Strategy-proof Mechanisms for Facility Location Games with Many Facilities⋆

Bruno Escoffier[1], Laurent Gourvès[2,1], Nguyen Kim Thang[1], Fanny Pascual[3], and Olivier Spanjaard[3]

1. Université Paris-Dauphine, LAMSADE, F-75775 Paris, France
2. CNRS, FRE 3234, F-75775 Paris, France
3. Université Pierre et Marie Curie, LIP6, F-75005 Paris, France

Abstract. This paper is devoted to the location of public facilities in a metric space. Selfish agents are located in this metric space, and their aim to minimize their *own* cost, which is the distance from their location to the nearest facility. A central authority has to locate the facilities in the space, but she is ignorant of the true locations of the agents. The agents will therefore report their locations, but they may lie if they have an incentive to do it. We consider two social costs in this paper: the sum of the distances of the agents to their nearest facility, or the maximal distance of an agent to her nearest facility. We are interested in designing *strategy-proof* mechanisms that have a small approximation ratio for the considered social cost. A mechanism is strategy-proof if no agent has an incentive to report false information. This setting has been previously studied for extreme cases where one and two facilities are to be opened. In this paper, we consider the other extreme case: given n agents, we wish to design strategy-proof mechanisms to locate $n - 1$ facilities. We study this problem in the general metric and in the tree metric spaces. We provide lower and upper bounds on the approximation ratio of deterministic and randomized strategy-proof mechanisms. Our work could be considered as one step toward the general case.

1 Introduction

We study *Facility Location Games* that model the following problem in economics. Consider installation of public service facilities such as hospitals or libraries within the region of a city, represented by a metric space. The authority announces that some locations will be chosen within the region and run a survey over the population; each inhabitant may declare the spot in the region that she prefers some facility to be opened at. Every inhabitant wishes to minimize her individual distance to the closest facility, possibly by misreporting her preference to the authorities. The goals of the authority are twofold: avoiding such misreports and minimizing some social objectives. The authority needs to design a *mechanism*, that maps the reported preferences of inhabitants to a set

---

of locations where the facilities will be opened at, to fulfill the purposes. The mechanism must be *strategy-proof*, i.e., it ensures that no inhabitant can benefit by misreporting her preference. At the same time, the mechanism should guarantee a reasonable approximation to the optimal social cost. The model has many applications in telecommunication networks where locations may be easily manipulated by reporting false IP addresses, false routers, etc.

1.1 Facility Location Games

We consider a metric space (Ω, d), where $d : \Omega \times \Omega \to \mathbb{R}$ is the metric function. Some usual metrics are the line, circle and tree metrics where the underlying spaces are an infinite line, a circle and an infinite tree, respectively. The distance between two positions in such metrics is the length of the shortest path connecting those positions. Let n be the number of agents, each agent i has a location $x_i \in \Omega$. A *location profile (or strategy profile)* is a vector $\mathbf{x} = (x_1, \ldots, x_n) \in \Omega^n$. Let k be the number of facilities that will be opened. A *deterministic* mechanism is a mapping f from the set of location profiles Ω^n to k locations in Ω. Given a reported location profile $\mathbf{x}$ the mechanism's output is $f(\mathbf{x}) \in \Omega^k$ and the individual cost of agent i under mechanism f and profile $\mathbf{x}$ is the distance from its location to the closest facility, denoted by $c_i(f, \mathbf{x})$:

$$c_i(f, \mathbf{x}) := d(f(\mathbf{x}), x_i) := \min\{d(F, x_i) : F \in f(\mathbf{x})\}$$

A *randomized* mechanism is a function f from the set of location profiles to $\Delta(\Omega^k)$ where $\Delta(\Omega^k)$ is the set of probability distributions over Ω^k. The cost of agent i is now the expected distance from its location to the closest facility over such distribution:

$$c_i(f, \mathbf{x}) := \mathbb{E}\left[d(f(\mathbf{x}), x_i)\right] := \mathbb{E}\left[\min\{d(F, x_i) : F \in f(\mathbf{x})\}\right]$$

We are interested in two standard *social objectives*: (i) the *utilitarian objective* defined as the total individual costs (total individual expected cost for a randomized mechanism), i.e., $C(f, \mathbf{x}) = \sum_{i=1}^{n} c_i(f, \mathbf{x})$; and (ii) the *egalitarian objective* defined as the maximal individual cost (expected maximal individual cost for a randomized mechanism), i.e., $C(f, \mathbf{x}) = \mathbb{E}\left[\max_{1 \le i \le n} d(f(\mathbf{x}), x_i)\right]$. This is thus simply $\max_{1 \le i \le n} c_i(f, \mathbf{x})$ for deterministic mechanisms.

We say that a mechanism f is *r-approximate* with respect to profile $\mathbf{x}$ if

$$C(f, \mathbf{x}) \le r \cdot OPT(\mathbf{x})$$

where $OPT(\mathbf{x})$ is the social cost of an optimal facility placement (for the egalitarian or utilitarian social cost). Note that since for a randomized mechanism the social cost is the expectation of the social cost on each chosen set of locations, there always exists an optimal deterministic placement.

We will be concerned with *strategy-proof* (SP) mechanisms, which render truthful revelation of locations a dominant strategy for the agents.

Definition 1. (Strategyproofness) *Let* $\mathbf{x} = (x_1, \ldots, x_n)$ *denote the location profile of n agents over the metric space (Ω, d). A mechanism f is strategy-proof (SP) if for every agent $1 \le i \le n$ and for every location $x_i' \in \Omega$, $c_i(f, (x_i', x_{-i})) \ge c_i(f, \mathbf{x})$ where x_{-i} denotes the locations of the agents other than i in $\mathbf{x}$.*

1.2   Previous works

The facility locations game where only *one* facility will be opened is widely-studied in economics. On this topic, Moulin [1] characterized all strategy-proof mechanisms in the line metric space. Subsequently, Schummer and Vohra [2] gave a characterization of strategy-proof mechanisms for the circle metric space. More recently, Procaccia and Tennenholtz [3] initiated the study of approximating an optimum social cost under the constraint of strategy-proofness. They studied deterministic and randomized mechanisms on the line metric space with respect to the utilitarian and egalitarian objectives. Several (tight) approximation bounds for strategy-proof mechanisms were derived in their paper. For general metric space, Alon et al. [4] and Nguyen Kim [5] proved randomized tight bounds for egalitarian and utilitarian objectives, respectively.

Concerning the case where *two* facilities are opened, Procaccia and Tennenholtz [3] derived some strategy-proof mechanisms with guaranteed bounds in the line metric space for both objectives. Subsequently, Lu et al. [6] proved tight lower bounds of strategy-proof mechanisms in the line metric space with respect to the utilitarian objective. Moreover, they also gave a randomized strategy-proof mechanism, called Proportional Mechanism, that is 4-approximate for general metric spaces. It is still unknown whether there exists a deterministic strategy-proof mechanism with bounded approximation ratio in a general metric space.

Due to the absence of any positive result on the approximability of multiple facility location games for *more than two* facilities, Fotakis and Tzamos [7] considered a variant of the game where an authority can impose on some agents the facilities where they will be served. With this restriction, they proved that the Proportional Mechanism is strategy-proof and has an approximation ratio linear on the number of facilities.

1.3   Contribution

Prior to our work, only extreme cases of the game where the authority opens one or two facilities have been considered. No result, positive or negative, has been known for the game with three or more facilities. In this paper, we consider another extreme case of the game where $k = n - 1$ facilities will be opened (note that the case when n facilities are open is trivial, opening a facility at each player's location is strategy-proof and optimal). Toward the general number of facilities, we need to understand and solve the extreme cases of the problem. This paper could be considered as one step toward this goal.

In this paper, we study strategy-proof mechanisms for the game with n agents and $n - 1$ facilities in a general metric space and in a tree metric space. Our

Objective	Tree metric space	General metric space
Utilitarian	UB: $n/2$ (rand)	UB: $n/2$ (rand)
	LB: 3/2 (det), 1.055 (rand)	LB: 3 (det), 1.055 (rand)
Egalitarian	UB: 3/2 (rand)	UB: n (rand)
	LB: 3/2 (rand) [3]	LB: 2 (det)

Table 1. Summary of our results. In a cell, UB and LB mean the upper and lower bounds on the approximation ratio of strategy-proof mechanisms. The abbreviations *det* and *rand* refer to deterministic and randomized strategy-proof mechanisms, respectively.

main results are the following ones. For general metric spaces, we give a randomized strategy-proof mechanism, called *Inversely Proportional Mechanism*, that is an $n/2$-approximation for the utilitarian objetctive and an n-approximation for the egalitarian one. For tree metric spaces, we present another randomized strategy-proof mechanism that particularly exploit the property of the metric. This mechanism is also an $n/2$-approximation under the utilitarian objective but it induces a 3/2-approximation (tight bound) under the egalitarian objective.

