Efficient and Equitable Deployment of Mobile Vaccine Distribution Centers

Da Qi Chen[1], **Ann Li**[1], **George Z. Li**[2], **Madhav Marathe**[1], **Aravind Srinivasan**[2], **Leonidas Tsepenekas**[2], and **Anil Vullikanti**[1]

[1]Biocomplexity Institute & Initiative, University of Virginia
[2]University of Maryland
wny7gj@virginia.edu, al7gc@virginia.edu, gzli929@gmail.com, mvm7hz@virginia.edu, srin01umd@gmail.com, leo.tsepe92@gmail.com, asv9v@virginia.edu

Abstract

Vaccines have proven to be extremely effective in preventing the spread of COVID-19 and potentially ending the pandemic. Lack of access caused many people not getting vaccinated early, so states such as Virginia deployed *mobile* vaccination sites in order to distribute vaccines across the state. Here we study the problem of deciding where these facilities should be placed and moved over time in order to minimize the distance each person needs to travel in order to be vaccinated. Traditional facility location models for this problem fail to incorporate the fact that our facilities are mobile (i.e., they can move over time). To this end, we instead model vaccine distribution as the Dynamic k-Supplier problem and give the first approximation algorithms for this problem. We then run extensive simulations on real world datasets to show the efficacy of our methods. In particular, we find that natural baselines for Dynamic k-Supplier cannot take advantage of the mobility of the facilities, and perform worse than non-mobile k-Supplier algorithms.

1 Introduction

Vaccines have played a vital role in reducing the negative health effects of COVID-19 and continue to be the best strategy to end the pandemic. Despite the effectiveness of vaccines, it remains difficult to vaccinate all eligible individuals in the population. There are three primary reasons: (i) lack of accessibility of vaccines; (ii) hesitancy due to perceived harm or mistrust; and (iii) strategic behavior or misinformation—see [Yan, 2021]. This paper focuses on addressing elements of the *accessibility* problem, which will be important in tackling future epidemics and remains a challenge in several low- and middle-income countries; see [Acharya *et al.*, 2021; Bayati *et al.*, 2022; Kim *et al.*, 2021; Wouters *et al.*, 2021] and [Wel,]. In the U.S., states such as Virginia have funded *mobile* vaccination sites to distribute vaccines throughout the state [Mehrab *et al.*, 2022; Shukla *et al.*, 2022a]. Here, we study how these mobile vaccination sites should be *dynamically deployed* through a county to reduce the amount people need to travel for vaccines. **Our methods take into account issues of equitable access by placing these mobile vaccine distribution sites strategically.** Specifically, lower income groups, elders and other groups have often found it hard to get vaccines due to the time it might take to reach a facility and waiting in the queue to get the vaccine. In an important article [Lu *et al.*, 2022], the authors point out this issue in Boston. Quoting them: *However, the state has not set up mass vaccination sites in some of the communities most disproportionately impacted by COVID-19 in Boston, like East Boston, Chelsea, or Hyde Park. In Chelsea — a city just outside of Boston — only 7% of Latino residents have received vaccines, even though Latino people are 68% of the population. Many individuals in these neighborhoods cannot easily travel to the mass vaccination sites due to disability, work schedules, or lack of transportation.* We formulate this as a *Dynamic Priority k-Supplier* problem wherein individuals can value each facility not just by distance but access to modes of transportation. This allows us to capture individual-level accessibility constraints and this leads to more equitable solutions. *These extensions have not been considered in prior work.*

The problem of deploying vaccine distribution sites is a facility location problem, and is classically formulated as the k-supplier problem and its variants, e.g., [Brubach *et al.*, 2021; Li *et al.*, 2022]. In this problem, our goal is to choose k facility locations in order to minimize the maximum distance between a client and their closest facility. However, the k-supplier problem fails to model a key component of our vaccine distribution application: our facilities can move over time since they are *mobile*. We can use this to our advantage to significantly reduce the distance agents need to travel in order to get vaccinated. Our contributions are:

- We formulate the problem of placing mobile vaccine distribution sites as the Dynamic k-Supplier problem that incorporates mobility considerations into the problem by developing a multi-timestep version of k-supplier. We also extend the problem formulation to incorporate practical considerations such as outliers and fairness. Since [Deng *et al.*, 2022] showed that it is NP-Hard to obtain any non-trivial approximation algorithm for the Dynamic k-Supplier problem, we turn to bicriteria approximation algorithms.
- We design two bicriteria approximation algorithms. The first algorithm is based on the Set Cover problem, and

obtains the optimal service cost but violates the budget k by a logarithmic factor. This algorithm gives rise to a simple heuristic, which performs extremely well in practice (see Section 6). Unfortunately, the algorithm must violate the number of facilities used by a logarithmic factor, which is theoretically unsatisfying. Our second algorithm is based on a network-flow approach, and obtains a constant approximation to the service cost but violates the movement constraint by a constant factor and requires a (realistic) density assumption on the candidate facility locations. In the Appendix, we detail how both of our algorithms can be extended to consider outliers and our cover-based algorithm can incorporate equality considerations discussed in the introduction.

- We evaluate our algorithms for populations of a city and a county in Virginia. Our primary goal was to understand how much the mobility aspect of the facility location model can improve the accessibility of vaccines. We find that neither our flow-based algorithm nor a natural heuristic exploits the additional flexibility introduced by having mobile facilities. In particular, the objective values of both of these algorithms are comparable to that of a standard k-supplier algorithm where facilities don't move. In contrast, our cover-based algorithm performs significantly better than all baselines. In particular, this illustrates the challenges in improving accessibility using the mobile facilities. We additionally explore the tradeoffs between the number of facilities, the movement constraint, and the proportion of the population served. Furthermore, we investigate a priority-based and a limited capacity variant, modelling different aspect of equity and fairness in allocation. All these results help inform policymakers concerning decisions about vaccine distribution.

Even though our results are stated in the context of vaccine distribution, they also have broad applications elsewhere, e.g., delivering other healthcare resources (such as cancer screening units and blood banks), and library outreach programs [Raghavan *et al.*, 2019]. More generally, Dynamic k-Supplier was introduced in the context of clustering dynamic points without moving the cluster centers too much, e.g., [Deng *et al.*, 2022]; our results apply there as well.

