Randomized Strategyproof Mechanisms for Multi-stage Facility Location Problem with Capacity Constraints

Chi Kit Ken Fong[1] ✉, Xingchen Sha[2], Hau Chan[3], Vincent Chau[4], and Wai-Lun Lo[5]

[1] School of Data Science, Division of Artificial Intelligence, Lingnan University, Hong Kong, China
kenfong@ln.edu.hk
[2] City University of Hong Kong, Hong Kong, China
xingchsha2-c@my.cityu.edu.hk
[3] University of Nebraska-Lincoln, Lincoln, USA
hchan3@unl.edu
[4] Southeast University, Nanjing, China
vincentchau@seu.edu.cn
[5] Hong Kong Chu Hai College, Hong Kong, China
wllo@chuhai.edu.hk

Abstract. We consider the multi-stage facility location problem with capacity constraints. In the problem, we seek to locate at most one capacity constrained facility in each stage to serve a subset of agents, who arrive over different stages and are located on a line. Our goal is to design randomized strategyproof mechanisms to elicit agents' true information and locate facilities that minimize the social cost and maximum cost, which are defined to be the sum and the maximum of the agents' costs, respectively. Because of the stages, an agent's cost depends on the agent's distance to their assigned facility and the agent's waiting cost. For different facility capacity settings with waiting cost, we provide randomized strategyproof mechanisms for the considered cost objectives. We also establish lower bounds for the approximation ratios given by any randomized strategyproof mechanisms.

Keywords: Facility location · Mechanism design · Algorithmic game theory.

1 Introduction

In recent years, facility location problems have received significant attention in mechanism design and social choice communities because of their ability to model real-world preference aggregation settings (e.g., voting [3, 4, 14] and site locations [6, 15]). In the standard mechanism design study of facility location, a social planner aims to locate facilities to serve strategic agents while eliciting agent preferences and optimizing the social or maximum cost [6].

Most existing mechanism design studies on facility location problem have focused on facilities without any capacity constraints or immobile over time. However, in real-world settings, facilities (e.g., mobile health clinics, mobile blood donation centers, mobile outreach programs) are both capacity constrained [1, 2, 18] and mobile (i.e., need to be relocated over several periods or stages [8, 17, 19]). In this paper, we consider the multi-stage facility location problem with capacity constraints (MSFLPs-CC) by merging the facility location problem with capacity constraints and the multi-stage facility location problem (with mobile facilities). The considered facility location problem can be used to model a wide range of multistage settings, including locating shuttle stops, providing mobile health services, scheduling education sessions, and other multi-stage extensions of scenarios (see e.g., [1, 2, 7, 9, 16, 18]).

Notice that both multi-stage models with [8] or without moving cost [17, 19] have been studied in the facility location literature. We focus on the latter settings because the (mobile) facility's moving cost are negligible/independent across each period and the social planner focuses on minimizing the agents' costs rather than their own costs.

For instance, consider an organization that is providing shuttle services from a central station/point (e.g., hotels/workplaces) at the beginning of each stage, where each shuttle will travel to its designated location and take their agents (e.g., clients or employees), who have location preferences and arrive over different time intervals, to a relatively distant destination. Because a shuttle often has a limited seating capacity to serve new agents arriving over different time intervals, the company needs to determine the pick-up points of the shuttle for each time period and how to assign agents to the shuttle at different time intervals to better serve the agents, accounting for agent travel distances to the shuttle locations and, possibly, agent waiting times over different time intervals. Moreover, the moving cost of the facilities between stages is negligible compared to the pre-determined round-trip or operation costs (which are independent and incomparable to the agents' costs). As such, moving cost is not considered and incorporated into the planner's objectives.

In addition, the considered multi-stage problems play an important role in assisting the mobile health services [7, 16]. At the beginning of each day, the medical facility (e.g., vaccine vehicles, mobile blood donation centers) travels from a central station/point (e.g., medical clinic) to serve residents of several areas. Due to the storage capacity and vaccination time restrictions, the medical facility can only serve a limited number of residents each day. Therefore, the government needs to decide where to locate the facility and which residents to serve over different periods.

1.1 Our Contributions

We study the multi-stage facility location problem with k capacity constrained facilities on the real line $\mathbb{I} = [0, 1]$, where k is the maximum number of available facilities, and the real line is the most widely studied setting [6] that abstractly represents the facility serving range (e.g., the location range of the mobile vaccine vehicle on a street, and the difficulty scale of the standardized materials).

In this paper, we characterize our mechanism design results (see Tables 1) for minimizing the two cost objectives by the number of agents (i.e., n), the number of facilities (i.e., k), facility capacities (i.e., $c_1, \ldots, c_k$), the last stage with arriving agents (i.e., T), and the penalty coefficient (i.e., d).

We consider two different facility capacity settings with waiting cost, i.e., $d > 0$ (see Table 1). For the equal capacity setting (e.g., when the facility capacity cannot change over time, such as shuttle buses) where $c_1 = \cdots = c_k = c$ and $k \cdot c = n$,

Table 1. Summary of our results. Notice that it reduces to the classic single facility location problems when $k = 1$.

