

Facility Reallocation on the Line

Bart de Keijzer, Dominik Wojtczak

University of Liverpool

B.De-Keijzer@liverpool.ac.uk, D.Wojtczak@liverpool.ac.uk

Abstract

We consider a multi-stage facility reallocation problems on the real line, where a facility is being moved between stages based on the locations reported by n agents. The aim of the reallocation mechanism is to minimize the social cost, i.e., the sum over the total distance between the facility and all agents at all stages, plus the cost incurred for moving the facility. We study this problem both in the offline setting and online setting. In the offline case the mechanism has full knowledge of the agent locations in all future stages, and in the online setting the mechanism does not know these future locations and must decide the location of the facility on a stage-per-stage basis. We derive the optimal mechanism in both cases. For the online setting we show that its competitive ratio is $(n+2)/(n+1)$. As neither of these mechanisms turns out to be strategyproof, we propose another strategyproof mechanism which has a competitive ratio of $(n+3)/(n+1)$ for odd n and $(n+4)/n$ for even n , which we conjecture to be the best possible. We also consider a generalization with multiple facilities and weighted agents, for which we show that the optimum can be computed in polynomial time for a fixed number of facilities.

1 Introduction

Facility location is one of the most well-studied problems in the literature due to its multitude of practical applications, e.g., to clustering of images [Song *et al.*, 2017] and document and image summarization [Lin and Bilmes, 2012; Tschatschek *et al.*, 2014]. In its simplest form, also referred to as the Weber problem [Weber, 1909], the aim is to locate a single point from which the sum of the transportation costs to n agents' locations is minimal. The generalization of this problem to placing k facilities in a way that the sum of the distances of each agent to its nearest facility is minimized is NP-hard already in two-dimensions [Megiddo and Supowit, 1984]. However, it is polynomial time solvable in the one-dimensional setting [Megiddo *et al.*, 1983], i.e., when the agents' and facilities' locations are all placed along a single real line. Such scenarios were studied, e.g., in the context of an optimal placement of

public facilities along a street [Miyagawa, 2001] or to analyse voting scenarios [Feldman *et al.*, 2016].

We generalise this classic facility location problem to the situation where the interaction between agents and facilities lasts over multiple rounds, the agents' locations may not be known in advance and the facilities can be moved if needed. In particular, let us consider the following motivating example. Assume there is a political party with k members that would like to win the next T consecutive parliamentary elections. In order to achieve this, the party would like its members to represent the political opinions of as many voters as possible to get their votes. A voter feels well-represented if at least one party member has a similar political stance as her. As a result, a party that would like to succeed should try to gather members with a diverse range of political opinions. (In reality these cannot be too diverse, as the party would not be taken seriously. This issue will be considered in our future work.) During each term, the political opinion of the voters may change and the party may need to refocus to better reflect the current political sentiments. At the same time, each time a politician changes his opinion, he loses a bit of credibility. To estimate such a difference in opinions, Downs [1957] proposed to model the political views from extreme-left to extreme-right as points along a single real line. The ultimate aim for the party is then to minimize the sum of the distances from its voters and the credibility that is lost when readjusting the party members' stance before each election.

In an alternative formulation, one can imagine a long and narrow beach where k ice-cream vendors (owned by the same company) are to be located. For the next T hours, the beach is visited by n customers and their location may change throughout the day. As each client will typically simply pick the closest vendor, it is best for the vendors to change their location throughout the day to adjust to the demand. The aim in this case is the minimization of the social cost, i.e., the total distance that the customers as well as the ice-cream vendors have to travel.

The models we described so far assumed the agents to report their location truthfully. However, since each agent would like to be as close as possible to one of the facilities, he may have an incentive to lie. As such a behaviour is highly undesirable, one typically strives to devise a *strategyproof mechanism*, where a mechanism is simply any algorithm that outputs facility assignment based on the locations reported by the agents, for

which no agent can gain by misreporting his location. For instance, if the preferences are quasi-linear then the celebrated VCG mechanism is optimal and strategyproof and it achieves this by collecting a payment from each agent. However, in several domains such payments are impossible or undesirable, e.g., in kidney exchanges, public projects, politics and voting, due to ethical, legal or privacy issues. In such settings, the mechanism proposed needs to be strategyproof without using any monetary transfers.

Our analysis starts in Section 3 with finding an optimal mechanism in the case the true locations of all the agents are known. We show that there is a mechanism that runs in linear time for one facility ($k = 1$) and another one that is polynomial for any fixed k . The mechanism for $k = 1$ is quite intricate and we then adapt it to the online setting in Section 4. In such a setting, we need to make decisions before we see the whole input, so we cannot hope to find an optimal solution but try to minimize the *competitive ratio* instead, which is the worst-case ratio of the cost returned by the online mechanism and the optimal one. We show a mechanism of which the competitive ratio is $(n+2)/(n+1)$ and prove that no other mechanism can do better. Finally, in Section 5, we show that neither of these one facility location mechanisms is strategyproof, and devise a new strategyproof mechanism without monetary transfers. We show that its competitive ratio is $(n+4)/n$ for odd n and $(n+3)/(n+1)$ for even n , and that these values are tight.

Related work.

Our work fits tightly into the literature of time-evolving optimization problems, where an instance of a computational problem changes over time and there is a cost incurred by implementing a change in the solution at each time step. See, for example, [An et al., 2017; Gupta et al., 2014; Eisenstat et al., 2014], where the latter two works consider two other variants of time-evolving facility reallocation problems.

