Hotelling-Downs with Facility Synergy: The Mall Effect

Elliot Anshelevich[1[0000-0001-9757-6839]], Jianan Lin[2], and Noah Prisament[3[0009-0006-0989-0689]]

[1] Rensselaer Polytechnic Institute (RPI), Troy, NY
eanshel@cs.rpi.edu
[2] Rensselaer Polytechnic Institute (RPI), Troy, NY
hcm6755@gmail.com
[3] Rensselaer Polytechnic Institute (RPI), Troy, NY
nprisament@gmail.com

Abstract. We consider a variation of the classic Hotelling-Downs model with the addition of facility synergies. Unlike in the classic model, where clients always use the facility closest to them, we study clients who prefer locations with many facilities to those with few facilities while simultaneously attempting to minimize their distance as well. We show that, in contrast with the classic model, Nash equilibria for our setting always exist, and, in fact, there always exists a Nash equilibrium such that the sum of client costs equals the cost of the optimum solution. Our main result is a bound of $\frac{225}{64} \approx 3.516$ on the Price of Anarchy for our model, showing that, although the client behavior is more complex in our model (and often more realistic depending on the application), the cost of Nash equilibrium solutions still cannot be much worse than the cost of the optimum facility placement.

Keywords: Price of Anarchy · Hotelling Games · Competitive Facility Location.

1 Introduction

Beginning with Hotelling's seminal work almost a century ago [22], the Hotelling-Downs model has grown to become one of the classics of game theory. Originally defined using the story of ice cream vendors choosing their placement along a beach, it has been since used to model a variety of applications, including the strategic placement of facilities and firms, as well as the formation of political parties and the ideological placement of strategic political candidates.

In most existing work on the topic, there are *clients* located along some interval X, and a set of k *facilities* or firms. Each facility can individually choose a location $x_i \in X$ where it will be placed; it is allowed to choose any location in X. In the classic model, each client simply utilizes the closest facility to them, i.e., a client located at x would shop at the position x_i with a facility such that $|x-x_i|$ is as small as possible. Note that in political and social choice applications, this corresponds to a voter/client supporting the party/candidate which is closest to them. Knowing this client behavior, the goal of each facility is typically to position themselves in order to maximize the total fraction of clients using their facility. If multiple facilities choose the same location, then they typically share the clients equally. More precisely, let u_i represent the fraction of clients which use location x_i, i.e., the fraction of clients for whom x_i is the closest location of any point with a facility at it; and let n_i be the number of facilities at x_i. Then, the *utility* of a facility located at x_i is defined to be u_i/n_i: it is the total fraction of clients which use that location, divided by the number of facilities located there.

Many important variations of the above classic model have been studied for many diverse applications, as we discuss in the Related Work section. However, an important aspect missing from existing work is that of *facility synergy*. In many realistic applications, a client does not simply wish to go to the closest facility, but prefers to go to a location with *multiple* facilities if possible. Consider, for example, someone who wishes to go shopping for a luxury good, such as designer shoes. They could go to a nearby store, but they may not find something that suits them. Or they could instead go to a collection of stores all near each other, even if these stores are farther away. Similarly, when people need to buy goods without knowing exactly what they are looking for, they are much more likely to go to a shopping mall with its many options than simply to the closest store. Such behavior does not only apply to shopping, of course. When a group of friends wants to go to a restaurant or for entertainment but does not want to decide on a place in advance, they are likely to go to a place with a large collection of options, even if it is farther away. When joining a group (such as patients choosing a hospital or a medical practice, or graduate students choosing a research group), there is a benefit to joining a larger group since there will be many options and more support, even if there is a smaller group which is technically closer to your interests. This also applies to political parties: while a small party may exist which is closer to your views, people will often join larger parties due to the benefits which come from having a large number of powerful representatives who are members of your party.

In this paper, we begin the study of facility synergy by considering the simplest model possible, which we believe captures the basic fundamental differences with the classic model without synergy. To model the fact that clients prefer to go to a location with more facilities but still care about the distance as well, we define the client cost function $c(x, x_i) = |x - x_i|/n_i$ to be the cost that a client positioned at x experiences for going to location x_i with n_i facilities located there. Each client then uses the location that minimizes their cost $c(x, x_i)$, not simply the distance. Thus, clients prefer to use closer facilities, but at the same time prefer locations which contain more facilities. u_i is still defined as before: it is the total fraction of clients who use location x_i, and thus the utility that each facility receives is still just u_i/n_i.

1.1 Our Contributions

As done previously for the classic Hotelling-Downs model, our goal in this paper is to understand the structure of Nash equilibrium solutions which result from the self-interested behavior of the facilities. We are especially interested in the relationship of Nash equilibrium solutions with the optimum facility placement, as quantified by Price of Anarchy and Price of Stability measures.[4] In other words, we are interested in quantifying how much the self-interest of the facilities hurts the clients as compared to facilities being placed by some central and altruistic decision maker. Here, the *cost* of a solution is defined as the total cost of all the clients, as in most previous work.[5]

Although similar to the classic model, the difference in client behavior in our model results in very different properties of Nash equilibria and behavior of the facilities. This mainly stems from the fact that, unlike in the classic model where a set of clients using a particular facility location always forms a contiguous interval, in our model these sets can be discontinuous. In other words, while some interval of clients next to a set of facilities at x_i will indeed use location x_i, there can be other intervals far away from x_i which will also use x_i, as we discuss in Section 2. We call such intervals *bubbles*. The existence of bubbles greatly changes the structure of Nash equilibrium solutions and requires new techniques for their analysis. Another notable difference which complicates the analysis is that in the classic model, an addition of an extra facility always decreases the utilities of all the other facilities; in our model, however, it can decrease the utilities of some, but improve the utilities of others due to clients taking into account facility synergy. Thus, the facility utilities are non-monotone with respect to the addition of other facilities.

