MATHEMATICS OF OPERATIONS RESEARCH
Vol. 35, No. 3, August 2010, pp. 513–526
ISSN 0364-765X | EISSN 1526-5471 | 10 | 3503 | 0513
DOI 10.1287/moor.1100.0457

Strategyproof Approximation of the Minimax on Networks

Noga Alon
Microsoft Israel R&D Center, 46725 Herzeliya, Israel, and School of Mathematics and Computer Science, Tel Aviv University, 69978 Tel Aviv, Israel, nogaa@tau.ac.il, http://www.math.tau.ac.il/~nogaa

Michal Feldman
Microsoft Israel R&D Center, 46725 Herzeliya, Israel, and Center for the Study of Rationality and School of Business Administration, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel, mfeldman@huji.ac.il, http://pluto.huji.ac.il/~mfeldman

Ariel D. Procaccia
School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, arielpro@seas.harvard.edu, http://people.seas.harvard.edu/~arielpro

Moshe Tennenholtz
Microsoft Israel R&D Center, 46725 Herzeliya, Israel, and Technion–IIT, 32000 Haifa, Israel, moshet@microsoft.com, http://iew3.technion.ac.il/Home/Users/Moshet.phtml

We consider the problem of locating a facility on a network represented by a graph. A set of strategic agents have different ideal locations for the facility; the cost of an agent is the distance between its ideal location and the facility. A *mechanism* maps the locations reported by the agents to the location of the facility. We wish to design mechanisms that are *strategyproof* (SP) in the sense that agents can never benefit by lying and, at the same time, provide a small approximation ratio with respect to the *minimax* measure. We design a novel "hybrid" strategyproof randomized mechanism that provides a tight approximation ratio of 3/2 when the network is a circle (known as a *ring* in the case of computer networks). Furthermore, we show that no randomized SP mechanism can provide an approximation ratio better than $2 - o(1)$, even when the network is a tree, thereby matching a trivial upper bound of two.

Key words: mechanism design; approximation; minimax; facility location
MSC2000 subject classification: Primary: 91A10; secondary: 90B80
OR/MS subject classification: Primary: games/noncooperative; secondary: facilities/location/continuous
History: Received November 11, 2009; revised April 12, 2010, and May 25, 2010.

1. Introduction. In facility location problems, one must locate a facility so as to serve a set of agents. Each choice of location for the facility has a *cost* (also known as *effect*) with respect to each agent. In this context, a *mechanism* (or *social choice function* or *location rule*) receives the locations of the agents as input and then outputs the facility location.

Mechanisms are often expected to optimize some quantitative measure; two such measures have been prominently featured in the operations research literature (see, e.g., Holzman [9]). Under the intuitive *minisum* measure, the goal is to minimize the sum of the costs of the agents, thereby maximizing efficiency. In this paper, however, we seek *equitable* solutions. Marsh and Schilling [13] list 20 different measures of equity that have been employed in the facility location literature. The "earliest and most frequently used" (Marsh and Schilling [13], p. 6) measure is the *minimax*, where the best location is one that minimizes the maximum cost of any agent. This measure was suggested as early as 1965 (Hakimi [6]) and is a paradigmatic example of the Rawlsian principle, which seeks to improve the situation of members of society that are worst off (Rawls [16]). The literature includes studies that argue for the equitability of the minimax measure in the context of public services (Hodgart [8]) and, in particular, fire station locations (Richard et al. [17]).

A natural assumption regarding the preferences of the agents is that they are *single-peaked*, that is, each agent has an *ideal location* (or *peak*) and the "closer" the facility is to an agent's ideal location, the lower the agent's cost (Black [4]). This setting has been extensively studied, especially by social choice theorists (as it easily lends itself to social choice interpretation, e.g., in terms of elections where the locations correspond to political points of view). A significant body of literature, starting with the work of Moulin [14], rigorously studies the game theoretic aspects of this setting. This line of work seeks *strategyproof* (SP) mechanisms in the sense that an agent cannot benefit by misreporting its ideal location regardless of the reports of the other agents. In other words, reporting truthfully must be a dominant strategy.

We will be particularly interested in the setup of Schummer and Vohra [19], where the agents are located on a *network* that is represented as a graph. The cost to an agent in its work is just the length of the shortest path between the agent's ideal location and the facility location. This abstract setting has many applications (e.g., traffic network): our running example will be telecommunications networks such as a local computer network or the Internet. In these cases, the agents are the network users or service providers, and the facility can be a filesharing server or a router, for example. This interpretation motivates attention to specific, common network topologies such as tree networks (also known as hierarchical networks), star networks (which are, graph theoretically speaking, a special case of trees), and ring networks. Interestingly, in computer networks, an agent's perceived network location can be easily manipulated, for example, by generating a false IP address or rerouting incoming and outgoing communication, etc.

Schummer and Vohra [19] give a characterization of SP mechanisms when the underlying network is a tree. Furthermore, they demonstrate that if the network contains a cycle, then any SP and onto mechanism is almost dictatorial in the sense that one fixed agent dictates the location of the facility. This result is analogous to the celebrated Gibbard-Satterthwaite impossibility theorem (Gibbard [5], Satterthwaite [18]) but holds for the case where the preferences of the agents are restricted by the topology of the network. It follows from the results of Schummer and Vohra [19] (but was also previously known) that when the network is a tree, then the minisum can be implemented by an SP mechanism. In contrast, there is no SP mechanism that implements the minimax (i.e., always selects a solution that minimizes the maximum cost) under essentially any nontrivial network topology.

In this paper, we ask whether the minimax can be *approximated* well by an SP mechanism, where approximation is defined in the usual sense by looking at the worst-case ratio between the maximum cost of the mechanism's solution and the maximum cost of the optimal solution. We note that, by resorting to approximation, we may lose some of the axiomatic properties that make the minimax appealing, e.g., Pareto optimality.[1] However, in some systems, performance is commonly quantified by the minimax measure; this is true in telecommunications networks where arguably the prominent performance measure is the *maximum network delay*, that is, the maximum time it takes a bit to travel from one node of the network to another (here, the distances in the graph correspond to network congestion). Therefore, in some settings, approximating the minimax while guaranteeing strategyproofness implies optimizing performance in the face of strategic behavior.

As it turns out, our answers are straightforward when restricted to deterministic mechanisms but become more intricate when randomization is allowed. We believe that the consideration of randomized mechanisms in the setting of Schummer and Vohra [19] is one of our main conceptual contributions.

1.1. Our results. We first observe that even the dictatorial rule yields a deterministic SP 2-approximation mechanism. Moreover, this bound is tight with respect to deterministic mechanisms, even if the network is a line.

In §3, we deal with randomized mechanisms for the case where the network is a circle. We present a mechanism that works by combining two mechanisms: One is applied when all the agents are located on one semicircle whereas the other is applied when the agents are not located on one semicircle. We show that this mechanism is an SP 3/2-approximation mechanism for the minimax when the network is a circle. This result matches a lower bound of 3/2. Although the theorem's proof is quite lengthy, we suggest that the truly striking aspect of this result is the mechanism itself and its strategyproofness. Indeed, the mechanism seems to be a coarse hybridization of two mechanisms where the combination is required for achieving the desired approximation ratio. Although each of the two mechanisms is SP in its own right, it seems to us quite extraordinary that the combined mechanism is SP as well.

In §4, we establish that, in contrast to circles, if the network is a tree, then randomization cannot significantly help us. Indeed, even though we show that there is a randomized SP mechanism with an approximation ratio of $2-2/(n+2)$ (where n is the number of agents), we establish the following lower bound: for every n, there exists a tree network such that no randomized SP mechanism can have an approximation ratio that is smaller than $2-o(1)$ for the minimax.

1.2. Additional related work. Our work builds on a previous paper by Procaccia and Tennenholtz [15]. Their basic setting is a special case of ours, where the network is a line. In this setting, if one wishes to minimize the maximum cost, then the best deterministic SP mechanism has an approximation ratio of 2 whereas the best randomized SP mechanism has a ratio of 3/2. Procaccia and Tennenholtz [15] focus on two extensions of the basic setting: location of two facilities on a line (where the cost of an agent is its distance to the nearest facility) and location of one facility on a line when each agent controls multiple points. Tighter results for these two extensions have been found by Lu et al. [11]. All the foregoing results by Procaccia and Tennenholtz [15] and by Lu et al. [11] hold only when the agents are located on a line. They do not consider more general networks. In a very recent paper that is partly motivated by a draft of this paper, Lu et al. [12] provide results with respect

[1] The minimax can be adapted to satisfy Pareto optimality by combining it in a multiobjective model with an efficiency objective (Marsh and Schilling [13]).

to locating two facilities and the minisum measure; in particular, they give a deterministic SP mechanism for locating two facilities on a circle that yields an $O(n)$-approximation for the minisum, thereby matching an $\Omega(n)$ lower bound. They also provide a randomized SP 4-approximation 2-facility mechanism for the minisum in any metric space.

