Pricing Social Visibility Service in Online Social Networks: Modeling and Algorithms

Shiyuan Zheng, Hong Xie, *Member, IEEE*, and John C.S. Lui, *Fellow, IEEE*

***Abstract*—Online social networks (OSNs) such as YouTube, Instagram, Twitter, Facebook, etc., serve as important platforms for users to share their information or content to friends or followers. Oftentimes, users want to enhance their social visibility, as it can make their contents, i.e., opinions, videos, pictures, etc., attract attention from more users, which in turn may bring higher commercial benefits to them. Motivated by this, we propose a mechanism, where the OSN operator provides a *"social visibility boosting service"* to incentivize "transactions" between requesters (users who seek to enhance their social visibility via adding new "neighbors") and suppliers (users who are willing to be added as a new "neighbor" of any requester when certain "rewards" is provided). We design a posted pricing scheme for the OSN provider to charge the requesters who use such boosting service and reward the suppliers who make contributions. The OSN operator keeps a fraction of the payment from requesters and distributes the remaining part to participating suppliers *"fairly"* via a rewarding rule based on Shapley value. We consider two different objectives/problems of the OSN provider, i.e., to select the optimal prices and supplier set to maximize (1) the revenue or (2) the welfare increase, under the requesters' budget constraint on suppliers. We first show that the problems are not simpler than NP-hard. We then decomposed each problem into two sub-routines, where one focuses on selecting the optimal set of suppliers, and the other one focuses on selecting the optimal prices. We prove the hardness of each sub-routine, and eventually design a computationally efficient approximation algorithm to solve the problems with provable theoretical guarantee on the revenue/welfare increase gap. We conduct extensive experiments on four public datasets to validate the performance of our proposed algorithms.**

***Index Terms*—Approximation algorithms, revenue maximization, social visibility, welfare maximization.**

I. Introduction

Online social networks (OSNs) such as YouTube, Instagram, Twitter, Facebook, etc., serve as important platforms for users to share their information or content with friends or followers, e.g., users on Facebook can share their opinions or status with their friends via the friendship network. Users on YouTube can share their videos with their subscribers via the subscriber network. A user's friends or followers can further share the information or content of this user with their friends or followers. Hence, the information or content of a user can be propagated to her direct friends or followers, or even multi-hop friends or followers. In other words, a user is "socially visible" to her direct friends or followers and some multi-hop friends or followers.

Oftentimes, users want to enhance their social visibility, as it can make their contents, i.e., opinions, videos, pictures, etc., attract attention from more users, which in turn may bring higher commercial benefits to them. We call a user who wants to enhance social visibility as a "*requester*". One way to enhance a requester's social visibility is to attract some new friends or followers. It is well known that the OSN is under the "the rich gets richer" phenomenon, which makes it difficult for requesters, especially those with low social visibility, to attract new friends or followers. Moreover, a user is reluctant to connect to or follow requesters that she is not interested in. But when the finical incentive is provided, some users would be willing to be added as requesters' new friends or followers. We call such users "*suppliers*". For example, Dou+,[1] an official service provided by TikTok, implements this finical incentive, where content creators (requesters) who want to increase visibility can pay to gain more followers. Designing appropriate finical incentives can benefit requesters, suppliers and the whole OSN ecosystem, which in turn benefits the OSN operator. However, no work has studied this finical incentive design problem before. The objective of this work is to fill in this blank.

We propose a mechanism, where the OSN operator provides a *"social visibility boosting service"* to incentivize the transaction between requesters and suppliers. In this mechanism, the visibility boosting service is provided via a posted normalized pricing scheme (p, q), where $p\in[0,1]$ and $q\in[0,1]$. Here, p is the price of unit social visibility improvement that the operator charges a participating requester, and q is the price of unit contribution to the total visibility improvement that the operator rewards a participating supplier.

We consider the case that each participating supplier adds links to all participating requesters and requesters have a budget to add at most $b \in \mathbb{N}_+$ new friends or followers. In this

[1] https://doujia.douyin.com/

Manuscript received 5 January 2022; revised 20 September 2022; accepted 11 November 2022. Date of publication 23 November 2022; date of current version 23 February 2023. The work of John C.S. Lui was supported by Research Grants Council (RGC) under Grant GRF-14200321. The work of Hong Xie was supported by Chongqing Natural Science Foundation under Grant cstc2020jcyj-msxmX0652. Part of this work was done when Hong Xie was working at Chongq-ing University. Recommended for acceptance by Dr. Yan Chen. *(Corresponding author: Hong Xie.)*

Shiyuan Zheng and John C.S. Lui are with the Department of Computer Science, Engineering, The Chinese University of Hong Kong, Central Ave, Hong Kong (e-mail: syzheng@cse.cuhk.edu.hk; cslui@cse.cuhk.edu.hk).

Hong Xie is with Chongqing Institute of Green, Intelligent Technology, Chinese Academy of Sciences, Beijing 100045, China (e-mail: xiehong2018@foxmail.com).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSE.2022.3223958, provided by the authors.

Digital Object Identifier 10.1109/TNSE.2022.3223958

work, we assume that users are all rational and requesters and suppliers decide whether to participate or not by comparing their valuations to the posted prices (p, q). In real-world OSN, valuations can be solicited by incentive-compatible mechanisms or estimated from users' history responses using online learning algorithms. In this work, we assume valuations of users are accessible to the OSN operator. To design the social visibility boosting service, the challenges the OSN operator needs to address are: (1) select the prices (p, q) and supplier set so as to optimize certain objectives; and (2) divide the reward "fairly" among all suppliers. In this paper, we consider the following two typical objectives for the operator:

- the revenue which is defined as the difference between the total payment from the requesters and the total reward to the suppliers.
- the welfare increase which is defined as the sum of the valuation increase of participating requesters. We also add the constraint of non-negative revenue (a.k.a weakly budget balanced) to this setting to make sure the operator will not lose money in the service.

Note that, each objective is technically challenging to achieve. One can observe that selecting the price (p, q) and supplier set involves a mixed optimization problem, i.e., with both continuous decision variables, i.e., p and q, and set decision variables, i.e., supplier set. This implies that gradient-based optimization methods do not work for this problem. To illustrate the hardness of this optimization problem, consider the case where p and q are given and our objective is to select the supplier set. The number of candidates may be much larger than budget b and the OSN operator needs to select at most b suppliers from them. Due to the network externality effect of social visibility, suppliers are not independent in enhancing the social visibility of a requester. In other words, the total visibility increase contributed by two suppliers does not necessarily equal to the sum of the increase made by each individual supplier. As one will see in Section III, this dependency among suppliers makes this simplified problem NP-hard already. Also, this dependency among suppliers also makes it challenging to fairly divide the reward to suppliers. We address these challenges and our contributions are:

- We formulate a mathematical model to quantify social visibility. To the best of our knowledge, we are the first to propose a posted pricing scheme and formulate revenue/welfare maximization problem for visibility boosting service.
- We prove that the revenue/welfare maximization problem is not simpler than an NP-hard problem. We decomposed it into two sub-routines, where one focuses on selecting the optimal set of suppliers and the other focuses on selecting the optimal price, and propose approximation algorithms for each sub-problem with theoretical guarantees. Finally, we prove that by combining these approximation algorithms, we obtain an algorithm to solve the revenue/welfare maximization problem with provable theoretical guarantee on the objective gap.
- We show how to divide the reward to suppliers fairly via a distribution rule based on the concept of Shapley value.
- We conduct experiments on real-world social network datasets, and the results validate the effectiveness and efficiency of our algorithms.

The remainder of this is organized as follows. Section II presents the social visibility model and the formulation of the revenue/welfare maximization problem. Section III presents the hardness analysis of the revenue maximization problem and approximation algorithms for the problem. Section IV presents the hardness analysis of the welfare maximization problem and approximation algorithms for the problem. Section V presents algorithms for the fair division of contribution among suppliers. Section VI presents the performance evaluation over real-world dataset. Section VII presents the related work and Section VIII concludes.

II. Model & Problem Formulation

In this section, we first present the social visibility model. Then we present the social visibility pricing problem to maximize the revenue/welfare.

A. The Social Visibility Model

Online Social Network. Consider an OSN which is characterized by an unweighted and directed graph $\mathcal{G} \triangleq (\mathcal{U}, \mathcal{E})$, where $\mathcal{U} \triangleq \{1, \ldots, U\}$ denotes a set of $U \in \mathbb{N}_+$ users and $\mathcal{E} \subseteq \mathcal{U} \times \mathcal{U}$ denotes a set of edges between users. Note that a directed edge from user $v \in \mathcal{U}$ to user $u \in \mathcal{U}$ is denoted by $(v, u) \in \mathcal{E}$, which can be interpreted as v following u in a Twitter-like OSN. The above graph can be directly used to model directed networks such as Instagram and Twitter. Note that it is also able to model undirected networks such as Facebook and DBLP, since each undirected edge between user u and v can be represented by two directed edges (u, v) and (v, u). We focus on the case that there is no self-loop edge, i.e., $(u, u) \notin \mathcal{E}, \forall u \in \mathcal{U}$.

