamic Bayesian game. In Section IV, we present the Perfect Bayesian Equilibrium analysis in Stages I and II and explain the results. In Section V, we explore the impact of the user profiling accuracy on personalized pricing. Finally, we conclude this paper in Section VI. **Most of the proofs are technically involved, thus presented in the online technical report [9] due to space limitations.**

II. Related Work

To enable personalized pricing, there is a growing literature focusing on users' product purchase behaviors in online market considering the implication to reveal their private information to product sellers (e.g., [10], [11], [12], [13], [14]). For example, Conitzer *et al.* in [10] analyzed a repeated purchases scenario, where users could take some measures to hide their past purchase records to hinder the personalized pricing. In [12], Aron *et al.* analyzed users' purchase behaviors by further considering the benefit from product/service customization through private information revelation. Koh *et el.* in [13] also considered the effect of privacy loss on users' product purchase decisions. Valletti *et al* in [14] studied how users conceal their private information, given the seller's investment on consumer profiling technology.

The existing studies did not consider the seller's proactive pricing strategy to profile users or affect their behaviors' in social networks. Overall, the linkage between user profiling in social networks and personalized pricing is missing in current studies. Thus, this paper faces unique challenges to analyze the double coupled interactions among users in social networks and between users and the seller under random user profiling. This leads to a rather involved Perfect Bayesian Equilibrium analysis of such dynamic Bayesian interactions.

Several works in personalized pricing studied the Perfect Bayesian Equilibrium (PBE) of their models (e.g., [13], [14], [15]). The standard procedure in analyzing PBE is as follows. First, a belief (usually of threshold structure) in certain players' strategies is proposed. Then, the remaining analysis about PBE is derived using *backward induction* based on the proposed belief. Finally, the equilibrium strategies and the proposed belief are verified considering belief consistency and sequential rationality. While in our work, to ensure consistent belief updates, we propose to *alternate* backward induction and forward induction to successfully analyze our dynamic Bayesian game, as motivated by the advanced game theory literature [16]. Specifically, we *explicitly* use *forward induction* to characterize the structure of the equilibrium belief.

III. System Model

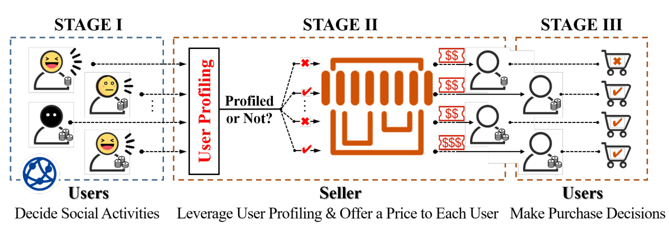

Fig. 1. System Model

We consider the coupled interactions among the seller and users in $\mathcal{N} \triangleq \{1, \dots, n\}$ as a three-stage Bayesian game, as illustrated in Fig. 1. The users' social network is a public platform to interact (e.g., Facebook and Twitter). The shared data is monitored by the platform or even accessible by the public, and the seller exploits users' private profiles from their social activities to enable personalized pricing. For convenience, without any gender bias, we hereafter use female pronouns to refer to the seller and male (or plural) pronouns to refer to each user (or users). The timing of the dynamic game is illustrated below.

- In Stage I, each user $i \in \mathcal{N}$ decides his (normalized) social activity level x_i in the normalized range $[0, 1]$, by trading off the social network benefit (through interacting with the other users) and the potential risk of revealing his private information to the seller. Here, x_i of each user i indicates the time or attention he devotes to social media. Specifically, the normalized maximum level of social activity $x_i = 1$ means user i spends all his available time using the social networks, and $x_i = 0$ implies the user is inactive to use the social media.
- In Stage II, the seller first leverages the profiling technology on users' activities in the social network to figure out their product valuations or willingness to pay (if possible).[1] Once the seller receives an accurate signal about a particular user i's valuation of the product v_i, she would tailor the price for that user based on it, offering a *personalized price* exactly the same as his true valuation $p_i = v_i$. Then the seller turns to those users that are not successfully profiled (due to their inactive social behaviors or the profiling technology constraints) by announcing a *uniform price* p_0 offered to them. Thus, the price offered to each user i is given by,

$$p_i = \begin{cases} v_i, & \text{if profiled by the seller,} \\ p_0, & \text{otherwise.} \end{cases} \tag{1}$$

- In Stage III, each user $i \in \mathcal{N}$ makes his binary purchase decision $d_i \in \{0, 1\}$ regarding whether to purchase the product by comparing his valuation v_i to the offered price p_i.

In the following, we first introduce the user model in the social network and the seller model under user profiling, and then formally formulate the dynamic Bayesian game between the seller and the users.

A. Users' Model under User Profiling

In this subsection, we first model users' social interaction in Stage I and the product purchase benefit in Stage III (see Fig.1), and then formulate the final payoff of each user.

