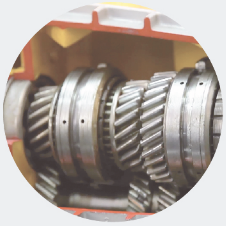
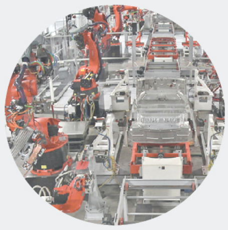
INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO
MESTRADO EM ENGENHARIA E GESTÃO INDUSTRIAL



URBAN NAVIGATION HANDLING OPENSTREETMAP DATA FOR AN EASY TO DRIVE ROUTE

ROSÁRIA SPINA
Outubro de 2022

POLITÉCNICO
DO PORTO

URBAN NAVIGATION – HANDLING OPENSTREETMAP DATA FOR AN EASY TO DRIVE ROUTE

Rosária Spina

2022

Instituto Superior de Engenharia do Porto

Mechanical Engineering Department

isep

P.PORTO

URBAN NAVIGATION – HANDLING OPENSTREETMAP DATA FOR AN EASY TO DRIVE ROUTE

Rosária Spina

Student n.º 1200199

Dissertation presented to the Instituto Superior de Engenharia do Porto to fulfil the requirements to obtain the Master's Degree in Industrial Management and Engineering, under the supervision of Doctor António José Galrão Ramos.

2022

Instituto Superior de Engenharia do Porto

Mechanical Engineering Department

P.PORTO

ACKNOWLEDGMENTS

In the first place, I would like to express my gratitude to my supervisor, Professor António Galrão Ramos, for his availability and guidance.

To my family, who have always supported me and my decisions. Their love and support made me believe in myself even in the most difficult moments.

To all my friends, for always being by my side and keeping me going through this process, for their friendship and encouragement.

To GoGlobal – Business Experts, my current employer, for giving me the opportunity to reconcile the work and the Master in Industrial Management and Engineering.

página propositadamente em branco

RESUMO

Atualmente, os cidadãos podem escolher as suas opções de viagem com base no tempo, distância, emissões, consumo, entre outros parâmetros. Não obstante, a literatura indica que os sistemas de planeamento de rotas atuais têm, maioritariamente, por base a distância e o tempo. Com efeito, verificou-se uma falta de sistemas de planeamento de rotas que se preocupem com as preferências dos utilizadores num ponto de vista mais qualitativo. Este projeto de investigação desenvolve um framework de planeamento de rotas com a integração de diferentes atributos da rede rodoviária como semáforos, passadeiras e paragens de autocarro, com o objetivo de providenciar aos utilizadores a opção de evitar estes mesmos atributos, oferecendo uma opção *easy drive*, nomeadamente em ambiente urbano. O estudo foi conduzido através de dados georreferenciados da cidade de Lisboa, Portugal. No entanto, é transferível para qualquer outra cidade. O algoritmo providencia alternativas para a rota mais curta, *easy drive* e rota balanceada, considerando apenas um modo de viagem: carro/mota.

O modelo foi desenvolvido no PostgreSQL com a extensão PostGIS e PgRouting, e os resultados foram visualizados no software QGIS. O software permite customizar pesos para cada uma das restrições para a escolha das rotas e estes pesos são modificados com o objetivo de encontrar o caminho ótimo consoante as preferências de cada utilizador.

PALAVRAS-CHAVE

Roteamento, Ambiente urbano, Dados georreferenciados, PgRouting

página propositadamente em branco

ABSTRACT

Currently, citizens can choose their travel options based on time, distance, consumption, emission, among other parameters. Nevertheless, the literature indicates that current route planning systems are based on distance and time. In fact, there is a lack of route planning systems which are concerned with users' preferences from a more qualitative point of view. This research project develops a route planning framework with the integration of different road network features like traffic lights, pedestrian crossings, and bus stops, to provide users with the option to avoid these features, offering an easy drive option, namely in an urban environment. The study was conducted using georeferenced data from the city of Lisbon, Portugal. However, it is transferable to any other city. The algorithm provides alternatives for the shortest route, easy drive, and balanced route, considering only one travel mode: car/motorbike.

The model was developed in PostgreSQL with the PostGIS extension and PgRouting, and the results were visualized in QGIS software. The software allows to custom weights for each of the constraints for route choices, and these weights are modified to find the optimal route according to the preferences of each user.

KEYWORDS

Routing, Urban environment, Georeferenced data, Easy Drive, PgRouting

página propositadamente em branco

INDEX

página propositadamente em branco

INDEX OF FIGURES

página propositadamente em branco

TABLE INDEX

página propositadamente em branco

LIST OF ACRONYMS

List of acronyms

CVRP	Capacitated Vehicle Routing Problem
DBMS	Database Management System
GIS	Geographic Information Systems
GVRP	Green Vehicle Routing Problem
HFVRP	Heterogeneous Fleet Vehicle Routing Problem
KPI	Key Performance Indicators
MDVRP	Multi-Depot Vehicle Routing Problem
OSM	OpenStreetMap
OVRP	Open Vehicle Routing Problem
PDVRP	Pick-up and Delivery Vehicle Routing Problem
PVRP	Periodic Vehicle Routing Problem
R&D	Research and Development
STO	Scientific and Technological Objectives
TDVRP	Time Dependent Vehicle Routing Problem
TSP	Travelling Salesman Problem
VGI	Volunteered Geographic Information
VRP	Vehicle Routing Problem
VRPB	Vehicle Routing Problem with Backhauls
VRPMT	Vehicle Routing Problem with Multiple Trips
VRPTW	Vehicle Routing Problem with Time Windows

página propositadamente em branco

1. INTRODUCTION

This first chapter introduces the research problem of this dissertation, highlights the main objectives of this work as well as its contribution to the existing literature and finally the methodological options adopted for the scientific and technological investigation of the problem.

1.1. Problem of research, framework and relevance

Nowadays, there is a variety of applications (Google maps, Maps, Waze, etc) that calculate different options for routes between the origin and the desired destination. When calculating routes, users can define certain parameters for their route, such as: avoiding highways, tolls and, more recently, choosing the type of car engine, to obtain the most economical route. Additionally, this type of navigation systems includes real-time information relevant to the route calculation, such as traffic information, accidents, or repair works. However, the current route planning systems are focused on (Engelmann *et al*., 2019; Ahmed *et al*., 2020): distance optimization (shortest path by distance), travel time optimization (fastest path) and, more recently, emission optimization (eco-friendly path).

A wide range of parameters that can be incorporated in the route calculation includes parameters such as (Huang *et al*., 2014): safety, ease of driving, panoramic route, fuel/energy consumption and more environmentally friendly route.

It is precisely in this notable need in the literature and in the market, that the present project is inserted, to develop a route planning solution incorporating Geographic Information System (GIS) objects and, additionally (differentiating feature of this project), integrate the drivers' preferences regarding comfort and ease of driving in an urban environment. In this project, driving comfort and ease are characterised by a set of road network features: junctions, roundabouts, stop signs, traffic lights, pedestrian crossings, bumps, and public transport stops.

Thus, this project undertakes to develop a new route option when choosing the route to travel between the origin and the desired destination.

In terms of scientific and technological motivation, the Research and Development (R&D) implicit in this project stands out, and its complexity, in itself, worked as an incentive.

Additionally, the author of the project intends to be active about scientific research, through rigorous scientific and technological surveillance and market practices and, with this, to strengthen his skills, promoting a practice of continuous improvement and innovation, as well as the development of studies that can lead to the generation of disruptive solutions.

In what specifically concerns this research project, the author intends to be able to explore GIS objects to develop a routing model that respects the preferences and needs of drivers in an urban environment, namely with respect to an easy drive mode (a mode that may vary depending on the profile of each driver).

The innovation of this work is the incorporation of such features of the road network like roundabouts, traffic lights, crossings, bus stops and bumps, that can make a route easier or harder to drive.

1.2. Research question and objectives

The objectives of the current study include integrating comfort and easy to drive mode as an optimizing parameter in route planning and providing route alternatives in the shortest, easy to drive and balanced scenarios (shortest plus easy to drive). Thus, the goal is to develop a routing model that offers a new functionality thar responds to the interests of users, developing a personalized routing model adapted to the needs and preferences of each user.

Considering the needs identified above, the following Scientific and Technologies Objectives (STO) were identified to promote the differentiation, innovation, and research of this project:

- Specify and develop a database structure for route planning needs, to include the following features:
    - Integration of multiple data sources with different characteristics;
    - Centralized and properly integrated data storage area that will provide added value for analysis;
    - Suitability of data for routing algorithms execution.
- Creation of a system capable of supporting the requirements for a smooth operation of the easy drive mode, of which the most important are:
    - Host different routing algorithms;
    - Performance, ensuring that value can be extracted from the data;
    - Scalability, ensuring that the system will not be limited by the number of drivers profiles;
    - Elasticity, making it possible to accommodate other data sources, i.e., data from different sources and formats, as well as relating to any geographical location on the globe;
- Develop a model that allows, at any time, to change route preferences according to each user;
- Develop an analytical service:
    - Visual analysis of the routes generated by executing routing algorithms;
    - Key Performance Indicators (KPI), to measure performance according to the parameters that are considered relevant to the quality of the system.

1.3. Methodological options

The research methodology adopted for the development of this work was based on the case studies methodology.

The case study is an investigative strategy through which one seeks to analyse, describe, and understand certain cases (of individuals, groups, or situations), which may subsequently lead to comparisons with other cases and formulate certain generalisations.

The case study fits into what Becker (1996) calls the craft model of science, to refer to a model in which the researcher acts as an intellectual craftsman, as he adapts and personalises the instruments according to his specific research objective.

The Case Study, due to its nature, can be represented by a methodology divided into three phases (Ludke and André, 1986):

1) Exploratory phase: phase where several issues coexist that will be made explicit as the research process evolves. The object of study may emerge from reading and analysing existing literature, as well as from the need to solve a particular problem or even from speculations based on the researcher's experience. It is at this stage that the investigation domain is specified, and the key elements and outlines of the problem are identified;
2) Data collection phase: after defining the research domain, the second phase consists of collecting the data and information necessary to conduct the study. This stage develops according to the object of study, the objectives and the data to be collected;
3) Analysis, interpretation, and dissemination of results phase: third and final stage which corresponds to the treatment, analysis and interpretation of the data collected, culminating in the presentation of the study's conclusions. In this phase, the analysis procedures, the interpretation of the data collected according to the object of study and the objectives defined prevail to guide the process and the presentation of the results.

For this specific research project, in the first phase, the state of the art was surveyed regarding the inclusion of road network features in route planning to identify gaps, determining the domain and research questions. The next step was to collect the necessary data to conduct the study. Finally, the architecture to analyse and interpret the data was developed to be able to discuss the results and answer the research questions.

1.4. Structure of work

After this introduction, the present dissertation is structures as follows. The literature review is demonstrated in section 2. Section 3 exhibits the methodology, describes the data and the architecture development to answer the research questions. Section 4 documents the results and findings. In conclusion (section **Error! Reference source not found.**), we highlight the main contributions and limitations of our study.

2. LITERATURE REVIEW

In this chapter, a review of the pertinent literature consulted for the elaboration of this dissertation can be found. The review essentially divides intro three parts: the vehicle routing problem, solution methods and a synthesis of recent studies that have developed models with disruptive constraints. The latter works as a theoretical bridge for the development of the scientific research. Regarding the vehicle routing, an intensive study of its history, characteristics, and its most common variants as well as the exact and approximate methods were carried out. A brief assessment concerning the characteristics of the presented papers is done as well.

2.1. Vehicle Routing

In this section, the variants that most logistics and distribution companies face in their daily operations regarding the routing of vehicles and scheduling of deliveries, as well the way the variants are related to real-life cases are analyzed.

2.1.1. Brief History of VRP

The Vehicle Routing Problem (VRP) is one of the most complex problems of combinatorial optimization and one of the most widely studied topics in the field of operations research. The paper by Dantzig et al. (1954) is known as the first study in the VRP literature, creating a mathematical VRP programming model through the truck dispatching problem. This problem implies the distribution of goods from a central warehouse to geographically dispersed clients. That study was followed by a huge amount of further VRP papers.

The VRP is a combinatorial optimization and integer programming problem, whose goal is to build and organize with a minimal total cost, a set of routes for a fleet of vehicles to serve a set of customers, each one with their own demand. Thus, the VRP has three common goals: to minimize the total distance travelled by the vehicles, to minimize the cost related with those travels or to maximize the occupation of the vehicles employed. VRP can be seen as a generalization of the Travelling Salesman Problem (TSP). Considering a set of locations representing cities and a set of distances between them, the TSP (Lin and Kernighan, 1973) consists of defining a route that starts in one of the cities, it passes through each city of the set exactly once and returns to the initial city of the route so that the total distance travelled is minimal.

In 1964, a new problem formulation was made by Clark and Wright (1964) that first incorporated more than one vehicle using a seminal heuristic method to solve it. Kulkarni and Bhave (1985) leads to a better understanding and a more complete overview of a variety of mathematical models for straightforward TSP and VRP problems.

Over the previous decades, many surveys on VRP models and their solutions have been performed and reported. Nevertheless, recent VRP models differ from those presented by Dantzig and Ramser and Clarke and Wright because since then, various constraints that influence the model have been proposed to integrate real world challenges. Kumar and Panneerselvam (2012) mention some examples of more complex constraints such time windows, time-dependent travel times, multiple depots, and heterogeneous fleets. These constraints added difficulty to the formulation of the routing problems and that's why solving VRPs is computationally expensive and categorized as NP-hard. Since the use of mathematical programming or combinatorial optimization is limited, thus the use of heuristics is common due to the size and frequency of the real-world problems.

