Fachhochschule Dortmund
University of Applied Sciences and Arts
Embedded Systems Engineering

Autonomous Transportation-on-Demand: Holistic Autonomous Transportation in Smart Cities

Master Thesis

Author: BRODO, Luca
Matriculation Number: 2780015

Supervisor: Henkler Stefan
Co-Supervisor: Name
Date: 00.2024

Abstract

Declaration

I hereby confirm, that I have written the Master Thesis at hand independently – in case of a group work: my respectively designated part of the work -, that I have not used any sources or materials other than those stated, and that I have highlighted any citations properly.

Contents

List of Figures

List of Tables

Chapter 1

Motivation

The economic impact of congestion in the United States is substantial, amounting to approximately \$121 billion per year or 1% of the country's GDP ([30]). This figure encompasses not only the staggering 5.5 billion hours lost to traffic congestion but also an additional 2.9 billion gallons of fuel burned. Moreover, these estimates overlook the potential costs associated with negative externalities such as vehicular emissions (including greenhouse gases and particulate matter) ([26]), travel-time uncertainty ([5]), and an elevated risk of accidents ([12]).

Chapter 2

Preliminaries

2.1 Model Predictive Control

Chapter 3

Literature Review

This chapter will center on an exploration of various pieces of literature and related works that have been thoroughly reviewed, aligning with the overarching purpose of this study.
To facilitate the reader, this chapter has been divided into multiple subsections, which follow the structure of this work.

3.1 Fleet and Transportation Network

Most of the literature reviewed during the development of this study focused on (Autonomous) Mobility-on-Demand, therefore only considering human mobility.In this case, *Frazzoli et al.* in [40] indentify three main mathematical models used to represent transportation networks: graph-theoretic, queuing theory and continuum models.

Graph-theoretic approach

In graph-theoretic models, the transportation network is modeled as a directed graph $G\langle V, R\rangle$, where V is the set of vertices and $E \subseteq V \times V$ is the set of edges. Usually, each node $v \in V$ represents a location, or a station. In [4], in a more abstract way, nodes are used to represent areas of interest. Each edge $\langle v_1, v_2\rangle \in E$ rapresents a connection, which could be a road or a collection of those, from v_1 to v_2. Furthermore, each link is usually associated with a metric, such as distance $d : E \rightarrow R_{\geq 0}$. Many studies adopt a dynamic fluid approach, representing AVs and customers not as individual entities but rather as flows moving between nodes. In this framework, models that do not vary with time often simplify to static network flow problems. Alternatively, graph-theoretic models can be integrated with vehicle-centric models, where both AVs and customers are modeled individually.

Queuing-theoretic approach

Queuing theory is concerned with the examination of waiting lines. In the context of AMoD systems, trips are conceptualized as queues between locations, enabling the analysis of various characteristics AVs, closely linked to waiting times. Formally, considering N stations situated at specific geographical locations with m AVs providing mobility services. Customers arrive at each station based on an exogenous stochastic process (usually a Poisson process with rate λ, like for example in [43]) and select destinations with certain probabilities. If AVs are accessible at a station, customers proceed to their respective destinations; otherwise, they exit the system (referred to as passenger loss). Travel times between stations are also stochastic and are typically modeled as exponentially distributed random variables. When formulating policies for AMoD systems, the focus lies in analyzing and regulating the movement of AVs from one queue to another. These models are commonly employed in time-invariant settings. The scenario of infinitely large fleets has garnered particular attention for deriving theoretical insights.

Battery model

This chapter delves into the exploration of various sources and solutions that have been examined in the development of the battery model utilized in this study. As the first work considered, *Montoya et al.* in [22] employed linear piecewise approximations to approximate the nonlinear charging function, reporting an error of approximately 1%. Additionally, they introduced the concept of different types of charging stations. *Froger et al.* in [9] and *Kancharla et al.* in [16] also employed piece-wise linear functions for their respective charging regimes. Conversely, *Nie et al.* in [24] did not consider the possibilities for vehicle recharging in their model.

In terms of modeling the battery charging profile, diverse approaches have been proposed. *Han et al.* in [11] developed a model based on the internal resistance and voltage of the battery; however, this model was disregarded in favor of others that better captured the battery charging profile based on more relevant factors. *Yu et al.* in [38] proposed a sophisticated model for charging constraints in general. Although their approach differs from the one presented in this work, as it incorporates charging during routing, this aspect was excluded in favor of an alternative approach. *Lee et al.* in [18] considered a simplified battery model that provides states of charge (SoC) as a function of the charging current. The author asserted the general applicability of their approach to every charging profile, provided SoC(t) is a concave and non-decreasing function with SoC(0) = 0, and an inverse function $\mathrm{SoC}^{-1}(a)$ exists for $0 < a < Q$.

3.2 AToD Challanges

AV Dispatching

For the vehicle dispatching problem, multiple solutions have been already proposed. For example, *Vasco et al.* in [1] model it as a linear minimum cost multicommodity flow problem; *Holzapfel et al.* in [13] model it as a MIP problem. Others, on the other hand, make use of other heuristics. For example, *Levin et al.* in [19] and *Mora et al.* in [3] use nearest neighbors. Recently, to improve real-time capabilites, some approaches have been proposed which make use of deep learning, like for e.g. the one proposed by *Yu et al.* in [39].

AV Routing

On a basic level, the routing problem can be reconducted to the Vehicle Routing Problem (VRP) (a thorough analysis and explanation can be found in [33]). Usually, the VRP is solved and analyzed as a static problem, which implies that requests are known in advance. As a result, origins and destination of each trip is also known a priori. As pointed out by *Frazzoli et al.* in [40], in AToD systems requests are dynamic, meaning they are not known in advance. More specifically, the task of managing an AToD system can be viewed as a specific case of the dynamic one-to-one pickup and delivery problem. As pointed out by *Zhang* in [42],to provide a more detailed characterization of these systems, several additional attributes and constraints must be considered:

- Immediate Service Requests: Requests are made for immediate service rather than being scheduled in advance with specified time windows.
- Stochastic Customer Arrivals and Travel Times: The timing of customer requests and the duration of travel are subject to stochastic variability.
- Multi-Occupancy Vehicles: Contrary to [42], vehicles must be considered to have a capacity, rather than being single-occupancy. This is motivated by the fact that recently developed AV are equipped with more than one seat (see for e.g. [10])

Wollenstein-Betech et al. in [36] propose a solution to the routing-rebalancing problem in mixed traffic situations. They propose an interesting two-graph solutions, i.e. one graph for the AToD system and one for the pedestrians which are interconnected. Furthermore, they also consider private and public vehicles. *Pavone et al.* in [27] also propose a solution for the rebalancing problem. Although they mainly consider MoD systems, their approach can be easily applied to AToD systems. Their solution, moreover, focuses only on the rebalancing issue, which is studied using a steady-state fluid model, as pointed out in [36]. Furthermore, this work does not take into consideration congestions in their model.

