Privacy-Aware Online Social Networking With Targeted Advertisement

Guocheng Liao, *Member, IEEE*, Xu Chen, *Senior Member, IEEE*, and Jianwei Huang, *Fellow, IEEE*

***Abstract*— In an online social network, users exhibit personal information to enjoy social interaction. The social network provider (SNP) exploits users' information for revenue generation through targeted advertisement, in which the SNP presents advertisements to proper users effectively. Therefore, an advertiser is more willing to pay for targeted advertisement to promote his product. However, the over-exploitation of users' information would invade users' privacy, which would negatively impact users' social activeness. Motivated by this, we study the privacy policy (policies) of the SNP(s) with targeted advertisement, in both monopoly and duopoly markets. We characterize the privacy policy in terms of the fraction of users' information that the provider should exploit, and formulate the interactions among users, advertiser, and SNP(s) as a three-stage Stackelberg game. By leveraging the model's supermodularity property, we prove the threshold structure of users' equilibrium information levels. We discover the overall information that can be exploited by an SNP is non-monotonic in the exploitation fraction. Monopoly (one SNP) study shows our proposed optimal privacy policy helps the SNP earn even more advertisement revenue than full exploitation policy does. The situation of the duopoly market is much more complicated. In that case, if the service quality gap between the two SNPs is large, the stronger SNP will choose a conservative privacy protection policy that drives the other SNP out of the market. However, if the service quality gap is small and the advertisement revenue is promising, the stronger SNP would choose an aggressive policy to exploit the advertisement revenue and both SNPs will have positive market shares.**

***Index Terms*— Privacy, online social networks, targeted advertisement.**

Manuscript received May 22, 2021; revised October 13, 2021; accepted December 12, 2021; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor K. Ren. Date of publication January 10, 2022; date of current version June 16, 2022. This work was supported in part by the Shenzhen Science and Technology Program under Grant JCYJ20210324120011032; in part by the Shenzhen Institute of Artificial Intelligence and Robotics for Society; in part by the Presidential Fund from The Chinese University of Hong Kong, Shenzhen; in part by the National Science Foundation of China under Grant U20A20159 and Grant 61972432; in part by the Program for Guangdong Introducing Innovative and Entrepreneurial Teams under Grant 2017ZT07X355; and in part by the Pearl River Talent Recruitment Program under Grant 2017GC010465. Part of the results was presented in INFOCOM 2020 [1] [DOI: 10.1109/INFOCOM41043.2020.9155500]. *(Corresponding author: Jianwei Huang.)*

Guocheng Liao is with the School of Software Engineering, Sun Yat-sen University, Zhuhai 519082, China (e-mail: liaogch6@mail.sysu.edu.cn).

Xu Chen is with the School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China (e-mail: chenxu35@mail.sysu.edu.cn).

Jianwei Huang is with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen 518172, China, and also with the Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen 518129, China (e-mail: jianweihuang@cuhk.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNET.2021.3137513, provided by the authors.

Digital Object Identifier 10.1109/TNET.2021.3137513

I. INTRODUCTION

A. Background and Motivation

ONLINE social network plays an important role in people's daily life. Users can enjoy a wide variety of services, such as chatting with their friends [2], [3], updating real-time status [4], sharing photos [5], watching news [6], and exchanging knowledge [7]. During these activities, users could leave traces of their personal information such as preferences and browsing history, which can be valuable to business providers.

This explosion of users' data generation and the advancement of big data analysis enable the social network provider (SNP) to attract more users and generate revenue through ways such as targeted advertisement. For example, Facebook's advertisement brought 98% of total revenue in 2020 [8]. This is because users' activities on the social network (e.g., browsing history and updated status) reveal their personal characteristics and preferences. By exploiting such valuable information, the SNP can present the advertisements effectively to those who are more likely to be interested in the related products [9], [10]. For example, showing luxury product advertisements to wealthy users would be more effective than to average users. Advertisements promoting sports shoes can be directed to users who share sports-related articles. Such effective targeted advertisement can significantly improve advertisement efficiency and the revenue for the advertisers, and hence the revenue for the SNP.

However, the exploitation of users' information would compromise users' privacy. Prior studies have demonstrated the privacy issues in the online social network (e.g., Facebook), especially in the context of targeted advertisement (e.g., [11]–[14]). For example, Cabañas *et al.* in [11] suggested that 73% European Union Facebook users are labeled with potential sensitive interests for advertisement. Privacy incidents (e.g., [15]–[17]) happened frequently, such as the Facebook's data leakage incident in which millions of users' data is illegally used by third-party Cambridge Analytica [15].

Users have been gradually realizing the potential privacy threat if the SNP possesses much of their personal information. They feel a lack of control over their personal information [18] or are concerned about underlying data misuse and privacy leakage [19] to an unknown third party. The privacy-aware users would take into account the privacy concerns when deciding how much information to expose during social interaction. Thus, it is important for the SNP to implement

an effective privacy protection policy, to balance the trade-off between users' privacy protection and data-driven business revenue generation.

In this work, we focus on the SNP's privacy policy optimization problem considering targeted advertisement, in the presence of users' privacy concerns. We quantify the privacy policy according to the fraction of users' information that the SNP would extract for his business.[1] We begin with the monopoly case involving a single SNP. We further extend our study to the duopoly case in which two major SNPs compete with each other. We aim at answering the following key questions:

- How much will a user expose himself on the social network considering the issue of privacy invasion?
- In the monopoly case, how should the SNP adopt a proper privacy policy to maximize his utility, balancing the social network benefit and the targeted advertisement revenue?
- In the duopoly case, how would users with diverse social preferences choose between the SNPs, and how should the SNPs decide the equilibrium privacy policies considering the market competition?

To answer these questions, we study the interactions among users, an advertiser, and SNP(s) as a three-stage Stackelberg game. In Stage-I, the SNP(s) decide the privacy policies and the advertisement prices. In Stage II, based on the privacy policy, each user decides their information exhibition levels (and which SNP to choose in the duopoly case). In Stage III, based on the users' information exhibition levels and the SNP's privacy policy and advertisement price, the advertiser decides whether to invest advertisement to the targeted users.

B. Contributions

The main contributions of this paper are as follows:

- *Problem formulation.* To the best of our knowledge, this is the first theoretic study of SNP's privacy policy optimization with targeted advertisement. The problem is practically critical, considering the frequent privacy incidents happening to major SNPs over recent years.
- *Users' equilibrium information exhibition level.* By exploiting the supermodularity property of users' social interactions, we reveal a threshold structure of users' equilibrium information exhibition level with respect to information exploitation fraction (i.e., SNP's privacy policy). We find that the overall amount of information that the SNP can obtain is non-monotonic in the information exploitation fraction.
- *SNP's optimal privacy policy in monopoly.* We solve the SNP's utility maximization problem to analytically derive the SNP's optimal privacy policy, under several general assumptions of the SNP's utility function. Extensive numerical studies show that compared with the full information exploitation benchmark, the SNP can earn even more advertising revenue and users can have higher utilities under the proposed optimal privacy policy.
- *SNPs' equilibrium privacy policies in duopoly.* In duopoly competition between two SNPs with different qualities of social service, theoretic study discovers two types of equilibria, depending on their service quality gap: (1) only the stronger SNP survives, and (2) both SNPs co-exists. If the gap is large, the stronger SNP will choose a conservative privacy policy to drive the other SNP out of the market. However, suppose the gap is small and the advertisement revenue is promising. In that case, the stronger SNP will instead choose an aggressive policy to exploit the advertisement revenue, and both SNPs will have positive market shares.

The rest of this paper is organized as follows. We review the related literature in Section II. We introduce the monopoly system model in Section III, and solve the model in Sections IV, V, and VI. In Section VII, we analyze duopoly competition between two SNPs. In Section VIII, we present some simulation results for both the monopoly and duopoly cases. We conclude this paper in Section IX, and provide all the proofs in the appendix.

II. Literature Review

We will review three groups of literature most related to our work: privacy protection considering targeted advertisement, privacy protection on social networks, and social-aware or quality-aware incentive mechanism.

The first group of studies (e.g., [21]–[23]) considered the choice of privacy policy in the context of targeted advertisement and discussed the corresponding impacts. For example, Goldfarb and Tucker in [21] characterized the degradation of targeted advertisement effectiveness due to privacy regulation that restricts user data collection. Johnson in [22] showed that even when users can choose to block targeted advertisement, increased ability of targeting can still benefit firms. Cummings *et al.* in [23] considered that a user's behavior information could be perturbed with noise in a differentially private manner. The authors identified a counter-intuitive case where increasing privacy level might enable the advertiser to know more about the user, which harms the user.

The above studies [21]–[23] mainly focused on the perspective of the advertiser, and investigate the impacts of privacy policy on the advertisement. They did not capture the SNP's valuation of the social network and its possible decisions. The determination of privacy policy involves many factors such as users' behaviors and SNP's social network benefit. We focus on the perspective of the SNP, who implements the privacy policy to balance the trade-off between social network benefit and targeted advertisement revenue. This enables a new angle of analyzing the impacts of the privacy policy.

The second group of studies (e.g. [24]–[26]) focused on privacy protection in the context of social interactions. For example, Gross and Acquisti in [24] characterized privacy risks in terms of identifiability and presented the amount of user information leakage in an online social network. Gradwohl in [25] focused on users' trade-off between social interaction and privacy loss, and studied the impacts of privacy

[1]Such a metric of privacy is general and is related to some more specific privacy metrics (such as differential privacy [20]) through more detailed characterization of privacy-utility trade-off.

enhancements on users' information sharing and utilities. Liao *et al.* in [26] considered that data reporters are concerned with others' privacy loss due to established social relationship. They characterized data reporters' privacy-preserving reporting and data collector's privacy-preserving mechanisms.

However, the above studies [24]–[26] did not consider the targeted advertisement business in social networks. Recently, targeted advertisement has become a significant revenue source in many online social network providers [27], such as Facebook and Wechat. Considering the impacts of targeted advertisement brings additional difficulties in more complicated interactions among users, the social network provider, and the advertiser. This further leads to a challenging multi-layer optimization problem, which involves game-theoretic analysis of users' behaviors in a social network.

The third group of studies (e.g., [28]–[31]) focused on incentive mechanism design considering users' social relationship or qualities. Nie *et al.* in [28] captured the social network effects in mobile crowdsensing. They analytically derived the service provider's incentive mechanisms, in both complete and incomplete social information scenarios. Nuo and Liu in [29] focused on the vehicular crowdsensing setting and studied the service provider's incentive mechanism considering the vehicular social network effects based on a deep reinforcement learning method. Yang *et al.* in [30] proposed to leverage the social ties among users to motivate their cooperation in mobile crowdsensing, which helps promote users' sensing levels. Han *et al.* in [31] proposed a quality-aware pricing scheme in mobile crowdsensing to recruit participants with reliable sensing quality by leveraging the sub-modularity of the problem. The mechanisms proposed in these studies attempted to exploit the crowd intelligence, while ours focuses on alleviating alleviate users' privacy concerns.

To the best of our knowledge, our work is the first to jointly capture the characteristics of targeted advertisement and social network and study their implications on privacy policy.

III. Monopoly System Model

We first focus on the modeling and analysis of the monopoly case. Fig. 1 illustrates the interactions among users, an advertiser, and a single SNP. The SNP offers the social network service to a group of users, who may expose their personal information (including user profiles and social relationships) when enjoying the social service (left part in Fig.1). This brings the opportunity for the SNP to launch a user-specific business. For example, the SNP enables an advertiser to offer a targeted advertisement service on the social network. In return, the advertiser needs to pay for the advertisement service (right part in Fig. 1). The detailed interactions are as follows:

1. The SNP creates a social network platform.
2. The users interact with each other on the social network through posting personal information, such as, photos or status. The activities constitute the user's personal profile.
3. The SNP provides the advertiser the opportunity of targeted advertisement and announces the corresponding price to the advertiser.

