Journal of Data Intelligence, Vol. 1, No. 1 (2020) 018–035

INFORMATION DIFFUSION ENHANCED BY MULTI-TASK PEER PREDICTION

KENSUKE ITO
The Graduate School of Interdisciplinary Information Studies,
The University of Tokyo, 7-3-1, Hongo
Bunkyo-ward, Tokyo, 113-0023, Japan
k-ito@g.ecc.u-tokyo.ac.jp

SHOHEI OHSAWA
The Interfaculty Initiative in Information Studies,
The University of Tokyo, 7-3-1, Hongo
Bunkyo-ward, Tokyo, 113-0023, Japan
s.ohsawa@iii.u-tokyo.ac.jp

HIDEYUKI TANAKA
The Graduate School of Interdisciplinary Information Studies,
The Interfaculty Initiative in Information Studies,
The University of Tokyo, 7-3-1, Hongo
Bunkyo-ward, Tokyo, 113-0023, Japan
tanaka@iii.u-tokyo.ac.jp

Our study aims to strengthen truthfulness of the two-path mechanism: an information diffusion algorithm to find an influential node in non-cooperative directed acyclic graphs (DAGs). This subject is important because the two-path mechanism ensures only weak truthfulness (i.e., nodes are indifferent between reporting true or false out-edges), which restricts node selection accuracy. To enhance the mechanism, we employed an additional reward layer based on a multi-task peer prediction, where an informative equilibrium provides strictly higher rewards than any other equilibrium in virtually all cases (strong truthfulness). Rewards, which are derived from a comparison of each report, encourage a node to report true out-edges without affecting its own probability of being selected by the original two-path mechanism. We have also experimentally confirmed that our proposed *strongly truthful two-path mechanism* can sufficiently elicit true out-edges from each node.

Keywords: peer prediction, information diffusion, mechanism design, directed acyclic graph

1. Introduction

When considering information-diffusion mechanisms used to find an influential node in a given non-cooperative network, *truthfulness* (also called *strategy-proofness* or *incentive compatibility*) is an important condition that intuitively represents the situation where no agent can obtain a higher utility by any possible strategy deviating from the agent's true preferences (e.g., [1]). An example to highlight its importance would be the selection from a directed network on intellectual contents, such as webpages and patents. If the adopted selection mechanisms[a] (e.g., PageRank [2], HITS [3]) were common knowledge for participants of the network, then a

[a]See definition 1.

selfish node (webpage, patent) would deliberately add or remove its out-edges (link, citation) to increase the probability of being selected by the mechanism. In other words, the nodes seeking to be influential might strategically misreport their own out-edges as long as the selection mechanism is not truthful.

The *two-path mechanism* [4] achieves truthfulness in information diffusion on directed acyclic graphs (DAGs). This mechanism specifically regards an influential node as the first intersection of two independent random paths drawn by letting each node report its own out-edges. This is particularly noteworthy because a node can no longer manipulate its own probability of becoming an intersection at the point it reports its out-edges. Such simple but effective modeling brought truthfulness to influential node selection in non-cooperative DAGs.[b] A remaining concern here is that the two-path mechanism ensures only *weak truthfulness*, which means that nodes are indifferent between reporting true or false out-edges. The accuracy of node selection would be restricted under this weak condition, where even rational nodes do not necessarily report true out-edges.

Accordingly, the aim of this study is to strengthen the truthfulness of the two-path mechanism, thereby improving its accuracy. To achieve this objective, we employed an additional reward layer based on the *peer-prediction* method, which is a mechanism used to elicit informative truthful reports for problems with no ground truth (e.g., customer review in Amazon, peer review of academic papers) by comparing them with other reports associated with the same task.

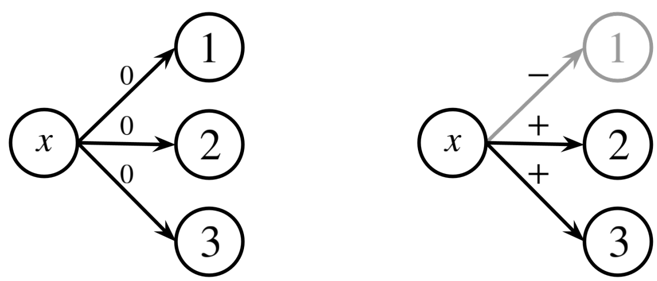


Fig. 1. Our model intends to elicit true DAG structure by giving positive (negative) rewards for truthful (false) reports on its edges.

We can present this approach using a simple example in Fig. 1, assuming that true out-edges for a node x are $(x, 2)$ and $(x, 3)$. Whereas x can report any out-edge to the powerset of $\{1, 2, 3\}$ without the reward layer, x would be induced to report truthfully if peer-prediction method could provide rewards (penalties) for the report on true (false) out-edges. Our model is an enhanced two-path mechanism in which reporting itself generates new rewards, and their expected amount is maximized if nodes report true edge structures.

This merger has significance from a peer-prediction perspective. Peer-prediction method, which assumes that a group of agents evaluates one independent task,[c] would be difficult to apply simply to an interdependent tasks such as network structure, where a new edge (a task to be verified) might affect the rewards of all nodes (agents) indirectly across group

[b] Babichenko et al. [4] mentions that two-path mechanism is also useful for directed cyclic graphs (DCGs) like Twitter's social network, by assigning random order to each node.

[c] See Fig. 3 as an example.

Table 1. Our model for overall diffusion ensures strong truthfulness.

	Weakly truthful	Strongly truthful
One-step diffusion (voting, reviewing)	Impartial selection mechanism e.g., [11, 12]	Multi-task peer prediction e.g., [5, 6]
Overall diffusion (information diffusion)	Two-path mechanism [4]	Our model

boundaries. Although peer-prediction works in the case of eliciting only network structure, each report would be biased if the elicited network were used as a criteria to select influential nodes (interdependent rewards). Consequently, to avoid the complicated mechanism designs, it would be practical for such a network to first secure weak truthfulness (i.e., an environment for impartial reporting) by the two-path mechanism and then introduce peer prediction.

Among many preceding peer-prediction methods, we adopt the mechanism presented by Dasgupta and Ghosh [5] for the following three reasons. First and foremost, this mechanism satisfies *strong truthfulness* [6], where an equilibrium by truthful reports has the highest rewards among any other realistic equilibria. Although many peer-prediction methods have multiple equilibria with homogeneous rewards, [5] achieved strong truthfulness for the first time because of its *multi-task* approach, which allocates several tasks to one agent and computes rewards based on various reports. Given the objective of our research, the addition of peer prediction is expected to satisfy a stronger condition than weak truthfulness to the greatest possible degree. Secondly, this mechanism, which only requires the report of assigned tasks, is *minimal* [6], thereby maintaining the original simplicity of the two-path mechanism.[d] In our proposed mechanism, a node only needs to additionally assess its own in-edges with binary signals. Thirdly, unlike some earlier mechanisms (e.g., [9, 10]), this mechanism is *detail-free* [6] in that the designer requires no knowledge about the probability distribution of model components. This condition is important for the two-path mechanism, originally designed to search an unobservable network structure.