Besides, several lower bounds on the approximation ratio of deterministic/randomized strategy-proof mechanisms are derived (see Table 1 for a summary). We proved that any randomized strategy-proof mechanism has ratio at least 1.055 even in the tree metric space. The interpretation of this result is that no mechanism, even randomized one, is both socially optimal and strategy-proof. Moreover, deterministic lower bounds of strategy-proof mechanisms are shown to be: at least 3/2 in tree metric space, utilitarian objective; at least 3 in a general metric space, utilitarian objective; and at least 2 in general metric space, egalitarian objective. Note that the lower bounds given for a the tree metric space hold even for a line metric space.

Organization We study the performance of randomized SP mechanisms in general metric spaces and in tree metric spaces in Section 2, and Section 3, respectively. Due to the lack of space, some claims are only stated or partially proved. We refer the reader to the appendix for the complete proofs of all claims.

2   SP mechanisms for general metric spaces

2.1   Inversely Proportional Mechanism

Consider the setting of n agents whose true locations are $\mathbf{x} = (x_1, \ldots, x_n)$. For each location profile $\mathbf{y} = (y_1, \ldots, y_n)$, define $P_i(\mathbf{y})$ as the placement of $(n-1)$ facilities at the reported locations of all but agent i, i.e., $P_i(\mathbf{y}) = \{y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n\}$. Moreover, $d(y_i, P_i(\mathbf{y}))$ is the distance between y_i and her closest location in $P_i(\mathbf{y})$. The idea of the mechanism is to choose with a given probability a location y_i where no facility is open (and to put $n-1$ facilities precisely on the $n-1$ locations of the other agents), i.e., to choose with a given probability the placement $P_i(\mathbf{y})$. The main issue is to find suitable probabilities

such that the mechanism is strategy-proof, and such that the expected cost is as small as possible.

Inversely proportional mechanism Let $\mathbf{y}$ be a reported location profile. If there are at most $(n-1)$ distinct locations in profile $\mathbf{y}$ then open facilities at the locations in $\mathbf{y}$. Otherwise, choose placement $P_i(\mathbf{y})$ with probability

$$p_i(\mathbf{y}) = \frac{\frac{1}{d(y_i, P_i(\mathbf{y}))}}{\sum_{j=1}^{n} \frac{1}{d(y_j, P_j(\mathbf{y}))}}$$

Lemma 1. *The Inversely Proportional Mechanism is strategy-proof in a general metric space.*

Sketch of the proof. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be the true location profile of the agents, and let $d_j := d(x_j, P_j(\mathbf{x}))$ for $1 \leq j \leq n$.

If there are at most $(n-1)$ distinct locations in profile $\mathbf{x}$ then the mechanism locates one facility on each position: no agent has incentive to misreports its location. In the sequel, we assume that all the agent locations in $\mathbf{x}$ are distinct. If all the agents report truthfully their locations, the cost of agent i is

$$c_i := c_i(f, \mathbf{x}) = \sum_{j=1}^{n} p_j(\mathbf{x}) \cdot d(x_i, P_j(\mathbf{x})) = p_i(\mathbf{x}) \cdot d_i = \frac{1}{\sum_{j=1}^{n} 1/d_j}$$

Thus $c_i < d_i$. Let us now suppose that i misreports its location and bids x_i'. Let $\mathbf{x}' = (x_i', x_{-i})$ be the location profile when i reports x_i' and the other agents report truthfully their locations. Let $d_j' = d(P_j(x_j, \mathbf{x}'))$ for $j \neq i$ and $d_i' = d(P_i(x_i', \mathbf{x}'))$. We will prove that $c_i' := c_i(f, \mathbf{x}') \geq c_i$. The new cost of agent i is:

$$c_i' = \sum_{j=1}^{n} p_j(\mathbf{x}') \cdot d(x_i, P_j(\mathbf{x}')) \geq p_i(\mathbf{x}') \cdot d_i + (1 - p_i(\mathbf{x}')) \min\{d_i, d(x_i, x_i')\}$$

where the inequality is due to the fact that in $P_j(\mathbf{x}')$ (for $j \neq i$), agent i can choose either some facility in $\{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n\}$ or the facility opened at x_i'. Define $T := \{j : d_j' \neq d_j, j \neq i\}$. Note that

$$p_i(\mathbf{x}') = \frac{1/d_i'}{\sum_{j \notin T} 1/d_j + \sum_{j \in T} 1/d_j' + 1/d_i'}$$

Let $e := d(x_i, x_i')$. Remark that i has no incentive to report its location x_i' in such a way that $e \geq d_i$ since otherwise $c_i' \geq p_i(\mathbf{x}') \cdot d_i + (1 - p_i(\mathbf{x}'))d_i = d_i > c_i$. In the sequel, consider $e < d_i$. In this case,

$$c_i' \geq p_i(\mathbf{x}') \cdot d_i + (1 - p_i(\mathbf{x}')) \cdot e$$

We also show that $e \geq |d_i' - d_i|$ by using the triangle inequality. Then, by considering two cases (whether $\sum_{j \in T} \frac{1}{d_j'}$ is larger than $\sum_{j \in T} \frac{1}{d_j}$ or not), we show that in both case $c_i' \geq c_i$: any agent i has no incentive to misreport its location, i.e., the mechanism is strategy-proof.   □

Theorem 1. *The Inversely Proportional Mechanism is strategy-proof, an $n/2$-approximation with respect to the utilitarian social cost and an n-approximation with respect to the egalitarian one. Moreover, there exists an instance in which the mechanism gives the approximation ratio at least $\frac{n}{2} - \epsilon$ for the utilitarian social cost, and $n - \epsilon$ for the egalitarian one, where $\epsilon > 0$ is arbitrarily small.*

Proof. By the previous lemma, the mechanism is strategy-proof. We consider the approximation ratio of this mechanism. Recall that $\mathbf{x} = (x_1, \ldots, x_n)$ is the true location profile of the agents. Let $P_i := P_i(\mathbf{x})$, $d_i := d(x_i, P_i)$ and $p_i = p_i(\mathbf{x})$. Let $\ell := \arg\min\{d_i : 1 \leq i \leq n\}$. For the egalitarian social cost, due to the triangle inequality at least one agent has to pay $d_\ell/2$, while the optimal solution for the utilitarian objective has cost d_ℓ (placement P_ℓ for instance).

The mechanism chooses placement P_i with probability p_i. In P_i, agent i has cost d_i and the other agents have cost 0. Hence, the social cost induced by the mechanism (in both objectives) is $\sum_j p_j(\mathbf{x})d_j = \frac{n}{\sum_j 1/d_j}$. For the utilitarian objective, the approximation ratio is $\frac{n}{d_\ell \sum_j 1/d_j} < \frac{n}{2}$ since in the sum of the denominator, there are two terms $1/d_\ell$. Similarly, it is at most $\frac{2n}{d_\ell \sum_j 1/d_j} < n$ for the egalitarian objective.