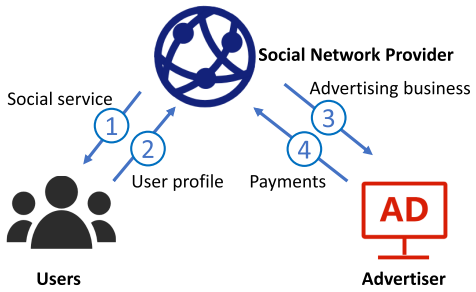


Fig. 1. System model.

4. The advertiser, who wishes to promote his product efficiently, decides whether to invest advertisement based on the price and targeting accuracy. If he decides the invest, the SNP will display the ads to the corresponding users.

Next we introduce the modeling of three parts in more details.

A. Users

We consider a set $\mathcal{N} \triangleq \{1, 2, \ldots, N\}$ of users. A user $n \in \mathcal{N}$ decides his information exhibition level $x_n \in \mathcal{X}_n \triangleq [0, 1]$ on the social network. The minimum value of $x_n = 0$ corresponds to the case where the user does not use the social network. The maximum value of $x_n = 1$ corresponds to the case where the user fully exposes himself to the network actively by always posting photos/status and sharing preferences.

Next, we characterize user n's utility, which consists of his network benefit, privacy loss, and subsidy. Here, we use $\boldsymbol{x}_{-n}$ to denote the decision profile of all other users in set $\mathcal{N}$ except for user n.

$$U_n(x_n, \boldsymbol{x}_{-n}) = b_n(x_n, \boldsymbol{x}_{-n}) - g(\delta x_n) + r \cdot q(\delta x_n). \quad (1)$$

Next we explain each part on the right hand side of (1) in details.

1) Network Benefit $b_n(x_n, \boldsymbol{x}_{-n})$: The user n experiences a *benefit* $b_n(x_n, \boldsymbol{x}_{-n})$ from social interactions, and such a benefit depends on both his own decision x_n and other users' decisions x_{-n}, known as the network effect [32]. For example, the user n feels happy not only when he shares his own photos but also when he sees others share photos. According to Zipf's law [32], the impact of network effect from other users follows a logarithmic formulation, which captures diminishing marginal impact as others' information exhibition levels increase. This motivates the following formulation of the benefit:

$$b_n(x_n, \boldsymbol{x}_{-n}) = \lambda_n x_n \ln\left(\sum_{j \neq n} x_j + \alpha_n\right). \quad (2)$$

Parameter λ_n is the user n's valuation of social interaction. The parameter α_n is larger than 1, meaning that the benefit b_n is always positive when $x_n \neq 0$. A special case is $x_j = 0$ for all $j \neq n$, in which $\lambda_n x_n \ln \alpha_n$ captures the user n's self-enjoyment.

2) Privacy Loss $g(\delta x_n)$: The SNP would exploits a fraction of user information for business, which causes *privacy loss* $g(\delta x_n)$ to the user n. Here $\delta \in [0, 1]$ is the SNP's decision, which represents the fraction of user information that the

SNP will exploit for business such as targeted advertisement. We will discuss more about this in Section III-C. Function $g(\cdot)$ is an increasing in δx_n, as the user's privacy leakage is more significant if the SNP extracts more information [15].

3) Expected Subsidy $r \cdot q(\delta x_n)$: The SNP offers subsidies to incentivize users to participate in the social interactions [33]. Here r is subsidy (in forms such as service discount, virtual currency, and red packet). The function $q(\delta x_n)$ represents the user n's probability of obtaining the subsidy r. Such a function is an increasing function in the user's information level x_n (under a given δ).

To account for the heterogeneity among users, we assume users have different valuation parameters λ_n. When analyzing users' information levels in Nash Equilibrium in Section V, we do not require a specific distribution of the parameter. But when analyzing the SNP's decisions in Section VI, we do require a discrete distribution to obtain some interesting insights, and we will discuss the distribution in detail in Section VI. Meanwhile, we keep the rest of the utility function parameters homogeneous across users.[2]

B. Advertiser

The advertiser wants to deliver advertisements to potential users on the social network to promote his products.[3] When the SNP exploits the user information at a level of δ, he will be able to reveal partially reveal the users' personal references and suggest more accurate advertisement targets to the advertiser. For example, for a sport-related advertiser selling sports shoes, the SNP can identify those users often playing sports according to their personal social activities. These users often have more willingness to purchase related sports products.

As putting the advertisement on the social network is not free, the advertiser needs to carefully trade-off the cost and the potential revenue, which depends on the user's information exhibition level x_n and the SNP's information exploitation level δ, before deciding his advertising investment. Let us denote the set of targeted users recommended by the SNP with the set $\mathcal{M} \subset \mathcal{N}$ with a size M,[4] which depends on the overall size N. We will elaborate on the set $\mathcal{M}$ later in Section III-C.2. Let $V(\delta x_m)$ be the advertiser's value (expected profit) that can be obtained from a user $m \in \mathcal{M}$ watching the advertisement [22], [37]:

$$V(\delta x_m) = v\left(p(\delta x_m) + (1 - p(\delta x_m))\pi\right), \tag{3}$$

where v is the actual profit from a user who end up purchasing buying the product, and $p(\delta x_m)+(1-p(\delta x_m))\pi$ is purchasing probability.

Next, we elaborate the probability $p(\delta x_m)+(1-p(\delta x_m))\pi$ in detail. Here, $p(\delta x_m)$ denotes the probability of successful targeting, which is increasing in both depends on the user's information exhibition level x_m and information exploitation level δ. If the targeting fails, the advertiser can still estimate the probability of users buying the product based on his prior belief $\pi \in [0, 1]$ (obtained from historical sales). A special case of prior $\pi = 0$ means that the advertiser will completely rely on the SNP's recommendation.

The above discussion suggests that the value function $V(\delta x_m)$ is an increasing function of exploited information δx_m, i.e., users with more exploited information are more valuable.

Based on the targeted users' values, the advertiser decides a binary investment vector $\boldsymbol{s} \triangleq [s_m, m \in \mathcal{M}]$, where $s_m = 1$ indicates that he will advertise to user m and $s_m = 0$ means not to user m. He will choose $\boldsymbol{s}$ to maximize the *total expected payoff* from all the targeted users:

$$U_{ad}(\boldsymbol{s}) = \sum_{m \in \mathcal{M}} \left(V(\delta x_m) - p_a\right) s_m. \tag{4}$$

Here the advertiser needs to pay a price p_a to the SNP for each advertisement (to a particular user).

C. Social Network Provider

The SNP needs to optimize his privacy policy δ and targeted advertisement price p_a by jointly considering the revenue obtained from providing social service and enabling targeted advertisement.

First we consider the privacy policy $\delta \in [0, 1]$. When $\delta = 0$, the SNP will not exploit any user information, and there is no user privacy leakage. A higher value of δ means more information exploitation by the SNP and more privacy leakage. In the extreme case of $\delta = 1$, the SNP will aggressively store all users' information, and thus leave no privacy for users. When users' activity levels $\boldsymbol{x}$ are fixed, a higher δ will make the users more valuable to the advertiser. However, a higher δ may discourage users' information sharing and hence may reduce the product of δx_n. Hence, the SNP needs to strike a delicate tradeoff.

Next, we model the SNP's utility U_p which consists of his social network benefit, targeted advertisement revenue, and privacy issues as follows in (5).

$$U_p(\delta, p_a) = (1 - \rho) b_s \left(\sum_{n \in \mathcal{N}} x_n \right) + p_a \cdot n_a - l(\delta). \tag{5}$$

Next, we explain each part on the right hand side of (5) in details.

1) Social Network Benefit $b_s(\sum_{n\in\mathcal{N}} x_n)$: The SNP can obtain more benefit (e.g., revenue from selling related to social network functionalities, charges to software developers using the infrastructure [8], and revenue from value-added services [38]), when users are more active. Thus $b_s(\cdot)$ is an increasing function of total users' information exhibition

[2]As we will show later on, such a one-dimensional heterogeneity already leads to significant analysis challenges as well as interesting engineering insights. We will extend to multidimensional heterogeneity in future work.

[3]When considering multiple advertisers, as long as these advertisers are independent of each other (for example, in different industries), our analysis in this paper is applicable to each of them. For the case of multiple advertisers competing with each other, we will consider a more sophisticated dynamic game model in future work.

[4]The SNP can construct the targeted user set based on users' demographic information and activities on the social network [34]–[36]. For example, Facebook uses its users' demographic and geographic information (e.g., age and location), activities (e.g., clicked ads, likes, sharing, and posting), and social connections, together with some prediction algorithms (e.g., logistic regression) to identify targeted users who are likely to be interested in the advertiser's product [36].

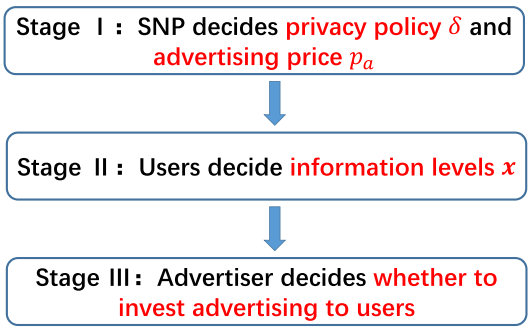


Fig. 2. Three stage Stackelberg game.

levels $\sum_{n\in\mathcal{N}} x_n$. To incentivize users' activeness, the SNP subsidizes would share ρ fraction of his benefit $b_s\left(\sum_{n\in\mathcal{N}} x_n\right)$ to the users [33] and retain the remaining $1-\rho$ fraction. This corresponds to the subsidy r in the users' utility functions (1).[5]

2) Targeted Advertisement Revenue $p_a \cdot n_a$: The SNP would select M users as the targeted users, which form the set $\mathcal{M}$.[6] He will charge the advertiser a price p_a for each single advertisement to a specific user. The number of users watching advertisements is n_a, which should be no larger than M.

3) Privacy Issues $l(\delta)$: The SNP would suffer loss in terms of monetary forfeit or damaged reputation when a privacy incident happens [15], [39], [40].[7]For example, according to IBM research [40], a company would suffer an average total cost of \$3.86 million from a data breach. The function $l(\delta)$ captures the expected loss of such a probabilistic event, which is increasing in the information exploitation level δ.

D. Stackelberg Game Formulation

We model the interactions among users, the advertiser, and the single SNP as a three-stage Stackelberg game, as shown in Figure 2. In Stage I, the SNP first determines the privacy policy δ and advertisement price p_a, by taking into account the impacts on both the users' and the advertiser's decisions. In Stage II, based on the privacy policy, each user $n \in \mathcal{N}$ decides his information exhibition levels x_n considering the possible choices of other users. We model this as a multi-user information exhibition game. In Stage III, based on the users' information exhibition levels and the SNP's decisions, the advertiser decides his advertising strategy $\boldsymbol{s}$. Notice that the advertiser needs to observe the users' information levels before making the advertisement decisions. Thus, the advertiser and the users will make decisions in Stage III and Stage II, respectively. We analyze this three-stage Stackelberg game through backward induction.

[5]Please note that the total budget, i.e., shared benefit to the users, should be more than the total subsidies received by all the users. However, this constraint can always be satisfied by properly designing the subsidy probability $q(\cdot)$ or subsidy r by the SNP. We do not explicitly consider this to avoid over-complication of the model.