A significant body of work (e.g., Hansen and Thisse [7], Labbé [10], Bandelt [2], Bandelt and Labbé [3]) deals with a model practically identical to ours and considers several variations on the following question: How bad can a Condorcet point be in terms of its sum of costs? A Condorcet point is a location in a network that is preferred by more than half the agents to any other location. The quality of the Condorcet point is measured by bounding the ratio between its sum of costs and the sum of costs of the minisum location. In this respect, the results are reminiscent of approximation (for the minisum) but the approach is descriptive rather than algorithmic. In addition, this line of work does not directly deal with incentives.

2. Model and deterministic mechanisms. We use the model of Schummer and Vohra [19]. Let $N = \{1, \ldots, n\}$ be the set of agents. The network is represented by a graph G, formalized as follows. The graph is a closed, connected subset of Euclidean space $G \subset \mathbb{R}^k$. The graph is composed of a finite number of closed curves of finite length, known as the *edges*. The extremities of the curves are known as *vertices*. The agents and the facility may be located anywhere on G.

The reader might feel that the traditional, discrete model of graphs is more appropriate. In addition, other related papers (e.g., Labbé [10]) consider a similar continuous model but allow the agents to be located only on the vertices (whereas the facility can be located anywhere). Crucially, all our results or slight variations thereof hold under both these alternative models as well.

The *distance* between two points $x, y \in G$, denoted $d(x, y)$, is the length of the minimum-length path between x and y, where a *path* is a minimal connected subset of G that contains x and y. The center of the path between x and y is denoted $\text{cen}(x, y)$, that is, it is a point z on the path such that $d(x, z) = d(y, z)$. We will also use this notation to denote the center of an interval in $\mathbb{R}$.

A *cycle* in G is defined to be the union of two paths such that their intersection is equal to the set of both their endpoints. A graph that does not contain cycles is called a *tree*.

We shall be especially interested in the graph that is a single cycle; we refer to such a graph as a *circle*. Given that G is a circle, we denote the shorter open arc between $x, y \in G$ by $[x, y]$ and the shorter closed arc between x and y by $[x, y]$.[2] For every $x \in G$, we denote by $\hat{x}$ the antipodal point of x on G, that is, the diametrically opposite point. For two points $x, y \in G$ and in the context of an arc of length less than half the circumference of G, we denote the "clockwise operator" by $\succeq$ and its strong version by $\succ$. Specifically, $x \succeq y$ means that x is clockwise of y on the circle. We believe that this operator is completely intuitive and requires no formal definition, but in case of need the reader may find such a definition in Schummer and Vohra [19].

Each agent $i \in N$ has an (ideal) *location* $x_i \in G$. The collection $\mathbf{x} = \langle x_1, \ldots, x_n \rangle \in G^n$ is referred to as the *location profile*.

A *deterministic mechanism* is a function $f\colon G^n \to G$ that maps the reported locations of the agents to the location of a *facility*. When the facility is located at $y \in G$, the cost of agent i is simply the distance between x_i and y:

$$\text{cost}(y, x_i) = d(x_i, y).$$

A *randomized mechanism* is a function $f\colon G^n \to \Delta(G)$, i.e., it maps location profiles to probability distributions over G (which randomly designate the location of the facility). If $f(\mathbf{x}) = P$, where P is a probability distribution over G, then the cost of agent i is the expected distance from x_i:

$$\text{cost}(P, x_i) = \mathbb{E}_{y\sim P}[d(x_i, y)].$$

A mechanism is said to be *strategyproof* (SP) if agents can never benefit by lying. Formally, for every $\mathbf{x} \in G^n$, $i \in N$, and deviation $x_i' \in G$, it holds that $\text{cost}(f(x_i', \mathbf{x}_{-i}), x_i) \geq \text{cost}(f(\mathbf{x}), x_i)$, where $\mathbf{x}_{-i} = \langle x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \rangle$ is the vector of locations excluding x_i.

We are interested in optimizing the minimax, that is, minimizing the maximum cost. The *maximum cost* of y with respect to $\mathbf{x}$ is $\text{mc}(y, \mathbf{x}) = \max_{i\in N} \text{cost}(y, x_i)$; the maximum cost of a distribution P with respect to $\mathbf{x}$ is naturally defined as $\text{mc}(P, \mathbf{x}) = \mathbb{E}_{y\sim P}[\text{mc}(y, \mathbf{x})]$.

[2] If x and y are antipodal, these arcs are ambiguously defined. When this is problematic, we specify to which arc we are referring.

2.1. Deterministic mechanisms. We note that the problem of designing an optimal SP mechanism is very simple with respect to deterministic mechanisms. Consider the mechanism given by $f(\mathbf{x}) = x_1$ for all $\mathbf{x} \in G^n$, that is, a dictatorship of agent 1. This mechanism is clearly SP. Crucially, this mechanism does quite well in terms of the minimax: It provides a 2-approximation. To see this, given $\mathbf{x} \in G^n$, let y be an optimal facility location. Then, for all $i \in N$,

$$d(x_1, x_i) \leq d(x_1, y) + d(y, x_i) \leq 2 \cdot \max\{d(y, x_1), d(y, x_i)\} \leq 2 \cdot \mathrm{mc}(y, \mathbf{x}).$$

On the other hand, a deterministic SP mechanism cannot achieve an approximation ratio better than two, even if the underlying graph G is a line (Procaccia and Tennenholtz [15]). Because in a general graph any edge is locally a line, this lower bound applies to any graph. In other words, dictatorship gives a tight SP upper bound. In the sequel, we shall therefore restrict our attention to randomized mechanisms.

3. Randomized mechanisms on circles. We presently consider the case where the graph G is a circle. An important remark is that, even using randomization, we cannot hope to achieve an SP approximation ratio better than 3/2. Indeed, Procaccia and Tennenholtz [15] have established that a randomized SP mechanism does not yield an approximation ratio smaller than 3/2 on a line.

Furthermore, they have provided a straightforward matching SP upper bound of 3/2 on a line using the *left-right-middle* (*LRM*) *mechanism*: Given $\mathbf{x} \in G^n$ with probability 1/4, return the leftmost agent $\min_{i\in N} x_i$; with probability 1/4, return the rightmost agent $\max_{i\in N} x_i$; and with probability 1/2, return the midpoint of the interval between them, that is,

$$\mathrm{cen}\left(\min_{i\in N} x_i, \max_{i\in N} x_i\right) = \frac{\min_{i\in N} x_i + \max_{i\in N} x_i}{2}.$$

The idea behind the strategyproofness of this mechanism is very simple: An agent can only affect the outcome of the mechanism by deviating to a location $x_i' < \min_{i\in N} x_i$ or $x_i' > \max_{i\in N} x_i$. In this case, the agent pushes the left or right boundaries *away* from its location by δ but, in doing so, may push the midpoint toward its own location by $\delta/2$. Because the midpoint is selected with probability exactly twice that of each of the boundaries, the two terms cancel out.

Of course, when the agents are on a circle, in general it is meaningless to refer to the "leftmost" or "rightmost" agent. However, because any semicircle can naturally be treated as an interval (in the sense that there are two extreme agent locations), it is plausible to think of a hybrid mechanism that uses the LRM mechanism when the agents are located on one semicircle and uses another SP mechanism otherwise.

An obvious candidate for an SP mechanism for the case in which the agents are not all on one semicircle is the *random point* (*RP*) *mechanism*, which chooses a point on the circle uniformly at random. The RP mechanism is trivially SP and gives an approximation ratio of 7/4 for that case. To informally see this, observe that the worst approximation ratio is obtained when many agents are uniformly distributed over slightly more than a semicircle. Assume without loss of generality that the circumference of the circle is one. Then, if the mechanism chooses a point outside the semicircle (this happens with probability roughly 1/2), the maximum cost is roughly 1/2 because there is an agent opposite that point on the circle. If the mechanism chooses a point on the semicircle (this also happens with probability roughly 1/2), the expected cost is roughly 3/8. Taken together, the expected maximum cost of the mechanism is roughly 7/16 (which is the sum $(1/2)\cdot(1/2) + (1/2)\cdot(3/8)$). Obviously, the optimal location is the center of the semicircle, which provides a maximum cost of roughly 1/4. This gives an approximation ratio of 7/4. Therefore, a hybrid mechanism that uses the LRM mechanism when the agents are located on one semicircle and uses the RP mechanism otherwise gives an approximation ratio of 7/4.

