Winner-Imposing Strategyproof Mechanisms for Multiple Facility Location Games*

Dimitris Fotakis and Christos Tzamos

School of Electrical and Computer Engineering,
National Technical University of Athens, 157 80 Athens, Greece.
Email: fotakis@cs.ntua.gr, ctzamos@gmail.com

Abstract. We study Facility Location games, where a number of facilities are placed in a metric space based on locations reported by strategic agents. A mechanism maps the agents' locations to a set of facilities. The agents seek to minimize their connection cost, namely the distance of their true location to the nearest facility, and may misreport their location. We are interested in mechanisms that are *strategyproof*, i.e., ensure that no agent can benefit from misreporting her location, do not resort to monetary transfers, and approximate the optimal social cost. We focus on the closely related problems of *k-Facility Location* and *Facility Location* with a uniform facility opening cost, and mostly study *winner-imposing* mechanisms, which allocate facilities to the agents and require that each agent allocated a facility should connect to it. We show that the winner-imposing version of the Proportional Mechanism (Lu *et al.*, EC '10) is stategyproof and $4k$-approximate for the k-Facility Location game. For the Facility Location game, we show that the winner-imposing version of the randomized algorithm of (Meyerson, FOCS '01), which has an approximation ratio of 8, is strategyproof. Furthermore, we present a deterministic non-imposing group strategyproof $O(\log n)$-approximate mechanism for the Facility Location game on the line.

1 Introduction

We consider *Facility Location games*, where a number of facilities are placed in a metric space based on the preferences of strategic agents. Such problems are motivated by natural scenarios in social choice, where the government plans to build a number of public facilities in an area. The choice of the locations is based on the preferences of local people, or *agents*. So each agent reports her ideal location, and the government applies a *mechanism* mapping the agents' preferences to a set of facility locations. The government's objective is to minimize the *social cost*, namely the total distance of the agents' locations to the nearest facility plus the *construction cost*, in case where the number of facilities is not fixed and may depend on the agents' preferences. On the other hand, the agents seek to minimize their *connection cost*, namely the distance of their ideal location to the nearest facility. In fact, an agent may report a false preference in an attempt of manipulating the mechanism. Therefore, the mechanism should be *strategyproof*, i.e., should ensure that no agent can benefit from misreporting her

* Research partially supported by an NTUA Basic Research Grant (PEBE 2009).

location, or even *group strategyproof*, i.e., should ensure that for any group of agents misreporting their locations, at least one of them does not benefit. At the same time, the mechanism should achieve a reasonable approximation to the optimal social cost.

In this work, we consider two closely related facility location problems, and present computationally efficient strategyproof approximate mechanisms for both. In the k-*Facility Location* game, we place k facilities in a metric space so as to minimize the agents' total connection cost. In the *Facility Location* game, there is a uniform facility opening cost, instead of a fixed number of facilities, and we place a number of facilities in a metric space so as to minimize the sum of the agents' total connection cost and the total facility opening cost. This problem is motivated by natural scenarios where the social planner is willing to trade off the agents' connection cost against its own construction cost, so that a socially more desirable solution is achieved, and has been widely used as a natural relaxation of the k-Facility Location problem (see e.g. [6, 14]).

Related Work. The problems of Facility Location and k-Facility Location, a.k.a. k-Median, are classical and have received considerable attention in Operations Research (see e.g. [14]), Approximation and Online Algorithms (see e.g. [6, 21, 8, 13, 2]), Social Choice (see e.g. [16, 5, 22, 15, 20, 4, 9]), and recently, Algorithmic Mechanism Design (see e.g. [19, 1, 11, 10, 18]). The related work in Social Choice mostly focuses on locating a single facility on the real line, where the agents' preferences are single-peaked. A classical result due to Moulin [16], Barberà and Jackson [5], and Sprumont [22] characterizes the class of generalized median voter schemes as the only strategyproof mechanisms when agents have single-peaked preferences on the line (see also [3, 23] and [17, Chapter 10]). Schummer and Vohra [20] extended this result to tree metrics, where the class of extended median voter schemes are the only strategyproof mechanisms. On the negative side, Schummer and Vohra proved that for non-tree metrics, only dictatorial rules can be both strategyproof and onto. The problem of designing mechanisms with desirable properties for multiple facility location games has also been considered (see e.g. [15, 4, 9]). This line of work however does not address the issue of designing strategyproof mechanisms that approximate the optimal social cost.

Our work fits in the framework of *approximate mechanism design without money*, recently initiated by Procaccia and Tennenholtz [19]. They suggested that for optimization problems, such as 2-Facility Location on the line and 1-Facility Location on non-tree metrics, where computing the optimal solution is not strategyproof, approximation can circumvent impossibility results and yield strategyproof mechanisms that do not resort to monetary transfers. Procaccia and Tennenholtz [19] applied this approach to several location problems on the real line, and obtained strategyproof approximate mechanisms and lower bounds on the best approximation ratio achievable by a strategyproof mechanism. For the 2-Facility Location game on the line, they presented a deterministic $(n-1)$-approximate mechanism, where n is the number of agents, proved a lower bound of $3/2$ on the approximation ratio of any deterministic strategyproof mechanism, and conjectured that the lower bound for deterministic mechanisms is $\Omega(n)$.

Subsequently, Lu, Wang, and Zhou [11] improved the lower bound for deterministic mechanisms to 2, established a lower bound of 1.045 for randomized mechanisms, and presented a simple randomized $n/2$-approximate mechanism. For locating two facilities on the line, Lu, Sun, Wang, and Zhu [10] improved the lower bound for deterministic mechanisms to $(n-1)/2$, thus settling the conjecture of [19]. Moreover, they presented a deterministic $(n-1)$-approximate mechanism for locating two facilities on the circle, and proved that a natural randomized mechanism, the *Proportional Mechanism*, is strategyproof and achieves an approximation ratio of 4 for 2-Facility Location on any metric space. Unfortunately, Lu *et al.* observed that the Proportional Mechanism is not strategyproof for more than two facilities. For 1-Facility Location, Alon, Feldman, Procaccia, and Tennenholtz [1] gave an almost complete characterization of the approximation ratios achievable by randomized and deterministic strategyproof mechanisms.