Despite the added complexity of the model, we show that in some ways the Nash equilibrium properties are improved in our setting. *We first prove that Nash equilibrium solutions always exist in our model, and in fact that the Price of Stability always equals 1, i.e., that there always exists a Nash equilibrium with the total client costs being as small as possible.* In contrast, it is well-known that a Nash equilibrium *does not* always exist in the classic Hotelling-Downs model [14], and further that even when it does exist, it can be as much as a factor of 2 different in cost from the optimum solution.

We then proceed to our main result, which is an analysis of Price of Anarchy for our model. In the classic Hotelling-Downs model, where clients simply go to the closest facility, the Price of Anarchy is at most 2 [19]. Despite the existence of bubble intervals and other differences in our model behavior, we are able to *prove a bound of* $\frac{225}{64} \approx 3.516$ *on the Price of Anarchy in our model.* Doing this

[4] The Price of Anarchy [23] is the ratio between cost of the worst Nash equilibrium and the cost the optimum facility placement, while the Price of Stability [3] is the same ratio but for the best Nash equilibrium instead of the worst one.

[5] Note that the total sum of utilities of the facilities for our model is always the same, since every client uses some facility location. This is in contrast to previous work such as [17]. However, the cost of the clients can vary greatly depending on the solution.

requires showing various structural results about the properties of Nash equilibrium solutions, such as that the size of bubble intervals is bounded compared to adjacent intervals. Our Price of Anarchy bound shows that although the client behavior is more complex in our model (and often more realistic depending on the application), the cost of Nash equilibrium solutions still cannot be too much worse as compared with the optimum facility placement.

1.2 Related Work

Since Hotelling's seminal work, *Stability in Competition* [22], countless variations of the original model have been studied in depth. Downs extended the Hotelling model to voting patterns in a space of political ideology, realizing the term Hotelling-Downs Model [11]. The model was again expanded upon by [14] who extended the model from strictly analyzing duopolies towards arbitrary n-facility markets and characterized when a pure Nash equilibrium exists and when it does not.[6] This established that the classic model lacks any pure Nash equilibrium for $n = 3$, has a unique pure Nash equilibrium for $n = 2, 4, 5$ and has an infinite number of pure Nash equilibria for $n > 5$. Due to the popularity of this model, we refer to the recent survey by [12] (as well as [2, 16, 21], and [28]) for a larger overview of the breadth of traditional Hotelling-Downs models that have been studied. These surveys denote trends in model variants, including Graitson's enumerated changes in the number of firms, the shapes of the demand curves, and the types of spaces [21]. Depending on the literature, these variants have sometimes been referred to as *Facility Location Games*, *Competitive Facility Location Games*, and *Voronoi Games* — all with essentially the same meaning; however, to minimize confusion, we will refer to them as Hotelling-Downs Games. However, throughout this extensive body of work, we are not aware of any analysis of our client behavior in which clients prefer locations with more facilities.

Many variations of the model involve changing the metric space, X, that clients and facilities occupy. This is often done with the extension to a circle instead of an interval, thereby inducing periodic boundary conditions, as studied by [14] and [29]. Further, a logical next step includes location games in 2-dimensional spaces, e.g. [4], or in nonlinear or discrete markets, including graphs [25], networks [19], and finite sets of locations [26].

Other variations change how facilities attract clients or how said clients choose the facility they will utilize. For example, [17] propose a variation on the original Hotelling-Downs model in which all facilities have a limited attraction interval and "the support of clients that fall in the attraction interval of several agents is randomly shared among the latter." Similarly, [30] study a model akin to that of [17] formalized with an attraction width w_i for each facility and later extend this model to results over arbitrary distributions of clients. These models can both be considered variants of the general set of probabilistic *Shapley*

[6] This characterization was recently amended by [7], who also gave numerous results about the computational complexity of finding a Nash equilibrium.

Facility Location Games from [5]. A modification from the client perspective is captured by [10], who analyze a model where each client has a randomly distributed tolerance interval and they utilize the nearest facility in this interval, if one exists. This captures the concept that clients may not utilize the closest facility if it is too far from them. One type of variation that is close in spirit to our work is that of Hotelling-Downs models that take into account network externalities, such as the client cost function being a linear combination of distance and facility congestion as in [18, 27], or of distance and facility popularity as in [20].

There are also similar models comprising of multi-unit and multi-stage games including those where single agents can place multiple facilities concurrently [6], or in consecutive rounds [1]. Another multi-stage variation includes [24] in which the first round involves facilities choosing locations on a graph and a second stage in which clients distribute their purchasing power. Likewise, multi-stage games include those where location choice is followed by additional differentiation such as price competition leading to "non-symmetric" facilities [15]. There is also a somewhat separate direction of research (see, e.g., [9] and the references therein) which uses a lot of similar terminology of facility location games but is concerned more with developing centralized mechanisms for placing all the facilities based on the locations reported by the clients, with the clients being able to lie about their true locations.