The remainder of this paper is organized as follows. After introducing related works, we present our model which is referred to as the *strongly truthful two-path mechanism* in Section 2, along with the outline of its component elements [4, 5]. Section 3 examines the practical utility of our proposition through simulation by both synthetic and real-world DAG data. Conclusions and future works are presented in Section 4.

1.1. *Related Work*

Work related to our approach can be generally organized into an impartial selection mechanism as a background of the two-path mechanism combined with the peer-prediction method. Moreover, the classification presented in Table 1 is expected to help clarify the finding in our work.

1.1.1. *Impartial selection mechanism*

[d]In other peer-prediction mechanisms, agents must do another type of reporting, such as the report of their belief prior to signal observation [7] and prediction of other agents' reports [8].

Previous representative studies of impartial selection mechanisms include those by Alon et al. [11] and Holzman and Moulin [12], which address the difficulty of strategic manipulation when a group of peers selects one (or more) winner using reciprocal voting. They proposed weakly truthful mechanisms[e], whereby no agent can improve its own probability of being selected by strategic voting, while formulating the problem as a directed graph, where nodes and edges denote agents and voting directions, respectively. This research issue has been developing into several themes, such as an analysis of optimal group sizes and winners [13] and an attempt to generalize binary voting to quality evaluations [14]. The two-path mechanism [4] is the first extension of these selection mechanisms to use information diffusion in non-cooperative networks. Actually, [4] described this novelty as a shift from *one-step diffusion* to *overall diffusion*. The former merely confirms the in-degree of each node as the number of advocates, whereas the latter (e.g., [2, 15]) uses the entire network structure to measure the influence of each node.

Following an earlier contribution by [4], our work further enhances the robustness of this context from weakly truthful to strongly truthful by leveraging multi-task peer prediction, as presented in Table 1 above.

1.1.2. *Peer-prediction method*

Peer prediction was first introduced by Miller et al. [9] as an application of the proper scoring rule [16] and game theory. To model the problem of eliciting private information, [9] allows each agent to report probabilistic but correlated signals associated with the tasks. The method calculates rewards (scores) based on how much the report will affect another reference report on the same task. A common problem in [9] and subsequent models, as examined explicitly in a report by [17], is that the mechanism has multiple Nash equilibria, including *uninformative* ones where reports are independent of the observed signals[f] (e.g., always report the same signals or random signals to avoid the effort of observation). In addition to increasing the number of reference reports [10], Dasgupta and Ghosh [5] proposed a more simplified *multi-task* mechanism as a solution to this problem. This mechanism realizes strong truthfulness to other uninformative equilibria under the assumption of positively correlated binary signals. More recent studies generalize multi-task peer prediction from binary to multiple signals [6] and computationally simulate its convergence to an equilibrium [18]. See the book [19], for more comprehensive review on peer-prediction method and other information elicitation models.

To the best of our knowledge, we use such multi-task peer prediction for information diffusion in non-cooperative networks for the first time. A paper by Mohite and Narahari [20] is the only report in the relevant literature that describes the use of a peer-prediction-like mechanism in a similar field of research. However, their model, with a single-task mechanism, is not strongly truthful and deals with an independent cascade [15], which is a different influence measurement from what is evaluated in this study.

2. Model

Before presenting a detailed description, we first present basic settings and definitions of the

[e] Weak truthfulness is often designated as *impartiality* in this research field.
[f] Uninformative equilibria are designated as a *blind agreements* in [5].

equilibrium concepts used in the development of our model.

Definition 1: A *selection mechanism* $\mathcal{M}$ is a function that gives for every $G(V, E)$ a probability distribution on $V \cup \{\emptyset\}$.

Our work is a selection mechanism $\mathcal{M}$ with the above definition, which is quoted from [4]. Here, $G(V, E)$ denotes a DAG (truthfully exists) which comprises nodes $V = \{1, 2, \cdots, v\}$ and directed edges $E \subseteq V \times V$. The empty-set $\{\emptyset\}$ implies that the mechanism might not select any node in some cases.

A common property in $\mathcal{M}$ is to make any node $x \in V$ reports on any potential edge $t \in V \times V$ in G (note that $t \in E$ does not always hold). As described herein, we designate t as *task*. Also, $T_{x,\mathcal{M}}$ is the set of tasks that x can report in $\mathcal{M}$. Each report r is specified as r_x^t when done from x to t, or just as r_x when its task need not be emphasized. In addition, let R_x denote the set of r that x accomplished by picking tasks from $T_{x,\mathcal{M}}$. Furthermore, R_x^* is specifically R_x, where the elements are all truthful reports.

Whether a report is truthful or not is represented by the stochastic *signal* s, which any x can observe from each t and can use it as input information for reporting. Our model assumes binary signals $s \in \{0, 1\}$ and binary reports $r(s) \in \{0, 1\}$ (0, disapproval; 1, approval), and uses notations s_x^t, s_x in the same manner with reporting. That is, node x supports the existence of an edge (task) t if $r_x^t(s_x^t) = 1$ and does not support it if $r_x^t(s_x^t) = 0$. This report is truthful in $r_x^t(0) = 0$ or $r_x^t(1) = 1$ case and is not truthful in $r_x^t(0) = 1$ or $r_x^t(1) = 0$ case.

We add two more assumptions, which are both common in the literature on peer prediction for binary signals (e.g., [17, 7, 5]). First, s, observed by each node from each task, is positively correlated. Therefore, when we pick another reference node $\hat{x} \in V$, both $Pr(s_x = 0|s_{\hat{x}} = 0) > Pr(s_x = 0)$ and $Pr(s_x = 1|s_{\hat{x}} = 1) > Pr(s_x = 1)$ hold for all x and $\hat{x}$.[g] Secondly, x adopts a uniform reporting *strategy* which R_x must follow. The set of strategies feasible in our model is depicted in Fig. 2, which shows the union of mapping strategies and uninformative (signal-independent) strategies. Mapping strategies literally follow a mapping rule from signals to reports; on the other hand, reports in uninformative strategies follow a given stochastic distribution independent of observed signals.[h] In the four possible mapping strategies under the assumption of binary signals, we especially define a strategy that always reports the truth as *truthful strategy*, and a strategy that always reports non-truth as *opposite strategy*. R_x^* is achieved when x adopts truthful strategy.

Here, assuming that x in $\mathcal{M}$ obtains some expected utility, $\mathbb{E}_{x,\mathcal{M}}$ through R_x (and $\mathbb{E}_{x,\mathcal{M}}$ might be affected by the reports done by other nodes), the truthfulness of a mechanism can be defined as follows:

Definition 2: Selection mechanism $\mathcal{M}$ has the following characteristics.

- *Truthful*, if $\mathbb{E}_{x,\mathcal{M}}(R_x^*, R_{-x}) \geq \mathbb{E}_{x,\mathcal{M}}(R_x, R_{-x})$ holds for all x, R_x, R_{-x},
- *Weakly truthful*, if $\mathbb{E}_{x,\mathcal{M}}(R_x^*, R_{-x}) = \mathbb{E}_{x,\mathcal{M}}(R_x, R_{-x})$ holds for all x, R_x, R_{-x},

[g]Accordingly, $Pr(s_x = 1|s_{\hat{x}} = 0) < Pr(s_x = 1)$ and $Pr(s_x = 0|s_{\hat{x}} = 1) < Pr(s_x = 0)$ hold, simultaneously.
[h]Note here that the strategy to keep reporting always 0 (and always 1) can be classified as both mapping and uninformative strategies.