2 Preliminaries

We now formalize the Dynamic k-Supplier problem. Let $[T]$ denote the set $\{1, 2, \ldots, T\}$. We assume that our distances are in a metric space $\mathcal{X}$, characterized by a distance function $d : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}_{\geq 0}$. We assume a time discretization into T time steps and that for each time step $t \in [T]$, we are given a set of locations $\mathcal{F}^t \subseteq \mathcal{X}$ where we can place our k available mobile facilities. At each time t, we are also given a set of clients: our goal is to serve a set of clients $\mathcal{C} \subseteq \mathcal{X}$ over the course of the T time steps. Of course, each client only needs to be vaccinated once so we don't need to serve each client at every time step. Instead, we assume each client/person chooses a time step $t \in [T]$ when they wish to be vaccinated. Let $\mathcal{C}^t \subseteq \mathcal{C}$ be the given set of clients who wish to be vaccinated at time step $t \in [T]$; we will only need to serve $\mathcal{C}^t$ at time step $t \in [T]$ with our mobile facilities. However, our facilities cannot move arbitrarily far between two consecutive time steps. We take as input a movement constraint M and require that no facility move more than distance M between consecutive time steps. Given these constraints, our goal is to find k locations $S_t \subseteq \mathcal{F}^t$ for each t so that we minimize the maximum service distance. Formally:

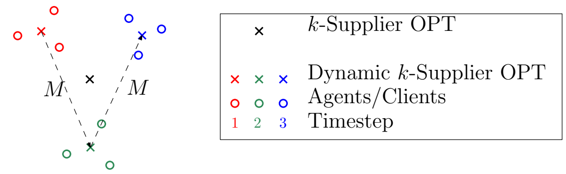

Figure 1: The example illustrates an instance of our vaccination problem where the optimal solutions of Dynamic k-Supplier and k-Supplier differ significantly. The different time steps are represented by different colors (red, green, and blue) and each colored circle represents a client. Given a single facility, the black "x" is the optimal k-Supplier solution and has large service cost. In contrast, allowing movement of M significantly reduces the service cost by travelling to the red, green, and blue "x" locations.

Problem 2.1. *Given a metric space $\mathcal{X}$, for each time step $t \leq T$, we are provided a set of clients $\mathcal{C}^t$, a set of potential facility locations $\mathcal{F}^t$, and an upper bound of movement M. For each time step t, find a sequence of k facilities $S_t = (s_{1,t}, s_{2,t}, \ldots, s_{k,t})$ each lying in $\mathcal{F}^t$, such that $d(s_{i,t}, s_{i,t+1}) \leq M$ for $1 \leq i \leq k$ and $1 \leq t \leq T-1$ to minimize the maximum service distance $R := \max_{t\in[T]} \max_{j\in\mathcal{C}^t} d(j, S_t)$, where $d(j, S_t) = \min_{i\in S_t} d(j, i)$. See Figure 1 for an example.*

To make our problem formulation more concrete, we illustrate an example of how it can be used in practice for deploying vaccine distribution sites. We assume that the government has a website on which people can register to be vaccinated at a specific hour (e.g., 2–3PM). For each day, let us discretize the time (8AM–8PM) into $T = 12$ hours; each mobile facility can use the first 50 minutes in the hour to vaccinate clients and the final 10 minutes to move to the next location. Now, let each person in the county choose a day and time to get vaccinated via the website/mobile app and let them input (approximately) where they'll be at that time; we can use this information to construct the client sets $\mathcal{C}^t$ for each time step $t \in [T]$. After obtaining the set of candidate facility locations $\mathcal{F}$ (for example, from the government), we can combine the facility and client locations and define the underlying metric to be the travel times between two locations (for example, given by Google maps). Now that we have all the components, we can apply one of our algorithms for the Dynamic k-Supplier problem for each day and obtain a near-optimal schedule for how each facility should move to minimize the maximum time any person needs to travel for their vaccine.

There are also many practical considerations for vaccine distribution that should also be modeled. For instance, some potentially-adversarially inserted clients may ask to be served at a location very far from everyone else. Since our goal is to maximize accessibility, we can regard these points as outliers and ignore them when choosing facility locations. Another practical consideration is that different people travel at different speeds (e.g., some people can drive but others need to bike or use public transit) so assuming the same metric for everyone leads to *unfairness*. A way to counteract this is to use a weighted objective; by increasing the weight of a client, we can increase their priority and reduce their service distance. We call these two variants Dynamic Robust k-Supplier and Dynamic Priority k-Supplier, respectively, and defer the formal definitions and their algorithms to the Appendix.

3 Related Work

Due to its broad applications in a number of domains, there has been much work on various facility location-type problems over the years. Facility-location-type optimization problems have also been studied in diverse contexts in the agents community: see [Kiekintveld *et al.*, 2009; Fang *et al.*, 2015; Chen *et al.*, 2022; Wada *et al.*, 2018; Moujahed *et al.*, 2006] and the references therein. The problem considered in our paper is a generalization of the k-center problem for which there exists simple 2-approximation algorithms [Hochbaum and Shmoys, 1985; Gonzalez, 1985] that are best-possible unless P=NP [Hochbaum and Shmoys, 1986]. Many generalizations of k-center have been studied: e.g., fair and stochastic versions [Kleindessner *et al.*, 2019; Brubach *et al.*, 2020; Chakrabarti *et al.*, 2022; Jia *et al.*, 2022; Anthony *et al.*, 2010; Huang and Li, 2017], but the work most closely related to ours is that of [Deng *et al.*, 2022] which introduced two dynamic clustering problems including the Dynamic k-Supplier problem that we study. They give a 3-approximation algorithm when $T = 2$ and show that the problem is NP-Hard to approximate to any factor when $T \geq 3$. Unfortunately, they were unable to give any algorithms when $T \geq 3$: our work is the first set of positive results in this more interesting regime where we give two bi-criteria approximation algorithms for the problem.

More broadly, there has been much recent and concurrent work on using facility location problems for improved vaccine distribution due to the pandemic. For example, [Bertsimas *et al.*, 2022; Bravo *et al.*, 2022; Shukla *et al.*, 2022b; Rader *et al.*, 2021; Roy *et al.*, 2021; Raghavan *et al.*, 2019; Shukla *et al.*, 2022a; Ou *et al.*, 2022; Nair *et al.*, 2022] used various facility location models to study how different policies/interventions will effect the pandemic via extensive large scale simulations. On the more theoretical side, [Li *et al.*, 2022] introduce a facility location problem which explicitly models the mobility of the clients by representing them as the set of points in the metric space they travel to during a day and develop two approximation algorithms for it. Our model is more realistic for vaccine distribution since we don't need to know each client's entire daily travel information; as a result, it can be more easily implemented. Most closely related is the work of [Mehrab *et al.*, 2022] which studied our current problem from a practical perspective. They give data-driven heuristics for the problem using real-time mobility data of the clients in order to decide the placement of the mobile vaccination sites. Overall, our work is largely complementary to the current literature, tackling a new theoretical problem motivated by vaccine distribution.

