Content Sharing Design for Social Welfare in Networked Disclosure Game (Supplementary Material)

Feiran Jia[1] **Chenxi Qiu**[2] **Sarah Rajtmajer**[1] **Anna Squicciarini**[1]

[1]Information Sciences and Technology, Pennsylvania State University,Pennsylvania, USA
[2]Computer Science and Engineering, University of North Texas,Texas, USA

1 OMITTED DETAILS OF THEORETICAL RESULTS

PROOF OF THEOREM 3.1

Proof. $\Rightarrow$: (1) Everyone who invests satisfies the **threshold condition**, implying that if $x_i = 1$, $U_i(x_i, \mathbf{x}_{-i}|\mathcal{G}) \geq U_i(1 - x_i, \mathbf{x}_{-i}|\mathcal{G}) \Rightarrow U_i(1, \mathbf{x}_{-i}|\mathcal{G}) \geq U_i(0, \mathbf{x}_{-i}|\mathcal{G})$. (2) other agents do not satisfy **threshold condition**, implying that if $x_i = 0$, $U_i(x_i, \mathbf{x}_{-i}|\mathcal{G}) < U_i(1 - x_i, \mathbf{x}_{-i}|\mathcal{G}) \Rightarrow U_i(0, \mathbf{x}_{-i}|\mathcal{G}) > U_i(1, \mathbf{x}_{-i}|\mathcal{G})$. Therefore, $\mathbf{x}$ is a PSNE.
$\Leftarrow$: If $\mathbf{x}$ is a PSNE and each user i breaks ties in favor of disclosing, then if $x_i = 1$, $U_i(x_i, \mathbf{x}_{-i}|\mathcal{G}) \geq U_i(1 - x_i, \mathbf{x}_{-i}|\mathcal{G})$, implying user i satisfies the threshold condition; if $x_i = 0$, $U_i(x_i, \mathbf{x}_{-i}|\mathcal{G}) < U_i(1 - x_i, \mathbf{x}_{-i}|\mathcal{G})$, implying user i doesn't satisfy the threshold condition.
The proof is completed. $\square$

PROOF OF THEOREM 3.3

Proof. We prove this theorem by constructing a polynomial time reduction from the NP-complete problem *subset sum* problem to OSDSP. Before the proof, we first introduce the decision problems of both subset sum and OSDSP:

Definition: The decision problem of subset sum.

- ▷ **Instance**: Given a set of M positive integers $\mathcal{A} = \{a_1, ..., a_M\}$ and a target sum value A.
- ▷ **Question**: Is there a subset $\mathcal{A}' \subseteq \mathcal{A}$, such that the sum of the elements in $\mathcal{A}'$ is equal to A, i.e., $\sum_{a_i \in \mathcal{A}'} a_i = A$

Definition: The decision problem of OSDSP.

- ▷ **Instance**: Given an input network $\mathcal{G}^{\text{in}} = (\mathcal{V}, \mathcal{E}^{\text{in}})$, impact coefficients $\{w_{i,j}\}_{i,j \in \mathcal{V}}$, cost coefficients $\{c_i\}_{i \in \mathcal{V}}$, a constant U, and a budget limit B.
- ▷ **Question**: Whether exists a solution $(\mathbf{x}, \mathcal{G})$ such that the social welfare $SW(\mathbf{x}|\mathcal{G}) \geq U$ and $|\mathcal{E}| \leq B$.

We now construct the following OSDSP instance that maps to any subset sum problem instance:
1) There are $2M$ users (nodes) in the network, i.e., $|\mathcal{V}| = 2M$;
2) As Fig. 1 shows, $\mathcal{G}^{\text{in}}$ is composed of M disjoint sub-graphs $\mathcal{G}_1^{\text{in}}, ..., \mathcal{G}_M^{\text{in}}$, where each $\mathcal{G}_i^{\text{in}}$ is composed of two nodes v_{2i-1} and v_{2i} connecting by an edge $e_{2i-1,2i}$ with edge cost $\lambda_{2i-1,2i} = a_i$;
3) In each $\mathcal{G}_i^{\text{in}}$, the two nodes v_{2i-1} and v_{2i} have their costs $c_{2i-1} = c_{2i} = 1.5a_i$, and $w_{2i-1,2i} = w_{2i,2i-1} = 2a_i$;
4) $U = A$ and $B = A$.

This reduction process from subset sum to OSDSP is performed in polynomial time. Before showing the correctness of this reduction, we first give Lemma 1.2 as a preparation:

Accepted for the 39th Conference on Uncertainty in Artificial Intelligence (UAI 2023).