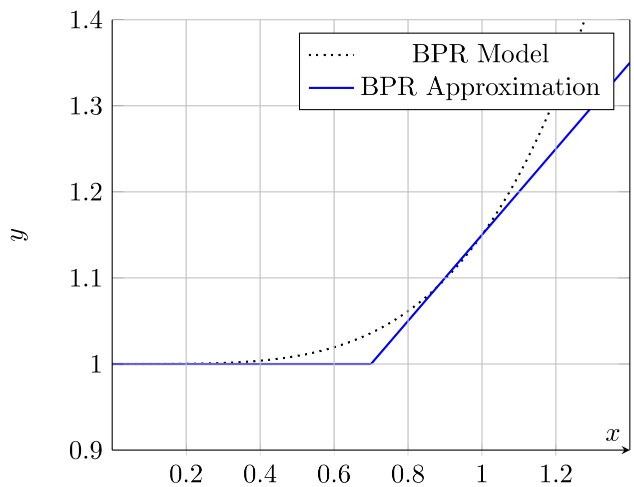


Figure 3.1: BPR Model and its Approximation as a piecewise linear function. In this example, $x_{th} = 0.7$,$x_{max} = 1.4$, $a = 1$, $b = 0.5$

AMoD Rebalancing

In the context of network flow models, rebalancing is usually formulated as a (integer) optimization problems, such as in the works of *Zgraggen et al.* in [41], *Carron et al.* in [6] and of *Smith et al.* in [31]. Crucially, in real-time texting, the rebalancing problem is solved leveraging the model-predictive control framework, such as in the works mentioned above.
Wollenstein-Betech et al. in [37] investigate effective pricing and rebalancing strategies for AMoD systems from a macroscopic planning standpoint. It aims to optimize profits while ensuring system load balance. Using a dynamic fluid model, the study demonstrates equilibrium attainment via pricing policies and develops an optimization framework for determining optimal pricing and rebalancing approaches. They model customers and vehicles located in regions using a queuing approach, i.e. two queues per region representing vehicles and clients.
Salazar et al. in [29] study the rebalancing problem while considering two important aspects, namely mixed traffic and congestions. The first is treated by modeling the road network and the walking network as two separate entites, which combined make up a supernetwork defining the entire transportation network. The latter are treated according to the Bureau of Public Roads (BPR) model ([25]) , more specifically by approximating their model linearly. The BPR model of congestions can be expressed by Equation 3.1 and is visualized in Fig 3.1 as a black dotted line.

$$f_{BPR}(x) = 1 + 0.15x^4 \tag{3.1}$$

Since this model would lead to an unconvex, i.e. intractable, problem, the authors propose a piece wise approximation of it, which take the form expressed in Equation 3.2

$$y = \begin{cases} a & \text{if } x \in [0, x_{th}] \\ a + b \cdot (x - x_{th}) & \text{if } x \in [x_{th}, x_{max}] \end{cases} \tag{3.2}$$

where a and b are used to model the height of the horizontal line and the slope of the second line, x_{th} is the threshold of the piecewise approximation and x_{max} defines the approximation window. This approximation can be seen in Fig 3.1 as a blue line.

Equation 3.2 is then used to calculate the traverse time for each edge. While this is a more representative congestion model than the one proposed in this paper in Section 4.1, one might argue that the two serve two different purposes. While the one proposed in [29] serves to express the travel time in terms of the amount of vehicles currently traveling on the node, the one proposed in this work is used to simply limit the amount of vehicles traveling on the link to avoid situations of stop-and-go traffic. Furthermore, it is worth mentionining that this model can be easily integrated into the model by modifying the definition of the travel time function.

In the abovementioned study, the energy consumption of the AVs is also modeled differently compared to this work. Within their network graph G, the energy consumption of each AV with efficiency η traveling through arc (i, j) with distance d_{ij} is expressed as:

$$e_{ij} = (\frac{\rho_a}{2} \cdot A_f \cdot c_d \cdot v_{ij} + c_r \cdot m_v \cdot g) \cdot \frac{d_{ij}}{\eta_{\text{AV}}} \quad \forall (i, j) \in G \tag{3.3}$$

where the aerodynamic drag is determined by the air density ρ_a, the frontal area A_f, the drag coefficient c_d and the moving velocity v_{ij}. The friction of the wheels on the road is determined by the rolling friction coefficient c_r, the mass of the vehicle m_v, and the gravitational acceleration g.

This is a more vehicle-based approach compared to the one adopted in this work, which only considers the battery dischargment, as it takes into consideration various other aspects such as drag and rolling friction. For the purpose of this work, however, it has been chosen to neglect those factors, which could be later introduced into the battery-dischargment model, to favour a more simplistic approach. Furthermore, given that in the aforementioned paper is assumed to have constant velocity $v_{ij} = \frac{d_{ij}}{t_{ij}}$, one can trivially adopt a similar solution for the model in this paper by making the same assumption.

Wallar et al. in [35] tackle the rebalancing problem independently from any other AToD challenge, while also considering ride-sharing possibilities. The main idea of the propsed approach is to assign free vehicles to regions, which are computed offline, according to the estimations of traveling request per region. In this case, the requests are estimated using the particle filter. Furthermore, the division in region for the graph is formulated as an integer linear programming problem using a rechebaility matrix R, which indicates whether a station j if recheable from i within a maximum time t_{max} ($R_{ij} = 1$ else $R_{ij} = 0$). Cleverly, the authors also took into consideration the fact that, although a vehicle is free, it requires some time to reach the assigned rebalancing region. In other words, if a vehicle requires eight minutes to reach a certain region, considering a time horizon of ten minutes, it will be available only for two minutes, i.e. 20% of the time. This is used to limit the oversaturation of vehicles in the respective rebalancing regions.

OThers

Fehn et al. in [8] modeled a fleet according to energy prices.

Chapter 4

Time-invariant Approach

This chapter focuses on tackling dispatching, routing and rebalancing problems for an AToD system. First, a time-invariant vehicle-centric model for the transportation network is developed. The model presented in this document expands upon the one formulated in [4], using graph theory and a vehicle-centric approach to provide flexibility in handling different vehicle capacities and types. Additionally, this model addresses the combination of mobility of people with goods transportation, allowing for a more comprehensive analysis of transportation networks. This model will then be used to formulate and solve the aforementioned problems in a real-time fashion. Finally, the proposed formualtin is evaluated on a ficticious use-case.

4.1 Vehicle-centric Model

The model provided in this section will expand the model formulated in [4]. It is worth noting that, although the model presented in [4] is specific for the scenario described in previous sections of the same work, it can be expanded to cover more general use cases.

It follows that the transportation network will be modeled using graph theory. Most of the literature makes use of (Eulerian) fluid-dynamic models ([40]), i.e. customers and AVs are represented as (non-integer) flow between nodes, in this work, this model will be used in addition to a vehicle-centric approach, where customers and AVs will be modeled individuall. Contrary to the solutions proposed by most of the literature, this work combines mobility of people with goods transportation, therefore modeling vehicles individually will provided the required flexibility to handle different capacities and the increased complexity given by the plethora and variety of vehicles, as we will see later in this section. The set of vehicles will be refered to as $\mathcal{A}$.

As previously stated, let $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$ being the directed graph representing corresponding transportation network of the city in question, where $\mathcal{V}$ is the set of vertices and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. Any vertex, or node, $v \in \mathcal{V}$ represents

a location. Following the notation in [4], each node consists of an area of interests. For the context of this paper, the two terms will be used intercheangeably, as they ultimately indicate the same thing. An edge $\langle v_i, v_j \rangle \in \mathcal{E}$ represents a connection, which can consist of a road or a combination of those, linking v_i to v_j. The level of abstraction is decided by the engineer and can be used to regulate the granularity of the model. Locations v_i and v_j can indicate low level key points, such as cross roads or trafficated traffic lights, as well as higher level areas such as residental areas or leisure centers. Similarly, an edge can contain multiple crossroads or joints or simply indicating a path from v_i to v_j, regardless of additional details. A correct granularity is not trivial to decide a priori and a general approach is likewise hard to delineate. Determining the level of abstraction remains, therefore, an engineer's prerogative and highly depends on the system in question. It is also worth noting that the, for conveniency and to better reflect the situation in real world applications, the graph is not a bidirect graph. To model a two-way road, one should simply make use of two edges from locations v_i and v_j.
As also described in [4], each edge is associated with multiple metrics and information.
Firstly, at each edge one must attribute a travel time T. T is a function $T : \mathcal{E} \times \mathcal{A} \rightarrow \mathbb{R}_{>0}$, which at each given a vehicle and edge maps a float value $T^a_{i,j} \in \mathbb{R}_{\geq 0}$ indicating the time required for the vehicle of type a to travel the path from v_i to v_j. In this work, the unit of $T^a_{i.j}$ is of particular interest, as it depends on the application.
Secondly, each edge is also associated with a distance $d : \mathcal{E} \rightarrow \mathbb{R}_{\geq 0}$, which maps an edge to a float value $d_{i,j}$ indicating the distance from v_i to v_j.
Another factor which is considered in this work, as stated above, is pollution. This metric is associated to each edge and is a function $f : \mathcal{E} \times \mathcal{A} \rightarrow \mathbb{R}_{\geq 0}$. To simplify the discussion for this work, f is assumed to produce a certain value refered to as pollution index. In other words, it is not supposed to represent a measurable element, such as CO_2 emissions, but rather a value which can represent multiple quantitative factors and is higher if the combination of path and vehicle type is highly polluting. While it can be argued that the pollution can also be simply a function of d and T, this abstraction does not account for other elements, such as road type, vehicle fuel and road slope. The index f, therefore, is a mathematical abstraction which allows a more flexible model for the system. To use this abstraction, however, each edge must include additional information, including the one mentioned before.
Similarly to what observed by *Zhang et al.* in [28], we can introduce the concept of congestions in the model by constraining the routing by the capacity of each road. In other words, we can associate each edge with a capacity $c : \mathcal{E} \rightarrow \mathbb{N}_{>0}$ indicating the limit in terms of car occupancy above which the traffic in that edge slows down eventually reaching a congestion. As pointed out in [28], this semplified model is adequate in this context as well as the aim of this work focusing solely on the control of vehicles to prevent congestion rather than analyzing their behavior in congested network. While it is understood that congestions behave differently in real world scenarios and other, more sophisticated approaches ex-

