RAFAEL OLIVEIRA PEREIRA

PERSONALIZED ROUTE GENERATION FOR MOUNTAIN BIKING CYCLING BASED ON THE USER'S PROFILE

Dissertation presented to the Computer Science Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Magister Scientiae*.

Advisor: Jugurta Lisboa Filho

VIÇOSA - MINAS GERAIS

2020

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T

P436p
2020

Pereira, Rafael Oliveira, 1992-
Personalized route generation for mountain biking cycling based on the user's profile / Rafael Oliveira Pereira. – Viçosa, MG, 2020.
65 f. : il. (algumas color.) ; 29 cm.

Orientador: Jugurta Lisboa Filho.
Dissertação (mestrado) - Universidade Federal de Viçosa.
Referências bibliográficas: f.64-65.

1. Ciclismo. 2. Sistemas de informação geográfica. 3. Trilhas para mountain bike. I. Universidade Federal de Viçosa. Departamento de Informática. Programa de Pós-Graduação em Ciência da Computação. II. Título.

CDD 22. ed. 005.74

RAFAEL OLIVEIRA PEREIRA

PERSONALIZED ROUTE GENERATION FOR MOUNTAIN BIKING CYCLING BASED ON THE USER'S PROFILE

Dissertation presented to the in Computer Science Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Magister Scientiae*.

APPROVED: 21st July 2020.

Assent:

Rafael Oliveira Pereira
Author

Jugurta Lisboa Filho
Advisor

*To everyone around me during this time. You were essential to the completion of this work***.**

ACKNOWLEDGEMENTS

This thesis was written in very hard times. I would not have been able to complete this work without the help of all those who were around me during this time. I apologize in advance if I forgot to explicitly mention your name here but be assured that I still have you in my mind. My deepest thank you to all of you.

I would like to thank my advisor Prof. Jugurta Lisboa Filho, for being with me during my bachelor's and also during my master's, and for pushing me until the end. His advice guided me to the best path, and he has always been a great person to talk to. He was the one who suggested and who convinced me to join this program and his presence has always been invaluable. During this journey Prof. Jugurta has been, and I hope will continue to be, a friend.

My co-advisor Prof. André Gustavo dos Santos, for providing me with the knowledge to make this project viable. He provided me with great opportunities and helped me immensely during all this time. Whenever I needed help, Prof. André stepped up. He was also someone I could talk to when in need.

My father Aloisio, and my mother, Mara, who have always been by my side supporting me. My brother Rodrigo, who has been living with me and standing my antics maybe more than anyone else. My girlfriend Joice, who has been with me since shortly after I entered this great university and has been especially supportive of me in these hard moments. I also cannot forget to mention my cat, Kissa, who was my only company for a long time.

My oldest friend Reuben, who has always been someone I can count on, no matter the time, no matter the topic. Ademir, Dinoel, Gustavo and Jean, who have been good friends since I met them, during my university years, sometimes in unexpected circumstances. My friend Pedro, who helped me with math and algorithms.

I am very grateful for all my UFV professors for teaching me everything I needed, not only for this project but also for my future, and for pushing me to learn more by myself as well. I am thankful for the countless workers of the university that maintained the building I worked in and were always already hard at work when I arrived in the morning. Jean Câmara, who was my *advisor* during the development of ClickOnMap. Victor Macedo, who shared the laboratory with me.

I also want to thank Cleiton, Vinício, Geissiane, Lidson, João Marcos among others from Univiçosa for understanding and supporting me, especially in the end, by allowing me to work from home.

An honorable mention to the Unimore professors Manuel Iori and Daniele Pretolani, whom I met in Italy and who gave me the directions I needed to guide my project.

If I were to mention every good attribute of every one of you this text would be endless. I owe this victory to all of you because, directly or indirectly, you made me grow.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

“"I wish it need not happened in my time," said Frodo. "So, do I," said Gandalf, "and so do all who live to see such times. But that is not for them to decide. All we have to decide is what to do with the time that is given us."”

J.R.R Tolkien, *The Fellowship of the Ring*.

ABSTRACT

PEREIRA, Rafael Oliveira, M.Sc., Universidade Federal de Viçosa, July, 2020. **Personalized route generation for Mountain Biking cycling based on the user's profile**. Advisor: Jugurta Lisboa Filho.

The popularity of cycling has been on the rise both as a sustainable transport alternative and as a leisure activity, through Mountain Biking (MTB). As its popularity increases, so does the need for tools to aid in the activity, such as route sharing tools. Most of these tools rely on Volunteered Geographic Information (VGI) both to acquire new trails for their databases and to rank them. While these tools are useful in most places, they can be of little to no help in regions with a smaller number of cyclists. This work proposes using data collected from multiple sources of VGI to automatically generate MTB routes based on user preferences, easing the decision making process of choosing new trails. Due to its vast availability, segments from Strava, a social network for athletes, were chosen as one of the data sources, the other being mapping data from the collaborative mapping tool OpenStreetMap(OSM), which was also used to select Points Of Interest (POI) relevant to the activity. An Integer Linear Programming model was developed to select sets of segments considering user preferences for terrain difficulty, elevation in the trail and total distance of the route, focusing on unpaved streets. The method developed in this work showed that it is possible, through the use of Strava segments, selected POIs and mapping data from OSM, to create pleasant circuits based on user preferences, with the main challenge being the quality of the collaborative data available in OSM. The circuits created with the method proved to be, in most part, pleasant, visiting relevant POIs and avoiding paved streets wherever possible.

Keywords: Cycling. MTB. VGI. Route Generation

RESUMO

PEREIRA, Rafael Oliveira, M.Sc., Universidade Federal de Viçosa, julho de 2020. **Geração de rotas personalizadas para ciclismo de Mountain Bike com base no perfil do usuário**. Orientador: Jugurta Lisboa Filho.

O ciclismo vem se popularizado tanto como uma alternativa de transporte sustentável, quanto uma alternativa de lazer, através do Mountain Biking (MTB). Com isso, o uso de ferramentas que auxiliam nesta prática, tais como ferramentas de compartilhamento de rotas, se faz cada vez mais necessária. Atualmente, a maior parte das ferramentas disponíveis fazem uso de Informação Geográfica Voluntária (VGI) tanto para adicionar novas trilhas aos seus bancos de dados quanto para avaliá-las. Porém, em lugares com poucos ciclistas elas perdem boa parte de sua utilidade. Este trabalho propõe utilizar dados de múltiplas fontes de VGI para gerar rotas de MTB automaticamente com base nas preferências de um usuário, facilitando o processo de tomada de decisão para a escolha de novas trilhas. Devido à sua vasta difusão, segmentos do Strava, uma rede social para atletas, foram escolhidos como uma das bases de dados, sendo a outra os dados de mapeamento da ferramenta de mapas colaborativos OpenStreetMap (OSM), que foi também utilizada para selecionar Pontos de Interesse (POI) relevantes à atividade. Foi desenvolvido um modelo de Programação Linear Inteira para selecionar um conjunto de segmentos e POIs de forma a respeitar as preferências do usuário em relação à dificuldade de terreno, elevação na trilha e distância total da rota, focando em estradas não pavimentadas. O método desenvolvido neste trabalho mostrou ser possível, através do uso de segmentos do Strava, POIs selecionados e dados de mapeamento do OSM, gerar circuitos agradáveis, baseados nas preferências de um usuário, tendo como principal desafio a qualidade dos dados colaborativos disponíveis no OSM. Os circuitos gerados com este método se mostraram, em sua maioria, agradáveis, visitando POIs relevantes à atividade e permanecendo em estradas não pavimentadas onde possível.