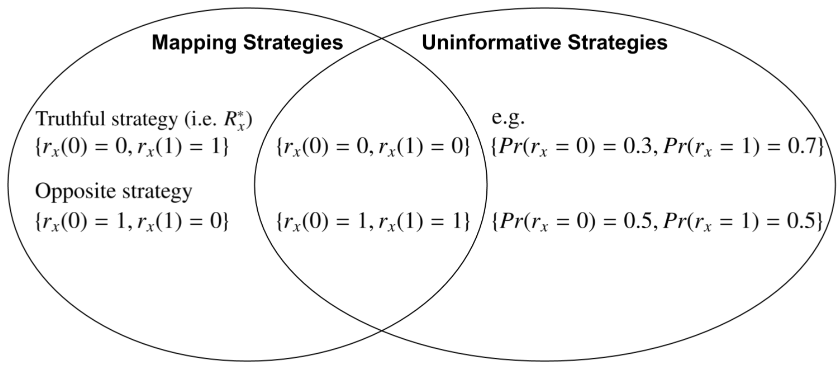


Fig. 2. Nodes (agents) can take either mapping or uninformative strategies for reporting.

where $R_{-x} = (R_1..., R_{x-1}, R_{x+1}..., R_v)$. Intuitively, using truthful selection mechanisms, a node can maximize its expected utility with a truthful strategy, irrespective of the reports done by other nodes. However, all nodes have the same expected utility no matter what they report when the selection mechanism is weakly truthful. Weak truthfulness is a necessary condition of truthfulness.

Although details are discussed in 2.2 and 2.3, the multi-task peer prediction that is part of the foundation of our model calculates the reward of a node x by comparison of one reference node whose reports that include at least one common task with x. If notation $\hat{x}$ is used again, then strong truthfulness [6] can be defined as follows:

Definition 3: A selection mechanism $\mathcal{M}$ is *strongly truthful* if $\mathbb{E}_{x,\mathcal{M}}(R_x^*, R_{\hat{x}}^*) \geq \mathbb{E}_{x,\mathcal{M}}(R_x, R_{\hat{x}})$ for all $x, \hat{x}, R_x, R_{\hat{x}}$. Equality might occur only when both x and $\hat{x}$ adopt the opposite strategy.[i]

Consequently, unlike weakly truthful selection mechanisms, strongly truthful selection mechanisms can assign strictly higher expected utility to the truthful strategy compared to any other strategy in virtually all cases. This equilibrium concept is weaker than truthfulness in terms of being dependent on the reports done by a reference node.

2.1. *Two-path mechanism*

In the two-path mechanism (we use the notation $\mathcal{M}_{2p}$, according to [4]), as already described, an influential node is the first intersection of the two independent random paths drawn on DAGs.

Specifically, $\mathcal{M}_{2p}$ follows the iterative process presented in Algorithm 1 (which is quoted from [4]), where the two random paths $\{P_1, P_2\}$ are drawn repeatedly until they intersect or until all nodes in the network are marked. This mark, denoted by the set U, is attached to all nodes on which the two paths have passed when they do not intersect. The marked nodes will never be selected as an influential node z. Since $\mathcal{M}_{2p}$ is designed for the environment

[i] The original definition by [6] generalizes both truthful strategy and opposite strategy as a *permutation strategy* in order to encompass the case of multiple (non-binary) signals.

Algorithm 1 Two-path mechanism $\mathcal{M}_{2p}$ by [4]

$U \leftarrow \emptyset$
while $U \neq V$ **do**
    Pick $x \in V$ uniformly at random
    $P_1 \leftarrow$ random path starting at x
    Pick $y \in V$ uniformly at random
    $P_2 \leftarrow$ random path starting at y
    if $P_1 \cap P_2 = \emptyset$ **then**
        $U \leftarrow U \cup P_1 \cup P_2$
    else
        $z \leftarrow$ the first vertex in $P_1 \cup P_2$
        if $z \in U$ **then**
            return $\emptyset$
        else
            return z
        end if
    end if
end while

where all nodes are selfish and wish to be selected as an influential one, we can write $\mathbb{E}_{x,\mathcal{M}_{2p}}$ as

$$\mathbb{E}_{x,\mathcal{M}_{2p}} = Pr(\mathcal{M}_{2p} = x),$$

where $Pr(\mathcal{M}_{2p} = x)$ stands for the probability that $\mathcal{M}_{2p}$ selects x.

It must be emphasized that $\{P_1, P_2\}$ are drawn by the random-walking on the out-edges elicited from each node because the designer behind $\mathcal{M}_{2p}$ cannot observe the true network structure. This random-walking enables us to define $T_{x,\mathcal{M}_{2p}}$ (the set of tasks for x in $\mathcal{M}_{2p}$) as x's potential out-edges, as

$$T_{x,\mathcal{M}_{2p}} = \{(x, v) \in P \times V \mid x \neq v\},$$

where we define $P \subseteq V$ as the set of nodes with total order, which denotes a path being drawn. Since x can report any out-edge except the self-edge, $|T_{x,\mathcal{M}_{2p}}| = |V| - 1$. Our study consistently assumes that x reports 0 for all tasks in $T_{x,\mathcal{M}_{2p}}$ other than the tasks it explicitly reported as 1. Namely, reporting is to divide $T_{x,\mathcal{M}_{2p}}$ into the two subsets: 0 (non-existing edges) and 1 (existing edges). The node would be the end point of the path if it reported 0 for all tasks in $T_{x,\mathcal{M}_{2p}}$.

An earlier report [4] presents the argument that $\mathcal{M}_{2p}$ consisting of the above settings is truthful in DAGs, and it can select influential nodes with sufficient accuracy.[j] However, strictly speaking, this mechanism satisfies only weak truthfulness.

Proposition 1: $\mathcal{M}_{2p}$ is weakly truthful in DAGs.

Proof: The proof of proposition 3.1 expressed in an earlier report [4] is illustrative. This proof describes that all nodes can no longer manipulate their probability of becoming the first

[j] See original paper [4] for the definition of *influential* and *accuracy*.

intersection at the point of reporting, regardless of the out-edges they report. In other words, $\mathbb{E}_{x,\mathcal{M}_{2p}}$ is neutral with respect to R_x. This outcome corresponds to weak truthfulness. □

Therefore, the $\mathcal{M}_{2p}$, where truthful reporting and misreporting are indifferent for all $x \in V$, cannot always draw random paths on true directed edges $E(G)$, even if all nodes report rationally to maximize their expected utility. Whereas [4] uses *given* DAGs for experiments to confirm the accuracy of $\mathcal{M}_{2p}$ on the grounds of its truthfulness, the possibility remains that nodes misreport their out-edges.