[6]Given the overall user set $\mathcal{N}$, we assume a fixed number of targeted users M, independent of other decisions p_a and δ. The SNP could always find up to M number of targeted users even when the exploitation level δ is low. It is up to the advertiser to decide whether to provide advertisements to all these M users. To make our analysis general, we assume that $\mathcal{M}$ can be any arbitrary subset of $\mathcal{N}$.

[7]blue

IV. Monopoly Stage III: Advertiser's Advertisement Decision

In this section, we obtain the advertiser's optimal advertisement decision that maximizes his utility in Stage III. Equation (4) suggests that each user's expected payoff $(V(\cdot)-p_a)$ is independent of each other. Hence we can derive the advertiser's decision for each recommended user separately.

Proposition 1: The advertiser's optimal decision with respect to a targeted user $m \in \mathcal{M}$ is

$$s_m = \begin{cases} 1, & \text{if } v\left((1-\pi)p(\delta x_m)+\pi\right) \geq p_a, \\ 0, & \text{otherwise}. \end{cases} \tag{6}$$

The advertiser decides to advertise to a target user if and only if the value generated by the advertisement is no lower than the price charged by the SNP. The advertiser is more likely to invest in advertisement if the targeting accuracy $p(\delta x_m)$ is higher.

V. Monopoly Stage II: Users' Information Exhibition Levels

In this section, we analyze the users' information exhibition levels in Stage II. We formulate the interaction among users as an information exhibition game, and characterize properties of the Nash Equilibrium.

A. Game Formulation

The users interact with each other in a game-theoretical fashion, as a user's utility in (1) not only depends on his own decision but also depends on other users' decisions.

Game 1 (Information Exhibition Game): The Information Exhibition Game $\left[\mathcal{N}, (\mathcal{X}_n)_{n\in\mathcal{N}}, (U_n)_{n\in\mathcal{N}}\right]$ is defined as follows:

- *Players: users in set $\mathcal{N}$.*
- *Actions: each user $n \in \mathcal{N}$ chooses his information exhibition level x_n in $\mathcal{X}_n = [0,1]$.*
- *Utilities: each user's utility function U_n is given in (1).*

We are interested in the Nash Equilibrium [41] (defined in Definition 1) of the game, in which every user is maximizing his utility given others' decisions. Nash Equilibrium (NE) represents a stable outcome as no user can be better off by unilaterally changing his decision.

Definition 1 (Nash Equilibrium): A decision profile $\boldsymbol{x}^*$ is a Nash Equilibrium if for every user, his decision maximizes his utility, i.e.,

$$U_n(x_n^*, \boldsymbol{x}_{-n}^*) \geq U_n(x_n, \boldsymbol{x}_{-n}^*), \quad \forall x_n \neq x_n^*, \quad \forall n \in \mathcal{N}. \tag{7}$$

The NE may not exist, and in general, characterizing the NE is NP-hard [42]. However, we discover the supermodularity property of the game, which enables tractable analysis of NE.

B. Supermodularity Property

We will show that Game 1 is one of strategic complements with supermodularity property [43], [44]. The property ensures the existence of the NE and the convergence of practical best response updates, under general utility function forms.

Definition 2 (Supermodular Game [43]): A game $[\mathcal{N}, (\mathcal{X}_n)_{n\in\mathcal{N}}, (U_n)_{n\in\mathcal{N}}]$ *is a supermodular game if 1)*

1) *the action set* $\mathcal{X}_n$ *of each user* n *is a compact set, and*
2) *for each player* n*, his utility function* U_n *is continuous and twice differentiable in all the players' decisions, and has increasing differences in* $(x_n, \boldsymbol{x}_{-n})$*, i.e.,*

$$\frac{\partial^2 U_n}{\partial x_n \partial x_j} \geq 0, \text{ for all } j \neq n. \tag{8}$$

More specifically, in a supermodular game, the marginal utility of a player choosing a higher decision (i.e., larger x_n) increases when other players also choose higher decisions (i.e., larger $\boldsymbol{x}_{-n}$). This implies that the best response of a player is a nondecreasing function of other players' decisions.

Theorem 1: The Information Exhibition Game is a supermodular game.

We obtain Theorem 1 by verifying that $\partial^2 U_n/\partial x_n \partial x_j = \lambda_n/(\sum_{k\neq n} x_k + \alpha_n) > 0$. For each user in the information exhibition game, if other users are more active and share their information more, it would be better for him to exhibit himself more as well.

Existing studies of the supermodular game (e.g., [43]) demonstrated some nice properties of the game.

Lemma 1: In the Information Exhibition Game,

1) *there exists at least one pure NE.*
2) *If there are multiple NEs, then there exists a component-wise smallest NE and largest NE.*
3) *Asynchronous best response updates converge to one of the NE.*

Once we confirm the existence of the NE, we further study the properties of the NE.

C. Properties of the NE

1) Assumptions Regarding Users' Utilities: We study some properties of the NE, under some minor assumptions regarding the users' utilities. Although best response updates can find the NE, it is still generally hard to obtain a closed-form characterization of the NE. In order to gain more insights, we consider Assumption 1 and Assumption 2 as discussed next. Note that the previous results in Theorem 1 and Lemma 1 do not require these assumptions.

We assume that each user's utility is a concave function of his own decision. This implies a diminishing marginal return associated with a higher information exhibition level. Assumption 1 can ensure the concavity of the user's utility function.

Assumption 1: The privacy loss function $g(\cdot)$ *is a convex function and the subsidy probability function* $q(\cdot)$ *is a concave function.*

Assumption 1 is relevant to the fact that as the extracted personal information increases, a user will experience much more significant increments in privacy loss [45], [46], and his opportunity of getting the subsidy will increase much more slowly.

Without loss of generality, we also make an assumption regarding the subsidy r, to avoid a trivial case of all-zero information exhibition levels at the NE.

Assumption 2: The subsidy r *satisfies*

$$r > \frac{g'(0)}{q'(0)}. \tag{9}$$

Without Assumption 2 (subsidy is not adequate), when all the users initialize their information exhibition levels from 0, it is possible that no one would like to switch to non-zero information exhibition level.[8] This all-zero information exhibition behavior at the NE is neither beneficial to the SNP nor practical. In other words, Assumption 2 does not restrict our analysis, but just rules our some trivial and impractical cases.

Proposition 2: Under Assumptions 1 and 2, the all-zero information exhibition behavior $\boldsymbol{x} = \mathbf{0}$ *is not an NE.*

2) Analysis of the NE: Then we focus on the non-zero NE and study its properties. We first discover the monotonicity property of $\boldsymbol{x}^*$ with respect to δ. This allows us to write $x_n^*(\delta)$ as a function of δ, representing the user n's *smallest* equilibrium information exhibition level in Stage II given the privacy policy δ. Since the NE is not necessarily unique, we focus on the smallest NE, which can be achieved by best response updates when all users start with the initial zero information levels.

Theorem 2: Under Assumption 1 and 2, there exists a value $\hat{\delta}_n \in (0, 1]$*, such that*

1) *if* $\delta \leq \hat{\delta}_n$*, then* $x_n^*(\delta) = 1$ *for each* $n \in \mathcal{N}$*;*
2) *if* $\delta > \hat{\delta}_n$*, then* $x_n^*(\delta) < 1$ *for each* $n \in \mathcal{N}$ *and* $x_n^*(\delta)$ *is decreasing in* δ*.*

The proof of Theorem 2 is in Appendix B. Theorem 2 shows the user n's equilibrium information exhibition level at the NE $x_n^*(\delta)$ is non-increasing in δ, as illustrated in Fig. 3. More specifically, the NE has a threshold structure with respect to the privacy policy δ. Recall that a higher value of δ means less privacy for users. A user would like to fully expose himself if the privacy policy is more conservative than (i.e., smaller than) a certain threshold. Otherwise, the user will choose a lower information exhibition level as the privacy policy becomes worse. The thresholds can be different for different users. A special case in Proposition 2 is full information exposition for any δ, in which the threshold $\hat{\delta}_n = 1$.

Next, we show that amount of information $\delta \cdot \boldsymbol{x}^*(\delta)$ that the SNP actually obtains is not monotonic in δ.

Proposition 3: Under Assumptions 1 and 2,

1) *when* $\delta \leq \hat{\delta}_n$*,* $\delta \cdot x_n^*(\delta)$ *is increasing in* δ*;*
2) *when* $\delta > \hat{\delta}_n$*,* $\delta \cdot x_n^*(\delta)$ *is decreasing in* δ*.*

The proof of Proposition 3 is in Appendix C. Fig. 4 illustrates the results in Proposition 3. When the privacy policy is more conservative than the threshold, the user will exhibit full information. Thus, a higher information extraction fraction brings more exploitation amount and more privacy leakage. However, when the privacy policy is worse than the threshold, the user chooses to partially exhibit himself and the total exploited information actually decreases in δ.

[8] When $r \cdot \delta q'(0) < \delta g'(0) - \lambda_n \ln \alpha_n$, for all $n \in \mathcal{N}$, all-zero information exhibition behavior is an NE.

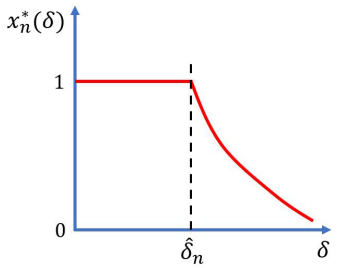


Fig. 3. $\delta \cdot x^*(\delta)$ vs. δ.

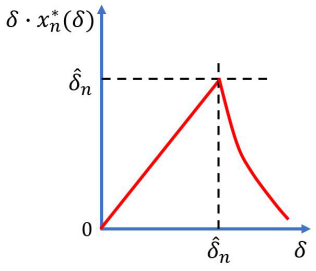


Fig. 4. $\delta \cdot x^*(\delta)$ vs. δ.

Finally, we characterize the impact of the social valuation parameter λ in the function (2) on the NE $\boldsymbol{x}^*$.

Proposition 4: For any two users $n, j \in \mathcal{N}$, we have $x_n^* = x_j^*$ if $\lambda_n = \lambda_j$, and $x_n^* \geq x_j^*$ if $\lambda_n > \lambda_j$.

The proof of Proposition 4 is in Appendix D. Users with higher social valuation are likely to expose themselves more, which makes them benefit more from social interaction.

With the general function forms of privacy loss $g(\cdot)$ and probability of obtaining subsidy $q(\cdot)$, it is difficult to provide a closed-form characterization of the NE. However, Theorem 2, Proposition 3, and Proposition 4 enable us to analyze the SNP's privacy policy and advertisement price without relying on the closed-form expression of NE.

VI. Monopoly Stage I: SNP's Advertisement Price and Privacy Policy

In this section, we study the SNP's optimal advertisement price and privacy policy in Stage I, given the advertiser's investment decision in Stage III and the users' information exhibition levels at the NE in Stage II.

A key challenge of analyzing Stage I lies in the characterization of the advertisement revenue $p_a \cdot n_a$ in the utility function (5). According to Proposition 1, the number of users watching the ads n_a, which is discrete, depends on both privacy policy δ and advertisement price p_a. As n_a is a discrete value, we cannot rely on the first-order condition to compute the solution. To resolve this issue, we incorporate the user type distribution which helps transform the SNP's problem from a discrete optimization problem to a continuous optimization problem. We will discuss this in details in Section VI-A.

Another key challenge of analyzing Stage I is the lack of closed-form solutions of users' NE in Stage II. As a result, we cannot write the SNP's utility explicitly as a function of its decision variables (by incorporating the analysis of Stages II and III). To address this issue, we leverage the threshold structure of NE and non-monotonicity property of the SNP's exploited information (i.e., Theorem 2 and Proposition 3) to characterize the monotonicity property of the SNP's utility function. We will discuss this in details in Section VI-B.