1) Social Interaction in Stage I: Users enjoy the social interaction on the platform, leaving traces or data online. Given the risk of being profiled and charged personalized pricing, each user $i \in \mathcal{N}$ needs to carefully decide his activity level $x_i \in [0, 1]$ (e.g., sharing photos, videos, posts, and comments) in social media. Let $\boldsymbol{x}_{-i}$ summarize the social activity levels of all other users except user i.

[1] We assume that the seller could not directly observe the social activity levels of users. Instead, she only observes the user profiling results. This corresponds to the fact that the seller can only leverage the profiling technology to mine users' social media traces without accessing the raw data due to some privacy regulations.

Each user i will gain a higher satisfaction level (e.g., feeling connected, gaining empathy and identity, and strengthening communication) in the social media if the social activity levels (by himself and other users) are higher [17]. Then, user i's *social network benefit* is modeled as below,

$$q_i(x_i, \boldsymbol{x}_{-i}) = x_i \ln\left(\sum_{j \neq i} x_j + \omega_0\right), \tag{2}$$

where we consider the positive effect of the other users' social activities on each user's social network benefit. The choice of logarithmic function is motivated by Zipf's law [18], which emphasizes that the marginal benefit that each user i experiences is diminishing as the others intensify their social activities.

Specifically, once user i is totally inactive in social networks ($x_i = 0$), he would have zero social network benefit in (2) no matter how active the others are. If x_i is positive, even if all the other users are inactive ($x_j = 0, \forall j \neq i$), user i still has positive intrinsic benefit in (2) with $\omega_0 > 1$ for using some basic services in social media without interaction with others.

2) Purchase Benefit in Stage III: The seller sells products or services to the users. Each user i's valuation for the commodity or service is v_i. For analytic convenience, we assume that users' valuations are independent and identically distributed (*i.i.d.*). Each user i's value v_i is uniformly distributed in the range $[0, \bar{v}]$ with an upper bound $\bar{v}$. The uniform distribution is assumed for ease of exposition of our results later, and our analysis method can be applied to some other distributions such as normal distribution.

We characterize users' purchase decisions in Stage III below, given the offered price p_i in (1),

$$d_i^*(v_i, p_i) = \mathbf{1}(v_i \geq p_i), \tag{3}$$

where $\mathbf{1}(\cdot)$ is an indicator function.

Hence, the final payoff π_i of each user i is modeled as below.

- If user i does not purchase the product ($d_i = 0$), he only has social network benefit $q_i(x_i, \boldsymbol{x}_{-i})$ in (2). That is,

$$\pi_i(x_i, \boldsymbol{x}_{-i} | d_i = 0) = q_i(x_i, \boldsymbol{x}_{-i}). \tag{4}$$

- If user i decides to purchase ($d_i = 1$), the final payoff of user i consists of both the social network benefit $q_i(x_i, \boldsymbol{x}_{-i})$ in (2) and the *purchase surplus* $v_i - p_i$ given by,

$$\begin{aligned} &\pi_i(x_i, \boldsymbol{x}_{-i} | d_i = 1) \\ &= \begin{cases} q_i(x_i, \boldsymbol{x}_{-i}), & \text{if profiled by the seller,} \\ v_i - p_0 + q_i(x_i, \boldsymbol{x}_{-i}), & \text{otherwise.} \end{cases} \end{aligned} \tag{5}$$

Especially, when user i is successfully profiled by the seller, he is left with zero purchase surplus (*valuation-pricing*) due to the personalized pricing. Otherwise, he can have a positive purchase surplus under the uniform pricing p_0.

B. Seller's Model under User Profiling

The seller employs user profiling technology (e.g., [6], [7]) to mine users' social activity data and statistically infer their willingness to pay (if possible). The profiling technology provides a signal about the user's valuation v_i ($\forall i \in \mathcal{N}$) of the product or service.

Similar to [11], [13], [14], we consider a typical binary profiler which will output a signal about users' true valuation if and only if it is sure about the result. We assume that the profiling result about user i follows a Bernoulli distribution. Specifically, user i's valuation v_i is revealed with a profiling probability $\lambda_i \in [0,1]$; otherwise, it is not revealed and non-profiled. Here, the *signal accuracy* λ_i is increasing in user i's social activity level x_i, i.e.,

$$\lambda_i = \delta x_i^{\alpha}, \tag{6}$$

where $0 < \alpha < 1$ and $0 \leq \delta \leq 1$.

Equation (6) is a concave function in x_i, implying that obtaining extra data from user i's social activity becomes less useful as the seller already harvests a lot of critical data. Note that (6) is upper bounded by δ (referred as *user profiling accuracy*). This practically captures the fact that due to the constraints in profiling technology or privacy-related regulations (e.g., General Data Protection Regulation (EU) [19]), whether a user will be successfully profiled is probabilistic even though he is fully active in social networks.