Objective	Upper Bound	Lower Bound
Equal Capacity Setting:		
Social Cost	$\frac{n}{2d} + 1$,$T(n-c) + 1$	$\frac{2}{4d+3} + 1$
Max Cost	$\frac{1}{d} + 1$, $max(T + k - 2, 2)$	$\frac{1}{4d+2} + 1$
Arbitrary Capacity Setting:		
Social Cost	$\frac{n}{2d} + 1$, $Tn - T + 1$	$\frac{2}{4d+3} + 1$
Max Cost	$\frac{1}{d} + 1$, $max(T + k - 2, 2)$	$\frac{1}{4d+2} + 1$

we first present a randomized strategyproof mechanism, which has a good performance when d is relatively large. It achieves approximation ratios of $\frac{n}{2d} + 1$ for social cost and $\frac{1}{d} + 1$ for maximum cost. We also provide another strategyproof mechanism which performs well when d is small, and notice that this mechanism can also work when there is spare capacity, i.e., $kc > n$. It achieves an approximation ratio of $T(n - c) + 1$ for social cost and $max(T + k - 2, 2)$ for maximum cost. We complement the result by giving lower bounds of $\frac{2}{4d+3} + 1$ for social cost and $\frac{1}{4d+2} + 1$ for maximum cost by any strategyproof mechanism.

For the arbitrary capacity setting (e.g., when the facility capacity can change over time such as a classroom in the education setting), we first provide a strategyproof mechanism with dynamic programming to optimize the waiting cost of agents, which performs well when the waiting cost is large. It achieves an approximation ratio of $\frac{n}{2d}+1$ for social cost and $\frac{1}{d} + 1$ for maximum cost. Then we provide another strategyproof mechanism with dynamic programming to optimize the distance cost of agents, which has a good performance when d is small. It has approximation ratios of $Tn - T + 1$ for social cost and max$\{T + k - 2, 2\}$ for maximum cost.

The remainder of the paper is organized as follows. We first formally define the problem in Section 2. Then, we study different capacity settings with waiting cost in Section 3. Finally, we conclude our work and discuss the open questions in Section 4.

1.2 Related Work

We focus on studies on mechanism design for facility location problem that are most related to ours. We note that there are optimization studies that consider facility location problem where facilities have capacity constraints (see, e.g., [5, 16]) or are mobile (see, e.g., [9, 16]).

In mechanism design for facility location problem, Moulin [14] first characterized strategyproof mechanisms for the classical single stage facility location problem on a line, where agents have single-peaked preferences. The work of Procaccia and Tennenholtz [15] initiated the study of approximate mechanism design without money and used the facility location problem as a case study. They obtained several approximately optimal (deterministic and randomized) strategyproof mechanisms for the single facility location problem under the social and maximum cost objectives. They also considered two homogeneous facilities or multiple locations per agent. Later numerous studies improved the bounds and complemented the results with k facilities [10–13].

Since then, different variants of the classical facility location problem have been proposed. Aziz et al. [2] first introduced the capacity constrained facility location problem from a mechanism design perspective, where the number of agents is larger than the total capacity of facilities. Later, Aziz et al. [1] provided negative results on several classical mechanisms for the uncapacitated settings in terms of the strategyproofness and the approximation ratio. They also provided a deterministic strategyproof mechanism, the INNERPOINT mechanism for two facilities with equal capacity constraint, and a deterministic strategyproof mechanism for two facilities with arbitrary capacity constraints. They also proved that the corresponding optimization problem with arbitrary capacity constraints is NP-hard. Besides the algorithmic and the mechanism perspective, Walsh [18] showed a strong characterization theorem that the INNERPOINT mechanism is the unique strategyproof mechanism that is both anonymous and Pareto optimal for two facilities location problem with equal capacity constraints.

There are also studies on the multi-stage facility location problem without any capacity constraints where agents arrive dynamically over different stages. The work of De Keijzer and Wojtczak [8] investigated the multi-stage facility reallocation problem on the real line, where the goal is to minimize the sum of distance costs between the facility and agents at all stages, plus the facility's moving cost. The work of Wada et al. [17] studied the dynamic facility location problem from the mechanism design perspective, where agents can decide their participations in each stage. Wang et al. [19] studied the multi-stage facility location problem with transient agents who arrive in arbitrary stage and stay for a number of consecutive stages. They analyze the problems from both algorithmic and mechanism design perspectives. To the best of our knowledge, we are the first to simultaneously consider the multi-stage settings, facility capacity constraints, and waiting time of each agent in facility location problem. For more details, please refer to a recent survey on mechanism design for facility location problem [6].

2 Preliminaries

In this section, we formally define the multi-stage facility location problem with capacity constraints (MSFLPs-CC) and the considered mechanism design problems.

Multi-Stage Facility location problem with Capacity Constraints. We are given a collection of agents $N := \{1, \ldots, n\}$, where agents arrive in different stages and $T \geq 1$ is the last stage with arriving agents. Each agent $j \in N$ has location $x_j \in \mathbb{I} = [0, 1]$. We denote the location profile of all agents as $X = (x_1, \ldots, x_n)$. We assume agents are ordered such that $x_1 \leq x_2 \leq \cdots \leq x_n$.

In our setting, we have $k \geq 1$ (mobile) facilities to serve these agents. At each stage, at most one facility can be placed due to resource constraints (such a setting with a single facility at each stage is also motivated and considered by [8, 9, 16, 19]).

In order to serve all agents, it is also possible to place the facility at stages beyond T. Therefore, we consider at most $T + k - 1$ stages when locating the facilities. When $k = 1$, the only way to serve all of the agents is to wait until everyone has arrived and serve them all at stage T, which is equivalent to the standard facility location problem with a single facility. Because $k = 1$ has been well-studied (see e.g., [6, 15]) for social and maximum costs, we focus on the cases of $k \geq 2$.

Each facility is indexed by $i \in [k] = \{1, \ldots, k\}$ with a capacity constraint c_i restricting the number of agents it can serve, and its location is denoted as $f_i \in \mathbb{I} = [0, 1]$. The k facilities' capacities are denoted by $C = (c_1, \ldots, c_k)$.