2.2. *Multi-task peer prediction*

To overcome such weak-truthfulness, we introduced multi-task peer prediction [5] to the DAGs as an additional reward layer. This model requires each node (agent) to report on at least two tasks, and calculates rewards to x when the two nodes (x and a reference node $\hat{x}$) report on a common task.[k] An important property on multi-task peer prediction is that its reward computation uses not only the reports that x and $\hat{x}$ did to a common task, but also the reports that x and $\hat{x}$ have done to other assigned tasks. According to earlier studies [5, 6], we can formulate the amount of reward σ_x^t when x has a common task t as

$$\sigma_x^t = \left[r_x^t \cdot r_{\hat{x}}^t + (1 - r_x^t)(1 - r_{\hat{x}}^t)\right] - (h_{x,0} \cdot h_{\hat{x},0} + h_{x,1} \cdot h_{\hat{x},1}),$$

where $h_{x,0} = \frac{|\{r_x \in R_x | r_x = 0\}|}{|R_x|}$ and $h_{\hat{x},0} = \frac{|\{r_{\hat{x}} \in R_{\hat{x}} | r_{\hat{x}} = 0\}|}{|R_{\hat{x}}|}$ are both empirical frequencies with which x and $\hat{x}$ report signal 0 to all assigned tasks (thus, $h_{x,1} = 1 - h_{x,0}$ and $h_{\hat{x},1} = 1 - h_{\hat{x},0}$). The terms in square brackets are the reward for agreement. It is apparent that the value of 1 is obtained when two reports return the same signal $(r_x^t, r_{\hat{x}}^t) = (0, 0)$ or $(1, 1)$; otherwise it is 0. The remaining terms are a kind of penalty, which increases as x and $\hat{x}$ repeat uninformative signal-independent reports. Presuming that x and $\hat{x}$ always report 1 for assigned tasks irrespective of the signals, $\sigma_x^t = 0$ holds because penalty terms become 1 even though a common task between x and $\hat{x}$ always engenders a reward of 1. A similar result would be derived in the case of fifty-fifty uninformative strategy (i.e., $Pr(r = 0) = Pr(r = 1) = 0.5$) because the expected value of reward terms and penalty terms both become 0.5. Note that the empirical frequency-based penalty can be interpreted as the expected value of the case where we randomly select $r_x \in R_x$ and $r_{\hat{x}} \in R_{\hat{x}}$, and assign penalty 1 if they are identical or assign 0 otherwise.[l] Actually, σ_x^t could be negative in case of disagreement because of the penalty terms.[m] Although not crucially important for our discussion, we can also make all rewards positive by adding 1 to all σ_x^t as a basic reward because penalty terms do not exceed 1.

The original report [5] indicated that the expected (net) reward for a common task $\mathbb{E}(\sigma_x^t)$ is maximized in the equilibrium where all nodes adopt the truthful strategy by exerting effort on signal observation, under the assumption of positively correlated signals. Furthermore, subsequent studies [6, 18] defined this contribution as strong truthfulness and pointed out the following necessary conditions to establish multi-task peer prediction without loss of generality: (i) two agents, (ii) three total tasks, (iii) two or more tasks per agent including at

[k] This type of peer prediction method is in some cases designated as an *output agreement* (OA).
[l] We use this interpretation for the proof of theorem 1.
[m] As described in Section 3, all nodes in our model would have negative total rewards.

least one common task, and (iv) the environment where each node cannot identify in advance which of the assigned tasks will be the common task.

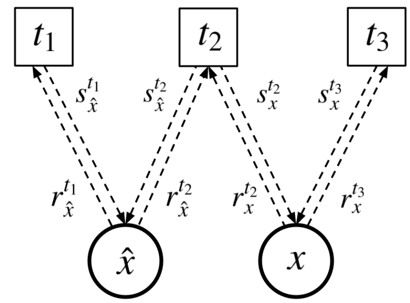


Fig. 3. In this minimum multi-task peer prediction, node x ($\hat{x}$) reports on signals $\{s_x^{t_2}, s_x^{t_3}\}$ ($\{s_{\hat{x}}^{t_1}, s_{\hat{x}}^{t_2}\}$) emitted from tasks $\{t_2, t_3\}$ ($\{t_1, t_2\}$), and common task is t_2.

Using the notation in our model, we can depict the minimum model reflecting these conditions as Fig. 3 above, where x and $\hat{x}$ can obtain reward or penalty to the common task t_2. Although σ_x^t uses all reports $\{r_{\hat{x}}^{t_1}, r_{\hat{x}}^{t_2}, r_x^{t_2}, r_x^{t_3}\}$ for reward computation, x and $\hat{x}$ do not know which of the two assigned tasks will generate rewards.

2.3. *Strongly truthful two-path mechanism*

We now present the strongly truthful two-path mechanism ($\mathcal{M}_{st2p}$), which is essentially a hybrid of methods described in Sections 2.1 and 2.2.

2.3.1. *Settings*

First, the expected utility $\mathbb{E}_{x,\mathcal{M}_{st2p}}$ is defined as a simple combination[n] of both components as

$$\mathbb{E}_{x,\mathcal{M}_{st2p}} = Pr(\mathcal{M}_{st2p} = x) + \mathbb{E}(\sigma_x^t).$$

However, we cannot have any common task under the assumption of $T_{x,\mathcal{M}_{2p}}$ where nodes report only their out-edges. $\mathcal{M}_{st2p}$ therefore newly allows each node to report its in-edges. This extension can be expressed as

$$T_{x,\mathcal{M}_{st2p}} = T_{x,\mathcal{M}_{2p}} \cup \{(v,x) \in V \times P \mid x \neq v\}.$$

In-edge reporting is done in a same manner as out-edge reporting: a node x reports 1 to the exsiting in-edges and out-edges and reports 0 to all other tasks in $T_{x,\mathcal{M}_{st2p}}$.[o] Because every nodes in a path needs to make $|T_{x,\mathcal{M}_{st2p}}| = 2(|V|-1)$ reports, x would obtain a number of common tasks as a result of the two-path drawings. $\mathcal{M}_{st2p}$ picks one common task t uniformly at random from them and computes σ_x^t according to multi-task peer prediction method by

[n]It might be unnatural to combine two different types of the terms: $Pr(\mathcal{M}_{st2p} = x)$ and $\mathbb{E}(\sigma_x^t)$, but this is not critical to our discussion, as we show in proposition 2.
[o]In the adjacency matrix for a DAG composed with $V = \{1, 2, \cdots, v\}$, the out-edge and in-edge reports by node x correspond to the declaration of x-th row and x-th column, respectively.

[5]. Here, we assume that $(r_x^t, r_{\hat{x}}^t) = (0,0)$ is excluded from the candidates of the common task for reward computation in order to make $\mathcal{M}_{st2p}$ practical in a DAG with low density.[p]

To establish compatibility between the two-path mechanism and multi-task peer prediction, we further add two assumptions to in-edge reporting. First, the result of in-edge reports have no priority over that of out-edge reports. Once an out-edge report $r_{\hat{x}}^{(\hat{x},x)} = 1$ is done, $\mathcal{M}_{st2p}$ accepts the existence of the edge $(\hat{x}, x)$ even if an in-edge report $r_x^{(\hat{x},x)} = 0$ is done before and after the out-edge report (and vice versa). In other words, an in-edge report is an assessment and does not directly determine the DAG structure. Secondly, a node cannot know the result of in-edge reports done by the nodes on the other path until it finishes reporting. This assumption is important for the two-path mechanism to prevent nodes from abusing in-edge reports to manipulate their probability of being selected as the first intersection of two paths. These are all settings to construct $\mathcal{M}_{st2p}$.

