On-demand Robotic Fleet Routing in Capacitated Networks with Time-varying Transportation Demand

Martin Schaefer[a], Michal Čáp[b], David Fiedler[c] and Jiří Vokřínek [d]
Artificial Intelligence Center, Computer Science Department, FEE, Czech Technical University in Prague, Czech Republic

Keywords: Multi-robot Systems, Fleet Routing, Coordination, Mobility-on-Demand System.

Abstract: In large-scale automated mobility-on-demand systems, the fleet manager is able to assign routes to individual automated vehicles in a way that minimizes formation of congestion. We formalize the problem of on-demand fleet routing in capacitated networks with time-varying demand. We demonstrate the limits of application of the steady-state flows approach in systems with time-varying demand and formulate a linear program to compute congestion-free routes for the vehicles in capacitated networks under time-varying demand. We evaluate the proposed approach in the simulation of a simplified, but characteristic illustrative example. The experiment reveals that the proposed routing approach can route 42% more traffic in congestion-free regime than the steady-state flow approach through the same network.

1 INTRODUCTION

Almost all world's metropolitan areas are plagued with traffic congestion that emerges when the number of vehicles attempting to travel along a specific road segment exceeds its capacity. Such a failure to ensure free traffic flow in an urban road network inevitably results in significant travel delays and higher emission of air pollutants. The levels of traffic congestion can be reduced either by increasing the capacity of the roads or by selecting routes for individual vehicles that use the available capacity more efficiently.

Today, on-line navigation tools, such as Google Maps or Waze, collect and share real-time data about travel delays to compute and recommend the fastest route to each of their users. The situation where every driver is perfectly informed about the delays and follows a route that optimizes its arrival time is described by so-called user equilibrium traffic assignment. On the other hand, the situation, where the individual vehicles are centrally assigned routes that optimize average travel time is described by so-called system optimal traffic assignment. In practical terms, the system-optimal assignment of traffic to road network was considered hard to achieve, because some of the drivers would have to voluntarily ”sacrifice” and follow longer routes to improve the overall system performance.

Recently, the self-driving technology emerged as a possible enabler for low-cost on-demand urban mobility. In fact, it has been argued (Spieser et al., 2014) that a fleet of shared automated vehicles is capable of providing personal point-to-point mobility at lower-cost than driving a privately-owned car. Such automated mobility-on-demand (MoD) systems consists of a large, centrally controlled fleet of shared self-driving vehicles that jointly service transportation needs of the users of the system. In contrast to conventional traffic, the routes for the robotic vehicles of the system can be coordinated centrally and thus the system-optimal assignment of routes to the road network can be, in principle, achieved. In result, such centrally-routed mobility-on-demand systems hold promise to improve the efficiency of utilization of existing road infrastructure, i.e., the system serves more transportation demand and at the same reduces or even completely prevents formation of congestion. In order to evaluate the potential of system-optimal fleet routing, it is important to understand how to compute the system-optimal routes for the fleet.

In the system-optimal on-demand fleet congestion-free routing problem, we seek coordinated routes for on-demand vehicles that serve the passengers' transportation requests such that the

[a] https://orcid.org/0000-0003-0578-3182
[b] https://orcid.org/0000-0001-6450-7976
[c] https://orcid.org/0000-0001-5374-1089
[d] https://orcid.org/0000-0003-3294-3030

Schaefer, M., Čáp, M., Fiedler, D. and Vokřínek, J.
On-demand Robotic Fleet Routing in Capacitated Networks with Time-varying Transportation Demand.
DOI: 10.5220/0010261009070915
In *Proceedings of the 13th International Conference on Agents and Artificial Intelligence (ICAART 2021) - Volume 2*, pages 907-915
ISBN: 978-989-758-484-8

capacities of the road segments are not exceeded and the average travel delay is minimized. Note that in contrast to the problem of system-optimal route assignment to privately-owned vehicles, where each passenger is tied to its own vehicle, in robotic fleet routing problem we can also optimize over the possible assignments of passengers (transportation requests) to individual shared robotic vehicles.

Traditionally, the research field of vehicle routing problems (VRP) (Pillac et al., 2013) has focused on finding the optimal assignment of transportation requests to vehicles in a fleet. Yet, fleets modeled in VRPs are typically small and thus the road capacity constraints are usually not modeled. However, when the controlled fleet consist of tens of thousands of vehicles, the capacity of the underlying road network must be considered. This led to emergence to the study of the problem of on-demand fleet routing in capacitated networks. In the seminal work of (Zhang et al., 2016a) and then in (Schaefer et al., 2019), (Solovey et al., 2019), and (Wollenstein-Betech et al., 2020), the on-demand system is modeled in the framework of network flows and analyzed under the steady-state assumption, i.e., the system-optimal routes are computed while the intensity of transportation demand is time-invariant.

However, one of the defining properties of urban transportation is that the transportation demand is highly time-dependent, both in its intensity and in its structure. Indeed, a typical urban passenger transportation demand is characterized by a morning and an evening peak: in the morning, the majority of passengers desires to travel from residential areas to business areas, while in the evening, the transportation demand is oriented in the opposite direction.

