Incentive-Compatible Selection for One or Two Influentials

Yuxin Zhao, **Yao Zhang** and **Dengji Zhao**
ShanghaiTech University
{zhaoyx5, zhangyao1, zhaodj}@shanghaitech.edu.cn

Abstract

Selecting influentials in networks against strategic manipulations has attracted many researchers' attention and it also has many practical applications. Here, we aim to select one or two influentials in terms of progeny (the influential power) and prevent agents from manipulating their edges (incentive compatibility). The existing studies mostly focused on selecting a single influential for this setting. Zhang *et al.* [2021] studied the problem of selecting one agent and proved an upper bound of $1/(1 + \ln 2)$ to approximate the optimal selection. In this paper, we first design a mechanism to actually reach the bound. Then, we move this forward to choosing two agents and propose a mechanism to achieve an approximation ratio of $(3 + \ln 2)/(4(1 + \ln 2))$ (≈ 0.54).

1 Introduction

Consider the scenario where we want to select influential agents in a network constructed by referral relationships (e.g., the following relationships in Twitter, the citations between academic papers, etc.). The selected agents may be rewarded with prizes or benefits (e.g., job opportunities [Kotturi *et al.*, 2020]). Hence, agents have the incentive to manipulate their relationships to make themselves selected. Therefore, selection mechanisms that can prevent agents from strategic manipulations (which is referred to as the property of incentive compatibility) are highly demanded [Alon *et al.*, 2011].

Many studies have investigated incentive-compatible selection mechanisms on different influence measurements for different purposes (see [Olckers and Walsh, 2022] for a complete survey). In this paper, we focus on the setting where an agent's influential power is measured by her progeny (the number of all agents who directly or indirectly follow her). For this setting, two studies have been conducted before. Babichenko *et al.* [2020b] proposed the first single agent selection mechanism for progeny maximization that can prevent agents from adding or hiding their out-edges. Their mechanism reaches an approximation ratio of about $1/3$ (i.e., the expected progeny of the chosen agent is about $1/3$ of the largest). However, their mechanism only works in forests. Therefore, Zhang *et al.* [2021] further studied the same problem in directed acyclic graphs (DAGs) with restricting manipulations in the scope of hiding edges (agents cannot add new edges). Their proposed mechanism achieves an approximation ratio of $1/2$. Moreover, they proved an upper bound $1/(1 + \ln 2)$ of the approximation ratio for any incentive-compatible and fair selection mechanism in the DAG setting.

In this paper, we follow the DAG setting of [Zhang *et al.*, 2021] and make the following contributions:

- For selecting one agent, we close the gap between the known approximation ratio of $1/2$ and the upper bound of $1/(1+\ln 2)$. We propose a mechanism to achieve the exact upper bound.
- For selecting two agents, we show that, for the class of mechanisms that only select agents from the 1-influential set[1] (most of the existing mechanisms belong to this class), the approximation ratio cannot exceed $1/2$ if the target is to select at most two agents. Moreover, we provide a deterministic mechanism in this class that exactly reaches the approximation ratio of $1/2$.
- We then propose a new incentive-compatible mechanism based on a 2-influential set for selecting two agents. The new mechanism achieves a higher approximation ratio of $(3 + \ln 2)/(4(1 + \ln 2))$ (≈ 0.54). We also provide a general upper bound ($23/27$) of any incentive-compatible mechanism for selecting two agents.

1.1 Other Related Work

Many studies on incentive-compatible selection mechanisms use in-degrees to measure agents' influential power, which is also referred to as peer selection. For the in-degree measurement, Alon *et al.* [2011] firstly proposed an incentive-compatible peer selection mechanism by a randomized partition method, which divides agents into two groups and chooses the agents according to their in-degrees from the other group. Following this work, there are two major directions. One is to characterize incentive-compatible peer selection mechanisms with axioms, which is initiated by Holzman and Moulin [2013]. Mackenzie [2015] continued this

[1] 1-influential set contains all agents each of whom has the largest progeny by hiding her out-edges.

study by adding symmetric axiomizations. The other direction is to improve the approximation ratios of the existing incentive-compatible peer selection mechanisms. Fischer and Klimm [2014] extended the idea of the partition mechanism to a permutation mechanism, which achieves the optimal approximation ratio ($1/2$) for selecting one agent with in-degree. Then, Bousquet *et al.* [2014] characterized a class of networks where the permutation mechanism selects an agent close to the optimal. Bjelde *et al.* [2017] generalized the permutation mechanism for selecting multiple agents, and gave both lower and upper bounds of the approximation ratio. Recent studies also started to consider an alternative evaluation called additive approximation, which focuses on the expected difference from the optimal rather than the worst-case ratio [Caragiannis *et al.*, 2021; Caragiannis *et al.*, 2022; Cembrano *et al.*, 2022]. There are also extensions on the networks, including a weighted network, where the influential power is a weighted in-degree [Kurokawa *et al.*, 2015; Wang *et al.*, 2018; Babichenko *et al.*, 2020a], and rank aggregation, where each agent assigns a rank to others [Kahng *et al.*, 2018; Mattei *et al.*, 2020].

There is also a rich body of work focusing on different measurements of influential power. Ghalme *et al.* [2018] designed a naturally strategy-proof score function to measure the popularity of agents, thereby simplifying the selection. In contrast, Babichenko *et al.* [2018] targeted a PageRank-like centrality [Page *et al.*, 1999] and offered a two-path mechanism which achieves a good approximation ratio of $2/3$. Moreover, Babichenko *et al.* [2020b] focused on the progeny of the selected agent and proposed a mechanism with an approximation ratio of $1/(4\ln 2)$ in forests. Zhang *et al.* [2021] then proposed a geometric mechanism with an approximation ratio of $1/2$ in DAGs, which is what we follow here.

2 Preliminaries

Let $\mathcal{G}_n$ be the set of all directed acyclic graphs (DAGs) with n nodes and $\mathcal{G} = \bigcup_{n\in\mathbb{N}^+} \mathcal{G}_n$ be the set of all DAGs. Consider a network represented by a graph $G = (N, E) \in \mathcal{G}$, where $N = \{1, 2, \dots, n\}$ is the node set and E is the edge set. Each node $i \in N$ represents an agent, and each edge $(i, j) \in E$ represents that agent i follows (votes for, or quotes) agent j. Let $E_i = \{(i, j) \mid (i, j) \in E, j \in N\}$ be the set of all edges from i. We say an agent j is influenced by the agent i if there exists a path from j to i in G. Let $P(i, G)$[2] be the set of agents who are influenced by agent i (including i herself), which is referred to as the progeny of agent i.