The facility that serves agent $j \in N$ is indicated by $a_j \in [k]$. We denote $N_i \coloneqq \{j | a_j = i\}$ as the group of agents that facility *i* serves in the same stage. For simplicity, denote $L_i \coloneqq \min_{j \in N_i} x_j$ and $R_i \coloneqq \max_{j \in N_i} x_j$ as the locations of the leftmost agent and the rightmost agent in group i, respectively. The arrival stage of agent j is denoted as $r_j \leq T$. We denote the arrival stage profile of the agents as $R = (r_1, \ldots, r_n)$. The serving stage of facility $i \in [k]$ (i.e., the stage in which the agents in N_i are served) is denoted as $s_i \leq T + k - 1$. Clearly, for N_i to be valid and feasible, $|N_i| \leq c_i$ and $r_j \leq s_i$ for any $j \in N_i$.

Mechanism Design Problems. We are interested in designing randomized strategyproof mechanisms that minimize the social or maximum cost.

A randomized mechanism is a function F that maps the profile (X, R, C) to $Y \coloneqq (f_{P_1}, \ldots, f_{P_k})$, $O \coloneqq (N_{P_1}, \ldots, N_{P_k})$, and $S \coloneqq (s_{P_1}, \ldots, s_{P_k})$, where each f_{P_i} is a set of probability distributions over $\mathbb{I}$, N_{P_i} is a set of probability distributions over N, and S_{P_i} is a set of probability distributions over $\{1, \ldots, T + k - 1\}$ such that for each pair of values (N_i, s_i) from the distribution (N_{P_i}, s_{P_i}), we have $r_j \leq s_i$ for any $j \in N_i$ and $|N_i| \leq c_i$. Given the penalty coefficient d, the cost of an agent is defined to be $cost(F(X, R, C), x_j, r_j) = E_{f_i \sim f_{P_i}, N_i \sim N_{P_i}, s_i \sim s_{P_i}}[|y_{a_j} - x_j| + d \cdot (s_{a_j} - r_j)]$, which is the expected sum of the distance cost and the waiting cost. We focus on the setting where at least one agent has to wait for at least one stage to get served. Otherwise, it can be solved and reduced to the static setting where each facility can serve all the arriving agents in a stage, and thus, agents will only have distance cost.

Denote (X_{-j}, x'_j) as the tuple X with x'_j in place of x_j, and (R_{-j}, r'_j) as the tuple R with r'_j in place of r_j. Below, we provide a formal definition of strategyproofness.

Definition 1. *A mechanism F is strategyproof if for all X, R, $x'_j \in \mathbb{I}$ and $r'_j \geq r_j$, we have*

$$\begin{aligned} &cost(F(X, R, C), x_j, r_j) \\ \leq &cost(F((X_{-j}, x'_j), (R_{-j}, r'_j), C), x_j, r_j). \end{aligned}$$

Given a mechanism F and a profile (X, R, C), the social cost function is defined as $SC[F(X, R, C)] = \sum_{j \in N} cost(F(X, R, C), x_j, r_j)$ and the maximum cost function is defined as $MC[F(X, R, C)]$=$\max_{j \in N} cost(F(X, R, C), x_j, r_j)$. A mechanism F achieves an approximation ratio of ρ for the social cost (resp. maximum cost), if for any profile (X, R, C), $SC[F(X, R, C)] \leq \rho \cdot SC[OPT(X, R, C)]$ (resp. $MC[F(X, R, C)] \leq \rho \cdot MC[OPT(X, R, C)]$) where $OPT(X, R, C)$ is the optimal solution that minimizes the social or maximum cost.

3 MSFLPs-CC With Waiting Cost

In this section, we study the randomized strategyproof mechanisms for multi-stage facility location problem with capacity constraints.

3.1 Equal Capacity Setting with Waiting Cost

In this subsection, we assume all facilities have equal capacity such that $c_1 = c_2 = \cdots = c_k = c = \frac{n}{k}$. We first design a randomized strategyproof mechanism that has a better performance if the penalty coefficient (i.e., d) is large. Mechanism 1 places all the facilities at the median of all agents. It will gradually allocate facilities from 1 to k as long as at least c agents have arrived but have not been served in each stage. Suppose the allocations of facilities from 1 to $i-1$ have been determined. Denote the agents that have arrived before or in stage t but have not been served by facilities at $f_1, f_2, \ldots, f_{i-1}$ as $M_{t,i} = \{j | r_j \leq t, j \in N - \sum_{k=1}^{i-1} N_k\}$.

Mechanism 1 *Let $f_1 = \cdots = f_k = x_{\lceil \frac{n}{2} \rceil}$. Starting from stage 1 and facility $i = 1$, we consider one of the following two cases. A function $q()$ is used to help randomly pick c agents with equal probability.*
Case 1. *If $|M_{t,i}| \geq c$, let $N_i = q(M_{t,i})$ and $s_i = t$. We then continue with stage $t+1$ and facility $(i+1)$.*
Case 2. *Otherwise, no agents are served in stage t. We continue with stage $t+1$ and facility i.*

Lemma 1. *Mechanism 1 is strategyproof. It has approximation ratios of $\frac{n}{2d} + 1$ for the social cost and $\frac{1}{d} + 1$ for the maximum cost.*

Proof. It is clear that no agent can decrease their distance cost by misreporting their location when all the facilities are placed at the median of all the agents. Besides, arrived agents are all selected with equal probability, and no agent can be served earlier by misreporting their location or a later arrival stage. We use OPT_d and OPT_w to indicate the target distance cost and the waiting cost of the agents in the optimal solution. Since the facility serves agents as soon as at least c agents have arrived in the current stage, the total waiting cost will not exceed the optimal solution's total waiting cost. And the sum of distance cost is at most $\frac{n}{2}$. Therefore, the approximation ratio for social cost is

$$\frac{Mechanism}{OPT} \leq \frac{\frac{n}{2} + OPT_w}{OPT_d + OPT_w} \leq \frac{n}{2d} + 1.$$

The maximum of distance cost is at most 1. Therefore, the approximation ratio for maximum cost is

$$\frac{Mechanism}{OPT} \leq \frac{1 + OPT_w}{OPT_d + OPT_w} \leq \frac{1}{d} + 1.$$

□

We also provide an example to show that these bounds are tight.