To facilitate the analysis, we first decompose the SNP's utility function in (5) into two parts as follows:

$$U_p(\delta, p_a) = U_a(\delta, p_a) + U_s(\delta), \tag{10}$$

where

$$U_a(p_a, \delta) = p_a \cdot n_a\left(\boldsymbol{x}^*(\delta), \delta, p_a\right), \tag{11}$$

$$U_s(\delta) = (1-\rho) b_s \left(\sum_{n \in \mathcal{N}} x_n^*(\delta)\right) - l(\delta). \tag{12}$$

Equation (11) corresponds to the advertisement revenue, which depends on advertisement price p_a and privacy policy δ. Equation (12) corresponds to the different the social network benefit and the privacy issues, which does not depend on the price p_a. This enables us to maximize the utility (10) in two steps. First, we focus on (11) and find the optimal price $p_a^*(\delta)$ that maximizes the advertisement revenue, given privacy policy δ. Then we further consider (12) and find the optimal policy δ^* that maximizes the total utility in (10).

A. Advertisement Pricing Problem

We first focus on the targeted advertisement pricing problem as follows, given a fixed privacy policy δ.

$$\max_{p_a \in [0,\infty)} \quad U_a(\delta, p_a) = p_a \cdot n_a(\boldsymbol{x}^*(\delta), \delta, p_a) \tag{13}$$

According to Proposition 1, the advertiser will choose to advertise to the user if the advertisement value associated with the user $V(\delta x_m^*(\delta))$ in (3) is no smaller than the price p_a. Thus, we have

$$n_a(\boldsymbol{x}^*(\delta), \delta, p_a) = \left|\{x_m^*(\delta) : V(\delta x_m^*(\delta)) \geq p_a, m \in \mathcal{M}\}\right|. \tag{14}$$

The number of users watching the advertisement $n_a(\boldsymbol{x}^*(\delta), \delta, p_a)$ depends on the targeted users' information exhibition levels $\boldsymbol{x}^*(\delta)$ at the NE. However, without the closed form of $\boldsymbol{x}^*(\delta)$, it is difficult to derive $n_a(\boldsymbol{x}^*(\delta), \delta, p_a)$. We get around this challenge by modeling the distribution of information exhibition level, which enables us to solve the problem (13) from a probabilistic perspective.

1) Distribution of Information Exhibition Level: We capture the distribution of information exhibition level by introducing the distribution of social interaction valuation λ. The SNP could estimate the information of λ distribution through market research. Lemma 4 shows that users with the same parameter λ will choose the same information exhibition level. This means that the distribution of parameter λ uniquely determines the distribution of information exhibition level. Assume that the parameter λ takes K values (types) from the set $\{\lambda_{(1)}, \lambda_{(2)}, \ldots, \lambda_{(K)}\}$ where $\lambda_{(1)} < \lambda_{(2)} < \ldots < \lambda_{(K)}$, with probabilities $P_{(k)} \triangleq Pr(\lambda = \lambda_{(k)})$ for $k = 1, 2, \ldots, K$. Let type-$\lambda_{(k)}$ user's information exhibition level

be $x^*_{(k)}(\delta)$.[9] Therefore, the information exhibition level in the population will follow the same probability distribution: $Pr(x^*(\delta) = x^*_{(k)}(\delta)) = P_{(k)}$, $k = 1, 2, \ldots, K$. With this distribution information, next we focus on an advertisement pricing problem of a single user.

2) Advertisement Pricing Problem Reformulation: We then transform the original targeted advertisement pricing problem in (13) to an optimization problem for a single user whose information exhibition level follows a certain distribution. The advertiser will show an advertisement to a user if the user's advertisement value $V(\delta x^*)$ is no lower than the price p_a, according to Proposition 1. Thus, the probability that the advertiser will advertise to a random targeted user is

$$P_v(p_a, \delta) \triangleq \sum_{k: V\left(\delta x^*_{(k)}(\delta)\right) \geq p_a} P_{(k)}. \tag{15}$$

Based on this, we reformulate the problem in (13) to the following problem:

$$\max_{p_a \in [0,\infty)} \quad \hat{U}_a(p_a, \delta) = M \cdot p_a \cdot P_v(p_a, \delta). \tag{16}$$

Recall that M is the number of targeted users. Thus, the objective $\hat{U}_a(p_a, \delta)$ in (16) is actually the total expected revenue from all the targeted users.

3) Solution of the Reformulated Problem: The optimal price solution to the reformulated problem in (16) is one of the K values, $V(\delta x^*_{(k)}(\delta))$, $1 \leq k \leq K$, depending on which one could yield the highest expected revenue. The insight may not be very clear when the number of types K is large. To get the closed-form solution and obtain more insights, hereafter, we focus on a simplified model with two types. We can generalize our analysis to the multi-type case with more tedious notation.

Assumption 3: Consider two types of the the social interaction valuation parameter λ: a low type λ_L with a probability P_L and a high type λ_H with a probability $P_H = 1 - P_L$.

Based on Proposition 4, let us denote the low-type users' information exhibition level as x^*_L and the high-type users' information exhibition level as x^*_H ($x^*_H \geq x^*_L$), according to Proposition 4. Thus, the advertiser's valuations on low-type users and high-type users are $V(\delta x^*_L(\delta))$ and $V(\delta x^*_H(\delta))$, respectively, as in (3). We show that the optimal price solution to the problem (16) is one of these two values.

Proposition 5: Under Assumption 3, given δ, the optimal price $p^*_a(\delta)$ that solves problem in (16) is

$$p^*_a(\delta) = \begin{cases} V\left(\delta x^*_H(\delta)\right), & \text{if } P_H V\left(\delta x^*_H(\delta)\right) > V\left(\delta x^*_L\right), \\ V\left(\delta x^*_L(\delta)\right), & \text{if } P_H V\left(\delta x^*_H(\delta)\right) \leq V\left(\delta x^*_L\right). \end{cases} \tag{17}$$

And the optimal expected revenue is

$$\hat{U}^*_a(\delta) = M \cdot \max\{P_H V\left(\delta x^*_H(\delta)\right), V\left(\delta x^*_L(\delta)\right)\}. \tag{18}$$

The proof of Proposition 5 is in Appendix E. If the probability P_H is higher than $V\left(\delta x^*_L\right)/V\left(\delta x^*_H\right)$, the optimal advertisement price is the valuation $V(\delta x^*_H(\delta))$ associated with high-type users, in which case the SNP can maximize its revenue by only presenting ads to high-type users.

[9]The main reason of considering discrete values is to derive the closed-form solution of the SNP's optimal advertisement price and privacy policy.

Algorithm 1 Searching Optimal Advertisement Price in the Case of N-Type Users

Input: Set of type indexes $\{1, \ldots, K\}$ with size K; probability of each type $P_{(k)}$, $k = 1, \ldots, K$; information exhibition level of each type $x^*_k(\delta)$, $n = 1, \ldots, N$; valuation function $V(\cdot)$.

1 **Initialization** Initialize revenue $R^* = V(\delta x^*_1(\delta))$, optimal advertisement price $p^*_a = V(\delta x^*_1(\delta))$, and probability $P_v = 1$.
2 **for** $k = 2$ *to* K **do**
3 $P = P - P_{k-1}$;
4 **if** $P \cdot V(\delta x^*_k(\delta)) \geq R^*$ **then**
5 $R^* = P \cdot V(\delta x^*_k(\delta))$;
6 $p^*_a = V(\delta x^*_k(\delta))$;
7 **return** p^*_a

For the general case of K types, we can also find the optimal price with a similar procedure outlined in Algorithm 1 with a complexity $O(K)$. We search the index $k \in \{1, \ldots, K\}$ with the highest revenue $\left(\sum_{j=k}^{K} P_j\right) \cdot V(\delta x^*_k(\delta))$ (Line 4-6) and output $V(\delta x^*_k(\delta))$ as the optimal price.

B. Privacy Policy Problem

Next, we focus on the privacy problem that combines the optimal expected revenue (18) and the utility (12) as follows.

$$\begin{aligned} \max_{\delta \in [0,1]} \quad & \hat{U}^*_a(\delta) + U_s(\delta) \\ & = M \cdot \max\left\{P_H V\left(\delta x^*_H(\delta)\right), V\left(\delta x^*_L(\delta)\right)\right\} \\ & \quad + (1-\rho) b_s \left(\sum_{n \in \mathcal{N}} x^*_n(\delta)\right) - l(\delta). \end{aligned} \tag{19}$$

The main challenge to solve problem (19) is that we do not have a closed-form solution of the NE $\boldsymbol{x}^*(\delta)$, due to intrinsic complexity of the game in Stage II. Further, to make our results general, we hope to solve problem (19) without restricting the value function $V(\cdot)$ and privacy issue function $l(\cdot)$ to specific functions.

1) Narrowing Down the Solution Set: We address the challenge by exploiting the threshold structure of the users' information exhibition levels, as shown in Theorem 2. This helps uncover the monotonicity of the total utility function $\hat{U}^*_a(\delta) + U_s(\delta)$. Recall in Theorem 2 that there always exists a threshold associated with a user. If the privacy policy δ is smaller than the threshold, the user will choose the full information exhibition level. We define the thresholds of low-type users and high-type users as $\hat{\delta}_L$ and $\hat{\delta}_H$, respectively. We narrow the set of solution to interval $[0, \hat{\delta}_H]$.

Lemma 2: Under Assumptions 1-3, the total utility $\hat{U}^*_a(\delta) + U_s(\delta)$ is strictly decreasing in $\delta \in [\hat{\delta}_H, 1]$. Thus the optimal privacy policy δ^* lies in the interval $[0, \hat{\delta}_H]$.

Proof: When the policy is higher than the threshold $\hat{\delta}_H$ (i.e., the privacy is worse), both users' information exhibition

TABLE I
HOW DOES THE SNP CHOOSE THE PRIVACY POLICY?

Cases	SNP's valuation on advertisement business relative to privacy issues	High-type users' probability	Provider's valuation on social network	Privacy policy
Case 1	Low	Does not matter	Does not matter	Most conservative
Case 2	Medium	Does not matter	Does not matter	Conservative
Case 3	High	Low	Does not matter	Medium
Case 4.1	High	High	High	Medium
Case 4.2	High	High	Low	Aggressive

levels $\boldsymbol{x}^*(\delta)$ and the SNP's exploited amount of information $\delta \cdot \boldsymbol{x}^*(\delta)$ would decrease in δ, according to Theorem 2 and Proposition 3. As a result, the total utility $\hat{U}_a^*(\delta)+U_s(\delta)$ would decrease in $\delta \in (\hat{\delta}_H, 1]$. □

Lemma 2 shows that the optimal δ^* is no larger than the threshold $\hat{\delta}_H$. As $\hat{\delta}_H$ is related to the high-type users' valuations on social network λ_H, Lemma 2 suggests that the range of the optimal privacy policy is constrained by the high-type valuation λ_H.

2) Further Convexity/Concavity Assumptions: The monotonicity of the total utility $\hat{U}_a^*(\delta) + U_s(\delta)$ in $[0, \hat{\delta}_H]$ relies on the properties of privacy issue function $l(\cdot)$ and value function $V(\cdot)$. To enable tractable analysis, we make the following assumptions.

Assumption 4: The privacy issue function $l(\cdot)$ is a convex function. The value function $V(\cdot)$ is a concave function.