Our goal is to select a group of agents in the network as delegates with larger progeny. Let $\mathcal{S}_k = \{S \mid S \subseteq N, |S| = k\}$ be the set of all subsets with k agents, and $\mathcal{S}_{\le k} = \bigcup_{t=0}^{k} \mathcal{S}_t$. A *$k$-selection mechanism* decides how to choose up to k agents as delegates.

Definition 1. *A k-selection mechanism for $\mathcal{G}$ is a family of functions $f : \mathcal{G}_n \to [0, 1]^{\mathcal{S}_{\le k}}$ for all $n \in \mathbb{N}^+$, that maps each graph to a probability distribution on all subsets with no more than k agents.*

[2]or simply $P(i)$ if no ambiguity. Additionally, for the sake of tidiness, the notation of the set also indicates its size if no ambiguity.

For a given graph $G \in \mathcal{G}$ and a k-selection mechanism f, denote $x_S(G) = (f(G))_S$ as the probability of the subset $S \in \mathcal{S}_{\le k}$ being selected, and $x_i(G) = \sum_{S\in\mathcal{S}_{\le k} \text{ and } i\in S} x_S(G)$ as the probability of the agent i being selected.

For an agent $i \in N$, she wants her probability to be selected (x_i) as large as possible, while for the owner of the mechanism, we want the influential power of the selected group (the sum of the progeny) as large as possible. Unfortunately, if we simply choose an optimal subset with k agents, i.e., $S_k^* \in \arg\max_{S\in\mathcal{S}_k} \sum_{i\in S} |P(i, G)|$, then agents will have incentives to hide their edges to increase their ranks to be selected. We want to avoid such a manipulation, which requires the mechanism to be *incentive-compatible*.

Definition 2. *A k-selection mechanism for $\mathcal{G}$ is incentive-compatible (IC) if for every $n \in \mathbb{N}^+$, and every two graphs $G = (N, E)$, $G' = (N, E') \in \mathcal{G}_n$, such that $E \setminus E_i = E' \setminus E'_i$ and $E_i \supseteq E'_i$ for $i \in N$, we have $x_i(G) \ge x_i(G')$.*

Intuitively, incentive compatibility implies that no matter how other agents follow each other, it is an undominated strategy for any agent not to hide her out-edges. Since an incentive-compatible k-selection mechanism cannot always choose a group with the highest influential power, we seek the approximation of the optimum, which guarantees a worst-case ratio between the expected progeny of the selected group and an optimal group for all DAGs.

Definition 3. *An incentive-compatible k-selection mechanism is α-optimal if*

$$\inf_{G\in\mathcal{G}} \frac{\mathbb{E}_{S\sim x_S(G)}[\sum_{i\in S} P(i, G)]}{\sum_{i\in S_k^*} P(i, G)} \ge \alpha.$$

For convenience, we can characterize an optimal group by defining a strict order of agents as follows.

Definition 4. *For a graph $G = (N, E) \in \mathcal{G}$, and agents i, $j \in N$, $i \ne j$, we say $i \succ j$ if either $P(i, G) > P(j, G)$ or $P(i, G) = P(j, G)$ with $i > j$.*

Let i_t^* be the agent with rank t such that $|\{j \mid j \succ i_t^*\}| = t - 1$, which must be unique since the order is strict. Then, we can order all the agents as the *ranking sequence* $i_1^* \succ i_2^* \succ \cdots \succ i_n^*$. Apparently, $\{i_1^*, \cdots, i_k^*\}$ is an optimal set for selecting k agents. Hence, for our strategic setting, we will pay attention to agents who can pretend to be the first k agents in the ranking sequence by hiding their out-edges.

Definition 5. *For a graph $G = (N, E)$, an agent i belongs to the k-influential set $S_k^{\mathsf{inf}\cdot}(G)$ if $|\{j \mid j \succ i\}| < k$ holds in the graph $G' = (N, E \setminus E_i)$.*

In this paper, we mainly focus on the cases for $k \in \{1, 2\}$. Hence, we use some observations about $S_1^{\mathsf{inf}\cdot}(G)$ and $S_2^{\mathsf{inf}\cdot}(G)$. To make it easy to follow, we present the observations about $S_1^{\mathsf{inf}\cdot}(G)$ below as preparation and present the observations of $S_2^{\mathsf{inf}\cdot}(G)$ in Section 4.2.

Observation 1 ([Zhang *et al.*, 2021])**.** *For any graph G, the set $S_1^{\mathsf{inf}\cdot}(G)$ can be written as $\{i_1, i_2, \cdots, i_m\}$, where $m \ge 1$, $i_1 = i_1^*$, and $i_{t+1} \in P(i_t) \setminus \{i_t\}$ for all $t < m$.*

Intuitively, the agent i_1^* who ranks the first is naturally in the 1-influential set. Furthermore, if there are more than one agent in the set, the agent who has a lower rank must be in the progeny of those who have higher ranks; otherwise, she will still have a lower rank when deleting her out-edges. Hence, in other words, we can find a path in G that passes through all agents in the set with the order of their ranks.

Observation 2 ([Zhang *et al.*, 2021])**.** *For any graph G, if the set $S_1^{\mathsf{inf}\cdot}(G) = \{i_1, i_2, \cdots, i_m\}$ has more than one agent, i.e., $m > 1$, then for any $1 < t \le m$, $2P(i_t) \ge P(i_1)$.*

Inferred from Observation 1, if there is another agent except for i_1^* is in the 1-influential set, she should hold at least half of i_1^*'s progeny to make herself rank the first after removing her out-edges.

3 Select One Agent

In this section, we present our result for only selecting $k = 1$ agent. Not only does it help us understand the proposed methods in the following section for $k = 2$, but also fills the gap between the existing mechanisms and the upper bound of approximation ratios for IC 1-selection mechanisms, which is $1/(1 + \ln 2)$ confirmed by Zhang *et al.* [2021].

Formally, our method can be viewed as a general variant of the modified Babichenko's mechanism [2020b].

