

Master Project

Experimental Comparison of Autonomous Vehicles Routing Optimization Algorithms

Author:	Prisca Aeby ¹	prisca.aeby@epfl.ch
Supervisors:	Bastien Rojanawisut ²	bastien.rojanawisut@bestmile.com
	Boi Faltings ³	boi.faltings@epfl.ch

July 6, 2017

¹ Computer Science, École Polytechnique Fédérale de Lausanne, Switzerland

² Scala Backend Software Engineer, BestMile, Switzerland

³ Artificial Intelligence Laboratory, School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne, Switzerland, liawww.epfl.ch

Abstract

Your abstract.

Contents

1	Introduction	4
2	Related Literature	4
2.1	Public Transport	4
2.2	Demand-Adaptive Transit Systems	6
2.3	Vehicles Assignment to Routes	6
3	Problem Formulation	7
3.1	Circular Route	7
3.2	Vehicles	7
3.3	Bookings	7
3.4	Output Metrics	7
4	Methodology	10
4.1	Simulation Framework	10
4.2	Scheduling	12
4.2.1	Vehicles' Activities Schedule	12
4.2.2	Headway	12
4.2.3	Dynamic Fleet Size	12
5	Numerical Experiments	12
5.1	Simulation Settings	12
5.2	Graph	12
5.3	Optimal Headway	13
5.4	13
5.5	Simulated Demand	13
5.6	13
6	Conclusion	13

1 Introduction

General problem: two types of services, on-demand and fixed-line => explain why we focus on fixed line (all constraints of the platform) Fixed line because now roads are not made for more complex systems

2 Related Literature

The scope of vehicle scheduling problems related to our specific case study can range from fixed-line bus services to on-demand systems. In fact, electric shuttles' routes and stops are predefined but timetables can be flexible to respond to the known real-time demand. The set of constraints differs from traditional bus transportation, for example workforce regulations can be eliminated as the shuttles are not human driven but a battery management component needs to be introduced into the planning process. As we will see in this chapter, current research often concentrates on one specific aspect of the scheduling strategy, simplifying some constraints, assessing the system's performance using various types of metrics and considering different inputs. In what follows, we describe in section 2.1 some techniques used in the traditional public transport industry to guarantee a good quality of service. In section 2.2 we describe the concept of Demand-Adaptive Transit Systems. Finally, we detail in section 2.3 an approach proposed by the Urban Transport Systems Laboratory at EPFL (LUTS) for scheduling autonomous vehicles activities.

2.1 Public Transport

Within the standard transit organization, policies and standards affect a lot the development of transport strategies and how people interact with the systems. The planning process of a fixed line bus transportation service is mainly composed of three different tasks: route planning, service frequencies and service timing. Route planning consists of defining a sequence of stops composing each route and how those routes are interconnected. Service frequency captures the number of vehicles per unit time which pass a given route (often expressed in vehicles per hour). A common measure used to express the ideal distance between vehicles is the headway, the inverse of the frequency, in other word the fixed interval at which vehicles are coming at a station. The service frequencies set by the transport organization can be chosen based on different policies (see Pine et al. (1998)):

1. **Fixed headway:** the agency establishes a fixed interval at which vehicles come to stations. It is convenient for customers as they have access to an exact schedule, but it is hard to keep the time between vehicles constant as it is vulnerable to external disturbances.
2. **Demand-based headway:** in existing transport services, agencies can adapt the timetables based on the observed demand at the stations (number of passenger boardings/deboardings) in order to reach the desired passenger load in vehicles.
3. **Performance-based headway:** the goal is to find the headway for which performance standards are optimized. Those costs are usually measured during a

service period, for example a day. It may include the service productivity (e.g. the revenue per passenger per hour), the cost effectiveness of the service or the overall effectiveness (e.g. net subsidy per passenger).

All those scheduling strategies suffer from the well-known bunching effect: two or more buses arrive at the same time at a stop with the first one being overcrowded and the other one empty. This phenomenon occurs because of external disruptions in the service (e.g. stochastic passenger arrivals at stops, traffic jams, etc) (see Camps and Romeu (2016)) slowing down one of the bus which pick up passengers who would have taken the next bus.