4 Algorithm via Set Cover

In this section, we give our bicriteria algorithm that violates the budget constraint by a factor of $H_n = \sum_{i=1}^{n} \frac{1}{i} \leq \ln n+1$, where n is the number of clients.

As in most k-supplier algorithms, we can assume (without loss of generality, via bisection search) that we know the optimal radius R^* and we want to choose k facilities at each time step to "cover" every client using the balls of radius R^*. Instead of fixing our budget k and trying to minimize the radius R, we consider the reverse problem where we fix the radius at R^* and minimize the number of facilities we need to place in order to cover every client. The primary observation is that this problem can be formulated as a Set Cover problem. The universe is the set of clients and each possible path a facility can take throughout the T time steps represents a set, where the set covers the clients which are within radius R^* of the locations in the path. We know by definition of R^* that the k paths/sets suffice to cover everyone. Since there exists a greedy H_n-approximation algorithm for Set Cover (where $H_n = \sum_{i=1}^{n} \frac{1}{i} \leq \ln n + 1$ is the n^{th} harmonic number), this is an algorithm that outputs at most $k \cdot H_n$ facilities such that each client needs to travel at most R^*.

We will now formalize the above intuition; we first define the Set Cover instance for a given radius R. Let a path p be represented by a sequence of T facilities $(p_1 \; \dots \; p_T)$, where each $p_t \in \mathcal{F}^t$; we say a path p is feasible if $d(p_t, p_{t+1}) \leq M$ for each $t \in [T-1]$. For each facility $i \in \mathcal{F}_t$, let $\mathcal{B}_R^t(i) = \{j \in \mathcal{C}_t : d(i,j) \leq R\}$ denote the ball of clients within radius R of i; then the set of clients a path p covers is exactly $\mathcal{B}_R(p) = \bigcup_{t\in[T]} \mathcal{B}_R^t(p_t)$. Given these definitions, our set system $\mathcal{S}_R$ will have a set $\mathcal{B}_R(p) \subseteq \mathcal{C}$ for each feasible path p and the goal is to choose the fewest sets/paths to cover the entire universe $\mathcal{U} := \mathcal{C}$. Now that the Set Cover instance has been formalized, we give our algorithm in Algorithm 1.

Algorithm 1 COVER

1: **Binary Search** R on $\{d(i,j) : i \in \mathcal{F}^t, j \in \mathcal{C}^t, t \in [T]\}$:
2:     Use R to create the Set Cover instance $(\mathcal{U}, \mathcal{S}_R)$.
3:     Let $\mathcal{U}_0 \leftarrow \mathcal{U}$, $\text{SOL}_R = \emptyset$, and $i \leftarrow 0$.
4:     **while** $\mathcal{U}_i \neq \emptyset$ and $i < k \cdot H_n$ **do**
5:         Add set $S_i = \arg\max_{S \in \mathcal{S}_R} |S_i \cap \mathcal{U}_i|$ to SOL_R.
6:         $\mathcal{U}_{i+1} \leftarrow \mathcal{U}_i - S_i$, $i \leftarrow i+1$
7:     **end while**
8:     **if** $\mathcal{U}_i \neq \emptyset$, increase R **else**, decrease R.
9: Output SOL_R for minimum R such that SOL_R covers $\mathcal{C}$.

Efficient implementation. Note that it isn't immediately obvious that Algorithm 1 runs in polynomial-time since the number of possible paths (and thus the number of sets) can be exponential in m. In particular, the naive implementation of the greedy approximation algorithm has exponential run-time since it's not obvious how to choose the set $S_i \in \mathcal{S}_R$ which covers the most additional elements (see line 5). We show that due to the structure of the set system, the greedy algorithm can still be implemented in polynomial time using a dynamic programming approach.

As mentioned above, we have a set system induced by paths and our goal is to choose the path which covers the most elements. Since there are exponentially many paths, our set system of paths is not explicitly given to us. Instead, we are given the facility locations $\mathcal{F}^t$ at each timestep and the movement constraint which implicitly defines the set system. We also have, for each facility location i, the corresponding set of clients which are within radius R of the location at time t, denoted $\mathcal{B}_R^t(i)$. Using this setup, we can now explain our dynamic programming approach.

Let $A[t,i]$ be the dynamic programming "matrix" indexed by a timestep $t \in [T]$ and a facility location $i \in \mathcal{F}^t$ which will store the maximum number of clients/elements a partial path $p[:t]$ which uses location $i \in \mathcal{F}^t$ can cover. Note that it is easy to change the algorithm so that it also stores the partial path which maximizes the coverage, but we omit this for cleaner presentation. To calculate $A[t,i]$, let $B_M^t(i)$ denote the set of facilities in timestep $t-1$ which can reach i (i.e., $B_M^t(i) = \{j \in \mathcal{F}^{t-1} : d(i,j) \le M\}$). Since we only need to consider locations in $B_M^t(i)$ for facility location $i \in \mathcal{F}^t$, we can calculate

$$A[t,i] = |\mathcal{B}_R^t(i)| + \max_{j \in B_M^t(i)} A[t-1,j] \tag{1}$$

with $A[1,i] = |\mathcal{B}_R^1(i)|$ as the base case. If we use this dynamic programming algorithm to implement line 5 of Algorithm 1, we have the following:

Theorem 4.1. *Algorithm 1 is a polynomial-time approximation algorithm which obtains the optimal service cost R^* while using at most $k \cdot H_n$ facilities.*

Remark 4.2. *In practice, it is not realistic to violate the budget constraint k by a factor of H_n. Fortunately, the budget violation comes from the rare worst case scenario when running the greedy set cover algorithm (see line 5 of Algorithm 1). In practice, researchers have observed that greedy set cover often obtains near-optimal solutions (see [Lan et al., 2007]). Thus, we can change line 5 of Algorithm 1 to instead require $i < k$. The resulting algorithm no longer violates the budget k but still serves as a high quality heuristic for the problem. We will use this heuristic in our experiments.*

5 Algorithm via Network Flow

We now provide another algorithm for Dynamic k-Supplier which obtains a constant approximation to the service cost but violates the movement constraint by a constant factor; in contrast to Algorithm 1, the number of facilities k is not violated. The technique generalizes the one used in [Deng *et al.*, 2022] by leveraging a density assumption on the candidate facility location to provide a constant bicriteria approximation for the Dynamic k-Supplier Problem when $T \ge 3$.