2.1.2. VRP and Its Variants

In recent decades, VRP variants have received increasing attention given the considerable volatility of conditions in the different real-world scenarios. In this sense, there is a need to include in the VRP variants certain constraints and assumptions to develop models more appropriate to reality (Braekers et al., 2016), with the objective of determining solutions with a greater accuracy and potential of applicability. Consequently, it is extremely important to be able to identify the VRP variants when one intends to solve a real case, the formulation of the problem and the objective function, in a way that it's possible to choose the appropriate solution method and thus, solve the problem with the best possible efficiency.

From the studies carried so far, it is possible to identify a wide range of VRP variants, with applicability in real challenges. Of the existing ones, the most significant variants are the following:

- Capacitated Vehicle Routing Problem (CVRP);
- Vehicle Routing Problem with Multiple Trips (VRPMT);
- Pick-up and Delivery Vehicle Routing Problem (PDVRP);
- Multi-Depot Vehicle Routing Problem (MDVRP);
- Open Vehicle Routing Problem (OVRP);
- Periodic Vehicle Routing Problem (PVRP);
- Vehicle Routing Problem with Time Windows (VRPTW);
- Vehicle Routing Problem with Backhauls (VRPB);
- Heterogeneous Fleet Vehicle Routing Problem (HFVRP);
- Time Dependent Vehicle Routing Problem (TDVRP).

Capacitated Vehicle Routing Problem

The CVRP is a static and deterministic version also considered as the most basic variant of the VRP. This variant seeks to build a set of routes to meet the customer demand without exceeding the capacity of the vehicles. This variant is an extension of the TSP so it's also classified as an NP-hard. However, it is possible to optimally solve small and medium sized instances (Toth and Vigo, 2002).

- CVRP lies on finding a set of vehicle routes that minimize the total cost, based on the following conditions (Lysgaard et al., 2004):
- Each customer is visited exactly once;
- The vehicles have a capacity restriction, which means that no vehicle can provide a set of customers whose demand exceeds its capacity;
- Each vehicle used leaves from and returns to the depot.

An example of applicability is given by Achuthan et al. (2003), who sought to solve a CVRP in which the objective function is to minimize the total distance travelled by the vehicles. To solve this problem, the authors proposed an exact solution algorithm in which new cut techniques based on the branch and cut algorithm were used.

Vehicle Routing Problem with Multiple Trips

The Vehicle Routing Problem with Multiple Trips is a variation of the classic VRP, which aims to minimize the costs related to transport using a homogeneous and limited fleet of vehicles from one depot to serve a set of customers. The difference between the classic VRP and this variant is that VRPMT allows multiple trips, which means that a vehicle can do more than one route during a planning period (Brandao and Mercer, 1997). VRPMT consists of build a set of vehicle routes that minimize the total cost, based on some conditions (Cattaruzza et al., 2014):

- each vehicle used leaves from and returns to the depot;
- the total demand for each route cannot exceed the vehicle's capacity;
- each vehicle can serve more than one route during the planning period.

Petch and Salhi (2003) developed a multi-phase constructive heuristic that creates an initial solution for the VRP using the classic Clarke and Wright heuristic. In their work, it was created a process to diversify the solution generated by this method, using a heuristic known as Tour Partition. To improve the solution, refinement heuristics like Exchange, 2-opt and 3-opt and heuristics based on route-grouping with the goal to minimize the penalties using the Bin Packing algorithm.

Pick-up and Delivery Vehicle Routing Problem

The Pick-up and Delivery VRP it was first presented by Min (1989) that developed this variant because of the need to transport books between bookstores with a depot, two vehicles and twenty-two customers. The PDVRP is a generalization of the classic VRP. However, in this specific variant, the customer requires a pick-up and delivery service at the same time. The deliveries are done from a single depot at the beginning of the route and the pick-up are taken to the same depot at the end of the service in a way that all requests are accomplished. The goal is to minimize the distance or cost of the routes.

Dell'Amico et al. (2006) propose an exact approach to the solution for the variant PDVRP. In this case, the consumer must be visited only once, and the demand for pick-up and delivery must be met simultaneously. In this study, the authors present a mathematical model that considers the minimization of the total distance travelled by the fleet of vehicles, using the branch and bound algorithm to face the challenge.

Multi-Depot Vehicle Routing Problem

The Multi-Depot Vehicle Routing Problem is a NP-hard problem, and it is a variant where there is a fleet of vehicles that serve several depots and it was studied by Renaud et al. (1996), who also proposed a new heuristic algorithm centered on a tabu search. MDVRP lies on building a set of vehicle routes, where the goal is to minimize the total cost of routing based on the following conditions:

- every route starts and ends at the same depot;
- each client is visited precisely once by a vehicle;
- the total demand for each route is less or equal than the vehicle's capacity;
- the entire time of each route does not exceed a limit (calculated in advance).

The MDVRP itself is subject to several variants. An example is presented by Crevier et al. (2007) that study a variant of the MDVRP, where the main difference focuses on the possibility that the depots can act as intermediate replenishment facilities along the vehicle route. The authors propose a heuristic as a solution method, which combines the principle of adaptive memory, tabu search and even integer programming.

Mirabi et al. (2010) proposed and hybrid approach for the MDVRP. The only goal considered by the authors in the formulation of the problem was the minimization of the distance travelled by the vehicles. To solve the problem, the authors combine a constructive and improvement heuristics.

Open Vehicle Routing Problem

Owning to transportation costs, several companies decide to hire outsourced transport to carry out the delivery of their products. Therefore, there is no interest by the company to care if the vehicles will or not return to the starting point. Li et al. (2007) described situations like this in their work, as the definition of the Open vehicle routing Problem, which is a variant of the classic VRP. According to MirHassani and Abolghasemi (2011) the OVRP is a combinatorial optimization problem that consists of finding the minimum number of vehicles and their respective routes, so that all customer demands are satisfied. In this variant, each customer is visited exactly once by the vehicle and capacity of each vehicle cannot be exceed. It is not required that vehicles should return to the depot. Therefore, the vehicle routes are not closed paths but open ones.

Sariklis and Powell (2000) proposed a heuristics method to solve this variant of the VRP that consists of applying the technique of the minimum spanning tree with penalties.

Periodic Vehicle Routing Problem

In the PVRP, besides the constraints presents in classical VRP, customers must be served during a certain period and not just only once. This period, usually defined in days is called time horizon. Each customer has a set of visit schedules or a combination of visits, which specify which days on the horizon he can be served. The goal is to plan the route for all the vehicles within this timeline, so that the total cost ate the end is the minimum possible and each client's schedule is satisfied (Angelelli & Speranza, 2002).

Vidal et al. (2012) propose a hybrid genetic algorithm for three classes of the VRP, including the PVRP. The tests run on the instances showed that the hybrid genetic search with adaptive diversity control metaheuristic is extremely robust and efficient.

Vehicle Routing Problem with Time Windows

The VRP with Time Windows refers to the variation of the VRP model that has restrictions on service time windows to customers. Approached with the restriction of deadlines or delivery times that many customers impose, the model includes, in addition to cost optimization, distance and travel times and the cost incurred in delaying or advancing the arrival of the vehicle at customers, that have restricted delivery times (Desrochers et al., 1992). In this sense, the VRPTM is a generalization of the VRP. These time windows require that each customer is served within a predetermined period. Being an extension of the VRP, the VRPTM algo belong to the class of NP-hard problems.

VRPTM not only has the goal to minimize the number of vehicles needed, but also the total travel time or total travel distance, based on the following terms (Bräysy and Gendreau, 2005):

- the construction of the routes must be carried out taking into account that each customer is visited precisely by one vehicle with a previously stipulated time window;
- each route starts and end at the depot;
- the total demand of all customers on the route must be less or equal than the vehicle's capacity.

Yu et al. (2011) introduce the objective of minimizing the total distance travelled by the vehicles in the VRPTM variant. In this approach to the problem, in addition to the constraints of time windows, it is also taken in account the maximum time constraints of the route. This last constraint ensures that vehicles do not exceed a pre-established maximum travel time. Therefore, to solve the problem, the authors suggested a hybrid approach that lies on an ant colony algorithm and tabu search.

Vehicle Routing Problem with Backhauls

Another variant of the classic VRP is the VRPB. It´s a variant that contemplates both delivery and pick-up points, focusing on the efficiency of the vehicle temporal utilization. The model considers two sets: one that gathers all the customers that propose a need for products, in other words, locations thar receive load from a depot, and another one that gathers all suppliers that offers availability of raw material to be collected. These sets are called linehaul and backhaul respectively (Pradenas et al., 2013). The constraints of the VRPN are the following (Goetschalckx & Jacobs-Blecha, 1989):

- each vehicle serves exactly one route;
- each route starts and ends at the depot;
- the load, both linehaul and backhaul cannot exceed the capacity of the vehicle;
- each route satisfies all linehaul customers first and it starts serving the backhaul clients just after the vehicle is empty;
- the total distance travelled by the vehicles is minimized.

Heterogeneous Fleet Vehicle Routing Problem

The HFVRP makes it possible to model the cases of applications where the use of different vehicles does not imply the same operating costs nor the same constraints ((Liu and Shen, 1999). Each type of vehicle has its own characteristics. In an urban environment, small-smized vehicles are preferred both because it is simpler to park and because the quantities to be delivered are generally more modest and fragmented. This can go through the lanes of the road network that the vehicle can take as well as the speed at which it can travel these lanes. A bicycle can borrow lanes closed to other vehicles. A truck has a very large capacity, a light vehicle has a higher average speed than other types of vehicles, for example.

Time Dependent Vehicle Routing Problem

The tours are planned before they are carried out. Once the rout has started, often only the visiting order of the nodes can change (in particular, the requests loaded are specific to the end customers of the route). It is then necessary to anticipate the congestion of the upstream road network, otherwise, the route could be, in whole or in part, unachievable. From the pois of view of users and customers, taking traffic into account is “simple”. Indeed, they have daily access to route calculations that consider the traffic variation in a forecast way and in real time. But this calculation is performed only between a starting point and a destination with a departure schedule or arrival fixed. The route calculation required many variables already fixed. In the case of the optimization of routes, the order of traversal of the nodes and the schedules on which the arcs of the graph will be traversed is not known in advance and may change at each step of the solution. This issue of route dependence on time adds complexity to an already very complex problem. However, research focuses on this theme to provide effective resolution methods and are grouped under the name TDVRP (Taş et al., 2014).

2.2. Solution Methods

In this section, the main resolution methods that have been applied in the VRP as well in its variants, are described. It is possible to find in the literature a wide range of methods and algorithms. Among the various methods developed that can be applied to the VRP, it is possible to divide it into two main categories (Goel and Maini, 2017):

- Exact methods: guarantee obtaining the optimal solution. However, these methods have great difficulties in solving problems of larger dimensions;
- Approximate methods: classical heuristics and metaheuristics. These algorithms traverse only a fraction of the search space, achieving high quality results but nor optimal or, when they find the overall optimum, there are no ways to prove it

2.2.1. Exact Algorithms

Exact algorithms are described as being exhaustive search algorithms that traverse the entire search space ensuring to find the optimal solution to the problem. However, these heuristics may not be right choice when it comes to problems of great dimension, taking too long to fins the optimal solution, which in most cases it becomes unfeasible. The following paragraphs briefly introduce the different methods that can be applied to the resolution of routing problems.

- Linear Programming (LP): is a modelling framework for continuous optimization problems using linear expressions (Vanderbei, 2020) and the most widely used method is the simplex algorithm. Combinatorial optimization problems, of which the VRP is part, consider only one finite set of objects. The search space is ten discrete – Integer Linear Programming (ILP). Since the resolution of LP is realized in a continuous space, it is then possible to apply the method of cutting plane, proposed by Gomory (2010), to introduce additional constraints to obtain integer solutions.

- Branch and bound: proposed by Land and Doig (2010) is based on the enumeration of the solutions of the search space. This space is broken down into sub-spaces (branching) and each sub-space is decomposed in another subspaces. Search space can then be represented by a decision tree. Each node of the tree represents a sub-problem. For each node, it is possible to evaluate (bounding) lower bounds and upper bounds of the corresponding sub-problem. Although it's difficult to directly apply a branch and bound to the VRP, it serves as the basis for the following approaches like branch and price, and branch and cut.
- Branch and Cut: Padberg and Rinaldi (1991) proposed combining the methods of secant and branch and bound planes (Mitchell, 2002). The branch and bound is used for dividing the search space. The method of secant planes allows when a solution has a non-integer variable to add two constraints. The first constraint requires that the variable be less than or equal to the usual integer part. The second constraint forces the variable to be greater than the integer part by excess. This id to force the search to obtain solutions in whole numbers.
- Column Generation: Large linear programs have a lot of variables. However, the optimal solution very often includes only a small number of variables, the others being zero. Dantzig and Wolfe (1960) then proposed a method which avoids the representation of all the columns. The problem is represented through a master problem which is the original problem with only a subset variable considered, and a sub-problem that generates new variables, we then speak of generation of columns, which have the potential to improve the objective function of the master problem. As soon as new variables are integrated into the master problem, a new ILP is resolved. The resolution is stopped as soon as no more variable can improve the objective function and the solution obtained is then optimal.
- Branch and Price: (Barnhart et al., 1998) combines the branch and bound column generation applied to each node of the tree of decision. A master problem is defined as a relaxed linear problem. At each node columns are added to the master problem by solving a subproblem.

2.2.2. Classical Heuristics

The developed heuristics can be categorized according to their method of operation. The classification presented next is based on constructive methods, two-phase methods, and improvement methods.

Constructive Methods

The starting point is an empty solution to which viable solutions are added, which means that it is built gradually, taking always into account the evolution of costs.