Several corrective measures based on different strategies have been developed to overcome this problem. There are mainly two different holding approaches to mitigate bus bunching as described in van Oort et al. (2010):

- **Headway-Based Holding:** vehicles arriving with a shorter headway at a stop (or holding point) wait to restore the headway distribution. If they arrive with a longer headway, it is possible to speed up the buses by skipping stations. The analytical study Cortés et al. (2010) proved the efficiency of using headway holding strategy and bus skipping (considering the extra waiting time of passenger whose station has been skipped) with the two-dimensional objective function composed of the regularization of bus headways on the one hand and the level of service with respect to a circle route scenario.
- **Schedule-Based Holding:** in the case of fixed timetables, schedule-based holding involves holding a vehicle at a stop if it is ahead of its schedule and dispatch it immediately otherwise.

Zhao et al. (2016) proposed a method using boarding limits at stations to control buses which does not involve bus accelerating or waiting at stations and thus does not influence the customers' travel time and does not disturb the traffic. They do not consider schedule and a priori target headway. However, they proved that their self-adjusting control stabilizes the headway spontaneously in a short time.

Adra et al. (2004) study the importance of the vehicle load effects on the emission of vehicles through the study of the load factor and the empty running rate. The passenger load factor measures the efficiency of a transportation system, representing the capacity utilization of the seats. It is often expressed as the ratio between the passenger-kilometres travelled to seat-kilometres available.

As stated by He (2015), the majority of earlier studies conducted on maintaining a balanced headway uses arrival time of the current bus at the current stop and arrival times of the preceding buses but does not take advantage of real time information like the vehicles' geolocation. Later methods proposed new approaches assuming availability of locations and even real-time arrivals of passengers to each bus stop. For example, Daganzo and Pilachowski (2011) proposed a solution taking into account the distance between the current bus and the preceding and following buses to adjust the speed of the current bus. It includes holding buses at stations, accelerating or decelerating.

The studies carried out within the traditional fixed-line bus services handle situations where the demand is consistently strong over the territory and where the fleet is

composed of high-capacity vehicles. When the demand is weaker, it is complicated to operate an economical and frequent transit system as the resources are shared by few people and are very costly (e.g. driver salaries, fuel expenses, etc.). The autonomous fleets are composed of vehicles with lower capacity, but it is easier to dynamically adapt their scheduling strategy at low cost and have access to real-time information.

2.2 Demand-Adaptive Transit Systems

Demand-Adaptive System (DAS), or Demand-Responsive Transit (DRT), is a personalized type of transportation displaying features from both fixed-line services and on-demand systems. In such systems, passengers are picked up and dropped off on-demand at predetermined stops, and unrequested stops are skipped. There are still compulsory stops, which are served within a predefined time window, but the routes are adapted based on the the requested stops. A method to determine the time windows at the compulsory stops has been proposed by Crainic et al. (2010). Li et al. (2007) depicted the advantages to substitute fixed-line bus services to DRT services in two cities in California with low demand density.

Gabriel et al. (2008) address the issues of evaluating DAS services in comparison to traditional transit services and fully on-demand systems. The transit systems' evaluation denotes the part of the planning process dedicated to the study of the behaviour and the performance of the line under various conditions with respect to demands, policies and costs. They mention the main steps of transit lines' evaluation in order to tune operating parameters impacting the system performance under different scenarios for cost-benefit analyzes: 1) the scenario, parameters and policies are specified, as well as the demand to serve 2) the line is designed 3) the operation of the line is simulated (during a specified time-horizon) 4) results are collected and performance measures are computed. The importance of demand generation in step 1) differs from one system to the other: whilst in fixed lines transits static methods are often used to simulate an average demand, dynamic methods are used to simulate the time-dependent stop-to-stop requests in on-demand systems. The performance measurements in step 4) vary as well from one system to the other: in fixed-lines the service quality is mainly represented by the buses' punctuality and the constant interarrival times whereas for on-demand systems it is expressed by the specific users' waiting time and journey time. Gabriel et al. (2008) propose a framework for evaluating DAS lines including the scenario input, the optimization modules for the design and operation of the line and the simulation module which yield the statistical information about the performance.