β-logarithmic Mechanism (β-LM)

1. Given a network $G = (N, E)$, find the 1-influential set $S_1^{\mathsf{inf}\cdot}(G) = \{i_1, \dots, i_m\}$, where $i_t \succ i_{t+1}$ for all $1 \le t < m$.
2. Assign the probability of each agent to be selected as follows:

$$x_j = \begin{cases} \beta, & j = i_m \\ (1-\beta)\log_2 \frac{P(i_t)}{P(i_{t+1})}, & j = i_t, t \ne m \\ 0, & j \notin S_1^{\mathsf{inf}\cdot}(G). \end{cases}$$

The total probability of selecting one agent in β-LM is at most $\beta + (1-\beta)\log_2(P(i_1)/P(i_m)) \le 1$ by the fact in Observation 2. Hence, the probabilities assigned by the mechanism are valid as long as $0 \le \beta \le 1$. Then, we show the mechanism is IC when $\beta \ge 1/2$.

Theorem 1. *A β-logarithmic mechanism is IC if $\beta \ge 1/2$.*

Proof. For any graph $G \in \mathcal{G}$, let $S_1^{\mathsf{inf}\cdot}(G) = \{i_1, \dots, i_m\}$. Then, we consider three different types of agents.

1. For an agent $i \notin S_1^{\mathsf{inf}\cdot}(G)$, by definition, she can never pretend to be the agent with rank 1 by hiding her out-edges. Hence, she will always not belong to the 1-influential set and will have 0 probability to be chosen.
2. For an agent $i_t \in S_1^{\mathsf{inf}\cdot}(G)$ such that $t < m$, no matter how she hides her out-edges, i_{t+1} will always belong to the 1-influential set because $i_{t+1} \in P(i_t)$ (Observation 1) and her progeny cannot be decreased. Hence, the probability of i_t to be chosen will not change.
3. For the agent $i_m \in S_1^{\mathsf{inf}\cdot}(G)$, if she hides some of her out-edges, there may happen two cases. (i) If there is no agent in $P(i_m)$ occurs in the new 1-influential set, the probability of i_m to be chosen will remain β. (ii) If there exists at least one agent in $P(i_m)$ that occurs in the new 1-influential set, let i_1' be the first agent in the new set. Let i_{m+1} be the one with the highest rank after i_m in the new set. Then, the probability of i_m to be chosen will become $(1-\beta)\log_2(P(i_m)/P(i_{m+1})) \le (1-\beta)\log_2(P(i_1')/P(i_{m+1})) \le (1-\beta) \le \beta$ by the fact of $\beta \ge 1/2$ and Observation 2.

Taking all the above together, no agent can increase her probability to be chosen by hiding her out-edges. □

Now we can compute the approximation ratios of IC β-LMs, from which we can find that the optimal β-LM is also an optimal IC selection mechanism for $k = 1$.

Theorem 2. *An IC β-logarithmic mechanism ($1/2 \le \beta \le 1$) is $\left(\min\left\{\frac{1}{2}\left(\beta + \frac{1-\beta}{\ln 2}\right), \beta\right\}\right)$-optimal.*

Proof. For any graph $G \in \mathcal{G}$, let $S_1^{\mathsf{inf}\cdot}(G) = \{i_1, \dots, i_m\}$. If $m = 1$, then we have

$$\mathbb{E}_{i\sim x_i}[P(i)]/P(i_1^*) = x_{i_1}P(i_1)/P(i_1) = \beta.$$

If $m > 1$, then we have

$$\begin{aligned}
&\mathbb{E}_{i\sim x_i}[P(i)]/P(i_1^*) \\
&= \frac{1-\beta}{P(i_1^*)}\sum_{t=1}^{m-1} P(i_t)\log_2\frac{P(i_t)}{P(i_{t+1})} + \beta\frac{P(i_m)}{P(i_1^*)} \\
&= \frac{1-\beta}{\ln 2 \cdot P(i_1^*)}\sum_{t=1}^{m-1} P(i_t)\int_{P(i_{t+1})}^{P(i_t)}\frac{\mathrm{d}z}{z} + \beta\frac{P(i_m)}{P(i_1^*)} \\
&\ge \frac{1-\beta}{\ln 2 \cdot P(i_1^*)}\sum_{t=1}^{m-1}\int_{P(i_{t+1})}^{P(i_t)}\mathrm{d}z + \beta\frac{P(i_m)}{P(i_1^*)} \\
&= \frac{1-\beta}{\ln 2 \cdot P(i_1)}\sum_{t=1}^{m-1}(P(i_t) - P(i_{t+1})) + \beta\frac{P(i_m)}{P(i_1)} \\
&= \frac{1-\beta}{\ln 2} + \left(\beta - \frac{1-\beta}{\ln 2}\right)\frac{P(i_m)}{P(i_1)} \\
&\ge \frac{1}{2}\left(\beta + \frac{1-\beta}{\ln 2}\right).
\end{aligned}$$

Therefore, the mechanism is $\left(\min\left\{\frac{1}{2}\left(\beta + \frac{1-\beta}{\ln 2}\right), \beta\right\}\right)$-optimal. □

It is not hard to find out that when $\beta = 1/(1 + \ln 2)$, the value $\min\left\{\frac{1}{2}\left(\beta + \frac{1-\beta}{\ln 2}\right), \beta\right\}$ takes it maximum as $1/(1 + \ln 2)$, i.e., the optimal β-LM is $(1/(1 + \ln 2))$-LM, which is $(1/(1 + \ln 2))$-optimal. Recall that Zhang *et al.* [2021] has proved that no IC and fair[3] selection mechanism can be α-optimal with $\alpha > 1/(1 + \ln 2)$. Hence, we can infer the optimality of $(1/(1 + \ln 2))$-LM.

[3]Fairness is a quite weak property for single agent selection that only requires i_1^* has the same probability to be chosen if the 1-influential set and the structure formed by $P(i_1^*)$ remain the same. It is clear to see that β-LM satisfies this fairness.

Corollary 1. *There is no other IC and fair selection mechanism for $k = 1$ that can have a higher approximation ratio than* $(1/(1 + \ln 2))$*-LM.*

At the end of this section, we give a running example of $(1/(1 + \ln 2))$-LM.