2.3.2. *Algorithms and properties*

Algorithm 2 Random path derivation for $\mathcal{M}_{st2p}$

```
function PATH(x, R)
    P ← ∅
    while true do
        P ← P ∪ {x}
        R_x ← reports by x
        if R_x ∉ R then
            R ← R ∪ R_x
        end if
        if R_x includes at least one out-edge then
            x ← random-walking according to R_x
        else
            return (P, R)
        end if
    end while
end function
```

Due to (weak) truthfulness, $\mathcal{M}_{2p}$ relies on the assumption that P_1 and P_2 are simply drawn at random for a given DAG. Actually, $\mathcal{M}_{st2p}$ would need a more precise definition of random path derivation because it calculates rewards based on the collected reports. We define the process of path derivation with the $\textsc{Path}(x, R)$ function shown in Algorithm 2, which returns (P, R): totally ordered set as the path and the updated set of finished reports, when x and existing R are given as arguments.

$\mathcal{M}_{st2p}$ is denoted by Algorithm 3, which inherits most of $\mathcal{M}_{2p}$ (Algorithm 1). It just replaces part of random path derivation with $\textsc{Path}(x, R)$ and adds several rows for the reward calculation.

[p]If any common task is a candidate, the common task used for reward computation is very likely to be $(0,0)$ in most DAGs because all tasks other than those explicitly reported as 1 are considered to be 0. We experimentally confirmed that this case diminishes the difference in expected rewards among the predetermined strategies. Note that excluding the $(0,0)$ case does not affect the strong truthfulness of multi-task peer prediction, as we will confirm in theorem 1.

Algorithm 3 Strongly truthful two-path mechanism $\mathcal{M}_{st2p}$

$U \leftarrow \emptyset$
$R \leftarrow \emptyset$ ▷ newly added
while $U \neq V$ **do**
    Pick $\{x, y\} \subset V$ uniformly at random
    $(P_1, R_1) \leftarrow \text{PATH}(x, R)$ ▷ replaced
    $(P_2, R_2) \leftarrow \text{PATH}(y, R_1)$ ▷ replaced
    $R \leftarrow R \cup R_2$ ▷ newly added
    if $P_1 \cap P_2 = \emptyset$ **then**
        $U \leftarrow U \cup P_1 \cup P_2$
    else
        $z \leftarrow$ the first vertex in $P_1 \cup P_2$
        if $z \in U$ **then**
            return $\emptyset$
        else
            return z
        end if
    end if
end while
Compute σ_x^t for all x according to R ▷ newly added
return σ_x^t for all x ▷ newly added

The two properties of $\mathcal{M}_{st2p}$ are worth mentioning. First, all edges composing P_1 and P_2 become common tasks for the two nodes connected by the edge, as depicted in Fig. 4 where t_2 is a common task for x and $\hat{x}$.

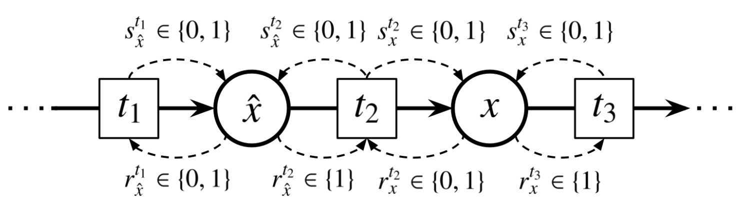


Fig. 4. All edges on a random path are subject to the reports by their endpoints. In this example, t_2 must be reported by x and $\hat{x}$ in this case.

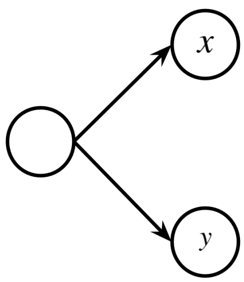
Secondly, $\mathcal{M}_{st2p}$ satisfies all necessary conditions to establish multi-task peer prediction as long as $|P_1 \cup P_2| \geq 3$ holds. Our mechanism actually can satisfy the necessary conditions when $|V| \geq 3$ holds in a DAG, if we ignore the assumption of excluding $(r_x^t, r_{\hat{x}}^t) = (0, 0)$ from the candidates of the common task for reward computation. In such a case, $\mathcal{M}_{st2p}$ satisfies:

- condition (i), because the mechanism selects at least two nodes as the starting point of the two paths.
- condition (ii), because the number of total tasks is ${}_3P_2 = 6$ if $|V| = 3$ (${}_2P_2 = 2$ if $|V| = 2$).
- condition (iii), because the number of tasks per node is $2(|V| - 1) = 4$ if $|V| = 3$. Also, a node x must obtain two common tasks $(x, \hat{x})$ and $(\hat{x}, x)$ as another node $\hat{x}$ reports

under the mechanism.

- condition (iv), because nodes have at least two common tasks, and they cannot know which of the common tasks will be used for reward computation.

However, when we exclude $(r_x^t, r_{\hat{x}}^t) = (0, 0)$, an exception is created. Fig. 5 depicts two minimum examples of $\mathcal{M}_{st2p}$ with three nodes and two edges, where x and y are the starting points of P_1 and P_2, respectively (as described in Algorithm 3). In example (a), where $|P_1 \cup P_2| = 2$, we can see that truthful reporting results in no common tasks available for reward computation, since the two non-existing edges (x, y) and (y, x) are all common tasks. Moreover, y having received x's out-edge report $r_x^{(x,y)} = 0$ can definitely make (y, x) the common task for reward computation, by misreporting $r_y^{(y,x)}$ as 1. (a) is thus contrary to the condition (iii) and (iv). To satisfy all necessary conditions, $\mathcal{M}_{st2p}$ needs $|P_1 \cup P_2| \geq 3$ such as example (b), where all nodes have at least two common tasks which they believe can both be used for reward computation at the point of reporting.[q]



(a) $|P_1 \cup P_2| = 2$ (b) $|P_1 \cup P_2| = 3$

Fig. 5. In the two examples of minimum $\mathcal{M}_{st2p}$, (a) does not satisfy the necessary conditions for multi-task peer prediction if we exclude $(r_x^t, r_y^t) = (0, 0)$ case from the candidates of the common task.

[q]As mentioned before, we here assume that a node cannot know the result of in-edge reports done by the nodes on the other path until it finishes reporting. Therefore, in both (a) and (b), y cannot know in advance whether x did in-edge report $r_x^{(y,x)} = 0$ or not.

2.3.3. *Strong truthfulness*

We can obtain the following proposition from $\mathcal{M}_{st2p}$.

Proposition 2: $Pr(\mathcal{M}_{st2p} = x)$ is weakly truthful in $\mathcal{M}_{st2p}$.