2.3 Vehicles Assignment to Routes

Bongiovanni et al. (2016) present an optimization approach for the autonomous vehicles scheduling within a transit system composed of circular routes. They separate the planning process in two main tasks: line planning and vehicle scheduling. Line planning corresponds to defining the routes and the required headway on each route over the planning horizon. Vehicle scheduling refers to computing a general assignment plan for each vehicle for each time period over the planning horizon of one day. It means that each vehicle is assigned to a specific route or to a recharging station for each time period. The underlying assumption is that the vehicles will not transfer between routes and charging stations too often which makes the time discretization of the planning horizon possible. They focus on the vehicle scheduling task which takes

as input the predefined routes, frequencies, fleet size and battery model. They present the scheduling problem as a mixed-integer linear programming (MILP) with battery constraints and the goal of the objective function being to maximize the total battery charge levels at the end of the planning horizon.

3 Problem Formulation

In this section we formulate the mathematical model describing the autonomous vehicles circular line system. All the variables are listed in Table 1. In addition to traditional bus transit system descriptions, we introduce the concept of battery stations and we include dynamic bookings from one station to an other station on the loop. As we have access to information during the service horizon time H , we denote a discrete timestamp t which takes values between $[0, H]$ at fixed sampling *interval*.

3.1 Circular Route

The circular line in this transportation system is a route which starts at a certain station, traverses other stations once and terminates at the starting station. The route is represented by set of stations and a set of directed edges E linking each station $s \in S$ to its following station. The set of charging stations C includes stations located on the stops within the loop: $C \subset S$. The maximum speed of an edge, representing the traffic at time t , is expressed as $speed_t^e$.

3.2 Vehicles

Each vehicle composing the fleet $v \in V$ moving on the circular closed line is characterized by a capacity c^v , a maximum charge q^v , a recharge rate α^v , a battery consumption rate β^v and a maximum speed s^v . The battery level of vehicle v at the timestamp t is $I_t^v \in [0, q^v]$. The battery change between the preceding discrete timestamp and the current timestamp t is given by δ_t^v . The edge it is currently traversing is e_t^v and the distance to the next vehicle in front of it on the loop is given by $next_t^v$. The occupancy of the vehicle is expressed as $o_t^v \in [0, c^v]$. The vehicle is travelling either at its maximum speed or at the limitation of the current edge. Its current speed is therefore $\max\{s^v, speed_t^{e_t^v}\}$. When a vehicle is charging at a station the boolean variable x_t^v is set to 1 and it can not carry any passengers $o_t^v = 0$.

3.3 Bookings

The set B represents the users' bookings made during the service horizon. A booking $b_{i,j} \in B$ requests a trip from station i to j at time t^b for a group of n^b persons. If the booking has been satisfied, the boolean variable D_b is set to 1 and we know that it has been executed by vehicle $v^b \in V$. The pickup time of booking b at station i is p^b and drop-off time at station j is d^b . The waiting time w^b equals $p^b - t^b$ and the journey time Δ^b is $d^b - p^b$. We denote B^* the set of bookings which have been satisfied, so the bookings for which $D_b = 1$.

3.4 Output Metrics

At the end of the service horizon T , some metrics can be derived in order to evaluate the performance of the overall system dispatching the vehicles to serve the requests.

	Signification
$S = \{1, \dots, n\}$	Set of stations on the route
$E = \{n + i, \dots, 2n\}$	Set of directed edges linking the stations
$V = \{1, \dots, v\}$	Set of vehicles composing the fleet
$C \subset S$	Set of charging locations
$B = \{1, \dots, b\}$	Set of bookings made during the scheduling horizon
$B^* \subset B$	Set of bookings which have been satisfied at the end of H
H	Length of the scheduling horizon
$interval \in \mathbb{R}$	Interval at which values are sampled
$T = \{t = interval * k, k \in \mathbb{N} 0 < t < H\}$	Set of timestamps
c^v	Maximum load capacity of vehicle $v \in V$
q^v	Maximum charge of vehicle $v \in V$
α^v	Battery recharge rate of vehicle $v \in V$
β^v	Battery consumption rate of vehicle $v \in V$
s^v	Maximum speed of vehicle $v \in V$
$speed_t^e$	Speed of edge e at time t
$I_t^v \in [0, q^v]$	Battery level of vehicle v at time t
δ_t^v	Battery change between t and its preceding timestamp
e_t^v	Edge $e \in E$ which vehicle v is traversing at time t
$next_t^v$	Distance to the next vehicle in front of v on the loop at time t
$o_t^v \in [0, c^v]$	Number of passengers in vehicle v at time t
$b_{i,j} \in B$	Request of trip from station i to station j , $i, j \in S$
t^b	Time at which the request $b \in B$ has been made
p^b	Pick-up timestamp of booking b
d^b	Drop-off timestamp of booking b
w^b	Waiting time of booking b
Δ^b	Journey time of booking b
C^s	1 if s is a charging station, 0 otherwise
x_t^v	1 if the vehicle v is charging at a station at time t , 0 otherwise
D_b	1 if request b has been satisfied