Palavras-chave: Ciclismo. MTB. VGI. Geração de rotas.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

Chapter I: Introduction

1.1. The problem and its relevance

Bicycles were introduced in the XIX century in Europe and now are the main means of transportation for many people around the world. Cycling provides many benefits to the cyclist: from benefits such as body aesthetic and conditioning improvement, and a general bettering on quality of life (MORRIS; GUERRA, 2015), to a reduction on automobile traffic and pollution (GRANVILLE et al., 2001), which in turn reduces the number of traffic accidents and the cyclists exposure to pollution. Even with the risks of inhaling air pollution, which has an even bigger toll than automobile accidents, and getting involved in traffic accidents ("Environment: New policy package to clean up Europe's air," [s.d.]), the benefits of the added physical activity by bike commuting still outweigh the risks (DE HARTOG et al., 2010). Some European capitals adopt policies to encourage people to use bikes as a form of commute by implementing bike-rental systems across cities, by considering cyclists when planning for public transport routes, be it by adding bike parking spots in places close to stations or by adding bike-carrying equipment in trains (PUCHER; BUEHLER, 2009). Cyclists are also considered during city planning, where suggestions to build cycling lanes are made with intent to reduce the perceived dangers of cycling in cities (UNWIN, 1995). The interest for cycling in the context of public health emerged in 1995 (CARVALHO; FREITAS, 2012) with a study recommending at least 30 minutes of physical activity every day of the week.

Aside from health benefits, the benefits listed above are also financial, since less pollution in cities and healthier people leads to less people using the health system and, with traffic congestion being bad for the economy, productivity increases too (GRANVILLE et al., 2001). Compared to cars and public transportation, cycling infrastructure is inexpensive and takes less space.

Besides being an option for commuting, cycling is a very popular sport, with racing competitions in many different categories, including road bicycle racing, time trial, cyclo-cross, mountain bike racing, track cycling, Bicycle Motocross (BMX) and cycle speedway, each with its own rules, environment and equipment. Casual riders usually opt for road bicycling and mountain bicycling, since those have the most common and accessible environments, paved roads and off-road terrain, respectively. This project has a focus on Mountain Biking.

Mountain Biking (MTB) consists of riding bicycles over off-road, often very rough terrain. For cyclists, finding a good trail to ride is an important part of the sport. Having a trail nearby may increase the amount of activity (TUCKER; MANAUGH, 2018), which makes local information important to the sport. There are many ways to find a trail to ride (REDDY et al., 2010): by checking maps and guides, exploring, asking other cyclists or, more recently, checking online tools for routes recommended by other athletes.

The process of improving existing trails is also a challenge (KOEMLE; MORAWETZ, 2016), given that an ideal trail has to vary in attributes to please riders with different levels of experience and, in some cases, share the space with hikers. Having a place with many attractive trails brings economic benefits to a region since cyclists will travel to ride in good trails (FIX; LOOMIS, 1997) and cyclists spend more money and visit local shops more often than car drivers (CLIFTON et al., 2013).

The evolution of the Internet gave us the Web 2.0, leaving a world of static pages behind and offering interactive pages to users. With interaction being a possibility, emerged the Voluntary Geographic Information (VGI), term proposed by Goodchild (GOODCHILD, 2007) to describe a way of collaboration in which the person gives, voluntarily, information about an area of its knowledge. Any type of information given voluntarily containing a geographic location is considered VGI. This development in technology is what made possible the union of cycling and VGI, in the shape of online platforms where cyclists share their own routes. With that being said, the *word of mouth* is still the standard way of sharing information between people who practice the sport.

The online sharing of routes between athletes, made possible by the Web 2.0, has been changing the way trails are chosen for an activity. The sharing of routes is as important for Mountain Bikers as it is for Road Cyclists, however, in places where not many people ride or places not known for having good trails, often there are not sufficient local routes in online tools, making it hard for people to find new places to ride.

Using other, more complete and diverse sources of VGI, may help with that problem by creating routes from the street network and segments of road that are frequently used by cyclists, without depending on people to share entire routes.

Online sharing tools often feature options to filter routes to a cyclist's preferences, filtering by distance, elevation, difficulty, among other options. Requirements for a trail may vary between people: “difficulty” for a road cyclist means the elevation of a segment, meanwhile for MTB riders it might mean a very irregular terrain, but otherwise flat; “security” for road cyclists means traffic, which does not apply to mountain bikers. Even with all these tools available, finding a good trail may still be toilsome, sometimes even boring, and even more when there are not many available trails nearby, thus becoming hard to find one that fits the cyclist's preferences. Therefore it is important to use many sources of collaborative information when building a route with the aid of decision tools.

1.2. Hypothesis

The use of multiple sources of VGI as a base of data on the automatic route generation for MTB, using optimization techniques to best attend to cyclists' preferences, may help or replace the process of choosing a new route to ride, which is made today mostly by individually checking routes made by other cyclists using apps and platforms specific for sharing complete routes, not necessarily with all the information about it, or by talking with other athletes.

1.3. OBJECTIVES

The objective of this project is to develop a model that uses voluntary collaboration of geographic data to help in process of choosing a route to be traveled for a Mountain Biking activity, taking into consideration the cyclist's profile.

The specific objectives are:

- Prepare a spatial database of segments, each segment having attributes such as distance, elevation, terrain, using VGI sources;
- Create a list of Points Of Interest (POI), which are points that might be relevant or interesting during a trail ride, for example, a waterfall;
- Develop a model that, from this spatial database, is able to suggest a list of routes according to the cyclist's profile;
- Develop a prototype of the model to analyze the efficiency and viability of the model and of the route generation algorithm.

1.4. Thesis Structure

This thesis is structured as such: Chapter 2 introduces VGI and VGI platforms, presents and analyzes some of the post popular tools used by cyclists, and introduces the concept of Linear Programming. Chapter 3 explains the project in depth and goes into what decisions were taken to make it work.

On Chapter 4 we run tests, evaluate and discuss the results and problems we ran into during testing. We also propose solutions to problems encountered and discuss possible changes. Chapter 5 exposes the conclusions and presents opportunities for future work and research.

Chapter II: Tools and Related Researches

In this chapter we review some of the existing literature on the subject of cycling routes, pointing out the differences and common points between available literature and this work. Then we introduce the concept of VGI through the literature about the subject and present OpenStreetMap[1], which is the most known VGI platform. We take a look at some of the existing route sharing tools, analysing their features and their faults. In the end we introduce the concept of Linear Programming, which is used to mathematically model the problem described in this work.

2.1. Literature Review

Most of the reviewed works on the subject of cycling routes approach the subject in a different manner than in this work, either by trying to identify a profile for an already existing network of cyclists, aiming to improve the quality of cycling infrastructure, or by creating a profile according to an individual cyclist's previous activities.

(BROACH; DILL; GLIEBE, 2012) uses GPS data from many cyclists in Portland, Oregon, in a choice set generation algorithm to discover their route preferences, (ZIMMERMANN; MAI; FREJINGER, 2017) does the same in Eugene, Oregon, using a recursive logit model and (MENGHINI et al., 2010) uses an improved mode detection algorithm to find the route choice preferences of cyclists in Zurich, Switzerland, all of them with a focus on improving cycling facilities on city planning. (HRNCIR et al., 2014) applies the A* routing algorithm in a model representing a city's transport network feature to optimize routes for travel time, comfort, quietness and flatness in the city of Prague.

(STROOBANT et al., 2018) works in a context of recreational routing to find a route that starts and ends in the same location, fitting a specified distance. It suggests a heuristic to the NP-hard problem and compares it to the optimal solution. One of the metrics it uses to ensure an enjoyable route is *roundness*, which represents how round a route is.

(KRISMER et al., 2016) takes into consideration user profiles to build a bicycle router that learns from the user's previous trips to suggest more appropriate routes.

---

[1] https://www.openstreetmap.org/

They use GraphHopper (which we describe better later in this chapter) to find the path, OpenStreetMap to plot the route and SRTM (Shuttle Radar Topography Mission[2]) to add elevation data to the planner.

2.2. Volunteered Geographic Information

Accurate and updated maps are a fundamental resource for a great set of applications. Through them one can have, for example, services based on geolocation that help on economical activity, services that help keeping society's health and well being and services that provide detailed maps of areas that need humanitarian help in cases of disasters (FOODY et al., 2017).