Following a more general agenda, McSherry and Talwar [12] suggested the use of differentially private algorithms as almost-strategyproof approximate mechanisms. Any agent has a limited influence on the outcome of differentially private algorithm, and thus a limited incentive to lie. McSherry and Talwar presented a general (randomized exponential-time) differentially private mechanism that approximates the optimal social cost within an additive logarithmic term. Subsequently, Gupta *et al.* [7] presented computationally efficient differentially private algorithms for several combinatorial optimization problems, including (k-)Facility Location.

Building on [12], Nissim, Smorodinsky, and Tennenholtz [18] developed the only known general technique for the design of strategyproof approximate mechanisms without money. Nissim *et al.* consider *imposing mechanisms*, namely mechanisms able to restrict how agents exploit their outcome. Restricting the set of allowable post-actions for the agents, the mechanism can penalize liars. For Facility Location games in particular, an imposing mechanism requires that an agent should connect to the facility nearest to her reported location, thus increasing her connection cost if she lies. Despite being stronger, imposing mechanisms do not circumvent the lower bounds of [11]. Nissim *et al.* combined the differentially private mechanism of [12] with an imposing mechanism that penalizes lying agents, and obtained a general imposing strategyproof mechanism. As a by-product, Nissim *et al.* obtained a randomized imposing mechanism for k-Facility Location with a running time exponential in k. The mechanism approximates the optimal average connection cost, namely the optimal connection cost divided by n, within an additive term of roughly $1/n^{1/3}$. Even though the error term is diminishing as n grows, it may happen that the optimal average cost decreases much faster. In fact, for the class of instances in [11, Theorem 3], the optimal average cost is $1/n$ and the mechanism's error is at least $1/n^{1/3}$. Thus, the additive approximation guarantee of [18] does not imply any constant approximation ratio for k-Facility Location.

Contribution. Our work is motivated by the absence of any positive results on the approximability of multiple facility location games by non-imposing mechanisms, and by the recent striking result of [18] on their approximability by imposing mechanisms. In fact, the only work prior to ours that addresses approximate mechanism design for location problems with more than two facilities is [18]. Throughout this work, we restrict our attention to computationally efficient strategyproof mechanisms without money[1] and to the standard multiplicative notion of approximation. We suggest two orthogonal

[1] We consider the problems of Facility Location and k-Facility Location, with k being part of the input, which are $\mathcal{NP}$-hard. Thus one cannot directly apply VCG payments (see e.g. [17, Chapter 9]) and obtain a computationally efficient strategyproof mechanism that minimizes the social cost.

ways of relaxing approximate mechanism design for the k-Facility Location game, and show that both lead to strong positive results.

We mostly consider a natural class of imposing mechanisms, which we call *winner-imposing* mechanisms. Such a mechanism operates by allocating facilities to the agents. If an agent is allocated a facility, the facility is placed to her reported location, and the agent should connect to it. Agents not allocated a facility connect to the facility closest to their ideal location. Thus a winner-imposing mechanism penalizes a lying agent only if she succeeds in manipulating the mechanism. Moreover, the "penalty" a lying agent receives equals the distance of her ideal location to her misreported location.

In contrast to the observation of [10] that the Proportional Mechanism is not strategyproof for more than two facilities, we prove that its winner-imposing version is strategyproof for any number of facilities (cf. Lemma 1). Establishing that its approximation ratio is at most $4k$ (cf. Lemma 2), we obtain a randomized winner-imposing strategyproof $4k$-approximate mechanism for k-Facility Location, for any k.

Next we consider the Lagrangian relaxation of the k-Facility Location game, namely the Facility Location game with a uniform facility opening cost, instead of a hard constraint on the number of facilities. In fact, considering the Facility Location problem as a relaxation of k-Facility Location, a.k.a. k-Median, has been a standard and quite successful approach in the fields of Operations Research (see e.g. [14]) and Approximation Algorithms (see e.g. [6, 8]).

For the Facility Location game, we first show that the winner-imposing version of Meyerson's randomized algorithm for Facility Location [13] is strategyproof (cf. Theorem 2). Combining this with [13, Theorem 2.1], we obtain a randomized winner-imposing strategyproof 8-approximate mechanism for the Facility Location game.

Moreover, we present a deterministic non-imposing mechanism for the Facility Location game on the line. The mechanism is based on a hierarchical partitioning of the line, and is motivated by the online algorithm for Facility Location on the plane by Anagnostopoulos, Bent, Upfal, and van Hentenryck [2]. We prove that the mechanism is group strategyproof (cf. Lemma 4) and $O(\log n)$-approximate (cf. Lemma 5). Notably, its approximation ratio is exponentially better than the lower bound of [10, Theorem 3.7] on the best ratio achievable by deterministic strategyproof mechanisms for the 2-Facility Location game on the line. Thus, our results demonstrate that the Facility Location game allows for some significantly (even exponentially) better approximation guarantees (by non-imposing strategyproof mechanisms) than the k-Facility Location game, and may suggest a potential connection between approximate mechanism design without money and online optimization.

We also consider (randomized) oblivious winner-imposing mechanisms, and derive a natural condition for them to be strategyproof. A mechanism is oblivious if conditional on the event that an agent is not allocated a facility, her presence has no impact on the mechanism's outcome. The Proportional Mechanism and Meyerson's algorithm are oblivious. We show that an oblivious winner-imposing mechanism for the (k-)Facility Location game on a continuous metric space is strategyproof iff it is *locally strategyproof*, i.e., no agent can benefit by reporting a location arbitrarily close to her true location (cf. Lemma 3). On the other hand, we note that local strategyproofness does not imply strategyproofness for the non-imposing version of Meyerson's algorithm.