- Nearest Neighbour: The node closest to the depot is assigned to a route. Then the node closest to the previous inserted point is added to the route, we proceed step by step until the tour is full. Until the node to be inserted makes the total vehicle requests greater to its capacity. As soon as a round is full, a new round is open, and we look for the unaffected node closest to the repository. The procedure is applied until there are no more nodes to insert or until all the tours are full. The method is simple and quick, however the quality of the solution obtained is often of poor quality

- Clarke and Wright: proposed a method where each node constitutes a tour. We calculate for each pair of routes the reduction in cost (called savings or gain) that the merger of their routes entails by replacing one of the arcs leading to the depot by an arc joining the two tours. The tours are merged until there are no more wins to be had.

Two-Phase Methods

In the two-phase heuristic, the problem is decomposed into two distinct components: the grouping of points to be connected into feasible routes and the route construction itself thus the order in which the group of points must be traversed.

Improvement Heuristics

Improvement heuristics, as the name propose, are capable to improve a given solution carrying out a sequence of exchanges between points in a route or between routes. These heuristics tend to be applied to other methods such as Clark & Wright, with the goal to be able to compete with modern metaheuristic algorithms, achieving solutions with a high level of quality.

2.2.3. Metaheuristics

Metaheuristics differ from classical heuristics because they allow the objective function to be temporally deteriorated, with the propose to prevent that the solutions found for the problem being stuck in local optimums. This allows to find different local optimums that may be better than the other found before an, in theory, a closer approximation to the global optimum.

Metaheuristics and their algorithms can be classified based on their characteristics, comparing population-based metaheuristics with metaheuristics based on trajectory methods. The first ones perform search processes that describe the evolution of a set of points in the search space. In other words, the population-based metaheuristics make use of several solutions simultaneously, with a focus on the optimization. On the other hand, metaheuristics based on trajectory methods describe a trajectory in the search space during the process and include local search metaheuristics such as: Tabu Search, Iterated Local Search and Variable Neighbourhood Search.

In the last decades, examples of metaheuristics have been applied due to their success in complex problems. The paper presented by Osman and Wassan (2002) consists of the application of two route-construction heuristics to solve the VRPB variant based on the saving-insertion and saving-assignment procedures. A reactive tabu search metaheuristic is used to improve the initial solutions. This procedure consists of switching between different neighbourhood structures for the intensification and diversification phases of the search. It was concluded that the metaheuristic proposed by the authors is competitive with the best existing approached in the literature.

Zachariadis and Kiranoudis (2010) address the OVRP variant (described previously in this dissertation), where they propose a metaheuristic to solve the problem based on the local search algorithm with an innovative position that besides exploring the wide neighbourhoods, the local search movements are statically encoded into static moves descriptor. To diversify the search, the authors also use a tabu scheme and a penalty strategy. As a result, it was possible to minimize the number of routes as well as minimize the routing cost, improving several solutions already known.

2.3. Synthesis of the empirical studies including map features to routing planning

The scientific articles consulted refer to the period from 2018 to the present, since it is a topic with a high recent complexity. Despite the progress of research in the apprehension and understanding of urban routing, nothing was found in the literature that refers to the incorporation of the characteristics of the road network in the formulation of the model, such as intersections, traffic lights, roundabouts, car parks, etc. as elements in determining routes in an urban environment.

In the context of city logistics, Zhu and Hu (2019) adapted the VRP model to satisfy the needs of green transportation in poor traffic conditions. Therefore, the authors considered fuel consumption as well as congestion, building a Time-Dependent Green Vehicle Routing Problem (TDGVRP) model. As a solution method, a hybrid algorithm was implemented based on the response surface method that mixes the genetic algorithm and particle swarm optimization. It was possible to conclude that the introduction of fuel consumption in the objective function can significantly decrease it, as well as the duration of congestion influence fuel consumption and total travel time.

Fleischmann et al. (2004) develop a VRP model with an innovative aspect that adapts to the real needs that arise with the evolution of technology. They consider a dynamic VRP system based on real-time traffic information. The problem is based on delivery to the customer that requires transportation from the pick-up location to the delivery destination, with time windows. The authors use a database of an urban traffic management center.

Tang and Yan (2007) have as main objective to develop a model that helps organizations to overcome the problem associated with the use of a fleet of vehicles with another type of fuel, which is the limitation of the vehicle's reach, from a perspective of more sustainable vehicles. Thus, the authors formulated a model based on the Green Vehicle Routing Problem (GVRP) as a mixed integer linear program. As a solution method, two construction heuristics were used: the modified Clarke and Wright savings heuristic and the density-based clustering algorithm.

To minimize variable costs, Kwon et al. (2013) consider the vehicle routing problem with heterogeneous fixed fleet and carbon emissions, through a mixed integer programming model. As a solution method, the authors resort to the tabu search algorithm with three neighbourhood generation methods.

Infrastructural information such as the location of traffic lights and their red/green periods, the location of roundabouts or street crossings and the common speed patterns while traversing those elements can be used to provide better estimations of the travel time to reach a particular destination using a routable map.

The metropolitan environment is characterized by both short road segments and long arterial streets, interconnected by different types of intersections (e.g., traffic lights and right-before-left priority). Additionally, the traffic comprises different types of vehicles, pedestrians, and unexpected obstacles.

Traditional views of the routing problem have tended to emphasise the vehicle and location constraints, without paying much attention to path constraints. Keenan (2008) recommends a broader definition of vehicle routing problems to accommodate problems where the path taken is an important component of the problem. The main constraints related to the path are: Underlying network—directed or undirected; Time to travel a given network segment and Vehicle type limitations on network segments. In urban applications additional concerns include, one-way streets, no right or left turns, and vehicle size restrictions. Claes et al.(2011) have described a routing strategy for anticipatory vehicle routing by using forecast information. Thus, traffic congestion can be not only be avoided by drivers, but also can be prevented. This model has revealed a positive impact on the total travel time, decreasing by 35% when compared with models that do not use data or real time data.

Hydén and Várhelyi (2000) have proven that although roundabouts are effective for speed reduction from the drivers, time consumption was reduced heavily when compared to a signalized intersection instead. Also, emissions at roundabouts, when compared to signalized intersection, show a reduction between 20% and 29%. Meneguzzer *et al*. (2017). carried out a study based on the comparison of CO2, NOX and CO emissions at a road intersection with a roundabout and with traffic lights, where they concluded that a road intersection with a roundabout is more energy efficient than a road intersection with traffic lights, because signalized intersections represent high CO2 emissions with longer travel time. Várhelyi (2002) also compared CO and NOX emissions, as well as fuel consumption in two scenarios: roundabout vs signalized junctions and roundabout vs yield regulated junctions. In the scenario where a signalized intersection was replaced by a roundabout, the speed at the intersection not only decreased but also became more uniform. The delay per car passing through the intersection decreased by an average of 11s. The number of cars stopped at the intersection decreased from 63% to 26% of the total. CO and NOX emissions and fuel consumption decreased significantly.

The probabilistic investigation conducted by Poli Jr and Monteiro (2005) reveals that traffic lights might improve traffic flow in both directions, although only above certain vehicle density. The author completes by showing that the presence of traffic lights, in below critical point situations, does not help increase traffic flow. Khanjary et al. (2011) proposed a shortest path algorithm where the traffic lights network is synchronized, using waiting times for green lights and the required time to pass. Time and space complexities are reduced, when compared to few similar algorithms. Also, Chen and Yang (2000) presents a novel time constraint, the traffic lights, simulating operations in intersections, when optimizing a time-constrained shortest path problem. Hu et al. (2016) aims to search for the optimal path by using an improved A*algorithm that reflects on an urban area with traffic lights. A simulation of the algorithm shows 5,8% savings, represented in time, when compared to the conventional A*Algorithm, finding the shortest path in urban areas. For the optimal route considering signalized intersections, their waiting time and an eco-driving model, Hu et al. (2018) achieves a model with good results, combined with an optimal vehicle energy consumption, the algorithm is effective, proposing better routes and a lesser amount of fuel consumption. The impact of virtual traffic light systems on carbon dioxide emissions is discussed by Ferreira and d'Orey (2011), concluded that these systems are beneficial in terms of carbon dioxide emissions for all traffic densities studied. Also, this traffic control system shows increasing average velocity. Regarding eco-routing, Sun and Liu (2015) algo gave their contribution developing an eco-routing algorithm for signalized traffic networks, incorporating vehicle emissions into a Markov decision process.

Concerning to speed bumps, Pau and Angius (2001) say that although speed bumps seem to affect drivers' behaviour to some extent, their effectiveness as speed reducing devices is quite far from optimal. On the same guiding thread, Pérez-Sansalvador et al. (2020) capture the effect of speed bumps on traffic flow and dioxide carbon plus other pollutants emissions, founding that higher number of speed bumps increase pollutants aggregation on low density traffic levels. When analyzing traffic speed, the author realized that for high densities of vehicles, the number of speed bumps have no significant impact on speed. Baltrėnas et al. (2017) analyzed the different types of speed bumps in various residential areas of Lithuania, with the aim of evaluating air pollution with particulate matter. The authors concluded that due to the existence of speed bumps, vehicles brake suddenly, which causes greater tire wear and, consequently, greater particle emissions. Lin et al. (2016) focus on vehicle scheduling and its cost, which is not only influenced by the customer's satisfaction and the effective use of vehicle's capacity. Thus, its related to traffic flow, road capacity, traffic lights and road rules, such as speedbumps. The new vehicle routing algorithm, when applied, showed results, increasing effectiveness.

Regarding crosswalks, Knoop and Daganzo (2018) showed that a larger number of crossings reduces the pedestrian load per crossing and facilitates both the pedestrian flow and the car flow.

Cohen and Dalyot (2021) developed an algorithm, using OSM data, for route planning especially designed for blind pedestrians, to obtain accessible and safe routes. The authors included in their algorithm road network features such as way type, traffic signals vibration and sound and tactile paving - attributes that are essential for a safe route. Regarding route planning for disabled people, Džafić et al. (2020) developed a model to improve route planning for people with wheelchairs, with the aim of obtaining more accessible and more efficient routes considering battery consumption for the case of electric wheelchairs. Some constraints of the algorithm are road surface, indoor and outdoor accessibility, road barriers and access to public transport.

Wang and Niu (2018) study on indoor and outdoor pedestrian route planning. The authors use OSM to extract the necessary attributes such as connections within buildings, the existence of lifts and stairs, and access to the outside. The algorithm can return the shortest path.

Wang et al. (2020) have dedicated themselves to developing a model to planarize quiet routes in traffic noise polluted environments for pedestrians. The thrust of this project is people's quality of life and health. Factors such as noise level and traffic volume were taken into consideration. Still on environmental issues, Nallur et al. (2015) present a route planning model focusing on obtaining four different types of routes, least noisy route, least air pollution, fastest and shortest route.

Sarraf and McGuire (2018) developed a route planning model in which the focus is to determine a route with minimal accident risk. To achieve this goal, the algorithm considers the following factors: road surface, lightening and visibility, speed limit, traffic volume and the possibility of crossing animals. The OSM is used as a data source. Similar to this project, Liao et al. (2022) developed route planning based on a crash risk prediction model, incorporating factors related to driver experience, vehicle, road, and environment.

Solé et al. (2022) are also concerned with driver safety, so they investigated a route planning model with the development of an application that allows the incorporation of different factors and criteria for the generation of routes. These factors are essentially related to the road, such as path length, road lightening, level of criminality and accident ratio.

Regarding route planning models that integrate in a qualitative way drivers' preferences, we highlight the studies of Wang et al. (2021) and Teng et al. (2019), which develop a customizable route planning system based on the preferences of each driver. The models can incorporate different factors, including the existence of traffic signals, sidewalks, traffic jam, travel time, road type, fuel consumption and turns and intersections. Jensen et al. (2020) also paid attention to drivers' preferences, dividing these preferences by the type of vehicle, electric or conventional.

Schambers et al. (2018) focused on the development of a route planning model specifically oriented for electric vehicles using the Bellman-Ford algorithm as it can include different criteria such as elevation, road length and friction and regenerative braking to obtain a more energy efficient route.

Hakeem et al. (2019) develop a deferential feature to the Multi-Destination Vehicle Routing Problem, which consists of including different factors such as traffic volume and information regarding car parks. This way, the model can present the shortest and fastest path with a higher accuracy.

Regarding route planning for cyclists, Hrncir et al. (2014) develops a route planning model considering cyclists' preferences, incorporating in their model a diverse set of factors such as road surface, obstacles, elevation, crossings, paths dedicated only to cyclists, paths shared with vehicles, among others. The authors commit themselves to obtain a route with focus on comfort, quietness, road levelling and travel time.

Krisp and Keler (2015) conduct a study using OSM data to develop a routing algorithm dedicated to citizens' preferences with respect to avoiding complicated crossings.

Felicio et al., (2022) developed a routing algorithm dedicated to preferences pedestrians have on their commute in urban environment through georeferenced data. The dimensions studied concern safety, comfort, accessibility, air quality, time, and distance.

Table 1 shows the summary presented so far, with a focus on the road network features, outcome variables and results.

Table 1:Synthesis of recent studies that analyze the impact of road network features on planning route, emissions, and traffic flow.