A mobile facility location problem, which can be seen as one stage version of our problem with k facilities, was introduced in [Friggstad and Salavatipour, 2011] and the problems was shown to be NP-hard in general. A polynomial $(3 + \epsilon)$ -approximation algorithm was shown in [Ahmadian et al., 2013].

The study of the k -facility location problem in an online setting, typically called the k -median problem in such a context, has been extensively studied (see, e.g., [Fotakis, 2011] for a survey). In particular, Divéki and Imreh [2011] studied an online model where the location of the facilities can be moved, but with a zero cost.

The field of approximate mechanism design without money was initiated by Procaccia and Tennenholtz [2009] where the facility location problem was considered. This research has attracted much attention in recent AI conferences. For example, Todo et al. [2011] study false-name strategyproof mechanisms on a real line, i.e., such mechanisms cannot be manipulated to their advantage by agents who replicate themselves. Sui et al. [2013] study strategyproof facility location in multi-dimensional space for different metrics and devises the *percentile mechanisms* for them. Zou and Li [2015] study strategyproof mechanisms for agents with dual preferences, i.e., some agents would like to be as close as possible to a

facility, while others would prefer to be as far as possible. Serafino and Ventre [2015] study the two facility problem where the cost function may differ between agents. Filos-Ratsikas et al. [2017] study strategyproof mechanisms for double-peaked preferences, e.g., each agent would like to be close to a facility but not too close. Procaccia et al. [2017] study the trade-off between variance and approximation factor for strategyproof mechanisms. Feldman et al. [2016] study the one-stage facility location problem in the context of voting under the constraint that the facilities can only be placed on agents' locations. [D. Fotakis and Tzamos, 2014] characterized completely the deterministic strategyproof mechanisms for the placement of two facilities on the line and showed that the best approximation ratio of such a mechanism is $n - 2$. Lu et al. [2010] showed there exists a 4-approximation randomized mechanism for the same problem, while a 1.045 lower bound is also known [Lu et al., 2009].

2 Preliminaries

For $a \in \mathbb{N}$, we will write $[a]$ to denote the set $\{1, \dots, a\}$. In this paper we will treat all sets as multisets, and all the operations are thus multiset operators. Due to space limitations, proofs are either replaced by proof sketches or omitted.

An instance of the *facility reallocation problem* is a quadruple (n, T, y^0, x) , where $n \in \mathbb{N}$ is the number of agents, T is the number of stages, y^0 is the starting location of facility, and $x = (x^1, \dots, x^T)$ are the vectors of agent locations in each stage, where $x^t = (x_1^t, \dots, x_n^t) \in \mathbb{R}^n$ are the locations of the agents at Stage $t \in [T]$. A *solution* of a given instance is a placement of the facility at each of the stages, i.e., a sequence $y = (y^1, \dots, y^T) \in \mathbb{R}^T$. A *mechanism* is a mapping from instances to solutions. The *cost* of a solution y is given by

$$C(y) = \sum_{t=1}^T \left(|y_j^{t-1} - y_j^t| + \sum_{i=1}^n |x_i^t - y_j^t| \right),$$

which is, in words, the sum of distances from each agent to the facility at each stage t , plus the total distance the facility moves across all stages. An *optimal solution* is a solution that minimizes C .

We define X^t as the multiset $\{x_1^t, \dots, x_n^t\}$. Let $t \in [T]$ be a stage, and let y^{t-1} be any location. We define $M_t(y^{t-1})$ as the *median* at Stage t with respect to y^{t-1} . That is, $M^t(y^{t-1})$ is the set of points z such that $\sum_{i=1}^n |x_i^t - z| + |y^{t-1} - z|$ is minimized. It is straightforward to verify that $M^t(y^{t-1})$ is the middle point of $\{y^{t-1}\} \cup X^t$ if n is even, and is the interval between (and including) the two middle points of $\{y^t\} \cup X^t$ if n is odd.

In Section 5, we study the *strategyproofness* property of our mechanisms. There, we assume that the input to the mechanism is provided by the agents, who are interested in minimizing their total distance to the facility. They may thus misreport their true locations, in case this results in facility placements closer to their true locations.

Let A be a mechanism. We define the *cost of Agent* $i \in [n]$ for a solution y as $c_i(y) = \sum_{t=1}^T |y^t - x_i^t|$. Let (\tilde{x}_S, x_{-S}) be a solution obtained from x by replacing the location vectors $\{x_i : i \in S\}$ where $x_i = (x_i^1, \dots, x_i^T)$, by different vectors

$\tilde{x}_S = \{\tilde{x}_i : i \in S\}$, where $\tilde{x}_i = (\tilde{x}_i^1, \dots, \tilde{x}_i^T)$ are the alternative locations corresponding to Agent $i \in S$. Mechanism A is *group-strategyproof* if for all $S \subseteq [n]$, for all \tilde{x}_S , there exists an $i \in S$ such that $c_i(A(x)) \leq c_i(A(\tilde{x}_S, x_{-S}))$. Mechanism A is *strategyproof* if for all i and for all \tilde{x}_i it holds that $c_i(A(x)) \leq c_i(A(\tilde{x}_i, x_{-i}))$.

3 Optimal Mechanisms

First, we consider the basic problem of computing an optimal solution to the facility reallocation problem when the complete instance is given to the mechanism in advance.

Let $I = (n, T, y^0, x)$ be a facility reallocation instance. The following lemmas show that in every Stage $t \in [T]$, putting the facility on a point in the interval $M^t(y^{t-1})$ is less expensive than putting the facility outside of $M^t(y^{t-1})$, regardless of the choice of facility locations in all the other stages.