2 Model, Definitions, and Notation

For an integer $m \geq 1$, we let $[m] = \{1, \ldots, m\}$. For an event E in a sample space, we let $\mathrm{I\!Pr}[E]$ be the probability of E happening. For a random variable X, we let $\mathrm{I\!E}[X]$ be the *expectation* of X.

We assume an underlying *metric space* (M, d), where $d : M \times M \mapsto \mathrm{I\!R}$ is the distance function, which is non-negative, symmetric, and satisfies the triangle inequality. For $x \in M$ and a non-empty $M' \subseteq M$, we let $d(x, M') = \inf\{d(x, y) : y \in M'\}$. For a location $x \in M$ and a positive real r, we let $\mathrm{Ball}(x, r) = \{y \in M : d(x, y) \leq r\}$. A metric space (M, d) is continuous if for any $x, y \in M$ with $d(x, y) \leq 2r$, there is a $z \in \mathrm{Ball}(x, r) \cap \mathrm{Ball}(y, r)$ such that $d(x, y) = d(x, z) + d(z, y)$.

For a tuple $\boldsymbol{x} = (x_1, \ldots, x_n)$, we let $\boldsymbol{x}_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ be the tuple without x_i. For a non-empty $S \subset [n]$, we let $\boldsymbol{x}_S = (x_i)_{i \in S}$ and $\boldsymbol{x}_{-S} = (x_i)_{i \in [n] \setminus S}$. We write $\boldsymbol{x} = (x_i, \boldsymbol{x}_{-i})$ and $\boldsymbol{x} = (\boldsymbol{x}_S, \boldsymbol{x}_{-S})$.

Mechanisms. Let $N = \{1, \ldots, n\}$ be a set of agents. Each agent $i \in N$ has a location $x_i \in M$, which is i's private information. Next we refer to $\boldsymbol{x} = (x_1, \ldots, x_n)$ as the *location profile*. A *deterministic mechanism* F maps a location profile $\boldsymbol{x}$ to a tuple of non-empty sets $(C, C^1, \ldots C^n)$, where $C \subseteq M$ is the facility set of F and each $C^i \subseteq C$ contains the facilities where agent i should connect. We write $F(\boldsymbol{x})$ to denote the facility set of F and $F^i(\boldsymbol{x})$ to denote the facility subset of each agent i. For the k-Facility Location game, $|F(\boldsymbol{x})| = k$, while for the Facility Location game, $|F(\boldsymbol{x})|$ can be any positive number. A *randomized mechanism* is a probability distribution over deterministic mechanisms.

A mechanism F is *non-imposing* if for all location profiles $\boldsymbol{x}$ and all agents i, $F^i(\boldsymbol{x}) = F(\boldsymbol{x})$, and *imposing* otherwise. We only consider imposing mechanisms where each agent i can connect to the facility in $F(\boldsymbol{x})$ closest to her reported location, namely where $\{z \in F(\boldsymbol{x}) : d(x_i, z) = d(x_i, F(\boldsymbol{x}))\} \subseteq F^i(\boldsymbol{x})$ for all i. A mechanism F is said to allocate facilities to the agents[2] if $F(\boldsymbol{x}) \subseteq \{x_1, \ldots, x_n\}$. A mechanism F that allocates facilities to the agents is *winner-imposing* if for every agent i, $F^i(\boldsymbol{x}) = \{x_i\}$ if $x_i \in F(\boldsymbol{x})$, and $F^i(\boldsymbol{x}) = F(\boldsymbol{x})$ otherwise. For a winner-imposing mechanism F and some location profile $\boldsymbol{x}$, we write either that F allocates a facility to agent i or that F places a facility at x_i to denote that F adds x_i in its facility set $F(\boldsymbol{x})$. Moreover, we write that F connects agent i to the facility at x_i to denote that $F^i(\boldsymbol{x}) = \{x_i\}$, as a result of $x_i \in F(\boldsymbol{x})$.

Individual and Social Cost. Given a deterministic mechanism F and a location profile $\boldsymbol{x}$, the cost of agent i is $\mathrm{cost}[x_i, F(\boldsymbol{x})] = d(x_i, F^i(\boldsymbol{x}))$. If F is a randomized mechanism, the expected cost of agent i is $\mathrm{cost}[x_i, F(\boldsymbol{x})] = \mathrm{I\!E}_{C^i \sim F^i(\boldsymbol{x})}[d(x_i, C^i)]$.

The *social cost* for the *k-Facility Location* game of a deterministic mechanism F for a location profile $\boldsymbol{x}$ is $\mathrm{SC}_k[F(\boldsymbol{x})] = \sum_{i=1}^{n} d(x_i, F(\boldsymbol{x}))$, subject to the constraint that $|F(\boldsymbol{x})| = k$. For the *Facility Location* game, there is a uniform facility opening cost $f > 0$, and the social cost of a deterministic mechanism F for a location profile

[2] To simplify and unify the presentation, we implicitly assume here that all locations x_i are distinct. This assumption does not affect the generality of our model and our results, and can be removed by letting the mechanism map each location profile to a tuple $(C, C^1, \ldots C^n)$, with $C \subseteq N$ and $C^i \subseteq C$, $C^i \neq \emptyset$, and place a facility at x_i for each $i \in C$.

$\boldsymbol{x}$ is $\text{SC}[F(\boldsymbol{x})] = f\,|F(\boldsymbol{x})| + \sum_{i=1}^{n} d(x_i, F(\boldsymbol{x}))$. Scaling the distances appropriately, we assume that the facility opening cost is equal to 1. The expected social cost of a randomized mechanism F for a location profile $\boldsymbol{x}$ is defined by taking the expectation of $\text{SC}_k[F(\boldsymbol{x})]$ (resp. $\text{SC}[F(\boldsymbol{x})]$) over the distribution of $F(\boldsymbol{x})$.