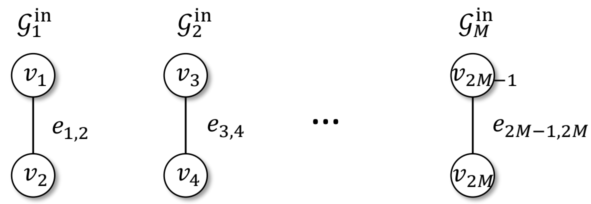

Figure 1: The OSDSP instance in the NP hard proof

Lemma 1.1. *For each sub-graph $\mathcal{G}_i^{in}$, there are only two possible PSNEs:*
(1) both nodes v_{2i-1} and v_{2i} self-disclose, or
(2) neither nodes disclose itself.

Proof of Lemma 1.2: Case (2) can be achieved when no edge is promoted.

For Case (1), we can promote the edge $e_{2i-1,2i}$ in $\mathcal{G}_i^{\text{in}}$. In this case, if both nodes disclose themselves, then $x_{2i} = 1$, and we can derive both $U_{2i}(x_{2i}, \mathbf{x}_{-2i}|\mathcal{G})$ and $U_{2i}(1 - x_{2i}, \mathbf{x}_{-2i}|\mathcal{G})$:

$$U_{2i}(x_{2i}, \mathbf{x}_{-2i}|\mathcal{G}) = w_{2i-1,2i}x_{2i-1}x_{2i} - c_{2i}x_{2i} = 2a_i - 1.5a_i = 0.5a_i, \tag{1}$$
$$U_{2i}(1 - x_{2i}, \mathbf{x}_{-2i}|\mathcal{G}) = 0, \tag{2}$$

indicating that $U_{2i}(x_{2i}, \mathbf{x}_{-2i}|\mathcal{G}) > U_{2i}(1 - x_{2i}, \mathbf{x}_{-2i}|\mathcal{G})$. Similarly, we can prove that $U_{2i-1}(x_{2i-1}, \mathbf{x}_{-(2i-1)}|\mathcal{G}) > U_{2i-1}(1 - x_{2i-1}, \mathbf{x}_{-(2i-1)}|\mathcal{G})$. Therefore, Case (1) is a PSNE.

Note that there exists no PNSE in $\mathcal{G}_i^{\text{in}}$ such that one node discloses and the other does not. For the sake of contradiction, consider the case that v_{2i} discloses and v_{2i-1} does not. Then, $x_{2i} = 1$ and $x_{2i-1} = 0$,

$$\begin{aligned} U_{2i}(x_{2i}, \mathbf{x}_{-2i}|\mathcal{G}) &= 0 - c_{2i} \le 0 - 1.5a_i = -1.5a_i & (3)\\ U_{2i}(1 - x_{2i}, \mathbf{x}_{-2i}|\mathcal{G}) &= 0 & (4) \end{aligned}$$

indicating that $U_{2i}(x_{2i}, \mathbf{x}_{-2i}|\mathcal{G}) < U_{2i}(1-x_{2i}, \mathbf{x}_{-2i}|\mathcal{G})$ and hence x_{2i} should not be 1 in this PNSE, which is a contradiction. The proof of Lemma 1.2 is completed.

Now we show the correctness of the polynomial reduction, i.e., *a solution exists for the subset sum instance if only if there exists a feasible solution for the OSDSP instance*. Note that in Case (1) of Lemma 1.2, the total social welfare of $\mathcal{G}_i^{\text{in}}$ is $\frac{a_i}{2} + \frac{a_i}{2} = a_i$ and the total cost of the promoted edges is a_i. In Case (2), the total social welfare of $\mathcal{G}_i^{\text{in}}$ is 0 (no node disclosed itself), and the total cost of the promoted edges is 0.

$\Rightarrow$: Assuming exists a solution $\mathcal{A}' = \{a_{i_1}, a_{i_2}, ..., a_{i_m}\}$ for the subset sum instance, i.e., $\sum_{l=1}^{m} a_{i_l} = A$, we can construct a feasible solution of the OSDSP instance: For each sub-graph $\mathcal{G}_{i_l}^{\text{in}}$ ($l = 1, ..., m$), we promote the edge $e_{2i-1,2i}$ and let the nodes disclose, i.e., $x_{2i-1} = x_{2i} = 1$. The total social welfare in $\mathcal{G}_{i_l}^{\text{in}}$ is equal to $\sum_{l=1}^{m} a_{i_l} = A \ge U$ and the total cost $\sum_{l=1}^{m} a_{i_l} = A \le B$.

