The Traveling k-Median Problem: Approximating Optimal Network Coverage

Dylan Huizing[1(✉)] and Guido Schäfer[1,2]

[1] Centrum Wiskunde and Informatica, Amsterdam, The Netherlands
{dylan.huizing,g.schaefer}@cwi.nl
[2] University of Amsterdam, Amsterdam, The Netherlands

Abstract. We introduce the *Traveling k-Median Problem (TkMP)* as a natural extension of the k-Median Problem, where k agents (medians) can move through a graph of n nodes over a discrete time horizon of ω steps. The agents start and end at designated nodes, and in each step can hop to an adjacent node to improve coverage. At each time step, we evaluate the coverage cost as the total connection cost of each node to its closest median. Our goal is to minimize the sum of the coverage costs over the entire time horizon.

In this paper, we initiate the study of this problem by focusing on the uniform case, i.e., when all edge costs are uniform and all agents share the same start and end locations. We show that this problem is NP-hard in general and can be solved optimally in time $\mathcal{O}(\omega^2 n^{2k})$. We obtain a 5-approximation algorithm if the number of agents is large (i.e., $k \geq n/2$). The more challenging case emerges if the number of agents is small (i.e., $k < n/2$). Our main contribution is a novel rounding scheme that allows us to round an (approximate) solution to the 'continuous movement' relaxation of the problem to a discrete one (incurring a bounded loss). Using our scheme, we derive constant-factor approximation algorithms on path and cycle graphs. For general graphs, we use a different (more direct) approach and derive an $\mathcal{O}(\min\{\sqrt{\omega}, n\})$-approximation algorithm if $d(s,t) \leq 2\sqrt{\omega}$, and an $\mathcal{O}(d(s,t) + \sqrt{\omega})$-approximation algorithm if $d(s,t) > 2\sqrt{\omega}$, where $d(s,t)$ is the distance between the start and end point.

Keywords: k-median problem · Network coverage · Routing over time · Approximation algorithms

1 Introduction

Background and Motivation. The *k-Median Problem (k-MED)* is a classic problem in combinatorial optimization that has been studied extensively, both because of its theoretical appeal as well as its practical relevance. In this problem, we are given a finite metric space (V, d) and a parameter k. Our goal is to select

J. Koenemann and B. Peis (Eds.): WAOA 2021, LNCS 12982, pp. 80–98, 2021.
https://doi.org/10.1007/978-3-030-92702-8_6

k points (called *medians*) such that the sum of the connection costs with respect to d from each node $u \in V$ to its nearest median is minimized.

In this paper, we consider the following natural extension of the k-Median Problem, which we term the *Traveling k-Median Problem (TkMP)*: Suppose we are given a fixed time horizon $T = \{0, \dots, \omega\}$ and a set $A = \{1, \dots, k\}$ of k medians (called *agents* in this context) which can move through the graph $G = (V, E)$ in discrete steps. Each agent $a \in A$ starts at time $\tau = 0$ at a designated start node $s_a \in V$, and ends at time $\tau = \omega$ at a designated end node $t_a \in V$. In each time step, an agent may either hop from her current location $u \in V$ to an adjacent node $v \in V$, or simply stay at u. Our goal is to determine a set of k walks (one for each agent) such that the total coverage costs (summed over all time steps) is minimized. Here, the *coverage cost* at time τ is defined the same way as in the k-Median Problem (with respect to the locations of the agents at time τ).

Intuitively, in the Traveling k-Median Problem one wants to distribute the agents over the graph such that they guarantee a low coverage cost, ideally positioning them at the optimal k-median locations. However, in many real-life applications we cannot simply assume that the agents can be 'teleported' to their ideal locations. Instead, the agents might have to move there over time, which in turn influences the achievable coverage. Assuming that the time horizon is fixed adds some new characteristics to the problem which are relevant to several applications in emergency logistics. For example, the agents might correspond to an 'emergency fleet' (e.g., police patrol force, emergency medical coverages, railway emergency responders) that one needs to schedule to provide a good emergency response time over a pre-specified time horizon. In fact, we arrived at the Traveling k-Median Problem introduced in this paper through our (more applied) investigations in the domain of emergency logistics; more specifically, the Traveling k-Median Problem considered here is a stylized variant of the so-called *Median Routing Problem* studied in [12].

Our Contributions. In this paper, we initiate the study of the Traveling k-Median Problem (TkMP) by focusing on the uniform case (referred to as U-TkMP), i.e., when all edge costs are uniform and all agents start from the same start location and end at the same end location. Our main contributions are as follows:

(1) We first argue that even U-TkMP is NP-hard in general by a simple reduction from the Set Cover Problem. We then show that the problem can be solved optimally in $\mathcal{O}(\omega^2 n^{2k})$ time. (These results are presented in Sect. 3.)
(2) We derive a 5-approximation algorithm for U-TkMP if the number of agents is large, i.e., $k \geq n/2$. This allows us to assume $k \leq n/2$ in the remainder. Basically, the idea here is to let each agent take care of an edge in a maximum matching of the underlying graph. (These results are given in Sect. 3).
(3) We introduce a novel way of approximating U-TkMP by rounding an (approximate) solution to a 'continuous movement' relaxation of the problem to a discrete solution, thereby incurring a bounded loss in the approximation factor. (This Rounding Theorem is presented in Sect. 4).

(4) We use this Rounding Theorem to derive 10-approximation algorithms for U-TkMP on path graphs and cycle graphs. (These results are given in Sect. 5).
(5) Finally, for U-TkMP on general graphs we obtain the following results: If the shortest path distance $d(s,t)$ between s and t satisfies $d(s,t) \leq 2\sqrt{\omega}$, then we obtain an $\mathcal{O}(\min\{\sqrt{\omega}, n\})$-approximation algorithm. If $d(s,t) > 2\sqrt{\omega}$, we instead have an $\mathcal{O}(d(s,t) + \sqrt{\omega})$-approximation algorithm. This is useful for those graph topologies for which we do not yet know how to solve or approximate their continuous counterparts. (These results are given in Sect. 6).

Related Work. To the best of our knowledge, the Traveling k-Median Problem introduced in this paper has not been studied in the literature before. The following are the key distinguishing features of TkMP with respect to other problems considered in the literature: (i) We consider a fixed time horizon and all agents have to start and end at designated nodes; in particular, this restricts the walks that can be chosen by each median. (ii) The coverage cost accounts for the total connection cost of each node to its closest median (k-MED-objective). However, there are several other optimization problems that are related to TkMP. Below is a partial list of related problems.

At the core of TkMP lies the k-Median Problem which has been studied extensively in the literature. The metric problem introduced above was first shown to be NP-hard by Kariv and Hakimi [15]. The more general variant of the problem where the connection cost is non-negative and symmetric (but does not necessarily satisfy the triangle inequality) was first shown to be APX-hard by Lin and Vitter [18]. On the positive side, the problem is polynomial time solvable on trees [15,21] and there exists a polynomial-time approximation scheme if (V,d) constitutes a Euclidean space [1]. The current best approximation algorithm for the problem achieves an approximation ratio of $\alpha = 2.675 + \varepsilon$ [4]. In terms of negative results, the current best lower bound is the $(1 + 2/e)$-inapproximability result by Guha and Khuller [9].

TkMP is also related to classic online problems such as the *k-Server Problem* [17] or, more generally, the *Metrical Task Systems* [3]. In the k-Server Problem [17], there are k servers that can move through a graph and have to serve a sequence of requests. Also here, the servers incur some movement costs (given by a metric). In each step a node is requested and a server has to move there. The main difference with respect to TkMP is that the time horizon is not limited and that there is no coverage cost. In Metrical Task Systems [3], an online algorithm starts at a designated node of a given graph and has to cover a sequence of tasks. Also here, the algorithm incurs some coverage cost and some movement cost (given by a metric) to serve a sequence of tasks. However, the coverage cost is given in terms of a non-negative cost incurred at the node in which the algorithm resides; this is different from TkMP, where the coverage cost accounts for the total connection cost of the nodes to the medians.