Authors (year)	Country	Issue	Road Network Features	Outcome Variables	Results
Keenan, P. (2008).	-	Route planning	Vehicle type limitations on network segments	Travel Time	-
Claes, Holvoet & Weyns (2011)	Belgium	Route planning	Traffic lights	Travel Time	-
			Speed limits		
			Turning rules		
			Traffic intensity		
			Intersections		
Hydén & Várhelyi (2000)	Sweden	Traffic flow	Roundabouts	Safety	+
				Travel Time	-
				Environmental cost	+
Meneguzzer, Gastaldi, Rossi, Gecchele & Prati (2017)	Italy	Emissions	Roudabouts	Environmental cost	-
			Signalized intersections	Environmental cost	+
Várhelyi (2002)	Sweden	Emissions	Signalized intersections	Environmental cost	+
			Roundabouts	Environmental cost	-
Poli Jr & Monteiro (2005)	Brazil	Traffic flow	Traffic lights	Traffic flow	+
Khanjary, Faez, Meybodi & Sabaei (2011)	Iran	Route planning	Traffic lights	Travel Time	+
Chen, Y., Yang, H., 2000	-	Route planning	Trafiic lights	Travel Time	+
Hu, L., Zhong, Y., Hao, W., Moghimi, B., Huang, J., Zhang, X., & Du, R (2018)	China	Route planning	Traffic lights	Fuel Consumption	-
				Travel Time	-
L. Hu, J. Yang, and J. Huang (2016)	China	Route planning	Traffic lights	Travel Time	-
Ferreira & d'Orey (2011)	Portugal	Emissions	Virtual traffic lights	Environmental cost	-
J. Sun and H. X. Liu (2015)	USA	Route planning	Traffic lights	Environmental cost	+

Pau, M., & Angius, S. (2001)	Italy	Traffic flow	Speed bumps	Vehicle Speed	/
Pérez-Sansalvador, Lakouari, Garcia-Diaz & Hernández (2020)	-	Emissions	Speed bumps	Environmental cost	+
				Traffic flow	/
Baltrėnas, H. P., Januševičius, T., & Chlebnikovas, A. (2017).	Lithuania	Emissions	Speed bumps	Particulate matter air concentration	+
Lin, X., Geng, F., Jin, Z., & Xu, Q. (2016, July).	China	Route planning	Intersections	Travel Time	-
			Traffic lights		
Knoop, V. L., & Daganzo, C. F. (2018)	USA	Traffic flow	Crosswalks	Traffic flow	-
Cohen, A., & Dalyot, S. (2021)	Israel	Pedestrian route planning	Way type	Safety for blind people	+
			Traffic signals vibration/sound		
			Tactile paving		
Wang, Z., & Niu, L. (2018)	Germany	Pedestrian indoor and outdoor route planning	Connections inside buildings	Distance	-
			Elevators/Escalators/Stairs		
Hrncir, J., Song, Q., Zilecky, P., Nemet, M., & Jakob, M. (2014)	Czech Republic	Bicycle route planning	Surface, Obstacles, Dismount	Travel time	+
			Surface, Obstacles, Dismount	Comfort	
			Bicycles roads and motor roads, Crossings	Quietness	
				Flatness	
Wang, Z., Novack, T., Yan, Y., & Zipf, A. (2020)	Germany	Pedestrian route planning	Noise level of roads	Route noise pollution	-
			Traffic volume		
Džafić, D., Candiotti, J. L., & Cooper, R. A. (2020)	Germany	Electric powered wheelchairs route planning	Road surface	Distance	+
			Indoor and outdoor accessibility	Comfort	+
			Road barriers	Battery Consume	-
Nallur, V., Elgammal, A., & Clarke, S. (2015)	Ireland	Route planning using participatory sensing	Noise	Least noisy route	-
			Air pollution	Least air pollution	-
			Time	Fastest route	-
			Distance	Shortest route	-

Sarraf, R., & McGuire, M. P. (2018)	USA	Safe route planning	Road surface	Risk of accidents	-
			Visibilaty		
			Road lightening		
			Speed limit		
			Traffic volume		
			Crossing animals		
Schambers, A., Eavis-O'Quinn, M., Roberge, V., & Tarbouchi, M. (2018)	USA and Canada	Electric vehicle route planning	Elevation changes	Energy Efficiency	+
			Friction/regenerative braking		
			Road length	Shortest distance	-
Wang, R., Zhou, M., Gao, K., Alabdulwahab, A., & Rawa, M. J. (2021)	Macau	Route planning	Road curvature	Network size	Qualitative analysis
			Congestion		
			Traffic flow		
			Number of pedestrians and bicycles		
			Fuel consumption		
			Toll fee		
			Traffic lights		
			Turns and intersections		
Liao, X., Zhou, T., Wang, X., Dai, R., Chen, X., & Zhu, X. (2022)	China	Route planning	Light conditions	Risk of accidents	-
			Atmospheric conditions		
			Roadway surface		
			Intersections	Travel time	/
			Traffic control device		
Hakeem, A., Gehani, N., Ding, X., Curtmola, R., & Borcea, C. (2019)	USA	Multi-destination vehicle route planning	Real-time traffic	Travel time	/
			Parking conditions		
Jensen, A. F., Rasmussen, T. K., & Prato, C. G. (2020)	Denmark	Route planning with driver preferences	Road length	Shortest route	Qualitative analysis
			Travel time		
			Left turns/Right turns	Fastest route	
			Electrical vehicle	Drivers’ preferences	

			Conventional vehicle		
Teng, L., Tian, J., & Izumi, T. (2019)	Japan	Route planning with driver preferences	Road width	Preferenced Route	+
			Travel time		
			Traffic jam	Travel time	+
			Road length		
			Traffic signals	Travel distance	+
			Sidewalk		
Solé, L., Sammarco, M., Detyniecki, M., & Campista, M. E. M. (2022)	England		Road lightening	Safety	+
			Accidents ratio		
			Criminality level		
			Path length	Total travel time	/
			Time		
Felício, S., Hora, J., Ferreira, M. C., Abrantes, D., Costa, P. D., & Galvão, T. (2022)	Portugal	Pedestrian route planning	Green areas	Safety and security	Qualitative analysis
			Crwon density	Confort	
			Traffic volume	Accessibility	
			Noise level	Air quality	
			Crosswalks	Time	
			Traffic lights	Distance	
Krisp, J.M, & Keler, A. (2015)		Route planning	Turn restrictions	Number of avoided complicated crossings	-
			Crossings		
			Distance		

Even though in the literature we can find numerous studies using road network features like traffic lights, speed bumps, intersections, or roundabouts, that focus on the effect on traffic flow and emissions (Baltrėnas *et al*., 2017; Meneguzzer *et al*., 2017; Várhelyi, 2002; Ferreira and d'Orey, 2011; Merkisz *et al*., 2014; Pérez-Sansalvador *et al*., 2020), studies that analyse these features as driver preferences are rare. This scarce empirical literature has focused on very specific features, namely traffic lights (Lin *et al*., 2016; Hu et al., 2018; Hu et al., 2016; Felício *et al*., 2022; Teng *et al*., 2019), intersections (Lin *et al*., 2016; Liao et al., 2022) and crossings (Krisp & Keler, 2015; Hrncir *et al*., 2014; Knoop & Daganzo, 2018).

After an exhaustive search of the state of the art, it is possible to verify that most studies that develop routing models focus on minimizing travel time or distance, consumption and exposure to noise and environmental pollution. Given the scarcity of studies of route planning models in which the focus is the preferences of each driver, i.e., the maximization of satisfaction and fulfilment of the needs of each driver, it would be scientific pertinent to enable the integration of each driver's preferences regarding road network features like traffic lights, roundabouts, crosswalks, intersections, and vehicle type and speed limitations on network segments, since these features, in an urban environment, are likely to influence the choice of route to take.

This dissertation contributes to fill this gap by developing a route planning model that incorporates the aforementioned features with a weight that is up to the driver to decide, respecting the needs of each driver. It is important to highlight that, unlike most of the studies presented in the previous table, the model to be developed in this project does not aim to minimize the total travel time, but to offer drivers different route options.

3. METHODS AND APPLICATION

This chapter describes the methodology adopted for the development of the project.

The complexity of this research project required the author to stipulate a highly organized work methodology, with numerous iterations and high standards of quality and control. In fact, the activities carried out are divided into two phases:

1) Exploratory phase. In this first phase, it is performed a survey of the needs of drivers in urban environment, particularly regarding the classification of an easy driving mode. In addition, an analysis of the state of the art in route planning models applied to the urban environment and that consider road network features of relevance to this project, as well as the preferences and needs of different drivers is performed.
2) Data collection and construction of the routing model. In this phase, the data needed to answer the research questions are collected and the routing model is built. To carry out the various activities that make up the architecture development of this phase, the technologies to be used are also chosen, based on rigorous selection criteria. The architecture development consists of a set of activities from the creation and structure of the database, the implementation of functions for calculating the different paths options between two points, and finally the visualization of routes.

Subsequently, in section 4 the results obtained are analyzed and discussed.

3.1. Scientific and technological uncertainties the project sought to address

About the scientific and technological uncertainties that the project intends to address, two main ones are pointed out, namely those related to data collection and the uncertainty about the effectiveness of the routing model with the integration of different weights (drivers' preferences).

1) Difficulties in collecting the necessary data sources to feed the model:

In a first phase of the project, some difficulty in accessing the necessary data sources to run the prototype may eventually prevail. However, having already clarified the objectives of the system (see section 1.2), the necessary steps will be taken to produce a viable data sample. This production will only occur at a stage when the requirements of the prototype are relatively stable.

2) Lack of effectiveness of routing models with integration of different weights:

Based on a given set of data, routing algorithms may be more or less effective, depending on the value and/or quality of that data. If the techniques and algorithms are not able to extract useful value from the data provided, or the quality of the data is not the desirable one, a mixed-model approach can be used to try to include as much as possible other knowledge that does not come from data, but from people.

3.2. Investigating the needs and preferences of users – defining an easy to drive route

In the last years there were several studies conducted to develop route planning models with several differentiating features. However, after exhaustive research of the literature, as well as of the market, most of the route planning models are focused on the shortest time or shortest distance, claiming that these are the features in a route that drivers value most. However, the profile of each driver may vary in attributes (value of time, time budgets, among others), driving experience, behavioural preferences (willingness to take risk, willingness to switch routes with potential savings), road knowledge, among other factors. This set of characteristics, according to Zhu and Levinson (2015), leads to significant heterogeneity at the moment of route choice decision. In this sense, drivers may choose a route option that does not contemplate the shortest or fastest route, but one that represents other factors such as simplicity, traffic congestion, environmental problem or simply user's satisfaction.

Given the above, the approach adopted in this work for the creation of the "easy to drive" concept contemplates the following factors that together can determine whether a route is easy/difficult to drive: roundabouts, pedestrian crossings, traffic lights, intersections, public transport stops and stop signs.

3.3. Architecture Development

To respond to the technological uncertainty that the present project seeks to resolve, the author defined the following methodology, divided into six stages:

1) Collecting georeferenced data;
2) Data Manager and data importation;
3) Formulating and quantifying the criteria;
4) Building of routing graph;
5) Execution of routing algorithms;
6) Routes visualization.

Each of the above steps is explained in detail in the following subchapters.

3.3.1. Colleting georeferenced data

The choice of platform for data collection was based on the following criteria: a comprehensive documentation and tutorials, free libraries, and a free access to data. After an extensive literature search, the three most used platforms were identified.

- Google Maps Platform is a platform geared towards providing public APIs that can easily be used and edit maps, and as far as existing documentation and tutorials are concerned, it can be said to be a very beneficial platform. However, it is not a free platform and additionally it is not clear that it is possible to use the provided data to create applications entirely with different features, namely that one can include user preferences.

- ArcGIS Developer framework resides on a public API supply, so it is quite similar to the Google Maps platform. Regarding data access, only a small part is free. Regarding the existing documentation, it only describes how to use the tools that are already implemented, being difficult to know how to access the data in order to create functionalities or expand libraries.
- OpenStreetMap framework is a digital map built through crowdsourced Volunteered Geographic Information (VGI) and can be visualised, queried, downloaded, and modified under open licenses. Regarding the existing documentation, OSM provides a wide description and classification of the data, offering also several tutorials on how to download the data and which tools can be used to work directly with OSM data and, on the other hand, which tools can import the data into a database.

In view of the above, the OpenStreetMap framework was the chosen platform since it is the only one that allows the user to work directly with the map data - a relevant aspect for the creation of a routing algorithm based on the preference of several users. Additionally, this platform respects the three requirements identified above: it has a wide range of documents and tutorials, and it is a digital map under open licenses. This last requirement is not verified in the other platforms.

The OSM data can be downloaded directly from the OpenStreetMap website, however it is not considered the best option to extract the data. Nevertheless, the data extraction can be performed using several online tools of which the following stand out: GeoFabrik, BBBike, HOT Export Tool and Planet OSM, tools that allow downloading the files in several formats, which are frequently updated. In this project, the BBBike platform was used for two reasons that, together, none of the other tools has: it allows to delineate the area from which the data is to be extracted and, additionally, all the tag keys and tag values that are available for free are extracted.

The data were extracted in ESRI Shapefile (.shp) format for the city of Lisbon, Portugal. The .shp format is composed by a zipped folder, which contains the following .shp files:

- Buildings: building outlines;
- Landuse: land use and land cover, like forests, parks, natuer reserves, etc;
- Natural: natural features like glaciers, volcanos, mines, beaches, etc;
- Places: location for cities, towns, among ohters;
- Points: health, leisure, public, catering, accommodation, shopping, Money, tourism, places of worship and traffic related;
- Railways;
- Roads: all kinds of roads from motorways to gravel tracks as well as cycleways and footpaths;
- Waterways.

For the present project the relevant files are points and roads, so only these two will be imported into the database.

3.3.2. Data manager and data importation

The Database Management System (DBMS) is a software that presents several functionalities with emphasis on the possibility of creating and defining the database structure, retrieving stored data, introducing and modifying data in the database, as well as processing the data. In effect, DBMS makes the process of working with databases more user friendly.