Example 1. Consider there are two facilities with an equal capacity of c, c agents at 0, and c agents at 1. All the agents at 0 arrive at stage 1, one agent at 1 arrives at stage 2,

[1] One implementation of $q()$ is to permute agents by a random order π, and output the agents $\pi(1), \ldots, \pi(c)$.

Algorithm 1 FindMinLen(X,c)

Input: Agent location profile X and the equal capacity c.
Parameter: Define $length(i, m)$ be the minimum length covering the first i agents with m intervals.
Initialize $length(i, 1) =$

$$\begin{cases} x_i - x_1 & \textbf{if } i \leq c, \\ \infty & \textbf{otherwise}. \end{cases}$$

Output:$length(n, k)$.

for $i \in \{2, \ldots, n\}, j \in \{i-c, \ldots, i-1\}, m \in \{2, \ldots, k\}$ **do**
  $length(i, m) = \min_{j \in \{i-c,\ldots,i-1\}}\{\max\{length(j, m-1), x_i - x_{j+1}\}\}$
end for
return $length(n, k)$

and $c - 1$ agents at 1 arrive at stage 3. Mechanism 1 places two facilities at the median point of all these agents. Therefore, both facilities are placed at 0. Agents at 0 will all be served in stage 1, and agents at 1 will all be served in stage 3. Assuming the waiting cost of each stage is d, the social cost of Mechanism 1 is $\frac{n}{2} + d$ and the maximum cost is $1 + d$. In the optimal solution, one facility is placed at 0, serving all the agents at 0 in stage 1, and the other facility is placed at 1 to serve all the agents at 1 in stage 3, which has a social cost and maximum cost of d. Hence, the approximation ratio is $\frac{n}{2d} + 1$ for social cost and $\frac{1}{d} + 1$ for maximum cost, which shows our analysis is tight.

We also consider the case when the penalty coefficient (i.e., d) is relatively small, and design a randomized strategyproof mechanism where no agents are served until they all arrived. Mechanism 2 is a nontrivial extension of the Equal-Cost randomized mechanism for the classical setting without any capacity constraints for $k \geq 2$ facilities [11]. Mechanism 2 aims to find at most k disjoint intervals such that each interval contains at most c agents (where c is the capacity of the facility), and has a minimum covering length Len. Notice that in order to minimize the maximum length of these intervals, it must be one of the distances between any two agents a and b, i.e., $Len = |x_a - x_b|$. However, different from the Equal-Cost randomized mechanism where each interval has an identical length, some of these lengths might contain more than c agents in which case we can determine and eliminate through the following algorithm $FindMinLen(X, c)$. Notice that Mechanism 2 achieves the same perfomance and guarantees strategyproofness even if there is spare capacity, i.e., $kc > n$. $FindMinLen(X, c)$ runs in $O(nkc)$ to determine a minimum covering length for each interval and covers at most c agents. Finally, it will wait until all the agents have arrived, and use a random map function $g : \{1, \ldots, k\} \to \{0, \ldots, k-1\}$ to generate, with equal probability, a permutation order in which facilities serve their agents.

Mechanism 2 *The mechanism performs the following steps.*

1. *Use* $FindMinLen(X, c)$ *to find a profile of allocations* $O = \{N_1, \ldots, N_k\}$ *such that* N_i *contains all the agents in the* $i-th$ *interval from left to right and* $Len = \max_{i\in[k]}\{R_i - L_i\}$ *is minimized where* L_i *and* R_i *are the locations of the leftmost agent and the rightmost agent served by facility* i *given* N_i*, respectively.*

2. *If* $L_i + Len \leq 1$*,* f_i *is placed at* L_i *with probability* $1/2$ *and* $L_i + Len$ *with probability* $1/2$*. Otherwise,* f_i *is placed at* R_i *with probability* $1/2$ *and* $R_i - Len$ *with probability* $1/2$*.*
3. *For each facility in the allocations, we pick an arbitrary stage in* $\{T, \ldots, T+k-1\}$ *for the facility to serve the assigned agents, such that* $S = \{T+g(1), \ldots, T+g(k)\}$*.*

Lemma 2. *Mechanism 2 is strategyproof. It has approximation ratios of* $T(n-c)+1$ *for social cost and max*$\{2, T+k-2\}$ *for maximum cost.*

Proof. We first show that $FindMinLen(X, c)$ returns minimum Len. Suppose there exists a division of intervals with $Len' < Len$ such that all the agents are covered while each interval contains at most c agents. However, Len' can still satisfy the requirements of $FindMinLen(X, c)$. Thus $FindMinLen(X, c)$ will not return Len, which contradicts our assumption.