Example 1. *Consider the network depicted in Figure 1, where $S_1^{\mathsf{inf.}}(G) = \{i_1, i_2, i_3, i_4\}$. For the last agent i_4 in the set, her selection probability x_{i_4} is $1/(1 + \ln 2) \approx 0.59$. For the agents i_3, i_2 and i_1, their selection probabilities are*

$$x_{i_3} = \frac{\ln 2}{1+\ln 2} \log_2 \frac{P(i_3)}{P(i_4)} = \frac{\ln 2}{1+\ln 2} \log_2 \frac{5}{4} \approx 0.13;$$
$$x_{i_2} = \frac{\ln 2}{1+\ln 2} \log_2 \frac{P(i_2)}{P(i_3)} = \frac{\ln 2}{1+\ln 2} \log_2 \frac{6}{5} \approx 0.11;$$
$$x_{i_1} = \frac{\ln 2}{1+\ln 2} \log_2 \frac{P(i_1)}{P(i_2)} = \frac{\ln 2}{1+\ln 2} \log_2 \frac{7}{6} \approx 0.09.$$

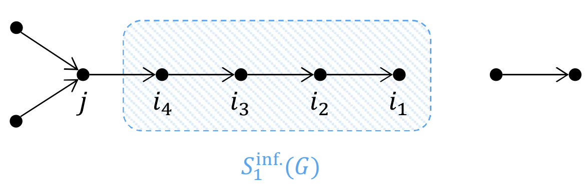

Figure 1: An example of the network, where the marked agents have the relationship as $i_1 \succ i_2 \succ i_3 \succ i_4 \succ j$. The 1-influential set is represented by a dashed border.

4 Select Two Agents

We start to consider selecting up to $k = 2$ agents as delegates. To select agents with larger progeny, one possible approach is to find the second delegate from the 1-influential set as well. However, this limits the performance of IC mechanisms.

4.1 Limitation of the 1-influential Set

The limitation of selecting agents from the 1-influential set for $k = 2$ mainly comes from the fact that there may be only a single agent in the set.

Theorem 3. *If an IC 2-selection mechanism only selects agents in the 1-influential set, then it cannot be α-optimal with $\alpha > 1/2$.*

Proof. Consider a two-star graph shown in Figure 2. Suppose $P(i_1) = P(j) = y$ and $i_1 > j$. Then the 1-influential set of this graph $S_1^{\mathsf{inf.}}(G)$ only contains i_1. Therefore, even if the mechanism can always select agent i_1 with probability 1, the approximation ratio in this graph is only $(1 \cdot y)/(y + y) = 1/2$. Hence, if only selecting agents in the 1-influential set, an IC 2-selection mechanism cannot achieve an approximation ratio of more than $1/2$. $\square$

We can also show that the limitation described in Theorem 3 is tight by providing the following mechanism.

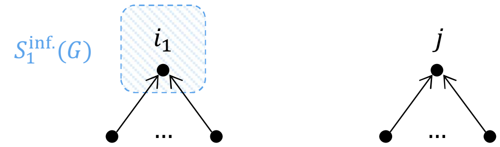

Figure 2: A two-star network where the hubs satisfy $i_1 \succ j$. Thus, the 1-influential set only contains i_1.

Least Deterministic Mechanism (LDM)

1. Given a network $G = (N, E)$, find the 1-influential set $S_1^{\mathsf{inf.}}(G) = \{i_1, \ldots, i_m\}$, where $i_t \succ i_{t+1}$ for all $1 \le t < m$.
2. Assign the probability of each agent to be selected as follows:

$$x_j = \begin{cases} 1, & j = i_m, \text{or } j = i_{m-1} \\ 0, & j = i_t, t < m-1, \text{or } j \notin S_1^{\mathsf{inf.}}(G). \end{cases}$$

Intuitively, LDM deterministically selects the last two agents in the 1-influential set or it selects the only agent in the set if $|S_1^{\mathsf{inf.}}(G)| = 1$.

Example 2. *We take the networks shown in Figure 1 and Figure 2 as running examples. In Figure 1, there are four agents, i_1, i_2, i_3, and i_4, in the 1-influential set. Hence, LDM deterministically selects the last two agents i_4 and i_3, i.e., $x_{i_4} = x_{i_3} = 1$. In Figure 2, there is a single agent i_1 in the 1-influential set. Hence, LDM deterministically selects the agent i_1 only.*

Now we prove that LDM is incentive-compatible and $1/2$-optimal as follows.

Theorem 4. *LDM is an IC 2-selection mechanism.*

Proof. For any graph $G \in \mathcal{G}$, suppose that $S_1^{\mathsf{inf.}}(G) = \{i_1, \ldots, i_m\}$. Then we consider the following cases.

1. For an agent $i \notin S_1^{\mathsf{inf.}}(G)$, same as the first point in the proof of Theorem 1, she always has no chance to be chosen by hiding her out-edges.
2. If $m \le 2$, the agents in $S_1^{\mathsf{inf.}}(G)$ will be deterministically selected. Hence, they have no incentive to hide their out-edges.
3. If $m > 2$, first, for agents i_m and i_{m-1} who will be deterministically selected, they have no incentive to hide their out-edges. Then, for any agent $i_t \in S_1^{\mathsf{inf.}}(G)$ with $i < m - 1$, no matter how she hides her out-edges, i_m and i_{m-1} will always belong to the 1-influential set because $\{i_m, i_{m-1}\} \subseteq P(i_t)$ (Observation 1) and their progeny cannot be decreased. Hence, the probability of i_t to be chosen will remain to be 0.

Taking all the above together, no agent can increase her probability to be chosen by hiding her out-edges. Therefore, the mechanism is IC. $\square$

Theorem 5. *LDM is* $1/2$*-optimal.*

Proof. For any graph $G \in \mathcal{G}$, if $|S_1^{\mathsf{inf.}}(G)| = 1$, then LDM deterministically selects the agent i_1^*. Therefore, in this case, we have

$$\frac{\mathbb{E}_S[\sum_{i\in S} P(i)]}{\sum_{i\in S_2^*} P(i)} = \frac{P(i_1^*)}{P(i_1^*)+P(i_2^*)} \geq \frac{P(i_1^*)}{P(i_1^*)+P(i_1^*)} = \frac{1}{2}.$$

If $|S_1^{\mathsf{inf.}}(G)| \geq 2$, then LDM deterministically selects the agent i_m and i_{m-1}. By Observation 2, in this case, we have

$$\begin{aligned}\frac{\mathbb{E}_S[\sum_{i\in S} P(i)]}{\sum_{i\in S_2^*} P(i)} &= \frac{P(i_m)+P(i_{m-1})}{P(i_1^*)+P(i_2^*)}\\ &\geq \frac{P(i_m)+P(i_{m-1})}{2P(i_1)}\\ &= \frac{1}{2}\left(\frac{P(i_m)}{P(i_1)} + \frac{P(i_{m-1})}{P(i_1)}\right)\\ &\geq \frac{1}{2}\left(\frac{1}{2}+\frac{1}{2}\right) = \frac{1}{2}.\end{aligned}$$

Therefore, LDM is $1/2$-optimal. □

If we consider general IC 2-selection mechanisms, we may have a higher upper bound of the approximate ratio. This suggests the limitation of only selecting agents from the 1-influential set when we target up to two delegates.