In the present research project, the author resorted to the comparative study between DBMS conducted by Poljak et al. (2017), which defined three comparative criteria when choosing a software: comparison of basic data, data types and speed performance.

According to the criteria mentioned above and taking into consideration the scientific/technological uncertainty that the project seeks to address, the DBMS chosen is PostgreSQL, since it meets the necessary requirements at the level of the three criteria and, most importantly, it is a software that because of its liberal license it can be used, modified, and distributed completely free for any purpose.

PostGIS is an open source, free to use spatial extension that inserts support for Geographic Information Systems (GIS) objects into PostgreSQL, allowing, for example, localization queries to be executed in SQL. Based on the definition in its manual (Zurbarán et al., 2018), PostGIS has support for spatial indexes, and functions for the analysis and processing of GIS objects.

This extension was chosen to work on top of PostgreSQL, which according to Hsu and Obe (2021) has an easy extensibility and provides some important benefits, such as transactional support, essential index support for spatial objects and a ready-to-use query plan generator. Still according to the author, with the provision of spatial indexing enhancements, operators, functions, and spatial data types to PostgreSQL, PostGIS in conjunction with these features and PostgreSQL, PostGIS becomes a very powerful tool for learning about GIS

This extension is added through the Stack Builder in PgAdmin 4, which is an open graphical tool that was installed with PostgreSQL. It is used for PostgreSQL administration, simplifying the creation, maintenance, and use of database objects through a powerful graphical interface.

Stack Builder is a tool that simplifies the download and installation of extensios for PostgreSQL. When installing a module with it, all software dependencies are resolved automatically. Its use was necessary for the installation of PostGIS, PostGis_topology and pgRouting.

To start the process of importing the data into PostgreSQL, it is first necessary to create a database (in this case, Lisbon_Routing). Subsequently, the PostGIS tool PostGIS Bundle 3 for PostgreSQL x64 14 Shapefile and DBF Loader Exporter is used, which allows importing .shp files into the created database.

For a better understanding,

Figure 1 shows a scheme with the different steps mentioned so far.

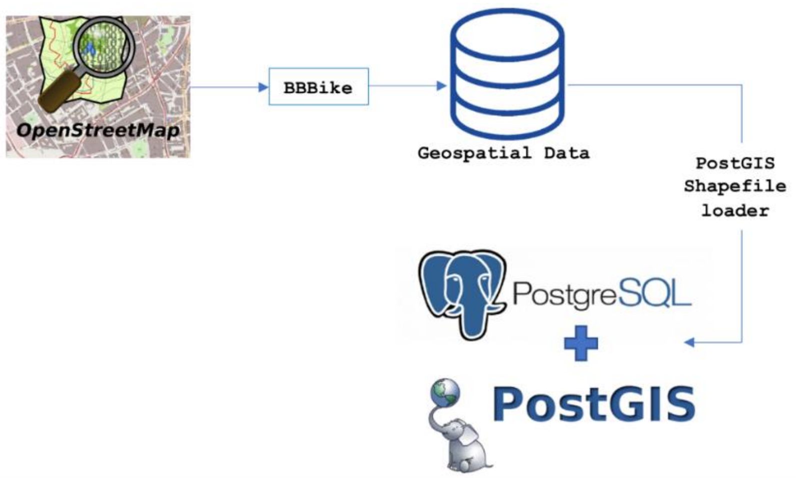


Figure 1:Extraction of data and importation to the DBMS.

3.3.3. Formulating and quantifying the criteria

Routing is the process of selecting the optimal path between two points (source and target) in a network, which consists of a set of edges connected through nodes. To select the optimal path, it is necessary to associate a cost to each edge. In effect, the path considered as optimal represents a combination of edges between source and target that represent a lower cost.

In the present research project, the problem that is sought to be solved is aimed at vehicle routing, so the edges will represent such features oh highways: motorway, trunk, primary ways, secondary ways, tertiary ways, residential, motorway link, trunk link, primary link, secondary link, tertiary link and living streets.

After researching users' needs and preferences (see section 3.2), was found that there is a group of people who, in an urban environment, prefer to opt for a quieter and easier to drive route, even if it means choosing a longer distance route. In fact, the author defined two dimensions: distance and Easy Drive mode. Regarding the first dimension, the criterion that defines it is the length of each edge, and the second dimension is composed of the following five criteria: bus stop, traffic signals, crossing, mini roundabout and stop signals. Criteria such as roundabouts, shared lane with cyclists and traffic calming have not been included as they are features that are not available in the free version. Nevertheless, they are features that have an impact on the easy drive mode.

It is defined a scale of three ratings for the criteria that define the easy drive mode:

- 0 - Neutral;
- 1 - Less preferred;
- 2 - Better to avoid.

Table 2 exposes the tag_key and tag_value (in the data is called type) essential for the approximation of the solution to the real-world scenario.

Table 2:Road network features.

Tag Key	Tag Value	Description
Highway	Motorway	A restricted access major divided highway, normally with 2 or more running lanes plus emergency hard shoulder. Equivalent to the Freeway, Autobahn, etc..
	Trunk	The most important roads in a country's system that aren't motorways. (Need not necessarily be a divided highway.
	Primary	The next most important roads in a country's system. (Often link larger towns.)
	Secondary	The next most important roads in a country's system. (Often link towns.)
	Tertiary	The next most important roads in a country's system. (Often link smaller towns and villages)
	Residential	Roads which serve as an access to housing, without function of connecting settlements. Often lined with housing.

	Motorway_link	The link roads (sliproads/ramps) leading to/from a motorway from/to a motorway or lower class highway. Normally with the same motorway restrictions.
	Trunk_link	The link roads (sliproads/ramps) leading to/from a trunk road from/to a trunk road or lower class highway.
	Primary_link	The link roads (sliproads/ramps) leading to/from a primary road from/to a primary road or lower class highway
	Secondary_link	The link roads (sliproads/ramps) leading to/from a secondary road from/to a secondary road or lower class highway.
	Tertiary_link	The link roads (sliproads/ramps) leading to/from a tertiary road from/to a tertiary road or lower class highway.
	Living_street	For living streets, which are residential streets where pedestrians have legal priority over cars, speeds are kept very low and where children are allowed to play on the street
	Bus_stop	A small bus stop. Optionally one may also use public_transport=stop_position for the position where the vehicle stops and
	Crossing	.k.a. Crosswalk. Pedestrians can cross a street here; e.g., zebra crossing
	Mini_roundabout	Similar to roundabouts, but at the center there is either a painted circle or a fully traversable island. In case of an untraversable center island, junction=roundabout should be used.
	Motorway_junction	Indicates a junction (UK) or exit (US). Ref=* should be set to the exit number or junction identifier. (Some roads – e.g., the A14 – also carry junction numbers, so the tag may be encountered elsewhere despite its name)
	Unclassified	Road
	Stop	A stop sign
	Traffic_signals	Lights that control the traffic
Cycleway	Shared_lane	Cyclists share a lane with motor vehicles, but there are markings indicating that they should share the lane with motorists. Paid file
Junction	Roundabout	Paid file
traffic_calming	Yes	Paid file
	Bump	Paid file
	Hump	Paid file

3.3.4. Building of routing graph

This section presents the methodology followed for building a routing graph.

To build a routing graph it was necessary to add the PgRouting extension. PgRouting is an extension that integrates with PostGIS / PostgreSQL and adds geospatial routing and other network analysis features to the database. It offers many routing algorithms. This extension is used to build a routable database as well as subsequently implement a routing algorithm.

1) Sort and filter unnecessary features and data base structure

This point explains how to filter and eliminate features that are not necessary for the development of the easy drive model (see section 3.3.3 for details).

As previously mentioned, the first point is that this is a routing problem intended for vehicles so, regarding the tag key, the following tag values are eliminated: track, pedestrian, busway, bridleway, cycleway, footway, path and steps. This step can be obtained through the following SQL query.

```
DELETE from roads
 WHERE
  type='track' or
  type='pedestrian' or
  type='busway'
  OR type='bridleway'
  OR type='cycleway'
  OR type='footway'
  OR type='path'
  OR type='steps';
```

Additionally, six tables were created, five of which relate to the criteria defined above and the last table relates to points of interest in the city of Lisbon, so that these points can be used as source and target (origin and destination).

Create table for criterion

```
CREATE TABLE [table name](
  osm_id double precision,
  type character varying(16),
  geo geometry);
```

Create table for points of interest

```
CREATE TABLE [table name](
  osm_id double precision,
  name character varying(48),
  type character varying(16),
  geom geometry,
  points_of_interest_id bigint);
```

Now there’s the need to insert data for criterion tables (retrieved from points) and for points of interest table.

```
INSERT INTO [table name]
  SELECT osm_id, type, geom
   FROM points
    WHERE points.type='[name of the criterion]';

INSERT INTO [table name]
 SELECT osm_id, name, type, geom
  FROM points
   WHERE points.type='museum'
    OR points.type='restaurant'
    OR points.type='castle'
    OR points.type='monument'
    OR points.type='attraction'
    OR points.type='attraction';
```

Since the source and target will be identified by the user as the name of a point of interest, it is necessary to eliminate the points of interest whose name is null.

```
DELETE FROM points_of_interest
 WHERE name IS NULL;
```

Once this step is complete, Figure 2 shows the tables that have been created.

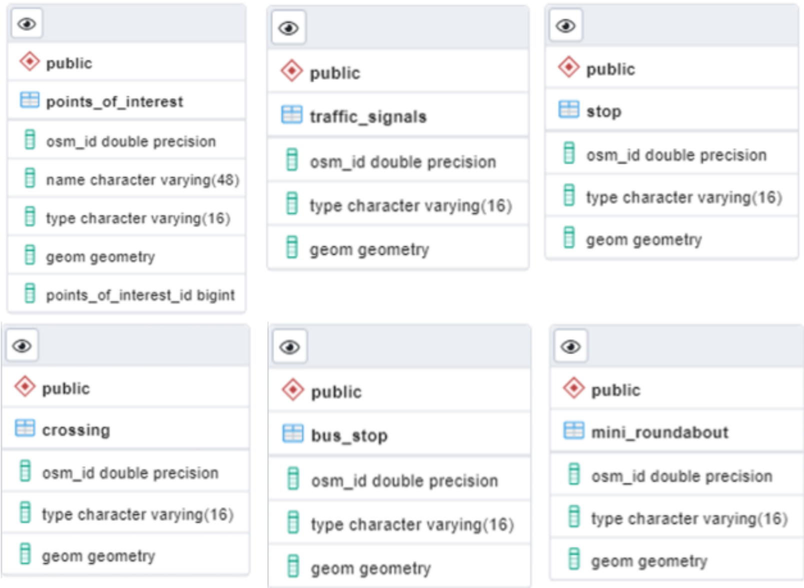


Figure 2:Database structure before routing graph construction.

2) Convert Multilinestring to linestring

PgRouting extends the PostGIS / PostgreSQL geospatial base and is prepared to work with OSM maps, providing geospatial routing and other network analysis functionalities. This extension provides various routing algorithms, such as functions for calculating the shortest path between two points.

After installing the PgRouting extension, we converted the roads geometry type from multilinestring to linestring since the pgRouting extension gets better results with this type of geometry. To convert the type of geometry, the following SQL query is executed.

```
ALTER TABLE roads
 ALTER COLUMN geom TYPE geometry(LineString,4326)
  USING ST_LineMerge(geom);
```

3) Fix topology

After the elimination of non-necessary features, namely regarding roads, it is possible to visualise the map obtained through the QGIS software, that is a geographic information system. Although the map seems visually correct, at this step the data do not have topology, which is required to connect nodes and to obtain real distances.

The topology can be created using the function pgr_createtopology using the PgRouting extension in PgAdmin 4. Once the map has a correct topology it will be possible to execute routing algorithms. In fact, the aim of this step is to create the topology and add columns to the roads table, source and target, to receive the nodes that will be created. Next, the function pgr_createtopology is used to attach a node to each end of every road segment in the roads table.

```
ALTER TABLE roads ADD COLUMN "source" bigint;
ALTER TABLE roads ADD COLUMN "target" bigint;
SELECT pgr_createtopology('roads',0.0001,'geom','gid’);
```

Once the topology has been created, it is necessary to analyse it and finalise the routable graph by performing the following functions:

- pgr_analyzeGraph: analyzes the edges for dead ends, segment breaks, isolated edges, and intersections. Additionally, statistics are created in the new table generated with the function pgr_createtopology executed earlier, such as the number of times a vertex is referenced, number of edges in the roads table that reference that same vertex (column cnt) and fills the column chk, which is an indicator of possible problems regarding the vertex;
- pgr_analyzeOneWay: in the table roads the oneway column is a binary value;
- pgr_nodeNetwork: creation of the roads_noded table (table on which the routing algorithm is executed). This table is created based on all the steps performed so far. In effect, the function breaks the segments into several different segments connected through the vertices that were generated in these same intersections. Thus, these vertices are seen as nodes connecting several edges.

```
SELECT pgr_analyzegraph('roads',0.0001,'geom','gid','source','target','true');

SELECT pgr_analyzeoneway('roads',
 ARRAY['1','0'],
 ARRAY['1','0'],
 ARRAY['1','0'],
 ARRAY['1','0']);

SELECT pgr_nodenetwork('roads',0.0001,'gid','geom');
```

Once the roads_noded table has been generated, the pgr_analyzegraph and pgr_analyzeOneway functions must be re-run on the newly generated table.

 shows the uploaded standard map from OSM (with the plugin QuickMapServices) in QGIS with the layer of data from the table roads_noded, obtained with a connection between QGIS and PostGIS.

Figure 3:Visualization of OSM with routable road network.