Lemma 1. *Let $y = (y^1, \dots, y^T)$ be a solution to I and let $t \in [T]$. The distance d between y^t and the nearest point $z \in M^t(y^{t-1})$ is at least*

$$\sum_{i=1}^n |x_i^t - y^t| + |y^{t-1} - y^t| - \left(\sum_{i=1}^n |x_i^t - z| + |y^{t-1} - z| \right).$$

The following lemma is proved using the former.

Lemma 2. *Let $y = (y^1, \dots, y^T)$ be a solution to I . Suppose that there is a Stage t such that y^t is not in $M^t(y^{t-1})$. Then, replacing y^t with the nearest point \tilde{y}^t to y^t that lies in $M^t(y^{t-1})$ results in a solution with a cost that is at most $C(y)$.*

This yields an easy and efficiently computable optimal mechanism when n is even: An optimal facility reallocation mechanism for $k = 1$ always places the facility at Stage $t \in [T]$ in the median interval $M^t(y^{t-1})$. Hence, when the number of agents is even, the optimal allocation vector is unique and can be computed in $O(nT)$ (i.e., linear) time.

For n odd, the above does not yet characterize the optimal mechanism, and it turns out that in that case the facility cannot be placed at just any point in the median without sacrificing optimality. This is due to the fact that the median $M^t(y^{t-1})$ of Stage t is dependent on the location y^{t-1} of the facility of the previous stage, and is therefore by recursion also dependent on the location the facility and all the agents at all previous stages. Because $M^t(y^{t-1})$ is generally an interval of points instead of a single point, there is a choice to be made that influences the medians of all the subsequent stages.

The following two example instances show that the optimal choice of facility at a given stage may depend on the locations of the agents in the next stage.

Example 3. Consider first the following example with $T = 3$ stages and $n = 3$ agents, depicted in Figure 1. Let $y^0 = 3$ be the initial facility location. The locations of the agents at each of the 3 stages are $x^1 = (3, 7, 7)$, $x^2 = (4, 5, 6)$, $x^3 = (1, 1, 2)$. The median in the first stage is the interval $[3, 7]$. The point in this median that we choose for y^1 influences the median in the second stage: When we set $y^1 \in [3, 4]$, the median in the second stage will be $[4, 5]$; when $y^1 \in (4, 5]$, the median in the second stage will be $[y^1, 5]$; when $y^1 \in [5, 6)$ the median in the second stage will be $[5, y^1]$; and when $y^1 \in [6, 7]$, the median in the second stage will be $[5, 6]$.

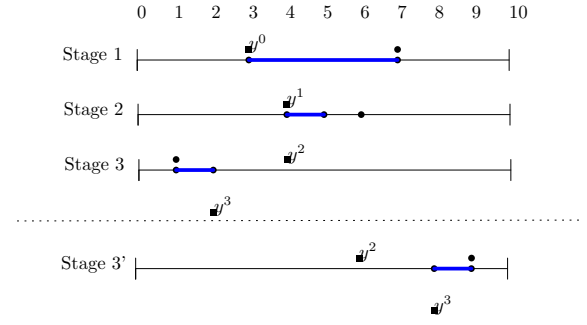


Figure 1: Depiction of the two facility reallocation instances of Example 3, one consisting of stages 1,2,3, and the other consisting of stages 1,2,3'. The dots indicate the locations of the agents at each stage. The squares indicate an optimal choice of facility locations, where the square at a given stage is the facility location at the previous stage. (At the first stage it is the starting location.) The square below the final stage is the facility location at the last stage. The blue part of the line at Stage t represents the median $M^t(y^{t-1})$.

The optimal solution is to set $y^1 \in [4, 5]$, and to not move the facility to a different location in the second stage.

However, if in Stage 3 the facilities of the three agents would be $\bar{x}^3 = (8, 9, 9)$, then the optimal choice of facility location for the first stage would be to set $y^1 \in [5, 6]$. \square

The above examples show that there may be infinitely many optimal solutions when n is odd. The analysis also suggest that it may always be optimal to put the facility at any given stage at the location of the central agent of the subsequent stage, whenever that is possible. We can prove this, and in fact we can refine this statement further, as follows.

Theorem 4. *Suppose that in instance I it holds that n is odd. There exists an optimal solution y for this instance such that at any Stage $t \in [T - 1]$, the facility is placed at the point in $M^t(y^{t-1})$ that lies closest to the location of the median agent at the next Stage $t + 1$, i.e., the median of $\{x_1^{t+1}, \dots, x_n^{t+1}\}$. At Stage T , the facility is placed anywhere in $M^T(y^{T-1})$.*

Proof sketch. We can show this through proving the following technical lemma:

Lemma 5. *Let I and \tilde{I} be two instances that differ only in the starting positions. Let y^0 and \tilde{y}^0 be these respective starting positions, let d be the distance between them, and assume that they lie on the same side of the middle agent location at Stage 1. Assume furthermore that I is the instance with the starting position that is closest to $x_{\lfloor n/2 \rfloor}^1$. Then, the optimal cost of I is less than the optimal cost of \tilde{I} , and is at most d lower.*

This is sufficient to prove the theorem, because Lemma 2 implies that the facility should always be placed in the median interval at every stage: Let p be the point in $M^1(y^0)$ at Stage 1 of instance I that is closest to the middle agent location. Consider any other point \tilde{p} in $M^1(y^0)$. Points p and \tilde{p} generate the same cost in the first stage, and Lemma 5 implies that the subinstance obtained by restricting I to Stages $2, \dots, T$ has a better optimal cost when p is the starting location rather than \tilde{p} , as p and \tilde{p} lie on the same side of the middle agent

location. Choosing point p is optimal in the first stage when in subsequent stages an optimal point is also chosen, hence by induction it is optimal to choose at any Stage $t \in [T - 1]$ the point in $M^t(y^{t-1})$ closest to the middle agent location.