$\Leftarrow$: Assuming exists a solution $(\mathbf{x}, \mathcal{G})$ in the OSDSP instance, where the edges (nodes resp.) in the sub-graphs $\mathcal{G}_{i_1}^{\text{in}}, \mathcal{G}_{i_2}^{\text{in}}, ..., \mathcal{G}_{i_m}^{\text{in}}$ are promoted (disclosed resp.). Hence,

$$SW(\mathbf{x}|\mathcal{G}) = \sum_{l=1}^{m} a_{i_l} \ge U = A \text{ and } \sum_{e_{i,j} \in \mathcal{E}} \lambda_{i,j} = \sum_{l=1}^{m} a_{i_l} \le B = A \tag{5}$$

indicating that $\sum_{l=1}^{m} a_{i_l} = A$. This implies that $\mathcal{A}' = \{a_{i_1}, a_{i_2}, ..., a_{i_m}\}$ is a feasible solution of the subset sum instance. $\square$

PROOF OF LEMMA 1.2

Lemma 1.2. *For each sub-graph $\mathcal{G}_i^{in}$, there are only two possible PSNEs:*
(1) both nodes v_{2i-1} and v_{2i} self-disclose, or
(2) neither nodes disclose itself.

Proof. Case (2) can be achieved when no edge is promoted.

For Case (1), we can promote the edge $e_{2i-1,2i}$ in $\mathcal{G}_i^{\text{in}}$. In this case, if both nodes disclose themselves, then $x_{2i} = 1$, and we can derive both $U_{2i}(x_{2i}, \mathbf{x}_{-2i}|\mathcal{G})$ and $U_{2i}(1 - x_{2i}, \mathbf{x}_{-2i}|\mathcal{G})$:

$$\begin{aligned} U_{2i}(x_{2i}, \mathbf{x}_{-2i}|\mathcal{G}) &= w_{2i-1,2i}x_{2i-1}x_{2i} - c_{2i}x_{2i} \\ &= 2a_i - 1.5a_i = 0.5a_i, \quad (6) \\ U_{2i}(1 - x_{2i}, \mathbf{x}_{-2i}|\mathcal{G}) &= 0, \quad (7) \end{aligned}$$

indicating that $U_{2i}(x_{2i}, \mathbf{x}_{-2i}|\mathcal{G}) > U_{2i}(1 - x_{2i}, \mathbf{x}_{-2i}|\mathcal{G})$. Similarly, we can prove that $U_{2i-1}(x_{2i-1}, \mathbf{x}_{-(2i-1)}|\mathcal{G}) > U_{2i-1}(1 - x_{2i-1}, \mathbf{x}_{-(2i-1)}|\mathcal{G})$. Therefore, Case (1) is a PSNE.

Note that there exists no PNSE in $\mathcal{G}_i^{\text{in}}$ such that one node discloses and the other does not. For the sake of contradiction, consider the case that v_{2i} discloses and v_{2i-1} does not. Then, $x_{2i} = 1$ and $x_{2i-1} = 0$,

$$\begin{aligned} &U_{2i}(x_{2i}, \mathbf{x}_{-2i}|\mathcal{G}) = 0 - c_{2i} \le 0 - 1.5a_i = -1.5a_i \\ &U_{2i}(1 - x_{2i}, \mathbf{x}_{-2i}|\mathcal{G}) = 0 \end{aligned}$$

indicating that $U_{2i}(x_{2i}, \mathbf{x}_{-2i}|\mathcal{G}) < U_{2i}(1-x_{2i}, \mathbf{x}_{-2i}|\mathcal{G})$ and hence x_{2i} should not be 1 in this PNSE, which is a contradiction. The proof of Lemma 1.2 is completed. □

PROOF OF LEMMA 4.2

Proof. **Base case** - In the first iteration, each node i deactivated by MaxInvest cannot disclose in any PSNE.
The induction step - Assuming that the nodes deactivated by MaxInvest in the first k iterations cannot disclose in any PSNE, then the nodes deactivated by MaxInvest in the $(k + 1)$th iteration cannot disclose in any PSNE. Therefore, any node deactivated by MaxInvest cannot disclose in any other PSNE. The proof is completed. □

PROOF OF THEOREM 4.2

Proof. For any PSNE $\mathbf{x}'$, we have $\mathbf{x}' \le \mathbf{x}$ (according to Lemma 4.2). We let $\mathcal{A}$ and $\mathcal{A}'$ denote the set of nodes disclosed in $\mathbf{x}$ and $\mathbf{x}'$, respectively, i.e., $\mathcal{A}' \subseteq \mathcal{A}$. Then,

$$\begin{aligned} SW(\mathbf{x}|\mathcal{G}) - SW(\mathbf{x}'|\mathcal{G}) &= \sum_{i\in\mathcal{A}} U_i\left(\mathbf{x}|\mathcal{G}\right) - \sum_{i\in\mathcal{A}'} U_i\left(\mathbf{x}'|\mathcal{G}\right) \quad (8) \\ &= \sum_{i\in\mathcal{A}\setminus\mathcal{A}'} \underbrace{U_i\left(\mathbf{x}|\mathcal{G}\right)}_{\ge 0 \text{ since each } i \text{ discloses}} + \sum_{i\in\mathcal{A}'} \left(U_i\left(\mathbf{x}|\mathcal{G}\right) - U_i\left(\mathbf{x}'|\mathcal{G}\right)\right) \\ &\ge \sum_{i\in\mathcal{A}'} \left(x_i \sum_{j\in\mathcal{N}_i} w_{j,i}x_j - c_ix_i - x_i' \sum_{j\in\mathcal{N}_i} w_{j,i}x_j' + c_ix_i' \right) \quad (9) \\ &= \sum_{i\in\mathcal{A}'} \left(\sum_{j\in\mathcal{N}_i} w_{j,i} \underbrace{\left(x_j - x_j'\right)}_{\ge 0} \right) \ge 0. \quad (10) \end{aligned}$$