As the exploited information increases, the users would suffer a higher marginal privacy loss [45], [46], which indicates the SNP's higher marginal loss in terms of financial fine or reputation degradation from privacy incidents. Meanwhile, the SNP would gain less additional information about users. We summarize the solution as follows:

Theorem 3: Under Assumptions 1-4:

- *Case 1: If*

$$M \cdot V'(0) - l'(0) \leq 0, \tag{20}$$

the optimal privacy policy is $\delta^* = 0$;

- *Case 2: If*

$$M \cdot V'(0) - l'(0) > 0 \tag{21}$$

and

$$M \cdot V'(\hat{\delta}_L) - l'(\hat{\delta}_L) \leq 0, \tag{22}$$

the optimal privacy policy is $\delta^* = \tilde{\delta}$, where $\tilde{\delta} \in (0, \hat{\delta}_L]$ satisfies

$$M \cdot V'(\tilde{\delta}) - l'(\tilde{\delta}) = 0; \tag{23}$$

- *Case 3: If*

$$M \cdot V'(\hat{\delta}_L) - l'(\hat{\delta}_L) > 0, \tag{24}$$

and

$$P_H \leq \frac{V\left(\hat{\delta}_H x_L^*(\hat{\delta}_H)\right)}{V\left(\hat{\delta}_H\right)}, \tag{25}$$

the optimal privacy policy is $\delta^* = \hat{\delta}_L$;

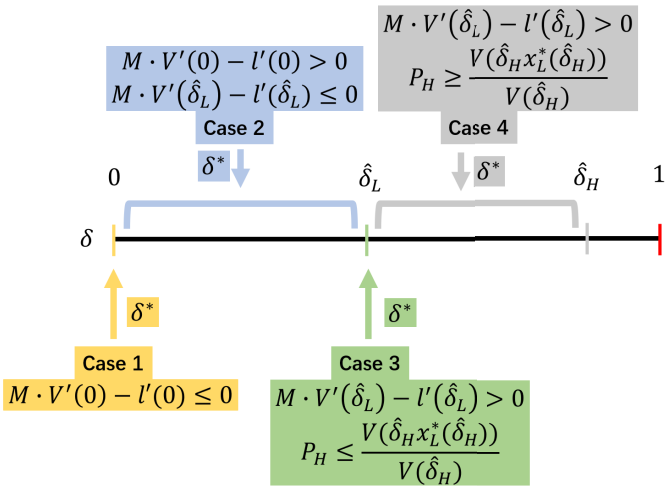


Fig. 5. Illustration of theorem 3.

- *Case 4: If*

$$M \cdot V'(\hat{\delta}_L) - l'(\hat{\delta}_L) > 0, \tag{26}$$

and

$$P_H > \frac{V\left(\hat{\delta}_H x_L^*(\hat{\delta}_H)\right)}{V\left(\hat{\delta}_H\right)}, \tag{27}$$

the optimal privacy policy is $\delta^* \in [\hat{\delta}_L, \hat{\delta}_H]$.

The proof of Theorem 3 is in Appendix F.

Based on Theorem 3, we elaborate how the SNP chooses the optimal privacy policy considering various system parameters, as summarized in Table I. Fig. 5 provides an illustration, in which different colors corresponds to different cases and solutions. Recall that a higher value of δ means worse privacy protection. In Case 1 where the utility loss due to privacy issues dominates the benefit of the advertisement ($l'(0) \geq MV'(0)$), [10] the SNP would choose the best (most conservative) privacy policy with a zero δ, regardless of other factors. However, as the benefit of advertisement becomes more significant (Case 2), the SNP would like to raise the information exploitation δ.

[10]The SNP can learn the advertiser's valuation $V(\cdot)$ through a carefully designed mechanism such as auction [47], [48]. For example, Google is now using first price auction to sell its advertisement spaces [49]. Thus, the SNP can estimate the advertiser's private valuation of advertisement from abundant previous practice of selling advertisement spaces through auctions.

Furthermore, when the advertisement is at a very critical role and dominates the privacy issues ($MV'(\hat{\delta}_L) > l'(\hat{\delta}_L)$), the SNP would heavily extract users' information with an aggressive privacy policy. How aggressive the privacy policy is depends on how the users and SNP value the social network relative to privacy. If high-type users are of a small proportion (Case 3) or the prosperity of social network is critical to the SNP (Case 4.1), the privacy policy is more conservative than in Case 2. Otherwise, in Case 4.2, the SNP would adopt an aggressive privacy policy, as both users and SNP do not care much about the negative impact of privacy issues.

VII. Duopoly Competition: Modeling and Analysis

So far, we have focused on the monopoly case with a single SNP. In this section, we study the duopoly case where there are two SNPs competing in attracting users and gaining revenue from the advertiser through targeted advertisement. We are interested in how users will choose one of the SNPs, and how both SNPs make the equilibrium privacy policy decisions. We first introduce the setup in Section VII-A. Then we study users' behaviors in Section VII-B. Finally, we discuss the SNPs' equilibria of duopoly competition in Section VII-C. We will extend our study to oligopoly competition in future work.

A. System Model

We consider two SNPs: SNP 1 and SNP 2, who compete in provisioning social network service and targeted advertisement business. They offer users the social network service, but with different fixed qualities. Without loss of generality, we assume that SNP 1 provides a better social network service than SNP 2. More specifically, a user can have a higher social benefit with SNP 1 than with SNP 2, given other conditions fixed.

We still model the interactions among SNPs, users, and an advertiser as a three-stage game. In Stage I, two SNPs simultaneously choose their price and privacy policy, respectively. In Stage II, each of the users simultaneously chooses which SNP to join, considering the choices of other users. In Stage III, the single advertiser decides his advertisement investment on both SNPs. Notice that the advertiser's decision of whether to advertise to each recommended targeted user on each platform is the same as in Section IV. In the following, we focus on how users choose SNPs in Stage II and the SNPs' decisions in Stage I.

B. Stage II: Users' Behaviors

A user would choose the SNP with which he can experience a higher utility. Let δ_1 and δ_2 be the privacy policy of SNP 1 and SNP 2, respectively. Similar as (1), we model the utilities of user n choosing SNPs 1 and 2 as follows, respectively,

$$U_{n,1}(\boldsymbol{x}_1^*) = \lambda_n \cdot w_1 \ln\left(\alpha + \sum_{j\neq n, j\in\mathcal{S}_1} x_j^*\right) x_n - g(\delta_1 x_n^*) + r\cdot q(\delta_1 x_n^*), \quad (28)$$

$$U_{n,2}(\boldsymbol{x}_2^*) = \lambda_n \cdot w_2 \ln\left(\alpha + \sum_{j\neq n, j\in\mathcal{S}_2} x_j^*\right) x_i - g(\delta_2 x_n^*) + r\cdot q(\delta_2 x_n^*), \quad (29)$$

where $\mathcal{S}_1$ and $\mathcal{S}_2$ are sets of users who choose SNP 1 and SNP 2, respectively.

Here, we model the heterogeneity in SNPs' social network service by two SNP-dependent parameters w_1 (of SNP 1) and w_2 (of SNP 2). Without loss of generality, we consider $w_1 \geq w_2$. Vector $\boldsymbol{x}_1^*$ and $\boldsymbol{x}_2^*$ are the users' information exhibition levels under NEs in SNP 1 and SNP 2, respectively. In order to enable comparison of user utilities in two SNPs and derive closed-form solution of $\boldsymbol{x}_1^*$ and $\boldsymbol{x}_2^*$, we consider linear format of $g(\cdot)$ and $q(\cdot)$ in the user's utility function as in Assumption 5, so as to obtain some insights. We adopt this assumption for the rest of Section VII. Later in Section VIII-B, we will use other types of function to run the simulations.

Assumption 5: User's privacy loss function is $g(t) = at$ with $a > r$, and the probability of obtaining subsidy is $q(t) = t$.

Next, we examine how users choose SNPs at the equilibrium. It is clear that a user will not choose an SNP if it brings him a non-positive utility. When both SNPs can provide positive utilities, a user will choose the SNP that offers a higher utility. Recall that under Assumption 3, there are two types of users, with low network benefit parameter λ_L and high parameter λ_H, respectively. The following Lemma shows the symmetric choice of users at the equilibrium.

Lemma 3: At the equilibrium, users of the same type would choose the same SNP.

The proof of Lemma 3 in Appendix G. The main idea is to show that users of the same type choosing different SNPs would reduce the network effect and reduce the users' payoffs. With Lemma 3, we can focus on the symmetric equilibrium in Stage II.

We still need to further explore which SNP will be chosen by which type of users. We denote users' action set as $\{0, 1, 2\}$, where 0, 1, 2 mean choosing neither SNP, SNP 1, and SNP 2, respectively. We model users' interactions as the *User Selection Game*.

Game 2 (User Selection Game): The User Selection Game $[\{L, H\}, (\mathcal{A}_n)_{n=L,H}, (U_n)_{n=L,H}]$ is defined as follows:

- *Players: low-type users in set L and high-type users in set H.*
- *Actions: each user n decides his action $a_n \in \mathcal{A}_n \triangleq \{0, 1, 2\}$.*
- *Utilities: a user obtains zero utility if he chooses neither SNP. Otherwise, a user's utilities of choosing SNP 1 and SNP 2 are given in (28) and (29), respectively.*

To derive the Nash Equilibrium of the *User Selection Game*, we first begin with the general case. After that, we will focus on some specific parameter choices that lead to more insights.

1) General Case: Theorem 4, shows the necessary and sufficient conditions of each equilibrium of the *User Selection Game*. We use the tuple (x, y) to denote the equilibrium, in which the first element x is the low-type user's action and the second element y is the high-type user's action.

TABLE II

EQUILIBRIUM OF *User Selection Game*

Necessary and sufficient conditions	Equilibrium
a1. $\delta_1 \le Z_1$, a2. $\delta_1 \le Z_2$, a3. $\delta_1 - \delta_2 \le Z_3$, if $\delta_2 \le Z_4$, a4. $\delta_1 - \delta_2 \le Z_5$, if $\delta_2 \le Z_6$.	$(1,1)$
b1. $\delta_2 \le Z_6$, b2. $\delta_2 \le Z_7$, b3. $\delta_1 - \delta_2 \ge Z_8$, if $\delta_1 \le Z_9$, b4. $\delta_1 - \delta_2 \ge Z_{10}$, if $\delta_1 \le Z_2$.	$(2,2)$
c1. $\delta_2 \le Z_4$, c2. $\delta_1 \le Z_2$, c3. $\delta_2 - \delta_1 \ge -Z_{10}$, c4. $\delta_2 - \delta_1 \le -Z_3$, if $\delta_1 \le Z_1$.	$(2,1)$
d1. $\delta_2 > Z_4$, d2. $\delta_1 > Z_1$, d3. $\delta_1 \le Z_2$, d4. $\delta_1 - \delta_2 \le Z_{11}$, if $\delta_2 \le Z_6$.	$(0,1)$
e1. $\delta_1 > Z_9$, e2. $\delta_2 > Z_7$, e3. $\delta_2 \le Z_6$, e4. $\delta_1 - \delta_2 \ge Z_{11}$, if $\delta_1 \le Z_2$.	$(0,2)$
f1. $\delta_1 \ge Z_2$, f2. $\delta_2 \ge Z_6$.	$(0,0)$

Theorem 4: There exists at least one equilibrium in the User Selection Game. Table II shows the necessary and sufficient conditions of each equilibrium. The constants Z_x, $x = 1, \ldots, 11$, in Table II are defined in Appendix H.

The proof of Theorem 4 is in Appendix H. We interpret each equilibrium and its associated conditions.