A priori, it seems that there is no reason to suppose that such hybridizations of two SP mechanisms may result in an SP mechanism. Nevertheless, quite surprisingly the hybrid mechanism just proposed is SP.

First, observe that under the RP mechanism the expected cost of each agent is exactly 1/4. Thus, if the agents are not all on one semicircle, so long as this remains the case no agent can benefit by deviating. Likewise, it is not difficult to verify that if all agents remain on one semicircle, then a deviation cannot be beneficial (although this fact does not follow directly from the strategyproofness of the LRM mechanism on a line). The unlikely phenomenon is that an agent cannot benefit from a deviation, even if by deviating it triggers a shift from the RP mechanism to the LRM mechanism or vice versa.

To see this, we first informally observe that if all the agents are located on one semicircle, no agent can benefit by deviating in a way that triggers a shift to the RP mechanism. Indeed, it suffices to show that an agent's cost under the LRM mechanism cannot exceed 1/4. Note that the highest cost is obtained when the agents are distributed along exactly one semicircle, in which case each one of the extreme agents incurs an expected cost of 1/4.

To deal with the opposite direction, it suffices to show that if the agents are not on one semicircle, an agent $i \in N$ deviating in a way that triggers the LRM mechanism will incur a cost of at least 1/4. Let l (for "left") and r (for "right") be the two extreme agent locations in the new profile, where $l \succeq r$. Note that because the agents were not on one semicircle under the original profile, we have that $x_i \in (\hat{l}, \hat{r})$. Let $y = \text{cen}(l, r)$ be the center of the arc (l, r) (see Figure A.3 later in this paper). We claim that $d(x_i, y) \geq 1/4$. This follows immediately from the fact that $d(\hat{l}, y) \geq 1/4$, $d(\hat{r}, y) \geq 1/4$, and $x_i \in (\hat{l}, \hat{r})$. Hence, the cost of agent i after the deviation is at least

$$\text{cost}(\text{lrm}(\mathbf{x}'), x_i) = \tfrac{1}{4} \cdot d(x_i, l) + \tfrac{1}{4} \cdot d(x_i, r) + \tfrac{1}{2} \cdot d(x_i, y) \geq \tfrac{1}{4} \cdot (d(x_i, l) + d(x_i, r)) + \tfrac{1}{2} \cdot \tfrac{1}{4} \geq \tfrac{1}{4},$$

where the last transition follows from the fact that $d(x_i, l) + d(x_i, r)$ is the length of the *long* arc between l and r. Therefore, the value of this sum is at least 1/2. In conclusion, this hybrid mechanism is SP and provides an approximation ratio of 7/4.

Can one do better by building on the above ideas? We next propose a more sophisticated mechanism that is, like the last mechanism, a hybridization of two mechanisms: the LRM mechanism when the agents are located on one semicircle and the *random center (RC) mechanism* (defined next) when the agents are not on one semicircle.

Mechanism 3.1. *Given* $\mathbf{x} \in G^n$:

(i) *If* $\mathbf{x}$ *is such that the agents are located on one semicircle, i.e., there exist* $y, z \in G$ *such that for all* $i \in N$, $x_i \in [y, z]$, *then we execute the LRM mechanism on the arc* $[y, z]$, *treating it as an interval with the boundaries* $y < z$.

(ii) *If* $\mathbf{x}$ *is such that the agents are not located on one semicircle, we execute the random center (RC) mechanism, defined as follows.*

(a) *Randomly select a point* $y \in G$.

(b) *Let* $\hat{x}_i$ *and* $\hat{x}_j$ *be the two antipodal points adjacent to* y, *that is,* $\hat{x}_i$ *is the first antipodal point encountered when walking clockwise from* y *and* $\hat{x}_j$ *is the first antipodal point encountered when walking counterclockwise from* y.

(c) *Return* $\text{cen}(\hat{x}_i, \hat{x}_j)$.

An equivalent way of thinking about the RC mechanism is letting the mechanism choose the center of an interval between two adjacent antipodal points with probability proportional to the length of the interval. The RC mechanism may seem unintuitive; the idea behind it is the observation that the optimal solution on a circle is the center of the longest arc between any two adjacent antipodal points.

Some technical comments regarding Mechanism 3.1 are in order. Regarding the first item, there may be many choices of y and z such that $x_i \in [y, z]$ for all $i \in N$, but the LRM mechanism is indifferent to the choice. In the context of the second item, it holds that $y \in [\hat{x}_i, \hat{x}_j]$ by the assumption that in $\mathbf{x}$ the agents are not on one semicircle, i.e., y is on the same arc whose center we return. Furthermore, the RC mechanism is ambiguously defined when the random point y is an antipodal point itself, but this happens with probability zero.

The following theorem shows that Mechanism 3.1 is SP and provides the best possible approximation ratio.

Theorem 3.1. *Assume that* G *is a circle. Then, Mechanism* 3.1 *is an SP 3/2-approximation mechanism for the minimax.*

The nontrivial part of the proof of Theorem 3.1 is the strategyproofness of the mechanism. Although the proof is long, it revolves around several basic properties of the RC mechanism. Very generally speaking, one of the ideas at the core of the proof is that the locations occupied by the agents in $\mathbf{x}$ are special, but only in the trivial sense that for every $i \in N$ the antipodal point $\hat{x}_i$ is among the antipodal points employed by the RC mechanism. For this reason, the cost of an agent (assuming that the circumference of the circle is one) under the RC mechanism is at most 1/4 (Lemma A.5 in this paper). A second important idea is that the expected cost of agent i under the RC mechanism is essentially the same as if the mechanism were choosing uniformly on the circle, except for its behavior on the arc between the two antipodal points adjacent to x_i (Lemma A.4 in this paper). The detailed proof is given in Appendix A.

4. Randomized mechanisms on trees. In the following, we assume that the graph G is a tree. We first observe that randomization allows us to do slightly better than dictatorship, especially when the number of agents is small. Indeed, given $\mathbf{x} \in G^n$, the *center of* G *with respect to* $\mathbf{x}$ is a point

$$y \in \arg\min_{z \in G} \text{mc}(z, \mathbf{x})$$

that optimizes the minimax. It is easy to verify that when G is a tree, the center is unique.[3] Therefore, we can denote the (unique) center of G with respect to $\mathbf{x}$ by $\text{cen}(G, \mathbf{x})$.

We consider the following mechanism: Given $\mathbf{x} \in G^n$, the distribution on the returned location gives probability $1/(n+2)$ to x_i for each $i \in N$ and probability $2/(n+2)$ to $\text{cen}(G, \mathbf{x})$. The fact that the mechanism is SP follows from the fact that when agent i deviates from x_i to x_i', it holds that

$$\left|d(x_i, \text{cen}(G, \mathbf{x})) - d(x_i, \text{cen}(G, \mathbf{x}'))\right| \leq \frac{d(x_i, x_i')}{2}$$

and, therefore, denoting the above mechanism by f, we have

$$\text{cost}(f(\mathbf{x}'), x_i) - \text{cost}(f(\mathbf{x}), x_i) \geq \frac{1}{n} \cdot d(x_i, x_i') - \frac{2}{n} \cdot \frac{d(x, x_i')}{2} = 0.$$

The approximation ratio of the mechanism satisfies

$$\frac{\text{mc}(f(\mathbf{x}), \mathbf{x})}{\text{OPT}} \leq \frac{(2/(n+2)) \cdot \text{OPT} + (n/(n+2)) \cdot 2 \cdot \text{OPT}}{\text{OPT}} = 2 - \frac{2}{n+2}.$$

Despite this small improvement over dictatorship, we shall demonstrate that we cannot do significantly better. In other words, our second major result asserts that an SP mechanism cannot achieve an approximation ratio that is bounded away from two for the minimax, even on trees.

THEOREM 4.1. *Let $N = \{1, \ldots, n\}$. Then, there exists a tree G such that no SP randomized mechanism can have an approximation ratio that is smaller than $2 - \Theta(1/2^{\sqrt{\log n}})$ for the minimax.*

The proof of Theorem 4.1 is given in Appendix B.