The outline of our algorithm is as follows. We binary search for the optimal radius R. Assuming we have guessed R correctly, we solve a system of linear constraints and obtain an optimal fractional solution to the problem. It remains to round the fractional solution into an integral one. To do this, we use known filtering techniques to aggregate the fractional facilities into balls of radius R and reduce our problem to rounding a fractional network flow while preserving some constraints. By rounding the fractional flow to an integral one, we obtain an integral approximate solution.

5.1 Linear Constraints

Assuming we know the optimal radius R, we construct an integer program as follows. For each timestep $t \in [T]$ and facility $i \in \mathcal{F}_t$, let $y_{i,t} \in \{0,1\}$ represent whether facility i is in the solution at time t. In addition, for each timestep $t \in [T]$ and each pair $i \in \mathcal{F}^t$ and $j \in \mathcal{C}^t$, let $x_{i,j}^t \in \{0,1\}$ represent whether j is served by i or not at time t. Let $\mathcal{B}_j^t \subseteq \mathcal{F}^t$ be the set of facilities at time t that is within distance R from client $j \in \mathcal{C}^t$. Finally, for each $t \in [T-1]$ and pairs $i_1 \in \mathcal{F}^t$ and $i_2 \in \mathcal{F}^{t+1}$, let $z_{i_1,i_2}^t \in \{0,1\}$ represent if facility i_1 will move to i_2 from time t to $t+1$. For simplicity of notation, let $\mathcal{E}^t \subseteq \mathcal{F}^t \times \mathcal{F}^{t+1}$ denote the set of possible facility movements at each timestep $t \in [T-1]$ (i.e., let $\mathcal{E}^t$ contain all pairs $(i_1,i_2) \in \mathcal{F}^t \times \mathcal{F}^{t+1}$ such that $d(i_1,i_2) > M$). Then, consider the following constraints.

$$\textstyle\sum_{i \in \mathcal{F}^t} y_i^t = k \qquad t \in [T] \tag{2}$$

$$\textstyle\sum_{i \in \mathcal{B}_j^t} x_{i,j}^t = 1 \qquad t \in [T], j \in \mathcal{C}^t \tag{3}$$

$$x_{i,j}^t \le y_i^t \qquad t \in [T], i \in \mathcal{F}^t, j \in \mathcal{C}^t \tag{4}$$

$$\textstyle\sum_{i_2} z_{i_1,i_2}^t = y_{i_1}^t \qquad i_1 \in \mathcal{F}^t,\ t \in [T-1] \tag{5}$$

$$\textstyle\sum_{i_1} z_{i_1,i_2}^t = y_{i_2}^{t+1} \qquad i_2 \in \mathcal{F}^{t+1},\ t \in [T-1] \tag{6}$$

$$z_{i_1,i_2}^t = 0 \qquad (i_1,i_2) \in \mathcal{E}^t,\ t \in [T-1] \tag{7}$$

The first constraint ensures only k facilities are chosen at each timestep. Constraint 3 forces every client is serviced by some facility within distance R. Constraint 4 relates the x and y variables ensuring that client j can be serviced by facility i only if facility i is built. Constraint 5 and 6 forces the k facilities between consecutive timesteps to form a matching, representing movement. Meanwhile, the last constraint ensures that no facility can move to another one that is farther than distance M. Observe that if R is guessed correctly, then the polytope formed by the above constraints is non-empty.

5.2 Filtering

From the previous observation, let us assume that we have correctly guessed R and have a feasible fractional solution to Constraints 2 - 7. For each timestep t, we use the following filtering technique to aggregate facilities and clients into balls of radius R. Formally, for each timestep t, we partition a subset of the facilities with non-zero y values into at most k sets $F_1^t, F_2^t, \dots F_{k_t}^t$ such that they satisfy:

$$\textstyle\sum_{i \in F_l^t} y_i^t \ge 1 \qquad \forall 1 \le l \le k_t \tag{8}$$

$$diam(F_l^t) \le 2R \qquad \forall 1 \le l \le k_t \tag{9}$$

Here, $diam(S)$ is the diameter of a set, defined as $\max\{d(i_1,i_2) : i_1,i_2 \in S\}$. We also partition the clients into disjoint sets C_l^t for $t \le T$ and $1 \le l \le k_t$ such that:

$$d(i,j) \le 3R \quad \forall i \in F_l^t, j \in C_l^t, 1 \le l \le k_t, t \le T \tag{10}$$

These constraints ensure that every client is close to some ball of facilities after filtering. We obtain these partitions using a standard subroutine, which we detail in the Appendix.

Lemma 5.1. *For each timestep $t \in [T]$, we can produce sets F_l^t and C_l^t satisfying Constraints 8 - 10.*

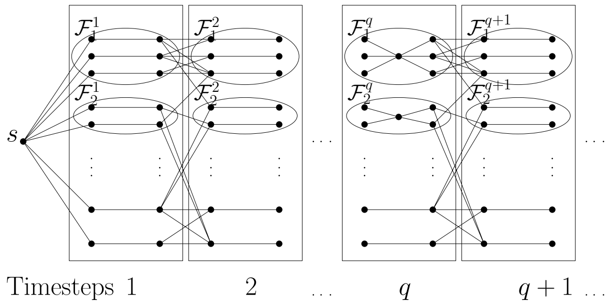

Figure 2: The construction of graph G'. Each timestep, represented by the boxes, contains layers of vertices where flow is required to go from left to right. Each set $\mathcal{F}_i^t$, is represented by an oval. Any facility hat does not belong to any sets $\mathcal{F}_i^t$ is found at the bottom of the boxes. Note that edges between timesteps connects all and only those facilities that are within distance M of each other.

5.3 Flow Rounding

Given these filtered facilities $\mathcal{F}_l^t$, we have a good approximation if we place at least one facility in each set F_l^t by Constraint 10. We can accomplish this by working with an auxiliary graph G', defined as follows (see Figure 2 for an illustration). For each timestep $t \in [T]$ and each facility $i \in F_l^t$, create a second copy of the facility called $i' \in F_l^{t\prime}$. Let $q = \lceil R/M \rceil$. Add an edge ii' for each $i \in \mathcal{F}_l^t$ if and only if t is not a multiple of q. For each t that is a multiple of q, for each set F_l^t, create a vertex v_l^t and connect it to all facilities in that set, including its copies. Thus far, for each timestep, we should have a collection of matchings (when t is not a multiple of q) or disjoint stars where the leafs are facilities and its copies, and the centers indicate which set they belong to. We then direct the edges from the original facilities to its copy (either directly or via the centers), creating two/three layers for each timestep. We connect the layers by adding an edge from $i_1' \in F_{l_1}^{t\prime}$ to $i_2 \in F_{l_2}^{t+1}$ if $d(i_1, i_2) \leq M$. For every facility $i \in \mathcal{F}^t$ that is not in any set F_l^t, we add a vertex i and a copy i' along with an edge ii'. We similarly connect i to any copies of facilities i_1' in time $t-1$ if they are within distance M. Lastly, we connect a source node to all facilities in the first layer and a sink to all facilities in the last layer.