Theorem 6. *There is no IC 2-selection mechanism that can be* α*-optimal with* $\alpha > 23/27$*.*

Proof. Consider three networks with four agents shown in Figure 3. Applying a generic IC 2-selection mechanism on these graphs, suppose the probabilities of each agent i being chosen in the three graphs are $x_i^{(a)}$, $x_i^{(b)}$ and $x_i^{(c)}$. Notice that network (b) can be obtained by agent 2 or 4 in network (a) hiding their out-edges (corresponding to agent 2 or 1 in network (b)), while network (c) can be obtained by agent 3 hiding her out-edge (corresponding to agent 1 or 3 in network (c)). Since the selection mechanism is IC, we have the following constraints:

$$x_2^{(b)} \leq x_2^{(a)}, \quad x_1^{(b)} \leq x_4^{(a)}; \tag{1}$$

$$x_3^{(c)} \leq x_3^{(a)}, \quad x_1^{(c)} \leq x_3^{(a)}. \tag{2}$$

Moreover, the mechanism selects at most 2 agents. Hence,

$$\sum_{i=1}^{4} x_i^{(a)} \leq 2, \quad \sum_{i=1}^{4} x_i^{(b)} \leq 2, \quad \sum_{i=1}^{4} x_i^{(c)} \leq 2. \tag{3}$$

The approximation ratio of the mechanism must be no more than the least ratio in these three graphs, i.e.,

$$\alpha \leq \min\left\{\frac{4x_1^{(a)}+3x_2^{(a)}+2x_3^{(a)}+x_4^{(a)}}{7}, \frac{x_1^{(b)}+3x_2^{(b)}+2x_3^{(b)}+x_4^{(b)}}{5}, \frac{2x_1^{(c)}+x_2^{(c)}+2x_3^{(c)}+x_4^{(c)}}{4}\right\}$$

With constraints (1) - (3), we can calculate the highest value of the minimum. Therefore, we have $\alpha \leq 23/27$ and the equations holds when

$$x_1^{(a)} = 2/3, x_2^{(a)} = 17/27, x_3^{(a)} = 19/27, x_4^{(a)} = 0;$$
$$x_1^{(b)} = 0, x_2^{(b)} = 17/27, x_3^{(b)} = 1, x_4^{(b)} = 10/27;$$
$$x_1^{(c)} = 19/27, x_2^{(c)} = 16/27, x_3^{(c)} = 19/27, x_4^{(c)} = 0.$$

□

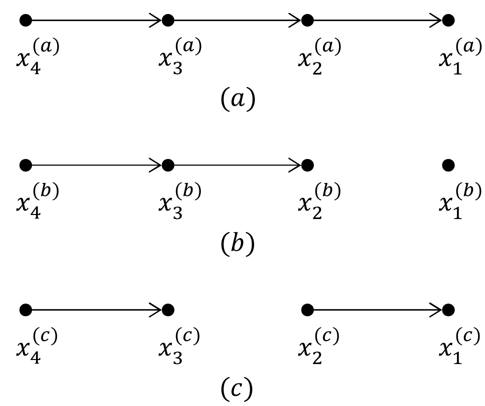

Figure 3: Three networks with four agents where (b) and (c) can be obtained by one of the agents in (a) hiding her out-edge. The probabilities of each agent being chosen by a generic IC 2-selection mechanism are attached beside the node.

4.2 Utilizing the 2-influential Set

To break through the limitation of the 1-influential set, one natural idea is to consider the 2-influential set. We first characterize the set by following observations.

Observation 3. *For any graph G,* $S_1^{\mathsf{inf.}}(G) \subseteq S_2^{\mathsf{inf.}}(G)$*.*

Observation 4. *For any graph G,* $\{i_1^*, i_2^*\} \subseteq S_2^{\mathsf{inf.}}(G)$*.*

According to Definition 5, agents who can pretend to be the first in the ranking sequence are definitely in the 2-influential set. The agents i_1^* and i_2^* who rank first and second are also naturally in the 2-influential set. Based on the relationship between i_1^* and i_2^*, the 2-influential set will have different forms.

Observation 5. *For any graph G, if* $i_2^* \in P(i_1^*)$*, the set* $S_2^{\mathsf{inf.}}(G)$ *can be written as* $\{i_1, i_2, \cdots, i_m\}$*, where* $i_1 = i_1^*$*,* $i_2 = i_2^*$*, and* $i_{t+1} \in P(i_t)$ *for all* $t < m$*.*

Proof. When there is another agent i_3, except for i_1^* and i_2^*, is in the 2-influential set, she must have the ability to decrease i_1^* or i_2^*'s progeny by hiding her out-edges, i.e., at least one of $i_3 \in P(i_1^*)$ and $i_3 \in P(i_2^*)$ is satisfied. If $i_2^* \in P(i_1^*)$, we will show i_3 must belong to $P(i_2^*)$ by contradiction as follows.

If $i_3 \notin P(i_2^*)$, then she cannot decrease i_2^*'s progeny by hiding her out-edges. Since $i_2^* \notin P(i_3)$, $i_2^* \in P(i_1^*)$ is always satisfied no matter how i_3 hides her out-edges. Hence, in order to rank first or second after hiding out-edges, $i_3 \succ i_2^*$ must be satisfied, which contradicts that i_2^* is with rank 2. Therefore, $i_3 \in P(i_2^*)$.