ist ([20], [34]), for the sake of simplicity, we will assume $T_{ija} = \infty$ if the number of vehicles in that edge from $\langle v_i, v_j \rangle$ to be larger than the capacity c_{ij}.
In addition, each edge $\langle v_i, v_j \rangle$ will also include information regarding traffic limitations, i.e. whether a vehicle type is allowed or not to travel that link. Intuitively, limitations will be represented as a function $s : \mathcal{E} \times \mathcal{A} \rightarrow \{0, 1\}$, where 1 implies the vehicle can traver that edge. Following the example in [4], limitations can be because of various factors, like for e.g. weight or height. For this reason, it is convenient to abstract away such details and just indicate wether a link is can be traversed or not. Notably, one could use this representation to transform a bidirect graph G' to be equivalent to G by letting $s(e, a) = 0$ for any one-way road $e = \langle v_i, v_j \rangle$ for all $a \in \mathcal{A}$. For convenience, this function will be incorporated in the definition of the capacity, updating it to become Equation 4.1.

$$c(e = \langle v_i, v_j \rangle) = \begin{cases} 0 & \text{if } s(e) = 0 \\ c(e) & \text{otherwise} \end{cases} \tag{4.1}$$

Following the discussion above, this is equivalent to setting $T_{ija} = \infty$ for all $a \in \mathcal{A}$ according to this model specifications. In this way, we can treat the road as being unaccessible without increasing the number of conditions and decrease the redeability of the model.
As mentioned above, the set of autonomous vehicles is indicated as $\mathcal{A}$ and each vehicle will be modeled as a tuple $\langle \underline{s_a}, \bar{t_a}, S_a, Q_a, I_a^b, R_a^-, R_a^+, \theta_a, B_a(t), \mathcal{R}_a, \mathcal{T}_a \rangle$. $\underline{s_a} \in \mathcal{V}$ and $\bar{t_a} \in \mathcal{V}$ are the starting and terminal node respectively; $Q_a \in \mathbb{R}_{>0}$ will be used to indicate battery capacity and charing rate and discharging rate will be represented as $R_a^+ \in \mathbb{R}_{>0}$ and $R_a^- \in \mathbb{R}_{>0}$ respectively; $\theta_a \in [0, 1]$ is used to model the battery breakpoint and $B_a(t) \in \mathbb{R}_{\geq 0}$ is the state of charge at time t; $\mathcal{R}_a$ is the set of requests assigned to vehicle a and $\mathcal{T}_a$ is the type. For a more detailed understanding of the vehicle type, the reader is encouraged to analyze the discussion in [4]. For the purpose of this work, the vehicle type will be understood as a tuple $\langle P_a, G_a, C_a, F_a \rangle$, where , $G_a \in \mathbb{R}_{\geq 0}$ and $P_a \in \mathbb{R}_{\geq 0}$ are the goods and people capacity and $C_a \in \mathbb{R}_{>0}$ and $F_a \in \mathbb{R}_{>0}$ indicate operational cost and pollution factor respectively. Furthermore, each set of assigned requests $\mathcal{R}_a$ is a subset of the set of all requests in the systems, i.e. $\mathcal{R}_a \subseteq \mathcal{R}$. For the purpose of this model, we will assume each request assignment to be unique, i.e. $\mathcal{R}_a \neq \mathcal{R}_b$ and $\mathcal{R}_a \subseteq \mathcal{R} \setminus \mathcal{R}_b$, for all $(a, b) \in \mathcal{A}$ with $a \neq b$.
The battery will be modeled according to the two operating mode, i.e. charging and discharging. While it is understood that these two operations are highly influenced by multiple factors, it is sensible to make assumptions in order to simplify the model. As it is also widely spread in industry, one can neglet the influence of external factors such as weather condition or intrinsic characteristics of the battery, such as temperature or age. The charging profile will be modelled taking inspiration from the model proposed by *Lee et al.* in [18]. As mentioned above, the vehicles have a charging rate, a battery capacity and a breakpoint, namely R_a^+, Q_a and θ_a respectively. The state of charge at time $t \in \mathbb{R}_{\geq 0}$ of a certain vehicle will be obtained according to the model described in Equation 4.2, which is derived from the CC-CV (Constant current - Constant Voltage)

scheme. (A more complete explenation can be found in [21]).

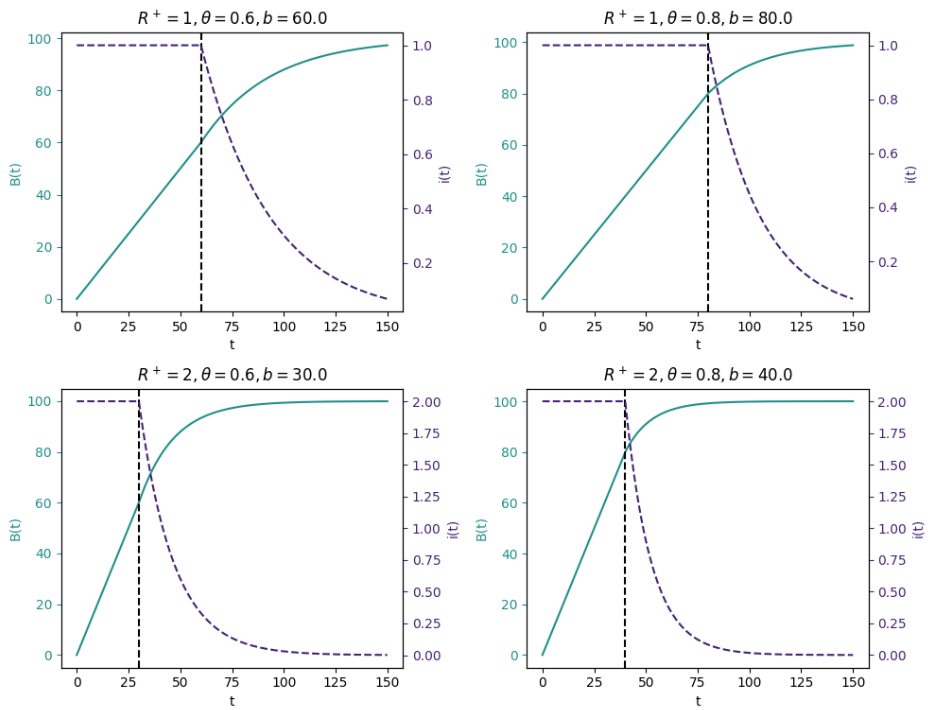


Figure 4.1: Different charging profile obtained according to the model developed in Equation 4.2. B is the State of Charge in Percentage, while $i(t)$ in function of the time unit t. The battery have $Q = 100$

$$B_a(t+1) = \begin{cases} B(t)_a + R_a^+ & \text{if } t \leq b_a \\ \\ Q_a - \dfrac{Q_a i(t)(1-\theta_a)}{R_a^+} \end{cases} \tag{4.2}$$

with $b_a = \dfrac{\theta_a Q_a}{R_a^+}$, $i(t) = \begin{cases} R_a^+ & \text{if } t \leq b_a \\ R_a^+ e^{-(t-b_a)/\tau} \end{cases}$ and $\tau \in \mathbb{R}_{\geq 0}$ being the time constant, representing how quickly the current decreases.
An example of the charging profiles from this model can be seen in Fig 4.1. Since $i(t)$ is assumed to be constant at time $t \leq b_a$, i.e. $i(t) = R_a^+$,the first

condition is derived as follows.