The proof of Lemma 5 also proceeds by induction on T and requires a case distinction based on the location of the middle agent at Stage 2 relative to the median interval $M^1(y^0)$. \square

The following corollary summarizes all of the above.

Corollary 6. *It is an optimal facility reallocation mechanism for $k = 1$ to place the facility at each stage $t \in [T - 1]$ at the point in the median interval $M^t(y^{t-1})$ that lies closest to the median agent of Stage $t + 1$, and to place the facility at Stage T at any point in the median interval. Hence, when the number of agents is even, the optimal allocation vector is unique and can be computed by an online mechanism. When the number of agents is odd, the optimal mechanism needs to look at each stage at the agent locations in Stage t and $t + 1$ only. The mechanism runs in both cases in $O(Tn)$ time.*

Thus, for n odd, we can compute the optimum efficiently, but we do need a one stage “look-ahead”. Thus, this result does not imply an optimal online mechanism. We give in Section 4 an online mechanism with an optimal competitive ratio.

We consider next a generalized variant of the problem where there are $k \geq 1$ facilities and the agents have weights. The cost of an agent $i \in [n]$ is her distance to the nearest facility at each stage, and her weight w_i is the factor by which her cost contributes to the objective function.

The problem of computing the optimal facility locations for such a generalized instance is considerably more complex. We prove that nonetheless, when the number of facilities k is fixed, this can be done in polynomial time.

Theorem 7. *There exists a mechanism that computes the optimal solution to a generalized facility reallocation problem in time $O(T^2(2 \max\{Tn, k\})^{k+1})$.*

Proof sketch. The main insight that we need (and the most challenging part of this proof) is that it suffices to consider only solutions where at each stage $t \in [T]$ each facility is placed on a location corresponding to one of the agent locations (at any stage) or to one of the starting facility locations y_1^0, \dots, y_k^0 .

Lemma 8. *Let $y = (y^1, \dots, y^T)$ be a solution to a generalized facility location instance (where y^t are k -dimensional vectors) with T stages. Then there exists a solution \tilde{y} such that $C(\tilde{y}) \leq C(y)$ and for all $t \in [T]$ and $j \in [k]$, it holds that $y_j^t \in \{y_1^0, \dots, y_k^0\} \cup X^1 \cup \dots \cup X^T$.*

We prove this lemma by studying the derivative of the cost function of any solution where a facility is not placed at one of the appropriate points. We show that the cost must increase when moving such a facility either to the left or right.

Using this lemma, a polynomial time algorithm with the claimed runtime can be constructed through standard dynamic programming techniques: For each possible vector \bar{y} of starting facility locations (there are at most $k + (Tn)^k$ of them by the above lemma), we can efficiently find a solution to a subinstance I on stages t, \dots, T with facility starting positions \bar{y} , by considering the optimal solutions to the subinstances on

stages $t + 1, \dots, T$ with all possible starting positions, and using the one that minimizes the cost for I . \square

4 The Online Setting

In this section we study again the basic facility reallocation problem, and we focus on the online variant of the problem, where for each stage $t \in [T]$ the agent locations $x_1^{t+1}, \dots, x_n^{t+1}$ of the next stage may only be read by the mechanism after the mechanism outputs the facility locations y_1^t, \dots, y_k^t for the current stage.

Example 3 implies that for an odd number of agents, the optimal mechanism necessarily needs to look one stage ahead. The following example shows that due to the lack of ability to look one stage ahead, no online mechanism has a competitive ratio better than $(n + 2)/(n + 1)$ when n is odd.

Example 9. Consider the following two instances $I_\ell = I'_\ell$, each with $2\ell + 1$ agents, for any $\ell \in \mathbb{N}$. Both instances have $T = 2$ stages. The agent locations of Stage 1 are 0 for the first ℓ agents and 1 for the remaining $\ell + 1$ agents. At Stage 2, all agents are located at 0 in instance I_ℓ , and at 1 in instance I'_ℓ . The initial facility location is 0.

The median $M^1(y^0)$ of the first stage is $[0, 1]$, so by Corollary 6 the optimal solution is to place the facility at 0 in Instance I_ℓ and at 1 in Instance I'_ℓ . In Stage 2, the facility then does not need to move.

However, as the instances differ only in the second stage, an online mechanism is restricted to place the facility at the same position in Stage 1, in both instances. Placing the facility at $1/2$ is the best that any online mechanism can choose, to minimize the maximum cost among those two instances. Therefore the cost of the optimal solution is $\ell + 1$ for both instances, while the cost of the solution generated by the optimal online mechanism is $\ell + 3/2$. The ratio of these two quantities is $(n + 2)/(n + 1)$, which is a lower bound on the competitive ratio achievable by an online mechanism. \square

We now provide an online mechanism of which the competitive ratio matches the lower bound on the competitive ratio of Example 9. The key idea behind this online mechanism is to try to place the facility at each stage as close as possible to the location where the optimum facility may be placed. While it is impossible to know the exact location of the optimum facility at the current stage t , an optimal mechanism can nonetheless derive at each stage the precise interval in which the optimum location may lie: The online mechanism can compute the precise optimum location \tilde{y}^{t-1} at Stage $t - 1$ (as it has access to the agent locations of Stage t), and by Corollary 6 the optimal location at Stage 2 can lie at any point in $M^t(\tilde{y}^{t-1})$, depending on the next stage. Our online mechanism will therefore place the facility y^t at the point $M^t(y^{t-1})$ that lies as close as possible to the middle point of $M^t(\tilde{y}^{t-1})$. Denote this mechanism as Mechanism A .