In this paper, we analyze the applicability of the steady-state approach to fleet routing in periods when the transportation demand changes in time, e.g., during the onset of morning peak. We demonstrate that the naive strategy of steady-state approach is limited and propose an extended model that accounts for time-varying demand and discuss research challenges that need to be tackled to allow system optimal fleet route planning in urban networks with time-varying transportation demand of practical sizes. In the simulated experiments with congestion model, we demonstrate the dominance of the proposed approach over the steady-state one in presence of the time-varying transportation demand. We show, that the proposed method can find congestion-free routes for by 42% more demand than the steady-state approach in the same illustrative road network.

2 PROBLEM DESCRIPTION

In this section, we state the on-demand fleet routing in capacitated network problem (OFRCNP). Consider an mobility-on-demand system consisting of a fleet of single-occupancy vehicles that jointly serve a point-to-point transportation requests over a road network. We assume that there is no other traffic, i.e, the road network is exclusively occupied by the centrally-controlled fleet.

The road network is modeled by a directed graph, where the nodes represent junctions or parking areas and the edges represent road links. The road links have time-invariant capacities describing the maximal flow rate along the link measured in the number of vehicles per time unit. We assume that if the flow over the link is below its capacity, the congestion will not occur and consequently the transit times over the link is time-invariant and deterministically known. The flow over the link is strictly caped by the capacity and any exceeding flow is not allowed.

We adopt a common assumption for the road networks that the nodes are without storage, i.e., we desire a vehicle to circulate in the network and to pick-up passengers by visiting the pick-up nodes without waiting there. Subsequently, the vehicle transports the passenger to the drop-off node from where the vehicle immediately continue to another passenger's pick-up node. In other words, the vehicles are not allowed to accumulate in any node, except in a set of pre-defined depots that model high capacity parking lots.

The transportation demand is deterministic and known apriori. It is described by a collection of transportation requests, each specified by the origin node, the destination node, the earliest pick-up time, and the latest drop-off time. We assume the passengers are willing to wait for pick-up as long as they are dropped-off before the drop-off deadline. The number of travel requests and a distribution of origins and destination can vary over time.

The goal is to plan a route for each of the fleet vehicles through the road network so that each transportation request is serviced within provided time constraints and at the same time ensure that link capacities are not exceeded. The optimization criterion is the cost of fleet operation, i.e., the total travelling time by the vehicles in the fleet. The quality of service is enforced by the drop-off time constraint of each transportation request.

3 NETWORK FLOW MODEL

We observe that OFRCNP is a multiple single-occupancy vehicle routing problem with pickup and deliveries, time windows and additional constraints on maximal allowed vehicle flow over road links. We note that the above problem is a generalization of the single vehicle with unit-occupancy pickup and delivery problem (Berbeglia et al., 2007), commonly referred to as a Stacker Crane Problem, which is known to be NP-hard (Frederickson et al., 1976).

Moreover, the fleets can easily contain tens of thousands of vehicles. For example, an mobility-on-demand system sufficient to serve the existing transportation demand in Prague with more than million inhabitants is estimated to require 60 000 of vehicles (Fiedler et al., 2017). Such fleet sizes are out of reach even for most heuristic VRP approaches.

One way to mitigate the complexity is to aggregate the transportation requests between the same origin and destination into a demand flow and vehicles that serve this demand flow into a vehicle flow. The problem can be then cast as a min-cost multi-commodity flow problem. This problem is in general still intractable, but the complexity can be eliminated with additional assumptions.

Firstly, due to indivisibility of individual vehicles the flow of vehicles is integral (it takes integer values), so the problem in consideration is in fact an integer multi-commodity flow problem, which is however NP-hard (Even et al., 1975). The relaxation to fractional (real-valued) flows is convenient as the fractional multi-commodity flow problem can be solved in polynomial time. In practice, it is sufficient to convert resulting fractional flows to integral flows by randomized rounding (Rossi et al., 2016).

Secondly, the demand is time-variant and thus the system should be properly modeled in the framework of flows over time. The explicit modeling of time can be avoided when the system is locally modeled as being in a steady-state, i.e., the demand flows are time-invariant. This approach has been proposed in (Zhang et al., 2016a) and it is argued to be applicable if the demand changes slowly relative to the average travel duration in the system. In practice, the steady-state formulation could be solved iteratively at different times of the day and the solution from the most recent computation is used to derive the routes for the vehicles in the fleet.

In the next section we introduce the formulation of steady-state flow model that is then extended to the dynamic model. The approaches are demonstrated and compared on an illustrative scenario.

4 STEADY-STATE APPROACH

In this section, we formally state the steady-state network flow model of MoD system that roughly corresponds to the approach introduced in (Zhang et al., 2016b). The road network is represented by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the node set and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the edge set. Nodes correspond to road junctions and edges to the road links. The capacity of a road link $(u, v) \in \mathcal{E}$ is denoted as $c(u, v) \in \mathbb{Z}_{>0}$ and corresponds to a maximal flow rate of vehicles that can enter the road link per time unit to keep free flow speed. We denote $\tau_{(u,v)} \in \mathbb{Z}_{\geq 0}$, as free flow traversal time in time units over edge (u, v).

The capacities and free flow traversal times are considered to be static. The set of demand flows (aggregated transportation requests) is $\mathcal{M} = \{(s_m, g_m)\}_m$, where for the m-th demand flow, $s_m \in \mathcal{V}$ is the origin of the flow, $g_m \in \mathcal{V}$ is the goal destination, and the intensity of the demand flow is λ_m, in passengers per time unit. The cardinality of the set $\mathcal{M}$ is denoted by M.