Nowadays changes are constant and fast, and citizens are an important source of geographical information. Recent technological advancements made GPS equipped devices available and accessible, facilitating the collaboration for common people without any expertise in the area. Until a few years ago, mapping was a costly and time consuming task, as it needed expensive equipment and skilled people to use them. Nowadays, with techniques requiring relatively small amounts of geographic data, it is possible to make a trustworthy and less costly mapping. One way to do that is by using the enormous potential of sensory citizens, individuals that collaborate with mapping systems through geolocation data, mainly acquired through their mobile devices. For example, if a person lives in an unmapped region, they can collaborate to a mapping system by adding their street network to this system or even photos of the places local to them. Collaborated data can be anything containing a geographic coordinate: from the single coordinate indication the position of something to a picture of a specific place.

This technological development allows VGI to be used in emergency situations, that require updated maps to aid in decision making and fast acting. Tools like HOT (Humanitarian OpenStreetMap Team[3] and Ushahidi[4] (OKOLLOH, 2009) are the embodiment of this evolution, both being humanitarian help tools.

[2] https://www2.jpl.nasa.gov/srtm/
[3] https://www.hotosm.org/
[4] https://www.ushahidi.com/

2.3. OpenStreetMap

OpenStreetMap (OSM) is a project through which people without any geography or cartography knowledge are able to collaborate in the creation of free geographical maps. OSM, as opposed to other similar projects, does not impose legal or technical restrictions to the way users want to utilize the platform, leaving them free to use it in creative and unexpected ways (OpenStreetMap Wiki, 2018).

Released in 2004, it is one of the biggest VGI projects that exist now. Similar to a Wikipedia for maps, through OSM it is possible to create and edit maps from anywhere on the planet. In addition to street maps, the platform also has cycling, transport and humanitarian maps.

The project started with the idea that if each user collaborates with geographical information about its own region, which is a place they know better than others, a spatial database of a region could be created and, if a lot of people do the same, it is possible to map the entire world. OSM has maintained a great growth and relative success since its creation, either because of the growth of the interactive web, of the great availability of low cost GPS devices or because of how easy it is to collaborate with data for the platform (MOONEY; MINGHINI; OTHERS, 2017).

In spite of the frequent lack of routing data, OSM manages to present itself as an attractive option to cyclists searching for new routes for an activity, either for leisure or for a commute. These routes can take into consideration different factors, such as elevation difference, distance, auto traffic and scenic routes (BERGMAN; OKSANEN, 2016). Data registered and sports apps increase even more this potential.

OSM organizes its data in nodes, ways and relations, which are described with tags in a typical `key=value` structure. A node is a single point in the world, comprised of a pair of (latitude, longitude) coordinates. A way, besides being used in the conventional sense of the word, as in to describe a street, is also used to determine the boundaries of a park or a forest. A relation, the most complex of the three, is an element used to combine the other two to represent a bigger entity, as in to comprise a complex of buildings into a single entity.

2.3.1. Tags

Tags are used to describe specific features on map elements. A tag consists of two items, a key and a value, represented as `key=value`, which are both text field, but also able to represent numeric fields. A lot of information about the edges on OSM comes form the tags, some of which are relevant for this work.

2.3.2. Surface

The surface tag provides additional information about the physical state of the roads described on OSM, mostly the material composition of a path. The material composition of a road is important in classifying the difficulty of a path, for example, an asphalt paved road offers no challenge at all, while gravel or ground surfaces are a little harder to traverse.

2.3.3. Smoothness

The smoothness tag provides a measure for the physical usability of a way, aimed at wheeled vehicles. The more regular and flat a way is, the smoothest it is. The values of this tag go from excellent, meaning skates and roller blades are able to traverse it, to impassable, meaning no wheeled vehicles are able to travel through it.

2.3.4. Tourism

This tag indicates places and things which might interest tourists, such as accommodations (hotel, hostel, motel etc), camp sites, art galleries and artwork sites. The value “attraction”, as in `tourism=attraction`, covers any type of touristic facility and is often used in conjunction with other tags, such as “amenity”, “building” and “historic”. For this project, the most relevant value for the tag would be `tourism=viewpoint`, which is an attraction more likely to be available on rural areas, compared to other values.

2.3.5. Amenity

The amenity tag is used to describe facilities that might be relevant for visitors and residents. In our case, it can be used to find Points of Interest (POIs) to the cyclist, which would be touristic attractions, such as churches and fountains, especially when paired with `tourism=attraction`, or supporting places, like a drinking fountain or a biking repair station. Despite being usually found in urban environments, some of these features, specially drinking fountains, are too important to be simply left out, because cyclists might want to replenish or fill up their water containers before, in the case of forgetting to fill it up at home, during, if they run out of water or want fresher water, or even after an activity, to have a more refreshing drink at a moment when the remaining water is usually warm.

2.3.6. Water

Water features are sought after in bike activities, especially waterfalls, even more when it is one with a swimmable water pool, thus, we add some water feature in the list of POIs. A lake can be represented in many ways, depending on its classification: assuming that all the following tags are paired with `natural=water`, a lake can be described with:

- `water=lake`, for a body of water located in a basin surrounded by land;
- `water=pond`, for non natural lakes, which are usually smaller;
- `water=lagoon`, which represents a shallow body of water separated from the sea;
- `water=stream_pool`, which describes “a small and usually deep collection of fresh water, supplied by a spring, or occurring along a stream or river”;
- `water=reservoir`, for a reservoir, a natural or man made lake contained by a dam.

For rivers, the only tags relevant as Points of Interest, coupled with water=river, are:

- `waterway=tidal_channel`, for tidal waterways, in which the direction of the stream changes with the tide;
- `waterway=canal`, for a canal built for use by vessels.

The current tagging convention for waterfalls is:

- `waterway=waterfall` (for nodes);
- `natural=cliff` (for ways).

All the elements can also have a `name=*` tag when the information is available.

2.4. GraphHopper

GraphHopper[5] is a fast and memory efficient Java routing engine. It has been released under Apache License 2.0. By default, it uses OpenStreetMap for mapping data and CGIAR for elevation data, but it can import other data sources like Mapbox for mapping and SRTM for elevation. GraphHopper is the routing engine used in the official OpenStreetMap website.

Through GraphHopper's Directions API it is possible to find the best route from point A to point B with turn instructions and additional data on the route, such as elevation difference and surface type. The API also has route optimization features, being able to solve the travelling salesman problem for up to 200 vehicles on paid plans and only 1 vehicle for free accounts.

Although it is a paid service, it has very limited free accounts, through which is possible to use all of the features offered by it, but with smaller number: for example, with a free account it is possible to make up to 500 daily requests to its routing API, versus 50.000 requests in a premium account.

2.5. Route sharing tools

With the rising popularity of the sport, many route sharing systems have arisen. While some of them are more focused on registering and sharing the activity with friends, working as a social network, others focus more on sharing the routes and less on the person who is sharing. Some of them use OpenStreetMap data to present the routes or even to plan them.