Dynamic facility location models are often on a strategic level [5]. Wesolowsky [23] extends several static facility location models, including k-MED, to a problem where facility locations have to be chosen for several periods. If the facility locations differ from one period to the next, then a period-dependent cost is incurred, independent of how many facilities are relocated and how far.

Of the many dynamic facility location problems reviewed and classified by Farahani et al. in 2012 [5], none fits our setting of one configuration determining the decision space in the next time-step. Distance-dependent relocation costs appear to be less commonly discussed than time-dependent costs, but they are studied by Huang et al. [11] and Gendreau et al. [7]. Galvão and Santibanez-Gonzalez, instead, charge period-dependent costs for having a facility open [6]. Melo et al. consider a combination of facility opening, closing and maintenance costs, and costs of reallocating capacity between facilities [19].

On a more operational level, ambulance relocation models seek to deploy an ambulance to a revealed emergency and relocate the rest for optimal response to the next emergency [8,13]. Even these operational models assume 'teleportation' from one steady-state configuration to the next. The same holds on the online level for the classical k-Server Problem [16]. Bertsimas and Van Ryzin coordinate vehicles over the Euclidean plane to minimize the completion time of dynamically revealed requests [2].

From a game-theoretical standpoint, Hotelling [10] discusses the *Ice-Cream Man Problem*, in which two competing ice cream vendors position themselves on the line to capture the largest portion for themselves. When moving to increase the portion of the line they cover, they surprisingly reach an equilibrium when their distance approaches zero. However, an equilibrium is no longer reached for three vendors [22].

2 Preliminaries

In the *Traveling k-Median Problem* (TkMP), we are given a connected graph $G = (V, E)$ with n nodes, a set $A := \{1, \dots, k\}$ of k agents, and a discrete time horizon $T := \{0, \dots, \omega\}$. Each agent $a \in A$ has their own start location $s_a \in V$, where they must be at time 0, and an end location $t_a \in V$, where they must be at time ω. Between each time-step and the next, agents may stay where they are or 'hop' to adjacent nodes.

At each time-step, we evaluate how well 'spread' the agents currently are. A *configuration* $\sigma = (\sigma_1, \dots, \sigma_k) \in V^A$ specifies for each agent $a \in A$ the node $\sigma_a \in V$ at which a resides. Let $d(u, v)$ denote the cost function for covering $v \in V$ from $u \in V$. We assume that every node will be covered from its nearest median in σ, i.e., the cost of node v is $D(\sigma, v) := \min_{a \in A} d(\sigma_a, v)$, and the total cost of this configuration is $D(\sigma) := \sum_{v \in V} D(\sigma, v)$. We say that a configuration σ is *optimal* if it minimizes the total cost $D(\sigma)$ (over all possible configurations).

A *feasible solution* $\mu \in V^{A \times T} = (\mu^0, \dots, \mu^\omega)$ is a sequence of $(T + 1)$ configurations, with μ^0 and μ^ω given by the start and end locations, respectively, and each configuration μ^τ being *adjacent* to the previous configuration $\mu^{\tau-1}$ for every $\tau \in \{1, \dots, \omega\}$. We say that μ^τ is adjacent to configuration $\mu^{\tau-1}$ if for each agent $a \in A$, either $\mu_a^{\tau-1} = \mu_a^\tau = u$ (agent a remains at node u), or $\mu_a^{\tau-1} = u$,

$\mu_a^\tau = v$ and $(u, v) \in E$ (agent a moves along edge (u, v) from u to v). For any feasible solution μ, we define the cost as $D(\mu) := \sum_{\tau=0}^{\omega} D(\mu^\tau)$. TkMP is the problem of finding a solution μ with minimal $D(\mu)$.

Given a feasible solution μ, we call the sequence $(\mu_a^0, \ldots, \mu_a^\omega)$ of nodes visited by agent a also the *walk* of a. The cost of the walk of agent $a \in A$ with respect to some fixed node v is defined as $\sum_{\tau=0}^{\omega} D(\mu_a^\tau, v)$.

In the *Uniform Traveling k-Median Problem* (U-TkMP), which is a special case of TkMP, we make the following three assumptions: (i) d is defined as the shortest path distance (with respect to the number of edges) induced by the underlying graph G, (ii) all agents start at the same node s, and (iii) all agents end at the same node t.

3 Hardness and Bounds

In this section, we will derive some lemmas and observations that will serve as instrumental building blocks for the approximation algorithms in this paper. As a first result, it deserves mentioning that even U-TkMP is NP-hard.

Theorem 1. *U-TkMP and its generalization TkMP are NP-hard.*

Proof. This follows from a reduction from the Set Cover Problem. See Appendix A. □

Theorem 2. *TkMP is solvable in $\mathcal{O}(\omega^2 n^{2k})$.*

Proof. This follows from formulating a shortest path problem on a digraph with one node per agent configuration and time-step. See Appendix A. □

However, if we want to determine a minimum cost walk of one agent with respect to a given node v, we can do this quite easily.

Lemma 1. *In U-TkMP, given any node v and an agent $a \in A$, we can find a minimum cost (s, t)-walk $\pi = (\pi^0, \ldots, \pi^\omega)$ of agent a with respect to v in $\mathcal{O}(n^4)$ time.*

Proof. See Appendix A.

We now provide several bounds on performance and parameters.

Lemma 2. *Consider U-TkMP. Then an optimal configuration σ^* has cost at least $D(\sigma^*) \geq (n - k)$. Further, any feasible solution μ has cost at least $D(\mu) \geq (\omega + 1)D(\sigma^*) \geq (\omega + 1)(n - k)$.*

Proof. In any configuration σ, at most k nodes can be occupied. All other nodes account for at least a cost of 1, implying $D(\sigma^*) \geq (n - k)$. Because any solution μ decomposes into $(\omega + 1)$ feasible configurations, $D(\mu) \geq (\omega + 1)D(\sigma^*)$. □

Theorem 3. *In U-TkMP, if $k \geq n/2$, we can find a 5-approximation in $\mathcal{O}(n^5)$ time.*

Proof. We construct the approximate solution μ as follows. Using a greedy $\mathcal{O}(n^2)$ algorithm, find any (inclusion-wise) maximal matching on G. For every edge in this matching, assign an agent to one of its end points: we have enough agents to do this, because no matching can consist of more than $n/2$ edges. Assign the remaining agents to t. For every agent, determine a minimum cost walk with respect to the assigned node by using Lemma 1 (which requires at most $\mathcal{O}(n^4)$ time). We need to determine at most $n/2 + 1$ such walks, resulting in an algorithm that needs at most $\mathcal{O}(n^5)$ time.

We can compare μ to the hypothetical solution μ^n where the number of available agents is n, and we assign each node its own agent. In μ^n, we again send each agent over a walk that minimizes the total distance to their picked node. Note that $D(\mu^n)$ is minimal, because every node that gets its own agent is served optimally.[1] If node v is 'picked' by μ, v contributes this same optimal amount to $D(\mu)$. If v is not 'picked', we show here that the cost increases by at most 4 per time-step.

Take any node v and time-step τ, and let us examine the contribution of (v, τ) to the cost function. Let a be the agent in μ^n assigned to v. Let $u \in V$ be the location of a at time τ. In μ, there may or may not be someone at u at time τ. If there is, the request (v, τ) has the same cost contribution to $D(\mu)$ as to $D(\mu^n)$, remarking again that nodes with their own agent are served optimally. If not, there is a node $u' \in V$ that was picked, at most 2 hops away from u; otherwise, the underlying matching would not be maximal. If the agent in μ assigned to u' is still headed for it, that agent is at most 2 hops away from u', because $d(s, u') \leq d(s, u) + d(u, u') \leq \tau + 2$. Similarly, if the agent is moving away from u' towards t, the agent is at most 2 hops away from u' because $d(u', t) \leq d(u', u) + d(u, t) \leq (\omega - \tau) + 2$. That is, there are only $\omega - \tau$ hops left to get to t, and while this is no problem when $d(u', t)$ is small, the agent will be pulled away by at most 2 hops if $d(u', t)$ is large. So indeed, by having less agents then in μ^n, each unpicked node v (at most $n - k$) contributes at most 4 extra per time-step. Thus, if μ^* is some optimal solution, then

$$D(\mu) \leq D(\mu^n) + 4(\omega + 1)(n - k) \leq D(\mu^*) + 4(\omega + 1)(n - k) \leq 5D(\mu^*). \quad \square$$

As a result of Theorem 3, we will often assume in the remainder that $k \leq n/2$, which yields the following useful refinement of Lemma 2.