For these game-theoretic models, facility utility functions are generally either the more popular "support maximizers" function (utility is proportional to clients received) that is employed by our model or a "winner-takes-all" function that is generally reserved for modeling political landscapes [17, 30]. Likewise, social welfare is generally modeled as minimizing the social cost due to client travel or prices, but it is occasionally defined as client "participation" in models that are not guaranteed to serve all clients as in [17]. Lastly, the distribution of clients in the space is generally uniform across this body of work; however, there are exceptions that derive results for arbitrary or random distributions of clients such as [26], [8], and others.

Generally, the results derived in existing work focus on the existence and uniqueness of Nash equilibria, the optimal solutions, and the efficiency of said equilibria. A useful measure of the efficiency of equilibria are what are known as the Price of Stability for measuring the best-case efficiency [3], and Price of Anarchy for measuring the worst-case efficiency [13,23]. This is our focus as well.

2 Model and Basic Results

Classic Hotelling-Downs Model First, let us recall the classic Hotelling-Downs model. We are given an interval X; without loss of generality, let us assume that $X = [0, 1]$. There are k facilities. Each facility can choose where in the interval X it should be placed; it is allowed to choose any location in X. In the version of the problem we study, multiple facilities are allowed to place themselves at the same location; in this case, we say that a *stack* of n_i facilities is located at

$x_i \in X$. A *facility placement* is a set of locations x_i and number of facilities n_i at each of these locations, so that $\sum_i n_i = k$.[7]

There are *clients* located in X; as in much of the previous work, we assume that the clients form a continuum and are uniformly distributed in X. In the classic model, the clients would utilize the closest facility to them, i.e., a client located at x would use position x_i with a facility so that $d(x, x_i)$ is as small as possible, where $d(x, x_i)$ is simply the distance $|x - x_i|$. For a given facility placement, let U_i be the set of clients who use the location x_i, i.e., $U_i = \{x \in X : i = \arg\min_j d(x, x_j)\}$. Then u_i (the size of U_i) is the total fraction of clients using location x_i.[8] The *utility* of a facility located at x_i is defined to be u_i/n_i: it is the total fraction of clients which use that location, divided by the number of facilities located there, as clients going to a location are assumed to be equally shared between all the facilities at that location.

The clients are non-strategic, and simply use the facility closest to them. The facilities, on the other hand, choose their locations x_i in order to maximize their utility as defined above. The choices made by the facilities on where they are positioned determines which locations the clients will use, and thus the utilities of the facilities themselves. In other words, the facilities are players in a game where they can choose any $x \in X$ as their strategy, and their utility is as defined above. A pure *Nash equilibrium* is a solution (i.e., facility placement) in which no single facility can increase their utility by changing their location.[9] It is well known that a Nash equilibrium may not exist for this classic model [14], and there are results about equilibrium quality when it does exist as well.

Adding Client Preferences for Facility Synergy As discussed in the Introduction, we instead consider clients who care both about the distance to x_i and the number of facilities at x_i. We choose the simplest model which captures the essence of such clients, and define the client cost function $c(x, x_i) = d(x, x_i)/n_i$ to be the cost that a client at x experiences for going to location x_i.[10] Clients now choose to go to x_i which minimizes the above cost, which increases with the distance but decreases with n_i. With this cost function, clients are indifferent between going to a location with 1 facility which is distance y away, and going to a location with 2 facilities which is $2y$ away. As a first step toward modeling facility synergy, we consider this natural, since the travel cost per facility visited remains the same in both. In this new model, U_i is still defined as the set of clients who choose to use x_i, that is, $U_i = \{x \in X : i = \arg\min_j c(x, x_j)\}$. u_i is still defined as before, and thus the utility that each facility located at x_i receives is still just u_i/n_i. The only difference is which locations the clients choose.

[7] i is formally defined as an arbitrary index from among all facility stacks.

[8] Note that U_i is usually an infinite set. By the "size of U_i", we mean the total fraction of clients in U_i compared to X. For example, if U_i consists of all clients in the interval $[\frac{2}{3}, 1]$, then u_i would equal $\frac{1}{3}$. Formally, u_i is defined as $u_i = \int_{x \in U_i} dx$.

[9] In this paper, we will only focus on pure Nash equilibria, in which facilities must pick a specific location instead of a randomized strategy.

[10] All our results hold in exactly the same way if we instead define $c(x, x_i) = d(x, x_i)/(\gamma n_i)$ for some constant γ.

Solution Cost and Price of Anarchy We will study the existence of pure Nash equilibria as well as their quality. Many different measures have been considered in the literature for the quality of solutions in Hotelling-Downs and similar models. For our model (as well as the classic model defined above), the total utility of the facilities always equals $|X| = 1$, since the clients always go somewhere.[11] The total cost of the clients, however, can be greatly impacted by how the facilities locate themselves. Because of this, as in the classic model, we consider as our objective function the total cost of the clients in the solution. More precisely, for fixed choice of locations by the facilities, we can define the cost of a client at x to be

$$c(x) = \min_i c(x, x_i).$$

Then the total client cost is simply

$$\int_{x\in X} c(x)\ dx.$$

Thus the optimum facility locations are the ones which minimize the above quantity. It is not difficult to see that in the classic model where $c(x, x_i) = d(x, x_i)$, the optimum solution is simply to equally space the facilities inside the interval, although that solution is not a Nash equilibrium. For our model, however, where $c(x, x_i) = d(x, x_i)/n_i$, there can be many optimum solutions, as discussed in the next section.