There are two types of flows considered: passenger flows and rebalancing flows. The passenger flows correspond to parts of vehicle routes with passengers on board and the rebalancing flows correspond to the parts of routes when vehicles drive without passengers. The flow rate on the edge $(u, v) \in \mathcal{V}$ for the passenger flow m is denoted as $f_m(u, v) : \mathcal{E} \times \mathcal{M} \mapsto \mathbb{Z}_{\geq 0}, (u, v) \in \mathcal{E}, m \in \mathcal{M}$. Further, the rate of rebalancing flow entering edge (u, v) is denoted by $f_R(u, v) : \mathcal{E} \mapsto \mathbb{Z}_{\geq 0}$.

For the modeling purposes, we introduce additional virtual nodes and edges. We add a virtual demand source node s_m^V for each $m \in \mathcal{M}$. Each s_m^V is connected to G by edge (s_m^V, s_m) that allows the serviced demand to join the transportation network.

Similarly, virtual demand sink nodes g_m^V and edges (g_m, g_m^V) are added. The edges (g_m^V, s_m^V) are added to virtually close the demand into a loop, which simplifies flow conservation constraints. The vehicles can park only in a set of depots $\mathcal{D} \subseteq \mathcal{V}$.

For each depot $v \in \mathcal{D}$ there is a virtual edge (dg_v^V, ds_v^V) representing the depot at node v that is connected to the graph G so that it allows the vehicles to join or to leave the transportation network, i.e., an edge (ds_v^V, v) and an edge (v, dg_v^V), $\forall v \in \mathcal{D}$.

Then, the set of all virtual nodes is $\mathcal{V}^V$ and the set of all virtual edges is $\mathcal{E}^V$.The resulting augmented graph is $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$, where $\mathcal{V}' = \mathcal{V} \cup \mathcal{V}^V$ and $\mathcal{E}' = \mathcal{E} \cup \mathcal{E}^V$.

Now we are in position to cast the problem of fleet routing in capacitated networks as an instance of minimum cost multi-commodity flow problem:

Problem 1. Steady-state OFRCNP. The task is to minimize:

$$\sum_{(u,v)\in\mathcal{E}'} \tau(u,v)\Big(\sum_{m\in\mathcal{M}} f_m(u,v) + f_R(u,v)\Big), \tag{1}$$

subject to:

$$\sum_{u\in\mathcal{V}'} f_m(u,v) = \sum_{w\in\mathcal{V}'} f_m(v,w), \tag{2}$$

$$\forall v\in\mathcal{V}', \forall m\in\mathcal{M}, \tag{3}$$

$$\sum_{j\in\mathcal{M}} f_j(g^V_m, s^V_m) = \lambda_m, \quad \forall m\in\mathcal{M}, \tag{4}$$

$$\sum_{v\in\{v:(s^V_m,v)\in\mathcal{E}')\}} f_m(s^V_m, v) = \lambda_m, \quad \forall m\in\mathcal{M}, \tag{5}$$

$$\sum_{m\in\mathcal{M}} f_m(u,v) + f_R(u,v) \le c(u,v), \quad \forall (u,v)\in\mathcal{E}, \tag{6}$$

$$\sum_{m\in\mathcal{M}} 1_{(v,g^V_m)\in\mathcal{E}'} f_m(v,g^V_m) + \sum_{u\in R_{S,v}} f_R(u,v) = \tag{7}$$

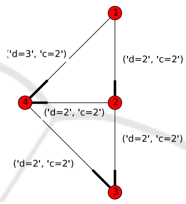
$$\sum_{m\in\mathcal{M}} 1_{(s^V_m,v)\in\mathcal{E}'} f_m(s^V_m,v) + \sum_{w\in R_{G,v}} f_R(v,w), \quad \forall v\in\mathcal{V}.$$

The objective is to minimize the total traveling time of the vehicles, i.e., the cost of the fleet operation. Note that, it also means that the average travel time of vehicles is minimized. The constraints enforce that the passenger flows are conserved and that the intermediate storage in junction nodes is forbidden (2), request flows are satisfied and consistent (4, 5), rebalancing flows in the road network transform to passenger flows, and vice versa, without loss[1] (7)[2], and the capacity of road links is not exceeded (6).

5 DEALING WITH TIME-VARYING DEMAND

The steady-state approach hinges on the assumption that the demand intensity is time-invariant. This assumption is however often not justified when demand patterns change rapidly, which occurs in practice in the transient periods around the traffic peaks. For example, between 6 a.m. and 7 a.m., the traffic intensity increases by 60% in the city of Prague (The Technical Administration of Roads of the City of Prague, 2020). To account for the time-varying demand, we extend the demand formulation by time. The set of demand flows (aggregated transportation requests) is then $\mathcal{M} = \{(s_m, g_m, t_m, d_m)\}_m$, where for the m-th demand flow, $s_m \in \mathcal{V}$ is the origin of the flow, $g_m \in \mathcal{V}$ is the goal destination, t_m is the earliest time when the requests in the flow can be served and d_m is the latest time to drop-off.