One interpretation of such a profiler assumption is to consider a binary hypothesis testing of user i's valuation at a value v_i with a confidence level λ_i. Our analysis method is also applicable for more general profiling models. For instance, we can extend our analysis to the case of noisy observation, where user i's valuation v_i is a random variable over support $[a_i, b_i]$ following a CDF $G(\cdot)$. Instead of choosing price $p_i^* = v_i$ in (1), the seller in this more general case can choose the optimal individualized price as $p_i^* = \arg\max_{p_i \geq 0} p_i(1 - G(p_i))$. For the clarity of presentation, in this paper, we will focus on the binary profiling model.

In Stage III of Fig.1, the seller gains *(sale) revenue* if users purchase the product at the pricing vector $\boldsymbol{p} = (p_i, \forall i \in \mathcal{N})$ announced for all users. Denote the sets of users who are not profiled and successfully profiled as $\mathcal{N}_0$ and $\mathcal{N}_1$, respectively. As the seller charges p_0 for non-profiled users and $p_i = v_i$ for each user $i \in \mathcal{N}_1$, we model the sale revenue of the seller as follows,

$$U(\boldsymbol{p}) = \sum_{i \in \mathcal{N}_0} p_0 \cdot \mathbf{1}(\text{User } i \text{ Purchase}) + \sum_{i \in \mathcal{N}_1} v_i. \tag{7}$$

C. Dynamic Bayesian Game Formulation

We model the interactions between the seller and users as the three-stage dynamic model in Fig.1, where users and the seller take turns to make their decisions. Besides the social interaction coupling among users, there is also a coupling between users and the seller. The users face the risk of being profiled and charged personalized prices by the seller when enjoying social media. The seller needs to balance the personalized pricing and the uniform price for non-profiled users since a higher uniform price would encourage more users to increase the activity levels but may discourage other non-profiled users from purchasing the product.

As for the game incomplete information structure, in Stage I, each user i initially only knows the common distribution of other users' valuation (instead of the precise private values). In Stage II, through user profiling, the seller obtains more information regarding the valuations of those profiled users; hence the level of incomplete information decreases in the system.

In the analysis of dynamic games, a standard approach is backward induction. However, if we only use the backward induction method, we only manage to solve the following two decisions.

- **Users' Purchase decisions in Stage III:** Each user i's optimal purchase decision given the offered price p_i is characterized as (3).
- **Seller's Personalized Pricing Scheme in Stage II:** Given knowledge of any profiled user i's valuation v_i, the seller's optimal personalized pricing scheme is to offer a personalized price exactly the same as the valuation, i.e., $p_i = v_i$.

Due to the double-coupled decision-making of users' social activities and the seller's uniform pricing (for non-profiled users) in Stages I and II, backward induction alone cannot solve the equilibrium. Instead, to ensure consistent belief updates over different stages, we further combine forward induction motivated by the advanced game theory literature [16]. That is, we propose to *alternate* backward induction and forward induction in the next section to analyze the Perfect Bayesian Equilibrium (PBE) of the dynamic game under incomplete information.

IV. Perfect Bayesian Equilibrium Analysis in Stage I & II

After determining users' purchase decisions in Stage III and the seller's personalized pricing schemes in Stage II, this section continues to analyze the perfect Bayesian equilibrium in Stage I and II.

We explain our PBE analysis as follows. First, in Section IV-A, we use backward induction to analyze the seller's uniform price decision in Stage II, based on a belief of users' social interaction structure. This belief is derived through forward induction considering the users' social activities in Stage I, which we will elaborate on in Section IV-B. In Section IV-C, we combine the analysis in forward induction and backward induction to derive the PBE.

A. Backward Analysis of the Seller's Uniform Price in Stage II

This subsection explores the seller's uniform pricing decision in Stage II. Specifically, we first propose a belief about users' social activity structure in Stage I, based on which we analyze the seller's uniform pricing.

We cannot directly analyze the seller's uniform pricing in Stage II through traditional backward induction. Specifically, a standard backward induction would depend on users' social interactions in Stage I. However, there exists a great diversity of possible outcomes of users' social activity decisions, which brings great complexity to the uniform pricing discussion through traditional backward induction in Stage II. To enable the seller's backward analysis of uniform pricing, we first present the structural result of the users' social decisions in Stage I by using forward induction to be explained in detail in Section IV-B.