Corollary 1. *Consider U-TkMP and assume that $k \leq n/2$. Then an optimal configuration σ^* has cost at least $D(\sigma^*) \geq \frac{1}{2}n$. Further, any feasible solution μ has cost at least $D(\mu) \geq (\omega + 1)D(\sigma^*) \geq \frac{1}{2}(\omega + 1)n$.*

[1] That is, for every node v, there exists an agent who, through Lemma 1, is always as nearby as possible, so far as the start and end conditions allow. While it may seem nontrivial which nodes this agent could give coverage to on the way to v, in fact those intermediate nodes have agents of their own serving them optimally, and the agent never has to cover anything else but v.

Lemma 3. *In U-TkMP, for any feasible solution* μ *and time-steps* τ_1 *and* τ_2*, we have* $D(\mu^{\tau_2}) \leq D(\mu^{\tau_1}) + |\tau_2 - \tau_1|n$.

Proof. For any $v \in V$, let a be the agent nearest to it in μ^{τ_1}. The contribution of v to $D(\mu^{\tau_2})$ cannot exceed the distance of v to a, who cannot have moved more than $|\tau_2 - \tau_1|$ hops away from their position in μ^{τ_1}. □

Theorem 4. *Suppose we have an* α*-approximation algorithm for k-MED that runs in* $\mathcal{O}(kmed(n))$ *time. In U-TkMP, if* $k \leq n/2$ *and* $\omega \geq 4n^2 + 2n$*, we can find an* $(\alpha + 1)$*-approximation in* $\mathcal{O}(kmed(n) + n^5)$ *time.*

The proof of Theorem 4 is given in Appendix A. Note that this theorem helps us in the following way: instances with $\omega \geq 4n^2 + 2n$ are 'easy', and we can assume in the remainder that $\omega \leq 4n^2 + 2n$. This, in turn, means that algorithms with runtime depending on ω (polynomially) can still be considered polynomial-time.

We also remark that by combining Theorem 2 and Theorem 4, U-TkMP can be solved in $\mathcal{O}(n^6)$ if there is only one agent.

4 Rounding Continuous Graph Movement

A related problem to U-TkMP is the *Continuous Uniform Traveling k-Median Problem* (CU-TkMP), which we can use to approximate U-TkMP. In CU-TkMP, the allowed positions for agents are not only the nodes, but also any point on the edges.

Technically, we define CU-TkMP as follows. Given an unweighted, undirected graph $G = (V, E)$, we construct its 'continuous version' as the following metric space $(\tilde{E}, c)$. Assuming an arbitrary orientation on the edges of E, we build $\tilde{E}$ out of V and all points 'strictly on' edges: $\tilde{E} := V \cup E \times (0, 1)$, where a point $((u, v), \theta) \in E \times (0, 1)$ represents the point that is on the oriented edge (u, v), at a fraction θ from u to v.

As for the distance metric c, we base it on a continuous version of the shortest path distance as follows. Take any $x \in \tilde{E}$ and $y \in \tilde{E}$. If both are in V, then $c(x, y) = d(x, y)$, where d is the shortest path distance on G. If $x \in V$ but $y \notin V$, let $y = ((u, v), \theta)$. Then $c(x, y) = \min\{\theta + d(u, x), 1 - \theta + d(v, x)\}$. We uphold an analogous definition if $y \in V$ but $x \notin V$. Finally, if both points are not in V, let again $y = ((u, v), \theta)$. Then $c(x, y) = \min\{\theta + c(u, x), 1 - \theta + c(v, x)\}$.

Now, in CU-TkMP, we are again given k agents and a discrete time horizon $\{0, \ldots, \omega\}$. The agents must start at $s \in V$ at time $\tau = 0$ and must be at their end location $t \in V$ at time $\tau = \omega$. If an agent is at position $x \in \tilde{E}$ at time τ, then at time $\tau + 1$, they may only be at a position $y \in \tilde{E}$ with $c(x, y) \leq 1$. The continuous cost of a configuration $\sigma \in \tilde{E}^A$ is not the discrete sum of distances to nodes, but the integrated distance to all points in $\tilde{E}$, i.e., $C(\sigma) := \int_{\tilde{E}} \min_{x \in C} c(x, y) \mathrm{d}y$. The continuous cost of a solution $\nu \in \tilde{E}^{A \times T}$ is again the sum over time, i.e., $C(\nu) := \sum_{\tau \in T} C(\nu^\tau)$.

Suppose G has maximum degree Δ, and $k \leq n/2$. Suppose also we can find a β-approximate solution ν to CU-TkMP. We prove that we can 'round' ν to a $(4\beta\Delta + 2)$-approximate solution $\tilde{\nu}$ for the original, 'discrete' problem.

We first show that the cost functions D and C are related. Remark that, if ρ is a continuous configuration, we mean by $D(\rho)$ that we sum distances to V rather than integrating over $\tilde{E}$: we define $D(\rho) := \sum_{v \in V} \min_{u \in \rho} c(u, v)$. If ν is a continuous solution, we define $D(\nu)$ similarly.

Lemma 4. *Let $\rho \in \tilde{E}^A$ be a continuous configuration. Then $D(\rho) \leq 2C(\rho) + k$.*

Proof. Given ρ, we split E into three classes. Denote by $E_0 \subset E$ all edges $\{u, v\}$ that have at least one agent 'strictly' on them, meaning we disregard agents on u and v. Note that $|E_0| \leq k$. Let $E_1 \subset E \setminus E_0$ denote all unoccupied edges $\{u, v\}$ with $d(\rho, u) \neq d(\rho, v)$, oriented as (u, v) if $d(\rho, u) < d(\rho, v)$ and as (v, u) otherwise. Let $E_2 \subset E \setminus E_0$ refer to all unoccupied edges $\{u, v\}$ for which $d(\rho, u) = d(\rho, v)$. We orient E_0 and E_2 arbitrarily.

For any edge $(u, v) \in E_0$, we have $d(\rho, u) + d(\rho, v) \leq 1$, as any one agent on (u, v) is $0 < \theta < 1$ away from u and $1 - \theta$ away from v. Notice that $\{u : (u, v) \in E_0\} \cup \{v : (u, v) \in E_0\} \cup \{v : (u, v) \in E_1\} = V$, as any node $v \in V$ is either incident to E_0, or has a neighbor u on the shortest (v, ρ)-path, thus $(u, v) \in E_1$. Thus summing over these three node sets is sufficient, though we may be over-counting:

$$D(\rho) \leq \sum_{(u,v) \in E_0} (d(\rho, u) + d(\rho, v)) + \sum_{(u,v) \in E_1} d(\rho, v) \leq |E_0| + \sum_{(u,v) \in E_1} (d(\rho, u) + 1)$$

Each $(u, v) \in E_1$ has $\int_u^v d(\rho, w) \mathrm{d}w = d(\rho, u) + \int_0^1 w \mathrm{d}w = d(\rho, u) + \frac{1}{2}$. So

$$C(\rho) = \sum_{(u,v) \in E} \int_u^v d(\rho, w) \mathrm{d}w \geq \sum_{(u,v) \in E_1} (d(\rho, u) + \tfrac{1}{2}) \geq \tfrac{1}{2}|E_1|$$

Putting this together, we find

$$D(\rho) \leq \sum_{(u,v) \in E_1} (d(\rho, u) + \tfrac{1}{2}) + \tfrac{1}{2}|E_1| + |E_0| \leq 2C(\rho) + k. \qquad \square$$

Lemma 5. *Let $\sigma \in V^A$ be a discrete configuration and $n \geq 2k$. Then $C(\sigma) \leq 2\Delta D(\sigma)$.*

Proof. Because σ is a discrete configuration, $E_0 = \emptyset$. Each $(u, v) \in E_1$ has $\int_u^v d(\sigma, w) \mathrm{d}w = d(\sigma, u) + \int_0^1 w \mathrm{d}w = d(\sigma, u) + \frac{1}{2}$. Each $(u, v) \in E_2$ has $\int_u^v d(\sigma, w) \mathrm{d}w = d(\sigma, u) + 2 \int_0^{\frac{1}{2}} w \mathrm{d}w = d(\sigma, u) + \frac{1}{4}$.