We describe an instance on a line metric space in which the bounds $n/2$ and n are tight. Let M be a large constant. Consider the instance on a real line in which $x_1 = 1, x_2 = 2, x_{i+1} = x_i + M$ for $2 \leq i \leq n$. We get $d_1 = d_2 = 1$ and $d_i = M$ for $3 \leq i \leq n$. An optimal solution chooses to put a facility in each x_i for $i \geq 2$ and to put the last one in the middle of $[x_1, x_2]$. Its social cost is 1 for the utilitarian objective and $1/2$ for the egalitarian one. The cost (in both objectives) of the mechanism is

$$\frac{n}{\sum_{j=1}^{n} 1/d_j} = \frac{n}{2 + (n-2)/M} = \frac{nM}{2M + n - 2}$$

Hence, for any $\epsilon > 0$, one can choose M large enough such that the approximation ratio is larger than $\frac{n}{2} - \epsilon$ for the utilitarian objectif and to $n - \epsilon$ for the egalitarian one. $\square$

2.2   Lower bounds on the approximation ratio for SP mechanisms

Proposition 1. *Any deterministic strategy-proof mechanism has approximation ratio at least $3 - 2\epsilon$ for the utilitarian objective and $2 - 2\epsilon$ for the egalitarian objective where $\epsilon > 0$ arbitrarily small.*

Proof. We consider the metric space induced by the graph in Figure 1. Note that this is a discrete space where agents and possible locations for facilities are restricted to be on vertices of the graph, i.e., $\Omega = V$. There are three agents in the game and two facilities to be opened. Let f be a deterministic strategy-proof mechanism. Let $\mathbf{x}$ be a profile where $x_1 = A_0, x_2 = B_0, x_3 = C_0$. For any (deterministic) placement of two facilities, there is one agent with cost at least
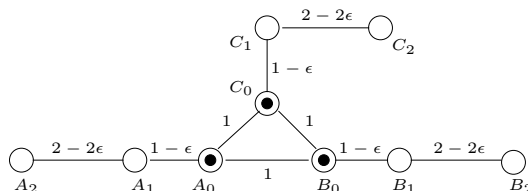
Fig. 1. Graph metric that gives a lower bound on the ratio of strategy-proof mechanisms in a general metric space (dots are the agents' locations in profile $\mathbf{x}$).

1. By symmetry of the graph as well as profile $\mathbf{x}$, suppose that agent 1 has cost at least 1.

Consider another profile $\mathbf{y}$ where $y_1 = A_1, y_2 = B_0, y_3 = C_0$ ($\mathbf{y}$ and $\mathbf{x}$ only differ on the location of agent 1). In this profile, no facility is opened neither at A_0 nor at A_1 since otherwise agent 1 in profile $\mathbf{x}$ could report its location as being A_1 and reduce its cost from 1 to $1 - \epsilon$ or 0. We study two cases: (i) in profile $f(\mathbf{y})$, there is a facility opened at A_2; and (ii) in profile $f(\mathbf{y})$, no facility is opened at A_2.

In the former, a facility is opened at A_2, no facility is opened at A_0, A_1. For the egalitarian objective, the social cost is at least $2 - 2\epsilon$. For the utilitarian objective, the total cost of agents 2 and 3 is at least 1 and the cost of agent 1 is $2 - 2\epsilon$, that induces a social cost at least $3 - 2\epsilon$. An optimal solution has cost 1 (for both objective) by opening a facility at A_1 and a facility at B_0.

In the latter, the cost of agent 1 is at least $2 - \epsilon$ (since no facility is opened at A_0, A_1, A_2). Consider a profile $\mathbf{z}$ similar to $\mathbf{y}$ but the location of agent 1 is now at A_2. By strategy-proofness, no facility is opened at A_0, A_1, A_2 in $f(\mathbf{z})$ (since otherwise, agent 1 in profile $\mathbf{y}$ can decrease its cost by reporting its location as A_2). So, the social cost induced by mechanism f in $\mathbf{z}$ is at least $4 - 3\epsilon$ (for both objectives), while optimal is 1 (for both objectives) by placing a facility at A_2 and other at B_0.

Therefore, in any case, the approximation ratio of mechanism f is at least $3 - 2\epsilon$ for the utilitarian objective and $2 - 2\epsilon$ for the egalitarian objective.   □

3   Randomized SP mechanisms on trees

We study in this section the infinite tree metric. This is a generalization of the (infinite) line metric, where the topology is now a tree. Infinite means that, like in the line metric, branches of the tree are infinite. As for the line metric, the locations (reported by agents or for placing facilities) might be anywhere on the tree. We first devise a randomized mechanism. To achieve this, we need to build a partition of the tree into subtrees that we call components, and to associate a status even or odd to each component. This will be very useful in particular to show that the mechanism is strategy-proof. In the last part of this section,

we propose a lower bound on the approximation ratio of any strategy-proof mechanism.

3.1 Preliminary tool: partition into odd and even components

Partition procedure Given a tree $\mathcal{T}$ and a set of vertices V on this tree, we partition $\mathcal{T}$ into subtrees with respect to V. For the ease of description, consider also some virtual vertices, named ∞, which represent the extremities of the branches in $\mathcal{T}$. We say that two vertices i and j are *neighbor* if the unique path in $\mathcal{T}$ connecting i and j contains no other vertex ℓ. A *component* T_t is a region of the tree delimited by a *maximal* set of pairwise neighbor vertices (see below for an illustration). The maximality is in the sense of inclusion: T_t is maximal means that there is no vertex $i \notin T_t$ such that vertex i is a neighbor of all vertices in T_t. The set $\{T_1, \ldots, T_m\}$ of all components is a cover of the tree $\mathcal{T}$. Note that a vertex i can appear in many sets T_t. As $\mathcal{T}$ is a tree, the set of all T_t's is well and uniquely defined.

For instance, in Figure 2, the components are the subtrees delimited by the following sets of vertices: $\{1, 2, 3\}$, $\{1, 4\}$, $\{2, 5\}$, $\{2, 6\}$, $\{6, 10\}$, $\{4, 7\}$, $\{4, 8, 9\}$, $\{3, \infty\}$, $\{5, \infty\}$, $\{7, \infty\}$, $\{8, \infty\}$, $\{9, \infty\}$, $\{10, \infty\}$.
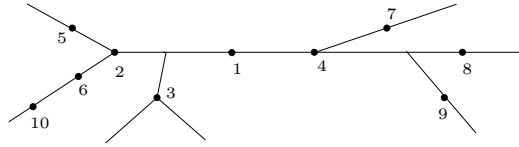

Fig. 2. An illustration of the partition procedure

Odd and even components Root the tree at some vertex i_0, and define the depth of a vertex j as the number of vertices in the unique path from i_0 to j (i_0 has depth 1). Then each component T corresponds to the region of the tree between a vertex j (at depth p) and some of its sons (at depth $p+1$) in the tree. We say that T is odd (resp. even) if the depth p of j is odd (resp. even). This obviously depends on the chosen root.

For instance, in Figure 2 vertices of the same depth are in the same horizontal position (the tree is rooted at vertex 1). Then the components corresponding to $\{1, 2, 3\}, \{1, 4\}, \{5, \infty\}, \{6, 10\}, \ldots$ are odd while the ones corresponding to $\{2, 5\}, \{2, 6\}, \{3, \infty\}, \{4, 8, 9\}, \ldots$ are even.

Note that each vertex except the root — and the ∞-vertices — is both in (at least) one even component and in (at least) one odd component. The root is in (at least) one odd component.

3.2   A randomized mechanism

Given a reported profile $\mathbf{y}$ and a tree $\mathcal{T}$ as a metric space, let $2\alpha := 2\alpha(\mathbf{y})$ be the minimum distance between any two neighbor agents. Let $i^* = i^*(\mathbf{y})$ and $j^* = j^*(\mathbf{y})$ be neighbor agents such that $d(y_{i^*}, y_{j^*}) = 2\alpha$ (if there are more than one choice, break ties arbitrarily). We partition $\mathcal{T}$ into its components as described previously, considering as vertices the set of locations $\mathbf{y}$. Let T^* be the component containing y_{i^*} and y_{j^*}, and let U be the set of agents in T^*. For instance, in Figure 3, the components are $\{7, 10, 11, 12\}$, $\{4, 6, 7, 8\}$, $\{6, 13\}$, $\{13, \infty\}, \cdots$ Suppose that $i^* = 4$ and $j^* = 7$. Then T^* is the component whose set of agents is $U = \{4, 6, 7, 8\}$.

We design a mechanism made of four deterministic placements P_1, P_2, P_3 and P_4; each P_i occurs with probability $1/4$. Intuitively, the mechanism satisfies the following properties: (i) all agents have the same expected cost α, and (ii) for any component in $\mathcal{T}$, with probability $1/2$, no facility is opened inside the component (but possibly at its extremities). To get this, each agent i different from i^* and j^* will have its own facility F_i open at distance α, while i^* and j^* will "share" a facility open either at y_{i^*}, or at y_{j^*}, or in the middle of the path between y_{i^*} and y_{j^*}. However, to ensure strategy-proofness, we need to carefully combine these positions.