We study both the Price of Anarchy and Price of Stability of this game. The Price of Anarchy is the ratio between the cost of the *worst* (largest cost) Nash equilibrium and the cost of the optimum solution. It represents the possible harm experienced by the clients due to the self-interest of the facilities: if the facilities form a Nash equilibrium, this is how bad it can be as compared to their optimum placement, as it would be created by an altruistic central authority. We also consider the Price of Stability, which looks at the same ratio but uses the *best* Nash equilibrium. This represents the cost increase experienced by the clients if the resulting solution is required to be an equilibrium but could be chosen in order to minimize client cost.

Properties of the optimum solution

We begin by considering the structure of the optimal facility placement which minimizes the total client cost. It is not difficult to see that any proportional spacing of facilities, as defined below, results in an optimum solution. Due to a lack of space, most of our proofs can be found in the full version of this paper.

Definition 1. *A* **proportional spacing** *is defined as follows. Given a list of stacks of facilities where the number of facilities in each stack is n_i, we can iterate through the list and place each stack at position:*

$$x_i = \frac{n_i}{2k} + \sum_{j=1}^{i-1} \frac{n_j}{k}.$$

[11] We will use the notation $|I|$ to refer to the length (size) of an interval I.

Proposition 1. *Any proportional spacing has cost* $\frac{1}{4k}$*. Moreover, no other solution has a smaller cost, so proportional spacings are optimal.*

Proportional spacings are easy to analyze: they have no bubbles (i.e., the set of clients using a stack is always a contiguous interval), each stack obtains a total number of clients equal to $u_i = n_i/k$, and thus each facility has utility exactly $1/k$. In particular, the following natural placements of facilities are optimal.

Corollary 1. *Placing all facilities in a stack of size* k *at position* $\frac{1}{2}$ *results in an optimum solution.*

Corollary 2. *Placing each facility in a separate stack, with the first at position* $\frac{1}{2k}$ *and the* i*'th at position* $\frac{i-1}{k} + \frac{1}{2k}$ *results in an optimum solution.*

The latter solution is exactly the optimum for the classic Hotelling-Downs model as well. The former solution is the simplest to analyze: it is simply all the facilities teaming up to form a giant mall in the middle of the interval, so the cost of a client located at x is exactly $|x - \frac{1}{2}|/k$.

Although the optimum solutions in this model are not difficult to analyze, they are *not necessarily* Nash equilibrium solutions. Nevertheless, we can establish the following claim, which shows the existence of Nash equilibria in contrast with the classic model.

Proposition 2. *The solution from Corollary 1 is a Nash equilibrium. Thus, Nash equilibria always exist, and the Price of Stability is 1.*

In fact, most of the proportional facility spacings are Nash equilibrium solutions. For example, if each stack is of size at least 2 (no facility is located by itself), then an optimal solution is a Nash equilibrium. On the other hand, the solution in Corollary 2 is not a Nash equilibrium, as the first and last facilities have incentive to move closer to the middle and thus obtain a larger share of the clients.[12]

Properties of Nash equilibria

While the best Nash equilibrium solutions are easy to analyze, since they are the same as optimum solutions, looking at the properties of general Nash equilibria becomes complex. The main difference compared to the classic model is the existence of *bubbles*. To define these formally, we first need to introduce the notion of core intervals:

Definition 2. Core intervals *are the maximal contiguous intervals* I_i *such that* $I_i \subseteq X$*,* $x_i \in I_i$ *and all clients* $x \in I_i$ *are using facility stack* i*.*

[12] Similarly to all Nash equilibria in the classic model, the first and last facilities can never be stacks of size 1 for an optimal solution to be a Nash equilibrium when $k > 1$. This is a necessary but not sufficient condition for optimal equilibrium.

In other words, core intervals are the sets of clients located next to a facility stack that use this stack. In the classic model, all clients belong to a core interval. In our model, however, client behavior results in bubbles, which are intervals of clients that do not belong to any core interval. To illustrate this, and show that this does, in fact, occur in Nash equilibrium solutions, consider the following simple example.

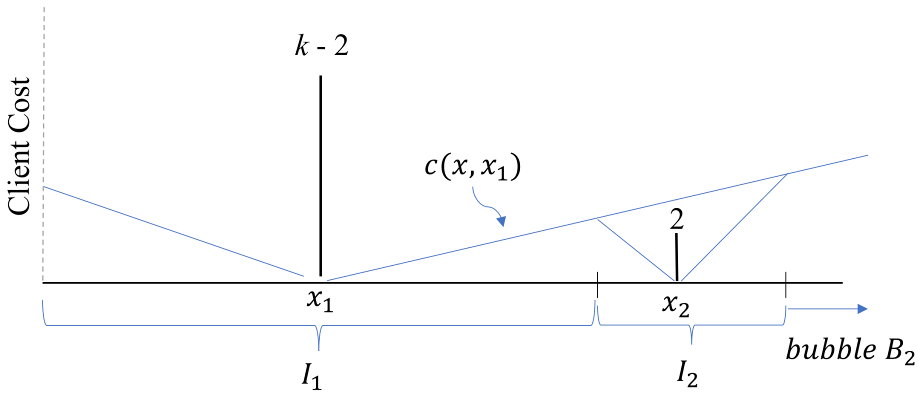

Fig. 1. A Nash equilibrium example with a bubble (Example 1). $k-2$ facilities are located at position x_1, and 2 facilities at location x_2. The blue lines show the client costs $c(x, x_1) = |x - x_1|/(k-2)$ to use facilities at x_1, and $c(x, x_2) = |x - x_2|/2$ to use facilities at x_2. The clients in interval I_2 use facilities at x_2, since that gives them the smallest cost. The clients both in intervals I_1 *and* in B_2 use facilities at x_1, since that has the smallest cost for them. Thus, the set B_2 forms a "bubble": a set of clients who use a facility stack such that they need to pass through another stack to get to it. In this figure, $u_1 = |I_1| + |B_2|$ and $u_2 = |I_2|$.