Table 1: Problem sets and parameters

Battery

The total battery consumed by a vehicle v is the sum of the battery changes over the scheduling horizon when the vehicle is not charging:

$$battery^v = \sum_{t \in T} \delta_t^v * (1 - x_t^v)$$

The total battery cost for the entire system is therefore the sum of the battery consumption of each vehicle composing the fleet:

$$batteryCost = \sum_{v \in V} battery^v$$

Occupancy

The average occupancy of a vehicle v over the scheduling horizon T is given by:

$$avgOccupancy = \frac{\sum_{t \in T} o_t^v}{|T|}$$

The average occupancy does not take into account of the time vehicles spend at charging stations and therefore can not carry any passenger. As explained in section 2.1, a common unit used in transportation measurement is the load factor, which expresses the efficiency of a vehicle based on the kilometres travelled by passengers over the total kilometres that could have been driven considering its maximum capacity. In our case, as we are mainly interested in how the battery of the vehicles has efficiently been used, we define the passenger-battery measure: it is the battery that has been consumed by the vehicle during the passenger journey time. The passenger-battery for one vehicle is therefore the sum of the passenger-battery of the bookings it has satisfied. The seat-battery available represents the maximum possible passenger-battery if the vehicle would have been always full: it is the battery consumption of the vehicle times its maximum capacity. The vehicles' load factor is its passenger-battery over its seat-battery:

$$loadFactor^v = \frac{\sum_{t \in T} \delta_t^v * (1 - x_t^v) * o_t^v}{battery^v * c^v}$$

We can compute the average load factor of the fleet as:

$$avgLoadFactor = \frac{\sum_{v \in V} loadFactor^v}{|V|}$$

We denote the stability load factor the difference between the maximum load factor among the vehicles composing the fleet and the minimum one. The closer it is to 0 the better as it implies that all vehicles are used equivalently: $stabilityLoadFactor = \max_{v^+ \in V} loadFactor^{v^+} - \min_{v^- \in V} loadFactor^{v^-}$.

Waiting Time

The average waiting time among the satisfied bookings is given by:

$$avgWaitingTime = \frac{\sum_{b \in B^*} w^b}{|B^*|}$$

The stability waiting time is the difference between the maximum waiting time and the minimum waiting time among the bookings: $stabilityWaitingTime = \max_{b^+ \in B^*} w^{b^+} - \min_{b^- \in B^*} w^{b^-}$.

Journey Time

The average journey time among the satisfied bookings is given by:

$$avgJourneyTime = \frac{\sum_{b \in B^*} \Delta^b}{|B^*|}$$

Completed Bookings Percentage

Travel time, booking time, At the end of the scheduling horizon, some bookings may not have been satisfied. The percentage of bookings which have been satisfied is given by:

$$completedBookings = \frac{|B^*|}{|B|}$$

4 Methodology

In order to evaluate the behaviour and performance of the system under various conditions with respect to the line configurations, the booking scenarios and the vehicles scheduling alternatives, we undertake analyses under laboratory conditions by simulating the the operation of the system on the platform dispatching the vehicles. One of our main goal in the analysis we conduct on the line is to get measures as close as possible to real world conditions. In order to achieve this, a dynamic system is required which handles the time-dependent events of the system. In section 4.1 we describe the behaviour of the different interconnected services simulating the vehicles' behaviour reacting to external conditions and how the overall system is evaluated. In section 4.2 we present the variations of scheduling strategies used to dispatch the vehicles on the line.