Proof: Presuming that P_1 reaches a node x, and that x reports its in-edges in addition to out-edges. As in the case of $\mathcal{M}_{2p}$, x can only be selected by the mechanism if P_2 reaches x before P_2 passes any other node of P_1. However, x cannot guide P_2 in a specific direction by abusing an in-edge report because the nodes in P_2 cannot observe x's in-edge reports until they finish reporting. Furthermore, although x strategically reports its out-edges considering that they would be assessed through future in-edge reports, no out-edge report affects x's probability of becoming the first intersection of P_1 and P_2 in $\mathcal{M}_{st2p}$ (as in $\mathcal{M}_{2p}$). These conditions are surely applicable to the nodes on P_2. Therefore, the additive in-edge reports have no influence on the weak truthfulness of original $\mathcal{M}_{2p}$. $Pr(\mathcal{M}_{st2p} = x)$ is still neutral with respect to R_x even if it includes in-edge reports. □

In other words, $Pr(\mathcal{M}_{st2p} = x)$ is independent of R_x and $\mathbb{E}(\sigma_x^t)$. Therefore, it is sufficient to examine $\mathbb{E}(\sigma_x^t)$ to prove the strong truthfulness of $\mathcal{M}_{st2p}$.

Theorem 1: $\mathcal{M}_{st2p}$ is strongly truthful in DAGs.

Proof: For the sake of convenience, this proof first assumes that $\mathcal{M}_{st2p}$ picks any common task for reward computation, and then removes $(r_x^t, r_{\hat{x}}^t) = (0, 0)$ case.[r] We also use the following Kronecker's delta to present the results of the comparison of two reports for given signals.

$$\delta_{r_x(s_x), r_{\hat{x}}(s_{\hat{x}})} = \begin{cases} 1 & (r_x(s_x) = r_{\hat{x}}(s_{\hat{x}})) \\ 0 & (r_x(s_x) \neq r_{\hat{x}}(s_{\hat{x}})) \end{cases}$$

The expected value of the reward terms in σ_x^t not only depends on $\delta_{r_x(s_x), r_{\hat{x}}(s_{\hat{x}})}$, but also on the probability distribution of input signals each node observes, as

$$\mathbb{E}\left[r_x^t \cdot r_{\hat{x}}^t + (1 - r_x^t)(1 - r_{\hat{x}}^t)\right] = \sum_{s_x=0}^{1} \sum_{s_{\hat{x}}=0}^{1} Pr(s_x, s_{\hat{x}}) \cdot \delta_{r_x(s_x), r_{\hat{x}}(s_{\hat{x}})},$$

where $Pr(s_x, s_{\hat{x}})$ is the joint probability distribution on the signals that x and $\hat{x}$ can receive from a common task.

As described already, penalty terms denote the expected value of the comparison between a random $r_x \in R_x$ and $r_{\hat{x}} \in R_{\hat{x}}$. This interpretation enables us to rewrite penalty terms into a similar form with the expected value of reward terms:

$$h_{x,0} \cdot h_{\hat{x},0} + h_{x,1} \cdot h_{\hat{x},1} = \sum_{s_x=0}^{1} \sum_{s_{\hat{x}}=0}^{1} Pr(s_x) Pr(s_{\hat{x}}) \cdot \delta_{r_x(s_x), r_{\hat{x}}(s_{\hat{x}})}.$$

[r] This proof is based on the proof of theorem 4.4 presented in an earlier report [6], which can deal with the case of multiple signals.

It uses product distribution $Pr(s_x)Pr(s_{\hat{x}})$ instead of $Pr(s_x, s_{\hat{x}})$ because the empirical frequency covers all R_x and $R_{\hat{x}}$: not only r^t_x and $r^t_{\hat{x}}$. Consequently, $\mathbb{E}(\sigma^t_x)$ can be expressed as

$$\mathbb{E}(\sigma^t_x) = \sum_{s_x=0}^{1} \sum_{s_{\hat{x}}=0}^{1} [Pr(s_x, s_{\hat{x}}) - Pr(s_x)Pr(s_{\hat{x}})] \cdot \delta_{r_x(s_x), r_{\hat{x}}(s_{\hat{x}})}.$$

In fact, terms in square brackets correspond to the correlation of s_x and $s_{\hat{x}}$. If one assumes that $Pr(s_x, s_{\hat{x}}) - Pr(s_x)Pr(s_{\hat{x}}) > 0$, for example, then both $Pr(s_x|s_{\hat{x}}) > Pr(s_x)$ and $Pr(s_{\hat{x}}|s_x) > Pr(s_{\hat{x}})$ hold because $Pr(s_x, s_{\hat{x}}) = Pr(s_x|s_{\hat{x}})Pr(s_{\hat{x}}) = Pr(s_{\hat{x}}|s_x)Pr(s_x)$. That is to say that s_x and $s_{\hat{x}}$ are positively correlated in this case.

Because $\mathcal{M}_{st2p}$ assumes positively correlated binary signals, the following condition holds in the expanded form of $\mathbb{E}(\sigma^t_x)$ as

$$\begin{aligned}\mathbb{E}(\sigma^t_x) = {} & [Pr(s_x = 0, s_{\hat{x}} = 0) - Pr(s_x = 0)Pr(s_{\hat{x}} = 0)]_{>0} \cdot \delta_{r_x(0), r_{\hat{x}}(0)} \\ & + [Pr(s_x = 0, s_{\hat{x}} = 1) - Pr(s_x = 0)Pr(s_{\hat{x}} = 1)]_{<0} \cdot \delta_{r_x(0), r_{\hat{x}}(1)} \\ & + [Pr(s_x = 1, s_{\hat{x}} = 0) - Pr(s_x = 1)Pr(s_{\hat{x}} = 0)]_{<0} \cdot \delta_{r_x(1), r_{\hat{x}}(0)} \\ & + [Pr(s_x = 1, s_{\hat{x}} = 1) - Pr(s_x = 1)Pr(s_{\hat{x}} = 1)]_{>0} \cdot \delta_{r_x(1), r_{\hat{x}}(1)},\end{aligned}$$

where $[x]_{>0}$ and $[x]_{<0}$ respectively indicate that x is positive and negative.[s]

It is apparent that $\mathbb{E}(\sigma^t_x)$ is maximized only when both x and $\hat{x}$ do truthful reports $(r(0) = 0, r(1) = 1)$ or opposite reports $(r(0) = 1, r(1) = 0)$. Any other pattern (i.e., nodes taking asymmetric strategies or always reporting the same signal) produces less-expected values. Under the assumption of reporting strategies, this outcome implies that $\mathbb{E}(\sigma^t_x)$ is maximized only when both x and $\hat{x}$ adopt either a truthful strategy or an opposite strategy.