Example 1. **[Bubble Example]** Consider the following solution, as shown in Figure 1. For ease of presentation, let $|X| = 2k - 7 - \varepsilon$; everything can be scaled equivalently to $|X| = 1$. Position a stack of $n_1 = k - 2$ facilities at location $x_1 = k - 6$, and a stack of $n_2 = 2$ facilities at location $x_2 = 2k - 10$. Now consider what stack the clients located to the right of x_2 end up using. It is not difficult to verify that the clients in the interval $[x_2, x_2 + 2]$ use the stack at x_2, but the clients in the interval $[x_2 + 2, x_2 + 3 - \varepsilon]$ use the stack at x_1 again.[13] Thus, the clients in the latter interval form a bubble: they use the stack which is farther away from them than x_2, and are not part of x_1's core interval. The clients using the stack at x_1 consist of many clients next to it (its core interval), as well as the clients on the other side of x_2's core interval. Moreover, we can

[13] It does not matter what facility clients at the boundary between intervals (i.e., a location y such that $c(y) = c(y, x_i) = c(y, x_j)$ where $i \neq j$) use, as they only contribute an infinitesimal amount of utility.

show that this example is a Nash equilibrium as long as k is large enough, by considering all possible types of deviations that individual facilities can perform.

Because of the existence of bubbles, we cannot use standard techniques to analyze the Price of Anarchy for our model. Moreover, even if there were no bubbles, the equilibrium solutions have different properties than those of the classic model. It is not difficult to verify, for example, that the solution in Corollary 1 remains a Nash equilibrium even if the stack is not located at exactly $\frac{1}{2}$, but instead deviates from $\frac{1}{2}$ by a small distance δ. Because of this, we must develop a new approach for showing our Price of Anarchy bounds, as we do in the following section.

3 Price of Anarchy

In this section, we give an outline of our arguments and techniques for establishing our bound on the Price of Anarchy, and thus on the quality of all Nash equilibrium solutions in our model. All detailed proofs can be found in the full version of this paper. We begin with the following very simple property.

Proposition 3. *In a Nash equilibrium, the utility of any facility is at most a factor of 2 away from the utility of any other. More precisely, given stacks of size n_i and n_j, it must be that*

$$\frac{u_i}{n_i} \geq \frac{u_j}{n_j + 1} \geq \frac{1}{2} \cdot \frac{u_j}{n_j}$$

Proof. This is simply because a facility from the first stack can deviate by moving on top of the second stack j (which would only increase the amount u_j), and get utility at least $\frac{u_j}{n_j+1}$. However, a facility should not be able to improve its utility in a Nash equilibrium. □

If there were no bubbles and each stack had to be located in the middle of its core interval, the above proposition could be used to form a nice bound on the Price of Anarchy. If u_i clients use a stack of size n_i, we could simply say that the total cost of these clients equals

$$\int_{x=0}^{u_i} \frac{|x - \frac{u_i}{2}|}{n_i} \, dx = \frac{u_i^2}{(4n_i)}$$

In our model, however, just because we know approximately how many clients use a stack does not mean that we have a good bound on the cost that these clients experience. These clients may be far away from the stack due to being in a bubble, or the stack could be off-center in its core interval, which increases the client cost. Because of this, our main goal in this section becomes that of proving a limit on how much bubbles can affect things, and in particular proving a limit on the possible sizes of the bubbles.

First, let us prove the following lemma, which will allow us to define bubbles more precisely.

Lemma 1. *The cost function* $c(x) = \min_i\{c(x, x_i)\}$ *between two neighboring core intervals is either:*

1. *monotonically increasing*
2. *monotonically decreasing*
3. *monotonically increasing and then decreasing*

Proof. This follows from the definition of the cost function, $c(x)$, because it is a minimum over a set of strictly monotone linear functions. Assume that the cost function between two core intervals does not follow one of the above patterns. This means that there must be a section of the cost function between these two points that is convex. Since the cost function is the minimum of linear functions, there must be a stack of facilities between the two core intervals in order for the cost function to be lower than it was previously if it will increase again before the second core interval. This implies that there must be another core interval in between our two, giving a contradiction. □

Definition 3. *A* **bubble** *is a maximal contiguous interval that is disjoint from all core intervals such that the bubble has either a monotonically increasing or a monotonically decreasing cost function. Note that the cost function in a bubble can correspond to multiple stacks.*

Definition 4. *An* **up bubble** *is a bubble where the cost function is monotonically increasing.*

Definition 5. *A* **down bubble** *is a bubble where the cost function is monotonically decreasing.*

Lemma 2. *Consider the following assignment of bubbles to adjacent core intervals. Assign up bubbles to the core intervals directly to the left of the bubble and assign down bubbles to the core intervals directly to the right of the bubble. Then, this assignment is an injective mapping (every core interval has at most one bubble assigned to it).*

Proof Sketch. To prove this, we show that if there is an up bubble directly to the right of a core interval, then the cost function $c(x)$ at the endpoint of the core interval adjacent to the bubble must be strictly larger than at the other endpoint of the core interval. Together with a similar result for down bubbles, this immediately implies that there cannot be both an up bubble directly to the right of a core interval and a down bubble directly to the left of it. □