Given the auxiliary graph G', we can produce an integral k-flow through G' such that every vertex corresponding to a set of facilities F_l^t has at least one unit of flow through it by rounding the fractional solution for Constraints 2 – 7 (see Appendix for more details). Since the k-flow is integral, this is equivalent to having k-paths where each path goes through each layer of G' by visiting three nodes $i_1^t, v_l^t, i_2^{t'}$. Naively, we would like to use these k-paths as our solution but i_1^t may not be the same center as i_2^t, so we need to be more careful when choosing locations for a particular time t. Our process, along with the entire network flow rounding algorithm, is described in Algorithm 2 and its guarantees are given in Lemma 5.2. For technical reasons, we require a density assumption on the set of facility locations (see Appendix for precise definition). In practice, we observe that this assumption is easily satisfied and only affects the approximation by a small constant factor. Additional details of the practical implications of the assumption can be found in the Appendix.

Algorithm 2 Network-Flow Algorithm

Require: An instance of Dynamic k-Supplier Problem
Ensure: A $(9, 3)$-approximate solution using k facilities.

1: Binary search through pairwise distances for the minimum R giving a feasible solution to Constraints 2 – 7
2: Apply the filtering algorithm (Lemma 5.1) on the feasible solution to obtain sets F_l^t and C_l^t, $t \in [T]$.
3: Construct auxiliary graph G' as described previously, and obtain an integral k-flow solution to G' with at least one unit of flow going through each ball F_l^t.
4: For each flow-path P, when t is a multiple of q, let i^t be the first vertex the path visits at time t
5: For each aq ($a \in \mathbb{N}$), find shortest route i^{aq} to $i^{(a+1)q}$ and subdivide the route into q equal segments, forming $q-1$ intermediate locations, $p^{aq+1}, p^{aq+2}, ..., p^{aq+q-1}$
6: Let $i^t \in \mathcal{F}^t$ be the closest facility location to the intermediate location p^t for $aq < t < (a+1)q$
7: **return** the facility locations i_1^t, i^t

Theorem 5.2. *For the Dynamic k-Supplier Problem, Algorithm 2 provides a $(9, 3)$-approximate solution, meaning the radius is at most $9 \cdot OPT$ but requires up to $3 \cdot M$ movement.*

6 Experiments

	Total Clients	Activity Locations	Residential Locations	Measured Diameter (km)
Charlottesville City	33156	5660	10038	8.12
Albemarle County	74253	9619	32981	61.62

Table 1: Network Information

Experimental Setup: We run experiments on synthetic mobility data for a city and county in Virginia. The dataset was constructed from the 2019 U.S. population pipeline (see [Chen *et al.*, 2022; Machi *et al.*, 2021] for details) and tracks the week-long activity of residents. Each resident has a record containing a sequence of activities, where each activity is described by duration, type, and location in the municipality. An overview of each municipality is given in Table 1. We set all non-residential locations within each municipality as potential facility locations. Motivated by the vaccination schedule and organization for COVID-19, we spread the vaccination process over a month, with a randomly selected subset of clients to be vaccinated on each day. Since the dataset contains weekly mobility patterns, we extract a day of the week and examine the hours of 6am-8pm for a total of $T = 14$ hour-long timesteps. For each individual in the selected day, we randomly select a visited location and the corresponding hour of the visit for them to be vaccinated. This random sampling and selection process mimics how individuals sign up for vaccination timeslots during the early stages of vaccine administration and allows individuals to indicate locations from which they would like to be serviced. This forms the demand input of our problem. We conduct all our experiments on 10 randomly generated instances of the demand input and show the 95% confidence interval for each result.

Baselines: We compare the performance of Algorithm 1 (COVER) and Algorithm 2 (FLOW) with that of two baselines:

- **Static k-supplier (STATIC)** assumes the facilities cannot move and thus attempts to cover all clients across all timesteps in one shot. The solution is obtained by running the standard k-supplier algorithm. Our goal is to evaluate how much improvement we obtain by allowing movement in our new model for vaccine distribution.
- **Iterative k-supplier (ITERATIVE)** first chooses a set of locations for the first timestep via a standard k-supplier algorithm. For each subsequent timestep, the algorithm chooses k new facility locations by solving a k-supplier problem for the new timestep, with the additional constraint that each facility can move at most M units away. We compare with this baseline since it is the natural greedy heuristic one would use when tackling the Dynamic k-Supplier Problem.

The formal definitions of the baseline heuristics and the details of our parameter settings are given in the Appendix. There, we also explore a heuristic for capacitated Dynamic k-Supplier, where each facility can only serve L people.

6.1 Budget and Objective Tradeoff

Depending on the severity of a disease and the demand for vaccines, policymakers may wish to adjust the number of vaccination sites to deploy. While an increase in the number of vaccination sites decreases the objective coverage radius, the resulting decrease may not be worth the incurred economic and labor cost. Therefore, it is important to explore the tradeoff between the objective and the budget. This experiment varies the budget k from $3-10$ in Charlottesville and $6-20$ for Albermarle, measuring the coverage radius for each of the solutions. As we see in Figure 3, both ITERATIVE and FLOW perform similarly to the static k-supplier algorithm, illustrating that it is not trivial to make good use of the facilities' mobility. The weak performance of these algorithms emphasizes the significance of our COVER algorithm. Additionally, COVER obtains the smallest 95% confidence interval, demonstrating robustness across different population instances.