This property holds even though $\mathcal{M}_{st2p}$ does not pick $(r^t_x, r^t_{\hat{x}}) = (0, 0)$ case, namely,

$$\begin{aligned}\mathbb{E}(\sigma^t_x) = {} & [Pr(s_x = 0, s_{\hat{x}} = 1) - Pr(s_x = 0)Pr(s_{\hat{x}} = 1)]_{<0} \cdot \delta_{r_x(0), r_{\hat{x}}(1)} \\ & + [Pr(s_x = 1, s_{\hat{x}} = 0) - Pr(s_x = 1)Pr(s_{\hat{x}} = 0)]_{<0} \cdot \delta_{r_x(1), r_{\hat{x}}(0)} \\ & + [Pr(s_x = 1, s_{\hat{x}} = 1) - Pr(s_x = 1)Pr(s_{\hat{x}} = 1)]_{>0} \cdot \delta_{r_x(1), r_{\hat{x}}(1)}.\end{aligned}$$

Hence, $\mathcal{M}_{st2p}$ is strongly truthful in DAGs. □

Accordingly, $\mathcal{M}_{st2p}$ can theoretically elicit truthful reports from each node on DAGs.

3. Experimental Study

In this section, we experimentally confirm the utility of strong truthfulness ensured by $\mathcal{M}_{st2p}$, with both synthetic and real-world data. The synthetic data were generated from the Barabási and Albert [21] (BA) model, and the real-world data were taken from a patent citation network. Experiments run $\mathcal{M}_{st2p}$ iteratively to the dataset, and show that the outcome satisfies our objective by comparing the derived reward distributions. Materials used for this experiment are uploaded on Github repository (https://github.com/knskito/Materials-for-ST2P-experiments).

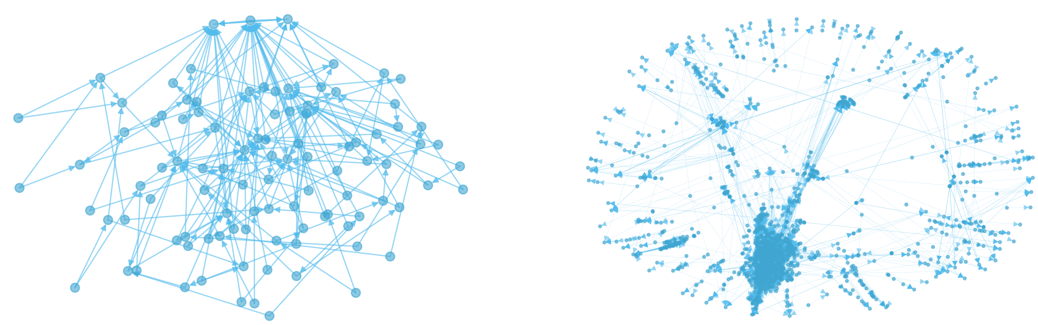
[s]Furthermore, if we designate $Pr(s_x = 0, s_{\hat{x}} = 0) - Pr(s_x = 0)Pr(s_{\hat{x}} = 0) = P_{00}, Pr(s_x = 0, s_{\hat{x}} = 1) - Pr(s_x = 0)Pr(s_{\hat{x}} = 1) = P_{01}, Pr(s_x = 1, s_{\hat{x}} = 0) - Pr(s_x = 1)Pr(s_{\hat{x}} = 0) = P_{10}, Pr(s_x = 1, s_{\hat{x}} = 1) - Pr(s_x = 1)Pr(s_{\hat{x}} = 1) = P_{11}$, they have the following relations: $P_{00} = P_{11}, P_{01} = P_{10}, P_{00} + P_{01} + P_{10} + P_{11} = 0$.

3.1. *Methods*

We specifically assume that every node in a DAG stochastically takes either a truthful strategy or aforementioned fifty-fifty uninformative strategy. The probability of strategy selection is exogenously determined by *randomness parameter* ϵ with which the node selects fifty-fifty uninformative strategy. This experiment first assigns the same ϵ to every node in a DAG, and then repeats $\mathcal{M}_{st2p}$ for a given number of times. After executing this process to all $\epsilon = \{0.0, 0.1, 0.2, \cdots, 0.9\}$ cases, we finally compare the 10 derived distributions for the total reward earned by each node. Note that the computed reward distributions do not include the nodes who did no reports as they have not been on any random paths (i.e., their rewards are zero.).

$\mathcal{M}_{st2p}$ is expected to give a node x the best expected reward $\mathbb{E}(\sigma_x^t)$ in the case of $\epsilon = 0.0$ (i.e., the equilibrium by truthful strategy), if the additional reward layer works well.

3.2. *Dataset*



(a) BA model (100 nodes)

(b) US patent citation network (1,579 nodes)

Fig. 6. Our experiments use two DAGs. (a) is generated from BA model with $m_0 = m = 2$ and $n = 100$, and (b) is extracted from the US patent citation network during years 2013-2017.

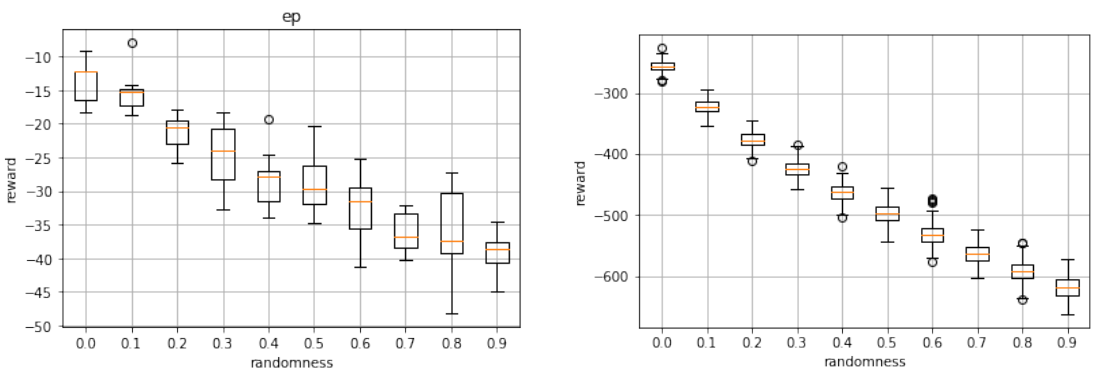
As mentioned above, we use the synthetic data from a BA model and the real-world data from the patent citation network.

The BA model is a network generative model in the basis of *preferential attachment* [21, 22]. At the beginning of the generative process, the model generates m_0 nodes. In each step, it generates one node with m edges and attaches the existing nodes randomly. The process ends after n steps are executed (i.e., n denotes the number of total nodes in the DAG). The distribution of attachment is determined by preferential attachment rule whose probability is proportional to the node's degree distribution. We set $m_0 = m = 2$ and $n = 100$ for the purposes of this experiment.

Real-world data were obtained from the US patent citation network. We have extracted a strongly connected component with 10,000 patents during years 2013–2017 from the database[t], and formatted them into a DAG structure with 1,579 nodes. Fig. 6 depicts the two DAGs generated as a result of the procudure above.

[t] http://www.patentsview.org/download/

3.3. *Results*



(a) BA model
(100 nodes and 100 iterations)

(b) US patent citation network
(1,579 nodes and 1,000 iterations)

Fig. 7. The box plot shows the median value as orange lines, 25/75 percentile as box, pseudo-maximum/minimum value as whiskers, and outliers as circles. (a) and (b) present the same trend that no-randomness provides the highest expected rewards.

Fig. 7 summarizes our experimental results, where we iterated $\mathcal{M}_{st2p}$ to 20 feasible patterns (2 DAGs · 10 parameters) and displayed the derived reward distributions as box plots.