Similarly, when there is another agent $i_4 \prec i_3$ in the 2-influential set, she must belong to $P(i_3)$, and this pattern extends to subsequent agents in the 2-influential set. □

Observation 6. *For any graph G, if $i_2^* \notin P(i_1^*)$, the set $S_2^{\text{inf.}}(G)$ can be written as $\{i_1, i_2, \cdots, i_m\}$, where $i_1 = i_1^*$, $i_2 = i_2^*$, and for others, at least one of the below is true:*

- *$i_3 \in P(i_1)$, $i_{t+1} \in P(i_t)$ for all $3 \le t < m$;*
- *$i_3 \in P(i_2)$, $i_{t+1} \in P(i_t)$ for all $3 \le t < m$.*

Proof. When there is another agent i_3, except for i_1^* and i_2^*, is in the 2-influential set, at least one of $i_3 \in P(i_1^*)$ and $i_3 \in P(i_2^*)$ is satisfied.

We first assume $i_3 \in P(i_1^*)$. Then, when there is another agent $i_4 \prec i_3$ in the 2-influential set, we show that $i_4 \in P(i_3)$ by contradiction as follows. If $i_4 \notin P(i_3)$, she cannot decrease i_3's progeny by hiding her out-edges. Since $i_3 \notin P(i_4)$, $i_3 \in P(i_1^*)$ is always satisfied no matter how i_4 hides her out-edges. Hence, $i_4 \succ i_3$ must be satisfied to make i_4 be in the 2-influential set, which makes a contradiction. Therefore, $i_4 \in P(i_3)$. Similarly, when there is another agent $i_5 \prec i_4$ in the 2-influential set, she must belong to $P(i_4)$, and this pattern extends to subsequent agents in the 2-influential set. The same results can also be obtained similarly if $i_3 \in P(i_2^*)$. □

From the above observations, we can see that the structure of the 2-influential set is much more complex than that of the 1-influential set. Furthermore, the progeny of the last agent in the 2-influential set might be much smaller than $P(i_1^*)$ since $P(i_2^*)$ might be small. These are the main difficulties to utilize the 2-influential set. Extending the idea of LDM and β-LM, we propose the following mechanism.

Logarithm After Least Deterministic (LALD)

1. Given a network $G = (N, E)$, find the 1-influential set $S_1^{\text{inf.}}(G)$ and the 2-influential set $S_2^{\text{inf.}}(G)$.
2. If $S_2^{\text{inf.}}(G) \setminus S_1^{\text{inf.}}(G) = \emptyset$, then $S_1^{\text{inf.}}(G) = S_2^{\text{inf.}}(G) = \{i_1, i_2 \ldots, i_m\}$, where $i_t \succ i_{t+1}$ for all $1 \le t < m$. Then, assign the probability of each agent to be selected as follows:

$$x_j = \begin{cases} 1, & j = i_m \\ \frac{1}{1+\ln 2}, & j = i_{m-1} \\ \frac{\ln 2}{1+\ln 2} \log_2 \frac{P(i_t)}{P(i_{t+1})}, & j = i_t, t < m-1 \\ 0, & j \notin S_2^{\text{inf.}}(G). \end{cases}$$

3. If $S_2^{\text{inf.}}(G) \setminus S_1^{\text{inf.}}(G) \neq \emptyset$, suppose $S_2^{\text{inf.}}(G) = \{i_1, \ldots, i_m\}$ where $i_t \succ i_{t+1}$ for all $1 \le t < m$. First, deterministically select the agent i_m, i.e., $x_{i_m} = 1$. Then, select the second agent by applying $(1/(1+\ln 2))$-LM on G.

Intuitively, LALD first deterministically selects the last agent in the 2-influential set. Then, it uses the same probability distribution as $(1/(1+\ln 2))$-LM to select another agent among the remaining agents in the 1-influential set.

Example 3. *We take the networks shown in Figure 1 and Figure 4 as running examples.*

In Figure 1, suppose $j < i_2$. Then, $S_2^{\text{inf.}}(G) = S_1^{\text{inf.}}(G) = \{i_1, i_2, i_3, i_4\}$. Hence, LALD first deterministically selects agent i_4, i.e., $x_{i_4} = 1$. For the remaining agents in $S_1^{\text{inf.}}(G)$, LALD assigns the probabilities as $x_{i_3} = 1/(1+\ln 2) \approx 0.59$, $x_{i_2} = \ln 2/(1+\ln 2)\log_2(P(i_2)/P(i_3)) \approx 0.11$, and $x_{i_1} = \ln 2/(1+\ln 2)\log_2(P(i_1)/P(i_2)) \approx 0.09$.

In Figure 4, suppose $i_2 > i_3$ and $i_4 > j$. Then, $S_1^{\text{inf.}}(G) = \{i_1\}$ and $S_2^{\text{inf.}}(G) = \{i_1, i_2, i_3, i_4\}$. Hence, LALD first deterministically selects agent i_4, i.e., $x_{i_4} = 1$. LALD then runs $(1/(1+\ln 2))$-LM, which assigns the probabilities among $S_1^{\text{inf.}}(G)$. Here, it assigns that $x_{i_1} = 1/(1+\ln 2) \approx 0.59$.

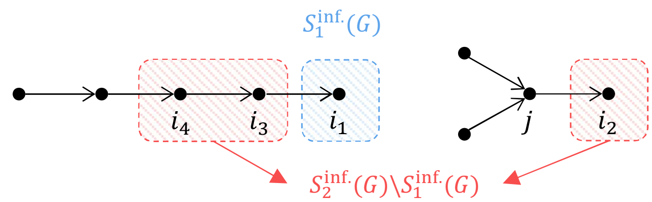

Figure 4: An example of the network, where the marked agents have the relationship as $i_1 \succ i_2 \succ i_3 \succ i_4 \succ j$. The 1-influential set and the 2-influential set are represented by dashed borders.

Theorem 7. *LALD is an IC 2-selection mechanism.*

Proof. For any graph $G \in \mathcal{G}$, We consider three different types of agents.