5. Discussion. The results we present are tight and rather exhaustive[4] as long as one is interested in optimizing the minimax. However, as noted in the introduction, the minisum is an equally interesting and prominent measure. In an expanded version of this paper, we investigate the minisum as well (Alon et al. [1]). Selecting the ideal location of one of the agents uniformly at random provides an approximation ratio of $2 - 2/n$ for the minisum. This mechanism is clearly SP. Moreover, we establish that when the network is a circle, it possesses a stronger game theoretic property called *group strategyproofness*; even a coalition of agents cannot benefit by jointly misreporting their locations. Specifically, there must be at least one member of the coalition that does not gain from the deviation. This last result is of theoretical interest, although it is not as meaningful as our minimax results from a mechanism design point of view.

Future work can take several directions. First, we note that our Mechanism 3.1 is not group SP, even when the agents are assumed not to be located on one semicircle before and after the deviation (that is, the RC mechanism is used in both cases). The counterexample is due to Dror Shemesh and Omri Sivan.[5] Is there a group SP 3/2-approximation mechanism on a circle?

Second, Procaccia and Tennenholtz [15] show that when two facilities must be located and the cost of each agent is its distance to the nearest facility, providing SP approximations for the minisum and minimax becomes much more complicated, even on a line. As mentioned in the introduction, Lu et al. [12] provide some results for this setting with respect to the minisum and general metric spaces or circles. It remains open to tighten the results of Lu et al. [12], extend them to the minimax, and, most importantly, generalize them to the case of k facilities for $k > 2$.

Third, recall that many measures were featured in the literature. For example, minimizing the sum of squared costs is known to be the unique solution that satisfies desirable axiomatic properties when the graph is a tree (Holzman [9]). Can our results be generalized in a way that captures multiple measures? In more detail, Marsh and Schilling [13] provide a framework in which the different equity measures can be organized; this framework consists of three dimensions. One could hope for a result that relates feasible SP approximation ratios to the parameters along the different dimensions.

[3] This would not be true in a discrete graph model but this issue can still be easily circumvented.

[4] One could consider, though, restricted graph topologies (e.g., trees with bounded branching factor) that circumvent Theorem 4.1 and possibly allow for additional positive results.

[5] Consider an instance with three agents located at a distance of 0.05, 0.5, and 0.75 from some point on a circle with circumference 1. The costs under truthtelling are 0.2475, 0.2475, and 0.21, respectively. Yet, if the first and second agents deviate to 0.1 and 0.45, respectively, the newly incurred cost by each one of them is reduced to 0.2425.

Fourth, the most interesting as well as most challenging question concerns the extension of the results of Schummer and Vohra [19] to randomized mechanisms, that is, obtaining a characterization of randomized SP mechanisms on a network. Presumably, such a characterization would immediately yield tight bounds on the feasible randomized SP approximation ratios under any conceivable measure. Because such a characterization would, in particular, have to capture nonintuitive mechanisms like Mechanism 3.1, this goal seems to be very ambitious; the deterministic results of Schummer and Vohra [19] are already quite intricate!

Acknowledgments. The authors thank the anonymous referees and the anonymous associate editor for very helpful comments. Alon is supported in part by a USA-Israeli BSF grant, by a grant from the Israel Science Foundation, by an ERC advanced grant, and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University. The work was done while Procaccia was at Microsoft Israel R&D Center.

Appendix A. Proof of Theorem 3.1. In the proof, we assume without loss of generality that the circumference of the circle G is one. In addition, we denote the outcome of the LRM mechanism and the RC mechanism given $\mathbf{x} \in G^n$ by $\text{lrm}(\mathbf{x})$ and $\text{rc}(\mathbf{x})$, respectively.

Lemma A.1. *Mechanism* 3.1 *is a* 3/2*-approximation mechanism for the minimax.*

Proof. Assume first that $\mathbf{x} \in G^n$ is such that not all agents are located on one semicircle. Let α be the length of the longest arc between two adjacent agents, and assume without loss of generality that these agents are agents 1 and 2. Because the agents are not located on one semicircle, it holds that $\alpha \leq 1/2$. It can be verified that the optimal facility location is $\text{cen}(\hat{x}_1, \hat{x}_2)$: hence, we have that $\text{OPT} = (1-\alpha)/2$. The mechanism selects the optimal solution with probability α, and with probability $1-\alpha$ selects a solution with maximum cost at most 1/2. Therefore, the approximation ratio is at most

$$\frac{\text{mc}(\text{rc}(\mathbf{x}), \mathbf{x})}{\text{OPT}} \leq \frac{\alpha \cdot (1-\alpha)/2 + (1-\alpha)\cdot(1/2)}{(1-\alpha)/2} = 1 + \alpha \leq \frac{3}{2}. \tag{A1}$$

If $\mathbf{x}$ is such that all the agents are located on one semicircle, then the LRM mechanism is applied. We choose the optimal location with probability 1/2 and a location with twice the optimal cost with probability 1/2. Hence, the approximation ratio is, once again, 3/2. □

In order to establish the strategyproofness of Mechanism 3.1, we must examine four types of "lies": an agent deviating such that before the deviation all the agents were located on one semicircle, and after the deviation they are located on one semicircle—"semicircle to semicircle" (the LRM mechanism is applied to both); not semicircle to semicircle (RC to LRM); semicircle to not semicircle (LRM to RC); and not semicircle to not semicircle (RC to RC). The semicircle to semicircle case is relatively straightforward, and we tackle it first.

Lemma A.2 (Semicircle to Semicircle). *Assume that* $\mathbf{x} \in G^n$ *is such that the agents are on one semicircle, and agent* i *deviates such that in the new location profile* $\mathbf{x}'$ *the agents are also on one semicircle. Then,*

$$\text{cost}(\text{lrm}(\mathbf{x}), x_i) \leq \text{cost}(\text{lrm}(\mathbf{x}'), x_i).$$

Proof. Let $x_1 \succeq x_2 \succeq \cdots \succeq x_n$ and denote $x_1 = l$ (for "left") and $x_n = r$ (for "right"). Suppose agent i deviates from x_i to x_i' such that the agents are on one semicircle. If (x_i', x_i) intersects with the new semicircle, the proof follows directly from the fact that the LRM mechanism is SP when applied to an interval (Procaccia and Tennenholtz [15]).

It is easy to verify that (x_i', x_i) may lie in the complement of the new semicircle only if (i) the deviating agent is r and $x_i' \succ l$ or (ii) the deviating agent is l and $x_i' \prec r$. We prove Lemma A.2 for the former case but note that the latter case is completely analogous.

Indeed, suppose without loss of generality that $x_i = r$ and $x_i' \succ l$. Let r' be the adjacent agent to r such that $r' \succ r$. It must hold that $x_i' \preceq \hat{r}'$ (otherwise, in the new location profile, the agents are not all on one semicircle). We denote $\alpha = d(r, r')$, $\beta = d(x_i', l)$, and $\gamma = d(l, r)$ (see Figure A.1). From the assumptions of Lemma A.2, it follows that $\gamma \leq 1/2$. The short arc from x_i' to x_i lies in the complement of the new semicircle if and only if $1 - \beta - \gamma \leq \beta + \gamma$, that is, $\beta + \gamma \geq 1/2$.

We first calculate the cost of x_i in the location profile $\mathbf{x}$:

$$\text{cost}(\text{lrm}(\mathbf{x}), x_i) = \frac{1}{4}\cdot\gamma + \frac{1}{4}\cdot 0 + \frac{1}{2}\cdot\frac{\gamma}{2} = \frac{\gamma}{2}.$$

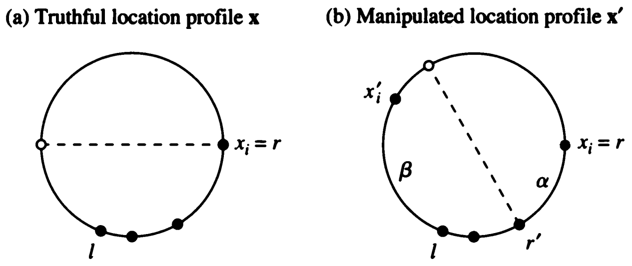

FIGURE A.1. Illustration of the proof of Lemma A.2.

We distinguish between two cases:

Case 1. $\alpha \le 1-\beta-\gamma$. In this case, the cost of the LRM mechanism in the new location profile with respect to x_i is

$$\text{cost}(\text{lrm}(\mathbf{x}'), x_i) = \frac{1}{4}\cdot\alpha + \frac{1}{4}\cdot(1-\beta-\gamma) + \frac{1}{2}\cdot\left(\alpha + \frac{\gamma-\alpha+\beta}{2}\right) = \frac{1}{4}+\frac{\alpha}{2}.$$

It holds that

$$\text{cost}(\text{lrm}(\mathbf{x}), x_i) \le \text{cost}(\text{lrm}(\mathbf{x}'), x_i) \quad\Leftrightarrow\quad \gamma \le \tfrac{1}{2}+\alpha,$$

but this holds because $\gamma \le 1/2$.