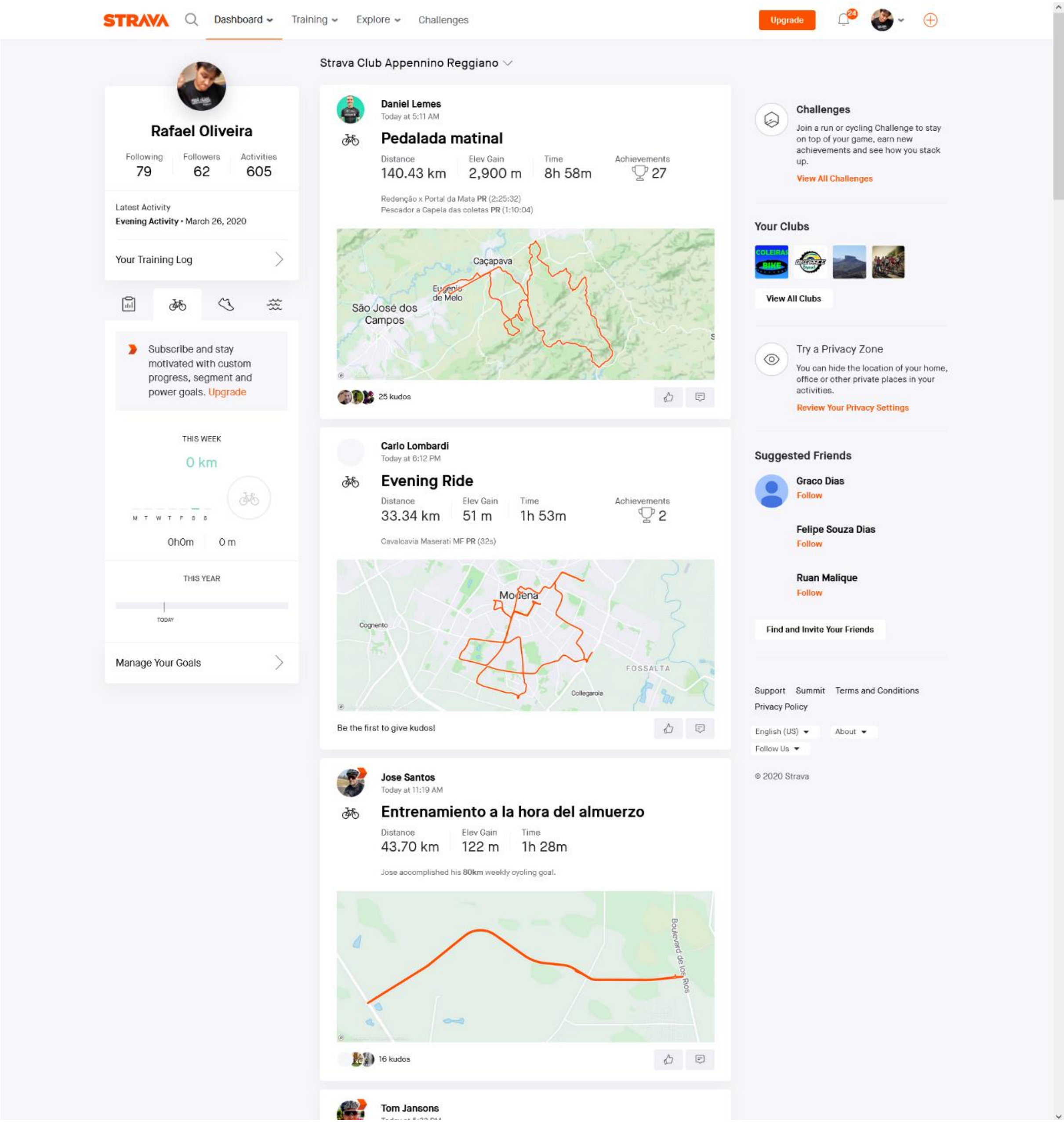
2.5.1. Strava

Strava[6] is the most popular social network for athletes, widely used by mainly cyclists and runners, who share their activities on the platform. It is possible to add photos to an activity that show up in friends timelines. People can create and be part of a group of athletes, in which one can view the status' of all participants such as number of activities and distance ridden, and it is also possible to create text posts and to interact with other people in groups. Strava uses GPS devices to track, record and

[5] www.graphhopper.com/
[6] https://www.strava.com/

share activities in a Facebook-like feed of activities, seen in Figure 1. Strava Dashboard.

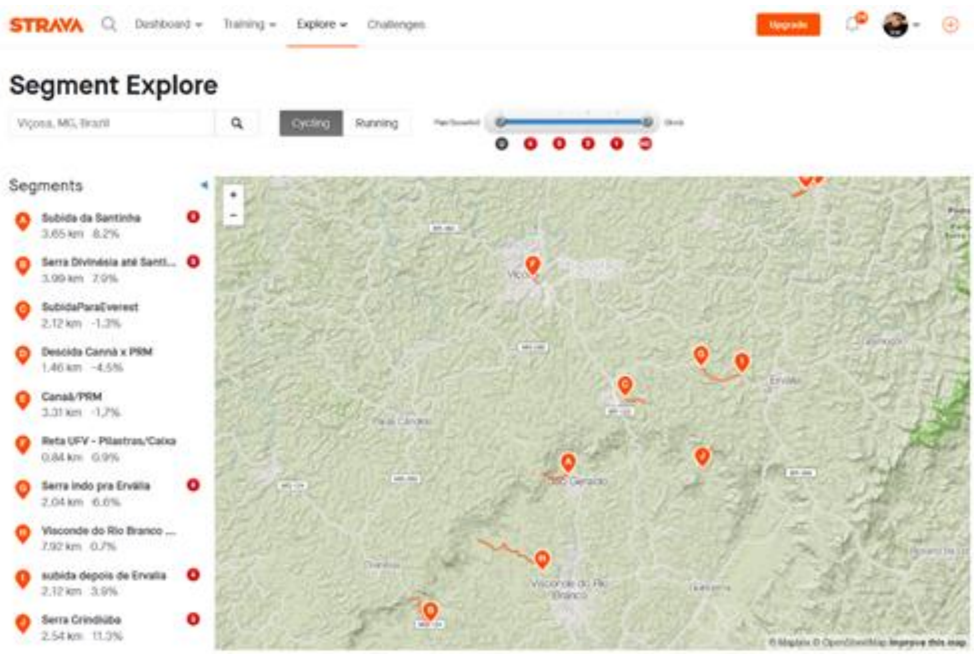
Figure 1. Strava Dashboard.



Source: Strava's website

An activity recorded in the app contains, for example, estimated power and energy output, calorie consumption, moving time, a map with the traveled route and elevation data about said route. Through the platform it is also possible to create, save and edit a custom route, however, it is not possible to automatically generate them. Although it is possible to see activities from people in your area and, thus, the route they traveled through, this is not the focus of the social network. Still in the social aspect, Strava has *segments*, which are segments of roads deemed relevant by other athletes, for which the app records times and ranks people for that place, instilling some competitivity in its users. The person with the best time in a segment becomes a King or Queen of the Mountain. It is possible to explore *segments* through Strava's Segment Explore tool, shown in Figure 2. Strava segments explore.

Figure 2. Strava segments explore.