Let B_i be defined as the interval of the bubble assigned to core interval I_i. Note that the clients in B_i do *not* use the facilities in I_i in this solution: they use some facilities farther away; I_i is simply the core interval next to the bubble B_i. The key component of our proof is the fact that the size of B_i cannot be too large compared to I_i. Consider Example 1. The bubble on the right has size $1-\varepsilon$, and would be assigned to the core interval I_2, which has size of approximately 4: 2 to the right of x_2 and a bit less than 2 to the left of x_2. Thus, the size of the bubble is only about 1/4 of the size of the core interval next to it. In the following, we prove that this is, in fact, the worst case, and no larger bubbles are possible. We do this through a series of lemmas: see the full version of this paper for detailed proofs.

Lemma 3. *If a stack i is assigned an up (down) bubble B_i, then stack i does not receive utility from any bubbles to its right (left) hand side.*

Lemma 4. *Single stacks (i.e., stacks with $n_i = 1$) cannot receive clients or utility from a bubble.*

Lemma 5. *If the current solution is a Nash equilibrium, then for each B_i we have that $n_i \geq 2$. In other words, a bubble will never be assigned to a single stack. Although some single stacks may exist in the solution, they will not be assigned any bubbles in our mapping.*

Using the above lemmas, we are able to prove the key property which makes our bound on the Price of Anarchy possible:

Lemma 6. *If the current solution is a Nash equilibrium, then for each bubble B_i we have that $|B_i| \leq \frac{1}{4}u_i$.*

The proof of this result is somewhat complex and requires a careful analysis of how much facilities from I_i would gain by deviating to somewhere in the middle of bubble B_i (which uses the fact that the stack i itself does not have any bubbles as derived from Lemma 3), together with the fact that $n_i \geq 2$ because of the above lemmas. Once we have established that bubbles cannot be too large, we can proceed to bound the costs of the clients at Nash equilibrium.

Consider an arbitrary Nash equilibrium solution called S, with stack locations x_i and stack sizes n_i. Let c_S represent the total cost of solution S, and let N be the set of all stacks for the solution S. Also, let c_O be the total cost of the optimal solution; we know by Proposition 1 that $c_O = \frac{1}{4k}$. To compare c_S and c_O, we compare them to several different intermediate quantities, as follows.

Definition 6. *Let c_1 be defined as:*

$$c_1 = \sum_{i \in N} \frac{u_i^2}{2n_i}$$

Where each u_i and n_i represent the same utilities and stack sizes as in solution S.

Definition 7. *Let c_2 be defined as:*

$$c_2 = \sum_{i \in N} \left[\int_{x \in I_i} c(x, x_i)\ dx + \int_{x \in B_i} c(x, x_i)\ dx \right]$$

Note that c_1 and c_2 do not represent values of any actual solutions; they are only convenient values for comparing the total cost of equilibrium and optimum solutions. For intuition on the quantity c_1, consider a new solution in which there are no bubbles, but each stack still receives the same amount of utility u_i as it does in the equilibrium solution S. Then, c_1 is an upper bound on the maximum possible cost of this solution, where all facilities are at the very start of their intervals of size u_i, and thus the cost of the clients in this interval is $\frac{u_i^2}{2n_i}$. For intuition on the quantity c_2, note that it is the total cost of S if clients in bubbles instead choose to utilize the stack that the bubble is assigned to by our mapping, instead of their preferred stack, even though their cost would be greater.

Lemma 7. $c_S \leq c_2$

Proof. The cost of all clients in core intervals remains the same between both total costs, thus the only difference is in the cost of the clients in bubbles. For any bubble B_i, the cost $c(x)$ of any given client $x \in B_i$ must be less than $c(x, x_i)$ or else that client would have chosen stack i which we know is not the case due to Lemma 3. Thus, the value of c_2 must be larger since it is increasing the cost of all clients in bubbles. □

Lemma 8. $c_2 \leq \frac{25}{16} c_1$

Proof. We will compare the values within both summations for any given i. Let $c_{2,i} = \int_{x \in I_i} c(x, x_i)\ dx + \int_{x \in B_i} c(x, x_i)\ dx$ and let $c_{1,i} = \frac{u_i^2}{2n_i}$. We will therefore show that $c_{2,i} \leq \frac{25}{16} c_{1,i}$. W.L.O.G. assume that all B_i are up bubbles. Here, $L(I_i)$ and $R(I_i)$ are defined as the leftmost and rightmost points of interval I_i.

$$
\begin{aligned}
c_{2,i} &= \int_{x \in I_i} \frac{|x - x_i|}{n_i}\ dx + \int_{x \in B_i} \frac{|x - x_i|}{n_i}\ dx \\
&= \frac{(x_i - L(I_i))^2}{2n_i} + \frac{(R(I_i) - x_i + |B_i|)^2}{2n_i} \\
&\leq \frac{(x_i - L(I_i) + R(I_i) - x_i + |B_i|)^2}{2n_i} \\
&= \frac{(R(I_i) - L(I_i) + |B_i|)^2}{2n_i} \\
&= \frac{(|I_i| + |B_i|)^2}{2n_i} && \text{(by the definition of } I_i\text{)} \\
&\leq \frac{(u_i + |B_i|)^2}{2n_i} && \text{(by the definition of } u_i\text{)} \\
&\leq \frac{(u_i + \frac{u_i}{4})^2}{2n_i} && \text{(by Lemma 6)} \\
&= \frac{25}{16} \times \frac{u_i^2}{2n_i}
\end{aligned}
$$

$$= \frac{25}{16} c_{1,i} \qquad \text{(by the definition of } c_{1,i}\text{)}$$

Since this is true for all $i \in N$, we have shown that $c_2 \leq \frac{25}{16} c_1$. □