Case 2. $\alpha > 1-\beta-\gamma$. In this case, the cost of the left-middle-right mechanism in the new location profile with respect to x_i is

$$\text{cost}(\text{lrm}(\mathbf{x}'), x_i) = \frac{1}{4}\cdot\alpha + \frac{1}{4}\cdot(1-\beta-\gamma) + \frac{1}{2}\cdot\left(1-\beta-\gamma + \frac{\gamma-\alpha+\beta}{2}\right) = \frac{3}{4}-\frac{\beta}{2}-\frac{\gamma}{3}.$$

It holds that

$$\text{cost}(\text{lrm}(\mathbf{x}), x_i) \le \text{cost}(\text{lrm}(\mathbf{x}'), x_i) \quad\Leftrightarrow\quad \frac{\gamma}{2} \le \frac{3}{4}-\frac{\beta}{2}-\frac{\gamma}{2} \quad\Leftrightarrow\quad \gamma \le \frac{3}{4}-\frac{\beta}{2}.$$

Because $\gamma \le 1/2$, it is sufficient to show that $1/2 \le 3/4 - \beta/2$, which holds if and only if $\beta \le 1/2$. The last inequality follows from the fact that all the agents are on one semicircle after the deviation; this concludes the proof of the Lemma A.2. □

The other three deviations are more complicated and their proofs require the laying of some foundations first. We start with two simple lemmas concerning partitions of intervals on the line, both of which will prove useful in several points in the sequel. On a circle, the arc between two points $x, y \in G$ such that $d(x,y) \le 1/2$ can be regarded as an interval.

LEMMA A.3. *Let* $x, y \in [0,1] \subset \mathbb{R}$. *It holds that*

$$d(x, 1/2) \le y\cdot d(x, \text{cen}(0,y)) + (1-y)\cdot d(x, \text{cen}(y,1)).$$

PROOF. Assume without loss of generality that $x \le 1/2$. We distinguish between two cases.

Case 1: $y/2 \ge x$. In this case,

$$y\cdot d(x,\text{cen}(0,y)) + (1-y)\cdot d(x,\text{cen}(y,1)) = y\cdot\left(\frac{y}{2}-x\right) + (1-y)\cdot\left(y+\frac{1-y}{2}-x\right) = \frac{1}{2}-x.$$

Case 2: $y/2 < x$. We have that

$$y\cdot d(x,\text{cen}(0,y)) + (1-y)\cdot d(x,\text{cen}(y,1)) = y\cdot\left(x-\frac{y}{2}\right) + (1-y)\cdot\left(y+\frac{1-y}{2}-x\right) = \frac{1}{2}-x+(2xy-y^2).$$

Because $y/2 < x$, it holds that $2xy > y^2$. Hence, $1/2 - x + (2xy - y^2) > 1/2 - x$. □

LEMMA A.4. *Let* $y_1, \ldots, y_{m+1} \in [0, 1/2]$ *such that* $y_1 = 0$, $y_{m+1} = 1/2$. *For all* $i = 1, \ldots, m$, *define* $d_i = y_{i+1} - y_i$. *Then,*

$$\sum_{i=1}^{m}\left(d_i\cdot\left(\sum_{j=1}^{i-1} d_j + \frac{d_i}{2}\right)\right) = \frac{1}{8}.$$

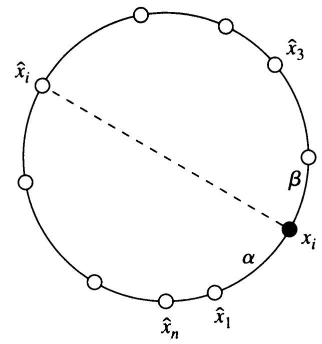

FIGURE A.2. Illustration of the proof of Lemma A.5.

The intuition behind Lemma A.4 is that choosing the center of each interval with probability equal to the length of the interval is like randomly choosing a point in $[0, 1/2]$ with probability $1/2$. The expected distance of a random point in $[0, 1/2]$ from 0 is $1/4$, multiplying by $1/2$ yields $1/8$.

PROOF. By reorganizing the terms, it can be verified that

$$\sum_{i=1}^{m}\left(d_i\cdot\left(\sum_{j=1}^{i-1}d_j+\frac{d_i}{2}\right)\right)=\frac{(\sum_{i=1}^{m}d_i)^2}{2}=\frac{1}{8}. \quad \square$$

The next lemma implies that under the RC mechanism, the cost of an agent is atmost $1/4$. It follows from the two previous lemmas.

LEMMA A.5. *For all* $\mathbf{x}\in G^n$ *such that* $\mathbf{x}$ *is not on one semicircle and for all* $i\in N$,

$$\operatorname{cost}(\operatorname{rc}(\mathbf{x}), x_i)\le \tfrac{1}{4}.$$

PROOF. Assume that the locations of the agents satisfy $\hat{x}_1\succeq\hat{x}_2\succeq\cdots\succeq\hat{x}_n\succeq\hat{x}_1$, and that there are no antipodal points in $(\hat{x}_i, \hat{x}_{i+1})$ for all $i\in N$. Let $i\in N$ and assume without loss of generality that $x_i\in[\hat{x}_1, \hat{x}_2]$. In addition, let $\alpha=d(\hat{x}_1, x_i)$ and $\beta=d(x_i, \hat{x}_2)$ (see Figure A.2).

We wish to calculate the cost of the RC mechanism with respect to agent i. We can break this cost down into two components: the cost when choosing a point in the arc $[\hat{x}_1, \hat{x}_2]$ (this happens with probability $\alpha+\beta$) and the rest of the cost. Notice that $x_i\in[\hat{x}_1, \hat{x}_2]$, i.e., the length of the arc between $\hat{x}_1$ and $\hat{x}_2$ that includes x_i is at most 1/2; this holds by the assumption that the agents are not on one semicircle. Therefore, we can treat $[\hat{x}_1, \hat{x}_2]$ as an interval. By normalizing $d(\hat{x}_1, \hat{x}_2)$, we get from Lemma A.3 that if we partition the arc $[\hat{x}_1, \hat{x}_2]$ into two arcs of length y and $1-y$ and choose the center of each with probability y and $1-y$, respectively, the cost of agent i can only increase. In particular, this is true when $y=\alpha$. Hence, we have that

$$\operatorname{cost}(\operatorname{rc}(\mathbf{x}), x_i)\le\operatorname{cost}(\operatorname{rc}(\mathbf{x}\cup\{\hat{x}_i\}), x_i), \tag{A2}$$

where $\mathbf{x}\cup\{\hat{x}_i\}$ is the profile $\mathbf{x}$ with an additional agent at $\hat{x}_i$, that is, an additional antipodal point at x_i. It is sufficient to show that

$$\operatorname{cost}(\operatorname{rc}(\mathbf{x}\cup\{\hat{x}_i\}), x_i)\le\tfrac{1}{4}.$$

The expression $\operatorname{cost}(\operatorname{rc}(\mathbf{x}\cup\{\hat{x}_i\}), x_i)$ is the expected distance from agent i when the center of one of the arcs

$$[x_i, \hat{x}_2], [\hat{x}_2, \hat{x}_3], [\hat{x}_3, \hat{x}_4], \ldots, [\hat{x}_n, \hat{x}_1], [\hat{x}_1, x_i] \tag{A3}$$

is chosen, where the probability of choosing the center of an arc is its length. In order to make this explicit, denote $d_i=d(\hat{x}_i, \hat{x}_{i+1})$ for $i=1,\ldots,n-1$, $d_n=d(\hat{x}_n, \hat{x}_1)$. We have

$$\operatorname{cost}(\operatorname{rc}(\mathbf{x}\cup\{\hat{x}_i\}), x_i)=\frac{\alpha^2}{2}+\frac{\beta^2}{2}+\sum_{j=2}^{i-1}\left(d_j\cdot\left(\beta+\sum_{k=2}^{j-1}d_k+\frac{d_j}{2}\right)\right)+\sum_{j=i}^{n}\left(d_j\cdot\left(\alpha+\sum_{k=j+1}^{n}d_k+\frac{d_j}{2}\right)\right). \tag{A4}$$