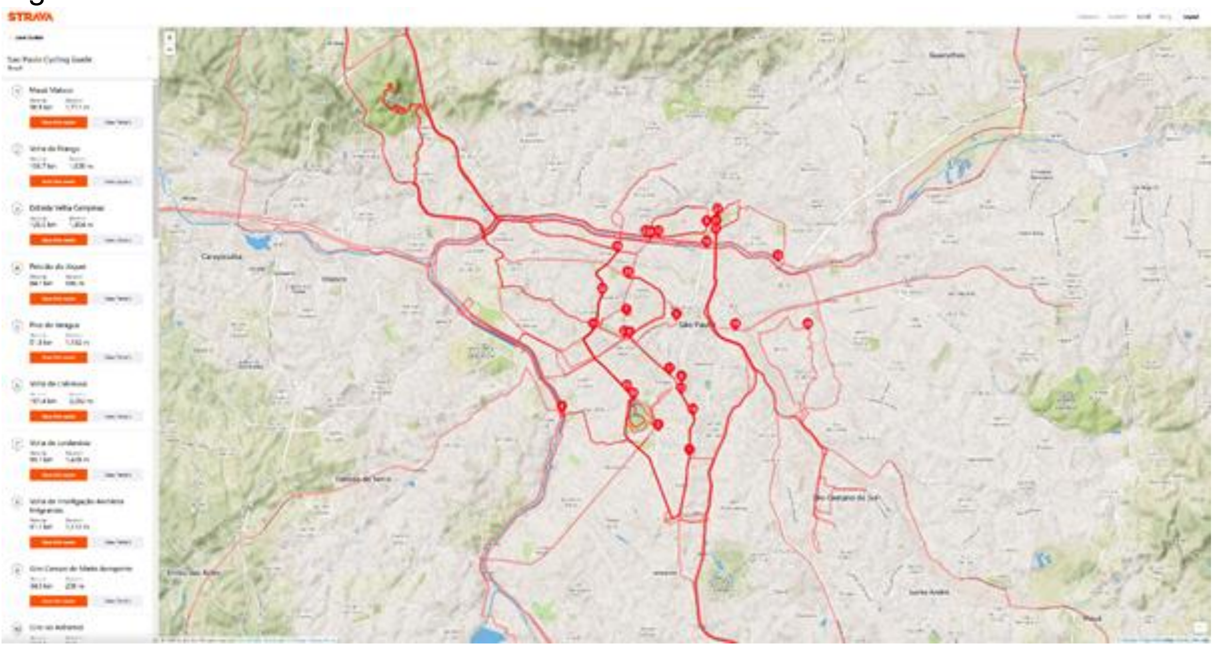


Source: Strava’s website

Strava also has curated routes, called “Local guides”, shown in Figure 3. Strava Local Guides.

Source: Strava’s website, for a few places in the world, such as Rio de Janeiro and São Paulo in Brazil. Curated routes have a better description and recommended stops along the way but otherwise have the same amount of detail as others.

Figure 3. Strava Local Guides.



Source: Strava’s website

Another interesting feature of the platform, geared to the most competitive or enthusiastic users, is the “Challenges”: there are challenges for many sports and many categories, such as complete 15 hours of walking or hiking, climb a total of 7500 metes on a bike or complete a 100 km ride. Challenges can also be used as Ads for companies, in which a company creates a challenge and offers products for the people who complete them.

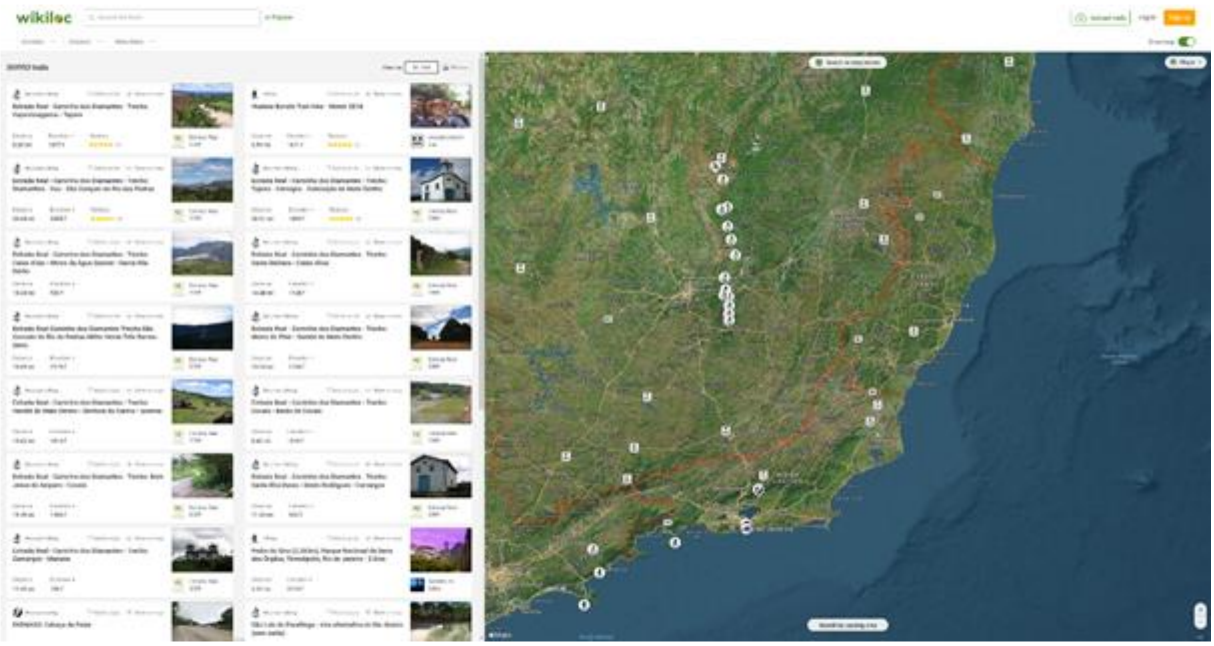
2.5.2. Wikiloc

Wikiloc[7] is a route sharing platform where athletes share the complete route they have traveled through in an activity. Wikiloc allows users to pick a type of route (a circuit or one-way only), a distance and a level of difficulty. It is also possible to track an activity using the app and a phone's GPS or to draw a route by hand in the website, for which the website calculates distance and climbing height. Also, the platform is not focused on a single kind of sport. It goes from mountain biking to bird watching to Joëlette (a one-wheeled chair held by two assistants). Every route in the platform was put there by one of its users via uploading a GPX file or by tracking an activity in real time using the app, and if another user has done the same trail, they can also add data to it.

In the website it is possible to search for trails in a map, as seen in Figure 4. Wikiloc map, filtering it by type of activity, distance, elevation, technical difficulty, among others.

Figure 4. Wikiloc map.