If we remove the component T^* (while keeping its vertices) from $\mathcal{T}$, we now have a collection of subtrees $\mathcal{T}_i$ for $i \in U$, where $\mathcal{T}_i$ is rooted at y_i (the location of agent i). For each rooted-subtree $\mathcal{T}_i$, assign the status odd or even to its components according to the procedure previously defined. In Figure 3 (B) if we remove T^* we have four subtrees rooted at 4, 6, 7 and 8. Bold components are odd.
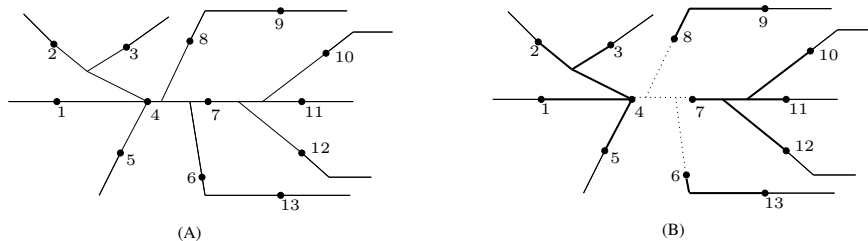

Fig. 3. (A) A tree $\mathcal{T}$ and a profile $\mathbf{y}$ where agents' locations are dots. (B) The four subtrees obtained after removing T^*. Bold components are the odd ones.

We are now able to define the four placements P_1, P_2, P_3, P_4. Nevertheless, recall that a node is in at least one odd component and at least one even component. Each agent $i \neq i^*, j^*$ is associated with a facility F_i, while i^* and j^* share a common facility. We describe in the following the placements of these
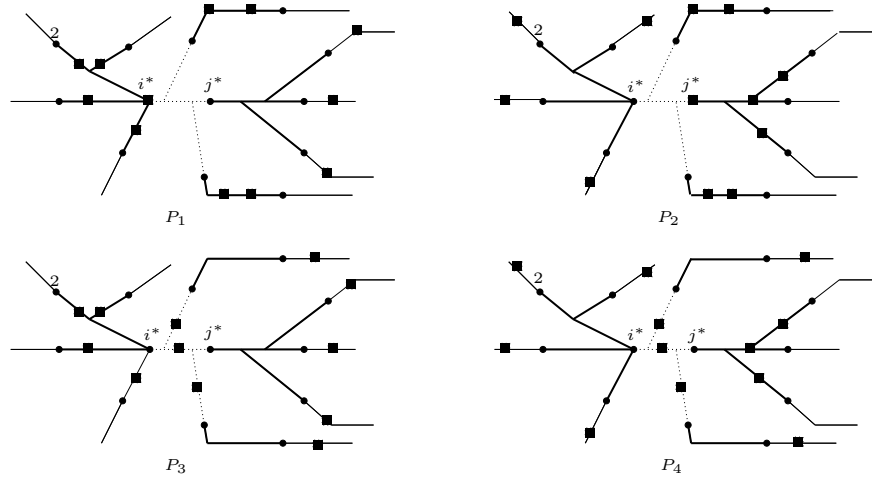
Fig. 4. Placements P_1, P_2, P_3, P_4 for the instance in Figure 3. Agents i^*, j^* are $4, 7$. Facilities are represented by squares.

facilities. We distinguish the agents with respect to the subtree $\mathcal{T}_i$ where they are.

Placement	i^*	$i \in \mathcal{T}_{i^*}$	j^*	$i \in \mathcal{T}_{j^*}$	$i \in U \setminus \{i^*, j^*\}$	$i \in \mathcal{T}_\ell \setminus U$ for $\ell \in U$
P_1	at y_{i^*}	O	no facility	E	O	O
P_2	no facility	E	at y_{j^*}	O	O	O
P_3	mid. y_{i^*}, y_{j^*}	O	no facility	E	T^*	E
P_4	no facility	E	mid. y_{i^*}, y_{j^*}	O	T^*	E

Table 2. Placements of facilities associated with agents

In this table, E (resp. O) means that we open a facility F_i in an even component (resp. odd component) at distance α of y_i for agent i; T^* means that the facility F_i is opened in the component T^*, with distance α from y_i. For the location of any facility, if there are several choices, pick one arbitrarily. In placements P_3 and P_4 "mid. i^*, j^*" means that the position in the middle in the path connecting y_{i^*} and y_{j^*}. Denote $F^*(\mathbf{y})$ the facility opened at this position. In this case, i^* and j^* share the same facility $F^*(\mathbf{y})$.

An illustration is shown in Figure 4. For instance, since y_2 is in the subtree $\mathcal{T}_4 = \mathcal{T}_{i^*}$, the facility F_2 associated with agent 2 is opened in an odd (bold) component in placements P_1 and P_3 and in an even one in placements P_2 and P_4.

Analysis By definition, all the placements P_1, P_2, P_3, P_4 are well defined, i.e., there are at most $n-1$ opening facilities in each placement (one associated to each agent $i \neq i^*, j^*$, plus only one shared by i^* and j^*). The following lemma shows some properties of the mechanism.

Lemma 2. *Given a reported profile* $\mathbf{y}$*, the expected distance between* y_i *and its closest facility equals* $\alpha(\mathbf{y})$ *for* $1 \leq i \leq n$*. Moreover, for any component, there are at least two placements in* $\{P_1, P_2, P_3, P_4\}$ *where the component does not contain any facility (but facilities can be at the extremities of the component).*

Now we prove the strategy-proofness of the mechanism. Suppose that an agent i strategically misreports its location as x_i' (while other agents' locations remain unchanged). Let $\mathbf{x}' = (x_i', x_{-i})$, where $\mathbf{x} = (x_1, \ldots, x_n)$ is the true location profile. Define the parameters $2\alpha := 2\alpha(\mathbf{x}), i^* := i^*(\mathbf{x}), j^* := j^*(\mathbf{x})$. For every agent i, $\mathcal{N}(i, \mathbf{x})$ denotes the set of its neighbors in profile $\mathbf{x}$ ($\mathcal{N}(i, \mathbf{x})$ does not contain i). The strategy-proofness is due to the two following main lemmas.

Lemma 3. *No agent* i *has incentive to misreport its location as* x_i' *such that* $\mathcal{N}(i, \mathbf{x}) \neq \mathcal{N}(i, \mathbf{x}')$*.*

Sketch of the proof. If $\mathcal{N}(i, \mathbf{x}) \neq \mathcal{N}(i, \mathbf{x}')$, then agent i is serviced by a facility outside T' with probability $1/2$ (by Lemma 2), where T' is the component formed from $\mathcal{N}(i, \mathbf{x})$. Since the distance from x_i to any agent in $\mathcal{N}(i, \mathbf{x})$ is at least 2α, agent i has no incentive to report x_i'. □

Lemma 4. *Agent* i *cannot strictly decrease its cost by reporting a location* $x_i' \neq x_i$ *such that* $\mathcal{N}(i, \mathbf{x}) = \mathcal{N}(i, \mathbf{x}')$*.*

Theorem 2. *The mechanism is strategy-proof and it induces a* $n/2$*-approximation according to the utilitarian objective and a tight* $3/2$*-approximation according to the egalitarian objective.*

Proof. The mechanism is strategy-proof by previous lemmas. The cost of each agent is α, so in the utilitarian objective, the cost induced by the mechanism is $n\alpha$. An optimal placement is to open facilities at the locations of all agents but i^*, which induces a cost 2α. Hence, the mechanism is $n/2$-approximation for the utilitarian objective.

Consider the egalitarian objective. By the mechanism, in P_3 and P_4 the maximum cost of an agent is α, while in P_1 and P_2 it is 2α. The average maximum cost of the mechanism is $3\alpha/2$. An optimal solution is to open facilities at locations of agents other than i^*, j^* and open one facility at the midpoint of the path connecting x_{i^*} and x_{j^*}; that gives a cost α. So, the approximation ratio is $3/2$ and this ratio is tight, i.e., no randomized strategy-proof mechanism can do better [3, Theorem 2.4]. □

3.3   Lower bounds on the approximation ratio of SP mechanisms

In this section, we consider only the utilitarian objective (as the tight bound for the egalitarian objective has been derived in the previous section).

Proposition 2. *No deterministic strategy-proof mechanism on a line metric space has an approximation ratio smaller than $3/2$.*

The following proposition indicates that even with randomization, we cannot get an optimal strategy-proof mechanism for the utilitarian objective.