4.1 Simulation Framework

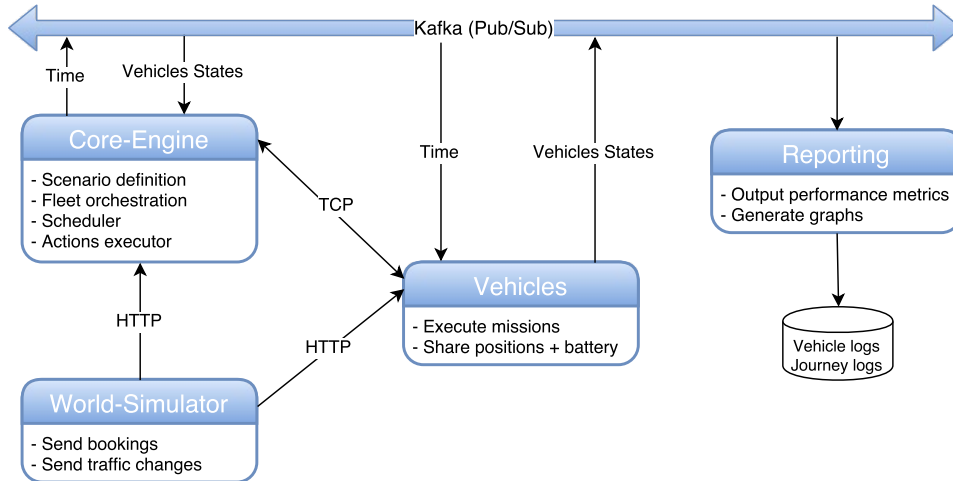


Figure 1: Interconnected modules composing the simulation framework

Figure 1 illustrates the different web services and how they are linked to each others. Services can communicate by exchanging messages via Kafka which is an asynchronous

publish-and-subscribe message broker. In what follows we describe the behaviour of each module and the simulation parameters which can be tuned.

World-Simulator

The world-simulator is a web service which is used to start the simulation, send the booking requests and send traffic updates information. At the beginning of the simulation, it sets the time control to a given ratio which is sent to the core-engine application. A day can be simulated in m minutes which corresponds to a ratio r of $m/24h$. The world-simulator is instantiated with a scenario describing the external events which affect the overall system. It mimics the bookings which would have been made by users with the booking application and sent through a REST API. Every second at time t it checks if there are bookings with $t^b = t$ and speed updates $speed_t^e$ to send to the core-engine application.

Simulated Vehicles

Vehicles are connected to the core-engine through an established TCP protocol. They can send status messages and receive missions, like going to a specific charging station or waiting at a stop for a given time. The vehicles are simulated concurrently following an actor model. Each vehicle is a different actor part of the actor system. When the vehicles web service receives an HTTP request to start the simulation with a given fleet size from the world-simulator, an actor is created for each vehicle, which announces its position to the core-engine through a TCP message and subscribes to the time topic. The vehicles start at equal distance of each others on the loop with the battery completely recharged and their position and battery level is then updated every $1000/rate$ milliseconds.

Core-Engine

The core-engine application is a distributed system that combines and processes real-time information for coordinating and optimizing the fleet of autonomous vehicles. Several scenario components are instantiated by the core-engine application: the map, the vehicles' configurations and fleet schedule through the day. The map is a GeoJSON file which is translated into the application to a directed graph with an edge between every stop and charging stations. The vehicles' configurations include all the parameters enumerated in section 3.2. The core-engine handles the bookings it receives from the world-simulator, updates the edges' speed of the graph and sends the actions to be executed by the vehicles. When it receives the starting simulation time and ratio r from the world-simulator, it publishes the updated time to the corresponding Kafka topic. The scheduler component manages the number of vehicles which need to be on service at different times during the day and sends corresponding missions to the vehicles.