In the following, we will use a simplified scenario to demonstrate in what situations can the steady-state approach fail. The scenario represents phenomena that may occur during the onset of afternoon traffic peak. Initially, we introduce the scenario, we describe how the steady-state approach deals with the time-varying demand. Then, when the limitations of the approach are shown, we introduce the flows over time approach to address the limitations of the former approach.



node / t	1	2	3	4
north-south	2	0	0	0
center-south	0	0	1	2

(a) (b)

Figure 1: *Road Network A* (a), demand over time (b).

5.1 Example Scenario

In Fig. 1a, we show a simplified road network, representing a city with central commercial area at Node 2 and north, south and west peripheries at N. 1, N. 3, and N. 4 respectively. The road network has a ring-radial topology, but we only consider the west half of the network and the north-to-south road directions to simplify the visualization. Also, for a matter of a clear demonstration of the influence of time-varying demand, we let each of the nodes be a depot node, i.e., no rebalancing is needed in these networks, the vehicles are immediately available in every node and can park in every node where a passenger gets off. In result, all the flows in the scenario are the passenger flows and we can easily observe and compare the passenger routes found by the compared algorithms. Also, we consider time to be discretized into a set of time steps $\mathcal{T} = \{1, 2, 3, \ldots\}$ and denote $T = |\mathcal{T}|$.

We consider two south-bound demand flows. North-south flow ($1 \to 3$) represents regular constant flow of through traffic that is present all day (in our example, it is active only at $t = 1$ because longer flows are difficult to visualize). The second, center-south flow ($2 \to 3$) represents commuters' flow leaving their

[1]Note that rebalancing flows origin in depot nodes or by dropping off passengers and sink by picking-up passengers.

[2]Function 1_x denotes the indicator of the Boolean variable $x = \{true, false\}$, that is 1_x equals one if x is true, and equals zero is x is false. For node v, the set $R_{S,v} = \{u : (u,v)\in\mathcal{E}' \land u \ne s^V_m \forall m\in\mathcal{M}\}$ and set $R_{G,v} = \{w : (v,w)\in\mathcal{E}' \land w \ne g^V_m \forall m\in\mathcal{M}\}$

work at the central commercial area that increases during afternoon and leads to a traffic peak. The demand flow intensities are time-varying as shown in Fig. 1b.

5.2 Applying Steady-state Approach

In this section, we will discuss an application of a steady-state approach on the road network with time-varying demand. The idea is to ignore the dynamism of demand and periodically recompute the solution under the assumption that the demand intensities observed at the current time are identical to past intensities and will remain constant in future. The steady-state solution is computed and all vehicles departing at the time step follow the corresponding steady-state flows solution computed using the current intensities. The vehicles are assigned a route at the time of their departure and this route remains unchanged until they reach their destination. This strategy is applied, for example, in (Zhang et al., 2016b).

We implemented the approach by repeatedly solving the steady-state flow problem in each time step. That is, we generated a linear program corresponding to Problem 1 and solved it using the CBC solver. The solution flows are visualized in Fig. 2a using a unit-length graph in which the original edges are split into unit-length edges. Then, the transit time of each edge in this graph is equal to one, which allows us to visualize the flows in time in the network by a sequence of the figures. The flow under the link capacity is depicted in blue, if the capacity is exceeded the proportional part of the flow is colored in red. The labels illustrate how much of the link capacity is consumed by the flow.

The steady-state approach operates as follows. In each time step, current demand flows are routed by min cost flows while ignoring previous flows and future demand. Path of each vehicle is fixed as soon as it departs.

The Fig. 2a illustrates the limit of the *steady-state approach* in the presence of a time-varying demand on the *Road network A*. We observe that at the time of the change in the demand, i.e., the increase of demand in the commercial center (N. 2), the approach ignores previous flows and sends the flow simultaneously with the north-south flow (edge $(2,3)$). In result, the flows exceed the road capacity.

We have demonstrated that *steady-state approach*, that ignores the dynamism of transportation demand, cannot be easily applied to the scenario with time-varying demand. An approach that allows finding a solution to scenario like *Road Network A* is to consider flows over time.

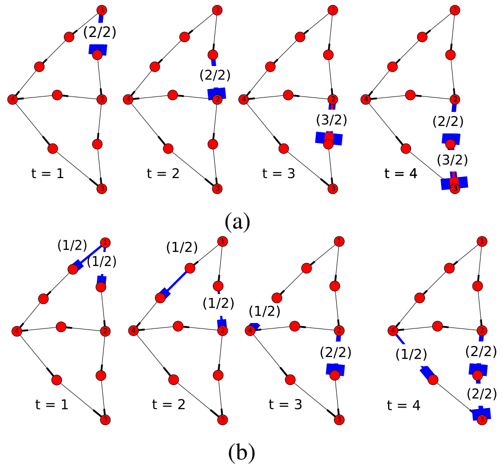


Figure 2: *Steady-state approach* (a) and *Dynamic approach* (b) solution for *Road Network A*.

6 FLOWS OVER TIME: DYNAMIC APPROACH

The main idea of the proposed *dynamic approach* is to construct a so-called time-expanded graph (Ford and Fulkerson, 1958) and apply the *steady-state approach* over such a graph. The resulting flows on the time-expanded graph then represent spatio-temporal routes on the original road network.