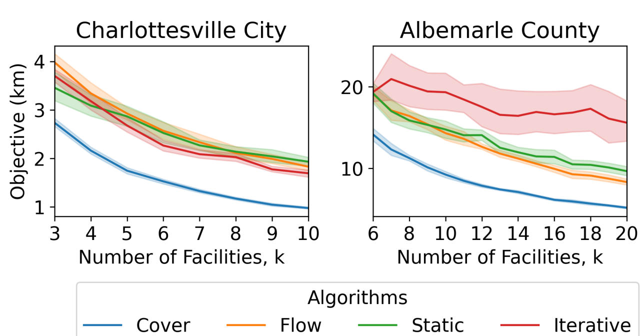

Figure 3: Budget Sensitivity

From the tradeoff curve, we can also provide advice to policymakers for the question: how many vaccination sites should be deployed to most effectively use the economic resources. There appears to be three segments to the tradeoff curve for Charlottesville City: (1) a steep drop in objective between budget values of $3-5$, (2) a moderate slope between $5-9$, and (3) a smaller change between $9-10$. We see the same pattern in Albemarle County for the intervals $6-10$, $10-16$, and $16-20$. As the budget increases beyond each of these intervals, the marginal decrease in the objective is diminished. Therefore, policymakers may benefit from setting the budget to be equal to the endpoints of these intervals to maximize budget and investment efficiency. For this reason, we will be deploying 5 facilities for Charlottesville and 10 facilities for Albemarle in our remaining experiments.

6.2 Percentile Coverage

When considering public health policy and targeting herd immunity, policymakers may wish to prioritize vaccine accessibility for the majority of the population at the expense of a few outliers. Since our objective is defined as the maximum distance a client needs to travel to reach its nearest vaccination site, it does not provide insight on how vaccine accessibility varies for our clients. In this experiment, we evaluate the quality of the vaccination site placements based on the objective values for percentiles $80-100$ of the clients.

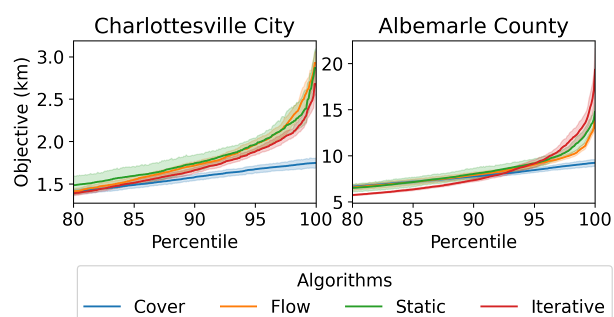

Figure 4: Percentile Coverage

In Figure 4, we observe that COVER generally results in a lower objective than the other algorithms, with the performance gap widening as the percentile increases. The steep increase in objective as percentile increases for FLOW, STATIC, and ITERATIVE indicates that the vaccination site placement is inaccessible for a small group of individuals, a signal of unfairness. In contrast, COVER has a near-linear relationship between objective and percentile, with no sharp increase nearing full coverage. Therefore, COVER not only outperforms the other algorithms at different levels of percentile coverage, but also chooses fairer vaccination site placements that results in few outliers. By utilizing COVER, policymakers can ensure high vaccine accessibility without compromising on fairness.

6.3 Inequality and Priority

Thus far, we assume that the population is homogeneous in behavior where they share similar values/utilities on the time and distance spent on travelling for vaccination. However, in reality, this assumption is far from the truth. For many different external reasons, such as limited mobility due to age, restrictive access to fast public transportation due to location, increased urgency for vaccination due to pre-existing conditions, and etc., policymakers may wish to assign different priorities to different clients to better represent their utility of getting a vaccine, resulting in a more fair planning. This is achieved by assigning different weights, w_j, to different clients. Then, the objective is to minimize $\max_j\{w_j R_j\}$ where R_j is the distance of client j to its closest facility. Note that higher w_j corresponds to higher priority since it forces R_j to be small in order to keep $w_j R_j$ small.

In this experiment, we use the household-income data associated with each client in the synthetic population and divide them into three groups: low, medium, and high income, according to tax-bracket divisions from [IRS, 2022]. A weight of 3, 2, 1 are given to the three groups respectively capturing the idea that due to limited time and/or access to fast transportation, a low-income individual may only wish to travel one-third of the distance to get their vaccination compared to a person with high income. Note that these weights can be easily adjusted to reflect real-life scenarios. We use an extension of COVER, detailed in the Appendix, for the Dynamic Priority k-Supplier problem. This experiment compares the distribution of R_j of different income groups between the unweighted and weighted priorities.

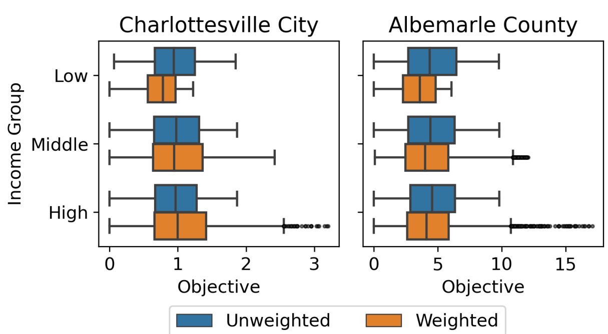

Figure 5: Priority Weighted Cover

Figure 5 illustrates that in both municipalities, using weighted priorities reduces the maximum radius needed to cover the low-income class, achieving the goal of benefiting the intended high-priority group. Even though some members of the other groups has to travel further in the weighted case, there is a decrease in the average coverage radius of each group. This implies that using weighted priorities has an additional benefit of reducing the overall average distance travelled by the population, helping the society as a whole. Thus, policy-makers can take advantage of the weighted priorities to provide targeted aid to those who need it the most.

6.4 Movement Sensitivity

In previous experiments, we set the movement constraint to $M = 5$km as a rough approximation of the distance that a mobile vaccination unit can travel in 10 minutes. However, policymakers can allocate more time for traveling between locations if this greatly decreases the amount people need to travel. Thus here, we explore: to what extent does the movement constraint influence the performance of the algorithms?

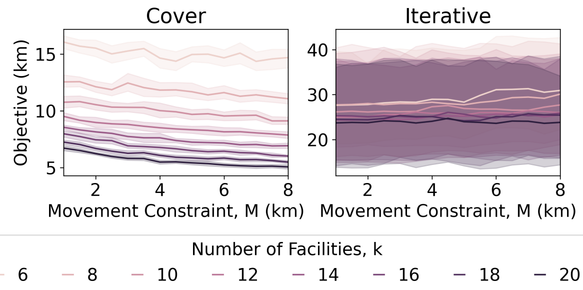

Figure 6: Movement Sensitivity: we run COVER and ITERATIVE on Albemarle for a range of budgets and movement constraints

In Figure 6, we see that even though ITERATIVE utilizes the mobility of vaccination sites, due to its short-sighted iterative design, the algorithm produces volatile objective values that do not exhibit strong trends with respect to the movement constraint. On the other hand, we see that there is a general decrease in objective as facilities are permitted to travel farther for COVER. However, both low and high values of budget exhibit relatively weaker movement sensitivity. This is somewhat reasonable since when only a few vaccination facilities are deployed, the restrictive budget cannot be alleviated by a relaxation of the movement constraint. Meanwhile, when a large number of vaccination facilities are deployed, there is a sufficient number of mobile units such that the movement constraint does not pose a significant limitation on the overall objective. Nonetheless, it is clear that the objective is more sensitive to the number of facilities deployed than the movement constraint. Just by introducing the multi-timestep formulation, we make significant coverage improvements even with restrictive movement constraints.