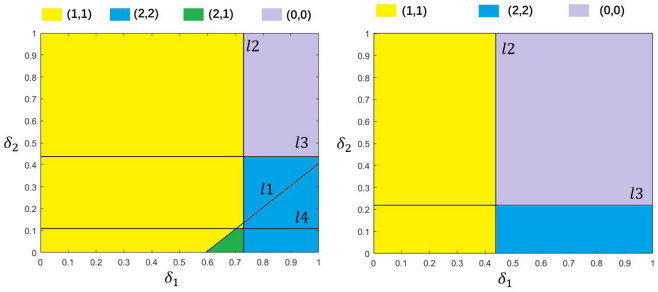
- Equilibrium $(1,1)$: This happens when SNP 1's policy is very conservative (Conditions a1 and a2), and is much more conservative than SNP 2's, if SNP 2's policy is also conservative (Conditions a3 and a4). This means that a conservative enough privacy policy can make one SNP dominate and drive the other SNP out of the market. The insight behind equilibrium (2,2) is similar.
- Equilibrium $(2,1)$: This happens when both SNP 1's and SNP 2's policies are very conservative (Conditions c1 and c2). Meanwhile, two SNPs' policies are similar (Conditions c3 and c4). Thus, neither of the SNPs is dominating, and both SNPs share the market.
- Equilibrium $(0,1)$: This happens when SNP 2's policy is aggressive (Condition d1). Thus, no users choose SNP 2. Meanwhile, SNP 1's policy is medium, i.e., neither too conservative nor too aggressive (Conditions d2 and d3), and is more conservative than SNP 2's (Condition d4). This makes SNP 1 fail to attract low-type users but manage to attract high-type users. The insight behind equilibrium (0,2) is similar.
- Equilibrium $(0,0)$: This happens when both SNPs' policies are very aggressive (Condition f1 and f2). Hence none of the users will choose any of the SNPs.

Notice that the conditions of all the equilibria are not necessarily mutually exclusive. Specifically, conditions of equilibrium $(1,1)$ and conditions of equilibrium $(2,2)$ might hold simultaneously. In that case, both (1,1) and (2,2) are valid equilibria.

TABLE III

LINE EXPRESSIONS IN FIG. 6

Line	Expression
$l1$	$\delta_1 - \delta_2 = Z_3$
$l2$	$\delta_1(a-r) = \lambda_H w_1 \log(\alpha)$
$l3$	$\delta_2(a-r) = \lambda_H w_2 \log(\alpha)$
$l4$	$\delta_2(a-r) = \lambda_L w_2 \log(\alpha)$



(a) Weak advantage of SNP 1. (b) Strong advantage of SNP 1.

Fig. 6. Users' equilibrium illustration.

2) Special Cases: Since the result in Theorem 4 is rather complicated, we provide more insights under some special parameter settings. Recall that SNP 1 provides a better advantage in social service than SNP 2. We wish to understand how such an advantage affects users' decisions, given different policies of both SNPs. Thus, we consider two advantage levels of SNP 1: weak advantage and strong advantage. W Let N_L and N_H represent the number of low-type users and high-type users, respectively. Corollary 1 presents the result in the weak advantage case.

Corollary 1: In the User Selection Game, if both the following conditions hold:

1. $N_H > \bar{N}_H$, *for a specific constant* $\bar{N}_H$;
2. $\hat{w}_2 < w_2 < \bar{w}_2$, *for some constants* $\hat{w}_2$ *and* $\bar{w}_2$;

where the constants $\bar{N}_H$, $\hat{w}_2$, *and* $\bar{w}_2$ *are given in Appendix I, there exists a unique Nash Equilibrium that is one of the following:* $(1,1)$, $(2,2)$, $(2,1)$, *or* $(0,0)$.

The proof of Corollary 1 is in Appendix I. Figure 6a illustrates Corollary 1. More specifically, the horizontal axis represents SNP 1's privacy policy δ_1 and the vertical axis represents SNP 2's privacy policy δ_1. Different colors represent different equilibria. More specifically, yellow, blue, green, and purple regions correspond to equilibrium $(1,1)$, $(2,2)$, $(2,1)$, $(0,0)$, respectively.

To interpret Corollary 1, notice that Condition 1 indicates that the number of high-type users is large enough, which is practical in a large network. Condition 2 in Corollary 1 means that SNP 2's service quality w_2 is at a medium level. In this case, when SNP 2's policy is conservative and SNP 1's policy is a little more aggressive than SNP 2's (green regime in Figure 6a), there exists a unique equilibrium of $(2,1)$, i.e., low-type users choosing SNP 2 and high-type users choosing SNP 1. This shows that when SNP 1's advantage in social service is weak, SNP 2 can leverage a more conservative privacy policy to obtain a positive market share.

Next, we show that the equilibrium of $(2,1)$ is not possible when SNP 1 has a strong advantage over SNP 2, as shown in Corollary 2.

Corollary 2: In the User Selection Game, if all the following conditions hold:

1. $N_H > \bar{N}_H$;
2. $N_L > \bar{N}_L$, *for a specific constant* $\bar{N}_L$;
3. $w_1 > \bar{w}_1$, *for a specific constant* $\bar{w}_1$;

where the constants $\bar{N}_L$ *and* $\bar{w}_1$ *are given in Appendix J, there exists a unique Nash Equilibrium that is one of the following:* $(1,1)$, $(2,2)$, *or* $(0,0)$.

The proof of Corollary 2 is in Appendix J. Figure 6b illustrates Corollary 2. To interpret Corollary 2, notice that Conditions 1 and 2 in Corollary 2 indicate that there are a large number of high-type users and low-type users, respectively. Conditions 3 in Corollary 2 indicates that SNP 1's service quality w_1 is significantly larger than w_2.[11] Corollary 2 shows that if SNP 1's social service is much better than SNP 2's, only one SNP would survive in the competition.

C. Stage I: SNPs' Privacy Policies

Recall that SNPs need to decide their advertisement prices and privacy policies. As the advertiser's advertisement decision in the duopoly case is similar to the monopoly case, the analysis of the SNPs' advertisement prices is also similar. It remains to study the SNPs' privacy policies in Stage I, based on the users' equilibrium choices in Stage II.

We formulate a *Duopoly Competition Game* between two SNPs to study their decisions of privacy policy.

Game 3 (Duopoly Competition Game): The Duopoly Competition Game $[\{1,2\},([0,1],[0,1]),(U_{pi})_{i=1,2}]$ *is defined as follows*:

- *Players: SNP 1 and SNP 2.*
- *Actions: each SNP* $i \in 1,2$*, decides his privacy policy* δ_i *in* $[0,1]$,.
- *Utilities: each SNP* i*'s utility is given in (5).*

We are interested in the equilibrium of the *Duopoly competition Game* (Definition 3), which represents a stable situation in which neither of them can benefit from unilateral deviation.

Definition 3: (Equilibrium of Duopoly Competition Game in Stage I): A profile (δ_1^*, δ_2^*) *is an equilibrium of duopoly competition, if*

$$U_{pi}(\delta_i^*, \delta_j^*) \geq U_{pi}(\delta_i, \delta_j^*), \tag{30}$$

for $\delta_i \neq \delta_i^*$, $i = 1,2$, *and* $j \neq i$.

Based on the results of Corollary 1 and Corollary 2, we present the corresponding the equilibrium of *Duopoly Competition Game*, to see how both SNPs react in both cases.

We first focus on the SNP 1's week advantage case in Corollary 1 and present an equilibrium of only SNP 1 surviving. We begin with some notations. Let $U^{p1}_{(1,1)}$ be the SNP 1's optimal utility when $\delta_1 \in [0, Z_3]$. In this case, $(1,1)$ is always the users' equilibrium regardless of SNP 2's policy. Let $\delta^{p1}_{(1,1)} \in [0, Z_3]$ be the corresponding optimal privacy policy of SNP 1. Let $U^{p1}_{(2,1)}$ be the supremum of SNP 1's utility given the users' equilibrium being $(2,1)$. The expressions of these notations are in Appendix K. We find that when $U^{p1}_{(1,1)} \geq U^{p1}_{(2,1)}$, SNP 1 would dominate and SNP 2 would be driven out of market.

Proposition 6: Assume that the conditions in Corollary 1 hold and $U^{p1}_{(1,1)} \geq U^{p1}_{(2,1)}$, $(\delta^{p1}_{(1,1)}, \delta_2)$ *for any* $\delta_2 \in [0,1]$ *is an equilibrium of the Duopoly Competition Game. SNP 2 obtains a zero market share under these equilibria.*

The proof is in Appendix K. Proposition 6 shows that if SNP 1 is able to generate a higher utility in user equilibrium $(1,1)$ than in $(2,1)$, SNP 1 would like to adopt a conservative privacy policy to attract all the users. This happens when the social benefit is strong or the advertisement benefit is low, which drives SNP 2 out of the market. Otherwise, it is possible for both SNPs to co-exist.

To further consider the possibility of co-existence, we present some additional notations. Let $\delta^{p2}_{(2,1)}$ be the SNP 2's optimal privacy policy given the users' equilibrium being $(2,1)$. Conditioned on SNP 2's policy being $\delta^{p2}_{(2,1)}$, let $\hat{U}^{p1}_{(1,1)}$ be the SNP 1's optimal utility when $\delta_1 \in [0, \delta^{p2}_{(2,1)} + Z_3]$, and $\hat{U}^{p1}_{(2,1)}$ be the SNP 1's optimal utility when $\delta_1 \in (\delta^{p2}_{(2,1)} + Z_3, Z_2]$.[12] The expressions of the above notations are given in Appendix L.

Proposition 7: Assume that the conditions in Corollary 1 hold, and $\hat{U}^{p1}_{(1,1)} < \hat{U}^{p1}_{(2,1)}$, $(\delta^{p1}_{(2,1)}, \delta^{p2}_{(2,1)})$ *is an equilibrium of the Duopoly Competition Game. Here,* $\delta^{p1}_{(2,1)} \in (\delta^{p2}_{(2,1)} + Z_3, Z_2]$ *is presented in details in Appendix L. Both SNPs have positive market shares under the equilibrium.*

Proposition 7 shows another market partition when SNP 1's advantage is weak: both SNPs co-exist. If SNP 1's utility in $(2,1)$ is higher than that in $(1,1)$, which happens when the social service benefit is weak or advertisement benefit is high, SNP 1 would like to adopt an aggressive privacy policy. Thus, SNP 2 is able to obtain a positive market share with a conservative policy.

Next, we show that co-existence is not possible when SNP 1's advantage is strong, corresponding to the case considered in Corollary 2.

Proposition 8: Assume that the conditions in Corollary 2 hold, $(\delta_1^{com}, \delta_2)$ *for any* $\delta_2 \in [0,1]$ *is an equilibrium of the Duopoly Competition Game. Here,* $\delta_1^{com} \in [0, Z_2]$ *is presented in details in Appendix M. SNP 2 obtains a zero market share these the equilibria.*

Proposition 8 shows only one market partition: SNP 1 dominates and SNP 2 is driven out of market, when SNP 1's advantage in social service is strong. This indicates that SNP 1 would leverage its advantage in service quality and adopt a conservative enough privacy policy to attract all the users.

[11] Corollary 1 corresponds to SNP 1's *weak* advantage over SNP 2, and Corollary 2 corresponds to SNP 1's *strong* advantage over SNP 2. To see this, recall that Condition 2 in Corollary 1 indicates that SNP 2's service quality w_2 is greater than a certain value $\hat{w}_2$. On the other hand, Condition 3 in Corollary 2 indicates that SNP 1's service quality is greater than a certain value $\bar{w}_1$. Combining both cases, we can see the gap between SNP 1's service quality w_1 and SNP 2's service quality w_2 is relatively small in Corollary 1 and is relatively large in Corollary 2.