Because all agents have the expected distance cost of $\frac{Len}{2}$. Any agent may only benefit by reducing the minimum cover length Len. If an agent misreports their location so that the minimum cover length becomes $Len', Len' < L$, the expected distance cost for him will increase to $\frac{Len+(Len-Len')}{2}$. Therefore, agents will not benefit by misreporting. The optimal social distance cost is at least Len since Len is the minimum distance that can cover all the agents in each stage. The total distance cost of Mechanism 2 is at most $\frac{n}{2} \cdot Len$. We use OPT_d and OPT_w to indicate the target distance cost and the waiting cost of the agents in the optimal solution. The worst case in terms of the total waiting cost is when $n-c$ agents arrive in stage 1 and c agents arrive in stage T. The sum of waiting cost of Mechanism 2 is at most $OPT_w + T(n-c)d \leq (T(n-c)+1) \cdot OPT_w$. The approximation ratio for social cost is

$$\frac{Mechanism}{OPT} \leq \frac{(T(n-c)+1) \cdot OPT_w + \frac{n}{2} \cdot OPT_d}{OPT_w + OPT_d}$$
$$\leq T(n-c)+1.$$

The optimal maximum distance cost is at least $\frac{Len}{2}$, and the expected maximum cost of Mechanism 2 is Len. The maximum of waiting cost is at most $(T+k-2)d \leq (T+k-2)\dot{O}PT_w$. Therefore, the approximation ratio for maximum cost is

$$\frac{Mechanism}{OPT} \leq \frac{(T+k-2)\dot{O}PT_w + 2OPT_d}{OPT_d + OPT_w} \leq max(2, T+k-2).$$

□

Lemma 3. *For* $k \geq 2$*, any randomized strategyproof mechanism has an approximation ratio of at least* $\frac{2}{4d+3} + 1$ *for social cost.*

Proof. We construct the first configuration such that $\frac{kc}{2} + 1$ agents locate at 0 and $\frac{kc}{2} - 1$ agents locate at $1, k \geq 2, c \geq 3$.There are k (k is an even number) facilities with an equal capacity of c. We first focus on the approximation ratio of the distance cost. Now one agent at 0 is moved to $\frac{c-2}{4(c-1)}$. Let $p(x)$ be the probability density function of the probability that the facility serving the agent at $\frac{c-2}{4(c-1)}$ is placed at x in the new configuration. Denote $\bar{x}_1$, $\bar{x}_2$ and $\bar{x}_3$ as the expected facility locations in the intervals $[0, \frac{c-2}{4(c-1)})$, $[\frac{c-2}{4(c-1)}, \frac{c-2}{2(c-1)}]$ and $(\frac{c-2}{2(c-1)}, 1]$ respectively. Let $P_1 = \int_0^{\frac{c-2}{4(c-1)}} p(x)dx$, $P_1 \cdot (\frac{c-2}{4(c-1)} - \bar{x}_1) = \int_0^{\frac{c-2}{4(c-1)}} p(x)xdx$, $P_2 = \int_{\frac{c-2}{4(c-1)}}^{\frac{c-2}{2(c-1)}} p(x)dx$, $P_2 \cdot (\frac{c-2}{4(c-1)} + \bar{x}_2) = \int_{\frac{c-2}{4(c-1)}}^{\frac{c-2}{2(c-1)}} p(x)xdx$, $P_3 = \int_{\frac{c-2}{2(c-1)}}^{1} p(x)dx$, $P_3 \cdot (\frac{c-2}{4(c-1)} + \bar{x}_3) = \int_{\frac{c-2}{2(c-1)}}^{1} p(x)xdx$. Suppose the approximation ratio of a randomized strategyproof mechanism for distance cost is α. In the first configuration, the distance cost of an agent at 0 is at most $(1-P) \cdot \frac{c-2}{4(c-1)} + P \cdot (1 - \frac{c-2}{4(c-1)})$, and the social cost is $P \cdot (\frac{kc}{2} + 1) \leq \alpha$, where P is the probability that agents at 0 are served by a facility placed at 1 and the optimal social cost in the first configuration is 1. By strategyproofness, the agent at $\frac{c-2}{4(c-1)}$ in the new configuration cannot benefit by misreporting to 0. Therefore, we have

$$\begin{aligned} P_1 \cdot \bar{x}_1 + P_2 \cdot \bar{x}_2 + P_3 \cdot \bar{x}_3 \leq & \frac{\alpha}{\frac{kc}{2}+1} \cdot (1 - \frac{c-2}{4(c-1)}) + (1 - \frac{\alpha}{\frac{kc}{2}+1}) \cdot \frac{c-2}{4(c-1)} \\ = & \frac{\alpha}{\frac{kc}{2}+1} + (1 - \frac{2\alpha}{\frac{kc}{2}+1}) \cdot \frac{c-2}{4(c-1)}. \end{aligned}$$

The optimal social cost of the new configuration is $\frac{3c-2}{4(c-1)}$. If the facility serving the agent at $\frac{c-2}{4(c-1)}$ is placed somewhere in $[0, \frac{c-2}{2(c-1)}]$, the optimal allocation is to serve $c-1$ agents at 0 with the agent at $\frac{c-2}{4(c-1)}$ together. If the facility serving the agent at $\frac{c-2}{4(c-1)}$ is placed in $(\frac{c-2}{2(c-1)}, 1]$, the optimal allocation is to serve $c-1$ agents at 1 and the agent at $\frac{c-2}{4(c-1)}$ together. Therefore, for the social cost of any randomized mechanism, we have