1. For an agent $i \notin S_2^{\text{inf.}}(G)$, by definition, she can never be in the set by hiding her out-edges. Hence, she will always have 0 probability to be chosen.
2. For an agent $i \in S_2^{\text{inf.}}(G) \setminus S_1^{\text{inf.}}(G)$ when $S_2^{\text{inf.}}(G) \setminus S_1^{\text{inf.}}(G) \neq \emptyset$, there are two cases. (i) If i is the last agent in the 2-influential set, her probability to be chosen is 1. Hence, she has no incentive to manipulate. (ii) If i is not the last agent, no matter how she hides her out-edges, both the last agent and herself are still in the set $S_2^{\text{inf.}}(G) \setminus S_1^{\text{inf.}}(G)$. Hence, her probability to be chosen remains 0.
3. For an agent $i \in S_1^{\text{inf.}}(G)$, suppose $S_1^{\text{inf.}}(G) = \{i_1, \ldots, i_q\}$ with $i_t \succ i_{t+1}$ for all $1 \le t < q$. There are three cases. (i) If i is the last agent i_q in the 1-influential set, she has no incentive to manipulate when $S_2^{\text{inf.}}(G) \setminus S_1^{\text{inf.}}(G) = \emptyset$ since $x_i = 1$. When $S_2^{\text{inf.}}(G) \setminus S_1^{\text{inf.}}(G) \neq \emptyset$, no matter how i hides her out-edges, agents in the set $S_2^{\text{inf.}}(G) \setminus S_1^{\text{inf.}}(G)$ will still be in the 2-influential set (even may be in the 1-influential set). After i hides some out-edges, if $S_2^{\text{inf.}}(G') \setminus S_1^{\text{inf.}}(G') \neq \emptyset$, then i cannot have higher probability by Theorem 1; if $S_2^{\text{inf.}}(G') \setminus S_1^{\text{inf.}}(G') = \emptyset$, then i can have at most $1/(1+\ln 2)$ probability since she will no longer be the

last agent in $S_1^{\mathsf{inf}\cdot}(G')$, which equals to her original probability. (ii) If $i = i_{q-1}$, she has no incentive to manipulate when $S_2^{\mathsf{inf}\cdot}(G) \setminus S_1^{\mathsf{inf}\cdot}(G) = \emptyset$. It is because i_q will always be in the 1-influential set no matter how i hides her out-edges, which makes increasing her probability to be chosen impossible. When $S_2^{\mathsf{inf}\cdot}(G) \setminus S_1^{\mathsf{inf}\cdot}(G) \neq \emptyset$, it is almost the same condition as that for $i = i_q$. Hence, she cannot increase her probability by manipulation, either. (iii) If $i = i_t$ with $t < q - 1$, no matter how she hides her out-edges, i_q and i_{q-1} are always in the 1-influential set. Hence, i's probability to be chosen will not change.

Taking all the above together, we can conclude that the mechanism is IC. □

Theorem 8. *LALD is* $\frac{3+\ln 2}{4(1+\ln 2)}$*-optimal.*

Proof. Suppose $S_2^{\mathsf{inf}\cdot}(G) = \{i_1, i_2, \cdots, i_m\}$ for any graph $G \in \mathcal{G}$,. There are two different cases that need consideration.

1. If $S_2^{\mathsf{inf}\cdot}(G) \setminus S_1^{\mathsf{inf}\cdot}(G) = \emptyset$, then according to Observation 1 and Theorem 2, we have
$$\begin{aligned}\frac{\mathbb{E}_S[\sum_{i\in S} P(i)]}{\sum_{i\in S_2^*} P(i)} &\geq \frac{P(i_m) + \frac{1}{1+\ln 2}P(i_1^*)}{P(i_1^*) + P(i_2^*)} \\ &\geq \frac{\left(\frac{1}{2} + \frac{1}{1+\ln 2}\right) P(i_1)}{P(i_1) + P(i_2)} \\ &\geq \frac{\frac{1}{2} + \frac{1}{1+\ln 2}}{2} = \frac{3 + \ln 2}{4(1 + \ln 2)}.\end{aligned}$$

2. If $S_2^{\mathsf{inf}\cdot}(G) \setminus S_1^{\mathsf{inf}\cdot}(G) \neq \emptyset$, we first consider the progeny $P(i_m)$. When $i_2^* \in P(i_1^*)$, the structure of the 2-influential set is characterized by Observation 5. Then after deleting i_m's out-edges, $i_1 \succ i_2$ holds. Hence, $P(i_m) \geq P(i_2) - P(i_m)$ which implies $2P(i_m) \geq P(i_2)$. When $i_2^* \notin P(i_1^*)$, the structure of the 2-influential set is characterized by Observation 6. Hence, at least one of $2P(i_m) \geq P(i_1)$ and $2P(i_m) \geq P(i_2)$ is satisfied, which all imply that $2P(i_m) \geq P(i_2)$. Therefore, let $P(i_2)/P(i_1) = \rho$ and we have
$$\begin{aligned}\frac{\mathbb{E}_S[\sum_{i\in S} P(i)]}{\sum_{i\in S_2^*} P(i)} &\geq \frac{P(i_m) + \frac{1}{1+\ln 2}P(i_1^*)}{P(i_1^*) + P(i_2^*)} \\ &\geq \frac{\frac{1}{2}P(i_2) + \frac{1}{1+\ln 2}P(i_1)}{P(i_1) + P(i_2)} \\ &= \frac{\rho/2 + 1/(1 + \ln 2)}{1 + \rho} \\ &= \frac{\frac{1}{2}(1 + \rho) + \frac{1-\ln 2}{2(1+\ln 2)}}{1 + \rho} \\ &= \frac{1}{2} + \frac{1 - \ln 2}{2(1 + \ln 2)} \cdot \frac{1}{1 + \rho} \\ &\geq \frac{1}{2} + \frac{1 - \ln 2}{4(1 + \ln 2)} = \frac{3 + \ln 2}{4(1 + \ln 2)}.\end{aligned}$$

Therefore, we can conclude that the mechanism is $\frac{3+\ln 2}{4(1+\ln 2)}$-optimal. □

5 Discussion

In this paper, we investigate the incentive-compatible selection mechanisms for one or two influentials, where an agent's influential power is defined by her progeny. The goal is to select agents with progeny as large as possible and to prevent them from hiding their out-edges at the same time. Based on the idea of assigning possibilities of being selected to those agents who can pretend to be the one with the largest or second largest progeny, we first propose the $1/(1 + \ln 2)$-LM mechanism for selecting one agent, which is optimal among all IC and fair single-agent selections. We then propose the LALD mechanism for selecting up to two influentials, which has an approximation ratio of $(3 + \ln 2)/(4(1 + \ln 2))$ (≈ 0.54). To the best of our knowledge, this is the first work to select more than one agent for progeny maximization. There are several interesting future directions worth investigating, and we provide some brief discussions here.