The time-expanded graph (Ford and Fulkerson, 1958) consists of T copies (layers) of nodes of the original road graph. The layers are connected according the traversal times of the corresponding edge. The time-expanded graph $\mathcal{G}^{\mathcal{T}} = (\mathcal{V}^{\mathcal{T}}, \mathcal{E}^{\mathcal{T}})$ is defined as $\mathcal{V}^{\mathcal{T}} = \{(u,t) : u \in \mathcal{V}, t \in \mathcal{T}\}$, $\mathcal{E}^{\mathcal{T}} = \{((u,t),(v,t+\tau(u,v)) : (u,v) \in \mathcal{E}\}$. Note that the capacities are static, i.e., $c((u,t_1),(v,t_2)) = c(u,v) \forall t_1,t_2 \in \mathcal{T}$ and $(u,v) \in \mathcal{G}$.

Additional virtual nodes are connected to the graph to model passenger waiting. We add a virtual demand source node s^V_m for $m \in \mathcal{M}$. The virtual source nodes are connected to the G^T by $(s^V_m,(s_m,t)), \forall t \in \{t \in \mathcal{T} : t_m \leq t \leq d_m\}, \forall m \in \mathcal{M}$, to let serviced demand to join the transportation network. Additionally, a virtual demand sink nodes g^V_m and edges $((g_m,t),g^V_m) \quad \forall t \in \{t \in \mathcal{T} : t_m \leq t \leq d_m\} \forall m \in \mathcal{M}$ are added. The edges $(g^V_m,s^V_m) \forall m \in M$ virtually close the demand circulation for simpler flow conservation constraints. The vehicles can be parked in depots $\mathcal{D} \subseteq \mathcal{V}$. Each depot $v \in \mathcal{D}$ generates edge (dg^V_v, ds^V_v) that is connected to the graph G^T so that it allows the vehicles to join or to leave the transportation network at each time, i.e., an edge $(ds^V_v,(v,t))$ and an edge $((v,t),dg^V_v) \forall v \in \mathcal{D} \forall t \in \mathcal{T}$. The set of virtual nodes is $\mathcal{V}^V$ and the virtual edges $\mathcal{E}^V$ The re-

Table 1: Cost and congested flow: *Road Network A*.

capacity constraints	**strict**	**relaxed**
Steady-state approach	infeasible	14 (6/4)
Dynamic approach	15	15 (0/0)

sulting graph is $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$, where $\mathcal{V}' = \mathcal{V}^T \cup \mathcal{V}^V$ and $\mathcal{E}' = \mathcal{E}^T \cup \mathcal{E}^V$. The traffic flows over time are defined analogically as steady-state flows, the difference is that these are defined on the time-expanded graph G^T. Finally, we define the solution of the on-demand fleet routing over time on road graph G to be the solution of the steady-state on-demand fleet routing over the time-expanded graph G^T. The described approach is later referred as the *dynamic approach*.

6.1 Dynamic Approach Solution

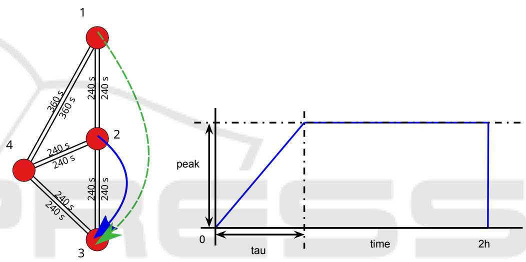
The model of flows over time, in contrast with the *steady-state approach*, considers both previous and predicted future flows. Fig. 2b shows the optimal solution of the Problem 1 with time provided by the *dynamic approach*. The *steady-state approach* ignores previous flows, so when the partial flows are put together in time the resulting flows can exceed the road capacity. We refer to such a solution as being *infeasible*. We already showed that for *Road Networks A*, the solution of the *steady-state approach* is infeasible.

We compare the solution quality of the considered approaches on the *Road Network A* in Table 1. We compare the approaches in two settings: 1) we consider problems strictly as defined, 2) we relax the road capacity constraints if no other feasible solution exists. In the strict setting, we can confirm that the basic steady-state approach that ignores previous flows is infeasible. The dynamic approach finds the optimal feasible solution in the both settings.

If we relax the capacity constraints in the case no feasible solution is found, we can observe lower costs in *steady-state approach*, but the rate of vehicle flows may exceed the capacity constraints of the road links. The exceeded capacity is described in brackets in the form (F/C) where F is a the sum of flows on the edges with exceeded capacity and the C is a capacity constraint over such edges. The values reveal by how much are the capacities exceeded, which also hints on the severity of congestion that would appear in practice. In Table 1 we observe that the *steady-state approach* can assign flows that consume up to 150% of the road capacity.

7 SIMULATION EVALUATION

In this section, we will evaluate the performance of the proposed fleet routing approach in simulation. For evaluation, we use an abstract model of a part of a city highway network shown in Fig. 3a. The road network connects the city center (N. 2) to three peripheries (N. 1, 3, 4) of a small metropolitan area. We study the interaction of a through-traffic demand flow (N. 1 $\rightarrow$ N. 3) with a local-traffic demand flow (N. 2 $\rightarrow$ N. 3) during morning traffic peak. The intensity of transportation demand in each of these demand flows increases in time up to a saturation level according to a trapezoid impulse function depicted in Fig. 3b. The shape of the demand "impulse" is controlled by *tau* and *peak* parameters. The former represents the time to reach full saturation, measured in seconds. The latter represents the intensity at full saturation relative to the free-flow capacity of a road segment.