$$\begin{aligned} \alpha \cdot \frac{3c-2}{4(c-1)} \geq & P_1 \cdot ((c-1)(\frac{c-2}{4(c-1)} - \bar{x}_1) + \bar{x}_1 + 1) \\ & + P_2 \cdot ((c-1)(\frac{c-2}{4(c-1)} + \bar{x}_2) + \bar{x}_2 + 1) \\ & + P_3 \cdot ((c-1)(1 - \frac{c-2}{4(c-1)} - \bar{x}_3) + \bar{x}_3) \\ \geq & (P_1 + P_2 + P_3)\frac{c+2}{4} + P_3 \cdot \frac{c-2}{2} + P_2 c\bar{x}_2 - (P_1\bar{x}_1 + P_3\bar{x}_3)(c-2) \\ \geq & \frac{c+2}{4} + P_2 c\bar{x}_2 + (c-2)(P_2\bar{x}_2 - \frac{\alpha}{\frac{kc}{2}+1} - (1 - \frac{2\alpha}{\frac{kc}{2}+1}) \cdot \frac{c-2}{4(c-1)}) \\ \geq & \frac{5c-6}{4(c-1)} - \frac{c(c-2)\alpha}{(kc+2)(c-1)}. \end{aligned}$$

Thus, we have $\alpha \geq \frac{5kc^2-(6k-10)c-12}{(3k+4)c^2-(2k+2)c-4}$. Suppose there exists a randomized strategyproof mechanism with an approximation ratio of β $(1 \leq \beta < \frac{5}{3})$ for distance cost. Let $k = c^2$ and $c > \frac{6+4\beta}{5-3\beta} \geq 5$. We have $\frac{5kc^2-(6k-10)c-12}{(3k+4)c^2-(2k+2)c-4} > \frac{5c^4-6c^3}{3c^4+4c^3} > \beta$, which contradicts the previous statement. Therefore, the sum of distance costs is at least $\frac{5}{3} \cdot OPT_d$. The sum of waiting cost is at least $OPT_w = d$. Let $k = c^2$ and c be as large as possible.

According to this instance, the approximation ratio is at least

$$\frac{Mechanism}{OPT} \geq \frac{\frac{5}{3} \cdot OPT_d + OPT_w}{OPT_d + OPT_w} \rightarrow \frac{2}{4d+3} + 1. \quad (1)$$

□

Lemma 4. *For $k \geq 2$, any randomized strategyproof mechanism has an approximation ratio of at least $\frac{1}{4d+2} + 1$ for maximum cost.*

Proof. We extend the proof of the lower bound of randomized strategyproof mechanisms in single facility location problem in [15] to this setting. Consider a configuration where we place $2(k-1)$ agents at 0, one agent at $1-2\epsilon$, and one agent at $1-\epsilon$. The capacity of each facility is 2 and all the agents arrive at the same stage. We assume ϵ is small enough such that the agent at $1-\epsilon$ and the agent at 1 are served by the same facility due to the approximation ratio. It is clear that at least one agent has the expected distance cost of at least $\frac{\epsilon}{2}$. Without loss of generality, we suppose the agent j at $1-\epsilon$ has the expected distance cost of $E(cost(j)) \geq \frac{\epsilon}{2}$. Now we move the agent at $1-\epsilon$ to 1, due to the strategyproofness, the expected distance between the facility and point $1-\epsilon$ in the new configuration should be at least $\frac{\epsilon}{2}$. Therefore, in the new setting, the expected maximum distance cost is at least $\frac{\epsilon}{2} + \epsilon$ while the optimal maximum distance cost is ϵ, which proves a minimum approximation ratio of $\frac{3}{2}$ for distance cost. The sum of waiting cost is at least $OPT_w = d$. Let $k = 2$ and move one agent at $1-2\epsilon$ to point 0. According to this instance, the approximation ratio is at least

$$\frac{Mechanism}{OPT} \geq \frac{\frac{3}{2} \cdot OPT_d + OPT_w}{OPT_d + OPT_w} \rightarrow \frac{1}{4d+2} + 1.$$

□

3.2 Arbitrary Capacity with Waiting Cost

In this subsection, we consider k facilities with arbitrary capacities such that $n = \sum_{i=1}^{k} c_i$, and provide two randomized strategyproof mechanisms for both social cost and maximum cost. Observe that the equal capacity setting is a special case of this setting. Therefore, we can inherit all the lower bounds in Subsection 3.1.

Similar to Mechanism 1, serving agents earlier can achieve a better performance when d is large. Thus, we present Mechanism 3 which will serve agents with an optimal allocation that minimizes the waiting cost. It can be proved by a similar method that finding an optimal allocation (for the total or the maximum waiting cost) in the arbitrary capacity setting is also NP-hard [1]. We introduce a dynamic programming algorithm $MinWaiting(R, C)$ to determine the allocation that minimizes the total (resp. maximum) waiting cost. The sub-routine is similar to $FindMinCover(X, C)$ by keeping track of an allocation that minimizes the target cost up to the i-th arriving agents (see the pseudo-code in our supplementary materials).

Let $Cap \leftarrow MinWaiting(R, C)$, and $f_1 = \cdots = f_k = x_{\lceil \frac{n}{2} \rceil}$. Starting from stage 1 and facility $i = 1$, we consider one of the following two cases. A function $q(Cap[i])$

Algorithm 2 MinWaiting(R,C)