A (randomized) mechanism F achieves an *approximation ratio* of $\rho \geq 1$, if for all location profiles $\boldsymbol{x}$, the (resp. expected) social cost of $F(\boldsymbol{x})$ is at most ρ times the optimal social cost for $\boldsymbol{x}$.

Strategyproofness and Group Strategyproofness. A mechanism F is *strategyproof* if for any location profile $\boldsymbol{x}$, any agent i, and any location y, $\text{cost}[x_i, F(\boldsymbol{x})] \leq \text{cost}[x_i, F(y, \boldsymbol{x}_{-i})]$. A mechanism F is *group strategyproof* if for any location profile $\boldsymbol{x}$, any non-empty set of agents S, and any location profile $\boldsymbol{y}_S$ for them, there exists some agent $i \in S$ such that $\text{cost}[x_i, F(\boldsymbol{x})] \leq \text{cost}[x_i, F(\boldsymbol{y}_S, \boldsymbol{x}_{-S})]$.

3 The Winner-Imposing Proportional Mechanism

We consider the winner-imposing version of the Proportional Mechanism [10] for the k-Facility Location game. Given a location profile $\boldsymbol{x} = (x_i)_{i \in N}$, the Winner-Imposing Proportional Mechanism, or WIProp in short, works in k rounds, fixing the location of one facility in each round. For each $\ell = 1, \ldots, k$, let C_ℓ be the set of the first ℓ facilities of WIProp. Initially, $C_0 = \emptyset$. WIProp proceeds as follows:

1st Round: WIProp selects i_1 uniformly at random from N, places the first facility at x_{i_1}, connects agent i_1 to it, and lets $C_1 = \{x_{i_1}\}$.

ℓ-th Round, $\ell = 2, \ldots, k$: WIProp selects $i_\ell \in N$ with probability $\frac{d(x_{i_\ell}, C_{\ell-1})}{\sum_{i \in N} d(x_i, C_{\ell-1})}$, places the ℓ-th facility at x_{i_ℓ}, connects agent i_ℓ to it, and lets $C_\ell = C_{\ell-1} \cup \{x_{i_\ell}\}$.

The output of the mechanism is C_k, and every agent not allocated a facility is connected to the facility in C_k closest to her true location. The proof of the following theorem follows from Lemma 1 and Lemma 2 below.

Theorem 1. WIProp *is a strategyproof $4k$-approximation mechanism for the k-Facility Location game on any metric space.*

Strategyproofness. Even though the non-imposing version of the Proportional Mechanism is not strategyproof for $k \geq 3$ [10], WIProp is strategyproof for any k.

Lemma 1. *For any $k \geq 1$,* WIProp *is a strategyproof mechanism for the k-Facility Location game.*

Proof. For each $\ell = 0, 1, \ldots, k$, we let $\text{cost}[x_i, F(y, \boldsymbol{x}_{-i}) | C_\ell]$ be the expected connection cost of an agent i at the end of WIProp, given that i reports location y and that the facility set of WIProp at the end of round ℓ is C_ℓ. For $\ell = k$, $\text{cost}[x_i, F(y, \boldsymbol{x}_{-i}) | C_k] = d(x_i, C_k)$. For each $\ell = 1, \ldots, k-1$, with probability proportional to $d(y, C_\ell)$ the next facility of WIProp is placed at i's reported location, in which case i is connected to y and incurs a connection cost of $d(x_i, y)$, while for each agent $j \neq i$, with probability proportional to $d(x_j, C_\ell)$ the next facility of WIProp is placed at x_j, in which case the expected connection cost of i is $\mathrm{cost}[x_i, F(y, \boldsymbol{x}_{-i})|C_\ell \cup \{x_j\}]$. Therefore:

$$\mathrm{cost}[x_i,\, F(y, \boldsymbol{x}_{-i})|C_\ell] = \\ = \frac{d(x_i, y)\, d(y, C_\ell) + \sum_{j \neq i} d(x_j, C_\ell)\, \mathrm{cost}[x_i, F(y, \boldsymbol{x}_{-i})|C_\ell \cup \{x_j\}]}{d(y, C_\ell) + \sum_{j \neq i} d(x_j, C_\ell)} \quad (1)$$

Similarly, for $\ell = 0$, the expected connection cost of agent i is:

$$\mathrm{cost}[x_i, F(y, \boldsymbol{x}_{-i})] = \frac{d(x_i, y) + \sum_{j \neq i} \mathrm{cost}[x_i, F(y, \boldsymbol{x}_{-i})|\{x_j\}]}{n} \quad (2)$$

By induction on ℓ, we show that for any y, any $\ell = 0, 1, \ldots, k$, and any C_ℓ,

$$\mathrm{cost}[x_i, F(y, \boldsymbol{x}_{-i})|C_\ell] \geq \mathrm{cost}[x_i, F(\boldsymbol{x})|C_\ell] \quad (3)$$

Thus agent i has no incentive to misreport her location, which implies the lemma.

For the basis, we observe that (3) holds for $\ell = k$. Indeed, if i's location is not in C_k, her connection cost is $d(x_i, C_k)$ and does not depend on her reported location y, while if i's location is in C_k her connection cost is $d(x_i, y) \geq d(x_i, x_i)$. We inductively assume that (3) holds for $\ell + 1$ and any facility set $C_{\ell+1}$, and show that (3) holds for ℓ and any facility set C_ℓ. If $\ell \geq 1$, we use (1) and obtain that:

$$\mathrm{cost}[x_i,\, F(y, \boldsymbol{x}_{-i})|C_\ell] \geq \\ \geq \frac{d(x_i, y)\, d(y, C_\ell) + \sum_{j \neq i} d(x_j, C_\ell)\, \mathrm{cost}[x_i, F(\boldsymbol{x})|C_\ell \cup \{x_j\}]}{d(y, C_\ell) + \sum_{j \neq i} d(x_j, C_\ell)} \\ = \frac{d(x_i, y)\, d(y, C_\ell) + \Big(d(x_i, C_\ell) + \sum_{j \neq i} d(x_j, C_\ell)\Big)\, \mathrm{cost}[x_i, F(\boldsymbol{x})|C_\ell]}{d(y, C_\ell) + \sum_{j \neq i} d(x_j, C_\ell)} \quad (4)$$