Belief 1. *(Belief of Social Interaction Structure) There exists a common valuation threshold $v^* \in [0, \bar{v}]$ in Stage I,*

- *Any user i with a valuation $v_i \geq v^*$ would be inactive, i.e., $x_i(v_i) = 0$;*
- *Any user i with a valuation $v_i < v^*$ would be active, i.e., $x_i(v_i) = 1$.*

Belief 1 will be formally proved in Section IV-B. Intuitively, users with higher valuations prefer to be non-profiled and charged with a low uniform price p_0 instead of the high personalized prices v_i. Thus, he refuses to reveal his profile and is inactive on the social network. If user i's valuation v_i is less than v^*, the user's consideration of social network benefit outweighs the potential of being profiled. Hence, he will choose $x_i = 1$, and the chance of being successfully profiled is δ according to (6).

In the backward induction analysis in Stage II, we will consider an arbitrary value of threshold v^*. Understanding this threshold-based structure of users' social decisions, the seller in Stage II updates her belief of the profiled and non-profiled users' valuations below.

Initially in Stage I, the seller only has prior knowledge about users' valuation distribution (i.e., uniform distribution in $[0, \bar{v}]$). After user profiling, the seller's posterior belief (probability density function) for a profiled user $f(v_i | i \in \mathcal{N}_1)$ is,

$$f(v_i | i \in \mathcal{N}_1) = \frac{1}{v^*}, \quad \text{if } 0 \leq v_i \leq v^*, \tag{8}$$

whereas the posterior probability density function for a non-profile user's valuation $f(v_i | i \in \mathcal{N}_0)$ is,

$$f(v_i | i \in \mathcal{N}_0) = \begin{cases} \frac{1-\delta}{\bar{v}-\delta v^*}, & \text{for } 0 \leq v_i \leq v^*, \\ \frac{1}{\bar{v}-\delta v^*}, & \text{for } v^* \leq v_i \leq \bar{v}. \end{cases} \tag{9}$$

where $\bar{v} > \delta v^*$ since $v^* \leq \bar{v}$ and $\delta \leq 1$. Noticed here, non-profiled users include two types of users: inactive users and active users whom the seller fails to profile.

Before we further derive the optimal uniform price p_0 in the general case of random user profiling $\delta \in (0, 1)$, we first introduce a benchmark of no profiling with $\delta = 0$.

Lemma 1. *If there is no user profiling ($\delta = 0$), the seller only considers the uniform price, which is*

$$p_0^* = \bar{v}/2, \tag{10}$$

and the valuation threshold for users' social interaction $v^* = \bar{v}$.

Next we analyze the general case of $\delta \in (0, 1)$ and compare with the benchmark with $\delta = 0$ in Lemma 1.

Proposition 1. *Given any users' valuation threshold v^* in Stage I, the seller's optimal non-profiled pricing is*

$$p_0^*(v^*) = \begin{cases} \frac{\bar{v}}{2}, & \text{if } v^* \leq \frac{\bar{v}}{2}, \\ v^*, & \text{if } \frac{\bar{v}}{2} < v^* \leq \frac{\bar{v}}{2-\delta}, \\ \frac{\bar{v}-\delta v^*}{2(1-\delta)}, & \text{if } \frac{\bar{v}}{2-\delta} < v^* \leq \bar{v}. \end{cases} \tag{11}$$

Proposition 1 suggests that, if relatively few users with smaller valuation than v^* prefer to reveal their profiles (i.e. $v^* < \frac{\bar{v}}{2}$), the seller mainly cares about the non-profiled revenue and the uniform pricing in (11) degenerates to (10) in Lemma 1. As v^* increases and more users with valuations less than v^* can be profiled, the non-profiled users' valuations under the seller's posterior belief increases according to (9). Thus, she will charge a higher uniform price from those users as in the second case of (11). Finally, as v^* becomes close to $\bar{v}$ and more users can be profiled, very few non-profiled users are with high valuations. Hence, the seller will reduce the non-profiled price to extract more sale revenue from users with lower valuation, as in the third case of (11).

B. Users' Social Activity Decisions in Stage I

In this subsection, we use *forward induction* to prove the threshold structure of users' social activity levels in Belief 1. Specifically, we characterize the equilibrium of users' social interaction game in Stage I by predicting the platform's uniform price p_0.

In Stage I, users cannot completely control their chances to be profiled due to randomness in (6). They also need to estimate the seller's pricing for the non-profiled case in Stage II, which in turn depends on users' social activities levels in Stage I. In this case, each user i in Stage I would maximize his expected utility as follows:

$$\tilde{\pi}_i(x_i, x_{-i}) = x_i \ln(\sum_{j \neq i} x_j + \omega_0) + (1 - \delta x_i^\alpha) \max\{v_i - p_0, 0\}. \tag{12}$$

Recall Belief 1, describing the threshold structure of users' social activity decisions. We prove it through the following three steps:

(i) We first identify users' social interaction in Stage I by predicting the uniform price p_0 as a static Bayesian game with strategic complementarities [20]. This then indicates the existence of Bayesian-Nash equilibrium in users' social activity levels.
(ii) We then verify the convexity of users' expected utility, which means the best response of each user i would be either $x_i = 1$ or $x_i = 0$ instead of social activity level somewhere between ($x_i \in (0, 1)$).
(iii) Finally, we prove that the social interaction equilibrium has a threshold structure.