[12] The optimality might not available in the open set in which the users' equilibrium is $(2,1)$. See Appendix L in details. We consider the case where the optimality exists to facilitate the analysis.

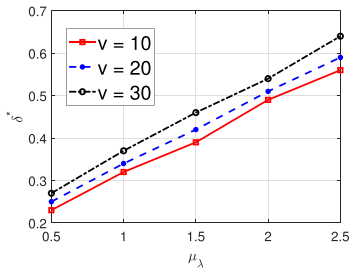


Fig. 7. Optimal privacy policy δ^*.

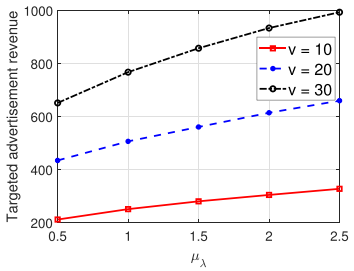


Fig. 8. SNP's advertisement revenue.

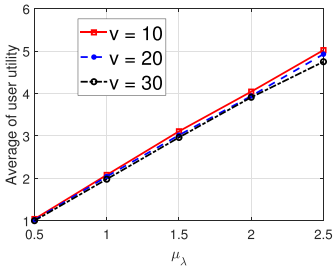


Fig. 9. Average of user utility.

VIII. Numerical Results

We conduct extensive numerical evaluations to gain more useful insights. The evaluations involve both cases: the monopoly case and the duopoly case.

A. Monopoly Case

We investigate the impacts of users' social valuation and SNP's targeted advertisement benefit on the system performance. We also compare the optimal privacy policy $\delta = \delta^*$ in Theorem 3 with other two policies:

- the perfect privacy policy without information invasion (i.e., $\delta = 0$);
- the worst privacy policy with full information exploitation (i.e., $\delta = 1$).[13]

The comparison shows our proposed policy provides more advertisement revenue to the SNP and provides better social interaction experiences and privacy protection to the users. In other words, our proposed policy achieves a win-win result comparing with the benchmarks.

We first introduce the simulation setup. We leverage the Facebook social data [50] obtained from SNAP datasets. Specifically, the data set shows the social connections among Facebook users (i.e., whether there exists an social relationship between them). As for the users' valuation on social interaction λ in user's utility function (1), we assume that λ follows a truncated normal distribution, with a mean of μ_λ and a variance of one in the interval of $(0, 2\mu_\lambda)$. A higher value of μ_λ represents a higher valuation of social interaction. We run our simulations under different values of μ_λ. For the rest parameters of the users' utility function, we use privacy loss $g(\delta x) = 3(\delta x)^2$ and expected subsidy $r \cdot q(\delta x) = 0.01 \times (1 - \exp(-2\delta x))$.

Next, we introduce the parameters of the SNP's utility function. For the SNP's targeted advertisement business, we consider the parameter v in (3), which reflects the importance of advertisement business to the SNP (although it is the advertiser's parameter). We run our simulations under different values of v. For the rest of parameters, we consider the privacy issues $l(\delta) = 50\delta^2$ and the social network benefit $b_s\left(\sum_{n\in\mathcal{N}} x_n\right) = \sum_n x_n \log(\sum_n x_n + 1)$, similar to users' social benefit in (2) based on Zipf's Law.

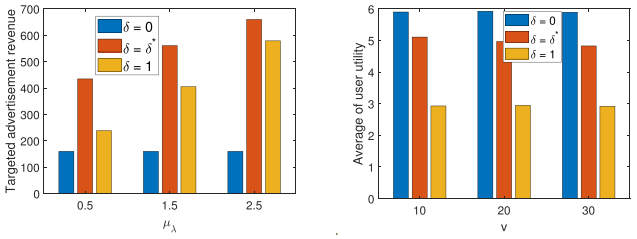
1) Impacts of the Users' Social Valuation and SNP's Advertisement Benefit: Fig. 7 illustrates the optimal privacy policy δ^* under different values of μ_λ and v. We can see that the optimal privacy policy δ^* is increasing in the mean of users' valuations on the social network μ_λ. *As users pay more attention to social interaction than potential privacy issues, the SNP can more aggressively to extract users' information.* Further, given fixed μ_λ, the optimal privacy policy δ^* also increases in the advertisement benefit v. This is expected, as the attraction of the advertisement business motivates the SNP's higher information exploitation to increase efficiency.

Fig. 8 illustrates the SNP's advertisement revenue under different parameters of μ_λ and v. The SNP can earn more advertisement revenue if users have higher valuations on social interaction. This is because *users are more willing to exhibit themselves, which enhances the advertisement targeting accuracy.* With a fixed μ_λ, a higher value of the advertiser's valuation v enables the SNP to earn more revenue.

Fig. 9 illustrates the user's utility depends on different values of μ_λ and v. We can see that the user's utility increases in the average social valuation μ_λ, since the user can derive more utility from social interaction. Furthermore, given a fixed μ_λ, the SNP's advertisement benefit v has very little impact on the user's utility. Although the SNP would like to extract more users' information due to the higher benefit of advertisement, the users can lower their information exhibition levels to offset the privacy loss. *This indicates that users achieve the best trade-off between social benefit and privacy loss.*

2) Optimal Privacy Policy Increases SNP's Advertisement Revenue: Fig. 10a compares the SNP's advertisement revenue under three privacy policies with different values of μ_λ. When $\delta = 0$, the SNP's advertisement revenue is independent of users' mean valuation μ_λ. However, as $\delta > 0$ (either δ^* or 1), the SNP can earn more advertisement revenue as μ_λ increases. With a fixed μ_λ, by comparing the revenue under no exploitation ($\delta = 0$) and full exploitation ($\delta = 1$), we can see that the full information exploitation can significantly

[13]In the future, we will gather more data from realistic social network platforms (such as Facebook and WeChat) to form an application-based benchmark.



(a) SNP's advertisement revenue. (b) Average of user utility.

Fig. 10. Performance under different privacy policies.

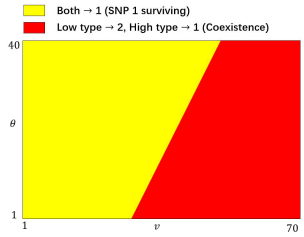


Fig. 11. Market partition.

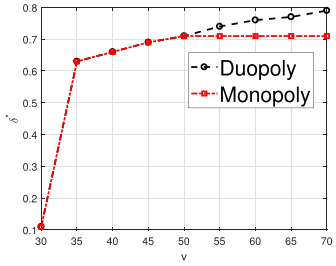


Fig. 12. Duopoly vs. monopoly.

improve the revenue by at most 105%. This validates the SNP's great interest in the targeted advertisement. *The SNP can earn even more revenue under the optimal privacy policy ($\delta = \delta^* < 1$). Although the exploitation fraction is lower than full exploitation, more conservative privacy protection can stimulate users' information exhibition.*

3) Optimal Privacy Policy Increases Users' Utilities: Fig. 10b compares the user's utility under three privacy policies with different values of v. The user's utility does not change greatly under different values of v. The perfect privacy policy ($\delta = 0$) yields the highest user utility. The worst privacy policy ($\delta = 1$) always gives rise to the lowest utility due to the significant privacy loss. *With the policy changed from the worst one to the optimal one, the users can experience significant utility gains, from both social interaction enhancement (due to threshold property of users' information levels in exploitation fraction) and privacy issue mitigation (due to increasing property of privacy loss in exploitation fraction).*

B. Duopoly Case

We study the SNPs duopoly competition equilibrium. More specifically, we investigate *how social network benefit and advertisement revenue affect the market partition*. We also *compare the duopoly case with the monopoly case to understand the impacts of market competition.*

We firstly introduce the simulation setup. Recall that we assume the SNP's social network benefit $b_s\left(\sum_{n\in\mathcal{N}} x_n\right) = \theta \sum_n x_n \log(\sum_n x_n + 1)$. Here, the social valuation parameter θ captures how beneficial the social network is to the SNPs. Meanwhile, we set the service quality parameters of the SNP 1 and the SNP 2 as $w_1 = 3$ and $w_2 = 2$, respectively. We set the social interaction valuation of high-type users and low-type users as $\lambda_H = 10$ and $\lambda_L = 0.1$, respectively. Furthermore, in the numerical study, we no longer require Assumption 5. Instead, we consider privacy loss $g(\delta x) = 90(\delta x)^2$ and subsidy probability $q(\delta x) = \delta x$.

1) Competition Equilibria: Fig. 11 shows two cases of equilibria (with two colors) under different values of advertisement benefit v and parameter θ.

1) **Yellow equilibrium of $(1, 1)$: SNP 1 dominates the market.** *This equilibrium exists when the advertisement benefit v is relatively low.* If the targeted advertisement does not generate much profit, SNP 1 would mainly rely on social network benefit. Hence it will decide to adopt a conservative privacy policy to attract both low-type and high-type users. This drives SNP 2 out of the market.
2) **Red equilibrium of $(1, 2)$: Both SNPs co-exist.** *This equilibrium exists when advertisement benefit v is relatively high.* SNP 1 has an incentive to adopt an aggressive privacy policy, which only attracts high-type users. SNP 2 can leverage this opportunity to attract low-type users with a more conservative privacy policy.

In addition, we find that as parameter θ increases, the boundary v that divides the above two cases increases. This implies that SNP 1 is more inclined of using a conservative policy to attract low-type users, as the benefit of more social interactions will help compensate for the loss of advertisement revenue. And greater benefit from social network is a good motivation of attracting low-type users. SNP 1 can leverage its advantage of social service to achieve that as parameter θ increases.

2) Duopoly vs. Monopoly: Fig. 12 shows that *SNP 1's privacy policy in the duopoly case is no conservative than that in the monopoly case.*

More specifically, when the advertisement benefit v is relatively low ($v \leq 50$), the optimal privacy policies and monopoly are the same. Under such a parameter setting, the duopoly equilibrium corresponds to the case where SNP 1 dominates the market, which is equivalent to monopoly.

However, when the advertisement benefit v is relatively high ($v > 50$), SNP 1's optimal privacy policy in duopoly is more aggressive than that in monopoly. In the monopoly equilibrium, only high-type users choose SNP 1 and exhibit high information levels. Thus, the SNP 1 can leverage this fact and adopt a more aggressive privacy policy to achieve a good advertisement revenue.

IX. Conclusion

We present the first theoretical study on the SNP's privacy policy (policies) with targeted advertisement, in both

monopoly and duopoly markets. In the monopoly case, we show that the SNP's optimal privacy policy not only yields good advertisement revenue but also encourages users' social interaction with promising privacy protection. In the duopoly case, we show that it is possible for the SNP with strong advantage of social service to choose a conservative privacy policy to drive the other SNP out of the market. However, if the advertisement revenue is significant, the SNP with the social service advantage would prefer to choosing an aggressive policy, and both SNPs co-exist in the market.

In the future, we plan to extend our study to a multi-SNP market, through three-stage Stackelberg game characterizing the interactions among multiple SNPs, the advertiser, and the users. We anticipate that several key intuitions from the current study will carry over: the SNPs with very weak qualities of social service will leave the market, and those with strong qualities will survive. Meanwhile, those with medium qualities of social service can also survive by adopting conservative enough privacy policies. We will be interested in further exploring the new insights in the multi-SNP market.

There are several other ways of further extending the work. For example, an SNP can adopt different privacy policies to different types of users. The fairness can be also an important new consideration.

REFERENCES

[1] G. Liao, X. Chen, and J. Huang, "Privacy policy in online social network with targeted advertising business," in *Proc. IEEE INFOCOM*, Jul. 2020, pp. 934–943.