Proposition 3. *No randomized strategy-proof mechanism on a line metric space has an approximation ratio smaller than $10 - 4\sqrt{5} \approx 1.055$.*

4   Discussion and Further Directions

The results presented in this paper are a first step towards handling the general case where one wishes to locate k facilities in a metric space with n agents (for $1 \leq k \leq n$). The general case is widely opened since nothing on the performance of strategy-proof mechanisms is known. Any positive or negative results on the problem would be interesting. We suggest a mechanism based on the *Inversely Proportional Mechanism* in which the k facilities are put on reported locations. Starting with the n reported locations the mechanism would iteratively eliminate a candidate until k locations remain. We do not know whether this mechanism is strategy-proof. For restricted spaces such as line, cycle or tree metric spaces, there might be some specific strategy-proof mechanisms with guaranteed performance which exploits the structures of such spaces. Besides, some characterization of strategy-proof mechanisms (as done by Moulin or Schummer and Vohra [1, 2]), even not a complete characterization, would be helpful.

References

1. Moulin, H.: On strategy-proofness and single peakedness. Public Choice **35** (1980) 437–455
2. Schummer, J., Vohra, R.V.: Strategy-proof location on a network. Journal of Economic Theory **104** (2001)
3. Procaccia, A.D., Tennenholtz, M.: Approximate mechanism design without money. In: ACM Conference on Electronic Commerce. (2009) 177–186
4. Alon, N., Feldman, M., Procaccia, A.D., Tennenholtz, M.: Strategyproof approximation of the minimax on networks. Math. Oper. Res. **35** (2010) 513–526
5. Nguyen Kim, T.: On (group) strategy-proof mechanisms without payment for facility location games. In: Proc. 6th Workshop on Internet and Network Economics (WINE). (2010) 531–538
6. Lu, P., Sun, X., Wang, Y., Zhu, Z.A.: Asymptotically optimal strategy-proof mechanisms for two-facility games. In: ACM Conference on Electronic Commerce. (2010) 315–324
7. Fotakis, D., Tzamos, C.: Winner-imposing strategyproof mechanisms for multiple facility location games. In: Proc. 6th Workshop on Internet and Network Economics (WINE). (2010) 234–245

Appendix

A   SP mechanisms for general metric spaces

Lemma 1. *The Inversely Proportional Mechanism is strategy-proof in a general metric space.*

Proof. Recall that $\mathbf{x} = (x_1, \ldots, x_n)$ is the true location profile of the agents. Let $d_j := d(x_j, P_j(\mathbf{x}))$ for $1 \leq j \leq n$.

If there are at most $(n-1)$ distinct locations in profile $\mathbf{x}$ (some d_j equals 0) then by the mechanism, every agent has cost 0 and no one has incentive to strategically misreports its location. In the sequel, assume that all the agent locations in $\mathbf{x}$ are distinct ($d_j > 0 \ \forall j$).

If all agents report truthfully their locations, the cost of agent i is

$$c_i := c_i(f, \mathbf{x}) = \sum_{j=1}^{n} p_j(\mathbf{x}) \cdot d(x_i, P_j(\mathbf{x})) = p_i(\mathbf{x}) \cdot d_i = \frac{1}{\sum_{j=1}^{n} 1/d_j}$$

By the formula, it is straightforward that $c_i < d_i$.

In the sequel, we suppose that i misreports its location as x_i'. Let $\mathbf{x}' = (x_i', x_{-i})$ be the location profile when i reports x_i' and the other agents report truthfully their locations. Let $d_j' = d(P_j(x_j, \mathbf{x}'))$ for $j \neq i$ and $d_i' = d(P_i(x_i', \mathbf{x}'))$. We will prove that $c_i' := c_i(f, \mathbf{x}') \geq c_i$.

If $x_i' = x_j$ for some other agent j then by the mechanism, all the facilities are opened in $P_i(\mathbf{x})$. That induces the cost d_i for agent i, which is strictly larger than c_i. So if i benefits by its strategical report, the new profile $\mathbf{x}'$ must consist of n distinct locations, i.e., $d_j' > 0 \ \forall 1 \leq j \leq n$.

The new cost of agent i is

$$c_i' = \sum_{j=1}^{n} p_j(\mathbf{x}') \cdot d(x_i, P_j(\mathbf{x}')) \geq p_i(\mathbf{x}') \cdot d_i + (1 - p_i(\mathbf{x}')) \min\{d_i, d(x_i, x_i')\}$$

where the inequality is due to the fact that in $P_j(\mathbf{x}')$ (for $j \neq i$), agent i can choose either some facility in $\{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n\}$ or the facility opened at x_i'. Define $T := \{j : d_j' \neq d_j, j \neq i\}$. Note that

$$p_i(\mathbf{x}') = \frac{1/d_i'}{\sum_{j \notin T} 1/d_j + \sum_{j \in T} 1/d_j' + 1/d_i'}$$

Let $e := d(x_i, x_i')$. Remark that i has no incentive to report its location x_i' in such a way that $e \geq d_i$ since otherwise $c_i' \geq p_i(\mathbf{x}') \cdot d_i + (1 - p_i(\mathbf{x}'))d_i = d_i > c_i$. In the sequel, consider $e < d_i$. In this case,

$$c_i' \geq p_i(\mathbf{x}') \cdot d_i + (1 - p_i(\mathbf{x}')) \cdot e \tag{1}$$

Observe that the right-hand side of the inequality above decreases if $p_i(\mathbf{x}')$ decreases (since $e < d_i$).

First, we claim that $e \geq |d'_i - d_i|$. Let F be a facility in placement $P_i(\mathbf{x})$ such that $d_i = d(x_i, F)$, where $d(x_i, F)$ is the distance between x_i and F. We have $d'_i - d_i = d(x'_i, P_i(\mathbf{x'})) - d_i \leq d(x'_i, F) - d(x_i, F) \leq d(x_i, x'_i) = e$ (the last inequality is due to the triangle inequality). Similarly, $d_i - d'_i \leq e$. Thus, $e \geq |d'_i - d_i|$.

We are now ready to prove that $c'_i \geq c_i$.

Case 1: $\sum_{j \in T} \frac{1}{d'_j} \leq \sum_{j \in T} \frac{1}{d_j}$.

In this case, $p_i(\mathbf{x'}) \geq \frac{1/d'_i}{\sum_{j \notin T} 1/d_j + \sum_{j \in T} 1/d_j + 1/d'_i}$. Let $A = \sum_{j \neq i} \frac{1}{d_i}$. Using (1) we have

$$c'_i \geq \frac{\frac{d_i}{d'_i}}{A + \frac{1}{d'_i}} + \frac{Ae}{A + \frac{1}{d'_i}} = \frac{d_i + Aed'_i}{Ad'_i + 1}$$

We show that $c'_i > c_i = \frac{1}{A + 1/d_i}$ by proving

$$\frac{d_i + Aed'_i}{Ad'_i + 1} > \frac{d_i}{Ad_i + 1}$$

$$\Leftrightarrow Ad_i^2 + d_i + A^2 ed_i d'_i + Aed'_i > Ad_i d'_i + d_i$$

$$\Leftrightarrow Ad_i^2 + A^2 ed_i d'_i + Aed'_i > Ad_i d'_i$$

Recall that $e \geq |d'_i - d_i| \geq d'_i - d_i$. Thus $Ad_i^2 + Aed'_i \geq Ad_i^2 + Ad_i'^2 - Ad_i d'_i \geq Ad_i d'_i$ (by Cauchy inequality). Therefore, $c'_i > c_i$.

Case 2: $\sum_{j \in T} \frac{1}{d'_j} > \sum_{j \in T} \frac{1}{d_j}$.