Detailed proof is in the online technical report [9].

Based on the threshold structure of users' social interaction equilibrium, we further characterize the valuation threshold as below. Denote the cumulative distribution function (CDF) each user i's valuation follows as $F(\cdot)$.

Theorem 1. *The common equilibrium valuation threshold v^* satisfies,*

$$v^* = p_0 + \sum_{m=0}^{n-1} \binom{n-1}{m} \frac{1}{\delta} \ln(m+\omega_0) F(v^*)^m [1-F(v^*)]^{n-1-m}. \quad (13)$$

Proof. (Sketch) To see why the equilibrium threshold is characterized as in (13), we consider the two social activity states in equilibrium: active ($x_i = 1$) and inactive ($x_i = 0$). That is, for a user with a valuation equal to the threshold, the expected utility (consisting of social network benefit and purchase surplus) he will gain when he is active and the expected purchase surplus he will gain if he is inactive would be the same. It follows then that user would be indifferent between active and inactive.

Detailed proof is in the online technical report [9]. □

Alternatively, the formulation above in (13) could be re-written considering the expectation over the number of users who choose to be fully active online except the user himself. Such a number, denoted by n_a, follows the binomial distribution $b(n-1, F(v^*))$.

$$v^* = p_0 + \mathbb{E}_{n_a} \frac{1}{\delta} \ln(n_a + \omega_0). \quad (14)$$

Proposition 2. *The Bayesian-Nash Equilibrium of users' social interaction game in Stage I is unique.*

Notice we are now using forward induction in analyzing users' social activity decisions in Stage I. Specifically, we derive the valuation threshold v^* given the seller's uniform price p_0 in Stage II, i.e., $v^*(p_0)$. We next study how the equilibrium threshold v^* would be affected by (how users predict) the uniform price p_0 in Stage II.

Proposition 3. *The equilibrium valuation threshold $v^*(p_0)$ is nondecreasing in p_0. Specifically,*

1) *If $p_0 < \bar{v} - \ln((n-1)+\omega_0)/\delta$, $v^*(p_0)$ increases in p_0;*
2) *If $p_0 \geq \bar{v} - \ln((n-1)+\omega_0)/\delta$, $v^*(p_0)$ does not change with p_0.*

When the uniform price p_0 is below the threshold, the potential loss in purchase surplus due to personalized pricing is large, especially for high-valuation users, compared to the social network benefit. Thus these high-valuation users choose to be inactive to avoid being profiled. As the uniform price increases, the difference in purchase surplus between being charged by a personalized price and a uniform price diminishes; hence high-valuation users gradually choose to be active online.

C. Perfect Bayesian Equilibrium

In this subsection, we derive the perfect Bayesian equilibrium (PBE) of the whole three-stage game by combining users' social interaction equilibrium threshold v^* in (13) and the seller's equilibrium uniform price p_0^* in (11). In this way, we can meet the two requirements of PBE: (i) belief consistency: the seller's belief in Stage II is derived using Bayes' rule based on users' social activities in Stage I; (ii) sequential rationality: the seller optimizes uniform pricing to maximize her sale revenue given her belief in Stage II [21].

For illustration convenience, we now denote the expected social network benefit of each user i as $\hat{q}(v^*)$ given the valuation threshold v^*, i.e.,

$$\hat{q}(v^*) = \mathbb{E}_{n_a} \ln(n_a + \omega_0). \quad (15)$$

Theorem 2. *The perfect Bayesian equilibrium of the three-stage game is as*

- *Case 1* (All Active in Social Networks) *If*

$$\frac{\bar{v}}{2} \leq \frac{1}{\delta} \ln((n-1)+\omega_0), \quad (16)$$

a) *Users' equilibrium valuation threshold is $v^* = \bar{v}$, i.e., all users will be active online ($x_i = 1, \forall i \in \mathcal{N}$);*
b) *The seller's optimal uniform price is $p_0^* = \frac{\bar{v}}{2}$.*

- *Case 2* (Partially Active in Social Networks) *If*

$$\frac{\bar{v}}{2} > \frac{1}{\delta} \ln((n-1)+\omega_0), \quad (17)$$

a) *Users' equilibrium valuation threshold is $v^* = v^\dagger < \bar{v}$, where $v^\dagger$ is the unique solution to $\frac{\bar{v}+(\delta-2)v^\dagger}{2(1-\delta)} + \frac{\hat{q}(v^\dagger)}{\delta} = 0$. i.e., there exist some users with high valuation choose to be inactive;*
b) *The seller's optimal uniform price is*

$$p_0^* = v^\dagger - \frac{\hat{q}(v^\dagger)}{\delta} < \bar{v} - \frac{1}{\delta} \ln((n-1)r+\omega_0). \quad (18)$$