$$
\begin{aligned}
B_a(t+\alpha) &= B_a(t) + \int_t^{t+\alpha} R_a^+ dt \\
B_a(t+\alpha) &= B_a(t) + R_a^+ \int_t^{t+\alpha} dt \\
B_a(t+\alpha) &= B_a(t) + R_a^+ \quad t\Big|_t^{t+\alpha} \\
B_a(t+\alpha) &= B_a(t) + R_a^+(t+\alpha-t) \\
B_a(t+\alpha) &= B_a(t) + \alpha R_a^+
\end{aligned}
$$

with $\alpha = 1$.
The discharging, on the other hand, will be modeled as it being directly proportional to the travel time T, as described in Equation 4.3.

$$B_a(t + T^a_{u,v}) = B(t)_a - R_a^- T^a_{u,v} \tag{4.3}$$

The equation describes the relationship between the initial and final battery levels on edge $\langle v_u, v_v \rangle$ during the transit. Following Equation 4.3, one can assign to each edge a rate of battery discharge by defining it as the difference between the state of charge at time t and at time $t + T^a_{u,v}$.

$$e^a_{u,v} = B(t)_a - B_a(t + T^a_{u,v}) = R_a^- T^a_{u,v} \tag{4.4}$$

Equation 4.4 allows to describe the discharging rate as a function of only the edge and the vehicle.
In terms of operational costs, multiple factors should be considered and the analysis must be extended from the one developed in [4], where the operational cost was only in function of the vehicle type. Similarly to the aforementioned work, this model assumes a certain operational cost depending on the vehicle type, however the discussion is also extended in terms of vehicle charging cost. In other words, we can decouple it from the general concept of the operational cost per vehicle and considering in function of the charging or discharging profiles described above. Furthermore, it also makes sense to assign the cost also depending on the distance traveled, i.e. the value d associated to the edges.
Compared to [4], in this work, the model also extends the information associated to the nodes. Nodes will be of three categories namely (i) normal nodes, (ii) charging stations and (iii) depots, whose sets are denoted as $\mathcal{V}_n$, $\mathcal{V}_c$, $\mathcal{V}_d$, respectively. We can, therefore, conclude that $\mathcal{V} = \mathcal{V}_n \cup \mathcal{V}_c \cup \mathcal{V}_d$. The information associated to the nodes depend from the category each node belongs to. Nodes belonging to $\mathcal{V}_n$ do not possess any specific information, as they are not of high importance for this work. Charging nodes $\mathcal{V}_c$ and depots $\mathcal{V}_d$, on the other hand, have in common the vehicle capacity they can host: charging nodes can only host as many vehicle as charging stations available, while depots can

only station as many vehicles as the number of free parking places. Additionally, charging nodes are assumed to have all the same models of charging stations. For simplicity, only three types will be considered, i.e. fast, normal and slow charging, which will be based on the model in Equation 4.2. The choice of parameters is explained in Table 4.1.

Charging station	R^+	θ
Slow	1	0.6
Medium	1	0.8
Fast	2	0.8

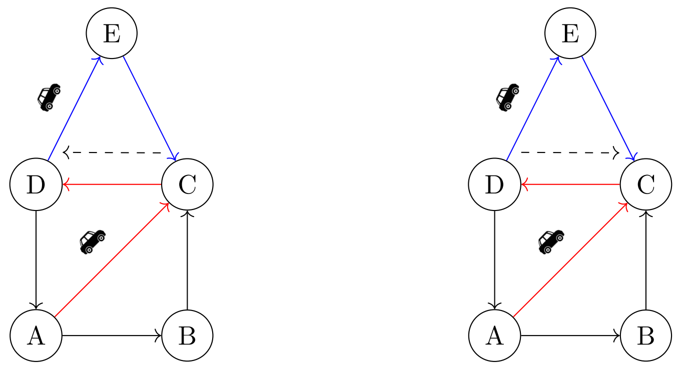
Table 4.1: Choice of parameters for the charging stations according to the type. For the charging profile, please refer to Fig 4.1

Depot nodes are also equipped with charging stations of the same three types. The difference between charging and depot stations lays on the fact that charging nodes are not meant to host vehicles for long periods of time, contrary to depots. The difference can be envisioned as if charging stations were gas stations and depots were bus depots.
It would be sensible to limit the domain of the starting and terminal node of each vehicle to the sets of charging and depot nodes, i.e. $\underline{s_a} \in \mathcal{V}_c \cup \mathcal{V}_d$ and $\bar{t_a} \in \mathcal{V}_c \cup \mathcal{V}_d$. In other words, this choice implies that each vehicle's destination will either be the depot at the end of the shift, for example, or a charging node, where its battery can be charged to accomodate future requests.
Requests are modeled as tuples $\langle \underline{s'}, \bar{t'}, Q', P', \lambda, a', b' \rangle$, where $\underline{s'} \in \mathcal{V}_n, \bar{t'} \in \mathcal{V}_n$ represent the pickup and delivery point respectively; $G' \in \mathbb{R}_{\geq 0}$ ($P' \in \mathbb{R}_{\geq 0}$) refers to the amount of goods (people) required to transport, and $\lambda \in \mathbb{R}_{>0}$ is the rate of requests, in customers per unit time, which, therefore, makes the requests stationary and deterministic. Additionally, requests must be delivered within a time window within $[a', b']$.

4.1.1 Model Evalutation

Some comments are in order. The model in Section 4.1 is time-invariant. According to the definition presented by *Frazzoli et al.* in [40], time invariance in the context of transportation modeling refers to the assumption that the number of requests remain constant over time allowing for the simplification of temporal dynamics and treating specific time intervals as homogeneous units. This modeling concept is employed when the rate of change in transportation demands is deemed slow compared to the average travel time of individual trips, often observed in stable urban environments ([23]). While this model is indeed applied in a relatively dense urban environment, some integrations are required in order to adapt it to time-varying scenarios. Moreover, customer requests are also assumed to be known. This requirement can be fulfilled in practice with requests made in advance or some techniques to estimate requests throughout

the day. It is important to note that request estimation might lead to suboptimal performance.
Secondly, as already mentioned in the previous section, the model used to describe congestions is indeed rather simple and more complex formulation might better capture the phenomena. However, the semplification is considered powerful enough for the purpose of this work and, as pointed out by *Zhang et al.* in [28] as well, more sophisticated models can be used offline using simulation techniques to derive the capacity metric used in this model.
Some comments are in order also regarding vehicles. Firstly, we assume the vehicles to be autonomous and fault proof. It is outside of the scope of this work to deal malfunctioning vehicles or exceptional situations outside of the normale functioning regime of the system.Additionally, we assume the vehicles to be fully electric. This assumption is widely used in literature and it is motivated, among other aspects, also by the recent trends in industry to transition towards electrical mobility. Furthermore, the model provided for the battery charging and discharging profiles is rather simplistics and more sophisticated approaches exist in literature (see Chapter 3 for more details). However, these models should be seen an addition to the corrent approach and, albeit with some potential modifications, it is plausible they can be integrated in the system model. While in [4] vehicles have been designed to be capable of transporting a potentially large number of people, some considerations should be made in this regard. While it is sensitive to consider vehicles capable of transporting up to 50 people in terms of environmental, economical and overall transportation efficiency and while it is also expected that the model designed would work with those vehicles as well, in terms of practical use, it might be more appropriate to make use of those vehicles in different ways. Notably, it should be studied whether it is efficient to tailor the routes of those vehicles according to passenger needs. Firstly, if we do not consider ride-sharing possibilities, the scenario of such large group of people traveling together is very unlikely. Additionally, if we consider ride-sharing, the possible route combinations, which increases drastically as the number of requests and nodes in the graph increase, would likely make routing such vehicle relatively inefficient. Motivated also by the fact that passengers might have common stops, it might be more sensible to treat those vehicles as buses are treated nowadays, i.e. with a pre-determined route among stops which are placed according to the most common stops, such as hospitals or train stations. Similarly, if we consider large vehicles to transport goods, such as large trucks, those are not used for home delivery, but rather used to transport goods among specialized centers. It must be noted that both situations do not invalidate this work, or any previous work. It is clear that the system described in this work in combination to the truck for goods transportation. Regarding people mobility, while buses are indeed an already established and efficient transportation mean, this systems aims at filling up the situations where a bus system is indeed lacking, such as transportation of people with special needs or, in general, more tailored to the specific needs of potential customers.
In this model, we also made certain assumptions on the nodes. First of all, we assumed the charging stations to be all of the same types. It might be argued