[7] https://www.wikiloc.com/

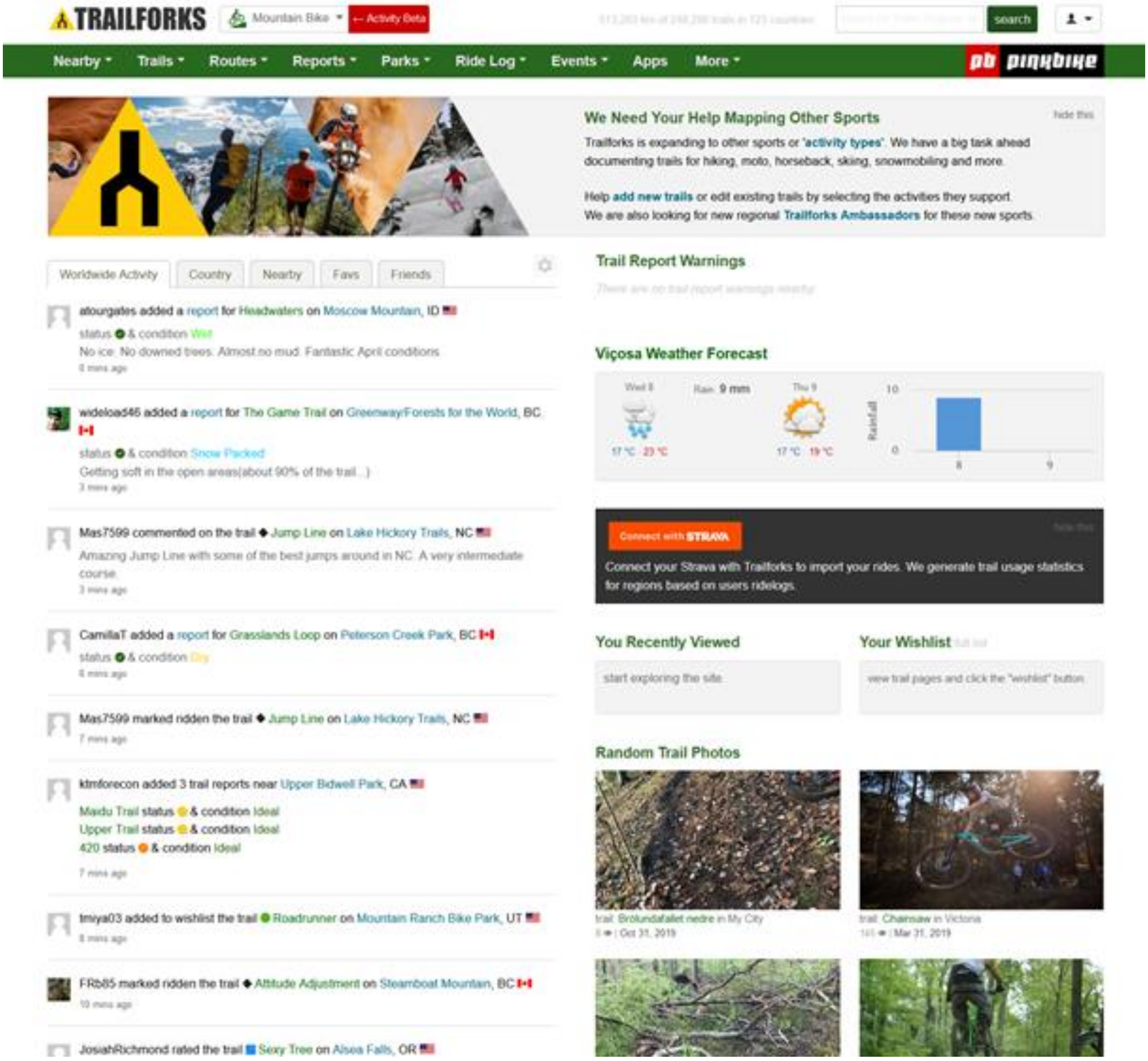


Source: Wikiloc’s website

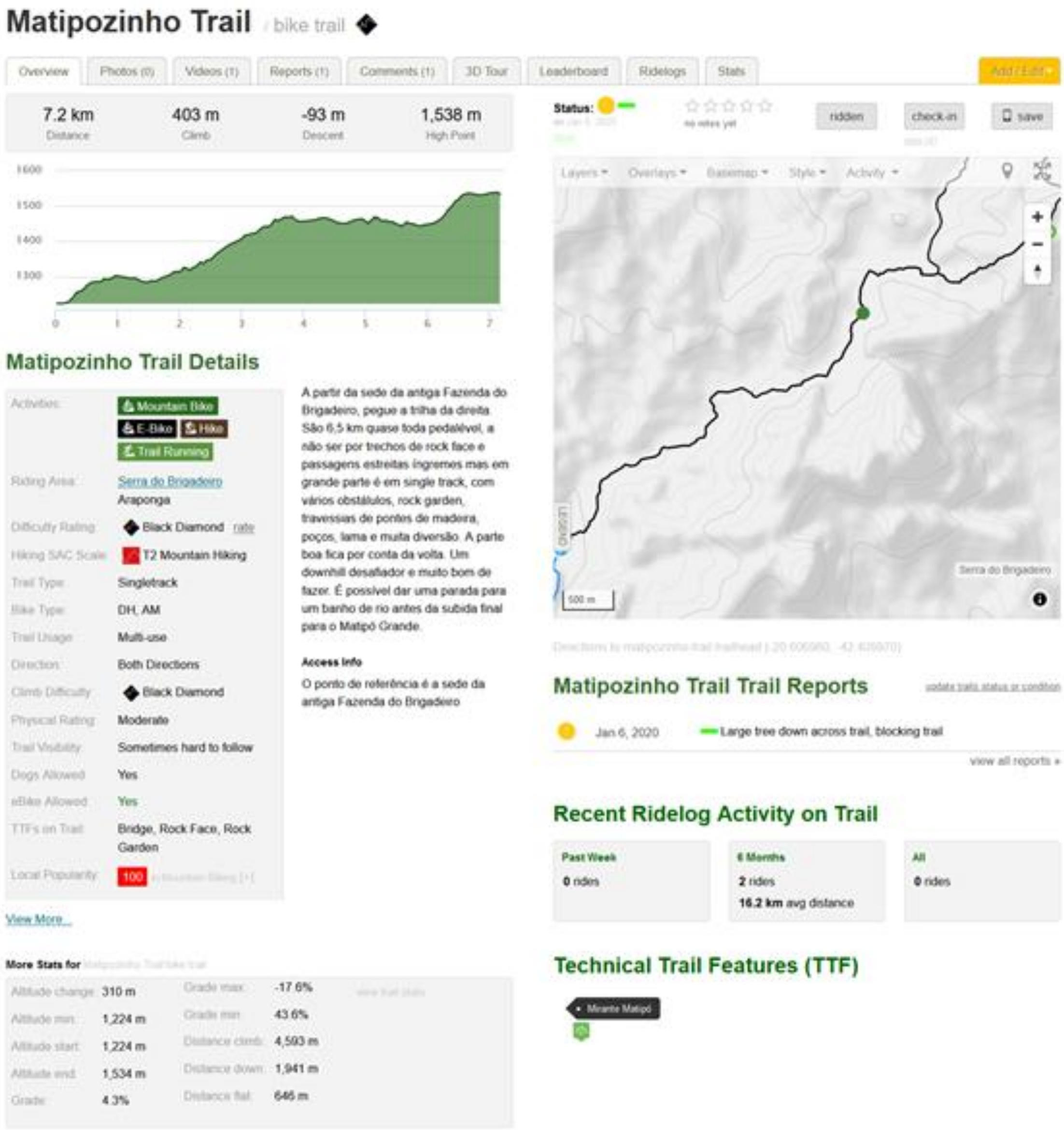
2.5.3. Trailforks

Trailforks[8], maintained by Pinkbike, is a trails and maps database where cyclists can upload their own trails, by tracking an activity or uploading a GPX file. The website's front page, seen in Figure 5. Trailforks’ website.