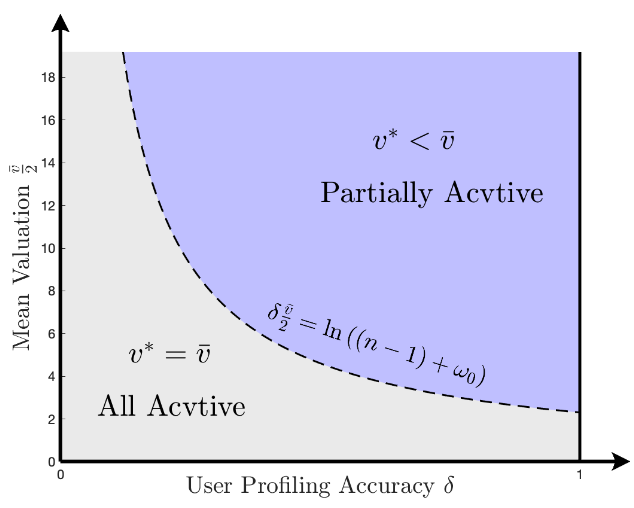

Fig. 2. Perfect Bayesian Equilibrium under Mean Valuation $\frac{\bar{v}}{2}$ and User Profiling Accuracy δ

Fig. 2 illustrates Theorem 2 which jointly considers the mean valuation and the user profiling accuracy. The insights are as follows.

- In Case 1 *(All active)*, with a small purchase surplus (small mean valuation $\frac{\bar{v}}{2}$), users are all going to fully expose themselves $x_i = 1$, without worrying too much about the purchase loss when being profiled. This, in turn, leads to the fact that non-profiled users would be dispersing uniformly in valuation. Hence, the platform maximizes the sale revenue from those non-profiled users by setting the uniform price exactly the mean valuation.
- In Case 2 *(Partially Active)*, where valuation for the commodity is significant, being profiled brings a certain extent of purchase surplus loss to the users so that some high-valuation users would choose to be inactive. In this case, the seller could set the uniform price higher than the mean valuation $\frac{\bar{v}}{2}$, in order to extract the sale revenue from those non-profiled users with higher valuations.

We can show that under the perfect profiling case with $\delta = 1$, the equilibrium valuation threshold $v^* = \bar{v}$, similar as case 1 in Theorem 2 with small δ (including the benchmark of $\delta = 0$). It then indicates that v^* does not monotonically changes in δ. This is because the valuation threshold $v^* = \bar{v}$ for both $\delta = 0$ and $\delta = 1$, and $v^* < \bar{v}$ for some $\delta \in (0, 1)$. Hence, the equilibrium valuation threshold cannot monotonically change in δ.

Corollary 1. *For perfect profiling with $\delta = 1$, the equilibrium valuation threshold is $v^* = \bar{v}$, i.e., all users are just active in equilibrium ($x_i = 1, \forall i \in \mathcal{N}$), and the platform's optimal uniform price is $p_0^* = \bar{v} - \ln((n-1) + \omega_0)$.*[2]

Compared with the no profiling case in Lemma 1, the optimal price p_0^* under $\delta = 1$ is different even though both cases induce all users to be active $v^* = \bar{v}$. The difference lies in the role of the uniform price set by the seller. In the no profiling case $\delta = 0$, uniform price is designed purely to maximize sale revenue, all from non-profiled users. However, in the perfect profiling case $\delta = 1$, the seller sets the uniform price not only to maximize the sale revenue from non-profiled users but also to encourage more users to be active online. The latter consideration would allow the seller to obtain more revenue through personalized pricing. This accounts for the higher optimal uniform price in the perfect profiling case.

V. Sensitivity Analysis of User Profiling Accuracy δ

After characterizing the PBE, we are ready to provide the sensitivity analysis on user profiling accuracy δ. We are also interested in examining how the seller's revenue is affected by profiling accuracy.

Proposition 4. *Users' equilibrium valuation threshold v^* is not monotonic in user profiling accuracy δ. Specifically,*

(1) (All Active) *If*

$$\delta \leq \frac{2}{\bar{v}} \ln\left((n-1) + \omega_0\right), \tag{19}$$

both users' equilibrium valuation threshold v^* and the seller's optimal uniform price does not change with δ.

(2) (Partially Active) *If*

$$\delta > \frac{2}{\bar{v}} \ln\left((n-1) + \omega_0\right), \tag{20}$$

users' equilibrium valuation threshold v^* first decreases and then increases in δ, whereas the seller's optimal uniform price increases in δ.

[2]Here any uniform price that is no smaller than $\bar{v} - \ln((n-1) + \omega_0)$ is optimal. Without loss of generality, we will choose the lowest value, and any value above it would not change the equilibrium payoff of the seller and users.