(a) (b)

Figure 3: Test Scenario. Road network and the two demand flows (left) and the shape of demand "impulse" (right).

7.1 Congestion Model

The travel velocity on the individual edges is modeled as density-dependent. We use an empirical two-regime link delay model for highway segments. The vehicles are able to drive at free-flow speed until the vehicle density at the link exceeds so-called free-flow capacity. In our case, the free-flow density is 20 veh./km, which corresponds to the flow of 0.55 veh./sec. at 100 km/h. At this point, the traffic enters so-called bound flow regime and the vehicles start decreasing their velocity. One of the crucial empirical observations made by transportation researchers is that, initially, the density of vehicles increases slightly faster then their speed reduces and thus the vehicle flow is typically observed to increase until the density reaches so-called critical capacity. The critical capacity in our model corresponds roughly to the density of 25 veh./km and flow of 0.56 veh./sec. at 80 km/h. After the critical capacity has been exceeded, both the velocity and the flow of traffic start rapidly declining. At this point, the system enters an unstable regime with self-reinforcing feedback loop that eventually leads to build-up of standstill or slowly moving queues.

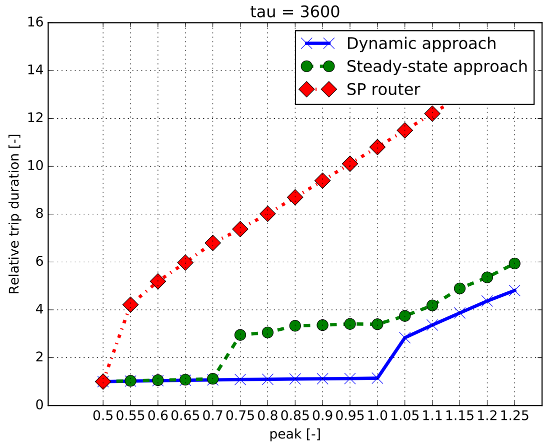
7.2 Experiment Setup

We compare the performance of three fleet-routing strategies: In *shortest-path approach*, all vehicles are routed along the shortest path from the origin of each demand request to its destination. In *steady-state approach*, we compute the routes by sequentially solving the steady-state formulation as described in Section 5.2. In *dynamic approach*, we compute the routes by solving the flows over time formulation as described in Section 6. The capacity constraint of each road segment is set to correspond to free-flow capacity of that edge, i.e, it is set to 0.55 veh./sec. In reality, the edge free-flow capacities might be too restrictive for high demand, and cause the approaches to fail to provide a feasible solution. For such cases, the capacity constraints are modeled as elastic constraints. The capacity can be exceeded for additional penalty that is linear in the capacity excess and the link free-flow traversal time. The penalty enforces that the capacity is exceeded only if no other feasible solution exists.

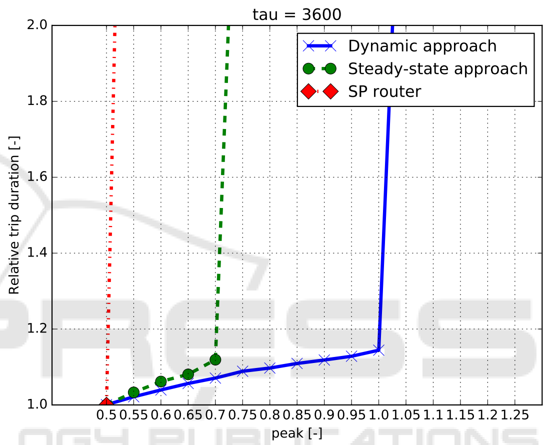
We generate two hours of through-traffic and local-traffic transportation demand discretized to 2 min. timesteps following trapezoidal shape with different combinations of *tau* and *peak* parameters. In particular, we create the combination such that *tau* takes values 20 min, 30 min, 45 min, and 60 min and *peak* takes values 0.5, 0.55, …, 1.25. This will allow us to study how does the dynamism and the scale of the demand affect the performance of the individual approaches. Then we let each approach to generate routes for all the vehicles serving the transportation demand. Finally, we simulate the movement of individual vehicles in the multi-agent traffic simulator Agentpolis employing the congestion model described in the previous section and record the travel delay of each vehicle.

7.2.1 Results

The average travel prolongation achieved by the three compared methods as a function of *peak* parameter is shown in Fig. 4 for slowly (1 h to saturation) evolving demand. To correctly interpret the results, it is useful to observe that there are two key thresholds for *peak* parameter. Firstly, since we are only working with two demand flows, for *peak* ≤ 0.5, it is impossible to exceed the link capacity (unless some vehicle travel along some segment multiple times) and thus all three approaches maintain zero prolongation. On the other



(a) Overall view



(b) Zoomed view on prolongation values for under-capacity flows

Figure 4: Average Travel Prolongation vs. *peak*, *tau* = 60 min.