Input: Agent arrival stage profile R in an increasing order and facilities' capacities C
Parameter: Initialize an array Waiting[n][2^k] to $n \cdot T$ and an array of empty tuples Cap[n][2^k], where Waitng[i][w] stores the minimum target waiting cost for agents $\{1, 2, \dots i\}$ using a set of facilities F, and Cap[i][w] stores the allocation used for agents $\{1, 2, \dots i\}$ with F , where $w = \sum_{f_j \in F} 2^{j-1}$. Denote $cost(N)$ as the target waiting cost function, i.e., sum or maximum.
Output: An allocation Cap that achieves the minimum target waiting cost.
Mechanism 3
for $i \in \{1, \dots, n\}, j \in \{1, \dots, k\}, w \in \{1, \dots, 2^k\}$ **do**
  if $i = c_j$ and $w = 2^{j-1}$ **then**
    Waiting[i][w] $\leftarrow cost(1, \dots, i)$
    Cap[i][w] $\leftarrow (c_j)$
  else if $c_j \notin$ Cap[$i - c_j$][$w - 2^{j-1}$] **then**
    Temp$\leftarrow cost$(Waiting[$i - c_j$][$w - 2^{j-1}$]$, i - c_j + 1, \dots, i$)
    if Waiting[i][w]>Temp **then**
      Waiting[i][w] $\leftarrow$ Temp
      Cap[i][w] $\leftarrow$ Cap[$i - c_j$][$w - 2^{j-1}$]$.append(c_j)$
    end if
  end if
end for
return Cap[n][$2^k - 1$]

is used to help randomly pick $Cap[i]$ agents with equal probability.
Case 1. *If $|M_{t,i}| \geq Cap[i]$, $N_i = q(M_{t,i})$ and $s_i = t$. We then continue with stage $t+1$ and facility $(i+1)$.*
Case 2. *Otherwise, no agents are served in stage t. We continue with stage $t+1$ and facility i.*

Lemma 5. *Mechanism 3 is strategyproof. It has approximation ratios of $\frac{n}{2d} + 1$ for the social cost, and $\frac{1}{d} + 1$ for the maximum cost.*

Proof. Notice that for any feasible solution, the amount of waiting cost decrease due to the misreporting is the same. The total waiting cost of Mechanism 3 will not exceed the optimal solution's total waiting cost. The sum of distance cost is at most $\frac{n}{2}$, and the maximum distance cost is at most 1. By a similar proof of Lemma 1, it is clear that the approximation ratios for the social cost and the maximum cost are $\frac{n}{2d} + 1$ and $\frac{1}{d} + 1$ respectively. □

Notice that finding an (asymptotically) optimal allocation is necessary to achieve bounded approximation ratios when d is small.

Example 2. Consider two facilities with capacities 1 and 2, and three agents at $\{0, x, 1\}$. If we do not always use the optimal one, there are three other possible allocations. The approximation ratio using any one of the three can be arbitrarily large i.e., $x/(1-x)$.

Thus, we then introduce a dynamic programming algorithm, $FindMinCover(X, C)$, used by Mechanism 4, to determine the allocations and the facilities' locations. In $FindMinCover(X, C)$, we use an array Len[n][2^k] to store the minimum covering

Algorithm 3 FindMinCover(X,C)

Input: Agent location profile X and facilities' capacities C

Parameter: Initialize an array Len[n][2^k] to 1 and an array of empty tuples Cap[n][2^k], where Len[i][w] stores the minimum covering length for agents $\{1, 2, \dots i\}$ using a set of facilities F, and Cap[i][w] stores the allocation used for agents $\{1, 2, \dots i\}$ with F , where $w = \sum_{f_j \in F} 2^{j-1}$.

Output: Minimum length Len covering agents in different groups and its corresponding allocations Cap.

for $i \in \{1, \dots, n\}, j \in \{1, \dots, k\}, w \in \{1, \dots, 2^k\}$ **do**
  if $i = c_j$ and $w = 2^{j-1}$ **then**
    Len[i][w] $\leftarrow x_i - x_1$
    Cap[i][w] $\leftarrow (c_j)$
  else if $c_j \notin$ Cap[$i - c_j$][$w - 2^{j-1}$] **then**
    Temp$\leftarrow \max\{$Len[$i - c_j$][$w - 2^{j-1}$]$, x_i - x_{i-c_j+1}\}$
    if Len[i][w]>Temp **then**
      Len[i][w] $\leftarrow$ Temp
      Cap[i][w] $\leftarrow$ Cap[$i - c_j$][$w - 2^{j-1}$]$.append(c_j)$
    end if
  end if
end for
return Len[n][$2^k - 1$], Cap[n][$2^k - 1$]

length and an array of tuples Cap[n][2^k] to store the capacities of the allocated facilities based on their locations from left to right. Within the algorithm, we represent each facility i as a binary vector with value 2^i and w is a value (sum of powers of two) to record which facilities have been used and fully occupied so that we will not allocate them again. Mechanism 4, will wait until all the agents have arrived, and use a random map function $g : \{1, \dots, k\} \to \{0, \dots, k-1\}$ to generate, with equal probability, a permutation order in which facilities serve their agents.

Mechanism 4 *Let* $(Len, Cap) \leftarrow FindMinCover(X, C)$. $N_i = \{1 + \sum_{w=1}^{i-1} Cap[w], \dots, Cap[i] + \sum_{w=1}^{i-1} Cap[w]\}$. *If* $L_i + Len \leq 1$*, the facility i is placed at* L_i *with probability* $\frac{1}{2}$ *and* $L_i + Len$ *with probability* $\frac{1}{2}$. *Otherwise,* f_i *is placed at* R_i *with probability* $\frac{1}{2}$ *and* $R_i - Len$ *with probability* $\frac{1}{2}$. *For each facility in the allocations, we pick an arbitrary stage in* $\{T, \dots, T+k-1\}$ *for the facility to serve the assigned agents, such that* $S = \{T + g(1), \dots, T + g(k)\}$.