7 Conclusion

In our work, we use Dynamic k-Supplier to formulate the problem of deploying mobile vaccine distribution sites. We give the first positive results for this problem via two bicriteria approximation algorithms. Through our experiments, we can see that our COVER algorithm substantially improves upon the standard k-supplier formulation of facility location in two primary ways: (i) it chooses more effective vaccination site placements with regards to budget and (ii) the distance different clients need to travel do not differ significantly, so the solutions are more fair (on an individual level). Both properties are essential for policymakers to consider due to the tradeoff between public health and the cost of vaccination site deployment. An additional advantage of our model is that our formulation is much more implementable than previous facility location models for vaccine distribution (e.g., [Li *et al.*, 2022]). In particular, [Li *et al.*, 2022] requires the location data of each client over the course of an entire day. In contrast, our algorithms only require the location where the client wishes to be served, which is the minimal amount needed to formulate a facility location problem. Due to the improved performance, fairness, and implementability, our dynamic formulation of facility location for mobile vaccination site deployment can greatly benefit public health.

Acknowledgements

Da Qi Chen, Ann Li, Madhav Marathe, and Anil Vullikanti were supported by DTRA (Contract HDTRA1-19-D-0007), University of Virginia Strategic Investment Fund award number SIF160, National Institutes of Health (NIH) Grants 1R01GM109718, 2R01GM109718, OAC-1916805 (CINES), CCF-1918656 (Expeditions), CNS-2028004 (RAPID), OAC-2027541 (RAPID), IIS-1908530, IIS-1955797, and IIS-2027848. George Li, Aravind Srinivasan, and Leonidas Tsepenekas were supported in part by NSF award number CCF-1918749.

References

[Acharya *et al.*, 2021] Krishna Prasad Acharya, Tirth Raj Ghimire, and Supram Hosuru Subramanya. Access to and equitable distribution of covid-19 vaccine in low-income countries. *npj Vaccines*, 6(1):1–3, 2021.

[Anthony *et al.*, 2010] Barbara M. Anthony, Vineet Goyal, Anupam Gupta, and Viswanath Nagarajan. A plant location guide for the unsure: Approximation algorithms for min-max location problems. *Math. Oper. Res.*, 35(1):79–101, 2010.

[Bayati *et al.*, 2022] Mohsen Bayati, Rayehe Noroozi, Mohadeseh Ghanbari-Jahromi, and Faride Sadat Jalali. Inequality in the distribution of covid-19 vaccine: a systematic review. *International journal for equity in health*, 21(1):1–9, 2022.

[Bertsimas *et al.*, 2022] Dimitris Bertsimas, Vassilis Digalakis Jr, Alexander Jacquillat, Michael Lingzhi Li, and Alessandro Previero. Where to locate covid-19 mass vaccination facilities? *Naval Research Logistics (NRL)*, 69(2):179–200, 2022.

[Bravo *et al.*, 2022] Fernanda Bravo, Jingyuan Hu, and Elisa Long. Optimal covid-19 vaccination facility location. *Available at SSRN 4008669*, 2022.

[Brubach *et al.*, 2020] Brian Brubach, Darshan Chakrabarti, John P. Dickerson, Samir Khuller, Aravind Srinivasan, and Leonidas Tsepenekas. A pairwise fair and community-preserving approach to k-center clustering. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, volume 119 of *Proceedings of Machine Learning Research*, pages 1178–1189. PMLR, 2020.

[Brubach *et al.*, 2021] Brian Brubach, Nathaniel Grammel, David G. Harris, Aravind Srinivasan, Leonidas Tsepenekas, and Anil Vullikanti. Approximating two-stage stochastic supplier problems. In Mary Wootters and Laura Sanità, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2021, August 16-18, 2021, University of Washington, Seattle, Washington, USA (Virtual Conference)*, volume 207 of *LIPIcs*, pages 23:1–23:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

[Chakrabarti *et al.*, 2022] Darshan Chakrabarti, John P. Dickerson, Seyed A. Esmaeili, Aravind Srinivasan, and Leonidas Tsepenekas. A new notion of individually fair clustering: α-equitable k-center. In Gustau Camps-Valls, Francisco J. R. Ruiz, and Isabel Valera, editors, *International Conference on Artificial Intelligence and Statistics, AISTATS 2022, 28-30 March 2022, Virtual Event*, volume 151 of *Proceedings of Machine Learning Research*, pages 6387–6408. PMLR, 2022.

[Chen *et al.*, 2022] Jiangzhuo Chen, Stefan Hoops, Achla Marathe, et al. Effective social network-based allocation of COVID-19 vaccines. In Aidong Zhang and Huzefa Rangwala, editors, *KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 14 - 18, 2022*, pages 4675–4683. ACM, 2022.

[Deng *et al.*, 2022] Shichuan Deng, Jian Li, and Yuval Rabani. Approximation algorithms for clustering with dynamic points. *J. Comput. Syst. Sci.*, 130:43–70, 2022.

[Fang *et al.*, 2015] Fei Fang, Peter Stone, and Milind Tambe. When security games go green: Designing defender strategies to prevent poaching and illegal fishing. In *Twenty-fourth international joint conference on artificial intelligence*, 2015.

[Gonzalez, 1985] Teofilo F. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theor. Comput. Sci.*, 38:293–306, 1985.

[Hochbaum and Shmoys, 1985] Dorit S. Hochbaum and David B. Shmoys. A best possible heuristic for the k-center problem. *Math. Oper. Res.*, 10(2):180–184, 1985.

[Hochbaum and Shmoys, 1986] Dorit S. Hochbaum and David B. Shmoys. A unified approach to approximation algorithms for bottleneck problems. *J. ACM*, 33(3):533–550, 1986.

[Huang and Li, 2017] Lingxiao Huang and Jian Li. Stochastic k-center and j-flat-center problems. In Philip N. Klein, editor, *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19*, pages 110–129. SIAM, 2017.