(a) (b)

Figure 4.2: Fig 4.2a and Fig 4.2b show a simplified example of a sensible request assignment. In red and blue are the paths the two AVs can traverse, while the dashed blue two different requests (In Fig 4.2a the custoer asks to go from C to D, while in Fig 4.2b the customer asks to go from D to C). In the case of Fig 4.2a, it is more sensible to assign the request to the red AV, while in the case of Fig 4.2b, the blue AV is a better choice.

that some stations might have different types of chargers. This characteristics can be reflected by the model simply by 'splitting' them. A station having, for instance, two types of chargers can be represented by two equivalent nodes in the graph having different category. Moreover, while in Table 4.1 we only considered three modes, this can be easily extended.

4.2 Problem Formulation

Before defining the problems for the different challenges faced by the AToD considered in this work, it is essential to lay down the motivations behind certain assumptions made during the conceptual phase.
Firstly, following the previously described model, it is assumed that each vehicle starts and ends at a charging or depot node, $\underline{s_a}$ and $\bar{t_a}$, which do not have to be necessarily the same. Furthermore, if a vehicle leaves from a depot, it is assumed to have a fully charged battery, i.e. $B_a(0) = 100$.

4.2.1 Dispatching

Informally, the dispatching problem can be defined as the task of assigning requests to the most suitable vehicle. There exist already multiple solutions proposed for the problem and we refer to Chapter 3 for a more thorough analysis. Dispatching is critical for the overall system performance and must be done in a way that can further facilitate the next steps. Clearly, dispatching can not be decoupled and solved as a stand-alone problem. For example, Fig 4.2 shows a

simplified situation where a sensible dispatching, which depends on a posterior step, will improve system performance. Nevertheless, during the dispatching problem, some additional elements must be considered as well.
Since the model proposed in Section 4.1 is a vehicle centric model, we can model the dispatching problem with the help of a binary variable x_{ar} defined in Equation 4.2.1.

$$x_{a,r} = \begin{cases} 1 & \text{if } r \text{ is assigned to } a \in \mathcal{A} \\ & \\ 0 & \end{cases} \qquad \forall r \in \mathcal{R}$$

According to the model, each vehicle has a capacity which must not be exceeded. Such constraint can be expressed as follows for both people and goods.

$$\sum_{r \in \mathcal{R}} P'_r \cdot x_{a,r} \leq P_a \quad \forall a \in \mathcal{A} \tag{4.5}$$

$$\sum_{r \in \mathcal{R}} G'_r \cdot x_{a,r} \leq G_a \quad \forall a \in \mathcal{A} \tag{4.6}$$

Furthermore, if a vehicle is already on the move, the request can be picked up only if the vehicle's charge is enough to satisfy such request as well. In this case, such requirement can be expressed as follows.

$$\sum_{r \in \mathcal{R}} e(\underline{s'_r}, \bar{t'_r}) x_{a,r} \leq B_a \quad \forall a \in \mathcal{A} \tag{4.7}$$

where $e : \mathcal{E} \times \mathcal{E} \to \mathbb{R}_{>0}$ is a function expressing the required energy to go from $\underline{s'_r}$ to $\bar{t'_r}$.
Accordingly, using Equation 4.5, 4.6 and Equation 4.7, one can simply formulate it as an integer programming problem by finding the appropriate cost function to minimize, such as minimizing waiting times. Thanks to this formulation, one can also make sure that each request has been served at most λ_r times , i.e. Equation 4.8.

$$\sum_{a \in \mathcal{A}} x_{a,r} \leq \lambda_r \quad \forall r \in \mathcal{R} \tag{4.8}$$

Furthermore, one must ensure that each request is assigned only to λ_r vehicles. This is insured by Equation 4.9.

$$\sum_{r \in \mathcal{R}} x_{a,r} = \lambda_r \quad \forall a \in \mathcal{A} \tag{4.9}$$

The above mentioned equations are based upon the work developed by *Hyland et al.* in [14].
This way allows to define the following cost function which expresses the number

of served requests.

$$\mathcal{J}_u = \sum_{r \in \mathcal{R}} (1 - \min_{a \in \mathcal{A}} (x_{a,r}, 1)) \tag{4.10}$$

The main strength of this approach is that can be integrated naturally in the formulation for the other steps, like for e.g. the one in Section 4.2.2.
Alternatively, the dispatching problem can be solved independently without being integrated in other steps. For example, in Section 3, some works are mentioned that make use of heuristics such as nearest neighbours. On the one hand,these approaches are known to obtain sub-optimal solutions for the problem; on the other hand, they provide flexibility and might result in less time or space complexity.

4.2.2 Routing

After being assigned to incoming requests, vehicles must be routed in such a way that can reach all customers and therefore satisfy all the requests. In other words, the routing problem consists of determining paths, i.e. a series of edges, each vehicle must travel in the graph to fulfill the requests while respecting all the requirements and minimizing some metrics, i.e. a cost function. Since this problem can be reconduced to the vehicle routing problem, in particular dynamic pickup and delivery problems (reviewd by *Toth et al.* in [32] and *Laporte et al.* in [17]), the formulation used in this work will be based on this family of problems.
Within our model, binary flow variables will be used to identify whether a vehicle should traverse a link. Formally, this can be expressed as

$$V^a_{u,v} = \begin{cases} 1 & \text{if } a \text{ traverses } (u,v) \in \mathcal{E} \\ \\ 0 & \end{cases} \qquad \forall a \in \mathcal{A}, \forall u, v \in \mathcal{V}$$

Furthermore, another variable will be needed to deal with time-related requirements. This concept is based on the work in [15]. The decision variable s^a_u indicates the service time of vehicle a at node u. This variable will be relevant only for nodes labeled as being terminal stations for the requests, i.e. $\underline{t'a}$ and will be irrelevant for other nodes.
Accordingly, one can express the abovementioned cost function using this variable. Classical examples of cost functions are for e.g. travel time and travel distance. These are frequently used in literature ([2]), as they are general, in a sense that many other metrics could be reconducted to them. For example, residual charging or operational costs are directly influenced by the two. Furthermore, the formualtion of the cost function derived from these metrics is rather trivial and intuitive, while at the same time producing desirable results in practice. The formulation is described in equation Equation 4.11 and 4.12.

$$\mathcal{J}_T = \sum_{a \in \mathcal{A}} \sum_{(u,v) \in \mathcal{E}} T^a_{u,v} V^a_{u,v} \tag{4.11}$$

$$\mathcal{J}_d = \sum_{a \in \mathcal{A}} \sum_{(u,v) \in \mathcal{E}} d_{u,v} V^a_{u,v} \tag{4.12}$$

In the effort to minimize the environmental impact of the AToD system, the pollution index discussed in previous sections can be utilized to formulate the cost function outlined in Equation 4.13.

$$\mathcal{J}_f = \sum_{a \in \mathcal{A}} \sum_{(u,v) \in \mathcal{E}} f^a_{u,v} V^a_{u,v} \tag{4.13}$$

Up to this point, the metrics under consideration have been geared towards minimization. Put differently, the goal is to reduce travel, encompassing both distance and time, to enhance system performance. Likewise, minimizing environmental impact is crucial in this scenario. However, there are instances where maintaining certain metrics at higher levels is preferable. For instance, closely tied to operational costs and environmental impact, it is advantageous to keep the state of charge at its maximum. Hence, it is imperative to maximize the cost function in Equation 4.14.