Social Visibility. We denote a directed path in a graph $\mathcal{G}$ by

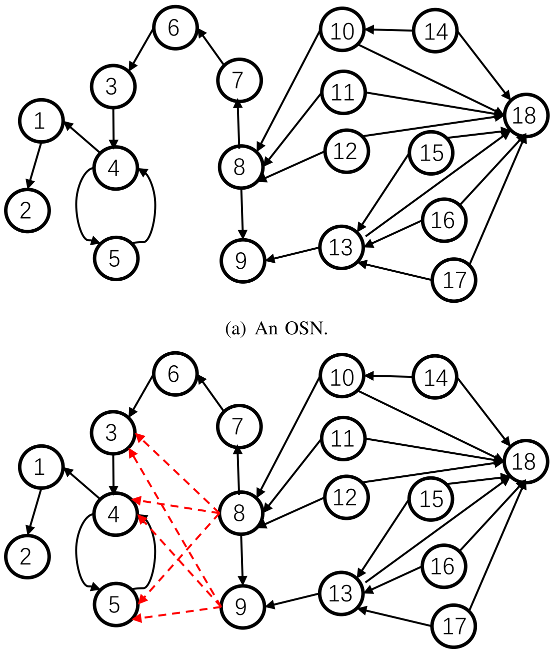
$$\vec{p} \triangleq (u_0 \rightarrow u_1 \rightarrow \ldots \rightarrow u_n),$$

where $(u_i, u_{i+1}) \in \mathcal{E}, \forall i \in \{0, \ldots, n-1\}$, and $u_i \neq u_j, \forall i, j$. Note that $u_i \neq u_j, \forall i, j$ captures that there is no self-loop edges or circles in the path. Denote a set of all directed edges on path $\vec{p}$ as

$$\mathcal{F}(\vec{p}) \triangleq \{(u_0, u_1), (u_1, u_2), \ldots, (u_{n-1}, u_n)\}.$$

Let $L(\vec{p})$ denote the length (i.e., number of hops) of path $\vec{p}$, which can be expressed as $L(\vec{p}) = |\mathcal{F}(\vec{p})|$. Let $\mathcal{P}(v, u)$ denote the set of all directed paths (without circles) from user v to user u in $\mathcal{G}$. Let $D(v, u)$ denote the distance from user v to user u in graph $\mathcal{G}$. We define $D(v, u)$ as the length of the shortest path from v to u, i.e.,

$$D(v, u) \triangleq \begin{cases} \min_{\vec{p} \in \mathcal{P}(v,u)} L(\vec{p}), & \text{if } \mathcal{P}(\text{v,u}) \neq \emptyset \\ +\infty, & \text{if } \mathcal{P}(\text{v,u}) = \emptyset.. \end{cases}$$



(a) An OSN.

(b) Updated OSN.

Fig. 1. A toy example.

Namely, when there is no directed path from v to u, the distance from v to u is infinite. Based on the distance between nodes, we define user u being d-visible to user v if

$$D(v,u) \le d.$$

The d-visible set of user u is defined as the set of all users to whom user u is d-visible, formally

$$\mathcal{V}(u,d) \triangleq \left\{ v \middle| v \in \mathcal{U}, D(v,u) \le d \right\}.$$

Let $\tau \in \mathbb{N}_+$ denote the social visibility threshold of an OSN. This notion of social visibility threshold captures that the information of a user can be propagated to its followers and its followers may further propagate their own followers and so on. Namely, user u is not visible to users whose distance to user u is larger than τ. For example, $\tau = 1$ models that each user is only visible to its own followers, while $\tau = 2$ models that each user is visible to its own one-hop and two-hop followers. Based on τ, we define the notation of visibility.

Definition 1: The visibility of user u is the cardinality of her τ-visible set.

For example, Fig. 1(a) shows that the 2-visible set of user 4 is $\{3,5,6\}$ and the 3-visible set of user 4 is $\{3,5,6,7\}$. If the social visibility threshold is $\tau = 2$, then the visibility of user 4 is 3 (cardinality of $\{3,5,6\}$).

B. Pricing the Social Visibility

The pricing scheme. A *"requester"* is a user in the set $\mathcal{U}$ who seeks to increase her visibility by requesting other users to be her new incoming neighbors. Let $\mathcal{R} \subseteq \mathcal{U}$ denote a set of all requesters. A *"supplier"* is a user in the set $\mathcal{U}$ who is willing to be a new incoming neighbor of any requester when a certain finical incentive is provided. Let $\mathcal{S} \subseteq \mathcal{U}$ denote a set of all suppliers. We assume the market is two-sided, i.e., $\mathcal{R} \cap \mathcal{S} = \emptyset$, which captures that a user can not be both requester and supplier. We consider the general case that there are some users who are neither requesters nor suppliers, i.e., $\mathcal{R} \cup \mathcal{S} \subseteq \mathcal{U}$. The OSN operator provides a *"social visibility boosting service"*to incentivize the "transaction" between requesters and suppliers. Specifically, the operator recruits participants with a normalized pricing scheme, specifying that a participating requester would be charged p for unit visibility increase and a participating supplier would be rewarded q for *unit contribution*to boost the visibility of participating requesters. We defer the details about quantifying contribution and distributing reward to Section V.

Requesters' decision model. Each requester $u \in \mathcal{R}$ has a valuation $p_u \in [0,1]$ of per unit visibility increase. In other words, p_u is the highest price that requester u is willing to pay for one unit visibility increase. Thus, a requester will use the social visibility boosting service if her per unit valuation is not lower than price p, i.e., $p_u \ge p$. Let $\widetilde{\mathcal{R}}(p)$ denote the set of all participating requesters under price p, formally

$$\widetilde{\mathcal{R}}(p) \triangleq \{u | p_u \ge p, u \in \mathcal{R}\}. \tag{1}$$

We assume each requester is restricted to add at most $b \in \mathbb{N}_+$ incoming neighbors. Note that, we consider the case that each participating supplier will add links to all participating requesters. Let $\mathcal{M}$ be the set of participating suppliers. The edges on the updated graph are $\mathcal{E} \cup (\mathcal{M} \times \widetilde{\mathcal{R}}(p))$. For example, in Fig. 1(a), the participating requesters are $\{3,4,5\}$ and the participating suppliers are $\{8,9\}$, then the updated network is Fig. 1(b) with dashed lines to denote the newly added links.

Based on the change of the network, we now analyze the visibility increase and welfare increase of requesters. For each requester $u \in \widetilde{\mathcal{R}}(p)$, we use $I_u(p,q,\mathcal{M})$ to denote the individual visibility increase of u under pricing scheme (p,q) with participating suppliers $\mathcal{M}$. The visibility increase of u can be derived as the following closed form:

$$I_u(p,q,\mathcal{M}) = \left| \bigcup_{l \in \mathcal{M} \times \widetilde{\mathcal{R}}(p)} \mathcal{V}(l^s, \tau - 1 - D(l^e, u)) \setminus \mathcal{V}(u,\tau) \right|,$$

where l^s and l^e are the start node and end node of link l respectively.

Next, we take the requester's valuation of unit visibility increase into consideration, and define the individual welfare increase for each requester. Given the set $\mathcal{M}$ of participating suppliers, the individual welfare increase of requester $u \in \widetilde{\mathcal{R}}(p)$, denoted by $W_u(p,q,\mathcal{M})$, is defined as the product of visibility increase $I_u(p,q,\mathcal{M})$ and valuation of unit visibility increase p_u, formally,

$$W_u(p,q,\mathcal{M}) \triangleq p_u I_u(p,q,\mathcal{M}).$$

The welfare $W_u(p, q, \mathcal{M})$ characterizes the happiness of requester u. For each requester $u \in \widetilde{\mathcal{R}}(p)$, the utility is defined as u's welfare increase minus u's payment:

$$U_u(p, q, \mathcal{M}) \triangleq (p_u - p) I_u(p, q, \mathcal{M}).$$

It is obvious that, no matter what $\mathcal{M}$ is, requesters in $\widetilde{\mathcal{R}}(p)$ would get non-negative utility by choosing to participate; and requesters in $\mathcal{R} \setminus \widetilde{\mathcal{R}}(p)$ would get negative utility if they participate. Thus the criterion of selecting participating requesters in (1) satisfies the assumption that all the users are rational.