3.3.5. Execution of routing algorithms

In the present project, three types of routes are calculated, according to the previously defined criteria that influence the considered optimal path. In effect, the three types of routes are defined as the shortest path (based on distance), the mode easy drive route (based on the existence of traffic lights, pedestrian crossings, mini roundabouts, stop signs and bus stops) and balanced route (a balanced weight system between shortest path and easy drive).

For each segment i a cost is calculated.

The cost for the shortest path is calculated by.Equation 1.

$$SP_i = L_i, \forall\, i \in I \qquad \text{Equation 1}$$

Where,
L_i – length of edge i

For the easy drive mode route, the cost is calculated by Equation 2.

$$ED_{ip}=\sum_{c=1}^{C} \alpha_{cp} N_{ic}, \forall\, i \in I, c \in C \qquad \text{Equation 2}$$

Where,

i – edge i

c – criterion c

α_{cp} – importance given by person p to criterion c

N_{ic} – number of elements of criterion c existing in edge i

For the balanced route, the cost is calculated by Equation 3.

$$B_i = D_{1p} L_i + D_{2p} \sum_{c=1}^{C} \alpha_{cp} N_{ic}\,, \forall\, c \in C \qquad \text{Equation 3}$$

Where,

D_{1p} – importance given by person p to dimension 1

D_{1p} – importance given by person p to dimension 2

α_{cp} – importance given by person p to criterion c

N_{ic} – number of elements of criterion c existing in edge i

L_i – length of edge i

After defining the weighting system for the effort that each person p has when crossing each edge, it becomes necessary to model the database once more.

First, columns are added to the roads_noded table for each defined criterion with the goal to populate these columns with the existing number of elements of each criterion in each segment:

- Trafic Signals (to receive the number of traffic signals that are in each road);

– Bus_Stop (to receive the number of bus stop that are in each road);
– Crossing (to receive the number of crossing that are in each road);
– Mini_Roundabout (to receive the number of mini roundabouts that are in each road);
– Stop.

```
ALTER TABLE roads_noded
 ADD criteria integer;
```

Afterwards, it is necessary to find out the number of elements of each of these criteria existing in each edge. To do so, we use the PostGIS function ST_Intersect.

```
UPDATE roads_noded SET [column name]=(
 SELECT count(*)
  FROM [table name]
   WHERE
    ST_Intersects([table name].geom, roads_noded.geom));
```

Additionally, a column called length is created to represent the length of each segment.

```
ALTER TABLE roads_noded
 ADD length double precision;
```

Finally, to include the cost values for each segment, three columns are created for the three cost types: distance, easy drive and balanced.

Figure 4 shows the processing through the various steps of the data contained in the table roads, which gave rise to the table roads_noded - routable data.

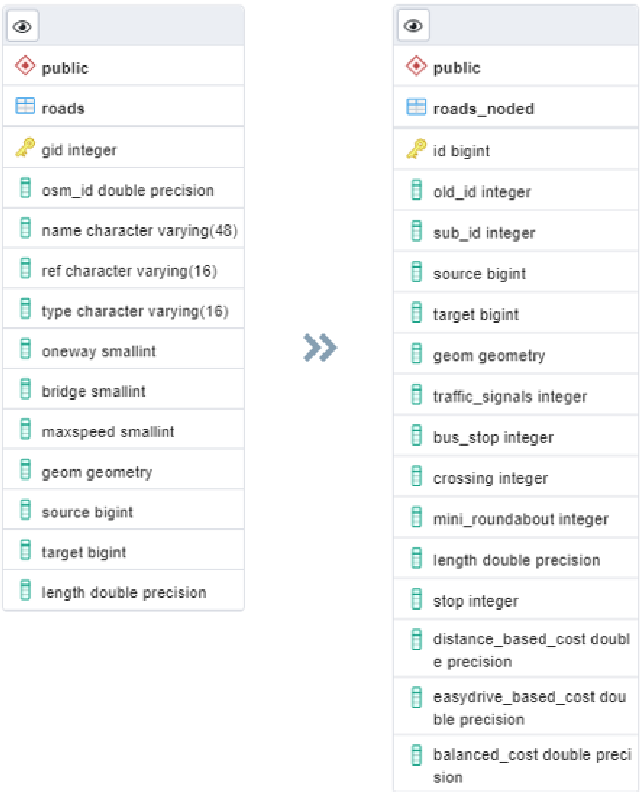


Figure 4:Building routable road network.

The algorithm chosen to be implemented is Djikstra's algorithm (pgr_Djikstra). The PgRouting extension has this same algorithm, being necessary to define the several parameters, namely the cost indication (in this project there are three types of cost), source and target.

Dijkstra's Algorithm (E.W. Dijkstra) is one of the algorithms that calculates the minimum cost path between vertices of a graph.

To make the indication of the source and target easier, the author proposes to indicate a point of interest (see section 3.3.4) instead of indicating coordinates. In this case, it is necessary to associate to each point of interest the nearest vertex. With an SQL query, it is possible to search for roads nodes that are within a very minimal distance from each location using the PostGIS function ST_DWithin and specifying an optional distance parameter.

```
UPDATE points_of_interest SET points_of_interest_id=v.id
 FROM roads_noded_vertices_pgr v
  WHERE ST_Dwithin(points_of_interest.geom, v.the_geom,0.001);
```

Some of the points of interest in the table do not have an id associated with them. The reason is because they are not within the specified distance to a road node. In this case, it is used the following SQL line to delete them.

```
DELETE from points_of_interest
 WHERE points_of_interest IS NULL;
```

Below is the algorithm developed in SQL.

```
SELECT s.path_seq, s.node, s.edge, s.cost, ST_astext(b.geom) AS
edge_geom_as_text, b.geom AS edge_geom, ST_astext(v.the_geom) AS
node_geom_as_text, v.the_geom AS node_geom
--edges
 FROM (SELECT seq , path_seq , node, edge, cost , agg_cost
 FROM pgr_dijkstra('SELECT id, source, target, [insert the desired column cost]
AS cost FROM
roads_noded',
--source
  (SELECT points_of_interest_id from points_of_interest WHERE
points_of_interest.name='[insert here name of point of interest]'),
--target
  (SELECT points_of_interest_id from points_of_interest WHERE
points_of_interest.name='[insert here name of point of interest'),FALSE
))AS s
LEFT JOIN roads_noded b
 ON (id = s.edge)
LEFT JOIN roads_noded_vertices_pgr v
 ON (v.id = s.node)
ORDER BY s.path_seq;
```

3.3.6. Routes visualization

The routes are visualised using QGIS software.

QGIS is a free and open-source software, which allows you to view, analyse, manage, edit and build maps. Additionally, this software can support various vector, raster and database formats and functionalities.

In this research project, the QGIS 3.26.3 Buenos Aires version was used.

This software was chosen because it enables the connection with PostGIS. In this way, it is possible to visualise the result of the execution of the Routing algorithm (using PgRouting) in QGIS.

First it is necessary to create tables to store the results of each route obtained with the implementation of the Djikstra algorithm.

Secondly, in the QGIS software using the QuickMapServices plugin, a layer of the OpenStreetMap standard map and a layer of the route obtained by connecting the QGIS software to the PostGIS software are added.

Figure 5 presents an overview of the architecture development.

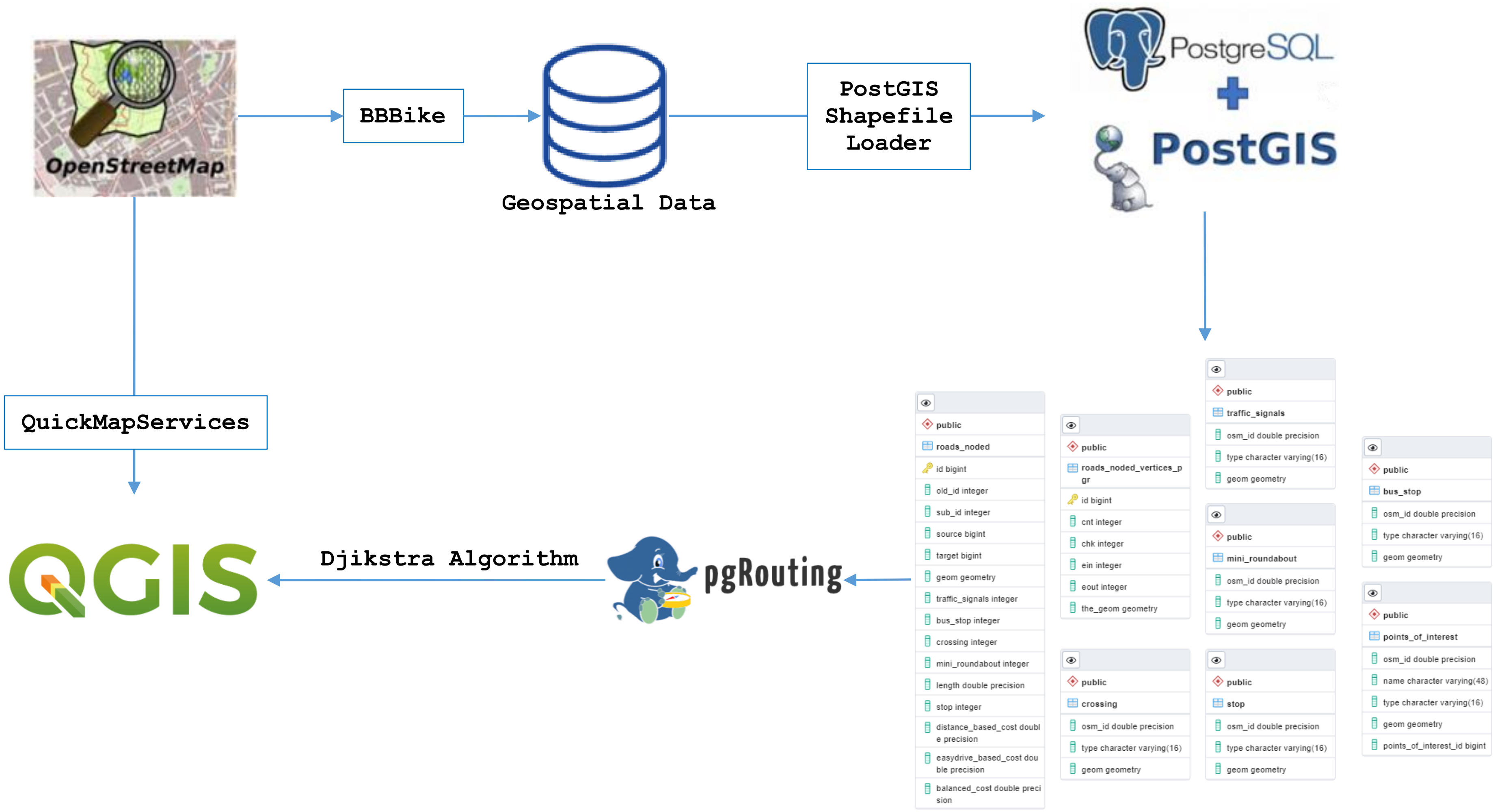


Figure 5:Overview of the architecture development.

4. RESULTS AND DISCUSSION

This section presents the results obtained with the development of the architecture and implementation of the model.

4.1. Presentation of results

In the reference period (2022), it was possible to create the necessary conditions to obtain results, namely:

- Design and implement the data warehouse;

- Create an operational database;

- Carry out the importation of data into the data warehouse;

- Develop a model with the drivers' preferences;

- Execute a routing algorithm;

Considering the results obtained from the tasks developed, it was possible to establish a solid basis for the continuation of the project.

The QGIS software allows visualizing the routes obtained through the result of Djikstra's algorithm implemented in PostgreSQL software from the PgRouting extension.

Firstly, the OSM map stantard was added through the QuickmapServices plugin.

Secondly, QGIS was connected to the PostgreSQL database (Lisbon_Routing) and the connection was made with the tables which stored the PgRouting results, namely route_p1 (contains the route obtained for person 1), route_p2 (contains the route obtained for person 2) and route_p3 (contains the route obtained for person 3).

Figure 6 shows an example of three routes obtained through different preferences given by people to different dimensions and criteria.

Equation 1, Equation 2 and Equation 3 were used to calculate the effort each person has when traversing each edge of the path.

The cost for the shortest path is calculated by.Equation 1.

$$SP_i = L_i, \forall\ i\ \in I \qquad \text{Equation 4}$$

Where,
L_i – length of edge i

For the easy drive mode route, the cost is calculated by Equation 2.

$$ED_{ip} = \sum_{c=1}^{C} \alpha_{cp} N_{ic}, \forall\ i \in I, c \in C \qquad \text{Equation 5}$$

Where,

i – edge i

c – criterion c

α_{cp} – importance given by person p to criterion c

N_{ic} – number of elements of criterion c existing in edge i

For the balanced route, the cost is calculated by Equation 3.

$$B_i = D_{1p}L_i + D_{2p}\sum_{c=1}^{C}\alpha_{cp}N_{ic}\,, \forall\, c \in C \qquad \text{Equation 6}$$

Where,

D_{1p} – importance given by person p to dimension 1

D_{1p} – importance given by person p to dimension 2

α_{cp} – importance given by person p to criterion c

N_{ic} – number of elements of criterion c existing in edge i

L_i – length of edge i

The preferences that were considered are reflected in Table 3 and

Table 4.

Table 3:Importance given to each dimension.

D_{jp} p	Shortest Path $D = 1$	Easy Drive $D = 2$
$p = 1$	1,0	-
$p = 2$	0,5	0,5
$p = 3$	-	1

Table 4:Importance given to each criterion.