Observe some oriented version $\vec{E}$ of E, such that if $(v, u) \in \vec{E}$, then $d(\sigma, u) \leq d(\sigma, v)$. For any $u \in V$, let $\delta^-(u)$ be the edges in E such that, in $\vec{E}$, they point at u. Then $|\delta^-(u)| \leq \Delta$ for any $u \in V$.

$$C(\sigma) = \sum_{(u,v)\in E_1} (d(\sigma,u) + \tfrac{1}{2}) + \sum_{(u,v)\in E_2} (d(\sigma,u) + \tfrac{1}{4}) \leq \sum_{u\in V} |\delta^-(u)|(d(\sigma,u) + \tfrac{1}{2})$$
$$\leq \Delta \sum_{u\in V} d(\sigma,u) + \tfrac{1}{2}\Delta n \leq \Delta D(\sigma) + \Delta \cdot (n-k) \leq 2\Delta D(\sigma) \qquad \square$$

Suppose we have some continuous solution $\nu = (\nu_0, \ldots, \nu_\omega)$, which has $C(\nu) \leq \beta C^*$, with C^* the optimal coverage solution value for CU-TkMP. We can for each agent a find a continuous movement line across $(\nu_0)_a, \ldots, (\nu_\omega)_a$ by taking shortest walks through $\tilde{E}$ from each $(\nu_\tau)_a$ to $(\nu_{\tau+1})_a$, waiting at $(\nu_{\tau+1})_a$ until $\tau+1$ if necessary. We can obtain a 'rounded solution' $\tilde{\nu}$ by setting $(\tilde{\nu}_\tau)_a$ equal to $(\nu_\tau)_a$ if that is a node. If instead it is a point on an edge (u,v), because the agent is moving from u to v in the continuous movement line, then we set $(\tilde{\nu}_\tau)_a$ to the closest of these two to $(\nu_\tau)_a$. If the point is exactly halfway between u and v, we set $(\tilde{\nu}_\tau)_a$ to the destination v. Assuming again that $k \leq n/2$ and $\omega \leq 4n^2$, we can perform this rounding in $\mathcal{O}(k\omega) = \mathcal{O}(n^3)$.

Theorem 5 (Rounding Theorem). *If $n \geq 2k$, then rounding ν to the solution $\tilde{\nu}$ yields a $(4\beta\Delta + 2)$-approximation with respect to $D(\mu^*)$, where μ^* is an optimal solution.*

Proof. Take some instance of U-TkMP with $n \geq 2k$. Let μ^* be an optimal solution for this instance, and ν^* an optimal solution for its CU-TkMP-equivalent. Let ν be a β-approximate solution to the CU-TkMP-instance, and let $\tilde{\nu}$ be the result of rounding ν in the way described above.

Solution $\tilde{\nu}$ is feasible for the discrete problem. That is, each agent still starts at their start point and ends at their end point; we show that they also only move from nodes to neighboring nodes. Take any agent a and any time-step $\tau < \omega$. Define $q := (\nu_\tau)_a$ and $r := (\nu_{\tau+1})_a$. By feasibility of ν in CU-TkMP, $d(q,r) \leq 1$. If q and r are on the same edge (u,v), including the end points, then they are rounded to u and/or v, which are neighboring. If q is on (u,v) and r is not, then r is on some (v,w). If q or r is rounded to v, then q and r are rounded to neighboring nodes. Suppose not, so q is rounded to u and r to w. Then $d(q,v) > \frac{1}{2}$ and $d(v,r) \geq \frac{1}{2}$, so $d(q,r) = d(q,v) + d(v,r) > 1$, but this contradicts feasibility of ν in CU-TkMP.

It remains to show the approximation factor. Using Corollary 1 for OPT, we have

$$OPT := D(\mu^*) \geq (\omega+1)(n-k) \geq (\omega+1)\tfrac{n}{2} \geq (\omega+1)k$$
$$D(\tilde{\nu}) \leq D(\nu) + (\omega+1)n \cdot \tfrac{1}{2} \leq D(\nu) + OPT \qquad (1)$$
$$D(\nu) \leq \sum_{\tau=0}^{\omega} (2C(\nu_\tau) + k) \leq 2C(\nu) + OPT \qquad (2)$$
$$C(\nu) \leq \beta C(\nu^*) \leq \beta C(\mu^*) \qquad (3)$$
$$C(\mu^*) \leq 2\Delta D(\mu^*) = 2\Delta \cdot OPT \qquad (4)$$

(1) follows from the fact that every demand point (v, τ) has its median shifted away at most $\frac{1}{2}$ in rounding. (2) follows from Lemma 4. (3) follows from the fact that ν is β-approximate for CU-TkMP, while μ^* is feasible. (4) follows from Lemma 5. □

5 Topologies with Constant Factor Guarantees

In this section, we show how to compute an optimal solution to the continuous problem CU-TkMP on two simple graph topologies, namely on a path and on a cycle. In both cases, we need to solve a linear program with a convex minimization objective, which can be done in polynomial time (see, e.g., [20]). Throughout this section, we use $convex(k_1, k_2)$ to refer to the time that is needed to solve such a program consisting of k_1 variables and k_2 constraints. Recall that in light of Theorem 4, we can assume that $k \leq n/2$ and $\omega \leq 4n^2 + 2n$; in particular, this implies that $\mathcal{O}(k\omega) = O(n^3)$

5.1 Path Case

We show that one can compute an optimal solution to the continuous problem CU-TkMP if the underlying topology is a line of length L, i.e., $\tilde{E} = [0, L]$. By applying our Rounding Theorem (Theorem 5) to the optimal solution, we obtain the result stated in Theorem 6 below (with $\beta = 1$ and $\Delta = 2$).

Theorem 6. *Let $n \geq 2k$. Then there is a 10-approximation algorithm for U-TkMP on the path graph that runs in $\mathcal{O}(convex(k_1, k_2) + n^3)$ time, where $k_1 = k_2 = \mathcal{O}(k\omega)$.*

A key observation that we can exploit in the path case is that we can impose an order on the positions of the agents. More precisely, let $\sigma \in \tilde{E}^A$ be a configuration. Assume that the agents in $A = \{1, \ldots, k\}$ are indexed such that $\sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_k$ (break ties arbitrarily). For notational convenience, we introduce two dummy agents 0 and $k+1$ which are fixed at $\sigma_0 = 0$ and $\sigma_{k+1} = L$, respectively. Further, let λ_a be a scalar that is defined as $\lambda_a = \frac{1}{4}$ for all $a \in \{1, \ldots, k-1\}$ and $\lambda_0 = \lambda_k = \frac{1}{2}$.

Note that we do not have to worry about agents 'switching' their index: because everyone has the same start and end location, there obviously exists an optimal solution in which they do not switch their index. This is because for any feasible solution with a switch, there also exists a feasible solution with identical cost in which the switch is not made, namely the solution in which the agents approach each other for the switch but go back again. Therefore, we are free to enforce in our convex program that agents will always keep their index.

The following lemma shows that the coverage cost reduces to a function that depends on the distances between consecutive agents.