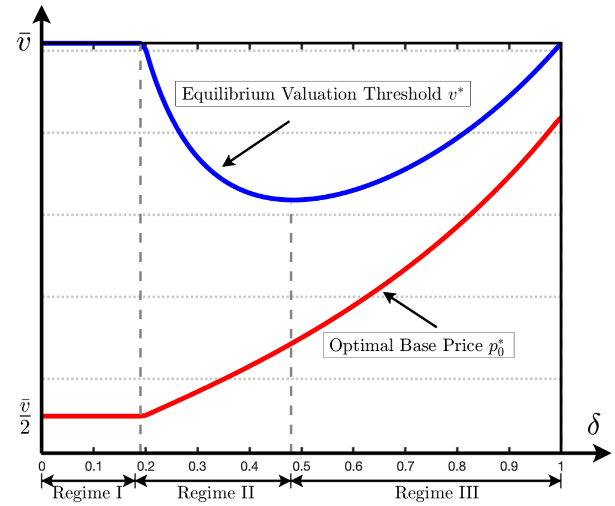

Fig. 3. Equilibrium Valuation Threshold v^* and Optimal Uniform Price p_0^* under Different Users Profiling Accuracy δ

Proposition 4 shows that users' activity levels first stay as a constant, then decrease, and finally increase as the user profiling δ improves. We illustrate such a trend in Fig.3, where the blue curve and the red curve denote users' equilibrium valuation threshold v^* and the seller's optimal uniform price p_0^* respectively. Based on Proposition 4, we further divide the range of δ into three regimes, which are discussed below.

- **Regime I** (All Active): In this regime, δ is so small that it's hard to profile users even when they are fully active in social networks. Thus, all users choose to be active. This, in turn, leads to the fact that non-profiled users disperse uniformly in their valuations. Hence, the seller maximizes the sale revenue from them by setting the uniform price equal to the mean valuation as in the no profiling case ($\delta = 0$).
- **Regime II** (Partially Active with Decreasing v^*): As δ continues to increase, users are more likely to be profiled, leading to a higher risk of being charged personalized prices when being active in the social network. Especially, a user i with a high valuation v_i would suffer a great loss in purchase surplus from it, motivating him to be inactive $x_i = 0$ to prevent from being profiled (v^* decreases). Hence, the seller would increase the uniform price to extract revenue from those non-profiled users with $v_i > v^*$, as the mean valuation of non-profiled users increases with δ.

- **Regime III** (Partially Active with Increasing v^*): After reaching the minimal equilibrium valuation threshold, inactive users in Regime II suffer less from being profiled due to the increasing of uniform price p_0^*. Thus, those users gradually choose to be active and enjoy the social networks again. In turn, the seller will keep increasing uniform price p_0^* to capture the sale revenue from those remaining high-valuation non-profiled users as well as motivate more users to share information online.

Although users' social activities may decrease for a specific range of profiling accuracy, as discussed for Regime II above, we find through numerical studies that the size of profiled users always increases in the accuracy.[3] Overall, the improvement of user profiling accuracy makes it easier to profile users for the seller.

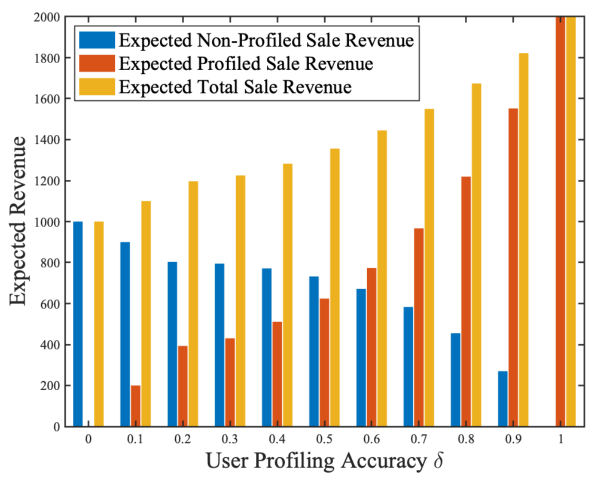

Fig. 4. Seller's Sale Revenue under Different User Profiling Accuracy δ

We finally explore how the profiling technology affects the seller's sale revenue. Fig. 4 numerically illustrates the seller's total sale revenue (including its *non-profiled* revenue part and *profiled* revenue part) versus the profiling accuracy δ at the PBE, where we consider 100 users with uniformly distributed valuations and fix the mean valuation $\frac{\overline{v}}{2}$ as 20. We can see that the seller's profiled sale revenue increases with the users' profiling accuracy because the size of profiled users increases in δ. Meanwhile, the non-profiled sale revenue decreases in δ. To some extent, the total sale revenue is transferring from the non-profiled part to the profiled part as the profiling technology advances. The total sale revenue is still increasing, hence benefits from the improvement of profiling accuracy.