Supplier's decision model. Each supplier $u \in \mathcal{S}$ has a unit valuation of $q_u \in [0, 1]$ on the unit contribution she made. In other words, q_u is the lowest price that supplier u is willing to "trade out" her unit contribution. Note that, all participating suppliers make a total contribution, i.e., the total visibility increase of participating requesters. Quantifying contribution among participating suppliers is a non-trivial problem due to the following reasons: (1) some participating suppliers may have a larger number of followers while others may have a smaller one; and (2) the network structure poses an externality effect, causing the contribution of participating suppliers to be correlated. To incentivize suppliers to participate, one needs to divide total visibility increase fairly among suppliers. We use ϕ to denote a *"fair"* division mechanism, which prescribes the contribution denoted by $\phi_u(p, \mathcal{M})$ for each participating supplier $u \in \mathcal{M}$. In order to avoid distracting readers, we defer the detailed explanation of ϕ to the next section. Given the fair division mechanism ϕ, a supplier is willing to participate in the visibility boosting service if her valuation of unit contribution is lower than price q. Let $\widetilde{\mathcal{S}}(p)$ be the set of potential participating suppliers under price q, formally

$$\widetilde{\mathcal{S}}(q) \triangleq \{u | q_u \leq q, u \in \mathcal{S}\}.$$

Optimal pricing. We have mentioned the two objectives of the OSN operator in Section I. With the definitions of the individual visibility increase and the individual welfare increase of a participating requester, we are ready to give the formal definitions of the two objectives. Given a pricing scheme (p, q) and participating suppliers $\mathcal{M}$, the visibility increase of all participating requesters, denoted by $I(p, q, \mathcal{M})$, is defined as the sum of the individual visibility increase of all the participating requesters, formally,

$$I(p, q, \mathcal{M}) \triangleq \sum_{u \in \widetilde{\mathcal{R}}(p)} I_u(p, q, \mathcal{M}).$$

Given the posted pricing scheme (p, q), since the total contribution is exactly the visibility increase $I(p, q, \mathcal{M})$, the revenue of the OSN operator is

$$R(p, q, \mathcal{M}) \triangleq pI(p, q, \mathcal{M}) - qI(p, q, \mathcal{M}). \quad (2)$$

The welfare increase is defined as the sum of the individual welfare increase of all the participating requesters, formally,

$$W(p, q, \mathcal{M}) \triangleq \sum_{u \in \widetilde{\mathcal{R}}(p)} W_u(p, q, \mathcal{M}).$$

Moreover, we have the constraint of non-negative revenue to make sure the operator will not lose money and such a service can be sustainable. According to (2), this constraint is satisfied if and only if $p \geq q$ hold. Finally, we use $R(p, q, \mathcal{M})$ and $W(p, q, \mathcal{M})$ as the objective functions to formulate the following pricing problems for the OSN operator:

Problem 1 (Optimal pricing for revenue): Given a budget b, select pricing scheme (p, q) and participating suppliers $\mathcal{M}$ to maximize the revenue $R(p, q, \mathcal{M})$:

$$\begin{aligned} \max_{p,q,\mathcal{M}} \quad & R(p, q, \mathcal{M}) \\ s.t. \quad & \mathcal{M} \subseteq \widetilde{\mathcal{S}}(q), |\mathcal{M}| \leq b, \\ & p \in [0, 1], \\ & q \in [0, 1]. \end{aligned}$$

Problem 2 (Optimal pricing for welfare): Given a budget b, select pricing scheme (p, q) and participating suppliers $\mathcal{M}$ to maximize the welfare increase $W(p, q, \mathcal{M})$.

$$\begin{aligned} \max_{p,q,\mathcal{M}} \quad & W(p, q, \mathcal{M}) \\ s.t. \quad & \mathcal{M} \subseteq \widetilde{\mathcal{S}}(q), |\mathcal{M}| \leq b, \\ & p \in [0, 1], \\ & q \in [0, 1], \\ & p \geq q. \end{aligned}$$

One can observe that Problems 1 and 2 are mixed optimization problems, i.e., with both continuous and set decision variables.

Remark. In this work, we focus on the setting where valuations of users are accessible to the OSN operator, for the purpose of studying the algorithmic aspect of the visibility pricing problem. We leave the game-theoretic aspect of this problem with unknown valuations as future work. It needs substantial amounts of work to design incentive-compatible mechanisms to make users report their valuations truthfully.

III. Algorithms for Optimal Pricing for Revenue

In this section, we present the analysis and the solution to Problem 1. We first show that Problem 1 is not simpler than an NP-hard problem. Then we decompose Problem 1 into two sub-problems. We prove the hardness of each sub-problem and propose approximation algorithms for each sub-problem. Finally, we prove that by combining these approximation algorithms, we can obtain an algorithm to solve Problem 1 with a provable theoretical guarantee on the revenue gap.

A. Hardness Analysis

Hardness analysis. Recall that Problem 1 is a mixed optimization problem with both continuous decision variables, i.e., p and q, and set decision variable, i.e., $\mathcal{M}$. This implies that gradient-based optimization methods does not work. To illustrate the hardness of Problem 1, we consider a sub-problem where pricing scheme (p,q) is given, which is stated as follows:

Problem 3 (Optimal supplier set $\mathcal{M}$ for revenue): Given pricing scheme (p,q) and budget b, select participating suppliers $\mathcal{M}$ from potential participating suppliers to maximize the revenue of the operator:

$$\begin{aligned} \max_{\mathcal{M}} \quad & R(p,q,\mathcal{M}) \\ s.t. \quad & \mathcal{M} \subseteq \widetilde{\mathcal{S}}(q), \\ & |\mathcal{M}| \leq b. \end{aligned}$$

Namely, Problem 3 only selects the optimal set of suppliers to maximize the revenue. Note that the set of potential participating suppliers $\widetilde{\mathcal{S}}(q)$ is determined as p and q are given. In the following theorem, we analyze its hardness.

Theorem 1: Problem 3 is NP-hard and Problem 1 is not easier than NP-hard.

Theorem 1 states that it is NP-hard to locate the optimal set of suppliers for Problem 3. In other words, it is computationally expensive to find an exact solution for Problem 3. Since Problem 3 is just a sub-problem of Problem 1, Problem 1 is not easier than Problem 3. Therefore, locating the exact optimal solution for Problem 1 is computationally expensive. We resort to approximation algorithms with theoretical guarantees to solve Problem 1.

Remark. The sub-problem is not a set cover problem and it can not be mapped into an equivalent Set Cover problem when there are multiple participating requesters. This is because a given supplier would have different effects on different participating requesters and the objective is to maximize the sum of cardinalities of all participating requesters' new visible set. Thus, one can not apply existing Set Cover algorithms to solve our sub-problem.

Our approach. To address Problem 1, we decompose it into two sub-problems, where each sub-problem serves as a subroutine. In particular, Problem 3 is the first sub-problem. Since Theorem 1 shows that Problem 3 is NP-hard, we aim to design an approximation algorithm for it, denoted by $\texttt{OptSupR}(p,q)$ (the detail is postponed to Section III-B). Algorithm $\texttt{OptSupR}$ (p,q) takes the pricing scheme (p,q) as an input, and returns an approximately optimal set of participating suppliers under (p,q). Then we use $\texttt{OptSupR}(p,q)$ as an oracle to search for the optimal pricing scheme (p,q). Formally, we aim to solve the second sub-problem defined as follows:

Problem 4 (Optimal pricing for revenue with oracle): Given the algorithm $\texttt{OptSupR}(p,q)$, select pricing scheme (p,q) so to maximize the revenue of the OSN operator:

$$\begin{aligned} \max_{p,q} \quad & R(p,q,\texttt{OptSupR}(p,q)) \\ s.t. \quad & p \in [0,1], \\ & q \in [0,1]. \end{aligned}$$

Note that $\texttt{OptSupR}(p,q)$ returns an approximately the optimal set of suppliers for each given (p,q), so Problem 4 returns an approximately optimal price. We also design an algorithm for Problem 4, denoted by OptPrice(OptSupR) (the detail is postponed to Section III-C). One needs to supply OptPri-cewith oracle OptSupR, and OptPrice(OptSupR) returns an approximately optimal (p,q). We next proceed to present the design and analysis of $\texttt{OptSupR}(p,q)$ and OptPrice(OptSupR).

Algorithm 1: OptSupR (p,q)

1: init $\mathcal{M} = \emptyset$
2: **for** $t = 1$ to b **do**
3: $u^* \leftarrow \arg\max_{u \in \widetilde{\mathcal{S}}(q)} R(p,q,\mathcal{M} \cup \{u\}) - R(p,q,\mathcal{M})$
4: $\mathcal{M} \leftarrow \mathcal{M} \cup \{u^*\}$
5: **end for**
6: **return** $\hat{\mathcal{M}}^* \leftarrow \mathcal{M}$

B. Design & Analysis of $\texttt{OptSupR}(p,q)$

Submodular analysis. First, note that once p and q are given, the set of participating requesters $\widetilde{\mathcal{R}}(p)$ and the set of potential participating suppliers $\widetilde{\mathcal{S}}(q)$ are fixed. Our objective is to select $\mathcal{M} \subset \widetilde{\mathcal{S}}(q)$ with constraint $|\mathcal{M}| \leq b$, so as to maximize the revenue. We first derive the closed form of the revenue. The revenue of the operator can be derived as

$$\begin{aligned} R(p,q,\mathcal{M}) &= (p-q) \sum_{u \in \widetilde{\mathcal{R}}(p)} I_u(p,q,\mathcal{M}) \\ &= (p-q) \sum_{u \in \widetilde{\mathcal{R}}(p)} \left| \cup_{l \in \mathcal{M} \times \widetilde{\mathcal{R}}(p)} \mathcal{V}(l^s, \tau - 1 - D(l^e, u)) \setminus \mathcal{V}(u,\tau) \right|. \end{aligned} \quad (3)$$

Based on the above closed-form expressions of the revenue, we have the following theorem which shows the sub-modularity and monotonicity of the revenue with respect to $\mathcal{M}$.

Theorem 2: Given the pricing scheme (p,q), the revenue $R(p,q,\mathcal{M})$ is monotonously increasing and submodular with respect to participating suppliers $\mathcal{M}$.

The $\texttt{OptSupR}(p,q)$ *algorithm.* With Theorem 2, we design Algorithm 1 to implement the oracle $\texttt{OptSupR}(p,q)$. The core idea of Algorithm 1 is that we select suppliers one by one and each time we select the supplier that achieves the largest marginal improvement in the revenue.