Lemma 6. *Consider a continuous configuration $\sigma \in \tilde{E}^A$ and suppose that the agents are indexed as stated above. Then the coverage cost is*

$$C(\sigma) = \sum_{a=0}^{k} \lambda_a (\sigma_{a+1} - \sigma_a)^2.$$

Proof. Consider the cost of an interior line segment $[\sigma_a, \sigma_{a+1}]$ between two consecutive agents σ_a and σ_{a+1} with $a \in \{1, \dots, k-1\}$. In this case, each point of the segment is covered by the closest agent among σ_a and σ_{a+1}. Thus, the coverage cost is

$$\int_{\sigma_a}^{\sigma_{a+1}} \min_{u \in \{\sigma_a, \sigma_{a+1}\}} c(u, v) \mathrm{d}v = 2 \int_0^{\frac{1}{2}(\sigma_{a+1} - \sigma_a)} v \mathrm{d}v = \frac{1}{4}(\sigma_{a+1} - \sigma_a)^2.$$

Next, consider the first line segment $[0, \sigma_1]$. In this case, each point of the segment is covered by σ_1. Thus, the coverage cost is

$$\int_0^{\sigma_1} c(\sigma_1, v) \mathrm{d}v = \int_{\sigma_0}^{\sigma_1} v \mathrm{d}v = \frac{1}{2}(\sigma_1 - \sigma_0)^2.$$

A similar argument proves that the coverage cost of the last line segment $[\sigma_k, \sigma_{k+1}]$ is $\frac{1}{2}(\sigma_{k+1} - \sigma_k)^2$. Summing over all line segments proves the claim. □

Exploiting the above observation, we can solve CU-TkMP for the path case simply by formulating the problem as a linear program with a convex (in fact quadratic) objective function.

We introduce a variable $x_a^\tau \in [0, L]$ for each $a \in \{0, \dots, k+1\}$ and time step $\tau \in T$. Our interpretation is that x_a^τ represents the position of the a-th leftmost agent (breaking ties arbitrarily). The following convex program then captures the continuous version of the problem:

$$\begin{aligned}
\min \quad & \textstyle\sum_{\tau \in T} \sum_{a=0}^{k} \lambda_a (\sigma_{a+1} - \sigma_a)^2 && \text{(QP1)} \\
\text{s.t.} \quad & 0 = x_0^\tau \le x_1^\tau \le \dots \le x_k^\tau \le x_{k+1}^\tau = L && \forall \tau \in T \\
& x_a^\tau - x_a^{\tau-1} \le 1 && \forall a \in A \;\; \forall \tau \in T \\
& -x_a^\tau + x_a^{\tau-1} \le 1 && \forall a \in A \;\; \forall \tau \in T \\
& x_a^\tau \in [0, L] && \forall a \in A \;\; \forall \tau \in T
\end{aligned}$$

Note that the second and third set of constraints together ensure that the distance that each agent $a \in A$ (more precisely, the a-th leftmost agent) moves in each time step $\tau \in T$ is at most 1, i.e., $|x_a^\tau - x_a^{\tau-1}| \le 1$.

(QP1) is a linear program with a convex minimization objective and can thus be solved in polynomial time, e.g., by using standard interior point methods (see [20]). Let $\nu \in \tilde{E}^{A \times T}$ be the continuous solution that we obtain by solving (QP1). If we then round ν according to our Rounding Theorem, we obtain a discrete solution which is a 10-approximation for the corresponding U-TkMP problem on the path graph. This completes the proof of Theorem 6.

5.2 Cycle Case

We show that ideas similar to the ones used for the path case above, can be used to compute an optimal solution for the continuous problem CU-TkMP if the underlying topology is a cycle of length L, i.e., $\tilde{E} = ([0, L] \mod L)$.[2]

We obtain the following result from our Rounding Theorem (with $\beta = 1$ and $\Delta = 2$):

Theorem 7. *Let $n \geq 2k$. Then there is a 10-approximation algorithm for U-TkMP on the cycle graph $\mathcal{O}(convex(k_1, k_2) + n^3)$ time, where $k_1 = k_2 = \mathcal{O}(k\omega)$.*

As before, given a configuration $\sigma \in \tilde{E}^A$ we assume that the agents in $A = \{1, \ldots, k\}$ are indexed such that $\sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_k$. We introduce a single dummy agent $k+1$ whose position is defined as $\sigma_{k+1} = L + \sigma_1$. We can again assume that the agents will never want to switch index, because there exists an optimal solution in which they never switch. Note that we have $(\sigma_{k+1} \mod L) = \sigma_1$.

Lemma 7. *Consider a continuous configuration $\sigma \in \tilde{E}^A$ and suppose that the agents are indexed as stated above. Then the coverage cost is*

$$C(\sigma) = \frac{1}{4} \sum_{a=1}^{k} (\sigma_{a+1} - \sigma_a)^2.$$

Proof. As in the proof of Lemma 6, we subdivide the cycle into k line segments and argue for each segment separately. Using the same arguments as in the proof of Lemma 6, the cost of each line segment $[\sigma_a, \sigma_{a+1}]$ with $a \in \{1, \ldots, k-1\}$ is $\frac{1}{4}(\sigma_{a+1} - \sigma_a)^2$.

It remains to consider the (cyclic) segment $([\sigma_k, \sigma_{k+1}] \mod L)$ covered by agents k and 1. It is not hard to see that the coverage cost of this segment is

$$\int_{\sigma_k}^{\sigma_{k+1}} \min_{u \in \{\sigma_k, \sigma_{k+1}\}} |v - u| \mathrm{d}v = 2 \int_{0}^{\frac{1}{2}(\sigma_{k+1} - \sigma_k)} v \mathrm{d}v = \frac{1}{4}(\sigma_{k+1} - \sigma_k)^2.$$

Summing over all line segments proves the claim. □

In light of the above lemma, we can now adapt the convex program for the path case to the cycle case as follows:

$$\begin{array}{lllll} \min & \frac{1}{4} \sum_{\tau \in T} \sum_{a=1}^{k} (\sigma_{a+1} - \sigma_a)^2 & & & \text{(QP2)} \\ \text{s.t.} & 0 \leq x_1^\tau \leq \cdots \leq x_k^\tau \leq L & & \forall \tau \in T & \\ & x_{k+1}^\tau = L + x_1^\tau & & \forall \tau \in T & \\ & x_a^\tau - x_a^{\tau-1} \leq 1 & & \forall a \in A \ \ \forall \tau \in T & \\ & -x_a^\tau + x_a^{\tau-1} \leq 1 & & \forall a \in A \ \ \forall \tau \in T & \\ & x_a^\tau \in [0, L] & & \forall a \in A \ \ \forall \tau \in T & \end{array}$$

As before, we can solve this program efficiently and use our Rounding Theorem to obtain a 10-approximation for U-TkMP in the cycle case.

[2] Note that the modulo operator ensures that 0 and L correspond to the same point.

Algorithm 1: A 'mediate at $\sqrt{\omega}$'-algorithm for U-TkMP if $d(s,t) \leq 2\sqrt{\omega}$.

Input: U-TkMP-instance, α-approximation algorithm for k-MED
Output: Solution μ

1 Find the reachable set $B := \{v \in V : d(s,v) \leq \theta \wedge d(v,t) \leq \theta\}$, with $\theta := \lfloor \sqrt{\omega} \rfloor$.;
2 Find medians σ by approximating k-MED with facilities V and clients V.;
3 Take any bijection between agents and medians in σ, then form solution μ by minimizing each agent's summed distance to their median, using Lemma 1.;
4 Find medians $\sigma^{[B]}$ by approximating k-MED with facilities B and clients B.;
5 Take any bijection between agents to medians in $\sigma^{[B]}$, then form solution $\mu^{[B]}$ by minimizing each agent's summed distance to their median, using Lemma 1.;
6 **return** $\mu \leftarrow \arg\min\{D(\mu), D(\mu^{[B]})\}$;

6 Approximation Algorithms for General Graphs

Our Rounding Theorem can be used whenever we can solve (or approximate) the corresponding continuous version CU-TkMP of the problem. In this section, we derive approximation algorithms that do not rely on this rounding scheme. We first derive an $\mathcal{O}(\min\{\sqrt{\omega}, n\})$-approximation algorithm if the distance between s and t satisfies $d(s,t) \leq 2\sqrt{\omega}$. We then show that we can obtain an $\mathcal{O}(d(s,t) + \sqrt{\omega})$-approximation algorithm if $d(s,t) > 2\sqrt{\omega}$.