Theorem 10. *Mechanism A runs in time $O(Tn)$ and has competitive ratio $(n + 2)/(n + 1)$ on instances with an odd number of n agents.*

Proof sketch. The bound on the runtime is obvious. We sketch our proof for the competitive ratio. Let $I = (n, T, y^0, x)$ be an instance where n is odd, let y be the output solution of A and

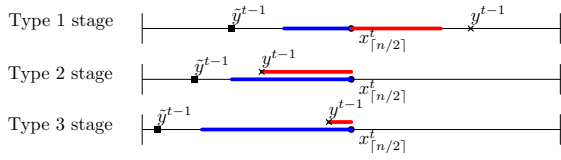


Figure 2: Depiction of stages belonging to each of the three types. The dot represents the middle agent location $x_{\lceil n/2 \rceil}^t$, the square represents the optimal facility location \tilde{y}^t , and the cross represents the facility location y^{t-1} output by the online mechanism. The blue interval represents the median $M^t(\tilde{y}^{t-1})$ associated to the optimal solution, while the red interval represents the median $M^t(y^{t-1})$ associated to the solution output by online mechanism A . Agents' locations other than the middle agent location are not depicted. Note that in a type 2 stage, it may either occur that y^{t-1} is to the left of \tilde{y}^{t-1} or to the right of \tilde{y}^{t-1} , although only the latter situation is displayed here.

let \tilde{y} be the optimal solution. We assume w.l.o.g. that agent locations are ordered non-decreasingly at each stage so that $x_{\lceil n/2 \rceil}^t$ is the location of the middle agent at stage t .

Note first that by Lemma 2, for each stage $t \in [T]$, it holds that $\tilde{y}^t \in M^t(\tilde{y}^{t-1})$ and by definition of A it also holds that $y \in M^t(y^{t-1})$. We define the *non-median cost* $C_{NM}^t(z)$ of solution z at stage t as the distance between all agents and the facility at stage t , except for the middle agent. We define the *residual cost* $C_R^t(z)$ of z at stage t of z as the distance between the middle agent and the facility and the movement of the facility. That is, $C_R^t(z) = |x_{\lceil n/2 \rceil}^t - z^t| + |z^{t-1} - z^t|$

It can be seen that $C_{NM}^t(z)$ is minimized and constant when z^t is in the interval $S^t = [x_{\lceil n/2 \rceil - 1}^t, x_{\lceil n/2 \rceil + 1}^t]$, which we refer to as the *supermedian* at Stage t . We denote its length by ℓ^t . The interval $M^t(z^{t-1})$ is always a subset of the supermedian at Stage t , regardless of its argument z^{t-1} . Hence, solutions y and \tilde{y} achieve the same non-median cost at every stage. Also, it follows that the residual costs for both solutions at any Stage $t \in [T]$ can be written as $C_R^t(y) = |x_{\lceil n/2 \rceil}^t - y^{t-1}|$ and $C_R^t(\tilde{y}) = |x_{\lceil n/2 \rceil}^t - \tilde{y}^{t-1}|$. Thus, we may derive that

$$C(y) - C(\tilde{y}) = \sum_{t=1}^T (|x_{\lceil n/2 \rceil}^t - y^{t-1}| - |x_{\lceil n/2 \rceil}^t - \tilde{y}^{t-1}|). \quad (1)$$

Therefore, we focus on bounding the right hand side.

Our proof roughly works as follows. We classify for each stage the behaviour of the mechanism into one of three *types*. A *type 1 stage* is a stage $t \in [T]$ such that y^{t-1} differs from \tilde{y}^{t-1} and lie on opposing sides of $x_{\lceil n/2 \rceil}^t$. Stage t is a *type 2 stage* if y^{t-1} and \tilde{y}^{t-1} lie on the same side of $x_{\lceil n/2 \rceil}^t$, and the middle point of $M^t(\tilde{y}^{t-1})$ is in $M^t(y^{t-1})$. Lastly, t is a *type 3 stage* if y^{t-1} and \tilde{y}^{t-1} lie on the same side of $x_{\lceil n/2 \rceil}^t$, and the middle point of $M^t(\tilde{y}^{t-1})$ is not in $M^t(y^{t-1})$ (implying that \tilde{y}^{t-1} is further away from $x_{\lceil n/2 \rceil}^t$ than y^{t-1}). Note that each stage is classified into exactly one of the three types. See Figure 2 for a visualization of the three types.

For each stage type we provide in separate propositions a meaningful bound. For stages t of type 2 we bound the distance between the facilities in distances in terms of the

length of the supermedian of Stage t and the optimal residual cost at Stage t .

Proposition 11. *For a type 2 stage $t \in [T]$ it holds that $|y^t - \tilde{y}^t| \leq \frac{n-1}{2(n+1)} \ell^t + \frac{1}{n+1} C_R^t(\tilde{y})$.*

For a type 3 stage t , Mechanism A yields a solution y with a better cost at Stage t than the globally optimal solution \tilde{y} . We prove that the distance $|y^t - \tilde{y}^t|$ between the facilities is at most equal to this profit.