The following theorem presents the theoretical guarantees for Algorithm 1.

Theorem 3: Given a pricing scheme (p,q), the output $\hat{\mathcal{M}}^*$ of Algorithm 1 satisfies:

$$R(p,q,\hat{\mathcal{M}}^*) \geq \left(1 - \frac{1}{e}\right) R(p,q,\mathcal{M}^*),$$

where $\mathcal{M}^*$ denotes the exact the optimal set of suppliers under the pricing scheme (p,q).

Theorem 3 states that Algorithm 1 is able to locate a set of suppliers with an approximation ratio of at least $1 - 1/e$. The proof of this approximation ratio is based on the property of submodularity and monotonicity. The technique used to prove Theorem 3 is original from [1].

C. Design & Analysis of OptPrice(OptSupR)

Perfect search. Now we assume the case that given oracle OptSupR(p, q) outlined in Algorithm 1, we are able to obtain the exact optimal pricing scheme for Problem 4. Next, we give the following theorem, which states the impact of the proximity of OptSupR(p, q) on finding the optimal pricing scheme. We state this impact in the following theorem.

Theorem 4: Let $(\hat{p}^*, \hat{q}^*)$ denote the exact optimal pricing scheme of Problem 4 using oracle OptSupR(p, q) outlined in Algorithm 1. We have:

$$R(\hat{p}^*, \hat{q}^*, \hat{\mathcal{M}}^*(\hat{p}^*, \hat{q}^*)) \geq \left(1 - \frac{1}{e}\right) R(p^*, q^*, \mathcal{M}^*),$$

where $(p^*, q^*, \mathcal{M}^*)$ denotes one ground truth optimal solution to the Problem 1.

Theorem 4 states that with OptSupR(p, q) outlined in Algorithm 1, an optimal solution of Problem 4 obtains an approximation ratio of $(1 - 1/e)$ of the optimal solution of Problem 1. However, the optimal solution of Problem 4 is not easy to obtain. One challenge is that the closed-form expression of the revenue with respect to variables p and q is not available, let alone its gradient. Thus, we resort to gradient-free methods to solve Problem 4, in particular, the discretized search method. We leave it as future work to study other advanced methods such as Monte Carlo optimization to address this challenge.

Discretized search. We therefore discretize the domain of p, i.e., [0,1], uniformly:

$$\mathcal{A}(\epsilon_p) \triangleq \left\{0, \epsilon_p, 2\epsilon_p, \ldots, \left\lfloor \frac{1}{\epsilon_p} \right\rfloor \epsilon_p, 1\right\},$$

and discretize the domain of q, i.e., [0,1], uniformly:

$$\mathcal{A}(\epsilon_q) \triangleq \left\{0, \epsilon_q, 2\epsilon_q, \ldots, \left\lfloor \frac{1}{\epsilon_q} \right\rfloor \epsilon_q, 1\right\},$$

where $\epsilon_p, \epsilon_q \in (0, 1]$ are the search step of p and q respectively. By varying ϵ_p, ϵ_q, the OSN operator can adjust the number of searched prices in $\mathcal{A}(\epsilon_p)$ and $\mathcal{A}(\epsilon_q)$. The OSN operator can search the discretized pricing space to locate the optimal pricing scheme, denoted by $(p^*_{\mathrm{D}}, q^*_{\mathrm{D}})$. Algorithm 2 outlines this discretized search algorithm. The set of suppliers is then $\hat{\mathcal{M}}^*(p^*_{\mathrm{D}}, q^*_{\mathrm{D}})$. The following theorem states the theoretical guarantee of this method.

Theorem 5: Given $\epsilon = \sqrt{\epsilon_p^2 + \epsilon_q^2}$, suppose the objective function $R(p, q, \hat{\mathcal{M}}^*(p, q))$ is β-Lipschitz with respect to $(p, q) \in [0, 1]^2$, then the output $(p^*_{\mathrm{D}}, q^*_{\mathrm{D}}, \hat{\mathcal{M}}^*(p^*_{\mathrm{D}}, q^*_{\mathrm{D}}))$ of Algorithm 2 satisfies

$$R(p^*_{\mathrm{D}}, q^*_{\mathrm{D}}, \hat{\mathcal{M}}^*(p^*_{\mathrm{D}}, q^*_{\mathrm{D}})) \geq R(\hat{p}^*, \hat{q}^*, \hat{\mathcal{M}}^*(\hat{p}^*, \hat{q}^*)) - \beta\epsilon.$$

Theorem 5 states that when $R(p, q, \hat{\mathcal{M}}^*(p, q))$ is β-Lipschitz with respect to (p, q), the revenue gap between the discretized search method and the perfect search method is bounded by $\beta\epsilon$. A favorable property of this algorithm is that the OSN operator can make this gap arbitrarily small by selection ϵ_p and ϵ_q which can induce small enough ϵ. However, smaller ϵ leads to a larger computational complexity. Thus, Theorem 5 serves as a building block for an OSN operator to make a trade-off between computational complexity and approximate optimal value achieved. Combining them all, we prove the revenue gap between the approximate optimal pricing scheme output by Algorithm 2 and the ground truth optimal pricing scheme for Problem 1.

Algorithm 2: OptPrice (OptSupR)

```
1: Opt ← 0
2: for (p, q) ∈ A(ε_p) × A(ε_q) do
3:     M ← OptSupR(p, q)
4:     if R(p, q, M) ≥ Opt then
5:         p*_D ← p, q*_D ← q, M̂*(p*_D, q*_D) ← M
6:         Opt ← R(p, q, M)
7:     end if
8: end for
9: return p*_D, q*_D, M̂*(p*_D, q*_D)
```

Corollary 1: Given OptSupR(p, q) outlined in Algorithm 1. Suppose $R(p, q, \hat{\mathcal{M}}^*(p, q))$ is β-Lipschitz with respect to p and q The output $(p^*_{\mathrm{D}}, q^*_{\mathrm{D}}, \hat{\mathcal{M}}^*(p^*_{\mathrm{D}}, q^*_{\mathrm{D}}))$ of Algorithm 2 satisfies that

$$R(p^*_{\mathrm{D}}, q^*_{\mathrm{D}}, \hat{\mathcal{M}}^*(p^*_{\mathrm{D}}, q^*_{\mathrm{D}})) \geq \left(1 - \frac{1}{e}\right) R(p^*, q^*, \mathcal{M}^*) - \beta\epsilon.$$

IV. Algorithms for Optimal Pricing for Welfare

In this section, following a similar flow as Section III, we present the hardness analysis of Problem 2 and design approximation algorithms for it.

A. Hardness Analysis

Hardness analysis. To illustrate the hardness of Problem 2, we consider a sub-problem where pricing scheme (p, q) is given, which is stated as follows:

Problem 5 (Optimal supplier set $\mathcal{M}$ for welfare increase): Given pricing scheme (p, q) and budget b, select participating suppliers $\mathcal{M}$ from potential participating suppliers to maximize the welfare increase:

$$\begin{aligned} \max_{\mathcal{M}} \quad & W(p, q, \mathcal{M}) \\ s.t. \quad & \mathcal{M} \subseteq \widetilde{\mathcal{S}}(q), \\ & |\mathcal{M}| \leq b. \end{aligned}$$

We have the following theorem about the hardness of Problems 5 and 2.

Theorem 6: Problem 5 is NP-hard is not easier than NP-hard.

Our approach. Theorem 6 indicates that Problem 5 is time-consuming to locate an exact optimal solution. We aim to design an approximation algorithm for Problem 5, denoted by

$\texttt{OptSupW}(p, q)$. Then we use $\texttt{OptSupW}(p, q)$ as an oracle to search for the optimal pricing parameter (p, q) for Problem 2. Formally, we define the second sub-problem as follows.

Problem 6 (Optimal pricing for welfare with oracle): Given the algorithm $\texttt{OptSupW}(p, q)$, select (p, q) so to maximize the welfare increase:

$$\begin{aligned} \max_{p,q} \quad & W(p, q, \texttt{OptSupW}(p, q)) \\ s.t. \quad & p \in [0, 1], \\ & q \in [0, 1], \\ & p \geq q. \end{aligned}$$

We also design an algorithm for Problem 6, denoted by OptPrice(OptSupW), which takes $\texttt{OptSupW}(p, q)$ as an oracle. We next proceed to present the design and analysis of $\texttt{OptSupW}(p, q)$ and OptPrice(OptSupW).

B. Design & Analysis of $\texttt{OptSupW}(p, q)$

Submodular analysis. Given p and q are given, we first derive the closed form of the welfare increase $W(p, q, \mathcal{M})$ as

$$\begin{aligned} & W(p, q, \mathcal{M}) \\ &= \sum_{u \in \widetilde{\mathcal{R}}(p)} p_u I_u(p, q, \mathcal{M}) \\ &= \sum_{u \in \widetilde{\mathcal{R}}(p)} p_u \big| \cup_{l \in \mathcal{M} \times \widetilde{\mathcal{R}}(p)} \mathcal{V}(l^s, \tau - 1 - D(l^e, u)) \setminus \mathcal{V}(u, \tau) \big|. \quad (4) \end{aligned}$$

The following theorem shows the submodularity and monotonicity of the welfare increase with respect to $\mathcal{M}$.