Lemma 9. $c_1 \leq \frac{9}{4} c_O$

Proof Sketch. Using Proposition 3, we know that u_i/n_i cannot be too different from each other, while the optimum solution is the total client cost when u_i/n_i are perfectly balanced, i.e., each equals $1/k$. This proof does not use game theory or anything beyond algebra and calculus and simply bounds the largest value that c_1 could have if u_i/n_i are not too different and if $\sum_{i \in N} u_i = 1$. This amount is at most $\frac{9}{16k}$, and since $c_O = \frac{1}{4k}$ we obtain a bound of $\frac{9}{4}$. □

Theorem 1. *The Price of Anarchy in our setting is at most* $\frac{225}{64} \approx 3.516$.

Proof.

$$\begin{aligned} c_S &\leq c_2 && \text{(by Lemma 7)} \\ &\leq \frac{25}{16} c_1 && \text{(by Lemma 8)} \\ &\leq \frac{25}{16} \times \frac{9}{4} c_O && \text{(by Lemma 9)} \\ &= \frac{225}{64} c_O \end{aligned}$$

Since our choice of S was arbitrary, this bound holds for all pure Nash equilibria. □

4 Conclusion and Future Directions

In this paper, we introduced and analyzed a new variation of the Hotelling-Downs model, in which clients do not simply use the closest facility, but instead prefer sites with many facilities. Because of this, interesting interactions between facilities begin occurring: the facilities want to be far away from other facilities so that they get more customers to themselves, but they also want to be close to other facilities since this might increase the amount of customers they get (by creating a "destination to visit"). This can also apply to social choice settings, where political candidates decide whether they want to position themselves as different from all others (to make themselves unique) or join an existing party (to "ride on their coat-tails"). Despite Nash equilibria having quite different structure in our setting than in the classic version, we showed that a good Nash equilibrium always exists and that the Price of Anarchy is bounded, thus establishing that Nash equilibrium solutions cannot be very bad as compared with the optimum facility placement.

Our work is only a first step in the study of what we call "The Mall Effect", however, and many open questions and future directions remain. The most immediate one is looking at other cost functions $c(x, x_i)$ which increase with the distance and decrease with the number of facilities at x_i. We believe that the function $c(x, x_i) = |x - x_i|/n_i$ used in this paper is natural, but many other cost functions make sense as well. For example, what changes if we consider $c(x, x_i) = |x - x_i|/(n_i)^p$ for some power p? For $p < 1$, the properties of this model become similar to the classic one, and a Nash equilibrium does not always exist. For $p > 1$, Nash equilibria always exist, and, in fact, this should help with the Price of Anarchy as well since this gives further incentive for facilities to form large stacks. Fully analyzing the Price of Anarchy in this model remains future work, however. A different type of cost function may give incentive for clients to go to facilities where many other facilities are *nearby*, instead of at exactly the same location.

Another promising direction involves looking at different metric spaces. Although we focused on X being a one-dimensional interval as in most existing work on this subject, all our results also hold if X were a circle instead. It would be interesting to consider more general metric spaces, such as two-dimensional spaces, or facilities which can be located on a graph such as in [19] and [25]. It would also be interesting to consider settings where the set of clients is not continuous, but instead there is a discrete set of client locations, as well as a discrete set of possible facility locations, as in [26]. Finally, it would be interesting to combine our client behavior with some of the other variations in the existing literature, such as putting limits on how far clients are willing to travel in order to use a facility: if all facilities are too far, they don't use any facility at all [17, 30].

Our work initiates the study of facility synergy, thereby shedding light on potential social processes that can lead to the creation of malls, shopping centers and downtowns while opening up a new avenue for future research.

Acknowledgments. This work was partially supported by National Science Foundation award CCF-2006286.

References

1. Ahn, H.K., Cheng, S.W., Cheong, O., Golin, M., Oostrum, R.: Competitive facility location: The voronoi game. Theor. Comput. Sci. **310**, 457–467 (01 2004). https://doi.org/10.1016/j.tcs.2003.09.004
2. Anderson, S.P., Palma, A.d., Thisse, J.F.: Discrete choice theory of product differentiation. MIT Press (1992). https://doi.org/10.7551/mitpress/2450.001.0001
3. Anshelevich, E., Dasgupta, A., Kleinberg, J., Tardos, É., Wexler, T., Roughgarden, T.: The price of stability for network design with fair cost allocation. SIAM Journal on Computing **38**(4), 1602–1623 (2008). https://doi.org/10.1137/070680096
4. Aoyagi, M., Okabe, A.: Spatial competition of firms in a two-dimensional bounded market. Regional Science and Urban Economics **23**(2), 259–289 (1993). https://doi.org/10.1016/0166-0462(93)90006-Z
5. Ben-Porat, O., Tennenholtz, M.: Shapley facility location games. In: Web and Internet Economics: 13th International Conference, WINE 2017, Bangalore, India, December 17–20, 2017, Proceedings 13. pp. 58–73. Springer (2017). https://doi.org/10.1007/978-3-319-71924-5_5