Proposition 12. *For a type 3 stage $t \in [T]$, it holds that $|y^t - \tilde{y}^t| \leq C_R^t(\tilde{y}) - C_R^t(y)$.*

For type 1, we define a *block* of stages $B \subseteq T$ as a maximal set of subsequent stages t, \dots, u such that Stages t to $u-1$ are all type 1 stages. We prove that that for such a block $B = \{t, \dots, u\}$ the total residual cost difference of block B is bounded by the distance between the facility locations $|y^{t-1} - \tilde{y}^{t-1}|$ in the previous stage $t-1$.

Proposition 13. *Let $B = \{t, \dots, u\} \subseteq [T]$ be a block. Then, $\sum_{s=t}^u \max\{0, C_R^s(y) - C_R^s(\tilde{y})\} \leq |y^{t-1} - \tilde{y}^{t-1}|$.*

Let $\{B_1, \dots, B_K\}$ be the unique partition of $[T]$ into blocks. We refer to t_k as the final stage of block $k \in [K]$. Let T_2 be the subset of stages $\{t_1, \dots, t_{K-1}\}$ that are type 2 stages (which are all type t stages in $[T]$), and let T_3 be the subset of $\{t_1, \dots, t_{K-1}\}$ that are type 3 stages. Propositions 12 combined with Proposition 13 states that, for a stage $t \in T_3$, the profit in residual cost at stage t exceeds the deficiency in residual cost at the block that follows it. Propositions 11 and 13 imply that for a stage $t \in T_2$ the deficiency in the residual cost of the subsequent block is at most $\frac{n-1}{2(n+1)} \ell^t + \frac{1}{n+1} C_R^t(\tilde{y})$. The claim follows by finally proving that the first of those two terms is at most $\frac{1}{n+1} C_{NM}^t(\tilde{y})$ and taking the sum over all stages in T_2 to derive that $C(y) - C(\tilde{y}) \leq \frac{1}{n+1} C(\tilde{y})$. \square

5 Strategyproofness

We investigate in this section the strategyproofness property of our mechanisms proposed in the previous sections. The results of Moulin [1980] yield a characterization of the class \mathcal{C} of strategyproof and group-strategyproof mechanisms for the classic (single-stage) facility allocation problem: They are those mechanisms that always place the facility at the median of the union of the set of agent locations and a set of points that are independent of the agent locations.

Unfortunately, the following examples show that the optimal mechanisms of Corollary 6 and Theorem 10 are not group-strategyproof, despite the fact that they can be seen as repeated applications of mechanisms in \mathcal{C} . This can be attributed to the interdependence of the facility locations among the stages.

Example 14. For even n , consider the following instance I with $T = 3$ stages and $n = 2$ agents. The starting location of the facility is $y^0 = 0$. The agent locations are $x^1 = (0, 1)$ and $x^2 = x^3 = (1, 0)$. If Agent 1 does not misreport, his total cost is 2 under the optimal solution, because the facility will not relocate at all. If Agent 1 reports instead that she is at location 1 in Stage 1, then the facility will be placed at location 1

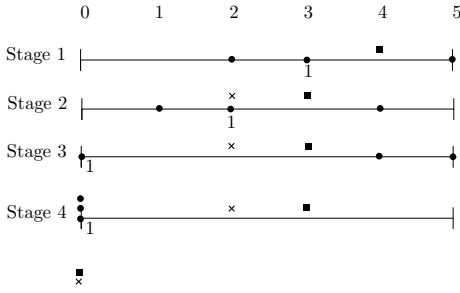


Figure 3: Depiction of the 3 agent instance of Example 14. The notation is identical to that of previous figures. Additionally, the locations of Agent 1 are labeled with her identity. The squares depict the optimal facility locations under truthful reporting. Crosses depict the optimal facility locations when Agent 1 misreports x_1^1 as 2.

in Stage 1, and will remain there for the remaining stages, reducing the cost of Agent 1 by 1.

For odd n , the example is slightly more complex, and is depicted in Figure 3: Let $T = 4$ and $n = 3$. The starting location is $y^0 = 4$. The agent locations are $x^1 = (1, 2, 5)$, $x^2 = (2, 1, 4)$, $x^3 = (0, 4, 5)$, $x^4 = (0, 0, 0)$. If Agent 1 does not misreport, then the optimal mechanism of Corollary 6 outputs the solution $(y^1, y^2, y^3, y^4) = (3, 3, 3, 0)$. This gives agent 1 a cost 4. If Agent 1 instead reports in the first stage that her location is 2, the vector of facility locations becomes $(2, 2, 2, 0)$, so the cost of Agent 1 reduces by 1. Moreover, instance I' demonstrates that Mechanism A of Theorem 10 is not strategyproof: when no agent misreports, the output solution is $(3.5, 2.5, 3.5, 0)$, which yields a total cost of 4.5 for Agent 1. If instead, Agent 1 misreports her location in the first stage as 2, the output solution becomes $(3, 2, 3, 0)$ which yields Agent 1 a cost of 3. \square

This establishes that there is a gap between the cost generated by the optimal mechanism and the cost generated by the optimal strategyproof mechanism. The following simple mechanism bounds this gap. It is an online mechanism performing slightly worse than the optimal online mechanism of Theorem 10, though its competitive ratio still tends to 1 as the number of agents grows.

Theorem 15. *The online mechanism that puts the facility in every stage at the location of the middle agent (breaking ties arbitrarily in case of even n , in a way that is independent of the reported agent locations) is group-strategyproof and has a competitive ratio of $(n+4)/n$ for even n , and $(n+3)/(n+1)$ for odd n .*

Proof sketch. Group-strategyproofness is straightforward to show, and follows in essence from the fact that the mechanism is a repetition of a mechanism from Moulin’s class \mathcal{C} , which reduces the induced game to a sum of T independent games where truth-telling is a dominant strategy in each of them.