Our first algorithm is described in Algorithm 1. The idea of this algorithm is as follows: We first find approximate medians, and then try to stay as close as possible to these medians. We do not only try to find good medians for all of V, but also for a smaller subset B that we can cover well without moving more than $\sqrt{\omega}$ hops from s and t. If the 'optimal medians' are indeed within $\sqrt{\omega}$ hops from s and t, then the V-medians will give us a constant-factor approximation; otherwise, the B-medians will give us an $\mathcal{O}(\sqrt{\omega})$-approximation.

We assume we have access to an α-approximation algorithm that runs in $\mathcal{O}(kmed(n))$ time for k-MED for some given facility and client sets; e.g., like the 6-approximation by Jain and Vazirani [14].

Lemma 8. *Algorithm 1 runs in $\mathcal{O}(kmed(n) + n^5)$ time.*

Proof. Line 1 costs $\mathcal{O}(n)$ time. Lines 2 and 4 cost $\mathcal{O}(kmed(n))$ time. In lines 3 and 5, we seek $k \leq n/2$ walks, which according to Lemma 1, each cost $\mathcal{O}(n^4)$ time to find.

Theorem 8. *If there exists an optimal configuration σ^* that is within $\lfloor\sqrt{\omega}\rfloor$ hops from the starting configuration and the ending configuration, then Algorithm 1 is an $(\alpha + 2 + \sqrt{2})$-approximation that runs in $\mathcal{O}(kmed(n) + n^5)$ time.*

Proof. Let OPT be the optimal solution value, and ALG the value of the solution from Algorithm 1. In μ, the agents are at some configuration σ (as determined by the algorithm) throughout $\tau = \theta, \theta+1, \ldots, \omega+1-\theta$. We assumed that there is some optimal configuration $\sigma^* \subseteq B$, so $D(\sigma) \leq \alpha D(\sigma^*)$. With Lemma 3 and Corollary 1, we obtain

$$\begin{aligned} D(\mu) &\leq (\omega+1)D(\sigma) + n \cdot \sum_{\tau=1}^{\theta} \tau + n \cdot \sum_{\tau=1}^{\theta} \tau = (\omega+1)D(\sigma) + (\theta+1)\theta n \\ &\leq \alpha \cdot (\omega+1)D(\sigma^*) + \omega n + \sqrt{\omega} n \leq (\alpha + 2 + \sqrt{2})OPT \end{aligned}$$

where the last inequality holds because $\frac{\sqrt{\omega}n}{\frac{1}{2}(\omega+1)n} \leq \frac{2}{\sqrt{\omega}} \leq \sqrt{2}$ for $\omega \geq 2$ (which we can assume). We conclude that $ALG \leq D(\mu) \leq (\alpha + 2 + \sqrt{2})OPT$. □

Theorem 9. *Assuming $d(s,t) \leq 2\sqrt{\omega}$, Algorithm 1 is an $\mathcal{O}(\sqrt{\omega})$-approximation algorithm that runs in $\mathcal{O}(kmed(n) + n^5)$ time.*

Proof. For any configuration σ and $U \subseteq V$, define $D(\sigma, U) := \sum_{u \in U} d(\sigma, u)$. Let $V' := V \setminus B$ be the nodes that are not reachable in 2θ hops, with B as defined in Algorithm 1. Also let $\vec{D} := \sum_{v \in V'} \min_{u \in B} d(u, v)$ be the summed distance of V' to B. Let σ^* be an optimal configuration, and σ^*_B be a minimizer of $D(\sigma, B)$.

We can split the value of some optimal solution μ^* over its contribution to B and V'. The former can be lower-bounded with a configuration σ^*_B that minimizes $D(\sigma, B)$ instead of $D(\sigma)$. The latter can be lower-bounded by the fact that there can be no agents in V' if $\tau \leq \sqrt{\omega} - 1$ or $\tau \geq \omega + 1 - (\sqrt{\omega} - 1)$, so in those time-steps, at least a cost $\vec{D}$ is incurred. If $\vec{D} = 0$ or $\theta = 0$, then Theorem 8 gives us a constant-factor approximation, so assume $\vec{D} > 0$. Then

$$OPT \geq (\omega+1)D(\sigma^*, B) + D(\mu^*, V') \geq (\omega+1)D(\sigma^*_B, B) + (\sqrt{\omega} + \sqrt{\omega})\vec{D}.$$

In $D(\mu^{[B]})$, the agents will be at $\sigma^{[B]}$ throughout $\tau = \theta, \theta+1, \ldots, \omega+1-\theta$. By Lemma 3, we again know that the total cost cannot exceed $(\omega+1)D(\sigma^{[B]}) + (\theta+1)\theta n$. While we know that $\sigma^{[B]}$ is favorable for serving B, we need to upper-bound the cost of V' as well. For any $v \in V'$, if u is the nearest node in B to v, then we know that there is a median m at most 2θ hops away, because $d(u,m) \leq d(u,s) + d(s,m) \leq 2\theta$. So $D(\sigma^{[B]}) \leq D(\sigma^{[B]}, B) + \vec{D} + |V'| \cdot 2\theta$. We have

$$\begin{aligned} ALG &\leq D(\mu^{[B]}) \leq (\omega+1)D(\sigma^{[B]}) + (\theta+1)\theta n \leq (\omega+1)D(\sigma^{[B]}) + \omega n + \sqrt{\omega} n \\ &\leq (\omega+1)\left(D(\sigma^{[B]}, B) + \vec{D} + |V'| \cdot 2\sqrt{\omega}\right) + \omega n + \sqrt{\omega} n \\ \frac{ALG}{OPT} &\leq \frac{(\omega+1)D(\sigma^{[B]}, B)}{(\omega+1)D(\sigma^*_B, B)} + \frac{(\omega+1)\vec{D}}{(2\sqrt{\omega})\vec{D}} + \frac{(\omega+1)|V'| \cdot 2\theta}{|V'| \cdot \frac{1}{2}(\theta+1)\theta} + \frac{\omega n + \sqrt{\omega} n}{\frac{1}{2}(\omega+1)n} \\ &\leq \alpha + \frac{\omega+1}{2\sqrt{\omega}} + \frac{4(\omega+1)}{\sqrt{\omega}+1} + 2 + \sqrt{2} = \mathcal{O}(\sqrt{\omega}) \end{aligned}$$

□

Though it is true that $\sqrt{\omega}$ can become arbitrarily large, we can always apply Theorem 8 when $\sqrt{\omega}$ exceeds n. Therefore, these results together make Algorithm 1 an $\mathcal{O}(\min\{\sqrt{\omega}, n\})$-approximation algorithm.

Algorithm 2: A 'mediate at $\sqrt{\omega}$'-algorithm for U-TkMP if $d(s,t) > 2\sqrt{\omega}$.

Input: U-TkMP-instance, α-approximation algorithm for k-MED
Output: Solution μ

Find the reachable set $H := \{v \in V : d(s,v) + d(v,t) \leq d(s,t) + \theta\}$, with $\theta := \lfloor\sqrt{\omega}\rfloor$.;
Find medians σ by approximating k-MED with facilities V and clients V.;
form solution μ by minimizing each agent's summed distance to their median, using Lemma 1.;
Find medians $\sigma^{[H]}$ by approximating k-MED with facilities H and clients H.;
Take any bijection between agents to medians in $\sigma^{[H]}$, then form solution $\mu^{[H]}$ by minimizing each agent's summed distance to their median, using Lemma 1.;
return $\mu \leftarrow \arg\min\{D(\mu), D(\mu^{[H]})\}$;

Next, we consider the case $d(s,t) > 2\sqrt{\omega}$. Then it may occur that B is empty. In that case, we resort to Algorithm 2, the performance of which is $\mathcal{O}(d(s,t) + \sqrt{\omega})$. Analogous statements of Lemma 8 and Theorem 8 clearly hold for Algorithm 2, but we explain the weaker performance bound in Theorem 10.