$$\mathcal{J}_B = \sum_{a \in \mathcal{A}} B_a \tag{4.14}$$

Finally, while it should also be explored whether the combination of those could improve the overall performance of the system. For this purpose, the cost functions can be combinated using weights as follows.

$$\mathcal{J}_{tot} = \lambda_T \mathcal{J}_T + \lambda_d \mathcal{J}_d + \lambda_f \mathcal{J}_f + \lambda_B \mathcal{J}_B \tag{4.15}$$

where $\lambda_i \in \mathbb{R}$ with $i \in \{T, d, f, B\}$ are the weights.
It is not uncommon to also find in literature cost functions which are developed in terms of operational costs, as it provides a general idea to reason about this problem and allows for a systematic evaluation of different dispatching strategies. Moreover, the inclusion of operational costs in the literature emphasizes the real-world impact of dispatching decisions, aligning theoretical models with practical considerations. This connection to tangible costs not only enhances the applicability of proposed solutions but also contributes to the development of more realistic and effective dispatching strategies. Considering that, within our model in Section 4.1, each vehicle is modeled to have an operational cost, one can also derive a cost function similarly to what already done in [4].

Unconstrained Version

Before formulating the routing problem formally, a small consideration must be made. For convenience, the set of nodes recheable from node u by traversing a single edge $\mathcal{N}_u^+$, i.e. the ingoing neighbours of u, and $\mathcal{N}_u^-$ as the set of outgoing neighbours of u.
Additionally, in order to ease the notation, the set containin all the initial stations of each request $r \in R_a$ will be indicated as $\underline{S_a'}$. Likewise, $\bar{T_a'}$ will be the set of terminal stations of each request $r \in R_a$.
On a basic level, i.e. without constraints, one can formulate the *Unconstrained Routing Problem (URP)* as follows. Given a transportation network $\mathcal{G}$, a set of vehicles $\mathcal{A}$ and a set of requests $\mathcal{R}$, defined within the description in section Section 4.1, solve:

$$\min \text{ (4.11), (4.12), (4.13) or (4.15)}$$

s.t.

$$\sum_{u \in \mathcal{V}} V_{u,v}^a - \sum_{w \in \mathcal{V}} V_{v,w}^a = 0 \quad \forall a \in \mathcal{A}, v \in \mathcal{V} \setminus \{\underline{s_a}, \bar{t_a}\} \tag{4.16}$$

$$\sum_{u \in \mathcal{N}_{\underline{s_a}}^+} V_{\underline{s_a},u}^a = 1 \quad \forall a \in \mathcal{A} \tag{4.17}$$

$$\sum_{u \in \mathcal{N}_{\bar{t_a}}^-} V_{u,\bar{t_a}}^a = 1 \quad \forall a \in \mathcal{A} \tag{4.18}$$

$$\sum_{u \in \mathcal{N}_{\underline{t_r'}}^-} V_{u,\bar{t_r'}}^a = 1 \quad \forall r \in \bar{R_a}, \forall a \in \mathcal{A} \tag{4.19}$$

$$\sum_{u \in \mathcal{N}_{\underline{s_a'}}^+} V_{\underline{s_a'},u}^a = 1 \quad \forall r \in \bar{R_a}, \forall a \in \mathcal{A} \tag{4.20}$$

(4.16) insures that for all edges which do not lead to a source or destination, if a reaches v from a road, an incoming flow will lead to an outgoing one. In other words, they guarantee connections between roads. (4.17) - (4.18) insure that each universal source and destination is reached ones. Similarly, (4.17) - (4.20) achieves the same result, but for each requests. It should be mentioned that to guarantee that those special nodes are reached at most once one can simply change from equalities to inequalites, like for e.g. $\sum_{u \in \mathcal{N}_{\underline{s_a}}^+} V_{\underline{s_a},u}^a \geq 1$.
In order to find the minimum number of vehicles required, one can use the method described *Brodo* in [4].

Constrained Version

While it could bring some interesting insights, the URP does not contain the necessary information to provide efficient routing for the scenario considered in this work. The goal is to identify the best possible path that (i) satisfies all the requests and ii is congestion free. The *Holistic Congestion-Free Routing Problem (HCRR)* is formally defined as follows.
Given a transportation network $\mathcal{G}$, a set of vehicles $\mathcal{A}$ and a set of requests $\mathcal{R}$, defined within the description in section Section 4.1, solve:

$$\min \text{ (4.11), (4.12), (4.13) or (4.15)}$$

$$\text{s.t.}$$

$$\text{(4.16)-(4.20)}$$

$$\sum_{a \in \mathcal{A}} V^a_{u,v} \leq c_{u,v} \quad \forall (u,v) \in \mathcal{E} \tag{4.21}$$

$$V^a_{u,v} \cdot (s^a_u + T^a_{u,v} - s^a_v) \leq 0 \quad \forall (u,v) \in \mathcal{E}, \forall a \in \mathcal{A} \tag{4.22}$$

$$a'_v \leq s^a_v \leq b'_v \quad \forall v \in \mathcal{E}, \forall a \in \mathcal{A} \tag{4.23}$$

$$\sum_{(u,v) \in \mathcal{E}} e^a_{u,v} \cdot V^a_{u,v} \leq B_a(0) \quad \forall a \in \mathcal{A} \tag{4.24}$$

(4.21) insures the number of vehicles in the link (u, v) do not exceed the capacity of that link. (4.22) establishes the relationship between the service time of each node, implying that the service time of a predecessor must be lower than the successor. (4.23) establish the time window constraints, indicating that it must be within the interval of the request. For nodes which are not associated with a termination node of a request, one will simply set $a'_v = 0$ and $b'_a = \infty$, also insuring that $s^a_v \geq 0$. (4.24) assures that the vehicle charge is enough to cover all path. $B_a(0)$ can be assumed to be 100, i.e. that the batteries are full at the beginning of service.

Combination with Dispatching

Multiple solutions have already been proposed in literature which solve the dispatching and routing problem in the same work.
In this work, however, in order to combine dispatching and routing into the same formulation, one must adapt some of the conditions specified in previous sections above. More specifically, each equation related to the requests must be changed to accomodate the fact that requests have not been previously assigned.
Accordingly, the *Holistic Routing and Dispatching Problem (HRDR)* can be formulated as

$$
\begin{aligned}
&\min \ (4.10), (4.11), (4.12), (4.13) \text{ or } (4.15) \\
&\text{s.t.} \\
&\quad (4.5),(4.6) \\
&\quad (4.8),(4.9) \\
&\quad (4.16)\text{-}(4.18) \\
&\quad (4.21)\text{-}(4.24) \\
&\quad \sum_{u \in \mathcal{N}^+_{\bar{t'_r}}} V^a_{u,\bar{t'_r}} = x_{a,r} \quad \bar{t'_r} \in r, \forall r \in \mathcal{R}, \forall a \in \mathcal{A} && (4.25) \\
&\quad \sum_{u \in \mathcal{N}^+_{\underline{s'_r}}} V^a_{u,\underline{s'_r}} = x_{a,r} \quad \underline{s'_r} \in r, \forall r \in \mathcal{R}, \forall a \in \mathcal{A} && (4.26) \\
&\quad x_{a,r} \cdot (s^a_{\underline{s'_r}} - s^a_{\bar{t'_r}}) \leq 0 \quad \forall r \in \mathcal{R}, \forall a \in \mathcal{A} && (4.27)
\end{aligned}
$$

(4.25) guarantees that when a request r is allocated to vehicle a, the vehicle must reach the terminal station of the request at least once. Similarly, equation (4.26) ensures that the vehicle traverses the starting station associated with the assigned request. Equation (4.27) serves the purpose of ensuring that the starting station is visited before the terminal station.