The inequality follows from (1) and the induction hypothesis. For the equality, we apply (1) with $y = x_i$. If $d(x_i, C_\ell) \geq d(y, C_\ell)$, (4) implies that $\mathrm{cost}[x_i, F(y, \boldsymbol{x}_{-i})|C_\ell] \geq \mathrm{cost}[x_i, F(\boldsymbol{x})|C_\ell]$. Otherwise, we continue from (4) and obtain that:

$$\mathrm{cost}[x_i, F(y, \boldsymbol{x}_{-i})|C_\ell] \geq \frac{d(x_i, y) + d(x_i, C_\ell) + \sum_{j \neq i} d(x_j, C_\ell)}{d(y, C_\ell) + \sum_{j \neq i} d(x_j, C_\ell)}\, \mathrm{cost}[x_i, F(\boldsymbol{x})|C_\ell] \\ \geq \mathrm{cost}[x_i, F(\boldsymbol{x})|C_\ell]$$

For the first inequality, we use that $d(y, C_\ell) > d(x_i, C_\ell) \geq \mathrm{cost}[x_i, F(\boldsymbol{x})|C_\ell]$. For the second inequality, we use that $d(x_i, y) + d(x_i, C_\ell) \geq d(y, C_\ell)$.

If $\ell = 0$, using (2) and the induction hypothesis, we obtain that:

$$\mathrm{cost}[x_i, F(y, \boldsymbol{x}_{-i})] \geq \frac{1}{n} \sum_{j \neq i} \mathrm{cost}[x_i, F(\boldsymbol{x})|\{x_j\}] = \mathrm{cost}[x_i, F(\boldsymbol{x})] \quad (5)$$

Thus we have established (3) for any location y, any $\ell = 0, 1, \ldots, k$, and any C_ℓ. □

Approximation Ratio. To establish the approximation ratio, we extend the ideas of [10, Theorem 4.2] to the case where $k \geq 3$.

Lemma 2. *For any $k \geq 1$,* WIProp *achieves an approximation ratio of at most $4k$ for the k-Facility Location game.*

4 A Randomized Mechanism for Facility Location

Next we consider the winner-imposing version of Meyerson's randomized algorithm for Facility Location [13], and show that it is strategyproof. Meyerson's algorithm, or OFL in short, processes the agents one-by-one in a random order, and places a facility at the location of each agent with probability equal to her distance to the nearest facility divided by the facility opening cost (which we assume to be 1). For simplicity, we assume that the agents are indexed according to the random permutation chosen by OFL. Also we let C_i denote the facility set of OFL just after agent i is processed.

Formally, given the locations $\boldsymbol{x} = (x_i)_{i \in N}$ of a randomly permuted set of agents, the (winner-imposing) OFL mechanism first places a facility at x_1, connects agent 1 to it, and lets $C_1 = \{x_1\}$. Then, for each $i = 2, \dots, n$, with probability $d(x_i, C_{i-1})$, OFL opens a facility at x_i, connects agent i to it, and lets $C_i = C_{i-1} \cup \{x_i\}$. Otherwise, OFL lets $C_i = C_{i-1}$. The output of the mechanism is C_n, and every agent not allocated a facility is connected to the facility in C_n closest to her true location.

Theorem 2. *The winner-imposing version of* OFL *is a strategyproof* 8*-approximation mechanism for the Facility Location game on any metric space.*

Proof. The approximation ratio follows from [13, Theorem 2.1]. Next, we show that the winner-imposing version of OFL is strategyproof for any permutation of agents.

Let i be any agent, and let x_i be i's true location. If $i = 1$ or $d(x_i, C_{i-1}) \geq 1$, OFL places a facility at x_i with certainty, so i has no incentive to lie about her location. So we restrict our attention to the case where $d(x_i, C_{i-1}) < 1$.

Let $\mathrm{cost}[x_i, F(y, x_{i+1}, \dots, x_n)|C]$ be the expected connection cost of agent i at the end of OFL, given that i reports location y, and that just before i's location is processed, the set of facilities is C. Similarly, let $\mathrm{cost}[x_i, F(x_{i+1}, \dots, x_n)|C]$ be the expected connection cost of agent i at the end of OFL, given that just after i's location is processed, the set of facilities is C. To establish the strategyproofness of OFL, we have to show that for any agent i located at x_i, for any location y, and for any C_{i-1},

$$\mathrm{cost}[x_i, F(x_i, x_{i+1}, \dots, x_n)|C_{i-1}] \leq \mathrm{cost}[x_i, F(y, x_{i+1}, \dots, x_n)|C_{i-1}] \qquad (6)$$

Calculating i's expected connection cost for x_i and y, we obtain that (6) holds iff

$$(d(y, C_{i-1}) - d(x_i, C_{i-1}))\, \mathrm{cost}[x_i, F(x_{i+1}, \dots, x_n)|C_{i-1}] \leq d(x_i, y)\, d(y, C_{i-1})$$

If $d(y, C_{i-1}) \leq d(x_i, C_{i-1})$, (6) holds because the lhs of the inequality above becomes non-positive. Otherwise, (6) holds because $d(y, C_{i-1}) - d(x_i, C_{i-1}) \leq d(x_i, y)$ and $\mathrm{cost}[x_i, F(x_{i+1}, \dots, x_n)|C_{i-1}] \leq d(x_i, C_{i-1}) < d(y, C_{i-1})$. □

Remark. The argument above fails to establish that the non-imposing version of OFL is strategyproof. This is demonstrated by a simple instance with n agents on the real line. The first agent is located at $-1/2$, the second at 0, the third at $1/2 - \varepsilon$, for some small $\varepsilon > 0$, and the remaining $n - 3$ agents are located at 0. For appropriately chosen n and ε and the particular permutation, the second agent can improve her expected connection cost in the non-imposing version of OFL by reporting $1/2$. On the other hand, no agent has an incentive to lie if the expectation of their connection cost is also taken over all random agents' permutations. Thus, our example does not exclude the possibility that the non-imposing version of OFL is strategyproof for the Facility Location game. □

5 Oblivious Winner-Imposing Mechanisms

Next we consider the class of oblivious winner-imposing mechanisms for (k-)Facility Location, and show that they are strategyproof iff they are locally strategyproof.