We partition the expression in the right-hand side of Equation (A4) into two sums, each corresponding to the cost of the mechanism on a semicircle of length 1/2, and we apply Lemma A.4 to each, with $y_1=x_i$ in both cases and $d_1=\alpha$ or $d_1=\beta$. In more detail, it holds that

$$\frac{\beta^2}{2}+\sum_{j=2}^{i-1}\left(d_j\cdot\left(\beta+\sum_{k=2}^{j-1}d_k+\frac{d_j}{2}\right)\right)=\frac{1}{8}$$

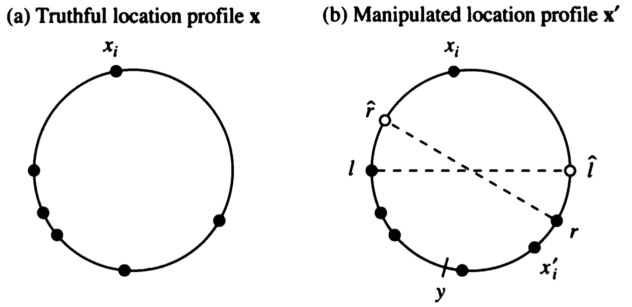

FIGURE A.3. Illustration of the proof of Lemma A.6.

and

$$\frac{\alpha^2}{2}+\sum_{j=i}^{n}\left(d_j\cdot\left(\alpha+\sum_{k=j+1}^{n}d_k+\frac{d_j}{2}\right)\right)=\frac{1}{8}.$$

We conclude that the expression on the right-hand side of Equation (A4) is exactly 1/4. □

LEMMA A.6 (NOT SEMICIRCLE TO SEMICIRCLE). *Assume that* $\mathbf{x}\in G^n$ *is such that the agents are not on one semicircle, but agent i deviates such that in the new location profile* $\mathbf{x}'$ *the agents are on one semicircle. Then,*

$$\operatorname{cost}(\operatorname{rc}(\mathbf{x}), x_i)\le \operatorname{cost}(\operatorname{lrm}(\mathbf{x}'), x_i).$$

PROOF. By Lemma A.5, we have that $\operatorname{cost}(\operatorname{rc}(\mathbf{x}), x_i)\le 1/4$. Therefore, it will be sufficient to prove that $\operatorname{cost}(\operatorname{lrm}(\mathbf{x}'), x_i)\ge 1/4$.

Let l (for "left") and r (for "right") be the two extreme agent locations in $\mathbf{x}'$, where $l\succeq r$. Note that because the agents were not on one semicircle under $\mathbf{x}$, we have that $x_i\in(\hat{l},\hat{r})$. Let $y=\operatorname{cen}(l,r)$ be the center of the arc (l,r) (see Figure A.3). We claim that $d(x_i,y)\ge 1/4$. This follows immediately from the facts that $d(\hat{l},y)\ge 1/4$, $d(\hat{r},y)\ge 1/4$, and $x_i\in(\hat{l},\hat{r})$. Hence, the cost of the mechanism is at least

$$\operatorname{cost}(\operatorname{lrm}(\mathbf{x}'), x_i)=\tfrac{1}{4}\cdot d(x_i,l)+\tfrac{1}{4}\cdot d(x_i,r)+\tfrac{1}{2}\cdot d(x_i,y)\ge\tfrac{1}{4}\cdot(d(x_i,l)+d(x_i,r))+\tfrac{1}{2}\cdot\tfrac{1}{4}\ge\tfrac{1}{4},$$

where the last transition follows from the fact that $d(x_i,l)+d(x_i,r)$ is the length of the *long* arc between l and r. Therefore, the value of this sum is at least 1/2. □

In order to deal with the last two deviations, we require one additional fundamental lemma. The lemma asserts that the cost of the RC mechanism with respect to a point y can only decrease if the point y is added to the vector of locations. This is, in fact, the mirror image of Equation (A2), which is itself a special case of Lemma A.3.

LEMMA A.7. *Let* $\mathbf{x}\in G^n$ *such that* $\mathbf{x}$ *is not on one semicircle and let* $y\in G$. *Then,*

$$\operatorname{cost}(\operatorname{rc}(\mathbf{x}), y)\ge\operatorname{cost}(\operatorname{rc}(\mathbf{x}\cup\{y\}), y).$$

PROOF. The cost incurred from the mechanism $\operatorname{rc}(\mathbf{x})$ with respect to y is identical to the cost incurred from $\operatorname{rc}(\mathbf{x}\cup\{y\})$ for all the intervals except for the interval the point $\hat{y}$ is on. Let p and q denote the antipodal points adjacent to $\hat{y}$ such that $p\preceq\hat{y}\preceq q$. Denote $\delta=d(p,\hat{y})$ and $\lambda=d(\hat{y},q)$, and assume without loss of generality that $\delta\le\lambda$ (see Figure A.4).

It is sufficient to show that the cost incurred when the random chosen point is on the arc $[p,q]$ is lower under $\operatorname{rc}(\mathbf{x}\cup\{y\})$ than under $\operatorname{rc}(\mathbf{x})$. The cost incurred by points on the arc $[p,q]$ under $\operatorname{rc}(\mathbf{x}\cup\{y\})$ is $\delta(d(y,p)+\delta/2)+\lambda(d(y,q)+\lambda/2)$. The cost incurred by points on the arc $[p,q]$ under $\operatorname{rc}(\mathbf{x})$ is $(\delta+\lambda)(d(y,q)+(\delta+\lambda)/2)$. It holds that

$$(\delta+\lambda)\left(d(y,q)+\frac{\delta+\lambda}{2}\right)\ge\delta\left(d(y,p)+\frac{\delta}{2}\right)+\lambda\left(d(y,q)+\frac{\lambda}{2}\right)\;\Leftrightarrow\; d(y,q)+\lambda\ge d(y,p).$$

The inequality on the right-hand side holds because $d(y,q)+\lambda=1/2$ while $d(y,p)\le 1/2$. □

LEMMA A.8 (SEMICIRCLE TO NOT SEMICIRCLE). *Assume that* $\mathbf{x}\in G^n$ *is such that the agents are on one semicircle, but agent i deviates such that in the new location profile* $\mathbf{x}'$ *the agents are not on one semicircle. Then,*

$$\operatorname{cost}(\operatorname{lrm}(\mathbf{x}), x_i)\le\operatorname{cost}(\operatorname{rc}(\mathbf{x}'), x_i).$$

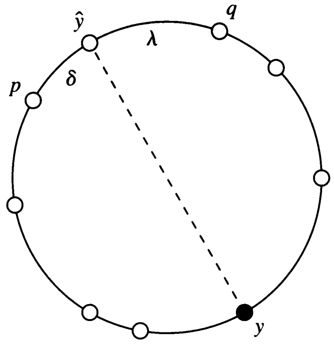

FIGURE A.4. Illustration of the proof of Lemma A.7.

PROOF. Let $x_1 \succeq x_2 \succeq \cdots \succeq x_n$ be the location of the n agents and denote $l = x_1$ (the leftmost agent) and $r = x_n$. Let $i \in N$ and assume without loss of generality that $d(x_i, r) \le d(x_i, l)$ or, equivalently, $d(x_i, \hat{l}) \le d(x_i, \hat{r})$. Let $\alpha = d(r, \hat{l})$ and $\beta = d(r, x_i)$. Finally, let x_i' denote the new location of agent i (see Figure A.5).

We first calculate the cost of x_i in the location profile $\mathbf{x}$ under the LRM mechanism. We have that $d(x_i, l) = 1/2 - \alpha - \beta$, $d(r, l) = 1/2 - \alpha$, and

$$d(x_i, \mathrm{cen}(l, r)) = \frac{1}{4} - \frac{\alpha}{2} - \beta.$$

Then,

$$\mathrm{cost}(\mathrm{lrm}(\mathbf{x}), x_i) = \frac{\beta}{4} + \frac{1/2 - \alpha - \beta}{4} + \frac{1/4 - \alpha/2 - \beta}{2} = \frac{1}{4} - \frac{\alpha}{2} - \frac{\beta}{2}. \tag{A5}$$

We now wish to give a lower bound on $\mathrm{cost}(\mathrm{rc}(\mathbf{x}'), x_i)$. First, by Lemma A.7, we have that

$$\mathrm{cost}(\mathrm{rc}(\mathbf{x}'), x_i) \ge \mathrm{cost}\big(\mathrm{rc}(\mathbf{x}' \cup \{x_i\}), x_i\big).$$

A subtle remark is that the above argument allows us to take $\hat{r}$ into account even if $x_i = r$ and agent i deviated to x_i', a fact that allows us to avoid distinguishing this extreme case.