Theorem 7: Given a pricing scheme (p, q), the revenue $W(p, q, \mathcal{M})$ is monotonously increasing and submodular with respect to participating suppliers $\mathcal{M}$.

The $\texttt{OptSupW}(p, q)$ algorithm. Based on Theorem 7, we modify Algorithm 1 slightly to implement $\texttt{OptSupW}(p, q)$, which is outlined in Algorithm 3. The only difference is that each time we select the supplier that achieves the largest marginal improvement in welfare increase rather than revenue.

The following theorem presents the theoretical guarantees for Algorithm 3.

Theorem 8: Given a pricing scheme (p, q), the output $\hat{\mathcal{M}}^*$ of Algorithm 3 satisfies:

$$W(p, q, \hat{\mathcal{M}}^*) \geq \left(1 - \frac{1}{e}\right) W(p, q, \mathcal{M}^*),$$

where $\mathcal{M}^*$ denotes the exact the optimal set of suppliers under the pricing scheme (p, q).

C. Design & Analysis of OptPrice(OptSupW)

To design OptPrice(OptSupW) to solve Problem 6, we can directly use the discretized search algorithm proposed in Section III-C. We can improve the search for this problem by the following strategy: if there exists a pricing scheme $(\bar{p}, \bar{q})$ which has been computed by Algorithm 3 and satisfies that

Algorithm 3: OptSupW (p, q)

1: init $\mathcal{M} = \emptyset$
2: **for** $t = 1$ to b **do**
3: $\quad u^* \leftarrow \arg\max_{u \in \widetilde{\mathcal{S}}(q)} W(p, q, \mathcal{M} \cup \{u\}) - W(p, q, \mathcal{M})$
4: $\quad \mathcal{M} \leftarrow \mathcal{M} \cup \{u^*\}$
5: **end for**
6: **return** $\hat{\mathcal{M}}^* \leftarrow \mathcal{M}$

$\widetilde{\mathcal{S}}(q) = \widetilde{\mathcal{S}}(\bar{q})$ and $\widetilde{\mathcal{R}}(p) = \widetilde{\mathcal{R}}(\bar{p})$, then we can directly take the result of $\texttt{OptSupW}(\bar{p}, \bar{q})$ as the result of $\texttt{OptSupW}(p, q)$. Moreover, Algorithms 3 and 1 have the same approximation ratio according to Theorems 8 and 3. Thus, the theoretical guarantees in Theorems 4, 5, and Corollary 1 still hold for the welfare increase version.

V. Algorithms for Fair Division of Contribution

In this section, we discuss how to divide the contribution, i.e., total visibility $I(p_D^*, q_D^*, \hat{\mathcal{M}}^*(p_D^*, q_D^*))$ fairly to suppliers. Problems 1 and 2 share the same challenge, so the solution proposed in this section applies to both of them. Note that fair division is a necessary part of the service. Without this fair division part, it is non-trivial for the operator to distribute rewards fairly, i.e., suppliers who contribute more should be rewarded more.

A. Shapley Value-Based Rewarding Rule

Our results so far can locate the approximate optimal price and supplier set, i.e., $(p_D^*, q_D^*, \hat{\mathcal{M}}^*(p_D^*, q_D^*))$. The remaining issue is how to divide the contribution to the network among participating suppliers fairly. As we mentioned in Section II that a *"fair"* division mechanism is important to incentivize the participation of suppliers. The OSN operator needs to divide the total contribution among all participating suppliers. Note that the naive equal division, i.e., participating suppliers equally share the total contribution is *not* a fair division. This is because: (1) some participating suppliers may have a larger number of followers while others may have a small number of followers; (2) the network structure poses an externality effect, causing the contribution of participating suppliers to be correlated. To achieve fair division, we apply the Shapley value [2]. The concept builds upon cooperative game theory where players are allowed to form coalitions in order to increase their payoffs in the game. One fundamental question in cooperative game theory is how to distribute the surplus achieved by cooperation among the players. To this end, Shapley in [2] proposed to reward agents with payoffs that correspond to their individual marginal contributions, known as the Shapley value.

With the concept of Shapley value, the key idea of our rewarding rule is to define a cooperative game on the solution $(p_D^*, q_D^*, \hat{\mathcal{M}}^*(p_D^*, q_D^*))$ output by our algorithms. Here, $\hat{\mathcal{M}}^*(p_D^*, q_D^*)$ is called grand coalition and shortened as $\hat{\mathcal{M}}^*$ in this section for simplicity. We next define a characteristic function Φ which assigns to every coalition $\mathcal{C} \subseteq \hat{\mathcal{M}}^*$ a real number representing the contribution of $\mathcal{C}$. We take the total visibility increase as the total contribution of the grand coalition, and the contribution of a coalition $\mathcal{C} \subseteq \hat{\mathcal{M}}^*$ is modeled as $\Phi(\mathcal{C}) = I(p_{\rm D}^*, q_{\rm D}^*, \mathcal{C})$. The OSN operator needs to divide a total contribution of $\Phi(\hat{\mathcal{M}}^*) = I(p_{\rm D}^*, q_{\rm D}^*, \hat{\mathcal{M}}^*)$ among the participating suppliers and reward them q per unit contribution.

Now we are ready to show how to compute the reward for each supplier. As we have mentioned, the computation of Shapley value is taken as the individual marginal contribution. More specifically, for a given player the individual marginal contribution is measured as the weighted average marginal increase in the payoff of any coalition that this agent could potentially join. We use $\phi_u(\hat{\mathcal{M}}^*)$ to denote the u's Shapley value when the grand coalition is $\hat{\mathcal{M}}^*$. Formally, in the game modeled for our problem, each participating supplier $u \in \hat{\mathcal{M}}^*$ gets the following share of Φ:

$$\phi_u(\hat{\mathcal{M}}^*) = \sum_{\mathcal{M} \subseteq \hat{\mathcal{M}}^* \setminus \{u\}} \frac{|\mathcal{M}|!(|\hat{\mathcal{M}}^*| - |\mathcal{M}| - 1)!}{|\hat{\mathcal{M}}^*|!} \times (\Phi(\mathcal{M} \cup \{u\}) - \Phi(\mathcal{M})).$$

Shapley value is a division scheme that meets the following desirable criteria in our problems: (1) efficiency: total reward $q\Phi(\hat{\mathcal{M}}^*)$ achieved by the grand coalition $\hat{\mathcal{M}}^*$ is distributed among them; (2) symmetry: the rewards to suppliers do not depend on their identity; and (3) null player: suppliers with zero marginal contributions to all coalitions receive zero reward. Readers can refer to [2] for more details about Shapley value.

B. Computational Complexity

One challenge is that the computational complexity of evaluating $\phi_u(\hat{\mathcal{M}}^*)$ is exponential in the cardinality of $\hat{\mathcal{M}}^*$. To address this computational challenge, we propose to use the sampling algorithm [3] to approximate $\phi_u(\hat{\mathcal{M}}^*)$. Let $\sigma = (u_1, \ldots, u_{\widetilde{b}})$ denote an ordering of the participating suppliers, where $\widetilde{b} = \min\{b, |\hat{\mathcal{M}}^*|\}$ and $u_i \in \hat{\mathcal{M}}^*$ denotes the participating supplier in the i-th order. Denote the set of players ranked before player u_i in the order σ as

$$\mathcal{S}_{u_i}^{\sigma} \triangleq \{\text{all players ranked before } u_i \text{ in the order } \sigma\}.$$

Based on [3], for each supplier $u \in \hat{\mathcal{M}}^*$, the Shapley value $\phi_u(\hat{\mathcal{M}}^*)$ can be rewritten as

$$\phi_u(\hat{\mathcal{M}}^*) = \mathbb{E}_{\sigma \sim \texttt{Uniform}(\Omega)}[\Phi(\mathcal{S}_u^{\sigma} \cup \{u\}) - \Phi(\mathcal{S}_u^{\sigma})], \quad (5)$$

where Ω denotes the set of all orderings of participating suppliers, and $\texttt{Uniform}(\Omega)$ denotes a uniform distribution over Ω. Based on Eq. (5), we use Algorithm 4 to approximate the Shapley value $\phi_u(\hat{\mathcal{M}}^*)$ of each supplier $u \in \hat{\mathcal{M}}^* \setminus \{\bar{s}\}$, where supplier $\bar{s}$ is selected randomly from $\hat{\mathcal{M}}^*$. Then, we set the contribution of $\bar{s}$ as $\hat{\phi}_{\bar{s}}(\hat{\mathcal{M}}^*) = I(p_{\rm D}^*, q_{\rm D}^*, \hat{\mathcal{M}}^*) - \sum_{u \in \hat{\mathcal{M}}^* \setminus \{\bar{s}\}} \hat{\phi}_u(\hat{\mathcal{M}}^*)$ to guarantee that the sum of the contributions of suppliers in $\hat{\mathcal{M}}^*$ is equal to $I(p_{\rm D}^*, q_{\rm D}^*, \hat{\mathcal{M}}^*)$.

The following theorem states the theoretical guarantee for the approximation accuracy of Algorithm 4.

Algorithm 4: Approximating $\phi_u(\hat{\mathcal{M}}^*)$.