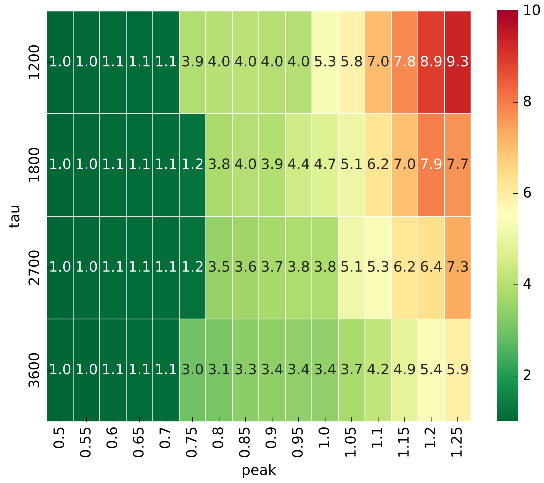
hand, for *peak* > 1, the link capacities will be necessarily exceeded by any routing policy. Thus, we are interested in the ability of the three routing policies to maintain reasonable delay in between these two extremes.

As we can see, the congestion-unaware shortest path policy routes both demand flows through the center and thus the traffic on link $(2,3)$ quickly exceeds the critical intensity and enters the unstable, highly congested regime, resulting in extremely high travel delays. The *steady-state approach* initially serves the through-traffic demand flow by two vehicle flows, one going through the center, the other one routed through the periphery. Since a portion of the vehicles travel via a route that is 25 % longer than the shortest route, the average travel prolongation is slightly increased. The fact that *steady-state approach* does not account for future changes in demand results in excess traffic on link $(2,3)$. This undesirable excess flow is larger for larger values of *peak* and for faster change in demand, i.e, for smaller values of *tau*. The flow eventually grows large enough to exceed the critical capacity, leading again to collapse of the traffic flow and subsequently to significant travel delays. Since the traffic collapses only for the through-flow routed via the city center and the vehicles driving via the periphery remain traveling at free-flow speed, the average prolongation remains smaller than the average prolongation for the shortest path strategy.

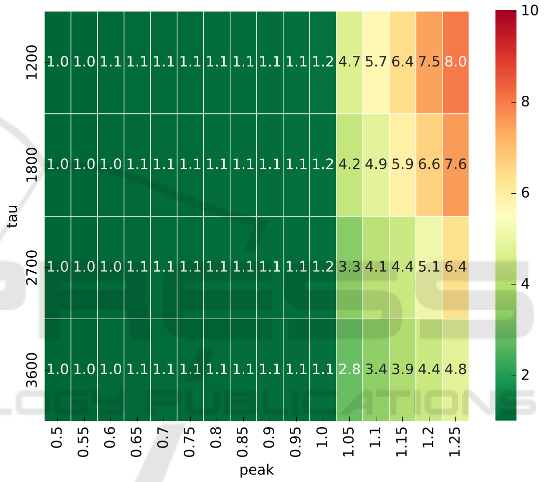
The dynamic approach successfully accounts for anticipated future demand by preemptively routing part of the through-flow via the periphery. Although this results in up to 1.2x travel time prolongation, the algorithm maintained sub-critical flows in the system for all peak values < 1.0. For completeness, the comparison of average travel delay for *steady-state* and *dynamic approach* for all combinations of values for *tau* and *peak* parameters are depicted in Fig. 5b. The same phenomena can be observed.

These results suggest that system optimal fleet routing that explicitly addresses time-varying demand has a potential to significantly increase the amount of transportation demand that can be efficiently transported through existing road infrastructures. Indeed, compared to the shortest path approach, the proposed technique can be used to service twice as many transportation request in the same road network without worsening the congestion. In comparison to the *steady-state approach*, the *dynamic* method is capable of servicing 42 % extra demand (1.0 vs. 0.7) through the same road network.

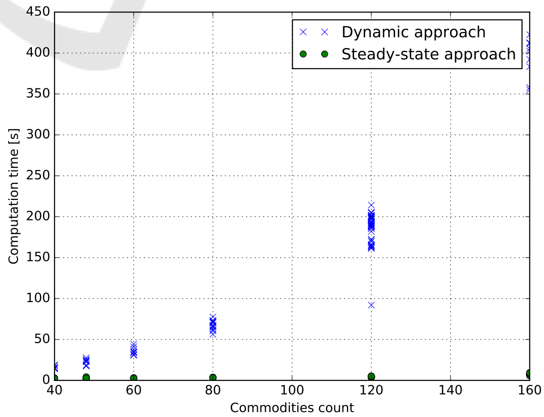
The dynamic approach results in a large linear program that is computationally demanding to solve optimally using general solution techniques as shown in Fig. 5c. The linear program can have up to $T \cdot M \cdot (|E| + M + |V|)$ variables and $T \cdot M \cdot |V| + M^2 + |E| \cdot T$ constraints. The most limiting factor in the sense of the scalability is M - the number of demand flows (i.e., commodities). Recall that a demand flow is an aggregation of requests that have the same origin, destination, the earliest pick-up and the latest drop-off time. In the worst case the M is in order of $V^2 \cdot T$ even under the additional assumption that the latest drop-off time is dependent on the three other demand flow parameters. Thus, the number of constraints may, in the worst case, grow with $|V|^4$, which makes this approach impractical for detailed model of large metropolitan networks that may contain tens or hundreds of thousands of vertices. Solving large-scale instances will thus likely require application of more advanced solution techniques. The results of (Schaefer et al., 2019) on the steady state problem variant indicate that Dantzig-



(a)



(b)



(c)

Figure 5: Average delay for combinations of *tau* and *peak* for the *steady-state* (a) and the *dynamic* (b) approach. Scalability (c): computation time for 2 h demand, the commodities count is varied by adapting the timestep.