Let $R = \{j \in T : d'_j < d_j\} \subseteq T$. For every agent $j \in R$, $d'_j = d(P_j(\mathbf{x'}), x_j) < d(P_j(\mathbf{x}), x_j) = d_j$. The difference is necessarily due to the new report x'_i, so $d'_j = d(x_j, x'_i)$. As $\frac{1}{d'_j} \leq \frac{1}{d_j} \; \forall j \in T \setminus R$, $p_i(\mathbf{x'}) \geq \frac{\frac{1}{d'_i}}{\sum_{j \notin R, j \neq i} \frac{1}{d_j} + \sum_{j \in R} \frac{1}{d'_j} + \frac{1}{d'_i}}$

Denote $B = \sum_{j \notin R, j \neq i} \frac{1}{d_i}$, we have using (1):

$$c'_i \geq \frac{\frac{1}{d'_i}}{B + \sum_{j \in R} \frac{1}{d'_j} + \frac{1}{d'_i}} \cdot d_i + \frac{B + \sum_{j \in R} \frac{1}{d'_j}}{B + \sum_{j \in R} \frac{1}{d'_j} + \frac{1}{d'_i}} \cdot e = \frac{\frac{d_i}{d'_i} + \left(B + \sum_{j \in R} \frac{1}{d'_j}\right) e}{B + \sum_{j \in R} \frac{1}{d'_j} + \frac{1}{d'_i}}$$

Again, we show $c'_i > c_i = \frac{1}{B + \sum_{j \in R} \frac{1}{d_j} + \frac{1}{d_i}}$ by proving

$$\frac{\frac{d_i}{d'_i} + \left(B + \sum_{j \in R} \frac{1}{d'_j}\right) e}{B + \sum_{j \in R} \frac{1}{d'_j} + \frac{1}{d'_i}} > \frac{1}{B + \sum_{j \in R} \frac{1}{d_j} + \frac{1}{d_i}}$$

$$\Leftrightarrow \left[\frac{d_i}{d'_i} + \left(B + \sum_{j \in R} \frac{1}{d'_j}\right) e\right]\left[B + \sum_{j \in R} \frac{1}{d_j} + \frac{1}{d_i}\right] > B + \sum_{j \in R} \frac{1}{d'_j} + \frac{1}{d'_i}$$

It is sufficient to prove that

$$\frac{Bd_i}{d'_i} + \frac{Be}{d_i} \geq B \tag{2}$$

and

$$\frac{d_i}{d'_i d_j} + \frac{e}{d_j d'_j} + \frac{e}{d_i d'_j} \geq \frac{1}{d'_j} \qquad \forall j \in R \tag{3}$$

Consider inequality (2). If $B = 0$ or if $d_i \geq d'_i$ then the inequality is straightforward. If $d_i < d'_i$ then the left-hand side of (2) is larger than $\frac{Bd_i + Be}{d'_i} > B$ since $e \geq d'_i - d_i$. The inequality follows.

Consider inequality (3) for a fixed agent $j \in R$. Observe that $e = d(x_i, x'_i) \geq d(x_i, x_j) - d(x'_i, x_j) \geq d_j - d'_j$. If $d_i \geq d'_i$ then

$$\frac{d_i}{d'_i d_j} + \frac{e}{d_j d'_j} = \frac{d_i d'_j + e d'_i}{d'_i d_j d'_j} \geq \frac{d_i d'_j + (d_j - d'_j) d'_i}{d'_i d_j d'_j} = \frac{d_j d'_i + (d_i - d'_i) d'_j}{d'_i d_j d'_j} \geq \frac{1}{d'_j}$$

Otherwise, if $d_i < d'_i$ then $\frac{e}{d_i d'_j} > \frac{e}{d'_i d'_j}$. Therefore,

$$\frac{d_i}{d'_i d_j} + \frac{e}{d_j d'_j} + \frac{e}{d_i d'_j} > \frac{d_i}{d'_i d_j} + \frac{e}{d_j d'_j} + \frac{e}{d'_i d'_j} = \frac{d_i d'_j + e d'_i + e d_j}{d'_i d_j d'_j} > \frac{d_i d'_j + e d_i + e d_j}{d'_i d_j d'_j} \geq$$

$$\geq \frac{d_i(d'_j + e) + (d'_i - d_i) d_j}{d'_i d_j d'_j} = \frac{d'_i d_j + d_i(d'_j + e - d_j)}{d'_i d_j d'_j} \geq \frac{1}{d'_j}$$

Thus, inequality (3) follows. Therefore, $c'_i > c_i$ since inequalities (2) and (3) hold.

In both cases, we have proved that any agent i has no incentive to misreport its location, i.e., the mechanism is strategy-proof. □

B   SP mechanisms on trees

Lemma 2. *Given a reported profile* $\mathbf{y}$*, the expected distance between* y_i *and its closest facility equals* $\alpha(\mathbf{y})$ *for* $1 \leq i \leq n$*. Moreover, for any component, there are at least two placements in* $\{P_1, P_2, P_3, P_4\}$ *where the component does not contain any facility (but facilities can be at the extremities of the component).*

Proof. Consider an agent $i \neq i^*(\mathbf{y}), j^*(\mathbf{y})$ where we recall that $i^*(\mathbf{y}), j^*(\mathbf{y})$ denote the two players whose reported locations are at minimum distance. In any placement, the closest facility is opened at distance $\alpha(\mathbf{y})$ from y_i. For agent $i^* = i^*(\mathbf{y})$, the distance from y_{i^*} to the closest facility is: 0 in P_1, $2\alpha(\mathbf{y})$ in P_2, $\alpha(\mathbf{y})$ in P_3 and P_4. Hence, the average is $\alpha(\mathbf{y})$, and similarly for agent $j^*(\mathbf{y})$.

Let T be the component containing the locations of agents $i^*(\mathbf{y})$ and $j^*(\mathbf{y})$. No facility is opened inside T under placements P_1 and P_2. Besides, by the definition of the mechanism, there are at least two placements in $\{P_1, P_2, P_3, P_4\}$

where a component does not contain a facility[1].   □

Lemma 3. *No agent i has incentive to misreport its location as x'_i such that $\mathcal{N}(i, \mathbf{x}) \neq \mathcal{N}(i, \mathbf{x}')$.*

Proof. Suppose that $\mathcal{N}(i, \mathbf{x}) \neq \mathcal{N}(i, \mathbf{x}')$. In this case, the locations of agents in $\mathcal{N}(i, \mathbf{x})$ form a component T' of tree $\mathcal{T}$ with respect to profile $\mathbf{x}'$. By Lemma 2, with probability at least $1/2$, no facility is opened in T', i.e., in those cases agent i is serviced by a facility outside T'. Note that the distance from x_i to the location of any agent in $\mathcal{N}(i, \mathbf{x})$ is at least 2α. Therefore, the new cost of agent i is at least α, meaning i has no incentive to report x'_i.   □

Lemma 4. *Agent i cannot strictly decrease its cost by reporting a location $x'_i \neq x_i$ such that $\mathcal{N}(i, \mathbf{x}) = \mathcal{N}(i, \mathbf{x}')$.*

Proof. As $\mathcal{N}(i, \mathbf{x}) = \mathcal{N}(i, \mathbf{x}')$, the path connecting x_i and x'_i contains no other agent's location. Hence, there is a component T'_i in the partition of $\mathcal{T}$ with respect to $\mathbf{x}'$ such that $x'_i \in T'_i$ and $x_i \in T'_i$. Let $2\alpha'$ be the minimum distance between two neighbors in $\mathbf{x}'$. Also let $e = d(x_i, x'_i)$.

Case 1: Consider the case where, with the new location x'_i, i is neither $i^*(\mathbf{x}')$ nor $j^*(\mathbf{x}')$. Hence, $\alpha' \geq \alpha$. By Lemma 2, with probability at least $1/2$, no facility is opened inside T'_i. In this case, the distance from x_i to the closest facility is at least $\min\{d(x_i, x'_i) + d(x'_i, F'_i), d(x_i, x_\ell) + d(x_\ell, F'_\ell)\}$ where: $\ell \in \mathcal{N}(i, \mathbf{x})$ and F'_ℓ is its associated facility; and F'_i is the facility opened at distance α' from x'_i, F'_i is in a component different from T'_i. In other words, this distance is at least $\min\{e + \alpha', 2\alpha\}$ since $d(x'_i, F'_i) = \alpha'$ and $d(x_i, x_\ell) \geq 2\alpha$. Besides, with probability at most $1/2$, the closest facility to x_i is either F'_i (the facility opened in component T'_i at distance α' from x'_i) or some other facility F'_ℓ in T'_i for some $\ell \in \mathcal{N}(i, \mathbf{x})$. The former gives a distance $d(x_i, F'_i) \geq \max\{d(x'_i, F'_i) - d(x'_i, x_i), 0\} = \max\{\alpha' - e, 0\}$ (by triangular inequality). The latter gives a distance $d(x_i, F'_\ell) \geq \max\{d(x_i, x_\ell) - d(x_\ell, F'_\ell), 0\} \geq \max\{2\alpha - \alpha', 0\}$. Hence, the cost of agent i is at least

$$\frac{1}{2}\left(\min\{e + \alpha', 2\alpha\} + \min\{\max\{\alpha' - e, 0\}, \max\{2\alpha - \alpha', 0\}\}\right) \geq \alpha$$

where the inequality is due to $\alpha' \geq \alpha$. Indeed, this is immediate if $e + \alpha' \geq 2\alpha$. Otherwise, the cost is either at least $e + \alpha' + \alpha' - e = 2\alpha'$, or $e + \alpha' + 2\alpha - \alpha' \geq 2\alpha$. Hence, $c_i(\mathbf{x}') \geq c_i(\mathbf{x})$.