Theorem 10. *Assuming $d(s,t) > 2\sqrt{\omega}$, Algorithm 1 is an $\mathcal{O}(d(s,t) + \sqrt{\omega})$-approximation algorithm that runs in $\mathcal{O}(kmed(n) + n^5)$ time.*

Proof. We follow a similar argument as in Theorem 9. That is, we separately bound the cost that H incurs in $\mu^{[H]}$ and that $\tilde{V}' := V \backslash H$ incurs. Denote $\vec{H} := \sum_{v \in \tilde{V}'} \min_{u \in H} d(u,v)$ the summed distance of $\tilde{V}'$ to H. Note first that any $u \in H$ and $v \in H$ can have distance at most $d(u,v) \leq d(s,t) + \sqrt{\omega}$:

$$2d(u,v) \leq d(u,s) + d(s,v) + d(u,t) + d(t,v) \leq 2(d(s,t) + \sqrt{\omega})$$

Let μ^* be an optimal solution, let $\sigma^{*[H]}$ be an optimal k-MED-configuration if the client set and facility set are both H, and let $\mu^{*[H]}$ be an optimal U-TkMP-solution for the induced subgraph $G[H]$. Note also that $OPT := D(\mu^*) \geq \sqrt{\omega}\vec{H}$, because for at least $\sqrt{\omega}$ time-steps, no agent can be in $\tilde{V}'$. Therefore:

$$ALG \leq D(\mu^{[H]}) = D(\mu^{[H]}, H) + D(\mu^{[H]}, \tilde{V}') \tag{5}$$

$$D(\mu^{[H]}, H) \leq (\omega+1)D(\sigma^{[H]}, H) + \frac{1}{2}(d(s,t) + \sqrt{\omega})(d(s,t) + \sqrt{\omega} + 1)n \tag{6}$$

$$(\omega+1)D(\sigma^{[H]}, H) \leq (\omega+1)\alpha D(\sigma^{*[H]}, H) \leq \alpha D(\mu^*) \tag{7}$$

$$\frac{D(\mu^{[H]}, H)}{OPT} \leq \frac{\alpha D(\mu^*)}{D(\mu^*)} + \frac{\frac{1}{2}n(d(s,t) + \sqrt{\omega})(d(s,t) + \sqrt{\omega} + 1)}{\frac{1}{2}n(\omega+1)} \tag{8}$$

$$\leq \alpha + \frac{d(s,t)^2 + 2d(s,t)\sqrt{\omega} + \omega + d(s,t) + \sqrt{\omega}}{\omega} \tag{9}$$

$$D(\mu^{[H]}, \tilde{V}') \leq (\omega+1)(\vec{H} + (d(s,t) + \sqrt{\omega})n) \tag{10}$$

$$\frac{D(\mu^{[H]}, \tilde{V}')}{OPT} \leq \frac{(\omega+1)\vec{H}}{\sqrt{\omega}\vec{H}} + \frac{(\omega+1)n(d(s,t) + \sqrt{\omega})}{\frac{1}{2}(\omega+1)n} \tag{11}$$

In (6), we use Lemma 3 repeatedly, knowing that movement occurs in at most $(d(s,t)+\sqrt{\omega})$ hops. In (10), we use the fact that the diameter of H does not exceed $d(s,t)+\sqrt{\omega}$. So if $v \in \tilde{V}'$ and $u = \min_{u \in H} d(u,v)$, then the cost contribution of v cannot exceed $d(u,v)+d(s,t)+\sqrt{\omega}$. So in every time-step, the nodes of $\tilde{V}'$ together cannot contribute more than $\vec{H}+n(d(s,t)+\sqrt{\omega})$ to the cost. Put together, this means that $\frac{ALG}{OPT} \leq \frac{D(\mu^{[H]},H)}{OPT} + \frac{D(\mu^{[H]},\tilde{V}')}{OPT}$ is an $\mathcal{O}(\sqrt{\omega}+d(s,t))$-approximation factor. □

Studying the approximation factor in the proof of Theorem 10, it would seem that Algorithm 2 performs poorly for high $d(s,t)$ and low ω. However, we remark that the most extreme version of this is observed only in the path graph, which is again 'easy'. It could be interesting to see if this 'gradual transition to path graphs' can help strengthen the performance bound.

7 Conclusions

We introduced TkMP and its uniform variant, U-TkMP. We provided several approximation algorithms for U-TkMP, including a novel method in which we round back an optimal or β-approximate solution to a 'continuous graph movement relaxation'.

As venues for future research, we suggest the following. There are likely more topologies than just path graphs and cycle graphs for which we can solve or approximate CU-TkMP. Extending the results of this paper to heterogeneous start and end locations seems imaginable, once an equivalent to Lemma 3 can be found that allows us to again lower-bound $D(\sigma)$ by some function of only n. (It seems more difficult to let go of the edge lengths being uniform, because this makes rounding back non-trivial, and we can no longer bound distances to unoccupied nodes to be between 1 and n.)

Acknowledgements. We thank the three anonymous reviewers for their valuable suggestions.

A Omitted proofs

A.1 Theorem 1

Proof. U-TkMP is obviously in NP. We will prove that U-TkMP is at least as hard as the Set Cover Problem, which is NP-hard. Let I_1 be the following instance of the decision version of the Set Cover Problem: given universe U and a collection S of subsets of U, can we find k (or less) members of S which have union U? Let I_2 be an instance of the decision version of U-TkMP that we construct as follows. Let the node set be $V := V_U \cup V_S \cup \{s\}$, where each element in U is represented by its own node in V_U, and each member in S is represented by its own node in V_S, and s is some additional node. In the graph $G=(V,E)$, we have $(u,v) \in E$ for nodes $u,v \in V$ exactly when either of these two conditions hold: $u, v \in V_S \cup \{s\}$, or $u \in V_U$ and $v \in V_S$ and the element that corresponds to u is contained in the subset that corresponds to v. In I_2, we set $\omega = 2$, and let all k agents start and end at s, and we ask: can we find a feasible solution with cost at most $2|S| + 4|U| + (n - k)$? I_2 can be constructed in polynomial time and space with respect to the input of I_1. In any feasible solution, all agents are at s at times $\tau = 0$ and $\tau = 2$, so we incur a cost of $|S| + 2|U|$. So the answer to I_2 is 'yes' if and only if we can find a configuration at $\tau = 1$ with cost $(n - k)$, meaning every node is adjacent to at least one agent position. The agents can only be in $V_S \cup \{s\}$, which is a clique. Adjacency to all nodes in V_U is achieved if and only if the chosen positions in V_S correspond to members of S which have union U. (If agents stay at s, this corresponds to choosing less than k subsets in I_1.) So any 'yes'-answer to I_1 supplies positions that will give a 'yes'-answer to I_2, and any 'yes'-answer to I_2 gives a 'yes'-answer to I_1 by reading out the positions at $\tau = 1$. So I_1 and I_2 are equivalent, implying U-TkMP is NP-hard. □

A.2 Theorem 2

Proof. Let $G = (V, E)$ be the original graph of our TkMP-instance. We construct a weighted, expanded graph $G^+ = (V^+, E^+)$ as follows. First, make a 'directed looping version' of G: take V, and for every $u, v \in V$, include arc (u, v) if edge $\{u, v\} \in E$ or $u = v$. Let this digraph be denoted by G'. Next, let $\vec{T}$ be the digraph with node set T, and arc set $\cup_{\tau \in T \setminus \{\omega\}} \{(\tau, \tau + 1)\}$, so $\vec{T}$ is a path digraph representing the time horizon.