A randomized mechanism F that allocates facilities to the agents is *oblivious* if for any location profile $\boldsymbol{x} = (x_i)_{i\in N}$, any agent i, and any location y (y may be x_i),

$$\mathrm{cost}[x_i, F(y, \boldsymbol{x}_{-i}) | i \notin F(y, \boldsymbol{x}_{-i})] = \overline{d}(x_i, F(\boldsymbol{x}_{-i}))\,,$$

where $\overline{d}(x_i, F(\boldsymbol{x}_{-i}))$ is the expected distance of x_i to the nearest facility in $F(\boldsymbol{x}_{-i})$, i.e., F's outcome on the locations of all agents other than i. Namely, F is oblivious if conditional on the event that an agent i is not allocated a facility, her presence has no impact on F's outcome. WIProp and OFL are oblivious mechanisms.

A mechanism F is *locally strategyproof* for the (k-)Facility Location game if there exists an $r > 0$, such that for any location profile $\boldsymbol{x} = (x_i)_{i\in N}$, any agent i, and any $y \in \mathrm{Ball}(x_i, r)$, $\mathrm{cost}[x_i, F(\boldsymbol{x})] \le \mathrm{cost}[x_i, F(y, \boldsymbol{x}_{-i})]$.

Lemma 3. *Let F be an oblivious winner-imposing mechanism for (k-)Facility Location on a continuous metric space. Then F is locally strategyproof iff it is strategyproof.*

Proof. Clearly, any strategyproof mechanism is locally strategyproof. For the other direction, let $\boldsymbol{x} = (x_i)_{i\in N}$ be any location profile. For any agent i and any location x, we let $p(x) = \mathbb{P}\mathrm{r}[i \in F(x, \boldsymbol{x}_{-i})]$ be the probability that i is allocated a facility by F if she reports location x. Similarly to the proof of Theorem 2, we observe that F is strategyproof iff for any agent i with true location x_i and any location y,

$$p(y)\,(\overline{d}(x_i, F(\boldsymbol{x}_{-i})) - d(x_i, y)) \le p(x_i)\,\overline{d}(x_i, F(\boldsymbol{x}_{-i})) \tag{7}$$

We show that if F is locally strategyproof for some $r > 0$, (7) holds for any location y.

Let i be any agent. If $r \ge \overline{d}(x_i, F(\boldsymbol{x}_{-i}))$, (7) holds for any location y, since any $y \notin \mathrm{Ball}(x_i, r)$ makes its lhs non-positive. Otherwise, we show that (7) holds for any location $y \in \mathrm{Ball}(x_i, 2r)$. Let $y \in \mathrm{Ball}(x_i, 2r) \setminus \mathrm{Ball}(x_i, r)$. Since the metric space is continuous, there is a $z \in \mathrm{Ball}(x_i, r) \cap \mathrm{Ball}(y, r)$ with $d(x_i, y) = d(x_i, z) + d(z, y)$. If $d(x_i, y) \ge \overline{d}(x_i, F(\boldsymbol{x}_{-i}))$, (7) holds because its lhs is non-positive. Otherwise,

$$\begin{aligned} p(y) \le\ & p(z)\,\frac{\overline{d}(z, F(\boldsymbol{x}_{-i}))}{\overline{d}(z, F(\boldsymbol{x}_{-i})) - d(z, y)} \le p(z)\,\frac{\overline{d}(x_i, F(\boldsymbol{x}_{-i})) - d(x_i, z)}{\overline{d}(x_i, F(\boldsymbol{x}_{-i})) - d(z, y) - d(x_i, z)} \\ =\ & p(z)\,\frac{\overline{d}(x_i, F(\boldsymbol{x}_{-i})) - d(x_i, z)}{\overline{d}(x_i, F(\boldsymbol{x}_{-i})) - d(x_i, y)} \le p(x_i)\,\frac{\overline{d}(x_i, F(\boldsymbol{x}_{-i}))}{\overline{d}(x_i, F(\boldsymbol{x}_{-i})) - d(x_i, y)} \end{aligned}$$

For the first inequality, we use that F is locally strategyproof for r, and apply (7) for locations z, y. For the first two inequalities, since $\overline{d}(x_i, F(\boldsymbol{x}_{-i})) > d(x_i, y)$, we have that $\overline{d}(z, F(\boldsymbol{x}_{-i})) - d(z, y) > 0$, that $\overline{d}(x_i, F(\boldsymbol{x}_{-i})) - d(x_i, z) > 0$, and that $\overline{d}(x_i, F(\boldsymbol{x}_{-i})) - d(z, y) - d(x_i, z) > 0$. For the last inequality, we use that F is locally strategyproof for r, and apply (7) for locations x_i, z. □

Remark. OFL is locally strategyproof for any permutation of agents and r equal to the minimum distance separating two different locations. On the other hand, we presented an instance where for certain permutations, the non-imposing version of OFL allows an agent to improve her expected cost by misreporting her location. Thus local strategyproofness does not imply strategyproofness for non-imposing OFL. □

6 A Deterministic Mechanism for Facility Location on the Line

We present a deterministic non-imposing group strategyproof $O(\log n)$-approximate mechanism for Facility Location on the real line. To simplify the presentation, we assume that the facility opening cost is 1, and that the agents are located in $\mathbb{R}_+ = [0, \infty)$.