α_{cp} p	Traffic Signals $c = 1$	Crossing $c = 2$	Mini Roundabout $c = 3$	Stop $c = 4$	Bus Stop $c = 5$
$p = 1$	-	-	-	-	-
$p = 2$	2	1	1	1	0
$p = 3$	1	2	2	1	1

As previously mentioned, our test consists of validating our method using an urban area. The location chosen was Lisbon, Portugal, where the origin is the *Museu Arqueológico do Carmo*, with destination to the *Museu do Dinheiro*.

The blue line refers to the route obtained for a weighting system based solely on distance, so the path obtained is the shortest path.

The pink line refers to the route obtained in which the user gives 50% importance to the distance to be travelled and 50% importance to the easy drive mode.

Finally, the orange line is the route obtained when the user gives 100% preference to the easy drive mode.

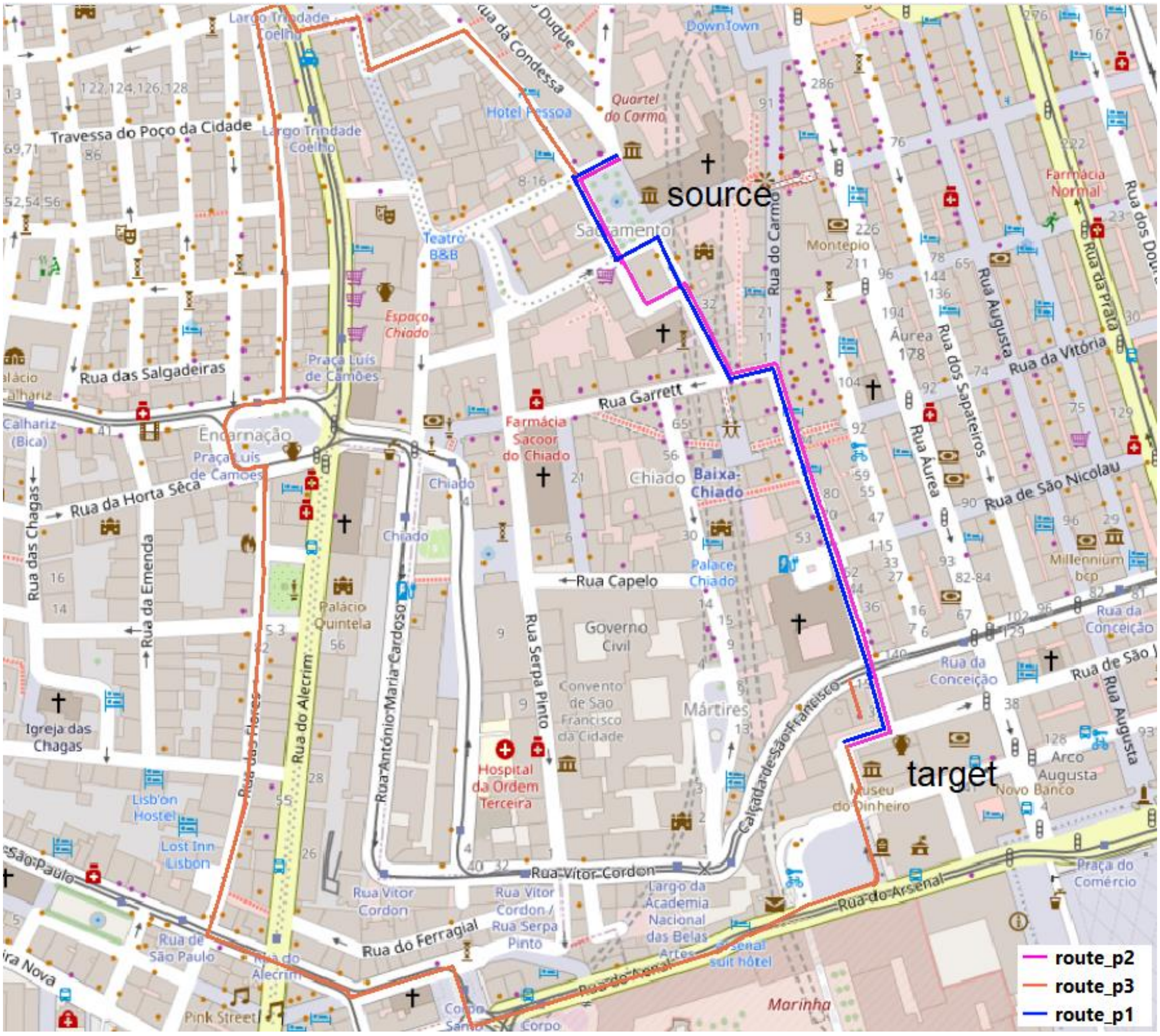


Figure 6:Visualization of routes with different weights.

Table 5 shows the quantitative outputs of the three routes obtained, namely, the distance and the number of elements existing in the route for each attribute (traffic lights, crossings, mini roundabouts, stop signals and bus stop).

Table 5:Quantitative output of routes.

	Route Person 1	Route Person 2	Route Person 3
Distance	846,09	846,24	2 420,60
Number of traffic lights	7	5	3
Number of crossings	7	6	3
Number of mini roundabouts	0	0	0
Number of stop signals	3	3	1
Number of bus stop	0	0	0

The following section analyses and discusses the results obtained.

4.2. Discussion of results

Obtaining a functional prototype of the data warehouse and route planning model, allowed us to test all the different components that make up this project, from the data import process to the creation of an operational database for the execution of routing algorithms, as well as the incorporation of the different features that characterise the innovation of the project and, finally, the performance of the routing algorithm.

With the implementation of the routing algorithm, three routes were generated for three different driver profiles. In this case, person 1 prefers to opt for the shortest route, person 2 prefers a balanced scenario, in which he attributed a 50% importance to distance and 50% to ease of driving, and person 3 gives importance only to ease of driving.

Considering the preferences of each person for each dimension (table 3) and for each attribute (table 4), it can be seen macroscopically in figure 6 that the algorithm allowed the generation of three different routes, which, in itself, reveals that the weighting system worked.

After analysing the quantitative output in table 5, it can be ascertained that the result is consistent with the weight of each route, that is, the route where the distance has the most weight, turns out to be, in fact, the route with the shortest path. Next, the balanced route (person 2) demonstrated the correct functioning of the algorithm, having reduced the passage through traffic lights and zebra crossings and increased the distance residually. Finally, the route where the driver's preference lies only in an easy-to-drive route, the algorithm looks for all possible paths between the origin and the destination to decrease the passage through traffic lights, zebra crossings, stop signs, mini-roundabouts, and bus stops, even if this implies an increase of about three times the distance to be travelled.

In fact, this model proves to be a proof of concept, by being able to consider different weights, namely, being able to offer a route to the driver that balances the distance to be travelled with the ease of condition.

Regarding data quality, it should be noted that it did not reach the desired level for the following reasons:

- Path characterisation: although it was possible to detect the existence of some attributes such as traffic lights, zebra crossings, roundabouts, bus stops and stop signs, we wanted to include other attributes such as intersections, speed bumps, priority lanes and left and right turns. Although OSM has several advantages, we must consider that it is a platform where any registered person can edit the map, so there may be poorly mapped attributes, as well as areas where the mapping is not at all the most complete.
- Routable Graph: through the functions pgr_analyzeGraph and pgr_nodeNetwork, it was possible to break the segments into several different segments connected through vertices, however, when the graph was analysed again it was possible to find that some segments remained disconnected.

With this research, we show that the model enables the driver a route where traffic lights, zebra crossings, roundabouts, stop signs and bus stops are avoided.

5. CONCLUSION

This chapter summarizes the conclusions of the dissertation regarding the research work developed. A critical analysis of the project is carried out, including deviations from objectives, its main contributions, the added value of the project for the scientific community and, subsequently, the limitations of the work are presented, followed by a proposal for future work.

5.1. Critical analysis of deviations from objectives

In terms of deviations regarding tasks that took longer than expected, due to the need felt to perform numerous iterations, deviations at the database implementation level stand out.

The implementation of a database that supports roads with a configurable number of attributes proved to be a challenge, which caused us some delays.

The simplest way to deal with this problem would be the denormalization of roads table, keeping the roads attributes information separate from the roads information itself, i.e., we would have a table with the list of roads and another table with the list of attributes that each road would have. However, such approach proved to be not feasible due to a very low performance.

After transforming the data to routable data, the solution was then to use only a single table of roads that would have a variable number of columns, one for each attribute. This implies having an automated process that creates and modifies the table based on the existence of each of the attributes in the roads. This solution proved to be viable and enabled the evolution to the next step which was to create columns for the cost that each edge represents.

Regarding the scientific and technological objectives that this project set out to achieve, a critical analysis of the deviations from the objectives is described.

- STO1 - Specify and develop a database structure for route planning needs, to include the following features:
    - Integration of multiple data sources with different characteristics;
    - Centralized and properly integrated data storage area that will provide added value for analysis;
    - Suitability of data for routing algorithms execution.

Regarding this goal, the project managed to build a centralized and properly integrated data storage area that will provide added value for analysis and data transformation to allow the execution of a routing algorithm. However, since only one data source was used, the current database structure does not allow the integration of more data sources with different characteristics.

- STO2 - Creation of a system capable of supporting the requirements for a smooth operation of the Easy Drive mode:
    - Host different routing algorithms;
    - Performance, ensuring that value can be extracted from the data;
    - Scalability, ensuring that the system will not be limited by the number of drivers profiles;

- Elasticity, making it possible to accommodate other data sources, i.e., data from different sources and formats, as well as relating to any geographical location on the globe.

Relatively to this goal, in this project only one routing algorithm was used and tested - the Djikstra algorithm. However, we believe that the database is prepared for the execution of other algorithms such as A* and Bellman Ford.

In terms of performance, although the data quality did not reach the desired level (due to external factors), the project reveals itself as a proof of concept and it was possible to extract value from the data. Regarding Scalability, the project reached this point, being possible to obtain a route for each set of preferences of each driver. Regarding Elasticity, as mentioned before, the current database structure would not work properly with data from another data source in a different format. However, the system is prepared to integrate another data source as long as it is in the same format, working for any other geographic place in the globe.

- STO 3 -Develop a model that allows, at any time, to change route preferences according to each user.

The project successfully achieved this goal.

- STO4 - Develop an analytical service:
  - Visual analysis of the routes generated by executing routing algorithms;
  - Key Performance Indicators (KPI), to measure performance according to the parameters that are considered relevant to the quality of the system.

This objective was successfully achieved, with the project achieving results in the reference period. Although limitations and improvements for future research can be identified (described in the next sections), it was possible to visualise the routes obtained on a map as well as to analyse and extract value from the results, according to the metrics defined.

5.2. Concluding remarks

Despite all the limitations associated with this project, it can be seen as a proof-of-concept of the incorporation of road network features to show that, despite a longer route, it represents a higher satisfaction for drivers as it meets their needs.

Many driving condition variables were kept constant in this model to directly compare a route based on driver preferences with a route based on distance.

The present project is aligned with a research and development strategy, attentive to the trends of incorporating geographic information systems for route planning, to investigate and develop solutions to meet the needs of drivers when in an urban environment. In this context, this project will be a solution that represents a good starting point for the development of route planning solutions that embrace the needs and preferences of each individual, to increase user satisfaction and gain a significant competitive advantage in the market by the disruptiveness and innovation of concept.

This research work contributes to literature by incorporating the driver's needs and requirements regarding factors that provide greater driving ease, in the route selection process.

5.3. Limitations and future research

Even though this project has some contributions to the routing literature, it presents several limitations. First, due to the data extracted from OpenStreetMap it was not possible for us to include all the features desired to bring the routing problem as close as possible to a real-world scenario, such as intersections, left and right turns, priority lanes, road sharing with cyclists, among others. Nevertheless, the work developed is the first step in the development of a system that embraces the drivers' preferences, suggesting an easy-to-drive route

Secondly, although our study is a proof of concept and has, in fact, presented different results according to different user preferences, it is necessary to take into account that OpenStreetMap, despite all its advantages already mentioned, is a map that is constantly being built and updated by volunteers, which implies that, in some areas, it may, on the one hand, not be complete in respect to all tag keys and tag values and, on the other hand, there are objects that may be wrongly mapped. Additionally, the quality of the routable graph cannot be guaranteed.

For future work it would be interesting to analyse other features to create different route options, namely the creation of a panoramic route and the creation of a route option essentially for tourism, with the inclusion of several points of interest in the route planning, such as monuments, gardens, parks, traditional restaurants, pleasant landscapes, among others, instead of providing only the shortest path.

REFERENCES

Achuthan, N. R., Caccetta, L., & Hill, S. P. (2003). An improved branch-and-cut algorithm for the capacitated vehicle routing problem. Transportation Science, 37(2), 153-169.

Angelelli, E., & Speranza, M. G. (2002). The periodic vehicle routing problem with intermediate facilities. European journal of Operational research, 137(2), 233-247.

Baltrėnas, H. P., Januševičius, T., & Chlebnikovas, A. (2017). Research into the impact of speed bumps on particulate matter air pollution. *Measurement*, *100*, 62-67.

Barnhart, C., Johnson, E. L., Nemhauser, G. L., Savelsbergh, M. W., & Vance, P. H. (1998). Branch-and-price: Column generation for solving huge integer programs. Operations research, 46(3),316-329.

Becker, H. S. (1996). The epistemology of qualitative research. *Ethnography and human development: Context and meaning in social inquiry*, 27, 53-71.

Braekers, K., Hartl, R. F., Parragh, S. N., & Tricoire, F. (2016). A bi-objective home care scheduling problem: Analyzing the trade-off between costs and client inconvenience. European Journal of Operational Research, 248(2), 428-443.

Brandao, J., & Mercer, A. (1997). A tabu search algorithm for the multi-trip vehicle routing and scheduling problem. European journal of operational research, 100(1), 180-191.