Reported Metrics

The reporting web-service subscribes to the vehicle and journey topics and writes the stream events to its database. An event for vehicle v at time t includes its speed, position, battery level and battery state x_t^v . There is an event only for a booking which has been satisfied $b \in B^*$ and it includes all the parameters listed in section 3.3. At the end of the simulation, events are fetched from the database at fixed sampling *interval* and then those logs are analysed in order to compare different scenarios in term of their

overall performance. All the metrics listed in section 3.4. In addition, we represent visually the simulation's performances with different graphs. In order to evaluate the different behaviours among the vehicles, we plot the hourly average occupancy of each vehicle during the scheduling horizon H . It is useful to see if all the vehicles are carrying passengers through the scheduling period or only few of them. We compare as well the number of vehicles which are at charging stations to the hourly average waiting time. It gives a good indicator to determine if the time at which the vehicle has been sent to charge was appropriate or if it impacts a lot the service in terms of quality of service provided to the passengers.

4.2 Scheduling

The approach proposed by Bongiovanni et al. (2016) described in section 2.3 is formulated as an optimization problem which considers the optimal headway for each route as an input but does not specify how to define it. Many real world aspects are not considered in their model. For example, it does not include transfers costs to charging stations or the time it takes to stop at stations and board passengers. It does not evaluate any demand scenarios neither. Our approach is to focus on scenarios as close as possible to real world conditions and find heuristics to define the number of vehicles needed on the fixed line in order to have a good balance between the battery consumption and the quality of service offered to the passengers. In what follows we describe how the core-engine handles the missions it sends to the vehicles based on the dynamic events it receives.

4.2.1 Vehicles' Activities Schedule

when they arrive at station, if they can they pick up the booking (based on booking size) they drop off bookings that finish here they wait based on headway computation *battery* check if it is under threshold finish to drop off everybody and go to charging station

4.2.2 Headway

compute headway based on fleetorchestrationservice

4.2.3 Dynamic Fleet Size

how choose vehicles to send to charge and which ones to activate

5 Numerical Experiments

Real environment

5.1 Simulation Settings

tuning speed of actors logs fetched from db

5.2 Graph

Unless stated, charging locations at same place

5.3 Optimal Headway

5.4

5.5 Simulated Demand

5.6

6 Conclusion

References

- Adra, N., Michaux, J. L., and Andre, M. (2004). Analysis of the load factor and the empty running rate for road transport. Artemis - assessment and reliability of transport emission models and inventory systems. Rapport de recherche.
- Bongiovanni, C., Kaspi, M., and Geroliminis, N. (2016). Scheduling autonomous vehicles activities. Technical report, Urban Transport Systems Laboratory EPFL.
- Camps, J. M. and Romeu, M. E. (2016). Headway adherence. detection and reduction of the bus bunching effect. *AET papers repository*.
- Cortés, C. E., Sáez, D., Milla, F., Núñez, A., and Riquelme, M. (2010). Hybrid predictive control for real-time optimization of public transport systems’ operations based on evolutionary multi-objective optimization. *Transportation Research Part C: Emerging Technologies*, 18(5):757–769.
- Crainic, T. G., Errico, F., Malucelli, F., and Nonato, M. (2010). Designing the master schedule for demand-adaptive transit systems. *Annals of Operations Research*, 194(1):151–166.
- Daganzo, C. F. and Pilachowski, J. (2011). Reducing bunching with bus-to-bus cooperation. *Transportation Research Part B: Methodological*, 45(1):267–277.
- Gabriel, T., Fausto, C., Malucelli, E. F., and Nonato, M. (2008). A proposal for the evaluation of demand-adaptive transit systems. *Annals of Operations Research*.
- He, S.-X. (2015). An anti-bunching strategy to improve bus schedule and headway reliability by making use of the available accurate information. *Computers & Industrial Engineering*, 85:17 – 32.
- Li, Y., Wang, J., and Chen, J. (2007). *Design of a Demand-responsive Transit System*. California PATH working paper. California PATH Program, Institute of Transportation Studies, University of California at Berkeley.
- Pine, R., Niemeyer, J., and Chisholm, R. (1998). *Transit Scheduling: Basic and Advanced Manuals*. Report (Transit Cooperative Research Program). National Academy Press.
- van Oort, N., Wilson, N., and van Nes, R. (2010). Reliability improvement in short headway transit services. *Transportation Research Record: Journal of the Transportation Research Board*, 2143:67–76.
- Zhao, S., Lu, C., Liang, S., and Liu, H. (2016). A self-adjusting method to resist bus bunching based on boarding limits. *Mathematical Problems in Engineering*, 2016:1–7.