Next, we would like to fix an "optimal" location for $\hat{x}_i'$. First, notice that because $\mathbf{x}'$ is not on one semicircle, it must hold that $x_i' \in (\hat{r}, \hat{l})$, that is, $\hat{x}_i' \in (r, l)$. Now, we ask what is the location of $\hat{x}_i'$ that yields the lowest cost with respect to x_i when the RC mechanism chooses a point in $G \setminus (\hat{l}, \hat{r})$ (the long arc between $\hat{l}$ and $\hat{r}$)? There are no antipodal points other than $\hat{x}_i'$ in $G \setminus (\hat{l}, \hat{r})$. Furthermore, we can treat this arc as an interval with respect to distances from x_i because the short arcs $[x_i, \hat{l}]$ and $[x_i, \hat{r}]$ are contained in $G \setminus (\hat{l}, \hat{r})$. Therefore, we can apply Lemma A.3 (by normalizing the length of $G \setminus (\hat{l}, \hat{r})$). In particular, let $\hat{x}_i^* \succeq x_i$ such that

$$d(\hat{x}_i^*, x_i) = d(x_i, \hat{l}) = \alpha + \beta$$

and let $\mathbf{x}^* = \langle x_i^*, x_{-i} \rangle$ (see Figure A.5 for an illustration). From (the proof of) Lemma A.3, it follows that

$$\mathrm{cost}\big(\mathrm{rc}(\mathbf{x}' \cup \{x_i\}), x_i\big) \ge \mathrm{cost}\big(\mathrm{rc}(\mathbf{x}^* \cup \{x_i\}), x_i\big).$$

Hence, it is sufficient to prove that

$$\mathrm{cost}\big(\mathrm{rc}(\mathbf{x}^* \cup \{x_i\}), x_i\big) \ge \frac{1}{4} - \frac{\alpha}{2} - \frac{\beta}{2}. \tag{A6}$$

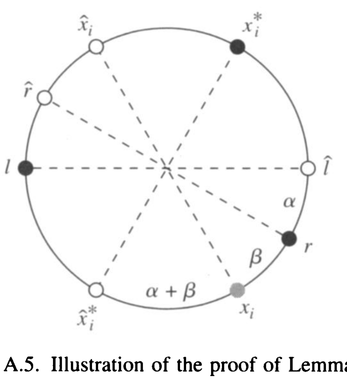

FIGURE A.5. Illustration of the proof of Lemma A.8.

We calculate the cost of the mechanism $\mathrm{rc}(\mathbf{x}^* \cup \{x_i\})$ with respect to agent i. Let y be the point chosen by the mechanism. By the choice of x_i^*, if y is on the arc $[\hat{l}, \hat{x}_i^*]$, the cost is zero. If $y \in [\hat{l}, \hat{x}_i]$, we apply Lemma A.4 with $y_1 = x_i$, $y_{m+1} = \hat{x}_i$, $y_j = \hat{x}_{j-1}$ for $j = 2, \ldots, m+1$, and $d_1 = \alpha + \beta$. We get

$$\frac{(\alpha+\beta)^2}{2} + \sum_{j=2}^{m}\left(d_j \cdot \left(\alpha+\beta+\sum_{k=2}^{j-1} d_k + \frac{d_j}{2}\right)\right) = \frac{1}{8}.$$

Thus,

$$\sum_{j=2}^{m}\left(d_j \cdot \left(\alpha+\beta+\sum_{k=2}^{j-1} d_k + \frac{d_j}{2}\right)\right) = \frac{1}{8} - \frac{(\alpha+\beta)^2}{2}. \tag{A7}$$

Similarly, if $y \in [\hat{x}_i, \hat{x}_i^*]$, we apply Lemma A.4 with $y_1 = x_i$, $y_2 = \hat{x}_i^*$, $y_{m+1} = \hat{x}_i$, $y_j = \hat{x}_{n-j+3}$ for $j = 3, \ldots, m+1$, and $d_1 = \alpha + \beta$. We get the same expected cost as in Equation (A7).

Taken together,

$$\mathrm{cost}(\mathrm{rc}(\mathbf{x}^* \cup \{x_i\}), x_i) = 2\left(\frac{1}{8} - \frac{(\alpha+\beta)^2}{2}\right) = \frac{1}{4} - (\alpha+\beta)^2.$$

It holds that

$$\frac{1}{4} - (\alpha+\beta)^2 \geq \frac{1}{4} - \frac{\alpha}{2} - \frac{\beta}{2} \quad \Leftrightarrow \quad \alpha+\beta \leq \frac{1}{2}$$

but the last inequality follows directly from the fact that $x_i \in [r, l]$. This establishes Equation (A6) and thus completes the proof of Lemma A.8. □

Lemma A.9 (Not Semicircle to Not Semicircle). *Assume that $\mathbf{x} \in G^n$ is such that the agents are not on one semicircle, and agent i deviates such that in the new location profile $\mathbf{x}'$ the agents are not on one semicircle. Then,*

$$\mathrm{cost}(\mathrm{rc}(\mathbf{x}), x_i) \leq \mathrm{cost}(\mathrm{rc}(\mathbf{x}'), x_i).$$

Proof. The proof of the lemma follows quite directly from the previous lemmata. Indeed, by Lemma A.7, we have that

$$\mathrm{cost}(\mathrm{rc}(\mathbf{x}' \cup \{x_i\}), x_i) \leq \mathrm{cost}(\mathrm{rc}(\mathbf{x}'), x_i).$$

Let $p, q \in G$ be the two antipodal points adjacent to x_i. By Lemma A.4, for any $\hat{x}_i' \in G \setminus [p, q]$,

$$\mathrm{cost}(\mathrm{rc}(\mathbf{x}), x_i) = \mathrm{cost}(\mathrm{rc}(\mathbf{x}' \cup \{x_i\}), x_i).$$

Hence, it sufficient to handle the case where $\hat{x}_i' \in [p, q]$. Notice that $x_i \in [p, q]$, that is, x_i is in the *short* arc between p and q by our assumption that the agents in $\mathbf{x}$ are not on one semicircle. Therefore, we can apply Lemma A.3 to this arc (by normalizing its length and replacing x in the lemma by x_i and y by $\hat{x}_i'$). It follows that the optimal location for $\hat{x}_i'$ is the edges of the arc $[p, q]$, which directly means that

$$\mathrm{cost}(\mathrm{rc}(\mathbf{x}), x_i) \leq \mathrm{cost}(\mathrm{rc}(\mathbf{x}' \cup \{x_i\}), x_i). \quad \square$$

Appendix B. Proof of Theorem 4.1. Let $m, k \in \mathbb{N}$, whose value remains to be determined. The construction of the graph G is recursive and depends on m and k. We start with an edge of length 1, which connects the vertices l^0 (for "left") and r^0 (for "right"). The vertex l^0 is connected to m vertices via edges of length 1; these vertices are called *left vertices on level* 1. Each left vertex on level 1 is connected to m vertices via m edges of length 2; these vertices are called *left vertices on level* 2. In general, each left vertex on level d is connected to m left vertices on level $d+1$ via edges of length 2^d. The maximum level is k, that is, the left vertices on level k are leaves. The construction is symmetric with respect to the right vertices, i.e., r^0 is connected to m right vertices on level 1 via edges of length 1 and so on.

Now, let $f: G^n \to \Delta(G)$ be a randomized SP mechanism. Assume for ease of exposition that $2m^k$ divides n and consider a location profile $\mathbf{x}^0$ where there are $n/2$ agents at l^0 and $n/2$ agents at r^0; this profile is illustrated in Figure A.6(a). Clearly, because the distance between l^0 and r^0 is 1, we have that $\mathbb{E}[d(f(\mathbf{x}^0), l^0)] \geq 1/2$ or $\mathbb{E}[d(f(\mathbf{x}^0), r^0)] \geq 1/2$; assume without loss of generality that the former statement is true.

Next, consider the location profile $\mathbf{x}^1$, where we have $n/2m$ agents in each left vertex on level 1 and $n/2$ agents in r^0; this profile is illustrated in Figure A.6(b). We claim that it still holds that $\mathbb{E}[d(f(\mathbf{x}^1), l^0)] \geq 1/2$. Indeed, $\mathbf{x}^1$ can be obtained from $\mathbf{x}^0$ by moving the agents, one by one, from l^0 to the left vertices on level 1. Because of strategyproofness, the expected distance from l^0 cannot decrease after each deviation. Because the expected distance was initially at least 1/2, this is true after the $n/2$ agents have all deviated.