[8] https://www.trailforks.com/

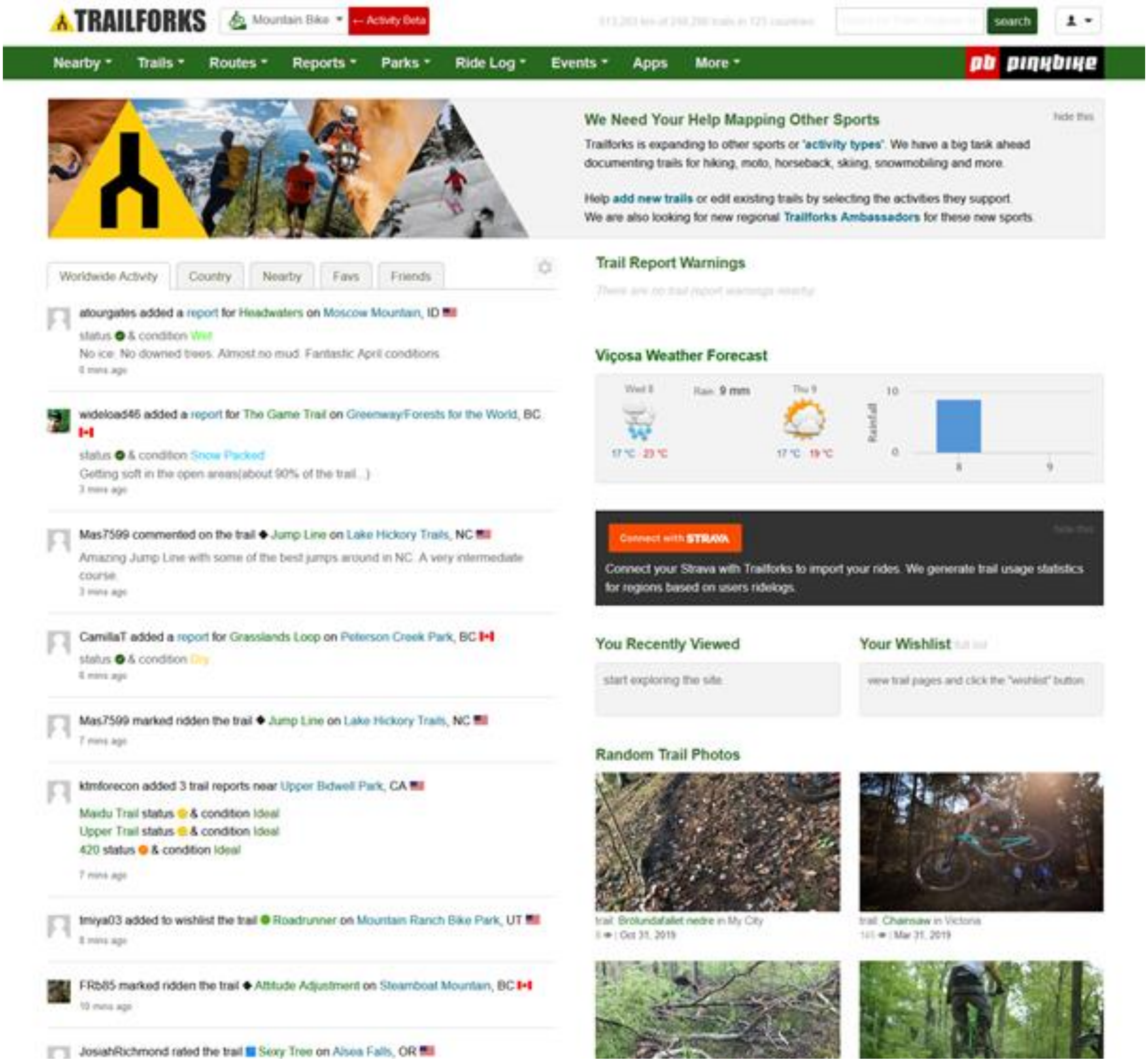


Source: Trailforks' website, shows recent activity as well as photos from random trails around the world. It is possible for anyone to mark any trail as ridden and to evaluate it in a scale from Easiest to Extremely Dangerous. When a trail is uploaded, the owner can add data such as distance, elevation grade, trail type, bike type and direction, among others, as seen inFigure 6. Trailforks' route.



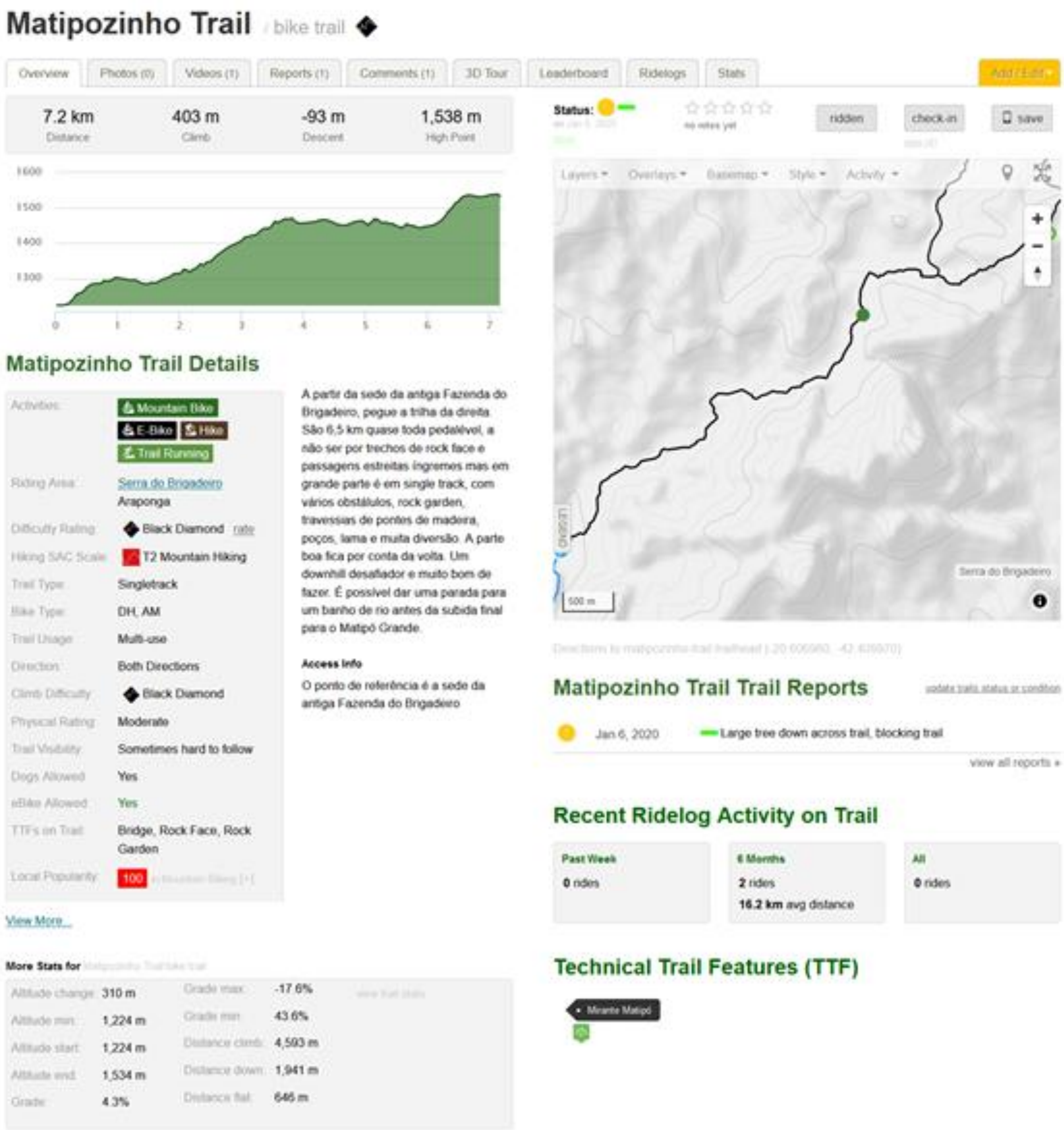
Source: Trailforks' website. Routes without enough detail are not accepted.

Figure 5. Trailforks' website.



Source: Trailforks' website

Figure 6. Trailforks' route.



Source: Trailforks' website

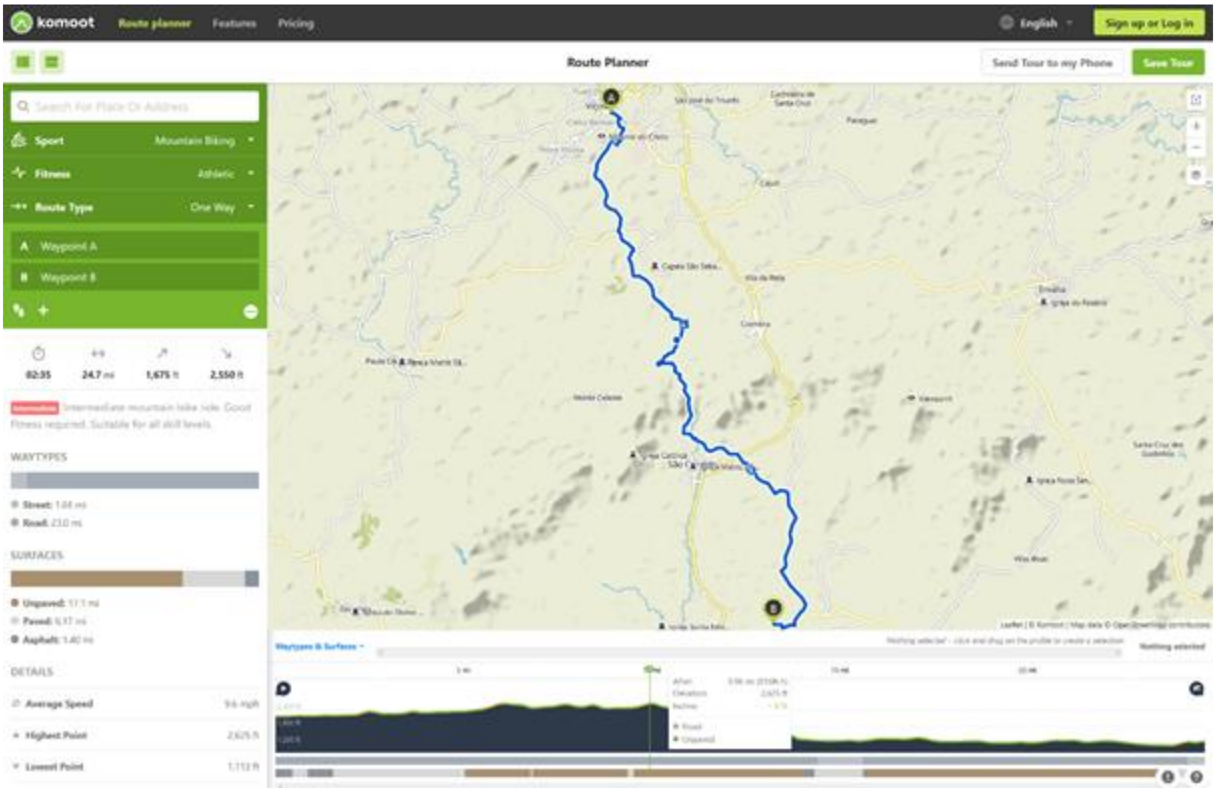
The Discover Routes feature recommends routes based on trails the user marked as ridden, activities tracked with the app or activities in the user's Strava account. Another way to find trails is to check the Ride Finder and filter by personal preferences, such as activity type, difficulty, distance and climbing height. One can also create their own Ride Plan by using the route planner, but unless the route goes through a trail already logged in the platform, it will not be able to populate the route's attributes.