indicating that $SW(\mathbf{x}|\mathcal{G}) \ge SW(\mathbf{x}'|\mathcal{G})$ for any PSNE $\mathbf{x}'$. □

PROOF OF THEOREM 4.5

Proof. Let $\mathbf{x}$ and $\mathbf{y}$ denote the returned profiles of $MaxInvest(\mathcal{S})$ and $MaxInvest(\mathcal{T})$.

(1) We prove $\mathbf{x} \le \mathbf{y}$ by induction. **Base Case:** In the first iteration of MaxInvest($\mathcal{T}$), for each node i satisfying $\sum_{e_{i,j}\in\mathcal{T}} w_{j,i} < c_i$, we set $y_i^{(0)} = 0$. Correspondingly, each $x_i = 0$ since $\mathcal{S} \subseteq \mathcal{T}$ and $\sum_{e_{i,j}\in\mathcal{S}} w_{j,i} \le \sum_{e_{i,j}\in\mathcal{T}} w_{j,i} < c_i$. Therefore $\mathbf{x} \le \mathbf{y}^{(0)}$.
The induction step: Assuming that in the kth iteration of MaxInvest($\mathcal{T}$), $\mathbf{x} \le \mathbf{y}^{(k)}$. We will then prove in the $k + 1$th

iteration, $\mathbf{x} \leq \mathbf{y}^{(k+1)}$. In the $(k+1)$th iteration, suppose that node i is popped off Q, we have $\sum_{e_{i,j} \in \mathcal{T}} w_{j,i} y_j^{(k)} < c_i$ and $y_i^{(k+1)} = 0$. Given $\mathbf{x} \leq \mathbf{y}^{(k)}$ and $\mathcal{S} \subseteq \mathcal{T}$, we have $\sum_{e_{i,j} \in \mathcal{S}} w_{j,i} x_j \leq \sum_{e_{i,j} \in \mathcal{T}} w_{j,i} y_j^{(k)} < c_i$. Since $\mathbf{x}$ is a PSNE, we have $y_i^{(k+1)} = x_i = 0$. Then in the iteration $k+1$, $\mathbf{x} \leq \mathbf{y}^{(k+1)}$.

We conclude that $\forall k: \ \mathbf{x} \leq \mathbf{y}^{(k)}$, and thus $\mathbf{x} \leq \mathbf{y}$.

(2) As $\sum_{i \in \mathcal{V}} x_i \leq \sum_{i \in \mathcal{V}} y_i \Rightarrow I(\mathcal{S}) \leq I(\mathcal{T})$.

(3) Then we show that $\sigma(\mathcal{T}) \geq \sigma(\mathcal{S})$. Let $\mathcal{A}(\mathcal{T})$ and $\mathcal{A}(\mathcal{S})$ denote the set of disclosed nodes in MaxInvest($\mathcal{T}$) and MaxInvest($\mathcal{S}$).

$$\begin{aligned}
\sigma(\mathcal{T}) - \sigma(\mathcal{S}) &= SW(\mathbf{y}|(\mathcal{V},\mathcal{T})) - SW(\mathbf{x}|(\mathcal{V},\mathcal{S})) && (11)\\
&= \sum_{i \in \mathcal{A}(\mathcal{S})} \Big(U_i(\mathbf{y}|(\mathcal{V},\mathcal{T})) - U_i(\mathbf{x}|(\mathcal{V},\mathcal{S})) \Big) + \sum_{i \in \mathcal{A}(\mathcal{T}) \setminus \mathcal{A}(\mathcal{S})} U_i(\mathbf{y}|(\mathcal{V},\mathcal{T})) && (12)\\
&\geq \sum_{i \in \mathcal{A}(\mathcal{S})} \Big(y_i \sum_{e_{i,j} \in \mathcal{T}} w_{j,i} y_j - c_i y_i - x_i \sum_{e_{i,j} \in \mathcal{S}} w_{j,i} x_j + c_i x_i \Big) && (13)\\
&= \sum_{i \in \mathcal{A}(\mathcal{S})} \Big(\sum_{e_{i,j} \in \mathcal{T} \setminus \mathcal{S}} w_{j,i} y_j + \sum_{e_{i,j} \in \mathcal{S}} w_{j,i} (y_j - x_j) \Big) && (14)\\
&\geq 0 && (15)
\end{aligned}$$