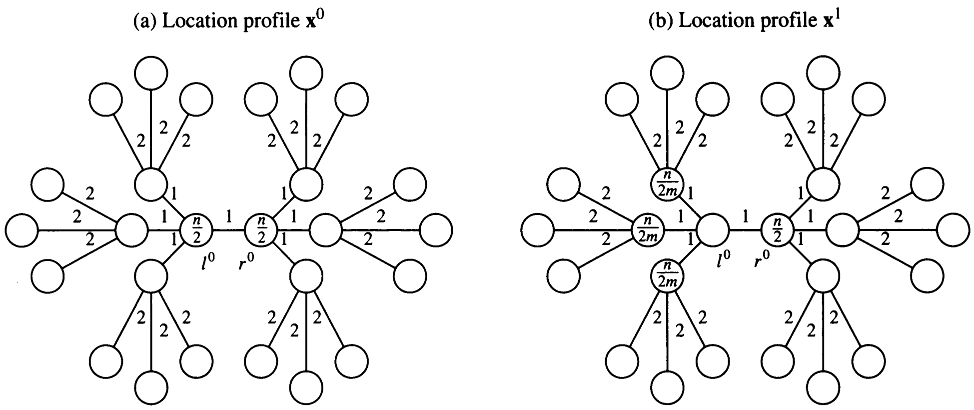

FIGURE A.6. An illustration of the proof of Theorem 4.1 for $m=3$ and $k=2$. A number inside a node indicates the number of agents that are located at this node.

Denote the left vertices on level 1 by $l_1^1, \ldots, l_m^1$. For any point $y \in G$ such that $d(y, l^0) \geq 1/2$, we have that

$$\sum_{j=1}^{m} d(y, l_j^1) \geq m + (m-2) \cdot \frac{1}{2} = \frac{3m-2}{2}.$$

Because this is true pointwise, it is also true that the sum of expected distances between $f(\mathbf{x}^1)$ and the left vertices on level 1 is at least $(3m-2)/2$, that is,

$$\sum_{j=1}^{m} \mathbb{E}[d(f(\mathbf{x}^1), l_j^1)] \geq \frac{3m-2}{2}.$$

By averaging over these m vertices, we conclude that there exists a left vertex on level 1 denoted l^1 such that $\mathbb{E}[d(f(\mathbf{x}^1), l^1)] \geq 3/2 - 1/m$.

We subsequently consider the location profile $\mathbf{x}^2$ that is obtained from $\mathbf{x}^1$ by moving the $n/(2m)$ agents from l^1 to its m neighbors on level 2 such that each left vertex on level 2 that is adjacent to l^1 has $n/(2m^2)$ agents. By similar arguments as before, we get that $\mathbb{E}[d(f(\mathbf{x}^2), l^1)] \geq 3/2 - 1/m$ and, therefore, there exists a left vertex on level 2 that is adjacent to l^1 (call it l^2) such that

$$\mathbb{E}[d(f(\mathbf{x}^2), l^2)] \geq \frac{2m + (m-2)(3/2 - 1/m)}{m} \geq \frac{7}{2} - \frac{4}{m}.$$

We inductively build location profiles $\mathbf{x}^3, \mathbf{x}^4, \ldots, \mathbf{x}^k$ in this fashion. We then have the following claim.

LEMMA B.1. *There exists a left vertex on level k denoted l^k such that*

$$\mathbb{E}[d(f(\mathbf{x}^k), l^k)] \geq \frac{2^{k+1}-1}{2} - \frac{2^{k+1}-(k+2)}{m}. \tag{B1}$$

PROOF. We prove Lemma B.1 by induction on the level d. For $d=0$, we have that

$$\frac{2^{0+1}-1}{2} - \frac{2^{0+1}-(0+2)}{m} = \frac{1}{2},$$

which is indeed the lower bound that we have obtained for $\mathbb{E}[d(f(\mathbf{x}^0), l^0)]$.

Let us assume that there exists a left vertex on level d denoted l^d such that

$$\mathbb{E}[d(f(\mathbf{x}^d), l^d)] \geq \frac{2^{d+1}-1}{2} - \frac{2^{d+1}-(d+2)}{m}.$$

Constructing $\mathbf{x}^{d+1}$ along the lines given earlier, we get that there is a left vertex on level $d+1$ denoted l^{d+1} such that

$$\begin{aligned}\mathbb{E}[d(f(\mathbf{x}^{d+1}), l^{d+1})] &\geq \frac{1}{m}\left[2^d \cdot m + (m-2)\left(\frac{2^{d+1}-1}{2} - \frac{2^{d+1}-(d+2)}{m}\right)\right] \\ &\geq \frac{2^{d+2}-1}{2} - \frac{2^{d+2}-(d+3)}{m}.\end{aligned}$$

From Lemma B.1, we obtain a location profile $\mathbf{x}^k$ and a left vertex on level k, l^k such that Equation (B1) holds. Hence, the expected maximum distance is at least the expression at the right-hand side of Equation (B1). On the other hand, under $\mathbf{x}^k$, the solution that locates the facility at l^{k-1} has a maximum cost of 2^{k-1}. The ratio is at least

$$\frac{\mathrm{mc}(f(\mathbf{x}^k), \mathbf{x}^k)}{\mathrm{mc}(l^{k-1}, \mathbf{x}^k)} \geq \frac{(2^{k+1}-1)/2 - (2^{k+1}-(k+2))/m}{2^{k-1}} \geq 2 - \frac{1}{2^k} - \frac{4}{m}.$$

Note that we have to choose m and k such that $n/(2m^k) \geq 1$ so that we still have at least one agent at each of m left vertices on level k in our construction. By taking $k = \Theta(\sqrt{\log n})$ and

$$m = \Theta(n^{1/k}) = \Theta(n^{1/\sqrt{\log n}}),$$

we satisfy the above constraint and get that the approximation ratio of f is as announced. □

References

[1] Alon, N., M. Feldman, A. D. Procaccia, M. Tennenholtz. 2009. Strategyproof approximation mechanisms for location on networks. http://arxiv.org/abs/0907.2049.

[2] Bandelt, H. J. 1985. Networks with Condorcet solutions. *Eur. J. Oper. Res.* **20**(3) 314–326.

[3] Bandelt, H. J., M. Labbé. 1986. How bad can a voting location be? *Social Choice Welfare* **3** 125–145.

[4] Black, D. 1958. *Theory of Committees and Elections*. Cambridge University Press, Cambridge, UK.

[5] Gibbard, A. 1973. Manipulation of voting schemes. *Econometrica* **41** 587–602.

[6] Hakimi, S. L. 1965. Optimal distribution of switching centers in a communications network and some related graph theoretic problems. *Oper. Res.* **13** 462–475.

[7] Hansen, P., J.-F. Thisse. 1981. Outcomes of voting and planning: Condorcet, Weber and Rawls locations. *J. Public Econom.* **16** 1–15.

[8] Hodgart, R. L. 1980. Optimizing access to public facilities: A review of problems, models and methods of locating central facilities. *Progress Human Geography* **2**(1) 17–48.

[9] Holzman, R. 1990. An axiomatic approach to location on networks. *Math. Oper. Res.* **15**(3) 553–563.

[10] Labbé, M. 1985. Outcomes of voting and planning in single facility location problems. *Eur. J. Oper. Res.* **20** 299–313.

[11] Lu, P., Y. Wang, Y. Zhou. 2009. Tighter bounds for facility games. *Proc. 5th Internat. Workshop Internet Network Econom.* (*WINE*), Springer, Berlin, 137–148.

[12] Lu, P., X. Sun, Y. Wang, Z. A. Zhu. 2010. Asymptotically optimal strategy-proof mechanisms for two-facility games. *Proc. 11th ACM Conf. Electronic Commerce* (*EC*), ACM Press, New York,

[13] Marsh, M. T., D. A. Schilling. 1994. Equity measurement in facility location analysis: A review and framework. *Eur. J. Oper. Res.* **74** 1–17.

[14] Moulin, H. 1980. On strategy-proofness and single-peakedness. *Public Choice* **35** 437–455.

[15] Procaccia, A. D., M. Tennenholtz. 2009. Approximate mechanism design without money. *Proc. 10th ACM Conf. Electronic Commerce* (*EC*), ACM Press, New York, 177–186.

[16] Rawls, J. 1971. *A Theory of Justice*. Harvard University Press, Cambridge, MA.

[17] Richard, D., H. Beguin, D. Peeters. 1990. The location of fire stations in a rural environment: A case study. *Environment Planning A* **22** 39–52.

[18] Satterthwaite, M. 1975. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *J. Econom. Theory* **10** 187–217.

[19] Schummer, J., R. V. Vohra. 2002. Strategy-proof location on a network. *J. Econom. Theory* **104**(2) 405–428.