For the bound on the competitive ratio, again we assume that the agent locations are ordered non-decreasingly at each stage. We also assume w.l.o.g. that ties are broken in favor of the left agent, when n is even, so that $y^t = x_{\lceil n/2 \rceil}^t$ by this assumption. For odd n , we define the supermedian S^t as in

the proof sketch of Theorem 10. For even n , we define the supermedian S^t as the interval $[x_{n/2}^t, x_{n/2+1}^t]$. We denote by ℓ^t the length of S^t . Note that $M^t(\tilde{y}^{t-1})$ is always contained in S^t . Thus, the optimal movement of the facility $|\tilde{y}^{t-1} - \tilde{y}^t|$ at stage t is at least $d(S^t, S^{t-1})$, i.e., the distance between S^t and S^{t-1} (where $S^0 = \{y^0\}$). On the other hand, the movement $|y^t - y^{t-1}|$ generated by our mechanism at stage t is at most $\ell^t + \ell^{t-1} + d(S^t, S^{t-1})$, where ℓ^t is the length of S^t .

Assume first that n is even. Then, the total distance between y^t and the agent locations is the minimum possible at each stage, which also holds for \tilde{y}^t . Thus, the difference $C(y) - C(\tilde{y})$ in cost is entirely attributed to the difference in total facility movement, which is at most $2 \sum_{t=1}^T \ell^t$. At stage t the distance between any two agents on opposite sides of the facility is at least ℓ^t , so that the total cost generated by the agents at stage t is at least $(n/2)\ell^t$. The latter implies that $(C(y) - C(\tilde{y}))/C(\tilde{y}) \leq \frac{4}{n}$, which yields the desired upper bound for even n .

Lastly, suppose that n is odd. The total inefficiency in facility movement is at most $2 \sum_{t=1}^T |\tilde{y}^t - x_{\lceil n/2 \rceil}^t|$. For odd n , the optimum does not always minimize the total distance between the facility and the agents at every stage in this case, though the facility is always placed in S^t , so the total distance from the agents to the facility at each stage t is $|\tilde{y}^t - x_{\lceil n/2 \rceil}^t|$ lower under y^t . We subtract this from our bound on the distance in facility movement, and we obtain that $C(y) - C(\tilde{y}) \leq \sum_{t=1}^T |\tilde{y}^t - x_{\lceil n/2 \rceil}^t|$. The quantity $|\tilde{y}^t - x_{\lceil n/2 \rceil}^t|$ is at most the length ℓ^t of the supermedian, which allows us to bound $C(y) - C(\tilde{y})$ as a convex combination between $|\tilde{y}^t - x_{\lceil n/2 \rceil}^t|$ and ℓ^t , so that

$$C(\tilde{y}) - C(y) \leq \frac{1}{\lfloor n/2 \rfloor + 1} \sum_{t=1}^T \left(|\tilde{y}^t - x_{\lceil n/2 \rceil}^t| + \lfloor n/2 \rfloor \ell^t \right).$$

The upper bound on the competitive ratio then follows because $\lfloor n/2 \rfloor = (n-1)/2$ and the latter summation can be shown to be a lower bound on $C(y)$. \square

Example 16. The following family of examples shows that the analysis of the competitive ratio in Theorem 15 is tight for all n . Let the starting facility location be $y^0 = 1$ and let there be two stages. In Stage 1, agents 1 to $\lfloor n/2 \rfloor$ are located at 1, and the remaining agents are located at 0. In Stage 2, all of the agents are located at 1. The optimal Mechanism (see Corollary 6) places the facility at location 1 in both stages, resulting in a cost of $n/2$ if n is even, and a cost of $(n+1)/2$ if n is odd.

The mechanism of Theorem 15 places the facility at location 0 in the first stage, and at location 1 in the second stage. This yields a total cost of $n/2 + 2$ if n is even, and a total cost of $(n-1)/2 + 2$ when n is odd. Thus, when n is even, the competitive ratio on these instances is $((n/2) + 2)/(n/2) = (n+4)/n$, and when n is odd, the competitive ratio is $((n+3)/2)/(n+1)/2 = (n+3)/(n+1)$. \square

6 Discussion

We studied a multi-stage variant of the classical facility location problem, where we characterized the optimal mechanisms

both in the offline setting, and in the online setting, and considered this problem under the constraint of strategyproofness as well. These mechanisms turn out to be elegant and simple in their definition, but are surprisingly challenging to analyze.

Interesting future directions are to design online and strategyproof mechanisms for the generalized variant of the problem that we briefly considered, and to characterize the class of (group)-strategyproof mechanisms for the basic version of the problem. We conjecture that the competitive ratio of Theorem 15 is the best achievable among the strategyproof mechanisms. Additionally, randomized mechanisms can be studied in this context as it is known that they outperform deterministic ones in the single stage case [Lu *et al.*, 2010].

An alternative generalization of the problem that would be interesting (and undoubtedly more complex) to study is to increase the dimension of the Euclidian space in which the locations lie, e.g. to consider facility reallocation on the plane instead of the line.

Acknowledgments

The first authors was partially supported by NWO grant 612.001.352, EPSRC grant EP/P020909/1, and University of Liverpool Visiting Fellowship. The second author was partially supported by EPSRC grants EP/M027287/1 and EP/P020909/1. We would like to thank Orestis Telelis and Guido Schäfer for helpful discussion that led to this paper.