Fig. 7 (a) is the result of running $\mathcal{M}_{st2p}$ 100 times to each case of BA model with different randomness values. We can see that the median is a maximum in $\epsilon = 0.0$ case, and that the median is decreasing as ϵ increases. Fig. 7 (b) is the result of running $\mathcal{M}_{st2p}$ 1,000 times to each case of the US patent citation network with different randomness values. It shows the same trend as that of the BA model. One feature in our experimental results is that any median value does not exceed zero.[u] This is probably due to the two assumptions: regarding all possible tasks other than report 1 as report 0, and excluding the (0, 0) case from the common tasks for reward computation. The former increases the penalty by making $h_{x,0}$ and $h_{\hat{x},0}$ closer to 1, and the latter simply reduces the probability that the selected common task has an agreement. The scale of these effects would depend on the density of a given DAG.

Although we need a further research on $\mathcal{M}_{st2p}$ and the density, these results imply that the additional reward layer based on multi-task peer prediction can work as an incentive for each node to adopt truthful strategy.

4. Conclusion

This study has assessed our *strongly truthful two-path mechanism*, which is a hybrid of the two-path mechanism [4] and multi-task peer prediction [5]. This mechanism improves the accuracy of the existing two-path mechanism by changing (strengthening) its weak truthfulness to strong truthfulness. We can achieve this enhancement simply with the additional reward layer, and the only new imposition on the nodes is the assessment of their own in-edges. Nevertheless, this mechanism encourages nodes to report truthfully, as indicated by the theorem and the results of simulations. Our these results have a contribution to the discussion

[u]We can adjust the median value by means of the basic reward mentioned in Section 2.2.

of information-diffusion mechanisms used to find an influential node from non-cooperative networks, such as Web and patent citation.

On the other hand, this approach limits an advantage of original two-path mechanism: the independence of expected utility against other nodes. This is inevitable because peer prediction requires reference nodes to elicit truthful reports for the problems with no ground truth (i.e., the edges on a DAG). It would be a goal for this research field to overcome this limitation; that is, to present a mechanism that provides strictly higher rewards to the truthful report, irrespective of the reports done by other nodes.

In addition, two more specific directions should be considered for future research. One is to extend the mechanism from binary to multiple signals (i.e., from a digraph to a weighted digraph in the context of network analysis). This would be one of the most required works because the extension has already been studied in each of preceding research fields, such as impartial selection mechanisms [14], information diffusion [20] and multi-task peer prediction [6]. The other important direction for this research is to resolve the budget problem. Our mechanism with peer prediction must pay an additional reward to the report by nodes, which imposes a burden for practical applications. Minimizing total payments while maintaining sufficient incentives for agents (nodes) is a branch of research issues relevant to peer prediction. It has been studied particularly by Jurca and Faltings [23, 24].

Acknowledgements

The authors would like to thank the anonymous referees for their valuable comments and helpful suggestions.

References

1. N. Nisan, T. Roughgarden, E. Tardos, and V.V. Vazirani (2007), *Algorithmic game theory*, Cambridge university press.
2. S. Brin and L. Page (1998), *The anatomy of a large-scale hypertextual web search engine*, Computer networks and ISDN systems 30, 1-7 (1998), 107117.
3. J.M. Kleinberg (1999), *Authoritative sources in a hyperlinked environment*, Journal of the ACM (JACM) 46, 5 (1999), 604632.
4. Y. Babichenko, O. Dean, and M. Tennenholtz (2018), *Incentive-Compatible Diffusion*, In Proceedings of the 2018 World Wide Web Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 13791388.
5. A. Dasgupta and A. Ghosh (2013), *Crowdsourced judgement elicitation with endogenous proficiency*, In Proceedings of the 22nd international conference on World Wide Web. ACM, 319330.
6. V. Shnayder, A. Agarwal, R. Frongillo, and D.C. Parkes (2016), *Informed truthfulness in multi-task peer prediction*, In Proceedings of the 2016 ACM Conference on Economics and Computation. ACM, 179196.
7. J. Witkowski and D.C. Parkes (2012), *Peer prediction without a common prior*, In Proceedings of the 13th ACM Conference on Electronic Commerce. ACM, 964981
8. G. Radanovic and B. Faltings (2013), *A robust bayesian truth serum for non-binary signals*, In Proceedings of the 27th AAAI Conference on Artificial Intelligence (AAAI" 13). 833839.
9. N. Miller, P. Resnick, and R. Zeckhauser (2005), *Eliciting informative feedback: The peer-prediction method*, Management Science51, 9 (2005), 13591373.
10. R. Jurca, B. Faltings (2009), *Mechanisms for making crowds truthful*, Journal of Artificial Intelligence Research34, 1 (2009), 209.
11. N. Alon, F. Fischer, A. Procaccia, and M. Tennenholtz (2011), *Sum of us: Strategyproof selection*

from the selectors, In Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge. ACM, 101110.
12. R. Holzman and H. Moulin (2013), *Impartial nominations for a prize*, Econo-metrica81, 1 (2013), 173196.
13. F. Fischer and M. Klimm (2015), *Optimal impartial selection*, SIAM J. Comput.44, 5 (2015), 12631285.
14. D. Kurokawa, O. Lev, J. Morgenstern, and A.D. Procaccia (2015), *Impartial Peer Review*, In IJCAI. 582588.
15. D. Kempe, J. Kleinberg, and . Tardos (2003), *Maximizing the spread of influence through a social network*, In Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 137146.
16. T. Gneiting and A.E. Raftery (2007), *Strictly proper scoring rules, prediction, and estimation*, J. Amer. Statist. Assoc.102, 477 (2007), 359378.
17. R. Jurca and B. Faltings (2005), *Enforcing truthful strategies in incentive compatible reputation mechanisms*, InInternational Workshop on Internet and Network Economics. Springer, 268277.
18. V. Shnayder, R. Frongillo, and D.C. Parkes (2016), *Measuring performance of peer prediction mechanisms using replicator dynamics*, (2016).
19. B. Faltings and G. Radanovic (2017), *Game theory for data science: eliciting truthful information*, Synthesis Lectures on Artificial Intelligence and Machine Learning 11, 2 (2017), 1151.
20. M. Mohite and Y. Narahari (2011), *Incentive compatible influence maximizationin social networks and application to viral marketing*, In The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 3. International Foundation for Autonomous Agents and Multiagent Systems, 10811082.
21. A.L. Barabsi and R. Albert (1999), *Emergence of scaling in random networks*, science 286, 5439 (1999), 509512.
22. D.S. Price (1976), *A general theory of bibliometric and other cumulative advantage processes*, Journal of the American society for Information science 27, 5(1976), 292306.
23. R. Jurca and B. Faltings (2006), *Minimum payments that reward honest reputation feedback*, In Proceedings of the 7th ACM conference on Electronic commerce. ACM, 190199.
24. R. Jurca and B. Faltings (2007), *Robust incentive-compatible feedback payments*, In Agent-Mediated Electronic Commerce. Automated Negotiation and Strategy Design for Electronic Markets. Springer, 204218.