Lemma 6. *Mechanism 4 is strategyproof, which achieves approximation ratios of* $Tn - T + 1$ *for social cost and max*$\{2, T+k-2\}$ *for maximum cost.*

Proof. The sum of distance cost of Mechanism 4 is at most $\frac{n}{2} \cdot OPT_d$. The total waiting cost is at most $OPT_w + T(n-1) \cdot d \leq (Tn - T + 1) \cdot OPT_w$. The approximation ratio for social cost is

$$\frac{Mechanism}{OPT} \leq \frac{(Tn - T + 1) \cdot OPT_w + \frac{n}{2} \cdot OPT_d}{OPT_w + OPT_d} \leq Tn - T + 1.$$

The maximum cost of Mechanism 4 is at most $(T+k-2)d + 2OPT_d$. Therefore, the approximation ratio for maximum cost is

$$\frac{Mechanism}{OPT} \leq \frac{(T+k-2)\dot{O}PT_w + 2OPT_d}{OPT_d + OPT_w} \leq max(2, T+k-2).$$

□

4 Conclusion

We initiate the study of the multi-stage facility location problem with capacity constraints from a mechanism design perspective. For settings with various capacity constraint configurations with waiting times, we provide randomized strategyproof mechanisms with approximation guarantees as well as lower bounds to the corresponding settings. There are several directions for future work. It will be intriguing to extend results beyond one dimension to other complex structures like trees and networks. Another approach to extend this work is to consider opening multiple facilities at each stage. Finally, one can also consider a setting where agents have preferences over heterogeneous facilities.

Acknowledgements

We gratefully acknowledge funding from Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. UGC/FDS13/E01/20), and by NSFC No.62202100. Hau Chan is supported by the National Institute of General Medical Sciences of the National Institutes of Health [P20GM130461], the Rural Drug Addiction Research Center at the University of Nebraska-Lincoln, and the National Science Foundation under grant IIS:RI #2302999. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

References

1. Aziz, H., Chan, H., Lee, B., Li, B., Walsh, T.: Facility location problem with capacity constraints: Algorithmic and mechanism design perspectives. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 34, pp. 1806–1813 (2020)
2. Aziz, H., Chan, H., Lee, B.E., Parkes, D.C.: The capacity constrained facility location problem. Games and Economic Behavior **124**, 478–490 (2020)
3. Black, D.: The decisions of a committee using a special majority. Econometrica: Journal of the Econometric Society pp. 245–261 (1948)
4. Blin, J.M., Satterthwaite, M.A.: Strategy-proofness and single-peakedness. Public Choice pp. 51–58 (1976)
5. Brimberg, J., Korach, E., Eben-Chaim, M., Mehrez, A.: The capacitated p-facility location problem on the real line. International Transactions in Operational Research **8**(6), 727–738 (2001)

6. Chan, H., Filos-Ratsikas, A., Li, B., Li, M., Wang, C.: Mechanism design for facility location problems: A survey. In: Zhou, Z.H. (ed.) Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21. pp. 4356–4365. International Joint Conferences on Artificial Intelligence Organization (8 2021). https://doi.org/10.24963/ijcai.2021/596, `https://doi.org/10.24963/ijcai.2021/596`, survey Track
7. Chen, D.Q., Li, A., Li, G.Z., Marathe, M., Srinivasan, A., Tsepenekas, L., Vullikanti, A.: Efficient and equitable deployment of mobile vaccine distribution centers. In: Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence. pp. 64–72 (2023)
8. De Keijzer, B., Wojtczak, D.: Facility reallocation on the line. Algorithmica pp. 1–28 (2022)
9. Demaine, E.D., Hajiaghayi, M., Mahini, H., Sayedi-Roshkhar, A.S., Oveisgharan, S., Zadimoghaddam, M.: Minimizing movement. ACM Transactions on Algorithms (TALG) **5**(3), 1–30 (2009)
10. Fotakis, D., Tzamos, C.: Winner-imposing strategyproof mechanisms for multiple facility location games. In: International workshop on internet and network economics. pp. 234–245. Springer (2010)
11. Fotakis, D., Tzamos, C.: Strategyproof facility location for concave cost functions. In: Proceedings of the fourteenth ACM conference on Electronic commerce. pp. 435–452 (2013)
12. Fotakis, D., Tzamos, C.: On the power of deterministic mechanisms for facility location games. ACM Transactions on Economics and Computation (TEAC) **2**(4), 1–37 (2014)
13. Lu, P., Sun, X., Wang, Y., Zhu, Z.A.: Asymptotically optimal strategy-proof mechanisms for two-facility games. In: Proceedings of the 11th ACM conference on Electronic commerce. pp. 315–324 (2010)
14. Moulin, H.: On strategy-proofness and single peakedness. Public Choice **35**(4), 437–455 (1980)
15. Procaccia, A.D., Tennenholtz, M.: Approximate mechanism design without money. ACM Transactions on Economics and Computation (TEAC) **1**(4), 1–26 (2013)
16. Raghavan, S., Sahin, M., Salman, F.S.: The capacitated mobile facility location problem. European Journal of Operational Research **277**(2), 507–520 (2019)
17. Wada, Y., Ono, T., Todo, T., Yokoo, M.: Facility location with variable and dynamic populations. In: AAMAS. pp. 336–344 (2018)
18. Walsh, T.: Strategy proof mechanisms for facility location with capacity limits. In: Raedt, L.D. (ed.) Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22. pp. 527–533. International Joint Conferences on Artificial Intelligence Organization (7 2022). https://doi.org/10.24963/ijcai.2022/75, `https://doi.org/10.24963/ijcai.2022/75`, main Track
19. Wang, X., Chau, V., Chan, H., Fong, C.K.K., Li, M.: Multi-stage facility location problems with transient agents. In: Proceedings of the AAAI Conference on Artificial Intelligence (2023)