$\square$

PROOF OF THEOREM 4.6

Proof. To demonstrate the super-modularity, we would like to prove that for any edge $e^* \in \mathcal{E}^{(in)}$, and all pairs of the set $\mathcal{S} \subseteq \mathcal{T} \subseteq \mathcal{E}^{(in)}$, $\sigma(\cdot)$ satisfies

$$\sigma(\mathcal{S} \cup \{e^*\}) - \sigma(\mathcal{S}) \leq \sigma(\mathcal{T} \cup \{e^*\}) - \sigma(\mathcal{T}) \tag{16}$$

Let $\mathcal{A}(\mathcal{T})$ and $\mathcal{A}(\mathcal{S})$ denote the set of disclosed nodes in MaxInvest($\mathcal{T}$) and MaxInvest($\mathcal{S}$). Suppose $e^* = (u, v)$, we discuss the following cases. According to the monontoncity, we have $\mathcal{A}(\mathcal{S}) \subseteq \mathcal{A}(\mathcal{T})$.
(1) If e^* connects the disclosed nodes in $\mathcal{A}(\mathcal{S})$, $\sigma(\mathcal{S} \cup \{e^*\}) - \sigma(\mathcal{S}) = \sigma(\mathcal{T} \cup \{e^*\}) - \sigma(\mathcal{T}) = w_{u,v} + w_{v,u}$.
(2) If e^* connects two nodes outside of $\mathcal{A}(\mathcal{T})$, then $\sigma(\mathcal{S} \cup \{e^*\}) - \sigma(\mathcal{S}) = \sigma(\mathcal{T} \cup \{e^*\}) - \sigma(\mathcal{T}) = w_{u,v} - c_v + w_{v,u} - c_w$.
(3) If e^* connects one node in $\mathcal{A}(\mathcal{S})$ and another node outside the set $\mathcal{A}(\mathcal{T})$ (i.e. $\mathcal{V} \setminus \mathcal{A}(\mathcal{T})$), $\sigma(\mathcal{S} \cup \{e^*\}) - \sigma(\mathcal{S}) = \sigma(\mathcal{T} \cup \{e^*\}) - \sigma(\mathcal{T}) = w_{u,v} - c_v + w_{v,u}$ or $w_{v,u} - c_w + w_{u,v}$.
(4) If e^* connects one node in $\mathcal{A}(\mathcal{T}) \setminus \mathcal{A}(\mathcal{S})$ and another node outside the set $\mathcal{A}(\mathcal{T})$, we have $\sigma(\mathcal{S} \cup \{e^*\}) - \sigma(\mathcal{S}) = w_{u,v} - c_v + w_{v,u} - c_w \leq \sigma(\mathcal{T} \cup \{e^*\}) - \sigma(\mathcal{T}) = w_{u,v} + w_{v,u} - c_w$ or $w_{u,v} - c_v + w_{v,u}$.
(5) If e^* connects two nodes in $\mathcal{A}(\mathcal{T}) \setminus \mathcal{A}(\mathcal{S})$, we have $\sigma(\mathcal{S} \cup \{e^*\}) - \sigma(\mathcal{S}) = w_{u,v} - c_v + w_{v,u} - c_w \leq \sigma(\mathcal{T} \cup \{e^*\}) - \sigma(\mathcal{T}) = w_{u,v} + w_{v,u}$. $\square$

PROOF OF THEOREM 1.1

Theorem 1.1. *The optimal investment function $I(\mathcal{E})$ is sub-modular when $\frac{c_i}{w_{j,i}} \leq 1, \forall (i,j) \in \mathcal{E}^{in}$.*

Proof. To demonstrate the sub-modularity, we would like to prove that for any edge $e^* \in \mathcal{E}^{(in)}$, and all pairs of the set $\mathcal{S} \subseteq \mathcal{T} \subseteq \mathcal{E}^{(in)}$, $\sigma(\cdot)$ satisfies

$$I(\mathcal{S} \cup \{e^*\}) - \sigma(\mathcal{S}) \geq I(\mathcal{T} \cup \{e^*\}) - \sigma(\mathcal{T}) \tag{17}$$