[IRS, 2022] IRS. Irs provides tax inflation adjustments for tax year 2022. https://www.irs.gov/newsroom/irs-provides-tax-inflation-adjustments-for-tax-year-2022, 2022. Accessed: 2023-01-16.

[Jia *et al.*, 2022] Xinrui Jia, Kshiteej Sheth, and Ola Svensson. Fair colorful k-center clustering. *Math. Program.*, 192(1):339–360, 2022.

[Kiekintveld *et al.*, 2009] Christopher Kiekintveld, Manish Jain, Jason Tsai, James Pita, Fernando Ordónez, and Milind Tambe. Computing optimal randomized resource allocations for massive security games. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 689–696, 2009.

[Kim *et al.*, 2021] Daniel Kim, Pelin Pekgün, İnci Yildirim, and Pınar Keskinocak. Resource allocation for different types of vaccines against covid-19: Tradeoffs and synergies between efficacy and reach. *Vaccine*, 39(47):6876–6882, 2021.

[Kleindessner *et al.*, 2019] Matthäus Kleindessner, Pranjal Awasthi, and Jamie Morgenstern. Fair k-center clustering for data summarization. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, volume 97 of *Proceedings of Machine Learning Research*, pages 3448–3457. PMLR, 2019.

[Lan *et al.*, 2007] Guanghui Lan, Gail W. DePuy, and Gary E. Whitehouse. An effective and simple heuristic for the set covering problem. *European Journal of Operational Research*, 176, 2007.

[Li *et al.*, 2022] George Z. Li, Ann Li, Madhav V. Marathe, Aravind Srinivasan, Leonidas Tsepenekas, and Anil Vullikanti. Deploying vaccine distribution sites for improved accessibility and equity to support pandemic response. In Piotr Faliszewski, Viviana Mascardi, Catherine Pelachaud, and Matthew E. Taylor, editors, *21st International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2022, Auckland, New Zealand, May 9-13, 2022*, pages 789–797. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2022.

[Lu *et al.*, 2022] Richard Lu, Suhas Gondi, and Alister Martin. Inequity in vaccinations isn't always about hesitancy, it's about access. https://www.aamc.org/news-insights/inequity-vaccinations-isn-t-always-about-hesitancy-it-s-about-access, 2022. American Association of Medical Colleges. Accessed: 2023-01-02.

[Machi *et al.*, 2021] Dustin Machi, Parantapa Bhattacharya, Stefan Hoops, et al. Scalable epidemiological workflows to support COVID-19 planning and response. In *35th IEEE International Parallel and Distributed Processing Symposium, IPDPS 2021, Portland, OR, USA, May 17-21, 2021*, pages 639–650. IEEE, 2021.

[Mehrab *et al.*, 2022] Zakaria Mehrab, Mandy L Wilson, Serina Chang, Galen Harrison, Bryan Lewis, Alex Telionis, Justin Crow, Dennis Kim, Scott Spillmann, Kate Peters, et al. Data-driven real-time strategic placement of mobile vaccine distribution sites. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 12573–12579, 2022.

[Moujahed *et al.*, 2006] Sana Moujahed, Olivier Simonin, Abderrafiâa Koukam, and Khaled Ghédira. A reactive agent based approach to facility location: application to transport. In *4th Workshop on Agants in Traffic and Transportation, located at Autonomous Agents and Multiagent Systems (AAMAS 2006)*, pages 63–69, 2006.

[Nair *et al.*, 2022] Vineet Nair, Kritika Prakash, Michael Wilbur, Aparna Taneja, Corrine Namblard, Oyindamola Adeyemo, Abhishek Dubey, Abiodun Adereni, Milind Tambe, and Ayan Mukhopadhyay. Adviser: Ai-driven vaccination intervention optimiser for increasing vaccine uptake in nigeria. *arXiv preprint arXiv:2204.13663*, 2022.

[Ou *et al.*, 2022] Han-Ching Ou, Christoph Siebenbrunner, Jackson Killian, Meredith B Brooks, David Kempe, Yevgeniy Vorobeychik, and Milind Tambe. Networked restless multi-armed bandits for mobile interventions. *arXiv preprint arXiv:2201.12408*, 2022.

[Rader *et al.*, 2021] Benjamin Rader, Christina M Astley, Kara Sewalk, Paul L Delamater, Kathryn Cordiano, Laura Wronski, Jessica Malaty Rivera, Kai Hallberg, Megan F Pera, Jonathan Cantor, et al. Spatial accessibility modeling of vaccine deserts as barriers to controlling sars-cov-2. *medRxiv*, 2021.

[Raghavan *et al.*, 2019] S. Raghavan, Mustafa Sahin, and F. Sibel Salman. The capacitated mobile facility location problem. *European Journal of Operational Research*, 277(2):507–520, 2019.

[Roy *et al.*, 2021] Satyaki Roy, Pratyay Dutta, and Preetam Ghosh. Generalizable multi-vaccine distribution strategy based on demographic and behavioral heterogeneity. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 1495–1498. IEEE, 2021.

[Shukla *et al.*, 2022a] S Shukla, F Fressin, M Un, H Coetzer, and SK Chaguturu. Optimizing vaccine distribution via mobile clinics: a case study on covid-19 vaccine distribution to long-term care facilities. *Vaccine*, pages 734–741, 2022.

[Shukla *et al.*, 2022b] Samta Shukla, Francois Fressin, Michelle Un, Henriette Coetzer, and Sreekanth K Chaguturu. Optimizing vaccine distribution via mobile clinics: a case study on covid-19 vaccine distribution to long-term care facilities. *Vaccine*, 40(5):734–741, 2022.

[Wada *et al.*, 2018] Yuho Wada, Tomohiro Ono, Taiki Todo, and Makoto Yokoo. Facility location with variable and dynamic populations. In *AAMAS*, pages 336–344, 2018.

[Wel,] Equitable access to vaccines, tests and treatments for covid-19. https://wellcome.org/what-we-do/infectious-disease/coronavirus-covid-19/access. Accessed: 2023-01-02.

[Wouters *et al.*, 2021] Olivier J Wouters, Kenneth C Shadlen, Maximilian Salcher-Konrad, Andrew J Pollard, Heidi J Larson, Yot Teerawattananon, and Mark Jit. Challenges in ensuring global access to covid-19 vaccines: production, affordability, allocation, and deployment. *The Lancet*, 397(10278):1023–1034, 2021.

[Yan, 2021] Holly Yan. Covid-19 vaccine myths: These reasons for not getting a shot don't hold up. in fact, they'll set the us back. *Cable News Network*, July 2021.