The *Line Partitioning* mechanism, or LPart in short, is motivated by the online algorithm of [2] for Facility Location on the plane. LPart assumes a hierarchical partitioning of $[0, \infty)$ with at most $1 + \log_2 n$ levels. The partitioning at level 0 consists of intervals of length 1. Namely, for $p = 0, 1, \ldots$, the p-th level-0 interval is $[p, p + 1)$. Each level-ℓ interval $[p\, 2^{-\ell}, (p + 1)2^{-\ell})$, $\ell = 0, 1, \ldots, \lfloor \log_2 n \rfloor - 1$, is partitioned into two disjoint level-$(\ell + 1)$ intervals of length $2^{-(\ell+1)}$, namely $[p\, 2^{-\ell}, p\, 2^{-\ell} + 2^{-(\ell+1)})$ and $[p\, 2^{-\ell} + 2^{-(\ell+1)}, (p + 1)\, 2^{-\ell})$. A level-0 interval is active if it includes the (reported) location of at least one agent. A level-ℓ interval, $\ell \geq 1$, is *active* if it includes the locations of at least $2^{\ell+1}$ agents, and *inactive* otherwise. Intuitively, an interval is active if it includes so many agents that the optimal solution opens a facility nearby.

LPart opens three facilities, two at the endpoints and one at the midpoint, of each level-0 active interval, and one facility at the midpoint of each level-ℓ active interval, for each $\ell \geq 1$. In particular, for each level-0 active interval $[p, p + 1)$, LPart opens three facilities at p, at $p + \frac{1}{2}$, and at $p + 1$. For each $\ell \geq 1$ and each level-ℓ active interval $[p\, 2^{-\ell}, (p + 1)2^{-\ell})$, LPart opens a facility at $p\, 2^{-\ell} + 2^{-(\ell+1)}$. LPart is non-imposing, so each agent is connected to the open facility closest to her true location.

Theorem 3. LPart *is a group strategyproof* $O(\log n)$*-approximate mechanism for the Facility Location game on the real line.*

Proof. We start with some observations regarding the structure of the solution produced by LPart. We observe that if an interval q is active, any interval containing q is active, while if an interval q is inactive, any interval included in q is inactive as well. Moreover, all level-$\lfloor \log_2 n \rfloor$ intervals are inactive, since each of them contains less than $2^{\lfloor \log_2 n \rfloor + 1}$ agents. So each agent is included in at least one active and at least one inactive interval. In the following, each agent i is associated with the maximal (i.e., that of the smallest level) inactive interval, denoted q_i, that contains her true location. The maximal inactive intervals q_i, q_j of two agents i, j either coincide with each other or are disjoint.

A simple induction shows that each active interval q has three open facilities, two at its endpoints and one at its midpoint. Moreover, if an active level-ℓ interval contains an inactive level-$(\ell+1)$ subinterval q', q' has two open facilities at its endpoints. Therefore, the connection cost of each agent i is equal to the distance of her true location to the nearest endpoint of her maximal inactive interval q_i. Furthermore, i's connection cost is at least as large as the distance of her true location to the nearest endpoint of any inactive interval containing her true location.

Group Strategyproofness. The above properties of LPart immediately imply that:

Lemma 4. LPart *is group strategyproof.*

Proof. Let $S \subseteq N$, $S \neq \emptyset$, be any coalition of agents who misreport their locations so as to improve their connection cost, and let $\boldsymbol{x}_S = (x_i)_{i \in S}$ and $\boldsymbol{y}_S = (y_i)_{i \in S}$ be the profiles with their true and their misreported locations respectively. If for some agent i,

i's maximal inactive interval q_i contains the same number of agents in $\mathrm{LPart}(\boldsymbol{x}_S, \boldsymbol{x}_{-S})$ and in $\mathrm{LPart}(\boldsymbol{y}_S, \boldsymbol{x}_{-S})$, q_i is inactive in $\mathrm{LPart}(\boldsymbol{y}_S, \boldsymbol{x}_{-S})$ as well, and i's connection cost does not improve. On the other hand, if q_i contains more agents in $\mathrm{LPart}(\boldsymbol{y}_S, \boldsymbol{x}_{-S})$ than in $\mathrm{LPart}(\boldsymbol{x}_S, \boldsymbol{x}_{-S})$, there are some agents in S whose maximal inactive interval is disjoint to q_i in $\mathrm{LPart}(\boldsymbol{x}_S, \boldsymbol{x}_{-S})$ and is included in q_i in $\mathrm{LPart}(\boldsymbol{y}_S, \boldsymbol{x}_{-S})$. Therefore, there is some agent $j \in S$ whose maximal inactive interval q_j contains less agents in $\mathrm{LPart}(\boldsymbol{y}_S, \boldsymbol{x}_{-S})$ than in $\mathrm{LPart}(\boldsymbol{x}_S, \boldsymbol{x}_{-S})$. Thus q_j is inactive in $\mathrm{LPart}(\boldsymbol{y}_S, \boldsymbol{x}_{-S})$ as well, and j's connection cost does not improve due to the deviation of S. □

Approximation Ratio. We proceed along the lines of [2, Theorem 1]. We first show that the optimal solution has a facility close to each active interval.