1: $\hat{\phi}_u = 0$
2: **for** $k = 1$ to K **do**
3: generate an ordering σ uniformly at random from Ω
4: $\hat{\phi}_u \leftarrow [(k-1)\hat{\phi}_u + \Phi(\mathcal{S}_u^{\sigma} \cup \{u\}) - \Phi(\mathcal{S}_u^{\sigma})]/k$
5: **end for**
6: **return** $\hat{\phi}_u$

Theorem 9: The estimated contribution satisfies

$$\sum\nolimits_{u \in \hat{\mathcal{M}}^*} \hat{\phi}_u = I(p_{\rm D}^*, q_{\rm D}^*, \hat{\mathcal{M}}^*).$$

Furthermore, for each $u \in \hat{\mathcal{M}}^* \setminus \{\bar{s}\}$, it holds that

$$|\hat{\phi}_u(\hat{\mathcal{M}}^*) - \phi_u(\hat{\mathcal{M}}^*)| \leq \frac{\max_{\sigma \in \Omega}[\Phi(\mathcal{S}_u^{\sigma} \cup \{u\}) - \Phi(\mathcal{S}_u^{\sigma})]}{\sqrt{K}} \sqrt{\frac{1}{2} \ln \frac{2}{\delta}},$$

with a probability of at least $1 - \delta$, where $\delta \in (0, 1]$.

Theorem 9 states that one can control the approximation error of Algorithm 4 arbitrarily small by selecting sufficiently large simulation rounds K.

C. Utility Performance

With the reward rule proposed, we are now ready to discuss the utility of suppliers. Using our proposed Shapley value-based rewarding rule, a supplier $u \in \hat{\mathcal{M}}^*$ would be distributed a reward of $q_{\rm D}^* \phi_u(\hat{\mathcal{M}}^*)$. The valuation q_u of a supplier u can be interpreted as the cost of the unit contribution. We then define the utility of supplier $u \in \hat{\mathcal{M}}^*$ as

$$U_u = (q_{\rm D}^* - q_u)\phi_u(\hat{\mathcal{M}}^*).$$

First, the condition for a supplier $u \in \mathcal{S}$ to be a potential participating supplier in $\tilde{\mathcal{S}}(q_{\rm D}^*)$ is $q_u < q_{\rm D}^*$ and participating suppliers $\hat{\mathcal{M}}^*$ are selected from $\tilde{\mathcal{S}}(q_{\rm D}^*)$, so we have the term $q_{\rm D}^* - q_u \geq 0, u \in \hat{\mathcal{M}}^* \subseteq \tilde{\mathcal{S}}(q_{\rm D}^*)$. Secondly, the Shapley value is always non-negative, so we have $\phi_u(\hat{\mathcal{M}}^*) \geq 0, u \in \hat{\mathcal{M}}^*$. Thus, the utility (i.e., the product of the above two terms) of each selected supplier must be non-negative for both of Problems 1 and 2. Thus, the fair division is a necessary part to guarantee that suppliers with q_u lower than q would all be willing to participate in the service since we assume users are rational. Other division methods, e.g., the naive average division would cause some suppliers to get negative utility and suppliers would be unwilling to participate in the service, which can cause a loss of revenue and social welfare increase in the long run.

VI. Performance Evaluation

In this section, we conduct experiments on real-world datasets to evaluate the performance of our algorithms, and results show their superior performance.

TABLE I
STATISTICS OF FOUR DATASETS

datasets	#nodes	#links	type	τ
Residence	217	2,672	directed	2
Blogs	1,224	19,025	directed	2
DBLP	10,000	55,734	undirected	2

A. Experimental Settings

Datasets. We evaluate our algorithms on four public datasets, whose overall statistics are summarized in Table I.

• *Residence [4].*This dataset contains friendship connections between 217 residents, who live at a residence hall in the Australian National University campus.

• *Blogs [4].*This dataset contains hyperlinks between blogs in the context of the 2004 US election. Blogs are mapped as nodes and hyperlinks are mapped as directed links.

• *DBLP [5].*This dataset contains a sub-network of the co-author network of the DBLP network. Scholars who have published papers in major conferences (those considered in DBLP) are mapped as nodes. Each co-author relationship between two scholars is mapped as two directed edges between these two scholars with different directions.

From Table I, one may argue that the scales of the above four datasets are not large. We intentionally make this choice because: (1) we have already proved the quality gap; (2) to compare with baseline algorithms such as the brute force method, the OSN has to be small to make it computationally feasible.

Parameter setting. To reflect the real-world setting that only a small portion of users in an OSN are interested in the social visibility boosting service, we select γ fraction of users uniformly at random from the user population $\mathcal{U}$ as requesters $\mathcal{R}$, and another γ fraction of users uniformly at random from the user population $\mathcal{U}$ as suppliers $\mathcal{S}$. We set γ as 0.25, 0.1 and 0.05 for dataset Residence, Blogs, and DBLP respectively. We set the social visibility threshold $\tau = 2$ by default and also vary τ from 1,2 to 3 in Section VI-E to study the impact of τ on the performance of our proposed algorithms.

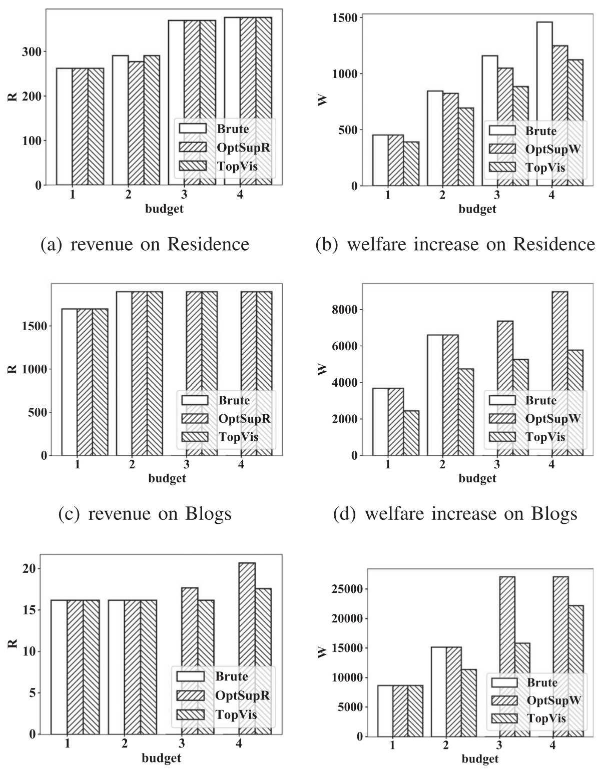
To generate synthetic valuation p_u for each requester $u \in \mathcal{R}$ and q_u for each supplier $u \in \mathcal{S}$, we define functions based on their visibility respectively. Specifically, we synthesize the valuation of a requester with visibility v as

$$p(v) \triangleq \frac{(1 + v/v_{max})^{\lambda}}{2^{\lambda}} \in [0, 1], \tag{6}$$

and synthesize the valuation of a supplier with visibility v as

$$q(v) \triangleq 1 - \frac{(1 + v/v_{max})^{\lambda}}{2^{\lambda}} \in [0, 1], \tag{7}$$

where v is the visibility of the user, v_{max} is the largest visibility in the network and $\lambda(\lambda > 0)$ is a parameter of the function to control the relationship between of u's visibility and p_u or q_u. For any $\lambda > 0$, (6) models that a requester with larger visibility tends to have a larger valuation p_u for unit visibility increase, and (7) models that a supplier with smaller visibility tends to have a larger valuation q_u for a unit contribution made.



(a) revenue on Residence (b) welfare increase on Residence

(c) revenue on Blogs (d) welfare increase on Blogs

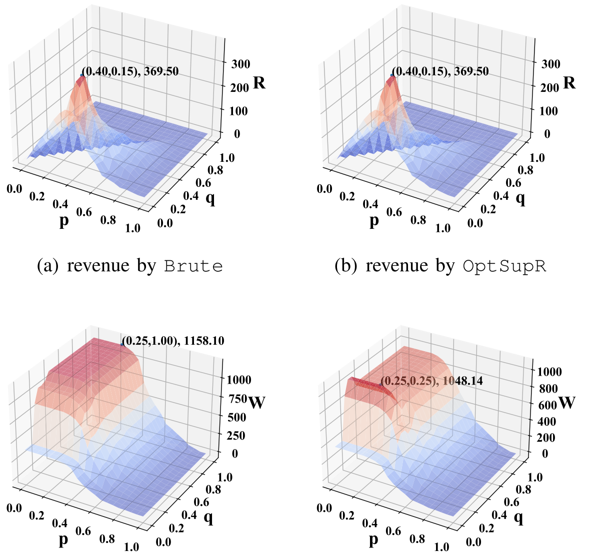
(e) revenue on DBLP (f) welfare increase on DBLP

Fig. 2. Algorithms to select optimal set ($\lambda = 2$, $\epsilon = 0.05$).