Then, let G^+ be the tensor product $(G')^k \times \vec{T}$. So every node $(u_1, u_2, \ldots, u_k, \tau_1)$ in G^+ represents a position for each agent, and a time-step. There is an arc in G^+ from node $(u_1, u_2, \ldots, u_k, \tau_1)$ to node $(v_1, v_2, \ldots, v_k, \tau_2)$ if and only if the arc (u_a, v_a) is in G' for each agent a, and $\tau_2 = \tau_1 + 1$.

Finally, let each such arc in G^+ have weight $D(\sigma_2)$, where σ_2 is the visited configuration $(v_1, v_2, \ldots, v_k)$. This means that, walking through G^+, we incur at each time-step $\tau = 1, \ldots, \omega$ the cost of the configuration currently visited. The cost made at $\tau = 0$ is a constant that is the same for each feasible solution. Therefore, finding an optimal TkMP-solution is equivalent to finding a cheapest $(s_1, s_2, \ldots, s_k, 0) - (t_1, t_2, \ldots, t_k, \omega)$-path. Because G^+ has $\mathcal{O}(\omega n^k)$ nodes, finding this path with Dijkstra's algorithm has time complexity $\mathcal{O}(\omega^2 n^{2k})$. □

A.3 Lemma 1

Proof. Suppose first that $\omega \geq 2n$. Then, certainly, a walk exists from s to v and then t. We take a shortest (s, v)-walk and a shortest (v, t)-walk, and spend all remaining time 'looping' at v. This is obviously a walk π that minimizes $\sum_{\tau=0}^{\omega} d(\pi^\tau, v)$, because at every time-step, the agent is as near to v as the start and end conditions allow.

Suppose now that $\omega < 2n$. In this case, we construct G^+ exactly as in the proof of Theorem 2, setting $k = 1$. Each arc $((u, \tau), (u', \tau + 1))$ in G^+ now has weight $d(u', v)$. Because $\omega < 2n$, we find our walk in $\mathcal{O}(n^4)$ by applying Dijkstra's algorithm on G^+ from $(s, 0)$ to (t, ω). □

A.4 Theorem 4

Proof. We construct our solution as follows. First, if σ^* is an optimal k-MED-configuration, we use the α-approximation algorithm to find a good configuration σ with $D(\sigma) \leq \alpha D(\sigma^*)$, requiring a running time of $\mathcal{O}(kmed(n))$. Next, we assign these medians to agents arbitrarily. Finally, for each of the $k \leq n/2$ agents, we find in $\mathcal{O}(n^4)$ time the (s, t)-walk that minimizes the total distance to that agent's median, using Lemma 1. This adds a running time of $\mathcal{O}(n/2 \cdot n^4)$, for a total of $\mathcal{O}(kmed(n) + n^5)$.

We now prove that this yields an $(\alpha + 1)$-approximation. Let μ be the found solution, and μ^* an optimal solution. Obviously, μ consists of shortest paths from s to points in σ, a lot of waiting in σ, then shortest paths to t.

Observe any time-step τ. If $n - 1 \leq \tau \leq (\omega + 1) - (n - 1)$, then clearly all agents have reached their position in σ and have not departed them yet. In each of those time-steps, we have $D(\mu^\tau) \leq \alpha D(\sigma^*)$. In the other time-steps τ, we at least know that $D(\mu^\tau) \leq n^2$, because each of the n nodes has an agent at most n hops away.

So, recalling Corollary 1, we find

$$\frac{D(\mu)}{D(\mu^*)} \leq \frac{2n \cdot n^2}{\frac{1}{2}\omega n} + \frac{(\omega + 1 - 2n) \cdot \alpha D(\sigma)}{(\omega + 1) D(\sigma^*)} \leq \frac{2n \cdot n^2}{\frac{1}{2} \cdot 4n^2 \cdot n} + \alpha = \alpha + 1$$

meaning μ is an $(\alpha + 1)$-approximation. □

References

1. Arora, S., Raghavan, P., Rao, S.: Approximation schemes for Euclidean k-medians and related problems. In: Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing, pp. 106–113. Association for Computing Machinery (1998)
2. Bertsimas, D.J., Van Ryzin, G.: Stochastic and dynamic vehicle routing in the Euclidean plane with multiple capacitated vehicles. Oper. Res. **41**(1), 60–76 (1993)
3. Borodin, A., Linial, N., Saks, M.E.: An optimal on-line algorithm for metrical task system. J. ACM **39**(4), 745–763 (1992)
4. Byrka, J., Pensyl, T., Rybicki, B., Srinivasan, A., Trinh, K.: An improved approximation for k-median and positive correlation in budgeted optimization. ACM Trans. Algorithms **13**(2) (2017)
5. Farahani, R.Z., Abedian, M., Sharahi, S.: Dynamic facility location problem. In: Zanjirani Farahani, R., Hekmatfar, M. (eds.) Facility Location. CMS, pp. 347–372. Physica-Verlag HD, Heidelberg (2009). https://doi.org/10.1007/978-3-7908-2151-2_15
6. Galvão, R.D., Santibanez-Gonzalez, E.D.R.: A Lagrangean heuristic for the pk-median dynamic location problem. Eur. J. Oper. Res. **58**(2), 250–262 (1992)
7. Gendreau, M., Laporte, G., Semet, F.: A dynamic model and parallel tabu search heuristic for real-time ambulance relocation. Parallel computing **27**(12), 1641–1653 (2001)

8. Gendreau, M., Laporte, G., Semet, F.: The maximal expected coverage relocation problem for emergency vehicles. J. Oper. Res. Soc. **57**(1), 22–28 (2006)
9. Guha, S., Khuller, S.: Greedy strikes back: improved facility location algorithms. J. Algorithms **31**(1), 228–248 (1999)
10. Hotelling, H.: Stability in competition. In: The Collected Economics Articles of Harold Hotelling, pp. 50–63. Springer (1990). https://doi.org/10.1007/978-1-4613-8905-7_4
11. Huang, W.V., Batta, R., Babu, A.: Relocation-promotion problem with Euclidean distance. Eur. J. Oper. Res. **46**(1), 61–72 (1990)
12. Huizing, D., Schäfer, G., van der Mei, R.D., Bhulai, S.: The median routing problem for simultaneous planning of emergency response and non-emergency jobs. Eur. J. Oper. Res. **285**(2), 712–727 (2020)
13. Jagtenberg, C.J., Bhulai, S., van der Mei, R.D.: An efficient heuristic for real-time ambulance redeployment. Oper. Res. Health Care **4**, 27–35 (2015)
14. Jain, K., Vazirani, V.V.: Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and Lagrangian relaxation. J. ACM (JACM) **48**(2), 274–296 (2001)
15. Kariv, O., Hakimi, S.L.: An algorithmic approach to network location problems. II: the p-medians. SIAM J. Appl. Math. **37**(3), 539–560 (1979)
16. Koutsoupias, E.: The k-server problem. Comput. Sci. Rev. **3**(2), 105–118 (2009)
17. Koutsoupias, E., Papadimitriou, C.H.: On the k-server conjecture. J. ACM **42**(5), 971–983 (1995)
18. Lin, J.H., Vitter, J.S.: *e*-approximations with minimum packing constraint violation. In: Proceedings of the Twenty-Fourth Annual ACM Symposium on Theory of Computing, pp. 771–782. Association for Computing Machinery (1992)
19. Melo, M.T., Nickel, S., Da Gama, F.S.: Dynamic multi-commodity capacitated facility location: a mathematical modeling framework for strategic supply chain planning. Comput. Oper. Res. **33**(1), 181–208 (2006)
20. Nesterov, Y., Nemirovskii, A.: Interior-point Polynomial Algorithms in Convex Programming. Society for Industrial and Applied Mathematics (1994)
21. Tamir, A.: An $O(pn^2)$ algorithm for the p-median and related problems on tree graphs. Oper. Res. Lett. **19**(2), 59–64 (1996)
22. Teitz, M.B.: Locational strategies for competitive systems. J. Regional Sci. **8**(2), 135–148 (1968)
23. Wesolowsky, G.O.: Dynamic facility location. Manage. Sci. **19**(11), 1241–1248 (1973)