4.2.3 Rebalancing

In simpler terms, the rebalancing problem revolves around efficiently redistributing autonomous vehicles (AVs) to optimize their responsiveness to new ride requests while minimizing any existing imbalances in the system. The goal is to fine-tune the positioning of AVs, ensuring they are strategically placed to promptly meet user demands and address any inherent irregularities in the distribution of service requests. This challenge is particularly crucial in ride-sharing systems and transportation systems alike, where the dynamic nature of user requests and varying demand across different locations can lead to imbalances in the fleet's distribution. Effectively tackling the rebalancing problem enhances the overall efficiency of the system, providing users with quicker response times and a more evenly distributed service, ultimately contributing to a smoother and more reliable autonomous transportation network.
The formulation of the rebalancing problem is partially inspired from the works of *Zhang et al.* in [28] and *Wallar et al.* in [35].
The model previosuly described allows to reason on the rebalancing model in a novel manner when compared to the literature reviewed in Chapter 3. More specifically one could deal with the rebalancing by leveraging the request arrival rate λ_r and the fact that vehicles are assumed to have a starting and terminal station, $\underline{s_a}$ and $\bar{t_a}$ respectively.

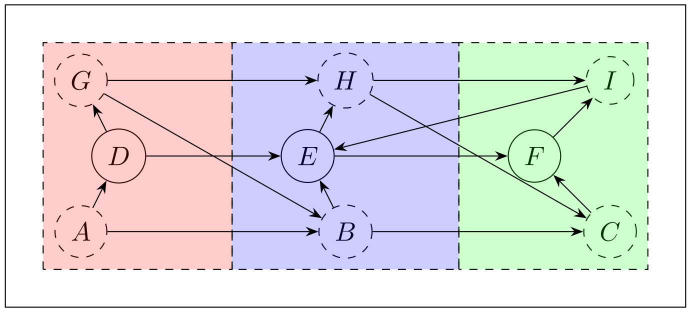


Figure 4.3: Simplfied example for the rebalancing strategy. Nodes D, E, F (continuos circles) to be the depot (or charging stations), while the other nodes (dashed circles) to be normal nodes. The graph has been previously divided in three areas (red, blue and green).

Considering the situation depicted in Fig 4.3 with a set of requests $\mathcal{R}'$ and a sets of vehicles $\mathcal{A}'$. According to this example, the graph is contained within the space Ω, which has been divided in three cells (in red, blue and green). This division has been done arbitrarily in this case, but different strategies can be used ([35], [44]). These regions must be specified, however, in such a way that they contain exactly one node belonging to $\mathcal{V}_d \cup \mathcal{V}_c$ and more than $n \in \mathbb{Z}_{\geq 1}$ nodes belonging to $\mathcal{V}_n$. Informally, this means that those regions are build around each charging station or depot and contain a certain amount of normal nodes. Different strategies could be used to determin this amount, like for e.g. all the nodes within a radius r from each terminal node or within a driving distance T_{max}. Nevertheless, as a result of this, one will obtain a set of regions R. Formally, each region around a node $v \in \mathcal{V}_d \cup \mathcal{V}_c$ can be defined as a subgraph $\mathcal{G}_v = \langle \mathcal{V}'_v, \mathcal{E}'_v \rangle$, where

$$\mathcal{V}'_v = \{u \in \mathcal{V}_n : (u, v) \in \mathcal{E} \wedge (v, u) \in \mathcal{E}, f(v, u) = 1\} \cup \{v\} \tag{4.28}$$
$$\mathcal{E}'_v = \{(u, w) \in \mathcal{E} : u, w \in \mathcal{V}'\} \tag{4.29}$$

In (4.28), the function $f : \mathcal{V}_d \cup \mathcal{V}_c \times \mathcal{V}_n \to \{0, 1\}$ is used to establish whether a node u belongs to the region of v ($f(v, u) = 1$) or not ($f(v, u) = 0$), as discussed before. Accordingly, one can obtain the total number of requests of region v ($|\mathcal{R}^v|$) by considering the requests which have a terminal node in $\mathcal{V}_n$. In other words

$$|\mathcal{R}^v| = \sum_{r \in \mathcal{R}'_v} \lambda_r \tag{4.30}$$

where

$$\mathcal{R}'_v = \{r \in \mathcal{R} : \underline{s_r} \in \mathcal{V}'{}_v\} \tag{4.31}$$

Requests, therefore, are distributed over those regions and it is not hard to envision scenarios where requests are unequally distributed, i.e. regions with a

higher $|\mathcal{R}^v|$ than others. As a result, due to the heterogenous nature of the nodes and the requests, AToD might experience imbalance, as AVs can be scattered through the whole graph if requests have different terminal nodes. Conversely, the opposite might happen, if all the requests have the same terminal node, but start from different areas.
Assuming all the requests in $\mathcal{R}'$ to be served, each vehicle's starting position is now either D, E, or F, i.e. $\underline{s_a} \in \{D, E, F\} \quad \forall a \in \mathcal{A}'$. This is due to the fact that, according to the transportation network model, since the requests have all been served, the vehicles have all reached their previous ending stations $\bar{t_a}$, which became their new starting nodes.
As a result, the solution of the rebalancing problem becomes ensuring that there is a sufficient number of vehicles at the depot (or charging) nodes to fulfill as many requests as possible in the given area, potentially covering all requests.

Unconstrained Rebalancing Problem

To formulate the different versions of the rebalancing problem binary flow variables will be introduced to help with the different formulations. Similarly to Section 4.2.2, those variables will be useed to indicate whether an idle vehicle, i.e. a vehicle not currently transporting customers or goods, should be moving from a node to another. Formally, this is expresseed as follows.

$$y^a_{u,v} = \begin{cases} 1 & \text{if } a \text{ traverses } (u,v) \in \mathcal{E} \\ \\ 0 & \end{cases} \qquad \forall a \in \mathcal{A}, \forall u, v \in \mathcal{V}$$

Within the assumptions made at the beginning of this section, it follows that a vehicle will be considered idle if it is currently at its final station.
For the sake of simplicity and brevity, a generic cost function $\mathcal{J}'$ will be considered for the entirity of this section since the choice of a cost function is similar to the one for Section 4.1.
Finally, within the model established in prior sections, the *Unconstrained Rebalancing Problem (UReP)* can be formulated as following.
Given a transportation network $\mathcal{G}$ and a set of idle vehicles $\mathcal{A}'$, defined within the description in section Section 4.1, solve:

$$\begin{aligned} \min \quad & \mathcal{J}' \\ \text{s.t.} \quad & \\ & \sum_{u \in \mathcal{V}} y^a_{u,v} - \sum_{w \in \mathcal{V}} y^a_{v,w} = 0 \quad \forall a \in \mathcal{A}', v \in \mathcal{V} \setminus \{\underline{s_a}, \bar{t_a}\} \end{aligned} \tag{4.32}$$

4.3 Real-time Formulation

4.4 Use Case/ Simulation / Evaluation

[7]

Chapter 5

Part 2 - TBD

Chapter 6

Summary and Outlook

In this work, we only considered one solution to solve the dispatching problem. While the solution found is guaranteed to be optimal, other algorithms could be used as well. For example, one possible approach is to develop an ad-hoc algorithm for it. For instance, requests could be assigned greedily to vehicles according to some heuristics, such as nearest neighbours, or according to the vehicle status, i.e. state of charge, capacity or traveling path. Requests should be assigned in such a way that are still in line with the previous path (see Fig 4.2). One could, for example, consider as input the previously calculated path for it.

Bibliography

[1] The dynamic vehicle allocation problem with application in trucking companies in brazil. *Computers and Operations Research*, 76:118–133, 2016.

[2] Ahmed Abdulaal, Mehmet H. Cintuglu, Shihab Asfour, and Osama A. Mohammed. Solving the multivariant ev routing problem incorporating v2g and g2v options. *IEEE Transactions on Transportation Electrification*, 3(1):238–248, 2017.

[3] Javier Alonso-Mora, Samitha Samaranayake, Alex Wallar, Emilio Frazzoli, and Daniela Rus. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences*, 114(3):462–467, 2017.

[4] Luca Brodo. Autonomous mobility-on-demand: A novel framework to study amod in smart cities. Paper not yet submitted for publication, 2023.

[5] Carlos Carrion and David Levinson. Value of travel time reliability: a review of current evidence. *Transportation Research Part A: Policy and Practice*, 46:720–741, 2012.