References

- [Ahmadian *et al.*, 2013] S. Ahmadian, Z. Friggstad, and C. Swamy. Local-search based approximation algorithms for mobile facility location problems. In *Proceedings of SODA'13*, pages 1607–1621. ACM-SIAM, 2013.
- [An *et al.*, 2017] H.-C. An, A. Norouzi-Fard, and O. Svensson. Dynamic facility location via exponential clocks. *ACM Trans. Algorithms*, 13(2):21:1–21:20, 2017.
- [D. Fotakis and Tzamos, 2014] Dimitris D. Fotakis and C. Tzamos. On the power of deterministic mechanisms for facility location games. *ACM Transactions on Economics and Computation*, 2(4):15, 2014.
- [Divéki and Imreh, 2011] G. Divéki and C. Imreh. Online facility location with facility movements. *Central European Journal of Operations Research*, 19(2):191–200, 2011.
- [Downs, 1957] A. Downs. An economic theory of political action in a democracy. *Journal of Political Economy*, 65(2):135–150, 1957.
- [Eisenstat *et al.*, 2014] D. Eisenstat, C. Mathieu, and N. Schabanel. Facility location in evolving metrics. In *Proceedings of ICALP'14*, pages 459–470. Springer, 2014.
- [Feldman *et al.*, 2016] M. Feldman, A. Fiat, and I. Golomb. On voting and facility location. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC'16)*, pages 269–286. ACM, 2016.
- [Filos-Ratsikas *et al.*, 2017] A. Filos-Ratsikas, M. Li, J. Zhang, and Q. Zhang. Facility location with double-peaked preferences. *Proceedings of AAMAS'17*, 31(6):1209–1235, 2017.
- [Fotakis, 2011] D. Fotakis. Online and incremental algorithms for facility location. *ACM SIGACT News*, 42(1):97–131, 2011.
- [Friggstad and Salavatipour, 2011] Z. Friggstad and M. R. Salavatipour. Minimizing movement in mobile facility location problems. *ACM Transactions on Algorithms (TALG)*, 7(3):28, 2011.
- [Gupta *et al.*, 2014] A. Gupta, K. Talwar, and U. Wieder. Changing bases: Multistage optimization for matroids and matchings. In *Proceedings of ICALP'14*, pages 563–575. Springer, 2014.
- [Lin and Bilmes, 2012] H. Lin and J. Bilmes. Learning mixtures of submodular shells with application to document summarization. In *Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence (UAI'12)*, pages 479–490. AUAI Press, 2012.
- [Lu *et al.*, 2009] P. Lu, Y. Wang, and Y. Zhou. Tighter bounds for facility games. In *International Workshop on Internet and Network Economics (WINE'09)*, pages 137–148. Springer, 2009.
- [Lu *et al.*, 2010] P. Lu, X. Sun, Y. Wang, and Z. A. Zhu. Asymptotically optimal strategy-proof mechanisms for two-facility games. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC'10)*, pages 315–324. ACM, 2010.
- [Megiddo and Supowit, 1984] N. Megiddo and K. J. Supowit. On the complexity of some common geometric location problems. *SIAM Journal on Computing*, 13(1):182–196, 1984.
- [Megiddo *et al.*, 1983] N. Megiddo, E. Zemel, and S. L. Hakimi. The maximum coverage location problem. *SIAM Journal on Algebraic Discrete Methods*, 4(2):253–261, 1983.
- [Miyagawa, 2001] E. Miyagawa. Locating libraries on a street. *Social Choice and Welfare*, 18(3):527–541, 2001.
- [Moulin, 1980] H. Moulin. On strategy-proofness and single peakedness. *Public Choice*, 35(4):437–455, 1980.
- [Procaccia and Tennenholtz, 2009] A. D. Procaccia and M. Tennenholtz. Approximate mechanism design without money. In *Proceedings of the 10th ACM Conference on Electronic Commerce (EC'09)*, pages 177–186. ACM, 2009.
- [Procaccia *et al.*, 2017] A. D. Procaccia, D. Wajc, and H. Zhang. Approximation-variance tradeoffs in facility location games. In *Proceedings of AAAI'17 (to appear)*. AAAI, 2017.
- [Serafino and Ventre, 2015] P. Serafino and C. Ventre. Truthful mechanisms without money for non-utilitarian heterogeneous facility location. In *Proceedings of AAAI'15*, pages 1029–1035, 2015.
- [Song *et al.*, 2017] H. O. Song, Stefanie Jegelka, R. Vivek, and M. Kevin. Deep metric learning via facility location. In *Proceedings of Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [Sui *et al.*, 2013] X. Sui, C. Boutilier, and T. Sandholm. Analysis and optimization of multi-dimensional percentile mechanisms. In *Proceedings of IJCAI'13*, pages 367–374, 2013.
- [Todo *et al.*, 2011] T. Todo, A. Iwasaki, and M. Yokoo. False-name-proof mechanism design without money. In *Proceedings of AAMAS'11*, pages 651–658, 2011.
- [Tschitschek *et al.*, 2014] S. Tschitschek, R. K. Iyer, H. Wei, and J. A. Bilmes. Learning mixtures of submodular functions for image collection summarization. In *Advances in Neural Information Processing Systems*, pages 1413–1421, 2014.
- [Weber, 1909] A. Weber. Über den standort der industrien, 1. teil: Reine theorie des standortes. (on the location of industries), 1909.
- [Zou and Li, 2015] S. Zou and M. Li. Facility location games with dual preference. In *Proceedings of AAMAS'15*, pages 615–623. International Foundation for Autonomous Agents and Multiagent Systems, 2015.