Proposition 1. *Let $q = [p\, 2^{-\ell}, (p{+}1)2^{-\ell})$ be an active level-ℓ interval, for some $\ell \geq 0$. Then, the optimal solution has a facility in $[(p-1)2^{-\ell}, (p+2)2^{-\ell})$.*

Proof. Let $q_l = [(p-1)2^{-\ell}, p\, 2^{-\ell})$ be the interval next to q on the left, let $q_r = [(p+1)2^{-\ell}, (p+2)2^{-\ell})$ be the interval next to q on the right, and let n_q be the number of agents in q. For sake of contradiction, we assume that the optimal solution does not have a facility in $q_l \cup q \cup q_r$. Then the connection cost of the agents in q is greater than $n_q 2^{-\ell}$. If $\ell = 0$, placing an optimal facility at the location of some agent in q costs 1 and decreases the connection cost of the agents in q to at most $n_q - 1$. If $\ell \geq 1$, placing an optimal facility at the midpoint of q decreases the connection cost of the agents in q to at most $n_q 2^{-(\ell+1)}$. Since q is active and $n_q \geq 2^{\ell+1}$ ($n_q \geq 1$ for $\ell = 0$), the total cost in the later case is less than the connection cost of the agents in q to a facility outside $q_l \cup q \cup q_r$, a contradiction. □

Lemma 5. LPart *has an approximation ratio of $O(\log n)$.*

Proof. Let k be the number of facilities in the optimal solution. By Proposition 1, there are at most 3 active intervals per optimal facility at each level $\ell = 0, 1, \ldots, \lfloor \log n \rfloor - 1$. The total facility cost for the three (neighboring) active level-0 intervals is 7, and the facility cost for each active level-ℓ interval, $\ell \geq 1$, is 1. Therefore, the number of active intervals is at most $3k \log_2 n$, and the total facility cost of LPart is at most $4k + 3k \log_2 n$.

To bound the connection cost of LPart, we consider the set of maximal inactive intervals that include the location of at least one agent (i.e., they are non-empty). This accounts for the connection cost of all agents, since each agent i is associated with her maximal inactive interval q_i. Each maximal inactive interval q at level ℓ, $\ell \geq 1$, contains less than $2^{\ell+1}$ agents and has two facilities at its endpoints. Thus the total connection cost for the agents in q is at most $2^{\ell+1} 2^{-\ell}/2 = 1$. Furthermore, q is included in some active level-$(\ell-1)$ interval. Thus, the total number of non-empty maximal inactive intervals, and thus the total connection cost of LPart, is at most $6k \log_2 n$. Overall, the total cost of LPart is at most $4k + 9k \log_2 n$, i.e. $O(\log_2 n)$ times the optimal cost. □

Acknowledgements: We wish to thank Angelina Vidali for helpful discussions and for bringing [18] to our attention.

References

1. N. Alon, M. Feldman, A.D. Procaccia, and M. Tennenholtz. Strategyproof approximation of the minimax on networks. *Mathematics of Operations Research*, to appear, 2010.
2. A. Anagnostopoulos, R. Bent, E. Upfal, and P. van Hentenryck. A simple and deterministic competitive algorithm for online facility location. *Information and Computation*, 194:175–202, 2004.
3. S. Barberà. An introduction to strategyproof social choice functions. *Social Choice and Welfare*, 18:619–653, 2001.
4. S. Barberà and C. Beviá. Locating public libraries by majority: Stability, consistency and group formation. *Games and Economic Behaviour*, 56:185–200, 2006.
5. S. Barberà and M. Jackson. A characterization of strategy-proof social choice functions for economies with pure public goods. *Journal of Economic Theory*, 61:262–289, 1994.
6. S. Guha. *Approximation Algorithms for Facility Location Problems*. PhD thesis, Stanford University, 2000.
7. A. Gupta, K. Ligett, F. McSherry, A. Roth, and K. Talwar. Differentially private combinatorial optimization. In *Proc. of the 21st ACM-SIAM Symposium on Discrete Algorithms (SODA '10)*, 2010.
8. K. Jain and V. Vazirani. Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and Lagrangian relaxation. *Journal of the ACM*, 48(2):274–296, 2001.
9. B.-G. Ju. Efficiency and consistency for locating multiple public facilities. *Journal of Economic Theory*, 138:165–183, 2008.
10. P. Lu, X. Sun, Y. Wang, and Z.A. Zhu. Asymptotically optimal strategy-proof mechanisms for two-facility games. In *Proc. of the 11th ACM Conference on Electronic Commerce (EC '10)*, pp. 315–324, 2010.
11. P. Lu, Y. Wang, and Y. Zhou. Tighter bounds for facility games. In *Proc. of the 5th Workshop on Internet and Network Economics (WINE '09)*, LNCS 5929, pp. 137–148, 2009.
12. F. McSherry and K. Talwar. Mechanism design via differential privacy. In *Proc. of the 48th IEEE Symposium on Foundations of Computer Science (FOCS '07)*, pp. 94–103, 2007.
13. A. Meyerson. Online facility location. In *Proc. of the 42nd IEEE Symposium on Foundations of Computer Science (FOCS '01)*, pp. 426–431, 2001.
14. P.B. Mirchandani and R.L. Francis (Editors). *Discrete Location Theory*. Wiley, 1990.
15. E. Miyagawa. Locating libraries on a street. *Social Choice and Welfare*, 18:527–541, 2001.
16. H. Moulin. On strategy-proofness and single-peakedness. *Public Choice*, 35:437–455, 1980.
17. N. Nisan, T. Roughgarden, É. Tardos, and V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, 2007.
18. K. Nissim, R. Smorodinsky, and M. Tennenholtz. Approximately optimal mechanism design via differential privacy. CoRR abs/1004.2888, 2010.
19. A.D. Procaccia and M. Tennenholtz. Approximate mechanism design without money. In *Proc. of the 10th ACM Conference on Electronic Commerce (EC '09)*, pp. 177–186, 2009.
20. J. Schummer and R.V. Vohra. Strategyproof location on a network. *Journal of Economic Theory*, 104:405–428, 2002.
21. D. Shmoys. Approximation algorithms for facility location problems. In *Proc. of the 3rd Workshop on Approximation Algorithms for Combinatorial Optimization (APPROX '00)*, LNCS 1913, pp. 27–33, 2000.
22. Y. Sprumont. The division problem with single-peaked preferences: A characterization of the uniform allocation rule. *Econometrica*, 49:509–519, 1991.
23. Y. Sprumont. Strategyproof collective choice in economic and political environments. *The Canadian Journal of Economics*, 28(1):68–108, 1995.