Metrics & baselines. We use the revenue and the total welfare increase achieved as the evaluation metrics. First, to understand the accuracy and efficiency of `OptSupR` (Algorithm 1) and `OptSupW` (Algorithm 3), we compare them with the following two baselines under each given pricing scheme (p, q): (1) `Brute`, which selects the optimal set of participating suppliers via exhaustive search; (2) `TopVis`, which selects potential participating suppliers who rank top-b by their social visibility. Secondly, to evaluate our `OptPrice` algorithm, i.e., Algorithm 2 and study the parameters of algorithms, we vary the search step $\epsilon_p = \epsilon_q = \epsilon$ from 0.2, 0.1 to 0.05. Then, we also study the impact of the functions to synthesize valuations for requesters and suppliers by varying the parameter λ from 1,2 to 3. Lastly, we study the impact of the social visibility threshold τ by varying τ from 1,2 to 3.

B. Evaluating `OptSupR`/`OptSupW`

We first compare our algorithms `OptSupR` and `OptSupW` with two baselines (1) `Brute` and (2) `TopVis`. In this subsection, we fix the search step as $\epsilon = 0.05$ and set the parameter $\lambda = 2$ for valuation functions in (6) and (7). Note that algorithm `Brute` is computationally expensive, so we do not run `Brute` for the setting with relatively large budgets and large network sizes. Fig. 2 shows the revenue and the welfare increase achieved by different methods of finding optimal supplier set on datasets Residence, Blogs, and DBLP. From Fig. 2 (a), we can observe that the revenue achieved by `OptSupR`



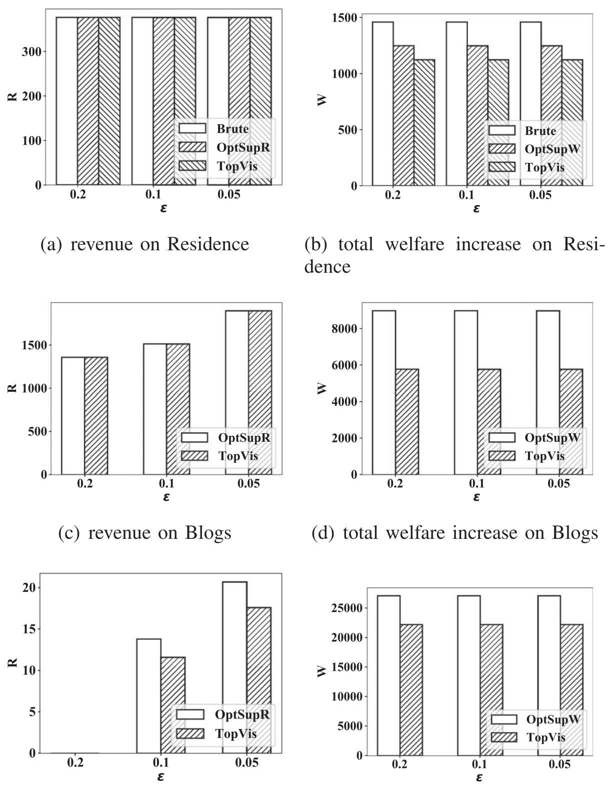
(a) revenue by `Brute` (b) revenue by `OptSupR`

(c) welfare increase by `Brute` (d) welfare increase by `OptSupW`

Fig. 3. Intermediate results on residence ($\lambda = 2$, $\epsilon = 0.05$, $b = 3$).

and `TopVis` are both nearly the same as `Brute` algorithm, i.e., ground truth optimum, which indicates that `OptSupR` and `TopVis` both perform very well on dataset Residence. From Fig. 2(b), we can observe that the welfare increase achieved by `OptSupW` is slightly lower than `Brute` and higher than `TopVis`, which shows the effectiveness of `OptSupW`. For the larger dataset Blog, since it is too time-consuming to run Brute for budgets larger than 2, we leave the corresponding bar blank. Fig. 2(c) shows that `OptSupR` and `TopVis` both perform closely on dataset Blogs and achieve the ground truth optimum for small budgets 1 and 2. Fig. 2(d) shows that `OptSupW` achieves the same welfare increase as `Brute` for small budgets 1 and 2, and achieves much higher values than `TopVis` on Blogs. Fig. 2(e) shows the revenue performance on dataset DBLP, where `OptSupR` performs similarly well as `Brute` and better than `TopVis` for larger budgets. Fig. 2(f) shows `OptSupW` achieves the same welfare increase as `Brute` for small budgets 1 and 2, and achieves much higher values than `TopVis` on DBLP.

Next, we will zoom in on all searched pricing schemes and take a closer look at `OptSupR`/`OptSupW` and `Brute`. Fig. 3 are the results on dataset Residence where we fix $\lambda = 2$, $\epsilon = 0.05$ and $b = 3$. Fig. 3(a) and (b) show the intermediate results when we search for the optimal pricing scheme to maximize the revenue using `OptSupR` and `Brute` respectively. For most searched pricing schemes the optimal revenue is very close. This shows that the subproblem Problem 3 are well solved by `OptSupR`. Fig. 3(c) and (d) show the intermediate results when we search for the optimal pricing scheme to maximize the welfare increase using `OptSupW` and `Brute` respectively. For most searched pricing schemes the optimal welfare increase is very close. This shows that the subproblem Problem 3 can be well solved by `OptSupW`.



(a) revenue on Residence (b) total welfare increase on Residence

(c) revenue on Blogs (d) total welfare increase on Blogs

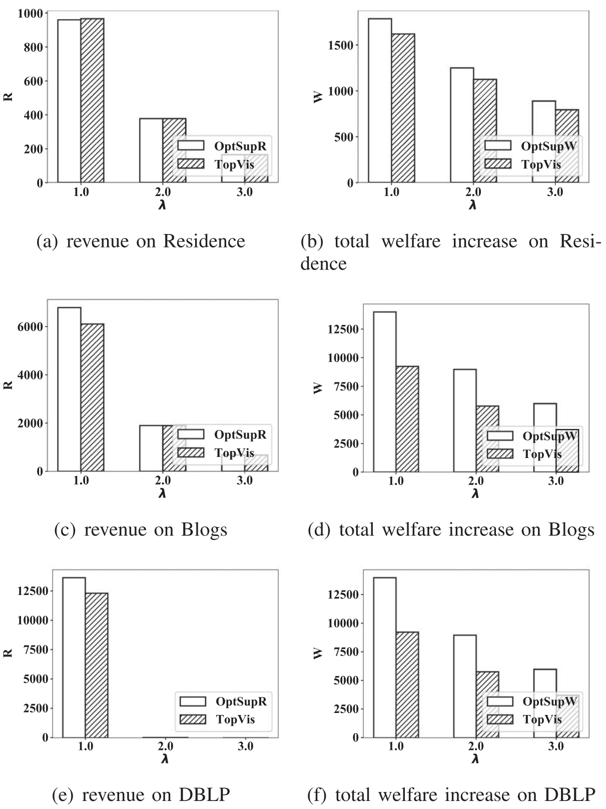
(e) revenue on DBLP (f) total welfare increase on DBLP

Fig. 4. Performance under different ϵ ($\lambda = 2$, $b = 4$).

C. *Evaluating* `OptPrice`

We study the impact of the search step on `OptPrice`, i.e., Algorithm 2. We vary the search step $\epsilon_p = \epsilon_q = \epsilon$ from 0.2, 0.1 to 0.05 on all datasets. From Fig. 4(a) we can observe that on dataset Residence, all the three algorithms `Brute`, `OptSupR` and `TopVis`, achieve the same optimal revenue under different search steps 0.2, 0.1, and 0.05. Fig. 4(c) and (e) show that the revenue achieved by all tested methods increase as ϵ decreases on dataset Blogs and DBLP. In contrast, the performance of welfare increase is not sensitive to ϵ. From Fig. 4(b), (d), and (f), we can observe that on all three datasets, all the three algorithms `Brute`, `OptSupW`, and `TopVis`, achieve almost the same optimal total welfare increase under different search steps 0.2, 0.1 and 0.05, which indicates a total welfare increase is not sensitive to the search step ϵ.

Note that smaller ϵ causes larger search space and larger computational costs. The above observation can be instructive to set an appropriate value for ϵ in Algorithm 2 due to the following reasons. (1) Problem 1 is more sensitive to search step ϵ than Problem 2. Reducing the step size can be effective to improve the performance when we search for the optimal pricing scheme for revenue while may not make a difference to the performance when we search for the optimal pricing scheme for welfare increase. (2) When we decrease ϵ in the setting of maximizing revenue, there is not much difference in revenue improvement for small datasets like Residence, while

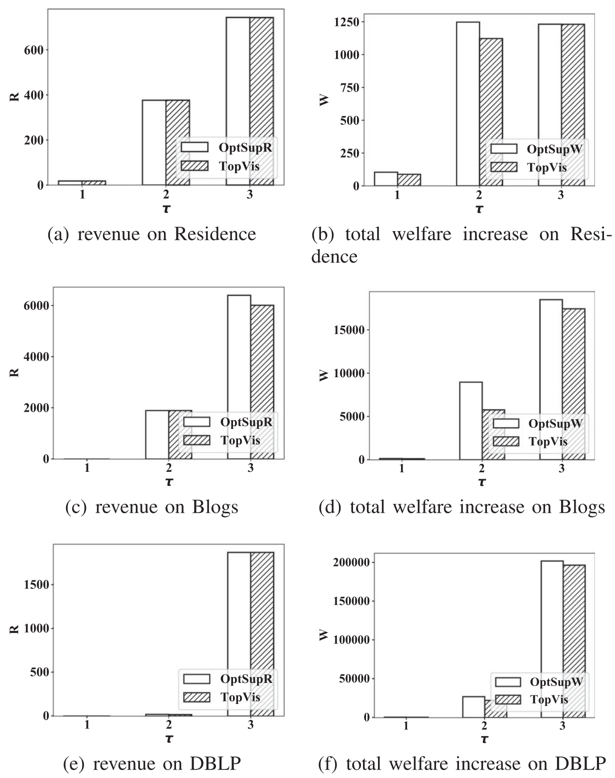


(a) revenue on Residence (b) total welfare increase on Residence

(c) revenue on Blogs (d) total welfare increase on Blogs

(e) revenue on DBLP (f) total welfare increase on DBLP

Fig. 5. Performance under different λ ($\epsilon = 0.05$, $b = 4$).