6. Ben-Porat, O., Tennenholtz, M.: Multiunit facility location games. Mathematics of Operations Research **44**(3), 865–889 (2019). https://doi.org/10.1287/moor.2018.0948, https://www.jstor.org/stable/27287863
7. Bhaskar, U., Pyne, S.: Equilibrium computation in the hotelling-downs model of spatial competition. arXiv preprint arXiv:2412.12523 (2024). https://doi.org/10.48550/arXiv.2412.12523
8. Blum, A., Hajiaghayi, M., Ligett, K., Roth, A.: Regret minimization and the price of total anarchy. In: Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing. p. 373–382. STOC '08, Association for Computing Machinery, New York, NY, USA (2008). https://doi.org/10.1145/1374376.1374430
9. 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
10. Cohen, A., Peleg, D.: Hotelling games with random tolerance intervals. In: Web and Internet Economics: 15th International Conference, WINE 2019, New York, NY, USA, December 10–12, 2019, Proceedings. pp. 114–128. Springer (2019). https://doi.org/10.1007/978-3-030-35389-6_9
11. Downs, A.: An economic theory of political action in a democracy. Journal of Political Economy **65**(2), 135–150 (1957), http://www.jstor.org/stable/1827369
12. Drezner, Z., Eiselt, H.: Competitive location models: A review. European Journal of Operational Research **316**(1), 5–18 (2024). https://doi.org/10.1016/j.ejor.2023.10.030
13. Dubey, P.: Inefficiency of nash equilibria. Mathematics of Operations Research **11**(1), 1–8 (1986). https://doi.org/10.1287/moor.11.1.1, https://www.jstor.org/stable/3690047
14. Eaton, B.C., Lipsey, R.G.: The principle of minimum differentiation reconsidered: Some new developments in the theory of spatial competition. The Review of Economic Studies **42**(1), 27–49 (1975). https://doi.org/10.2307/2296817, http://www.jstor.org/stable/2296817
15. Economides, N.: Minimal and maximal product differentiation in hotelling's duopoly. Economics Letters **21**(1), 67–71 (1986). https://doi.org/10.1016/0165-1765(86)90124-2
16. Eiselt, H.A., Laporte, G., Thisse, J.F.: Competitive location models: A framework and bibliography. Transportation Science **27**(1), 44–54 (1993). https://doi.org/10.1287/trsc.27.1.44, http://www.jstor.org/stable/25768567
17. Feldman, M., Fiat, A., Obraztsova, S.: Variations on the hotelling-downs model. Proceedings of the AAAI Conference on Artificial Intelligence **30**(1) (Feb 2016). https://doi.org/10.1609/aaai.v30i1.10054
18. Feldotto, M., Lenzner, P., Molitor, L., Skopalik, A.: From hotelling to load balancing: Approximation and the principle of minimum differentiation. In: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems. p. 1949–1951. AAMAS '19, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2019). https://doi.org/10.48550/arXiv.1903.04265
19. Fournier, G., Scarsini, M.: Location games on networks: Existence and efficiency of equilibria. Mathematics of Operations Research **44**(1), 212–235 (2019). https://doi.org/10.1287/moor.2017.0921
20. Fournier, G., Schröder, M.: Popularity in location games. arXiv preprint arXiv:2401.07732 (2024). https://doi.org/10.48550/arXiv.2401.07732

21. Graitson, D.: Spatial competition à la hotelling: A selective survey. The Journal of Industrial Economics **31**(1/2), 11–25 (1982). https://doi.org/10.2307/2098001
22. Hotelling, H.: Stability in competition. The Economic Journal **39**(153), 41–57 (1929). https://doi.org/10.2307/2224214
23. Koutsoupias, E., Papadimitriou, C.: Worst-case equilibria. Comput. Sci. Rev. **3**(2), 65–69 (may 2009). https://doi.org/10.1016/j.cosrev.2009.04.003
24. Krogmann, S., Lenzner, P., Molitor, L., Skopalik, A.: Two-stage facility location games with strategic clients and facilities. In: Zhou, Z.H. (ed.) Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21). pp. 292–298 (8 2021). https://doi.org/10.24963/ijcai.2021/41
25. Mavronicolas, M., Monien, B., Papadopoulou, V.G., Schoppmann, F.: Voronoi games on cycle graphs. In: Proceedings of the 33rd International Symposium on Mathematical Foundations of Computer Science. p. 503–514. MFCS '08, Springer-Verlag, Berlin, Heidelberg (2008). https://doi.org/10.1007/978-3-540-85238-4_41
26. Núñez, M., Scarsini, M.: Competing over a finite number of locations. Economic Theory Bulletin **4**, 125–136 (2016). https://doi.org/10.1007/s40505-015-0068-6
27. Peters, H., Schröder, M., Vermeulen, D.: Hotelling's location model with negative network externalities. International Journal of Game Theory **47**, 811–837 (2018). https://doi.org/10.1007/s00182-018-0615-0
28. ReVelle, C., Eiselt, H.: Location analysis: A synthesis and survey. European Journal of Operational Research **165**(1), 1–19 (2005). https://doi.org/10.1016/j.ejor.2003.11.032
29. Salop, S.C.: Monopolistic competition with outside goods. The Bell Journal of Economics **10**(1), 141–156 (1979). https://doi.org/10.2307/3003323
30. Shen, W., Wang, Z.: Hotelling-downs model with limited attraction. In: Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems. pp. 660–668 (2017). https://doi.org/10.48550/arXiv.1611.05959