[6] Andrea Carron, Francesco Seccamonte, Claudio Ruch, Emilio Frazzoli, and Melanie N. Zeilinger. Scalable model predictive control for autonomous mobility-on-demand systems. *IEEE Transactions on Control Systems Technology*, 29(2):635–644, 2021.

[7] B. Donovan and D. B. Work. New york city taxi trip data (2010-2013). `http://dx.doi.org/10.13012/J8PN93H8`, 2014.

[8] Fabian Fehn, Florian Noack, and Fritz Busch. Modeling of mobility on-demand fleet operations based on dynamic electricity pricing. In *2019 6th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pages 1–6, 2019.

[9] Aurélien Froger, Ola Jabali, Jorge E. Mendoza, and Gilbert Laporte. The electric vehicle routing problem with capacitated charging stations. *Transportation Science*, 56(2):460–482, 2022.

[10] German Aerospace Center (DLR). Nemo Bili - Mobility from the First to the Last Mile, Year of access. Accessed on: 04.01.2024.

[11] Kyoungseok Han, Tam W. Nguyen, and Kanghyun Nam. Battery energy management of autonomous electric vehicles using computationally inexpensive model predictive control. *Electronics*, 9(8), 2020.

[12] Deborah A. Hennessy and Donald L. Wiesenthal. Traffic congestion, driver stress, and driver aggression. *Aggressive Behavior*, 25:409–423, 1999.

[13] Andreas Holzapfel, Heinrich Kuhn, and Michael G. Sternbeck. Product allocation to different types of distribution center in retail logistics networks. *European Journal of Operational Research*, 264(3):948–966, 2018.

[14] M. Hyland and H. S. Mahmassani. Dynamic autonomous vehicle fleet operations: optimization-based strategies to assign avs to immediate traveler demand requests. *Transportation Research Part C: Emerging Technologies*, 92:278–297, 2018.

[15] Brian Kallehauge, Jesper Larsen, Oli Madsen, and Marius Solomon. *Vehicle Routing Problem with Time Windows*, pages 67–98. 03 2006.

[16] S. R. Kancharla and G. Ramadurai. Electric vehicle routing problem with non-linear charging and load-dependent discharging. *Expert Systems With Applications*, 160:113714, 2020.

[17] G. Laporte. Fifty years of vehicle routing. *Transportation Science*, 43:408–416, 2009.

[18] Chungmok Lee. An exact algorithm for the electric-vehicle routing problem with nonlinear charging time. *Journal of the Operational Research Society*, 72:1–24, 03 2020.

[19] Michael W. Levin, Kara M. Kockelman, Stephen D. Boyles, and Tianxin Li. A general framework for modeling shared autonomous vehicles with dynamic network-loading and dynamic ride-sharing application. *Computers, Environment and Urban Systems*, 64:373–383, 2017.

[20] C Robin Lindsey and Erik T Verhoef. Congestion modelling. Technical report, Tinbergen Institute Discussion Paper, 1999.

[21] Haining Liu, Ijaz Haider Naqvi, Fajia Li, Chengliang Liu, Neda Shafiei, Yulong Li, and Michael Pecht. An analytical model for the cc-cv charge of li-ion batteries with application to degradation analysis. *Journal of Energy Storage*, 29:101342, 2020.

[22] Alejandro Montoya, Christelle Guéret, Jorge E. Mendoza, and Juan G. Villegas. The electric vehicle routing problem with nonlinear charging function. *Transportation Research Part B: Methodological*, 103:87–110, 2017. Green Urban Transportation.

[23] Heinz Neuburger. The economics of heavily congested roads. *Transportation Research*, 5(4):283–293, 1971.

[24] Zi-Hao Nie, Qiang Yang, En Zhang, Dong Liu, and Jun Zhang. Ant colony optimization for electric vehicle routing problem with capacity and charging time constraints. In *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 480–485, 2022.

[25] Bureau of Public Roads. Traffic assignment manual. Technical report, U.S. Dept. of Commerce, Urban Planning Division, 1964.

[26] P. Pant and R. M. Harrison. Estimation of the contribution of road traffic emissions to particulate matter concentrations from field measurements: a review. *Atmospheric Environment*, 77:78–97, 2013.

[27] Marco Pavone, Stephen L Smith, Emilio Frazzoli, and Daniela Rus. Robotic load balancing for mobility-on-demand systems. *The International Journal of Robotics Research*, 31(7):839–854, 2012.

[28] Federico Rossi, Rick Zhang, and Marco Pavone. Autonomous vehicle routing in congested road networks. *CoRR*, abs/1603.00939, 2016.

[29] Mauro Salazar, Matthew Tsao, Izabel Aguiar, Maximilian Schiffer, and Marco Pavone. A congestion-aware routing scheme for autonomous mobility-on-demand systems. pages 3040–3046, 06 2019.

[30] David Schrank, Bill Eisele, and Tim Lomax. Texas transportation institute 2012 urban mobility report. Technical report, Texas Transportation Institute, Texas A&M University, College Station, TX, 2012.

[31] Stephen L. Smith, Marco Pavone, Mac Schwager, Emilio Frazzoli, and Daniela Rus. Rebalancing the rebalancers: optimally routing vehicles and drivers in mobility-on-demand systems. In *2013 American Control Conference*, pages 2362–2367, 2013.

[32] P. Toth and D. Vigo. *Vehicle Routing: Problems, Methods, and Applications*. Society for Industrial and Applied Mathematics, Philadelphia, 2014.

[33] Paolo Toth and Daniele Vigo. *The Vehicle Routing Problem*. Society for Industrial and Applied Mathematics, 2002.

[34] Erik T Verhoef. Time, speeds, flows and densities in static models of road traffic congestion and congestion pricing. *Regional Science and Urban Economics*, 29(3):341–369, 1999.

[35] Alex Wallar, Menno Van Der Zee, Javier Alonso-Mora, and Daniela Rus. Vehicle rebalancing for mobility-on-demand systems with ride-sharing. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4539–4546, 2018.

[36] Salomón Wollenstein-Betech, Arian Houshmand, Mauro Salazar, Marco Pavone, Christos G. Cassandras, and Ioannis Ch. Paschalidis. Congestion-aware routing and rebalancing of autonomous mobility-on-demand systems in mixed traffic. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–7, 2020.

[37] Salomón Wollenstein-Betech, Ioannis Ch. Paschalidis, and Christos G. Cassandras. Joint pricing and rebalancing of autonomous mobility-on-demand systems. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 2573–2578, 2020.

[38] James J. Q. Yu and Albert Y. S. Lam. Autonomous vehicle logistic system: Joint routing and charging strategy. *IEEE Transactions on Intelligent Transportation Systems*, 19(7):2175–2187, 2018.

[39] James J. Q. Yu, Wen Yu, and Jiatao Gu. Online vehicle routing with neural combinatorial optimization and deep reinforcement learning. *IEEE Transactions on Intelligent Transportation Systems*, 20(10):3806–3817, 2019.

[40] Gioele Zardini, Nicolas Lanzetti, Marco Pavone, and Emilio Frazzoli. Analysis and control of autonomous mobility-on-demand systems. *Annual Review of Control, Robotics, and Autonomous Systems*, 5(1):633–658, 2022.

[41] Jannik Zgraggen, Matthew Tsao, Mauro Salazar, Maximilian Schiffer, and Marco Pavone. A model predictive control scheme for intermodal autonomous mobility-on-demand. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 1953–1960, 2019.

[42] Rick Zhang. *Title of the Thesis*. Ph.d. thesis, Stanford University, Department of Aeronautics and Astronautics, 2016. Submitted to the Department of Aeronautics and Astronautics.

[43] Rick Zhang and Marco Pavone. A queueing-network approach to the analysis and control of mobility-on-demand systems. *Proceedings of the American Control Conference*, 2015, 09 2014.

[44] Pengbo Zhu, Isik Sirmatel, Giancarlo Trecate, and Nikolas Geroliminis. Idle-vehicle rebalancing coverage control for ride-sourcing systems. pages 1970–1975, 07 2022.