---

[1] There are facilities in T under P_3 and P_4 but facilities are put on the extremities under placements P_1 and P_2. Notice that a component may never receive a facility if there are two components named $\{i, \infty\}$ and i is located at the intersection of two branches of the tree, see location 3 in Figure 2.

Case 2: Consider the case where with the new location x'_i agent $i = i^*(\mathbf{x}')$ (the case where $i = j^*(\mathbf{x}')$ is completely similar)[2]. Let $j = j^*(\mathbf{x}')$. Let d_1, d_2, d_3, d_4 be the distance from x_i to the closest facility in placements P_1, P_2, P_3, P_4 (in $\mathbf{x}'$), respectively. Let T' be the component in $\mathcal{T}$ with respect to $\mathbf{x}'$ that contains x'_i and x_j. By the triangle inequality, we know that

$$e + 2\alpha' = d(x_i, x'_i) + d(x'_i, x_j) \geq d(x_i, x_j) \geq 2\alpha \qquad (4)$$

We study the two sub-cases and prove that $\sum_{t=1}^{4} d_t \geq 4\alpha$ always holds, meaning that agent i's deviation cannot be profitable since its cost is α when it reports its true location x_i.

(a) The true location x_i belongs to T'.

For each agent $\ell \neq i, j$, let F'_ℓ be its associated facility. The facility opened in the middle of $[x'_i, x_j]$ is denoted by $F^*(\mathbf{x}')$. We have:

$$d_1 = \min\{d(x_i, x'_i), d(x_i, F'_\ell)\} = \min\{e, d(x_i, x_\ell) + d(x_\ell, F'_\ell)\} \geq \min\{e, 2\alpha + \alpha'\}$$
$$(5)$$
$$d_2 = \min\{d(x_i, x_j), d(x_i, F'_\ell)\} \geq \min\{d(x_i, x_j), 2\alpha + \alpha'\} \geq 2\alpha \qquad (6)$$
$$d_3 = \min\{d(x_i, F^*(\mathbf{x}')), d(x_i, F'_\ell)\} \geq \min\{2\alpha - \alpha', e + \alpha', 2\alpha + \alpha'\} \qquad (7)$$
$$d_4 = \min\{d(x_i, F^*(\mathbf{x}')), d(x_i, F'_\ell)\} \geq \min\{2\alpha - \alpha', e + \alpha', 2\alpha + \alpha'\} \qquad (8)$$

where $\ell \neq i, j$ is some agent in $\mathcal{N}(i, \mathbf{x}')$ (note that agents ℓ in the expressions above are not necessarily the same). The first equality in (5) is due to the fact that in placement P_1, agent i goes either to the facility opened at x'_i or to a facility (outside T') associated to some other agent. In placement P_2, agent i can either choose a facility opened at x_j or another one outside T', that is translated in the equality in (6). In placement P_3 and P_4, agent i can go either to facility $F^*(\mathbf{x}')$ opened in the midpoint connecting x'_i and x_j, or to the facility associated with some agent ℓ (inside and outside T' respectively).

If $e + \alpha' < 2\alpha - \alpha'$ then $\sum_{t=2}^{4} d_t \geq 2\alpha + 2e + 2\alpha' \geq 4\alpha$ (since $e + 2\alpha' \geq 2\alpha$). In the sequel, assume $e + \alpha' \geq 2\alpha - \alpha'$. If $e \geq 2\alpha + \alpha'$ then $d_1 + d_3 \geq 4\alpha$. Otherwise, $\sum_{t=1}^{4} d_t \geq e + \min\{d(x_i, x_j), 2\alpha + \alpha'\} + 2\max\{2\alpha - \alpha', 0\}$. Note that by the triangle inequality $e + d(x_i, x_j) = d(x'_i, x_i) + d(x_i, x_j) \geq d(x'_i, x_j) = 2\alpha'$. Therefore, $\sum_{t=1}^{4} d_t \geq \min\{2\alpha' + 4\alpha - 2\alpha', 2\alpha + \alpha' + 2\alpha - \alpha'\} = 4\alpha$. Hence, the new cost of i is at least α.

(b) The true location x_i does <u>not</u> belong to T'.

---

[2] Contrasting with Case 1, $\alpha' \leq \alpha$ does not necessarily hold.

Let T_i' be the component in $\mathcal{T}$ with respect to profile $\mathbf{x}'$ such that T_i' contains x_i and x_i'. Similar to the previous case, we have:

$$
\begin{aligned}
d_2 &= \min\{d(x_i, x_j), d(x_i, F_\ell')\} = \min\{d(x_i, x_i') + d(x_i', x_j), d(x_i, x_\ell) + d(x_\ell, F_\ell')\} \\
&\geq \min\{e + 2\alpha', 2\alpha + \alpha'\} \geq 2\alpha \tag{9} \\
d_3 &= \min\{d(x_i, F^*(\mathbf{x}')), d(x_i, F_\ell')\} \\
&\geq \min\{d(x_i, x_i') + d(x_i', F^*(\mathbf{x}')), d(x_i, x_\ell) - d(x_\ell, F_\ell')\} = \min\{e + \alpha', 2\alpha - \alpha'\} \tag{10} \\
d_4 &= \min\{d(x_i, F^*(\mathbf{x}')), d(x_i, F_\ell')\} \geq \min\{e + \alpha', 2\alpha + \alpha'\} \tag{11}
\end{aligned}
$$

where $\ell \neq i, j$ is some agent in $\mathcal{N}(i, \mathbf{x}')$ (again agents ℓ in the expressions above are not necessarily the same). In placement P_2, agent i can choose either a facility opened at x_j or another one outside T_i'. The last inequality of (9) is due to $e + 2\alpha' \geq 2\alpha$ (Inequality 4). In placement P_3 and P_4, agent i can go either to facility $F^*(\mathbf{x}')$ opened in the midpoint connecting x_i' and x_j, or some facilities associated with some agent ℓ.

If $e + \alpha' < 2\alpha - \alpha'$ then $\sum_{t=2}^{4} d_t \geq 2\alpha + 2e + 2\alpha' \geq 4\alpha$ (since $e + 2\alpha' \geq 2\alpha$). Otherwise, $\sum_{t=2}^{4} d_t \geq \min\{e + 4\alpha, 4\alpha\} \geq 4\alpha$. Again, the new cost of agent i is at least α.

In conclusion, no agent has incentive to strategically misreport its location.   □

Proposition 2. *No deterministic strategy-proof mechanism on a line metric has an approximation ratio smaller than* $3/2$.

Proof. We prove the theorem for four agents (the construction can be straightforwardly generalized for any even number of agents). Let f be a deterministic SP mechanism with finite approximation ratio. Consider a profile $\mathbf{x}$ in Figure 5 in which $x_1 = A, x_2 = B, x_3 = C, x_4 = D$ where $d(A, B) = d(C, D) = 1$ and $d(B, C)$ is very large. For any placement of three facilities, there is one agent with cost at least $1/2$. By symmetry, suppose that agent 4 has cost at least $1/2$.

Let $\bar{f} : (D, +\infty) \to (D, +\infty) \cup \{+\infty\}$ be the function indicating the facility opened in $(D, +\infty)$ with smallest coordinate while the location of agent 4 varies in interval $(D, +\infty)$ (the other agents' locations remain unchanged at A, B, C). Conventionally, $\bar{f}(x) = +\infty$ if there is no facility opened in $(D, +\infty)$. Let $F^* := \inf Image(\bar{f})$. As f has a finite approximation ratio, F^* is finite (while the location of agent 4 is far enough from D, f must open a facility closed to that location).