(a) revenue on Residence (b) total welfare increase on Residence

(c) revenue on Blogs (d) total welfare increase on Blogs

(e) revenue on DBLP (f) total welfare increase on DBLP

Fig. 6. Performance under different τ ($\epsilon = 0.05$, $b = 4$, $\lambda = 2$).

it can bring larger improvement on larger datasets like Blogs and DBLP. Thus, one can choose a larger ϵ for smaller datasets to balance the computational cost and the improvement.

D. The Impact of Valuation Functions

We compare the different functions to synthesize p_u and q_u. We fix budget $b = 4$ and search step $\epsilon = 0.05$, and vary the function parameter λ from 1,2 to 3. Different λ means different sensitivity to visibility. Fig. 5 shows the results on the three datasets. On all three datasets, we can observe that for both revenue and welfare increase and all algorithms larger λ causes lower optimal objective values. This is because λ has an impact on the distribution of p_u and q_u. Specifically, with larger λ, the requesters tend to have lower valuations for unit visibility increase and the suppliers tend to have higher valuations for unit contribution, which means there tend to be fewer suppliers and requesters under the same pricing scheme.

E. The Impact of τ

We study the impact of social visibility threshold τ. We fix budget $b = 4$ and search step $\epsilon = 0.05$, and vary τ from 1 to 3. Fig. 6(a), (c), and (e) show the revenue achieved on three datasets. One can observe that a larger threshold τ leads to higher revenue for both algorithms on all three datasets. Fig. 6(b), (d), and (f) show the total welfare increase on three datasets. On datasets Blogs and DBLP, one can observe that both algorithms achieve higher total welfare increase under larger τ. However, on the dataset Residence, `OptSupW` achieves higher revenue in $\tau = 2$ than $\tau = 3$. The reasons are as follows. It is very likely that most users in 2-visible set of participating suppliers are already in the requester's 3-visible set, which leads to a smaller visibility increase. Such overlap tends to occur in a small network. Moreover, from Fig. 6(a) and (b), one can observe a jump in both revenue and total welfare increase on the largest dataset DBLP when we increase τ from 2 to 3, which indicates the parameter τ has a larger impact on larger datasets.

VII. Related Work

The notion of social visibility defined in this paper is closely related to social influence [6], [7], [8], [9]. One key difference is that when a user is visible to a set of users, it does not mean this user can influence this set of users. The objective of the influence maximization problem is to find a subset of nodes that could maximize the spread of information under certain influence diffusion models. But our problem focuses on the pricing of the social visibility service. The idea of adding new links to enhance social visibility is closely related to link prediction [10], [11], [12], and friend recommendation [13], [14], [15], [16], The objectives of link prediction and friend recommendation are to predict future or missing links. Our work adds links that can improve social visibility. Note that such links may have nothing to do with predicting the future or missing links.

Technically, our work is closely related to revenue management [17]. Different from classical revenue management literatures [17], which focuses on understanding the structure of optimal pricing, our work formulates a new revenue maximization framework and we focus on designing approximation algorithms to solve this problem. Similar to influence maximization [6], [7], [9], the core technique in selecting the supplier set is submodular analysis. Our contribution is in proving that our problem has the submodular property and show how the submodular property impacts the search optimal pricing.

VIII. Conclusion

This article proposes a posted pricing scheme for the OSN operator to price its social visibility boosting service. We formulate revenue/welfare increase maximization problems for the OSN operator to select the optimal pricing scheme. We show that revenue/welfare increase maximization problems are not simpler than an NP-hard problem. We decomposed revenue/welfare increase maximization problems into two sub-routines respectively, where one focuses on selecting the optimal set of suppliers, and the other one focuses on selecting the optimal prices. We prove the hardness of each sub-routine, and eventually design a computationally efficient approximation algorithm to solve the problems with provable theoretical guarantee on the revenue/welfare increase gap. We conduct extensive experiments on four public datasets to validate the performance of our proposed algorithms.

References

[1] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, “An analysis of approximations for maximizing submodular set functions,” *Math. Program.*, vol. 14, no. 1, pp. 265–294, 1978.

[2] L. S. Shapley, “A value for N-person games,” in *Contributions to the Theory of Games*, vol. II. Princeton, NJ, USA: Princeton Univ. Press, 1953, pp. 307–317.

[3] Y. Bachrach, E. Markakis, E. Resnick, A. D. Procaccia, J. S. Rosenschein, and A. Saberi, “Approximating power indices: Theoretical and empirical analysis,” *Auton. Agents Multi-Agent Syst.*, vol. 20, no. 2, pp. 105–122, 2010.

[4] J. Kunegis, “KONECT–The Koblenz network collection,” in *Proc. Int. Conf. World Wide Web Companion*, 2013, pp. 1343–1350.

[5] W. Nawaz, K.-U. Khan, Y.-K. Lee, and S. Lee, “Intra graph clustering using collaborative similarity measure,” *Distrib. Parallel Databases*, vol. 33, no. 4, pp. 583–603, 2015.

[6] D. Kempe, J. Kleinberg, and É. Tardos, “Maximizing the spread of influence through a social network,” in *Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2003, pp. 137–146.

[7] P. Domingos and M. Richardson, “Mining the network value of customers,” in *Proc. 7th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2001, pp. 57–66.

[8] Y. Lin, W. Chen, and J. C. Lui, “Boosting information spread: An algorithmic approach,” in *Proc. IEEE 33rd Int. Conf. Data Eng.*, 2017, pp. 883–894.

[9] M. Richardson and P. Domingos, “Mining knowledge-sharing sites for viral marketing,” in *Proc. 8th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2002, pp. 61–70.

[10] L. Lü and T. Zhou, “Link prediction in complex networks: A survey,” *Physica A: Stat. Mech. its Appl.*, vol. 390, no. 6, pp. 1150–1170, Mar. 2011.

[11] V. Martínez, F. Berzal, and J.-C. Cubero, “A survey of link prediction in complex networks,” *ACM Comput. Surv.*, vol. 49, no. 4, 2017, Art. no. 69.

[12] D. Liben-Nowell and J. Kleinberg, “The link-prediction problem for social networks,” *J. Amer. Soc. Inf. Sci. Technol.*, vol. 58, no. 7, pp. 1019–1031, 2007.

[13] P. Resnick and H. R. Varian, “Recommender systems,” *Commun. ACM*, vol. 40, no. 3, pp. 56–59, 1997.

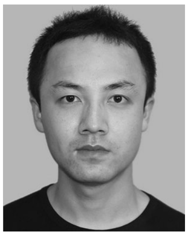
[14] H. Ma, “On measuring social friend interest similarities in recommender systems,” in *Proc. 37th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2014, pp. 465–474.

[15] H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King, “Recommender systems with social regularization,” in *Proc. 4th ACM Int. Conf. Web Search Data Mining*, 2011, pp. 287–296.

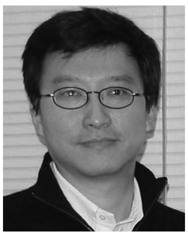
[16] X. Xie, “Potential friend recommendation in online social network,” in *Proc. IEEE/ACM Conf. Green Comput. Commun. Int. Conf. Cyber, Phys. Social Comput..*, 2010, pp. 831–835.

[17] K. T. Talluri, G. Van Ryzin, and G. Van Ryzin, *The Theory and Practice of Revenue Management*, vol. 1. Berlin, Germany: Springer, 2004.

Shiyuan Zheng received the Ph.D. degree in computer science and engineering from the Chinese University of Hong Kong, Hong Kong, in 2022, under the supervision of Prof. John C. S. Lui. Her research interests include social network analysis, game theory, and multi-armed bandit problems.



Hong Xie (Member, IEEE) received the Ph.D. degree from the Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong, in 2015, under the supervision of Prof. John C.S. Lui. He received his B.Eng. degree from the School of Computer Science and Technology, The University of Science and Technology of China, Hefei, China, in 2010. He was a Postdoctoral Research Fellow with The Chinese University of Hong Kong and National University of Singapore, Singapore.



John C.S. Lui (Fellow, IEEE) received the Ph.D degree in computer science from the University of California at Los Angeles, Los Angeles, CA, USA. From 2005 to 2011, he was the Chairman with the CSE Department. He is currently the Choh-Ming Li Chair Professor with the Department of Computer Science and Engineering, The Chinese University of Hong Kong. His research interests include quantum Internet, online machine learning. He was the recipient of the various departmental teaching awards and the CUHK Vice-Chancellors Exemplary Teaching Award. He is an Elected Member of the IFIP WG 7.3 and the Croucher Senior Research Fellow.