According to the monontoncity, we have $\mathcal{A}(S) \subseteq \mathcal{A}(\mathcal{T})$. Suppose $e^* = (u, v)$, we discuss the following cases. (1) If e^* connects the disclosed nodes in $\mathcal{A}(S)$, then there are no new disclosing nodes. $I(S \cup \{e^*\}) - I(S) = I(\mathcal{T} \cup \{e^*\}) - I(\mathcal{T}) = 0$. (2) If e^* connects two nodes outside of $\mathcal{A}(\mathcal{T})$, then $I(S \cup \{e^*\}) - I(S) = I(\mathcal{T} \cup \{e^*\}) - I(\mathcal{T}) = 2$. (3) If e^* connects one node in $\mathcal{A}(S)$ and another node outside the set $\mathcal{A}(\mathcal{T})$ (i.e. $\mathcal{V} \setminus \mathcal{A}(\mathcal{T})$), $I(S \cup \{e^*\}) - I(S) = I(\mathcal{T} \cup \{e^*\}) - I(\mathcal{T}) = 1$. (4) If e^* connects one node in $\mathcal{A}(\mathcal{T}) \setminus \mathcal{A}(S)$ and another node outside the set $\mathcal{A}(\mathcal{T})$, we have $I(S \cup \{e^*\}) - I(S) = 2 \geq I(\mathcal{T} \cup \{e^*\}) - I(\mathcal{T}) = 1$. (5) If e^* connects two nodes in $\mathcal{A}(\mathcal{T}) \setminus \mathcal{A}(S)$, we have $I(S \cup \{e^*\}) - I(S) = 2 \geq I(\mathcal{T} \cup \{e^*\}) - I(\mathcal{T}) = 0$. $\square$

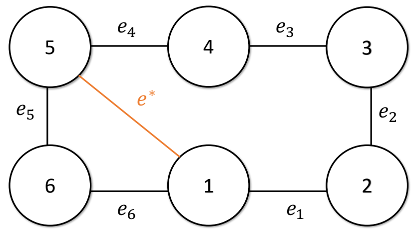

Figure 2: Instance of Remark 1.2 and Remark 1.3.

EXAMPLES OF REMARK 1.2 AND REMARK 1.3

Remark 1.2. *$\sigma(\cdot)$ is in general non-supermodular.*

Remark 1.3. *$I(\cdot)$ is in general non-submodular.*

Figure 2 gives an example of Remark 1.2: We assume that $\mathcal{T} = \{e_1.e_2, e_3, e_4, e_5, e_6\}$, $S = \{e_1.e_2, e_3, e_4\}$. All the nodes have the same cost c, and all the weights are $w \geq 0$. We assume that $c = 1.2w$. The node will disclose only when there are more than two neighbours disclose. We have $\sigma(S \cup \{e^*\}) - \sigma(S) = 5(2w - c) = 4w > \sigma(\mathcal{T} \cup \{e^*\}) - \sigma(\mathcal{T}) = 2w$.

Figure 2 also gives an example of Remark 1.2: Assuming $\mathcal{T} = \{e_1.e_2, e_3, e_4\}$, $S = \{e_1\}$, $\frac{c_i}{w_{j,i}} = 2, \forall (i,j) \in \mathcal{E}^{in}$, we have $I(S \cup \{e^*\}) - I(S) = 0 < I(\mathcal{T} \cup \{e^*\}) - I(\mathcal{T}) = 5$.

2 OMITTED DETAILS OF EXPERIMENTS

SOCIAL INTERACTION GRAPHS

Figures 3, 4, 5 illustrate the respective social interaction graphs. Node size is proportional to node degree. Red nodes represent users with a high self-disclosure rate (between 0.95 and 1).

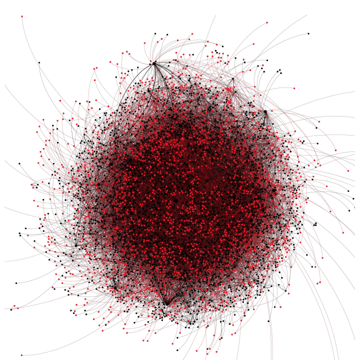

Figure 3: April 2021.

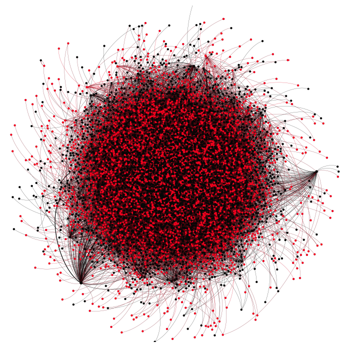

Figure 4: August 2020.

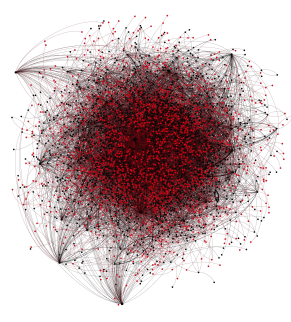

Figure 5: September 2020.

OMITTED DETAILS OF SECTION 5.2 (LABEL GENERATION)

In the original dataset, each sentence is associated with 6 labels: informational disclosure, emotional disclosure, support, general support, informational support, and emotional support. We use three of them in our task: informational disclosure, emotional disclosure, and emotional support. For example, the following training sentence is labeled for emotional disclosure and emotional support:

$\triangleright$ *I hope this chapter results in a better, healthier, more fulfilled you!!*

While our initial study only focused on the disclosure label, we believe that the labeled dataset created for this study can be of great value for future research in this area.

COMPUTATION TIME

In table 1, we present the computation time (in seconds) of our algorithm for the set of experiments in Section 5.2, which demonstrates that our heuristics work efficiently in large-scale networks. We conducted 20 trials for each experiment. Notably, our heuristics can solve the problem extremely quickly, particularly when the budget is large.

	b=0.2	**b=0.4**	**b=0.6**	**b=0.8**	**b=1.0**	**b=1.5**	**b=2.0**	**b=3.0**
Aug	464.27 ± 65.74	282.78 ± 68.12	126.40 ± 59.00	28.78 ± 28.96	1.42 ± 0.38	1.16 ± 0.38	1.06 ± 0.28	1.03 ± 0.28
Sep	77.63 ± 15.03	36.91 ± 11.78	9.90 ± 6.92	0.52 ± 0.15	0.52 ± 0.14	0.38 ± 0.11	0.38 ± 0.11	0.38 ± 0.11
Apr	198.19 ± 32.51	132.17 ± 23.53	58.84 ± 10.57	18.45 ± 14.19	1.44 ± 2.05	0.65 ± 0.17	0.65 ± 0.17	0.63 ± 0.17

Table 1: The run-time (in secs).

MODEL VALIDATION

We provided a preliminary attempt to validate our assumption that users' actions (disclose or not) depend on the threshold function (Equ. (2)). The modeling method is the same as Section 5.2 (Algorithmic Results). However, we only consider the self-disclosure information of the posts weekly due to the lack of time data regarding comments. Users are more conservative about posting than commenting. In each month, we pick the users who have at least one response record each week (41 users in August 2020, 52 users in September 2020, and 22 users in April 2021). We first estimate users' cost coefficients in the first week and use the estimated cost coefficients to predict users' responses in the remaining weeks of each month (4 weeks in April 2021, August 2020, and 2 weeks in September 2020). After, we compare whether the predicted responses are consistent with the users' actual responses, of which the results are listed in Table 2 (FPR and FNR stand for *false positive rate* and *false negative rate*). The table demonstrates that the threshold condition (Equ. (2)) can accurately predict users'

Month	FPR	FNR	F1	precision	recall
August 2020	0.1388	0.4607	0.6427	0.7953	0.5393
September 2020	0.1675	0.6595	0.4516	0.6703	0.3405
April 2021	0.0921	0.8127	0.2928	0.6723	0.1873

Table 2: Model validation results.

responses, especially in indicating the disclosure of users (as FPR is low). This provides empirical evidence to incentivize the disclosure of desired users by satisfying their threshold condition through edge promotion.

This simple modeling method is implicitly based on the assumptions that (1) the user behaves strategically, and (2) we assume that the has converged to a PSNE. Filling the gap between the theoretical model and real data (real-world user behavior) is non-trivial.

3 ADDITIONAL RELATED WORKS

Binary Networked Public Good Games. One line of relevant game theoretic research is the literature on binary networked public goods games (BNPG).The binary networked public goods game (BNPG) is a variant of a graphical game, where players' utilities depend on the strategies of their neighbors in the social graph. Benefits are a function of accumulated efforts Bramoullé and Kranton [2007] and investment strategies are binary Yu et al. [2020]. In Kempe et al. [2020], authors study network design to induce equilibria in BNPGs. Altruism modeling has also been considered to achieve desired investment profiles Yu et al. [2021]. Our model considers a setting similar to the BNPG but with different benefit functions designed for the self-disclosure application.

Network Design. The outer OSDSP (i.e. content-sharing network design) is intrinsically a network structure design problem. Despite the most similar work Corò et al. [2019], Coró et al. [2021], Yu et al. [2021], Kempe et al. [2020] mentioned before, types of network designs include removing edges Kimura et al. [2008, 2009] or nodes Jia et al. [2020], adding nodes or edges Sheldon et al. [2012], Amelkin and Singh [2019], and edge manipulation Chen et al. [2016], Castiglioni et al. [2020], etc.

4 ADDITIONAL ANALYSIS OF PSNES IN NDG

In this section, we discover several PSNEs in NDG that can be obtained with low time complexity.

Theorem 4.1. *If* $c_i > 0, \forall i$, $x_i = 0, \forall i$ *is a PSNE.*

Proof. When $x_j = 0, \forall j \neq i$, we have $g_i = 0$ and $U_i = -c_i x_i$, from which we can obtain that $U_i(1, \mathbf{0}) < U_i(0, \mathbf{0})$. Due to symmetric, $x_i = 0, \forall i$ is a PSNE.

To show that $x_i = 0, \forall i$ is a PSNE, we need to demonstrate that no player can gain by deviating from this strategy, given that all other players are also using this strategy.

For any player i, if $x_i = 0$, then $U_i = 0$. If $x_i = 1$, then player i incurs a cost of c_i. Therefore, player i has no incentive to deviate from $x_i = 0$, and this applies to all players.

Therefore, $x_i = 0, \forall i$ is a PSNE. □

Remark 4.2. *In cases where* $c_i > 0$ *for all* i*, we can generate PSNEs by considering each combination of connected components (CCs) in the optimal PSNE graph. Specifically, setting* x_i *to 1 for nodes within these CCs and to 0 for nodes outside of them results in a PSNE.*

To begin, we define the optimal PSNE graph G^* as a subgraph of the promotion network G that includes only the nodes who invest in the optimal solution $\mathbf{x}^*$ and the corresponding edges among them. The algorithm MaxInvest is used to determine $\mathbf{x}^*$. As $\mathbf{x}^*$ is the optimal solution, nodes with $x_i^* = 0$ will not disclose in any case (From Lemma 4.2).

As per the claim of the trivial PSNE, it is evident that $x_i = 0$ for all i in a connected component C constitutes a PSNE among the agents in that component (trivial PSNE of this connected component). This strategy does not impact the strategies of other connected components. This is because the utility functions of nodes in different connected components are independent of one another.

Therefore, to generate PSNEs, we can select a set of connected components in the optimal PSNE graph G^* and apply the strategy of setting x_i to 1 for nodes within the selected connected components and 0 for other nodes in G. This strategy results in a PSNE since no player can unilaterally change their strategy and increase their payoff.

Noted that we can use this approach to find PSNEs given a PSNE (not necessarily the optimal one).

Remark 4.3. *When* $\exists c_i = 0$*, we can use a spread process to find a PSNE.*

(Spread Process) We consider the set of agents with $c_i = 0$ as a "seed set" which are always disclosed as $U_i(1, \mathbf{x}_{-i}) \geq U_i(1, \mathbf{x}_{-i}) = 0$. We can start from them and do a process similar to the diffusion of the Linear Threshold model Kempe et al. [2003] to find a PSNE. The process can be described as follows: At each step, all nodes that have already disclosed in the previous step remain disclosed. Any agent who satisfies the threshold condition in the current step will also disclose. We continue iterating this process until no more agents can disclose. At this point, we have found a PSNE for the game.

References

Victor Amelkin and Ambuj K Singh. Fighting opinion control in social networks via link recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 677–685, 2019.

Yann Bramoullé and Rachel Kranton. Public goods in networks. *Journal of Economic theory*, 135(1):478–494, 2007.

Matteo Castiglioni, Diodato Ferraioli, and Nicola Gatti. Election control in social networks via edge addition or removal. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 1878–1885, 2020.

Chen Chen, Hanghang Tong, B Aditya Prakash, Tina Eliassi-Rad, Michalis Faloutsos, and Christos Faloutsos. Eigen-optimization on large graphs by edge manipulation. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 10 (4):1–30, 2016.

Feiran Jia, Kai Zhou, Charles Kamhoua, and Yevgeniy Vorobeychik. Blocking adversarial influence in social networks. In *Decision and Game Theory for Security: 11th International Conference, GameSec 2020, College Park, MD, USA, October 28–30, 2020, Proceedings 11*, pages 257–276. Springer, 2020.

David Kempe, Sixie Yu, and Yevgeniy Vorobeychik. Inducing equilibria in networked public goods games through network structure modification. *arXiv preprint arXiv:2002.10627*, 2020.

Masahiro Kimura, Kazumi Saito, and Hiroshi Motoda. Solving the contamination minimization problem on networks for the linear threshold model. In *PRICAI 2008: Trends in Artificial Intelligence: 10th Pacific Rim International Conference on Artificial Intelligence, Hanoi, Vietnam, December 15-19, 2008. Proceedings 10*, pages 977–984. Springer, 2008.

Masahiro Kimura, Kazumi Saito, and Hiroshi Motoda. Blocking links to minimize contamination spread in a social network. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 3(2):1–23, 2009.

Daniel Sheldon, Bistra Dilkina, Adam N Elmachtoub, Ryan Finseth, Ashish Sabharwal, Jon Conrad, Carla P Gomes, David Shmoys, William Allen, Ole Amundsen, et al. Maximizing the spread of cascades using network design. *arXiv preprint arXiv:1203.3514*, 2012.

Sixie Yu, Kai Zhou, Jeffrey Brantingham, and Yevgeniy Vorobeychik. Computing equilibria in binary networked public goods games. In *Proc. of the AAAI Conference on Artificial Intelligence*, volume 34, pages 2310–2317, 2020.

Sixie Yu, David Kempe, and Yevgeniy Vorobeychik. Altruism design in networked public goods games. *arXiv preprint arXiv:2105.00505*, 2021.