Wolf decomposition (Dantzig and Wolfe, 1960) could be also employed to solve larger scale instances of our problem with time-varying demand.

8 CONCLUSION

To conclude, in this paper, we identified the practical limitations of steady-state approach to solve on-demand fleet routing when the demand is time-varying. We proposed an alternative approach that explicitly models the evolution of system in time. We implemented both steady-state and dynamic approaches and compared them on the simplified, but characteristic illustrative example. While the steady-state approach fails to find a solution or generate routes that exceed route link capacities by up to 50 %, the proposed approach is able to solve such instances without exceeding the road link capacities. Consequently, the simulation experiments with the congestion model reveal that the proposed approach that uses the flows over time model outperforms the steady-state approach in presence of the time-varying demand. The dynamic solution can transfer 42% more demand in the congestion-free regime than the steady-state approach on the same illustrative road network. The proposed model that explicitly models time is larger and generally harder to analyze and compute than the steady-state model. Therefore, in future work, we will study the applicability of specialized solution techniques for large-scale linear programs, e.g. the applicability of Dantzig-Wolfe decomposition method as applied in (Schaefer et al., 2019), to improve the scalability of the proposed approach.

ACKNOWLEDGEMENTS

The authors acknowledge the support of the OP VVV MEYS funded project CZ.02.1.01/0.0/0.0/16_019/0000765 "Research Center for Informatics" and TACR NCK project TN01000026 "Josef Bozek National Center of Competence for Surface Vehicles".

REFERENCES

Berbeglia, G., Cordeau, J.-F., Gribkovskaia, I., and Laporte, G. (2007). Rejoinder on: Static pickup and delivery problems: a classification scheme and survey. *TOP*, 15(1):45–47.

Dantzig, G. B. and Wolfe, P. (1960). Decomposition Principle for Linear Programs. *Operations Research*, 8(1):101–111.

Even, S., Itai, A., and Shamir, A. (1975). On the complexity of time table and multi-commodity flow problems. In *Foundations of Computer Science, 1975., 16th Annual Symposium on*, pages 184–193. IEEE.

Fiedler, D., Čáp, M., and Čertický, M. (2017). Impact of mobility-on-demand on traffic congestion: Simulation-based study. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–6. IEEE.

Ford, L. R. and Fulkerson, D. R. (1958). Constructing Maximal Dynamic Flows from Static Flows. *Operations Research*, 6(3):419–433.

Frederickson, G. N., Hecht, M. S., and Kim, C. E. (1976). Approximation algorithms for some routing problems. In *17th Annual Symposium on Foundations of Computer Science (sfcs 1976)*, pages 216–227.

Pillac, V., Gendreau, M., Guéret, C., and Medaglia, A. L. (2013). A review of dynamic vehicle routing problems. *European Journal of Operational Research*, 225(1):1–11.

Rossi, F., Zhang, R., and Pavone, M. (2016). Congestion-aware randomized routing in autonomous mobility-on-demand systems. *arXiv preprint arXiv:1609.02546*.

Schaefer, M., Čáp, M., Mrkos, J., and Vokřínek, J. (2019). Routing a fleet of automated vehicles in a capacitated transportation network. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 8223–8229.

Solovey, K., Salazar, M., and Pavone, M. (2019). Scalable and congestion-aware routing for autonomous mobility-on-demand via frank-wolfe optimization. In *Robotics: Science and Systems*, Freiburg im Breisgau, Germany.

Spieser, K., Treleaven, K., Zhang, R., Frazzoli, E., Morton, D., and Pavone, M. (2014). Toward a Systematic Approach to the Design and Evaluation of Automated Mobility-on-Demand Systems: A Case Study in Singapore. In *Road Vehicle Automation*, pages 229–245. Springer.

The Technical Administration of Roads of the City of Prague (2020). Prague transportation yearbook 2019. http://www.tsk-praha.cz/static/udi-rocenka-2019-en.pdf. Accessed: Nov 26, 2020.

Wollenstein-Betech, S., Houshmand, A., Salazar, M., Pavone, M., Cassandras, C. G., and Paschalidis, I. C. (2020). Congestion-aware routing and rebalancing of autonomous mobility-on-demand systems in mixed traffic. In *Proc. IEEE Int. Conf. on Intelligent Transportation Systems*, Rhodes, Greece. In press.

Zhang, R., Rossi, F., and Pavone, M. (2016a). Routing autonomous vehicles in congested transportation networks: Structural properties and coordination algorithms. In *Robotics: Science and Systems XII, University of Michigan, Ann Arbor, Michigan, USA, June 18 - June 22, 2016*.

Zhang, R., Rossi, F., and Pavone, M. (2016b). Routing Autonomous Vehicles in Congested Transportation Networks: Structural Properties and Coordination Algorithms. In *Robotics: Science and Systems XII*, volume 42, pages 1427–1442. Robotics: Science and Systems Foundation.