VI. Conclusion

In this paper, we study personalized pricing through user profiling in social networks. We formulate the interactions between the seller and users as a dynamic Bayesian game. It is challenging to analyze the game given the double coupled structure and the randomness in user profiling. By alternating backward and forward induction, we characterize the Perfect Bayesian Equilibrium, and find the surprising result that users' activity levels first decrease and then increase as the profiling accuracy improves. Meanwhile, the uniform price is kept on rising by the seller to stimulate users' social activities.

For future work, we will investigate the case where the seller will experience a cost when profiling users. We will see how the seller will balance the tradeoff between the profiling cost and the sale revenue from more profiled users.

[3]Details are in the online technical report [9] due to space limitations.

References

[1] "On orbitz, mac users steered to pricier hotels," https://www.wsj.com/articles/SB10001424052702304458604577488822667325882, 2012.

[2] "Uber charges more if they think you're willing to pay more," https://www.forbes.com/sites/nicolemartin1/2019/03/30/uber-charges-more-if-they-think-youre-willing-to-pay-more, 2019.

[3] M. Bilal, A. Gani, M. I. U. Lali, M. Marjani, and N. Malik, "Social profiling: A review, taxonomy, and challenges," *Cyberpsychology, Behavior, and Social Networking*, vol. 22, no. 7, pp. 433–450, 2019.

[4] J. Tang, L. Yao, D. Zhang, and J. Zhang, "A combination approach to web user profiling," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 5, no. 1, pp. 1–44, 2010.

[5] G. Farnadi, J. Tang, M. De Cock, and M.-F. Moens, "User profiling through deep multimodal fusion," in *Proc. of the Eleventh ACM International Conference on Web Search and Data Mining*, 2018, pp. 171–179.

[6] M. Kosinski, D. Stillwell, and T. Graepel, "Private traits and attributes are predictable from digital records of human behavior," *Proceedings of the national academy of sciences*, vol. 110, no. 15, pp. 5802–5805, 2013.

[7] D. Garcia, M. Goel, A. K. Agrawal, and P. Kumaraguru, "Collective aspects of privacy in the twitter social network," *EPJ Data Science*, vol. 7, pp. 1–13, 2018.

[8] S. Liang, X. Zhang, Z. Ren, and E. Kanoulas, "Dynamic embeddings for user profiling in twitter," in *Proc. of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 1764–1773.

[9] Q. Lin, L. Duan, and J. Huang, "Online technical report," https://www.dropbox.com/sh/ah69t9tzcigh8ao/AABmXOGywWdJo8pBzx-gP1sWa?dl=0, Tech. Rep., 2021.

[10] V. Conitzer, C. R. Taylor, and L. Wagman, "Hide and seek: Costly consumer privacy in a market with repeat purchases," *Marketing Science*, vol. 31, no. 2, pp. 277–292, 2012.

[11] P. Belleflamme and W. Vergote, "Monopoly price discrimination and privacy: The hidden cost of hiding," *Economics Letters*, vol. 149, pp. 141–144, 2016.

[12] R. Aron, A. Sundararajan, and S. Viswanathan, "Intelligent agents in electronic markets for information goods: customization, preference revelation and pricing," *Decision Support Systems*, vol. 41, no. 4, pp. 764–786, 2006.

[13] B. Koh, S. Raghunathan, and B. R. Nault, "Is voluntary profiling welfare enhancing?" *Management Information Systems Quarterly*, vol. 41, no. 1, pp. 23–41, 2017.

[14] T. Valletti and J. Wu, "Consumer profiling with data requirements: Structure and policy implications," *Production and Operations Management*, vol. 29, no. 2, pp. 309–329, 2020.

[15] S. N. Ali, G. Lewis, and S. Vasserman, "Voluntary disclosure and personalized pricing," in *Proc. of the 21st ACM Conference on Economics and Computation*, 2020, pp. 537–538.

[16] P. Battigalli, "Rationalization in signaling games: Theory and applications," *International Game Theory Review*, vol. 8, no. 01, pp. 67–93, 2006.

[17] C. Seiter, "The secret psychology of facebook: Why we like, share, comment and keep coming back," *Date: April*, vol. 23, p. 2015, 2015.

[18] B. Briscoe, A. Odlyzko, and B. Tilly, "Metcalfe's law is wrong-communications networks increase in value as they add members-but by how much?" *IEEE Spectrum*, vol. 43, no. 7, pp. 34–39, 2006.

[19] "General data protection regulation," https://gdpr-info.eu, 2019.

[20] P. Milgrom and J. Roberts, "Rationalizability, learning, and equilibrium in games with strategic complementarities," *Econometrica: Journal of the Econometric Society*, pp. 1255–1277, 1990.

[21] D. Fudenberg and J. Tirole, *Game Theory*. MIT Press, 1991.