Bräysy, O., & Gendreau, M. (2005). Vehicle routing problem with time windows, Part II: Metaheuristics. Transportation science, 39(1), 119-139.

Cattaruzza, D., Absi, N., Feillet, D., & Vidal, T. (2014). A memetic algorithm for the multi trip vehicle routing problem. European Journal of Operational Research, 236(3), 833-848.

Chen, Y. L., & Yang, H. H. (2000). Shortest paths in traffic-light networks. T*ransportation Research Part B: Methodological*, 34(4), 241-253.

Claes, R., Holvoet, T., & Weyns, D. (2011). A decentralized approach for anticipatory vehicle routing using delegate multiagent systems. *IEEE Transactions on Intelligent Transportation Systems*, *12*(2), 364-373.

Clarke, G., & Wright, J. W. (1964). Scheduling of vehicles from a central depot to a number of delivery points. Operations research, 12(4), 568-581.

Cohen, A., & Dalyot, S. (2021). Route planning for blind pedestrians using OpenStreetMap. *Environment and Planning B: Urban Analytics and City Science*, 48(6), 1511-1526.

Crevier, B., Cordeau, J. F., & Laporte, G. (2007). The multi-depot vehicle routing problem with inter-depot routes. European journal of operational research, 176(2), 756-773.

Dantzig, G., Fulkerson, R., & Johnson, S. (1954). Solution of a large-scale traveling-salesman problem. Journal of the operations research society of America, 2(4), 393-410.

Dantzig, G. B., & Wolfe, P. (1960). Decomposition principle for linear programs. Operations research, 8(1), 101-111.

Dell'Amico, M., Righini, G., & Salani, M. (2006). A branch-and-price approach to the vehicle routing problem with simultaneous distribution and collection. Transportation science, 40(2), 235-247.

Desrochers, M., Desrosiers, J., & Solomon, M. (1992). A new optimization algorithm for the vehiclerouting problem with time windows. Operations research, 40(2), 342-354.

Džafić, D., Candiotti, J. L., & Cooper, R. A. (2020, July). Improving wheelchair route planning through instrumentation and navigation systems. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 5737-5740). IEEE.

Ludke, M., & André, M. (1986). Pesquisa em educação: abordagens qualitativas. In *Aberto*, 5(31).

Felício, S., Hora, J., Ferreira, M. C., Abrantes, D., Costa, P. D., Dangelo, C., ... & Galvão, T. (2022). Handling OpenStreetMap georeferenced data for route planning. *Transportation Research Procedia*, 62, 189-196.

Ferreira, M., & d'Orey, P. M. (2011). On the impact of virtual traffic lights on carbon emissions mitigation. *IEEE Transactions on Intelligent Transportation Systems*, 13(1), 284-295.

Fleischmann, B., Gnutzmann, S., & Sandvoß, E. (2004). Dynamic vehicle routing based on online traffic information. Transportation science, 38(4), 420-433.

Hakeem, A., Gehani, N., Ding, X., Curtmola, R., & Borcea, C. (2019, November). Multi-destination vehicular route planning with parking and traffic constraints. In *Proceedings of the 16th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services* (pp. 298-307).

He, Y., Wang, X., Zhou, F., & Lin, Y. (2019). *Dynamic vehicle routing problem considering simultaneous dual services in the last mile delivery.* Kybernetes.

Hrncir, J., Song, Q., Zilecky, P., Nemet, M., & Jakob, M. (2014). Bicycle route planning with route choice preferences. In *ECAI 2014* (pp. 1149-1154). IOS Press.

Hsu, L. S., & Obe, R. (2021). *PostGIS in action.* Simon and Schuster.

Hu, L., Yang, J., & Huang, J. (2016). The real-time shortest path algorithm with a consideration of traffic-light. *Journal of Intelligent & Fuzzy Systems*, 31(4), 2403-2410.

Hu, L., Zhong, Y., Hao, W., Moghimi, B., Huang, J., Zhang, X., & Du, R. (2018). Optimal route algorithm considering traffic light and energy consumption. Ieee Access, 6, 59695-59704.

Jensen, A. F., Rasmussen, T. K., & Prato, C. G. (2020). A route choice model for capturing driver preferences when driving electric and conventional vehicles. *Sustainability*, 12(3), 1149.

Keenan, P. (2008). Modelling vehicle routing in GIS. *Operational Research*, 8(3), 201-218.

Knoop, V. L., & Daganzo, C. F. (2018). The effect of crosswalks on traffic flow. *European Journal of Transport and Infrastructure Research*, 18(2).

Krisp, J. M., & Keler, A. (2015). Car navigation–computing routes that avoid complicated crossings. *International Journal of Geographical Information Science*, 29(11), 1988-2000.

Goel, R., & Maini, R. (2017). Vehicle routing problem and its solution methodologies: a survey. International Journal of Logistics Systems and Management, 28(4), 419-435.

Gomory, R. E. (2010). Outline of an algorithm for integer solutions to linear programs and an algorithm for the mixed integer problem. In 50 Years of Integer Programming 1958-2008 (pp.77-103). Springer, Berlin, Heidelberg

Khanjary, M., Faez, K., Meybodi, M. R., & Sabaei, M. (2011, May). Shortest paths in synchronized traffic-light networks. In *2011 24th Canadian Conference on Electrical and Computer Engineering (CCECE)* (pp. 000882-000886). IEEE.

Kulkarni, R. V., & Bhave, P. R. (1985). Integer programming formulations of vehicle routing problems. European journal of operational research, 20(1), 58-67.

Kumar, S. N., & Panneerselvam, R. (2012). A survey on the vehicle routing problem and its variants.

Kwon, Y. J., Choi, Y. J., & Lee, D. H. (2013). Heterogeneous fixed fleet vehicle routing considering carbon emission. Transportation Research Part D: Transport and Environment, 23, 81-89.

Land, A. H., & Doig, A. G. (2010). An automatic method for solving discrete programming problems. In 50 Years of Integer Programming 1958-2008 (pp. 105-132). Springer, Berlin, Heidelberg.

Liao, X., Zhou, T., Wang, X., Dai, R., Chen, X., & Zhu, X. (2022). Driver route planning method based on accident risk cost prediction. *Journal of advanced transportation*, 2022.

Li, F., Golden, B., & Wasil, E. (2007). The open vehicle routing problem: Algorithms, large-scale testproblems, and computational results. Computers & operations research, 34(10), 2918-2930.

Lin, S., & Kernighan, B. W. (1973). An effective heuristic algorithm for the traveling-salesman problem. Operations research, 21(2), 498-516.

Lin, X., Geng, F., Jin, Z., & Xu, Q. (2016). The optimization of delivery vehicle scheduling considering the actual road network factors. In 2016 *International Conference on Logistics, Informatics and Service Sciences (LISS)* (pp. 1-7). IEEE.

Lysgaard, J., Letchford, A. N., & Eglese, R. W. (2004). A new branch-and-cut algorithm for the capacitated vehicle routing problem. Mathematical Programming, 100(2), 423-445.

Nallur, V., Elgammal, A., & Clarke, S. (2015, May). Smart route planning using open data and participatory sensing. In *IFIP International Conference on Open Source Systems* (pp. 91-100). Springer, Cham.

Meneguzzer, C., Gastaldi, M., Rossi, R., Gecchele, G., & Prati, M. V. (2017). Comparison of exhaust emissions at intersections under traffic signal versus roundabout control using an instrumented vehicle. *Transportation research procedia*, *25*, 1597-1609.

Min, H. (1989). The multiple vehicle routing problem with simultaneous delivery and pick-up points. Transportation Research Part A: General, 23(5), 377-386.

Mirabi, M., Ghomi, S. F., & Jolai, F. (2010). Efficient stochastic hybrid heuristics for the multi-depot vehicle routing problem. Robotics and Computer-Integrated Manufacturing, 26(6), 564-569.

MirHassani, S. A., & Abolghasemi, N. (2011). A particle swarm optimization algorithm for open vehicle routing problem. Expert Systems with Applications, 38(9), 11547-11551.

Mitchell, J. E. (2002). Branch-and-cut algorithms for combinatorial optimizationproblems. Handbook of applied optimization, 1, 65-77.

Ogden, K. W. (1992). Urban goods movement: a guide to policy and planning.

Osman, I. H., & Wassan, N. A. (2002). A reactive tabu search meta-heuristic for the vehicle routing problem with back-hauls. Journal of Scheduling, 5(4), 263-285.

Padberg, M., & Rinaldi, G. (1991). A branch-and-cut algorithm for the resolution of large-scale symmetric traveling salesman problems. SIAM review, 33(1), 60-100.

Pau, M., & Angius, S. (2001). Do speed bumps really decrease traffic speed? An Italian experience. *Accident Analysis & Prevention*, *33*(5), 585-597.

Pérez-Sansalvador, J. C., Lakouari, N., Garcia-Diaz, J., & Hernández, S. E. P. (2020). The effect of speed humps on instantaneous traffic emissions. *Applied Sciences*, 10(5), 1592.

Petch, R. J., & Salhi, S. (2003). A multi-phase constructive heuristic for the vehicle routing problem with multiple trips. Discrete Applied Mathematics, 133(1-3), 69-92.

Poli Jr, J., & Monteiro, L. H. (2005). Improving vehicle flow with traffic lights. *Advances in complex systems*, 8(01), 59-63.

Poljak, R., Poščić, P., & Jakšić, D. (2017, May). Comparative analysis of the selected relational database management systems. In *2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 1496-1500). IEEE.

Pradenas, L., Oportus, B., & Parada, V. (2013). Mitigation of greenhouse gas emissions in vehicle routing problems with backhauling. Expert Systems with Applications, 40(8), 2985-2991.

Renaud, J., Laporte, G., & Boctor, F. F. (1996). A tabu search heuristic for the multi-depot vehicle routing problem. *Computers & Operations Research*, 23(3), 229-235.

Sariklis, D., & Powell, S. (2000). A heuristic method for the open vehicle routing problem. *Journal of the Operational Research Society*, 51(5), 564-573.

Sarraf, R., & McGuire, M. P. (2018). A data integration and analysis system for safe route planning. In *Proceedings of the International Conference on Information and Knowledge Engineering (IKE)* (pp. 111-117). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

Schambers, A., Eavis-O'Quinn, M., Roberge, V., & Tarbouchi, M. (2018, April). Route planning for electric vehicle efficiency using the Bellman-Ford algorithm on an embedded GPU. In *2018 4th International Conference on Optimization and Applications (ICOA)* (pp. 1-6). IEEE.

Solé, L., Sammarco, M., Detyniecki, M., & Campista, M. E. M. (2022). Towards drivers' safety with multi-criteria car navigation systems. *Future Generation Computer Systems*, 135, 1-9.

Sun, J., & Liu, H. X. (2015). Stochastic eco-routing in a signalized traffic network. *Transportation Research Procedia*, 7, 110-128.

Tang, C. H., & Yan, S. (2007). A routing and scheduling framework incorporating real-time adjustments for inter-city bus carriers under stochastic travel times and demands. *Journal of the Chinese Institute of Engineer*s, 30(4), 635-649.

Taş, D., Dellaert, N., van Woensel, T., & De Kok, T. (2014). The time-dependent vehicle routing problem with soft time windows and stochastic travel times. T*ransportation Research Part C:Emerging Technologies*, 48, 66-83.

Teng, L., Tian, J., & Izumi, T. (2019, December). A Proposal of Dynamic Route Search Method to Consider the Individual Driver Preferences. In *2019 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1226-1232). IEEE.

Toth, P., & Vigo, D. (Eds.). (2002). *The vehicle routing problem*. Society for Industrial and Applied Mathematics.

Vanderbei, R. J. (2020). *Linear programming: foundations and extensions* (Vol. 285). Springer Nature.

Várhelyi, A. (2002). The effects of small roundabouts on emissions and fuel consumption: a case study. *Transportation Research Part D: Transport and Environment*, 7(1), 65-71.

Vidal, T., Crainic, T. G., Gendreau, M., Lahrichi, N., & Rei, W. (2012). A hybrid genetic algorithm for multidepot and periodic vehicle routing problems. *Operations Research*, 60(3), 611-624.

Wang, R., Zhou, M., Gao, K., Alabdulwahab, A., & Rawa, M. J. (2021). Personalized route planning system based on driver preference. *Sensors*, 22(1), 11.

Wang, Z., & Niu, L. (2018). A data model for using OpenStreetMap to integrate indoor and outdoor route planning. *Sensors*, 18(7), 2100.

Wang, Z., Novack, T., Yan, Y., & Zipf, A. (2020). Quiet route planning for pedestrians in traffic noise polluted environments. *IEEE Transactions on Intelligent Transportation Systems*, 22(12), 7573-7584.

Yu, B., Yang, Z. Z., & Yao, B. Z. (2011). A hybrid algorithm for vehicle routing problem with time windows. *Expert Systems with Applications*, 38(1), 435-441.

Zachariadis, E. E., & Kiranoudis, C. T. (2010). An open vehicle routing problem metaheuristic for examining wide solution neighborhoods. *Computers & Operations Research*, 37(4), 712-723.

Zhu, L., & Hu, D. (2019). Study on the vehicle routing problem considering congestion and emission factors. *International Journal of Production Research*, 57(19), 6115-6129

Zhu, S., & Levinson, D. (2015). Do people use the shortest path? An empirical test of Wardrop's first principle. *PloS one*, 10(8), e0134322.

Zurbarán, M., Kraft, T., Mather, S. V., Park, B., & Wightman, P. (2018). *PostGIS Cookbook: Store, organize, manipulate, and analyze spatial data*. Packt Publishing Ltd.