[2] *Whatsapp*. Accessed: Dec. 2021. [Online]. Available: https://www.whatsapp.com/

[3] *Wechat*. Accessed: Dec. 2021. [Online]. Available: https://www.wechat.com/en/

[4] *Facebook*. Accessed: Dec. 2021. [Online]. Available: https://www.facebook.com/

[5] *Instagram*. Accessed: Dec. 2021. [Online]. Available: https://www.instagram.com/

[6] *Twitter*. Accessed: Dec. 2021. [Online]. Available: https://twitter.com/

[7] *Quora*. Accessed: Dec. 2021. [Online]. Available: https://www.quora.com/

[8] (2021). *Facebook's Annual Report*. [Online]. Available: http://d18rn0p25nwr6d.cloudfront.net/CIK-0001326801/4dd7fa7f-1a51-4ed9-b9df-7f42cc3321eb.pdf

[9] J. Yan, N. Liu, G. Wang, W. Zhang, Y. Jiang, and Z. Chen, "How much can behavioral targeting help online advertising?" in *Proc. Int. Conf. World Wide Web*, 2009, pp. 261–270.

[10] (2018). *An Understanding of Data-Driven Targeted Advertising Basics and Effectiveness*. [Online]. Available: https://www.nusparkmarketing.com/2018/08/an-understanding-of-data-driven-targeted-advertising-basics-effectiveness/

[11] J. G. Cabañas, A. Cuevas, and R. Cuevas, "Unveiling and quantifying Facebook exploitation of sensitive personal data for advertising purposes," in *Proc. USENIX Secur. Symp.*, 2018, pp. 479–495.

[12] J. M. Carrascosa, J. Mikians, R. Cuevas, V. Erramilli, and N. Laoutaris, "I always feel like somebody's watching me: Measuring online behavioural advertising," in *Proc. ACM Conf. Emerg. Netw. Exp. Technol.*, 2015, pp. 1–13.

[13] A. Andreou, G. Venkatadri, O. Goga, K. P. Gummadi, P. Loiseau, and A. Mislove, "Investigating ad transparency mechanisms in social media: A case study of Facebook's explanations," in *Proc. Netw. Distrib. Syst. Secur. Symp.*, 2018, pp. 1–15.

[14] G. Venkatadri *et al.*, "Privacy risks with Facebook's PII-based targeting: Auditing a data Broker's advertising interface," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2018, pp. 89–107.

[15] (2018). *Facebook-Cambridge Analytica Data Scandal*. [Online]. Available: https://en.wikipedia.org/wiki/Facebook-Cambridge_Analytica_data_scandal

[16] (2019). *Facebook Reportedly Gets Deeply Personal Info*. [Online]. Available: https://www.cnbc.com/2019/02/22/facebook-receives-personal-health-data-from-apps-wsj.html

[17] (2018). *Uber Settles Data Breach Investigation for $148 Million*. [Online]. Available: https://www.nytimes.com/2018/09/26/technology/uber-data-breach.html

[18] (2019). *News*. [Online]. Available: https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/

[19] X. Wang, A. Continella, Y. Yang, Y. He, and S. Zhu, "Leakdoctor: Toward automatically diagnosing privacy leaks in mobile applications," in *Proc. ACM Ubicomp*, 2019, pp. 1–25.

[20] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Proc. Theory Cryptography Conf.*, 2006, pp. 265–284.

[21] A. Goldfarb and C. E. Tucker, "Privacy regulation and online advertising," *Manage. Sci.*, vol. 57, no. 1, pp. 57–71, Jan. 2011.

[22] J. P. Johnson, "Targeted advertising and advertising avoidance," *RAND J. Econ.*, vol. 44, no. 1, pp. 128–144, 2013.

[23] R. Cummings, K. Ligett, M. M. Pai, and A. Roth, "The strange case of privacy in equilibrium models," in *Proc. ACM Conf. Econ. Comput.*, Jul. 2016, pp. 1–21.

[24] R. Gross, A. Acquisti, and H. J. Heinz, "Information revelation and privacy in online social networks," in *Proc. ACM workshop Privacy Electron. Soc. (WPES)*, 2005, pp. 71–80.

[25] R. Gradwohl, "Information sharing and privacy in networks," in *Proc. ACM Conf. Econ. Comput.*, Jun. 2017, pp. 349–350.

[26] G. Liao, X. Chen, and J. Huang, "Social-aware privacy-preserving correlated data collection," in *Proc. ACM MoBiHoc*, 2018, pp. 11–20.

[27] (2020). *Facebook's Financial Report*. [Online]. Available: https://investor.fb.com/investor-news/press-release-details/2020/Facebook-Reports-Fourth-Quarter-and-Full-Year-2019-Results/default.aspx

[28] J. Nie, J. Luo, Z. Xiong, D. Niyato, and P. Wang, "A Stackelberg game approach toward socially-aware incentive mechanisms for mobile crowdsensing," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 724–738, Jan. 2018.

[29] Y. Zhao and C. H. Liu, "Social-aware incentive mechanism for vehicular crowdsensing by deep reinforcement learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 4, pp. 2314–2325, Apr. 2021.

[30] G. Yang, S. He, Z. Shi, and J. Chen, "Promoting cooperation by the social incentive mechanism in mobile crowdsensing," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 86–92, Mar. 2017.

[31] K. Han, H. Huang, and J. Luo, "Quality-aware pricing for mobile crowdsensing," *IEEE/ACM Trans. Netw.*, vol. 26, no. 4, pp. 1728–1741, Aug. 2018.

[32] 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 Spectr.*, vol. 43, no. 7, pp. 34–39, Jul. 2006.

[33] (2016). *Privacy and Information Sharing*. [Online]. Available: https://www.pewinternet.org/2016/01/14/privacy-and-information-sharing/

[34] W.-S. Yang, J.-B. Dia, H.-C. Cheng, and H.-T. Lin, "Mining social networks for targeted advertising," in *IEEE Annu. Hawaii Int. Conf. Syst. Sci.*, Jan. 2006, p. 137.

[35] S. Nath, "MAdScope: Characterizing mobile in-app targeted ads," in *Proc. 13th Annu. Int. Conf. Mobile Syst., Appl., Services*, May 2015, pp. 59–73.

[36] (2018). *How Facebook Ads Target You*. [Online]. Available: https://www.cnbc.com/2018/04/14/how-facebook-ads-target-you.html

[37] A. Ghosh, M. Mahdian, R. P. McAfee, and S. Vassilvitskii, "To match or not to match: Economics of cookie matching in online advertising," *ACM Trans. Econ. Comput.*, vol. 3, no. 2, pp. 1–18, 2015.

[38] (2021). *Tencent's Annual Report*. [Online]. Available: https://static.www.tencent.com/uploads/2021/04/08/960eae1f18dd716fd3a7d704e123d7a5.pdf

[39] (2019). *The US Government is Fining Facebook 5 Billion for Privacy Violations*. [Online]. Available: https://www.engadget.com/2019/07/12/facebook-ftc/

[40] (2021). *Cost of Data Breach Report*. IBM Security. [Online]. Available: https://www.ibm.com/security/digital-assets/cost-data-breach-report/#/

[41] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA, USA: MIT Press, 1994.

[42] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic Game Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2007.

[43] D. M. Topkis, *Supermodularity and Complementarity*. Princeton, NJ, USA: Princeton Univ. Press, 1998.

[44] P. Milgrom and J. Roberts, "Rationalizability, learning, and equilibrium in games with strategic complementarities," *Econometrica*, vol. 58, no. 6, pp. 1255–1277, Nov. 1990.
[45] W. Wang, L. Ying, and J. Zhang, "The value of privacy: Strategic data subjects, incentive mechanisms, and fundamental limits," *ACM Trans. Econ. Comput.*, vol. 6, no. 2, pp. 1–26, 2018.
[46] A. B. Akbay, W. Wang, and J. Zhang, "Data collection from privacy-aware users in the presence of social learning," in *Proc. IEEE Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 2019, pp. 679–686.
[47] V. Krishna, *Auction Theory*. New York, NY, USA: Academic, 2009.
[48] (2018). *How do First-Price and Second-Price Auctions Work in Online Advertising?*. [Online]. Available: https://clearcode.cc/blog/first-price-second-price-auction/
[49] (2018). *Rolling Out First Price Auctions to Google Ad Manager Partners*. [Online]. Available: https://blog.google/products/admanager/rolling-out-first-price-auctions-google-ad-manager-partners/
[50] J. Leskovec and A. Krevl. (Jun. 2014). *SNAP Datasets: Stanford Large Network Dataset Collection*. [Online]. Available: http://snap.stanford.edu/data

Guocheng Liao (Member, IEEE) received the B.E. degree from Sun Yat-sen University in 2016 and the Ph.D. degree from the Department of Information Engineering, The Chinese University of Hong Kong, in 2021. He is currently an Assistant Professor with the School of Software Engineering, Sun Yat-sen University. His current research lies in data privacy, game theory, and mechanism design. He received the 2017 IEEE GLOBECOM ComSoc Young Professional Best Paper Award.

Xu Chen (Senior Member, IEEE) received the Ph.D. degree in information engineering from The Chinese University of Hong Kong in 2012. He was a Post-Doctoral Research Associate with Arizona State University, Tempe, USA, from 2012 to 2014, and a Humboldt Scholar Fellow with the Institute of Computer Science, University of Göttingen, Germany, from 2014 to 2016. He is a Full Professor with Sun Yat-sen University, Guangzhou, China, and the Vice Director of the National and Local Joint Engineering Laboratory of Digital Home Interactive Applications. He was a recipient of the Prestigious Humboldt Research Fellowship awarded by the Alexander von Humboldt Foundation of Germany, the 2014 Hong Kong Young Scientist Runner-Up Award, the 2016 Thousand Talents Plan Award for Young Professionals of China, the 2017 IEEE Communication Society Asia–Pacific Outstanding Young Researcher Award, the 2017 IEEE ComSoc Young Professional Best Paper Award, the Honorable Mention Award of 2010 IEEE International Conference on Intelligence and Security Informatics, the Best Paper Runner-Up Award of 2014 IEEE International Conference on Computer Communications (INFOCOM), and the Best Paper Award of 2017 IEEE International Conference on Communications. He is currently an Area Editor of IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY and an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE INTERNET OF THINGS JOURNAL, and IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC) series on network softwarization and enablers.

Jianwei Huang (Fellow, IEEE) received the Ph.D. degree in ECE from Northwestern University in 2005. He worked as a Post-Doctoral Research Associate with Princeton University during 2005 to 2007. From 2007 to 2018, he was on the Faculty of the Department of Information Engineering, The Chinese University of Hong Kong. Since 2019, he has been with the Faculty of The Chinese University of Hong Kong, Shenzhen, where he is currently the Presidential Chair Professor and an Associate Dean of the School of Science and Engineering. He also serves as the Vice President for the Shenzhen Institute of Artificial Intelligence and Robotics for Society. He has published more than 300 papers in leading venues, with a Google Scholar citation of more than 14000 and an H-index of 59. He has coauthored seven books, including the textbook *Wireless Network Pricing*. His research interests are in the area of network optimization, network economics, and network science, with applications in communication networks, energy networks, data markets, crowd intelligence, and related fields. He was an IEEE ComSoc Distinguished Lecturer and a Clarivate Web of Science Highly Cited Researcher. He has coauthored articles which received over ten best paper awards, including the 2011 IEEE Marconi Prize Paper Award in Wireless Communications. He was the Associate Editor-in-Chief of IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY and is the Editor-in-Chief of IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING.