Sybil-proof Answer Querying Mechanism

Yao Zhang, **Xiuzhen Zhang** and **Dengji Zhao**
Shanghai Engineering Research Center of Intelligent Vision and Imaging, ShanghaiTech University
{zhangyao1, zhangxzh1, zhaodj}@shanghaitech.edu.cn

Abstract

We study a question answering problem on a social network, where a requester is seeking an answer from the agents on the network. The goal is to design reward mechanisms to incentivize the agents to propagate the requester's query to their neighbours if they don't have the answer. Existing mechanisms are vulnerable to Sybil-attacks, i.e., an agent may get more reward by creating fake identities. Hence, we combat this problem by first proving some impossibility results to resolve Sybil-attacks and then characterizing a class of mechanisms which satisfy Sybil-proofness (prevents Sybil-attacks) as well as other desirable properties. Except for Sybil-proofness, we also consider cost minimization for the requester and agents' collusions.

1 Introduction

The development of online social networks has offered many opportunities for people to collaborate remotely in real time, such as P2P file-sharing network (e.g., BitTorrent) and Q&A platforms (e.g., Quora and Stack Overflow). Inspired by these applications, there are rich theoretical studies to look at the mechanism design problems on social networks [Rahman, 2009; Emek *et al.*, 2011; Li *et al.*, 2017].

In this paper, we focus on the answer/resource querying mechanisms via a social network where a requester is searching for a single answer from the network. Here we have two main challenges. The first is that the requester is only connected to a few agents (her neighbours) on the network and she needs to find a way to inform the other agents on the network if her neighbours do not have the answer. This can be achieved by incentivizing her neighbours to propagate the query to their neighbours (diffusion incentives). Kleinberg and Raghavan [2005] first formulated diffusion incentives in query networks. Later on, Arcaute *et al.* [2007] and Kota and Narahari [2010] studied the threshold of rewards needed to incentivize the query network. Kleinberg [2007] further discussed whether agents will act as the requester expects under different reward settings. Similar approach has been applied in advertising [Li and Shiu, 2012], auctions [Zhao *et al.*, 2018], recommendations [Margaris *et al.*, 2016] and others.

Once the first challenge is sovled, then an immediate new challenge is Sybil attacks, where an agent pretends to be multiple agents to gain more rewards [Conitzer *et al.*, 2010]. Sybil attack has been studied in many other applications such as multi-level marketing [Emek *et al.*, 2011; Drucker and Fleischer, 2012; Shen *et al.*, 2019], social choice [Conitzer and Yokoo, 2010] and blockchains [Babaioff *et al.*, 2012; Ersoy *et al.*, 2018]. For query networks, Sybil attack has been investigated under various settings. For example, Seuken and Parkes [2014] studied the trade-off between the transitive trust and Sybil attacks in P2P file-sharing networks. Chen *et al.* [2013] proposed a mechanism based on the query incentive network [Kleinberg and Raghavan, 2005] which achieves Sybil-proofness in expectation. Nath *et al.* [2012] identified a collusion-proof mechanism with approximated Sybil-proofness. However, all the existing work has only discussed Sybil-proofness in approximation or expectation on query networks.

Therefore, in this paper, we aim to solve both challenges in a dominant strategy implementation for query networks. We first demonstrate the difficulties to solve the challenges by proving the impossibility results. We then characterize a class of mechanisms, called double geometric mechanism (DGM), to satisfy the desirable properties. Except for the diffusion incentive and Sybil-proofness, we also look at the cost minimization problem for the requester and preventing agents' collusion. We will show that Sybil-proofness and collusion-proofness are not compatible.

1.1 Related Work

The query incentive network model was first proposed by Kleinberg and Raghavan [2005], where each agent in a d-ary tree network has the same probability to hold the answer and the actual query network is generated by a branching process (the cost in the query process is negligible). They considered a decentralized reward mechanism where each agent strategically chooses a fixed amount of reward to offer to her children if she can receive the answer from them. Different from the fixed reward mechanism, Cebrian *et al.* [2012] proposed a split contract mechanism in the same setting, which was motivated by the success of the winning strategy in the DARPA 2009 Network Challenge [Pickard *et al.*, 2011]. In their split contract mechanism, each agent should determine the splits of the reward received from her parent to offer to her

children. In both studies, they only considered the Nash equilibrium implementation, while we consider dominant strategy implementation here.

Moreover, the studies mentioned above mainly focused on propagating the query in the network and they did not consider agents' Sybil attacks. However, when we consider centralized query mechanisms where the reward distribution is decided by the requester, Sybil attack is a problem if the requester cannot verify their indenties [Douceur, 2002]. In the previous work, Chen *et al.* [2013] proved that split contract mechanisms cannot prevent Sybil attacks. They proposed direct referral mechanism to deal with this problem by allocating the majority of the reward to the winner (the agent who holds the answer) as well as her parent, i.e., the direct referral. However, the Sybil-proofness of the direct referral mechanism is only in expectation, which means that Sybil attack might be beneficial for an agent if the agent can acquire more knowledge about the network.

Again for Syblil-proofness, in the setting of dominant strategy implementation, Nath *et al.* [2012] studied the split contract mechanism design. They identified a set of desirable properties including Sybil-proofness and collusion-proofness and proved that no mechanism can satisfy them simultaneously under some conditions. They also examined the well-known geometric mechanism and showed that it satisfies collusion-proofness and approximated Sybil-proofness.

Different from the above, in this paper, we investigate centralized Sybil-proof reward mechanism design in the query network under dominant strategy implementation. We characterize a class of mechanisms to achieve Sybil-proofness and other properties in the query network.

The remainder of the paper is organized as follows. Section 2 describes the model of the query network and introduces the desirable properties of the reward mechanism. Then we prove the impossibility results in Section 3. Following that, we propose our double geometric mechanism and characterize its uniqueness under different properties in Section 4. Finally, we conclude and discuss the future work in Section 5.

2 The Model

We consider a question answering setting where a requester r is seeking the answer of a question from a set of agents. The agents are connected via their social connections such as friendship and r connects to a subset of them. r will first ask her neighbours for the answer and if her neighbours do not have the answer, r wants her neighbours to further propagate the question to their neighbours and so on. We assume that there is at least one agent who holds the answer and the answer is unique/verifiable. The goal of r is to design a reward mechanism to find the answer from the network. Ideally, we want each agent to offer the answer if she has, otherwise, to propagate the question to her neighbours if there is any. Formally, the propagation process will build a query tree $T = (V, E)$ rooted at r, where V is the set of all agents, including the requester r, who have been asked for the answer, and each edge $e = (i, j) \in E$ means that agent i has propagated the query to agent j and j has either offered the answer or propagated the question to her neighbours. For each agent i in T, let $p(i)$ be i's direct parent and $s(i)$ be i's direct children set.

We assume that there is one agent in T who offered the answer, which is called the winner, denoted by w. It is clear that $w \neq r$ and if there are multiple agents offered the answer, we choose the one with the smallest depth with random tie-breaking. We call the path from r to w a winning path, denoted by $p_w = (i_1, i_2, \ldots, i_n)$, where $i_1 \in s(r)$, $i_n = w$, $(i_j, i_{j+1}) \in E$ for all $1 \leq j < n$ and n is the length of the winning path.

Given the above setting, the requester needs to design a reward mechanism $M : \mathcal{T} \to \mathbb{R}^{V \setminus \{r\}}$ to incentivize the agents to find the answer, where $\mathcal{T}$ is the space of all possible resulting query trees and the output is the reward allocation for each player in the tree. Denote the reward allocated to agent i by x_i. In this paper, we focus on path mechanisms, which are mostly studied in the literature [Kleinberg and Raghavan, 2005; Chen *et al.*, 2013].

Definition 1. *A reward mechanism is called a **path mechanism** if*

1. *it only assigns non-zero rewards to the agents on the winning path, i.e., $x_i = 0$ for all $i \notin p_w$,*
2. *the reward distributed to an agent on the winning path only depends on her depth and the length of the path.*

*That is, a path mechanism M can be represented by a **reward function** $x : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ for the agents on the winning path, where the first parameter is the agent's depth and the second parameter is the length of the winning path.*

A path mechanism only rewards the agents on the winning path in the resulting query network. This will not weaken our results since the agents on the winning path made the actual contribution for seeking the answer. We assume that the cost for the query propagation is negligible as the propagation is often easy or automated [Yu and Singh, 2003]. Therefore, the literature has also focused on path mechanisms.

In the following, we define the desirable properties for path mechanisms. Firstly, it should incentivize agents to offer the answer or further propagate the query to all their neighbours; otherwise, the query will stop at the requester's neighbours. We call this property incentive compatibility.

Definition 2. *A path mechanism is **incentive compatible (IC)** if for all agent i, $x_i \geq x'_i$, where*

- *x_i is the reward i receives if i truthfully reports her answer, if i has the answer, or propagates the query to all her neighbours if i does not have the answer,*
- *x'_i is the reward i receives if i behaves differently.*

An incentive compatible mechanism guarantees that it will always find the answer if there is one in the network. Secondly, we also require that every agent is not forced to participate in the mechanism, i.e., their reward should not be negative, which is called individual rationality.

Definition 3. *A path mechanism is **individually rational (IR)** if for all agent $i \in p_w$, $x_i \geq 0$, i.e., $x(j, n) \geq 0$ for all $j, n \in \mathbb{Z}^+$, $j \leq n$. It is **strongly individually rational (SIR)** if for all agent $i \in p_w$, $x_i > 0$, i.e., $x(j, n) > 0$ for all $j, n \in \mathbb{Z}^+$, $j \leq n$.*

Notice that IR property can be easily satisfied even if the mechanism does not give any reward to any agent. Therefore, we also look for strongly IR path mechanisms, which at least reward something for every agent in the winning path. At the same time, the requester may also want to control the total reward distributed.

Definition 4. *A path mechanism is **budget constrained (BC)** if there exists a constant $B < \infty$ such that*

$$\sum_{i \in p_w} x_i = \sum_{j=1}^{n} x(j,n) \leq B$$

for all resulting query network T, for all winning path p_w of length n in T.

Next, we consider the most important property of Sybil-proofness, which requires that the mechanism should be resistant to Sybil-attacks. Since the resulting query network is a tree and we focus on path mechanisms, fake identities which are not on the winning path benefit nothing. So there is only one kind of meaningful Sybil-attacks for agent i, i.e., pretending to be multiple agents from agent $p(i)$ to agents in $s(i)$ (extending the paths from $p(i)$ to $s(i)$).

Definition 5. *A path mechanism is **Sybil-proof (SP)** if for any winning path $p_w = (i_1, i_2, \ldots, i_n)$, if an agent $i_j \in p_w$ extends p_w by making m copies of herself to get the new winning path $p'_w = (i_1, \ldots, i_{j-1}, i_j^0, i_j^1, \ldots, i_j^m, i_{j+1}, \ldots, i_n)$, we have:*

$$x_{i_j} = x(j,n) \geq \sum_{k=0}^{m} x'_{i_j^k} = \sum_{k=0}^{m} x(j+k, n+m) \quad (1)$$

Intuitively, if an agent pretends to be multiple agents along the winning path, the property of Sybil-proofness ensures that the total rewards she can get from the multiple agents is not more than what she will get originally. In our setting, Sybil-proofness can be easily verified as stated in Proposition 1.

Proposition 1. *A path mechanism is Sybil-proof if and only if*

$$x(j,n) \geq x(j,n+1) + x(j+1,n+1)$$

for all $j, n \in \mathbb{Z}^+$, $j \leq n$.

Proof. ("$\Rightarrow$") If a path mechanism is Sybil-proof, then in Inequality (1), let $m = 1$, it can be easily derived that $x(j,n) \geq x(j,n+1) + x(j+1,n+1)$ for all $j, n \in \mathbb{Z}^+$, $j \leq n$.

("$\Leftarrow$") If a path mechanism satisfies $x(j,n) \geq x(j,n+1) + x(j+1,n+1)$ for all $j, n \in \mathbb{Z}^+$, $j \leq n$, then

$$\begin{aligned} x(j,n) &\geq x(j,n+1) + x(j+1,n+1) \\ &\geq (x(j,n+2) + x(j+1,n+2)) \\ &\quad + (x(j+1,n+2) + x(j+2,n+2)) \\ &\geq \sum_{k=0}^{m} \binom{m}{k} x(j+k, n+m) \\ &\geq \sum_{k=0}^{m} x(j+k, n+m) \end{aligned}$$

for all $j, n, m \in \mathbb{Z}^+$, $j \leq n$. Hence, the mechanism is Sybil-proof. □

Sybil-proofness says that an agent cannot gain by making multiple copies of herself, while another property says multiple agents could also not collude together to receive a better reward, which is called collusion-proofness [Chen *et al.*, 2018]. We will show that these two properties cannot be satisfied together in general.

Definition 6. *A path mechanism is λ-**collusion proof (λ-CP)** ($\lambda \in \mathbb{Z}^+$ and $\lambda > 1$) if*

$$x(j,n) \leq \sum_{k=0}^{\lambda'-1} x(j+k, n+\lambda'-1) \quad (2)$$

for all $1 \leq \lambda' \leq \lambda$, all $j, n \in \mathbb{Z}^+$ such that $j \leq n - \lambda' + 1$. If λ-collusion proof holds for all $\lambda > 1$, i.e.,

$$x(j,n) \leq \sum_{k=0}^{m} x(j+k, n+m)$$

*for all $j, n, m \in \mathbb{Z}^+$ such that $j \leq n$, then, we call this mechanism **collusion-proof (CP)**.*

Intuitively, λ-collusion proofness indicates that any group of agents with size less than λ cannot get more reward if they pretend to be a single agent. Collusion proofness requires this to be held for all group sizes. In real-world applications, agents' ability to form a collusion is always limited as, for example, it is not easy to collude for agents far from each other in the network. Thus, λ-collusion proofness is an applicable approximation of collusion proofness in practice.

Lastly, to guarantee the reward distributed to agents on the winning path is sufficient, we require that each agent on the winning path can get at least a certain fraction of the winner's reward. This property is inspired by the efficiency of split contracts [Cebrian *et al.*, 2012].

Definition 7. *A path mechanism is ρ-**split secure (ρ-SS)**, $0 < \rho < 1$, if for all agent $i_j \in p_w \setminus \{w\}$, $x_{i_j} \geq \rho x_{i_{j+1}}$, i.e., $x(j,n) \geq \rho x(j+1,n)$ for all $j, n \in \mathbb{Z}^+$, $j < n$.*

Intuitively, the property of ρ-split security ensures that each agent on the winning path other than the winner has at least ρ fraction of the reward distributed to her children on the winning path. This guarantee will encourage agents to propagate the query in reality.

3 Impossibility Results

Before characterizing the reward mechanisms that satisfy the desirable properties, we prove several impossibility results in this section.

The first impossibility is that if strongly IR is required (i.e., each agent on the winning path is rewarded more than zero), Sybil-proofness and collusion-proofness cannot be held together.

Lemma 1. *An SIR path mechanism cannot be both Sybil-proof and λ-collusion-proof for all $\lambda \geq 3$.*

Proof. If such a mechanism exists, then according to the in-

equality in Proposition 1, let $m = \lambda - 1$ and we have

$$x(j,n) \geq \sum_{k=0}^{\lambda-1} \binom{\lambda-1}{k} x(j+k, n+\lambda-1)$$
$$\geq \sum_{k=0}^{\lambda-1} x(j+k, n+\lambda-1) \quad (3)$$

for any $j, n \in \mathbb{Z}^+$, $j \leq n-\lambda+1$. Together with Inequality (2), we know that

$$x(j,n) = \sum_{k=0}^{\lambda-1} x(j+k, n+\lambda-1)$$

Hence, the middle part of Inequality (3) is equal to both the left hand side and the right hand side.

$$\sum_{k=0}^{\lambda-1} \binom{\lambda-1}{k} x(j+k, n+\lambda-1) = \sum_{k=0}^{\lambda-1} x(j+k, n+\lambda-1)$$

Then, since $\lambda \geq 3$ and $\binom{\lambda-1}{0} = \binom{\lambda-1}{\lambda-1} = 1$,

$$\sum_{k=1}^{\lambda-2} \binom{\lambda-1}{k} x(j+k, n+\lambda-1) = \sum_{k=1}^{\lambda-2} x(j+k, n+\lambda-1)$$

from which we can derive that $x(j+k, n+\lambda-1) = 0$ for all $1 \leq k \leq \lambda-2$ since $\binom{\lambda-1}{k} > 1$ for all $1 \leq k \leq \lambda-2$. This is a contradiction with SIR. □

Following Lemma 1, we can conclude that under the requirement of SIR, being both Sybil-proof and collusion-proof is impossible for a path mechanism.

Proposition 2. *An SIR path mechanism cannot be both Sybil-proof and collusion-proof.*

However, we will show that the property of 2-collusion-proof, which says that two agents cannot collude to make a gain, can be satisfied with Sybil-proofness in Section 4.

The impossibility assumes strongly IR. However, even if we weaken the condition to be IR, Theorem 1 shows that the only mechanisms to satisfy both SP and CP are limited to very special mechanisms called two-headed mechanism.

Mechanism 1. *A path mechanism is a **two-headed mechanism** if its reward function satisfies*

$$x(j,n) = \begin{cases} a+b & \text{if } j=1 \text{ and } n=1 \\ a & \text{if } j=1 \text{ and } n>1 \\ b & \text{if } j=n \text{ and } n>1 \\ 0 & \text{otherwise} \end{cases}$$

where $a, b \geq 0$ are constants.

Intuitively, a two-headed mechanism only allocates positive rewards to the first agent and the winner on the winning path and all the other agents receive a zero reward.

Theorem 1. *A path mechanism is IR, SP and CP if and only if it is a two-headed mechanism.*

Proof. ("⇐") First, it is easy to show that a two-headed mechanism is IR, SP and CP since $x(j,n) \geq 0$ and

$$\sum_{j=1}^{n} x(j,n) = a+b$$

for all $j, n \in \mathbb{Z}^+$ and $j \leq n$.

("⇒") Then we show that these three properties determine a two-headed mechanism. From the definition of CP, we know that the mechanism is λ-CP for all $\lambda > 1$. Then from Lemma 1 we know that a mechanism with SP and λ-CP satisfies $x(j,n) = 0$ for all $1 < j < n$ and $n > 1$. Hence, from SP and 2-CP we can derive that

$$x(1,2) = x(1,3) + x(2,3) = x(1,3) = \cdots = x(1,n)$$

and

$$x(2,2) = x(2,3) + x(3,3) = x(3,3) = \cdots = x(n,n)$$

for all $n > 1$.

Hence, there exist two constants $a \geq 0$ and $b \geq 0$ such that $x(1,n) = a$, $x(n,n) = b$ for all $n > 1$ and $x(1,1) = x(1,2) + x(2,2) = a+b$. □

4 Double Geometric Mechanism

In this section, we characterize a class of reward mechanisms with the desirable properties. The mechanism we will characterize is called **Double Geometric Mechanism (DGM)**, which is a path mechanism defined by two parameters.

Mechanism 2. *A path mechanism is an (α, δ) double geometric mechanism ((α,δ)-DGM), for $0 < \alpha < 1$ and $\delta > 0$, if its reward function satisfies*

$$x(j,n) = (1-\alpha)^{j-1}\alpha^{n-j}\delta$$

for all $j, n \in \mathbb{Z}^+$ such that $j \leq n$.

Intuitively, (α,δ)-DGM has two fractions α^{n-j} and $(1-\alpha)^{j-1}$, which are controlled by the distance to the winner and the requester respectively. Note that if $\alpha < 1/2$, the reward is strictly monotone decreasing with the depth on the winning path, while if $\alpha > 1/2$, the reward is strictly monotone increasing with the depth on the winning path. Theorem 2 shows that (α,δ)-DGM can satisfy all the desirable properties defined in our model.

Theorem 2. *If a path mechanism is an (α,δ)-DGM with $\frac{\rho}{1+\rho} \leq \alpha < \frac{1}{2}$ $(0 < \rho < 1)$, then it is IC, SIR, BC, SP, 2-CP and ρ-SS.*

Proof. Suppose the winning path is $(i_1, i_2, \ldots, i_n)$ of length n when the agents behave truthfully. For an agent i_j with $1 \leq j < n$, $x_{i_j} = x(j,n)$, if she did not query all her neighbours, then she will be either on the winning path of length $\geq n$ or not on the winning path. By doing so, her reward is either $x(j,n+k) = (1-\alpha)^{j-1}\alpha^{n-j+k}\delta \leq (1-\alpha)^{j-1}\alpha^{n-j}\delta = x(j,n)$ for some $k \geq 0$, or zero, which is not better than behaving truthfully. For the winner $i_n = w$, if she did not provide the answer and further queried her children, then she would be either on the winning path of length $> n$ or not on the winning path, which gives her a reward either $x(n, n+$

$k) = (1-\alpha)^{n-1}\alpha^k\delta < (1-\alpha)^{n-1}\delta = x(n,n)$ for some $k \geq 1$ or zero. Hence, (α, δ)-DGM is IC.

For all $j, n \in \mathbb{Z}^+$, $j \leq n$, we have $x(j,n) = (1-\alpha)^{j-1}\alpha^{n-j}\delta > 0$. Hence, (α, δ)-DGM is SIR.

For all $j, n \in \mathbb{Z}^+$, $j \leq n$ and $0 < \alpha < \frac{1}{2}$, we have

$$\begin{aligned}\sum_{j=1}^{n} x(j,n) &= \sum_{j=1}^{n}(1-\alpha)^{j-1}\alpha^{n-j}\delta \\ &= \frac{\alpha^n\delta}{1-\alpha}\sum_{j=1}^{n}\left(\frac{1-\alpha}{\alpha}\right)^j \\ &= \frac{\alpha^n-(1-\alpha)^n}{2\alpha-1}\cdot\delta \leq \delta\end{aligned}$$

Hence, (α, δ)-DGM is BC.

For all $j, n \in \mathbb{Z}^+$, $j < n$, we have

$$\begin{aligned}x(j,n) &= \alpha x(j,n) + (1-\alpha)x(j,n) \\ &= \alpha(1-\alpha)^{j-1}\alpha^{n-j}\delta + (1-\alpha)(1-\alpha)^{j-1}\alpha^{n-j}\delta \\ &= (1-\alpha)^{j-1}\alpha^{n+1-j}\delta + (1-\alpha)^j\alpha^{n+1-(j-1)}\delta \\ &= x(j,n+1) + x(j+1,n+1)\end{aligned}$$

Hence, (α, δ)-DGM is 2-CP. According to Proposition 1, it is also SP.

Finally, for all $j, n \in \mathbb{Z}^+$, $j < n$, we have

$$\frac{x(j,n)}{x(j+1,n)} = \frac{\alpha}{1-\alpha} \geq \frac{\rho/(1+\rho)}{1-\rho/(1+\rho)} = \rho$$

Hence, (α, δ)-DGM is ρ-SS. □

(α, δ)-DGM satisfies all the desirable properties. Then we wonder are they the only mechanisms to satisfy these properties. Under some mild conditions, we will prove that (α, δ)-DGM is indeed the only mechanism to satisfy all the properties.

Note that if there are only two agents on the winning path, it is a very special case that has almost no constraints (the reward can be assigned arbitrarily), which suggests that it acts like an initial condition to the reward function. To satisfy the property of ρ-SS, we should have $x(1,2) \geq \rho x(2,2)$ and we let $x(1,2) = \rho x(2,2)$, which is the simplest way to construct the initial condition. We say a mechanism uses ρ-base condition if its reward function x satisfies $x(1,2) = \rho x(2,2)$. Theorem 3 proves that under the base condition, (α, δ)-DGM is the only kind of mechanism to satisfy all the properties.

Theorem 3. *If a path mechanism is IC, SIR, BC, SP, 2-CP, ρ-SS and uses ρ-base condition, then it is an (α, δ)-DGM with $\alpha = \frac{\rho}{1+\rho}$.*

Proof. First consider the value $x(j,n)$, $x(j+1,n)$, $x(j,n+1)$, $x(j+1,n+1)$ and $x(j+2,n+1)$ for some $j, n \in \mathbb{Z}^+$ and $j < n$. Suppose that $x(j,n) = \gamma x(j+1,n)$, $x(j,n+1) = \gamma_2 x(j+1,n+1)$ and $x(j+1,n+1) = \gamma_1 x(j+2,n+1)$.

Then according to the property of SP and 2-CP, we know that $x(j,n) = x(j,n+1) + x(j+1,n+1)$ and $x(j+1,n) = x(j+1,n+1) + x(j+2,n+1)$, from which we have

$$\begin{aligned}x(j,n) &= (1+\gamma_2)\gamma_1 x(j+2,n+1) \\ x(j+1,n) &= (1+\gamma_1)x(j+2,n+1)\end{aligned}$$

Hence, by $x(j,n) = \gamma x(j+1,n)$, we have

$$\gamma_2 = \gamma\left(1+\frac{1}{\gamma_1}\right) - 1$$

where according to the property of ρ-SS, we have $\gamma \geq \rho$, $\gamma_1 \geq \rho$ and $\gamma_2 \geq \rho$.

If $x(j,n) = \rho x(j+1,n)$, i.e., $\gamma = \rho$, then

$$\gamma_2 = \rho\left(1+\frac{1}{\gamma_1}\right) - 1 \geq \rho$$

which suggests that $\gamma_1 \leq \rho$. Hence $\gamma_1 = \rho$ and then $\gamma_2 = \rho$. Note that we have $x(1,2) = \rho x(2,2)$ as the base condition, so from above we have $x(1,3) = \rho x(2,3)$ and $x(2,3) = \rho x(3,3)$. Then, by induction, $x(j,n) = \rho x(j+1,n)$ holds for any $j, n \in \mathbb{Z}^+$ and $j < n$. Therefore, the following recursive relation holds for any $j, n \in \mathbb{Z}^+$ and $j \leq n$:

$$\begin{cases}x(j,n+1) = \frac{\rho}{1+\rho}x(j,n) \\ x(j+1,n+1) = \frac{1}{1+\rho}x(j,n)\end{cases}$$

Denote $\delta = x(1,1) > 0$, then we can derive that

$$x(j,n) = \left(\frac{1}{1+\rho}\right)^{j-1}\left(\frac{\rho}{1+\rho}\right)^{n-j}\delta = (1-\alpha)^{j-1}\alpha^{n-j}\delta$$

with $\alpha = \rho/(1+\rho)$ for any $j, n \in \mathbb{Z}^+$ and $j \leq n$, which is an (α, δ)-DGM. Also according to Theorem 2, the property of IC, SIR and BC will not be hurt. Therefore, these properties will uniquely determine the (α, δ)-DGM. □

So far, we have shown that (α, δ)-DGM is the only kind of mechanism to satisfy the properties under the ρ-base condition. Note that, without the ρ-base condition, we may get a different mechanism to satisfy all the properties, but it is still a DGM-like mechanism with a bounded difference. It suggests that the space of mechanisms removed by this base condition is also limited. Therefore, the base condition does not significantly hurt the generality of our result.

As Proposition 2 shows that SP and CP cannot be held together under SIR, the above mechanism can only satisfy 2-CP. In the following, we show how much extra gain a group of agents could get if they collude together and a weaker concept of CP. Furthermore, we will investigate the cost of the requester and show how to minimize it.

4.1 Approximation of Collusion-proofness

Definition 8. *A path mechanism is called β-**approximate collusion-proof** (β-**ACP**) if its reward function x satisfies*

$$x(j,n) \leq \beta\sum_{k=0}^{m} x(j+k, n+m)$$

for all $j, n, m \in \mathbb{Z}^+$, $j \leq n$.

Intuitively, the property of β-approximate collusion-proof ensures that if some agents pretend to be a single agent, they can achieve at most β times of their original reward. However, we show that a constant approximation cannot be achieved along with the other properties.

Proposition 3. *An IC, SIR, BC, SP, 2-CP and ρ-SS path mechanism with ρ-base condition cannot be $(1+\epsilon)$-ACP for any $\epsilon > 0$.*

Proof. Note that an IC, SIR, BC, SP, 2-CP and ρ-SS path mechanism with ρ-base condition must be an (α, δ)-DGM with $\alpha = \rho/(1+\rho)$. Then $(1+\epsilon)$-ACP implies that

$$(1-\alpha)^{j-1}\alpha^{n-j}\delta \leq (1+\epsilon)\sum_{k=0}^{m}(1-\alpha)^{j+k-1}\alpha^{n+m-j-k}\delta$$

from which we can derive that

$$(1-\alpha)^{m+1} - \alpha^{m+1} \geq \frac{1-2\alpha}{1+\epsilon}$$

However, for any given $\epsilon > 0$ and $0 < \alpha < 1/2$, there exists $N > 0$ such that for any $m > N$, $(1-\alpha)^{m+1} - \alpha^{m+1} < (1-2\alpha)/(1+\epsilon)$ since the left hand side approaches to 0 when m approaches to ∞. Hence, it cannot be $(1+\epsilon)$-ACP. □

On the other hand, an exponential approximation is easy to be achieved by an (α, δ)-DGM.

Theorem 4. *If a path mechanism is an (α, δ)-DGM with $0 < \alpha < \frac{1}{2}$, then it is 2^m-ACP, where m is the number of agents who collude together.*

Proof. If an (α, δ)-DGM is 2^m-ACP, it suggests that

$$(1-\alpha)^{j-1}\alpha^{n-j}\delta \leq 2^m\sum_{k=0}^{m}(1-\alpha)^{j+k-1}\alpha^{n+m-j-k}\delta$$

from which we can derive the equivalent inequality that

$$(1-\alpha)^{m+1} - \alpha^{m+1} \geq \frac{1-2\alpha}{2^m}$$

or being rearranged as

$$(1-\alpha)^{m+1} - \frac{1-\alpha}{2^m} \geq \alpha^{m+1} - \frac{\alpha}{2^m}$$

Notice that for function $f(y) = y\left(y^m - \frac{1}{2^m}\right)$, $f(y) < 0$ for any $0 < y < 1/2$ and $f(y) > 0$ for any $1/2 < y < 1$. Since $0 < \alpha < 1/2$, the above inequality always holds. □

4.2 Cost Minimization

Next, we consider the case where the requester is willing to minimize her cost to query the answer.

Definition 9. *A path mechanism is **of minimum cost** over a class of path mechanisms $\mathcal{X}$ if its reward function x satisfies*

$$x \in \arg\min_{x\in\mathcal{X}} \sum_{j=1}^{n} x(j,n)$$

for all $n \in \mathbb{Z}^+$.

Intuitively, a path mechanism is of minimum cost if it can minimize the total reward distributed to the agents on the winning path. Notice that when we minimize the cost, we consider time-critical mechanisms, which are essential to the real-world application [Pickard *et al.*, 2011]. A path mechanism is time-critical if the winner always takes the maximum reward, i.e., $x(n,n) = \max_j x(j,n)$ for any $n \geq 1$. Now we show that the (α, δ)-DGM also characterizes the space of time-critical mechanisms of minimum cost.

Theorem 5. *A path mechanism is of minimum cost over a class of time-critical path mechanisms that satisfy IC, SIR, BC, SP, 2-CP, ρ-SS and have $x(1,1) = \delta$ if and only if it is an (α, δ)-DGM with $\alpha = 1/2$.*

Proof. First we show a lower bound of the total cost before we prove the statement. According to the property of SP and 2-CP, denote $\sum_{j=1}^{n} x(j,n) = R_n$, then we have

$$2R_{n+1} = R_n + (x(1,n+1) + x(n+1,n+1))$$

with $R_1 = R_2 = \delta$. Hence, to minimize the total cost is to minimize the term $x(1,n+1) + x(n+1,n+1)$. Suppose that $x(1,2) = \gamma x(2,2)$, $x(1,3) = \gamma_2 x(2,3)$ and $x(2,3) = \gamma_1 x(3,3)$ where $\rho \leq \gamma, \gamma_1, \gamma_2 \leq 1$ with $\gamma = \rho + \epsilon$. Then similar to the process in Theorem 3, we can derive that $\gamma_2 = \gamma(1 + 1/\gamma_1) - 1$ and $\gamma_1 \leq (\rho+\epsilon)/(1-\epsilon)$. Hence,

$$x(1,3) + x(3,3) \geq \left(1 - \frac{2(\rho+\epsilon)}{(\rho+1)(1+\rho+\epsilon)}\right)\delta$$

As we want to minimize the cost, the minimum value will be reached if $\gamma = \gamma_1 = \gamma_2 = 1$, which means $x(1,2) = x(2,2)$ and $x(1,3) = x(2,3) = x(3,3)$. By induction, if for some n, $x(j,n) = x(j+1,n)$ for any $j < n$, suppose $x(j,n+1) = \gamma_j x(j+1,n+1)$ for any $j \leq n$, then we have $1 + \gamma_n = \gamma_n(1+\gamma_{n-1}) = \cdots = \gamma_n \cdots \gamma_2(1+\gamma_1)$ with $\rho \leq \gamma_j \leq 1$. So that the solution could only be $\gamma_1 = \cdots = \gamma_n = 1$. Therefore, $x(j,n) = x(j+1,n)$ for any $j, n \in \mathbb{Z}^+$ and $j < n$.

("⇐") First, an (α, δ)-DGM is IC, SIR, BC, SP, 2-CP, ρ-SS and have $x(1,1) = \delta$ by Theorem 2. Then, for an (α, δ)-DGM with $\alpha = 1/2$, the total reward distributed is

$$\sum_{j=1}^{n} x(j,n) = \sum_{j=1}^{n}(1-\alpha)^{j-1}\alpha^{n-j}\delta = n\left(\frac{1}{2}\right)^{n-1}\delta$$

for any $n \geq 1$. Besides, from the above we know $2R_n \geq R_{n-1} + \delta/2^{n-2}$ for $n \geq 3$ and $R_2 = \delta$. Then $R_n \geq (n\delta)/2^{n-1}$. Since (α, δ)-DGM with $\alpha = 1/2$ meets this lower bound, then it is the mechanism of minimum cost.

("⇒") By the equation of $x(j,n) = x(j+1,n)$ and $x(1,1) = \delta$, we can derive that $x(j,n) = \delta/2^{n-1}$, which is the same as for the $(1/2, \delta)$-DGM. Therefore, the mechanism of minimum cost over these properties uniquely determines the $(1/2, \delta)$-DGM. □

5 Conclusion

We have investigated Sybil-proof answer querying mechanisms on networks in a dominant strategy implementation. We proposed a class of double geometric mechanisms (DGM) to against Sybil attacks and characterized its uniqueness under other important properties such as IC, IR and 2-CP. We also characterized the mechanisms for minimizing the requester's reward expenses and illustrated the performance of the mechanisms in terms of the approximation of collusion-proofness. There are several other interesting aspects worth further investigation. There is a gap between constant approximation and exponential approximation of collusion-proofness. The characterization of Sybil-proof mechanisms in a general directed graph is also missing.

References

[Arcaute *et al.*, 2007] Esteban Arcaute, Adam Kirsch, Ravi Kumar, David Liben-Nowell, and Sergei Vassilvitskii. On threshold behavior in query incentive networks. In *Proceedings of the 8th ACM conference on Electronic commerce*, pages 66–74. ACM, 2007.

[Babaioff *et al.*, 2012] Moshe Babaioff, Shahar Dobzinski, Sigal Oren, and Aviv Zohar. On bitcoin and red balloons. In *Proceedings of the 13th ACM conference on electronic commerce*, pages 56–73. ACM, 2012.

[Cebrian *et al.*, 2012] Manuel Cebrian, Lorenzo Coviello, Andrea Vattani, and Panagiotis Voulgaris. Finding red balloons with split contracts: robustness to individuals' selfishness. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 775–788. ACM, 2012.

[Chen *et al.*, 2013] Wei Chen, Yajun Wang, Dongxiao Yu, and Li Zhang. Sybil-proof mechanisms in query incentive networks. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 197–214. ACM, 2013.

[Chen *et al.*, 2018] Peng-Peng Chen, Hai-Long Sun, Yi-Li Fang, and Jin-Peng Huai. Collusion-proof result inference in crowdsourcing. *Journal of Computer Science and Technology*, 33(2):351–365, 2018.

[Conitzer and Yokoo, 2010] Vincent Conitzer and Makoto Yokoo. Using mechanism design to prevent false-name manipulations. *AI magazine*, 31(4):65–78, 2010.

[Conitzer *et al.*, 2010] Vincent Conitzer, Nicole Immorlica, Joshua Letchford, Kamesh Munagala, and Liad Wagman. False-name-proofness in social networks. In *International Workshop on Internet and Network Economics*, pages 209–221. Springer, 2010.

[Douceur, 2002] John R Douceur. The sybil attack. In *International workshop on peer-to-peer systems*, pages 251–260. Springer, 2002.

[Drucker and Fleischer, 2012] Fabio A Drucker and Lisa K Fleischer. Simpler sybil-proof mechanisms for multi-level marketing. In *Proceedings of the 13th ACM conference on Electronic commerce*, pages 441–458. ACM, 2012.

[Emek *et al.*, 2011] Yuval Emek, Ron Karidi, Moshe Tennenholtz, and Aviv Zohar. Mechanisms for multi-level marketing. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 209–218. ACM, 2011.

[Ersoy *et al.*, 2018] Oğuzhan Ersoy, Zhijie Ren, Zekeriya Erkin, and Reginald L Lagendijk. Transaction propagation on permissionless blockchains: incentive and routing mechanisms. In *2018 Crypto Valley Conference on Blockchain Technology (CVCBT)*, pages 20–30. IEEE, 2018.

[Kleinberg and Raghavan, 2005] Jon Kleinberg and Prabhakar Raghavan. Query incentive networks. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 132–141. IEEE, 2005.

[Kleinberg, 2007] Jon Kleinberg. Cascading behavior in networks: Algorithmic and economic issues. *Algorithmic game theory*, 24:613–632, 2007.

[Kota and Narahari, 2010] Nagaraj Kota and Y Narahari. Threshold behavior of incentives in social networks. In *Proceedings of the 19th ACM international conference on Information and knowledge management*, pages 1461–1464. ACM, 2010.

[Li and Shiu, 2012] Yung-Ming Li and Ya-Lin Shiu. A diffusion mechanism for social advertising over microblogs. *Decision Support Systems*, 54(1):9–22, 2012.

[Li *et al.*, 2017] Bin Li, Dong Hao, Dengji Zhao, and Tao Zhou. Mechanism design in social networks. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

[Margaris *et al.*, 2016] Dionisis Margaris, Costas Vassilakis, and Panagiotis Georgiadis. Recommendation information diffusion in social networks considering user influence and semantics. *Social Network Analysis and Mining*, 6(1):108, 2016.

[Nath *et al.*, 2012] Swaprava Nath, Pankaj Dayama, Dinesh Garg, Yadati Narahari, and James Zou. Mechanism design for time critical and cost critical task execution via crowdsourcing. In *International Workshop on Internet and Network Economics*, pages 212–226. Springer, 2012.

[Pickard *et al.*, 2011] Galen Pickard, Wei Pan, Iyad Rahwan, Manuel Cebrian, Riley Crane, Anmol Madan, and Alex Pentland. Time-critical social mobilization. *Science*, 334(6055):509–512, 2011.

[Rahman, 2009] Muntasir Raihan Rahman. A survey of incentive mechanisms in peer-to-peer systems. *Cheriton School of Computer Science, University of Waterloo, Tech. Rep. CS-2009-22*, 2009.

[Seuken and Parkes, 2014] Sven Seuken and David C Parkes. Sybil-proof accounting mechanisms with transitive trust. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 205–212. International Foundation for Autonomous Agents and Multiagent Systems, 2014.

[Shen *et al.*, 2019] Wen Shen, Yang Feng, and Cristina V Lopes. Multi-winner contests for strategic diffusion in social networks. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI-19). AAAI Press*, 2019.

[Yu and Singh, 2003] Bin Yu and Munindar P Singh. Searching social networks. In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, pages 65–72. ACM, 2003.

[Zhao *et al.*, 2018] Dengji Zhao, Bin Li, Junping Xu, Dong Hao, and Nicholas R Jennings. Selling multiple items via social networks. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pages 68–76. International Foundation for Autonomous Agents and Multiagent Systems, 2018.