One direction is to narrow the gap between the current lower bound (given by our proposed mechanism) and the current upper bound ($23/27$ as we proved) of the approximation ratio for an optimal IC 2-selection mechanism. For the side of upper bounds, notice that our provided upper bound does not require additional properties like fairness defined in [Zhang *et al.*, 2021]. This is because the fairness for selecting a single agent does not apply in selecting multiple agents (e.g., in LALD, the probability of choosing i_1^* may be also related to the structure of the 2-influential set). If we extend the definition of fairness to k-fairness like

Definition 6 (sketch)**.** *i_1^* (or also with i_2^* to i_k^*) has the same probability to be chosen when the k-influential set and the structure formed by $P(i_1^*)$ (or also with $P(i_2^*)$ to $P(i_k^*)$) remain the same.*

Then, we can observe that k-fairness will become weaker when k becomes larger. Zhang *et al.* [2021] conjectured that dropping (1-)fairness will not affect the upper bound they characterized. If it can be proven to be true, then we can also draw a corollary that introducing k-fairness will not affect the upper bounds of approximation ratios for IC k-selection mechanisms. For the side of improving lower bounds, one may consider to utilize more agents in the 2-influential set but not in the 1-influential set. The main difficulty here is these agents may have too small progeny when $P(i_2^*) \ll P(i_1^*)$.

The other direction is to extend the mechanisms for selecting more agents ($k \geq 3$). Similar to the case of selecting two agents, only selecting agents in the k'-influential set with $k' < k$ may limit the performance. A natural idea is to select k agents in the k-influential set. The main difficulty here is that the structure of the k-influential set will become more and more complex when k becomes larger. Intuitively, the structure of the k-influential set depends on the relationships among agents $i_1^*, \ldots, i_k^*$. The number of different cases of the structure will grow exponentially with k, which is roughly $2^{O(n^2)}$. A possible way to handle this challenge may be recursively considering the influential set with lower k.

Finally, in terms of other applications, such as recruiting agents to promote some advertisements, designing selection mechanisms to maximize the expected cardinality of the union of progeny is also a promising future direction.

Acknowledgements

This work is supported by Science and Technology Commission of Shanghai Municipality (No. 22ZR1442200 and No. 23010503000), and Shanghai Frontiers Science Center of Human-centered Artificial Intelligence (ShangHAI).

References

[Alon *et al.*, 2011] Noga Alon, Felix Fischer, Ariel Procaccia, and Moshe Tennenholtz. Sum of us: Strategyproof selection from the selectors. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 101–110, 2011.

[Babichenko *et al.*, 2018] Yakov Babichenko, Oren Dean, and Moshe Tennenholtz. Incentive-compatible diffusion. In *Proceedings of the 2018 World Wide Web Conference*, pages 1379–1388, 2018.

[Babichenko *et al.*, 2020a] Yakov Babichenko, Oren Dean, and Moshe Tennenholtz. Incentive-compatible classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 7055–7062, 2020.

[Babichenko *et al.*, 2020b] Yakov Babichenko, Oren Dean, and Moshe Tennenholtz. Incentive-compatible selection mechanisms for forests. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 111–131, 2020.

[Bjelde *et al.*, 2017] Antje Bjelde, Felix Fischer, and Max Klimm. Impartial selection and the power of up to two choices. *ACM Transactions on Economics and Computation (TEAC)*, 5(4):1–20, 2017.

[Bousquet *et al.*, 2014] Nicolas Bousquet, Sergey Norin, and Adrian Vetta. A near-optimal mechanism for impartial selection. In *International Conference on Web and Internet Economics*, pages 133–146. Springer, 2014.

[Caragiannis *et al.*, 2021] Ioannis Caragiannis, George Christodoulou, and Nicos Protopapas. Impartial selection with prior information. *arXiv preprint arXiv:2102.09002*, 2021.

[Caragiannis *et al.*, 2022] Ioannis Caragiannis, George Christodoulou, and Nicos Protopapas. Impartial selection with additive approximation guarantees. *Theory of Computing Systems*, pages 1–22, 2022.

[Cembrano *et al.*, 2022] Javier Cembrano, Felix A. Fischer, David Hannon, and Max Klimm. Impartial selection with additive guarantees via iterated deletion. In *EC '22: The 23rd ACM Conference on Economics and Computation, Boulder, CO, USA, July 11 - 15, 2022*, pages 1104–1105. ACM, 2022.

[Fischer and Klimm, 2014] Felix Fischer and Max Klimm. Optimal impartial selection. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 803–820, 2014.

[Ghalme *et al.*, 2018] Ganesh Ghalme, Sujit Gujar, Amleshwar Kumar, Shweta Jain, and Y Narahari. Design of coalition resistant credit score functions for online discussion forums. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pages 95–103, 2018.

[Holzman and Moulin, 2013] Ron Holzman and Hervé Moulin. Impartial nominations for a prize. *Econometrica*, 81(1):173–196, 2013.

[Kahng *et al.*, 2018] Anson Kahng, Yasmine Kotturi, Chinmay Kulkarni, David Kurokawa, and Ariel Procaccia. Ranking wily people who rank each other. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

[Kotturi *et al.*, 2020] Yasmine Kotturi, Anson Kahng, Ariel Procaccia, and Chinmay Kulkarni. Hirepeer: Impartial peer-assessed hiring at scale in expert crowdsourcing markets. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 2577–2584, 2020.

[Kurokawa *et al.*, 2015] David Kurokawa, Omer Lev, Jamie Morgenstern, and Ariel D Procaccia. Impartial peer review. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

[Mackenzie, 2015] Andrew Mackenzie. Symmetry and impartial lotteries. *Games and Economic Behavior*, 94:15–28, 2015.

[Mattei *et al.*, 2020] Nicholas Mattei, Paolo Turrini, and Stanislav Zhydkov. Peernomination: Relaxing exactness for increased accuracy in peer selection. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*, pages 393–399. ijcai.org, 2020.

[Olckers and Walsh, 2022] Matthew Olckers and Toby Walsh. Manipulation and peer mechanisms: A survey. *arXiv preprint arXiv:2210.01984*, 2022.

[Page *et al.*, 1999] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999.

[Wang *et al.*, 2018] Yufeng Wang, Hui Fang, Chonghu Cheng, and Qun Jin. Tsp: Truthful grading-based strategyproof peer selection for moocs. In *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, pages 679–684. IEEE, 2018.

[Zhang *et al.*, 2021] Xiuzhen Zhang, Yao Zhang, and Dengji Zhao. Incentive compatible mechanism for influential agent selection. In *International Symposium on Algorithmic Game Theory*, pages 79–93. Springer, 2021.