For $\epsilon > 0$, let $x^*(\epsilon) = \inf\{\bar{f}^{-1}([F^*, F^* + \epsilon))\}$, i.e., for any $\delta > 0$, there exists $x_\delta(\epsilon) \in (D, +\infty)$ such that $|x_\delta(\epsilon) - x^*(\epsilon)| \leq \delta$ and $\bar{f}(x_\delta(\epsilon)) \in [F^*, F^* + \epsilon)$. If $Image(\bar{f}) \cap (D, D + 1/2) \neq \emptyset$ then agent 4 may report a false location and reduce its cost, contradicting the strategy-proofness of f. Therefore, $Image(\bar{f}) \subseteq [D + 1/2, +\infty) \cup \{+\infty\}$.

If there exits ϵ_0 such that $x^*(\epsilon_0) = D$. Then consider a profile in which agent 4 locates at $x_\delta(\epsilon_0)$ (δ arbitrarily small) and the other agents locate at A, B, C.
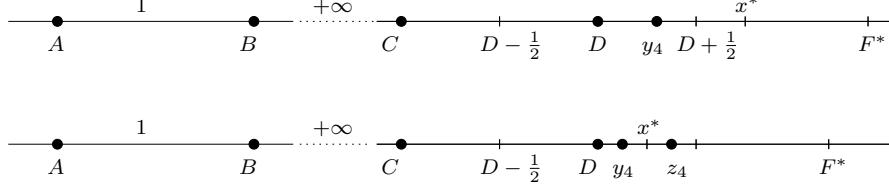
Fig. 5. Instance which gives the lower bound on the ratio of a deterministic strategy-proof mechanism in a line metric space. The cases (i) and (ii) are illustrated in the upper and lower lines, respectively.

We have $\bar{f}(x_\delta(\epsilon_0)) \in [D+1/2, +\infty)$ and no facility is opened by f in the interval $(D - 1/2, D)$ (by strategy-proofness of f). Thus, the cost of agent 4 is at least $1/2 - \delta$. Moreover, there are at most two facilities in $(-\infty, D - 1/2)$ so the total cost of the other agents is at least 1. Therefore, the total cost is at least $3/2 - \delta$, while OPT is 1. In the sequel, assume that $x^*(\epsilon) > D$ for all ϵ.

Fix $\epsilon > 0$ and denote $x^* := x^*(\epsilon)$. Consider a profile $\mathbf{y}$ in which agent 4 locates at $y_4 = \min\{x^*, D + 1/2\} - \gamma$ ($\gamma > 0$ small enough such that $y_4 > D$) and the other agents locate at A, B, C. By definition of x^*, $\bar{f}(y_4) > F^*$. Observe that the closest facility to agent 4 is not $\bar{f}(y_4)$ since otherwise, the agent could report its location as in $\bar{f}^{-1}\left([F^*, F^* + \frac{\bar{f}(y_4) - F^*}{2})\right)$ to strictly reduce its cost. Hence, the closest facility to y_4 is in $(-\infty, D - 1/2)$. In other words, no facility is opened in the ball with center y_4 and radius $y_4 - (D - 1/2)$.

Consider two cases: (i) $D + 1/2 \leq x^*$; and (ii) $D + 1/2 > x^*(> D)$. (An illustration is shown in Figure 5.)

(i) In this case $y_4 = D + 1/2 - \gamma$. In profile $\mathbf{y}$, no facility is opened in the ball with center y_4 and radius $1 - \gamma$. Hence, for any placement of three facilities with that property, the total cost is at least $3/2 - \gamma$. The optimal cost is 1 by placing facilities at locations of all agents but agent 1.

(ii) In this case $y_4 = x^* - \gamma$. By strategy-proofness, F^* is not in the ball with center y_4 and radius $y_4 - (D - 1/2) - \epsilon$. (Otherwise, agent 4 can reduce the cost by reporting its location in $\bar{f}^{-1}([F^*, F^* + \epsilon))$). Hence, $F^* \geq y_4 + (y_4 - D + 1/2) - \epsilon = 2(x^* - \gamma) - D + 1/2 - \epsilon$.

Let $\delta \geq 0$ such that $x^* + \delta < D + 1/2$ and $\bar{f}(x^* + \delta) \in [F^*, F^* + \epsilon)$. Consider a profile $\mathbf{z}$ in which agent 4 locates at $z_4 = x^* + \delta$ and the other agents locate at A, B, C. In this profile, no facility is opened in $(D - 1/2, F^*)$ and there is (at least) one facility opened in $[F^*, F^* + \epsilon)$. Therefore, the cost of agent 4 is at least $|F^* - z_4| > 1/2 - \delta - 2\gamma - \epsilon$ (since $x^* > D$). For any placement of the two other facilities in interval $(-\infty, D - 1/2)$, the total cost of the other agents is at least 1; so the total cost is at least $3/2 - \delta - 2\gamma - \epsilon$. The optimal cost is 1 by placing facilities at locations of all agents but agent 1.

In any case, we can choose parameter ϵ, δ, γ arbitrarily small, so the approximation ratio is arbitrarily close to $3/2$. □

Proposition 3. *No randomized strategy-proof mechanism on a line metric space has an approximation ratio smaller than $10 - 4\sqrt{5} \approx 1.055$.*

Proof. Let f be a randomized strategy-proof mechanism with an approximation ratio strictly better than $1 + \varepsilon > 1$. Consider a profile $\mathbf{x}$ where the positions of the agents are $x_1 = A, x_2 = B, x_3 = C, x_4 = D$ (Figure 6). For any placement of three facilities, the total cost is at least 1. Hence, there exists an agent with (expected) cost at least $1/4$. Without loss of generality, suppose that agent 1 (with $x_1 = A$) has cost $c_1(f, \mathbf{x}) \geq 1/4$.
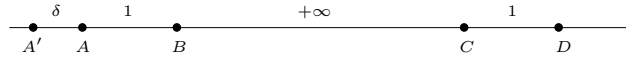

Fig. 6. Instance which gives the lower bound on the ratio of a randomized strategy-proof mechanism in a line metric space.

Let $0 < \delta < 1/4$ be a constant to be defined later. Let $A' \notin [A, B]$ be a location at distance δ from A. Let $\mathbf{y}$ be the profile in which agent 1 is located at $y_1 = A'$ and the other agents' locations are the same as in $\mathbf{x}$. By strategy-proofness, $c_1(f, \mathbf{x}) \leq \delta + c_1(f, \mathbf{y})$. Hence, $c_1(f, \mathbf{y}) \geq 1/4 - \delta$. In $\mathbf{y}$, an optimal solution has cost 1 (e.g. place the facilities at the locations of the agents other than agent 4). As f is a $(1 + \varepsilon)$-approximation, the total cost of the solution returned by the mechanism is $c_1(f, \mathbf{y}) + c_2(f, \mathbf{y}) + c_3(f, \mathbf{y}) + c_4(f, \mathbf{y}) \leq 1 + \varepsilon$. Thus, $c_3(f, \mathbf{y}) + c_4(f, \mathbf{y}) \leq 3/4 + \varepsilon + \delta$.

In outcome $f(\mathbf{y})$, let p be the probability that the closest facility of agent 3 is also the closest facility of agent 4 (in other words, agents 3 and 4 share one facility with probability p; and with probability $(1 - p)$ there is at most one facility between A' and B). We have $c_3(f, \mathbf{y}) + c_4(f, \mathbf{y}) \geq p \cdot 1 = p$. Therefore, $p \leq 3/4 + \varepsilon + \delta$.

Besides, the social cost of $f(\mathbf{y})$ is at least $p + (1 - p)(1 + \delta) = 1 + \delta - p\delta$. This is lower bounded by $1 + \delta - (3/4 + \varepsilon + \delta)\delta$. Hence, $1 + \delta - (3/4 + \varepsilon + \delta)\delta \leq C(f, \mathbf{y}) \leq 1 + \varepsilon$. We deduce that $\varepsilon \geq \frac{\delta/4 - \delta^2}{1 + \delta}$.

The function $\frac{\delta/4 - \delta^2}{1 + \delta}$ for $\delta \in (0, \frac{1}{4})$ attains maximal value $9 - 4\sqrt{5}$ for $\delta = \sqrt{5}/2 - 1$. Therefore, the approximation ratio is at least $1 + \varepsilon \geq 10 - 4\sqrt{5} \approx 1.055$. □