The reports section in the website serves to indicate the current and past status of a trail, which can be dry, wet, muddy or icy, among others. It can also be used to indicate if a trail is closed. The parks section is used to show the location of bike parks and riding areas around the world.

Although initially the platform focused only on mountain biking, it has been expanding to accommodate other types of off-road activities, so a trail can be suitable to more than one activity, like mountain biking, hiking and trail running, but be primarily used for mountain biking.

2.5.4. Komoot

Komoot[9] is a social network for athletes with activity tracking and route planning features. Komoot's route planner, shown in Figure 7. Komoot's route planner.

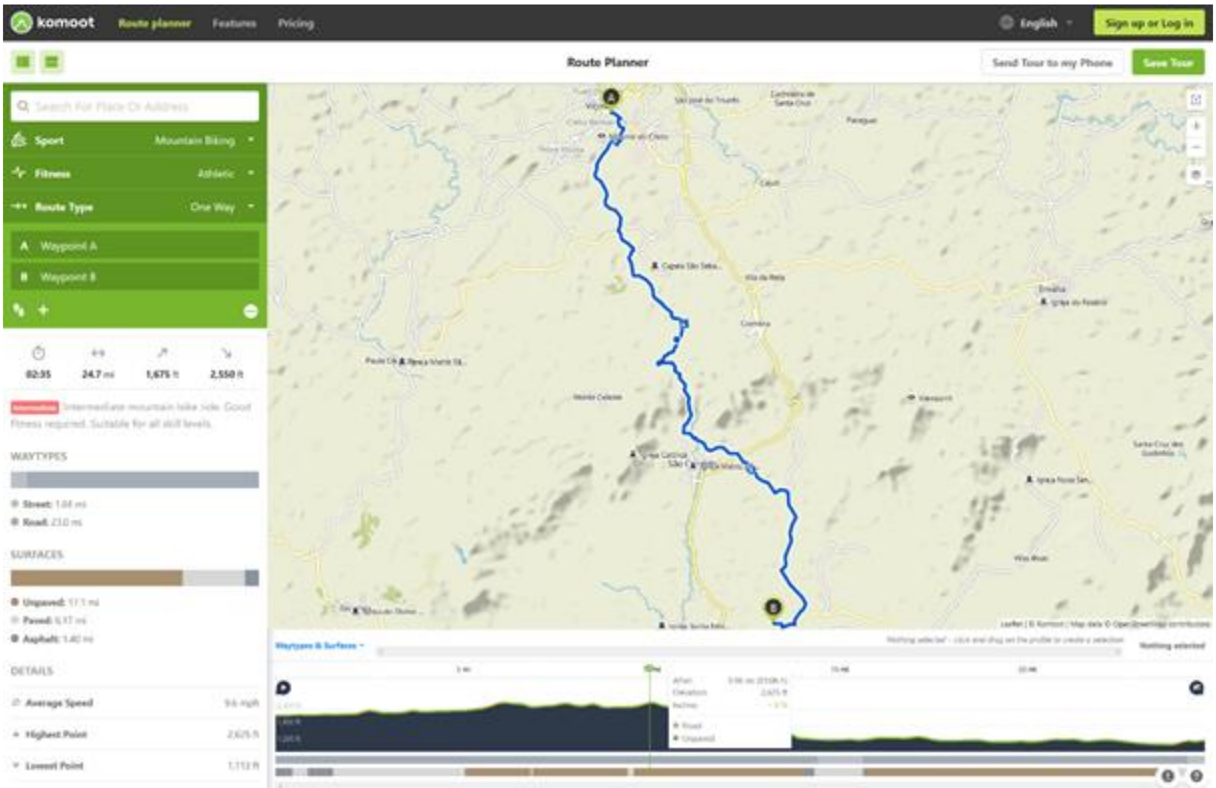


Source: Komoot's website allows for planning a route (a "Tour"') and gives detailed information such as the percentages and absolute distances of the tour that is asphalt, paved and unpaved surfaces, the percentage and absolutes distances of roads and streets and also the uphill and downhill distances. Still in the route planner, it is possible to choose a starting point, a destination and as many waypoints as wished, and it creates a route that goes through all those waypoints. Depending on the type of activity chosen, the calculated route might change (e.g.: a road cycling tour cannot go through

[9] https://www.komoot.com/

unpaved roads). It is also possible to choose a level of fitness between Couch Potato and Pro, so the planner is able to give the tour a rating between easy, intermediate and difficult and an average time to complete.

Figure 7. Komoot's route planner.



Source: Komoot's website

Besides route planning, Komoot also has editors who compile recommendation lists, multi day route compilations and, in some places, sponsored route lists: in Berlin, Germany, the public transport company S-Bahn has lists for routes that are accessible via public transport.

Komoot uses GraphHopper to help planning the routes and presents the data on OpenStreetMap.

2.6. Linear Programming

Linear Programming (LP) is a mathematical technique for finding optimal solutions to problems that can be expressed using linear equations and inequalities. Most complex problems in the real world cannot be expressed in linear equations, but if one can, the optimal solution to this problem can be found using this method.

If a LP problem has at least one of the variables restricted to integer values, it is an Integer Linear Programming (ILP) problem, although this term is used interchangeably with Mixed-Integer Programming, with Pure Integer Programming being used for problems in which all the variables are restricted to integer values.

This work uses ILP to model the problem in question, with a few simplifications, so it can be expressed as a linear problem. For example, when a user specifies a distance for the route, given that it is very unlikely that a route with the exact specified distance exists, a tolerance value was added so that any distance above or below the specified one by that percentage still fits the resolution of the problem. A measure of quality that penalizes any amount below or above that range can be modeled using linear functions.

2.7. Tools used in this work

One of the takeaways from analyzing the most popular tools used in cycling is that none of them are able to generate a route only based on user preferences. While Wikiloc and Trailforks have tools to filter uploaded routes according to the preferences of the user, if no routes are available in a region then there is no alternative, except to look for routes in other websites. In Komoot, while the route planner shows the way type, the surface and has the option to add a fitness level, if the user does not know the region well, it is very hard to create a route. The analysis of these tools shows the importance of this work.

Of all these tools, Strava is the most relevant as a data source for this work. With the incentive of gamification adding competitiveness to virtual segments, Strava segments are the most prevalent crowdsourced bike routing data. Besides being mostly everywhere, Strava segments already contain an accurate distance and elevation data, only missing the road surface data among the data we need. Road surface data of all streets is freely available in OSM, along with the location of POIs. Through GraphHopper it is possible to link Strava segments to each other, adding length, elevation and terrain data to these links and also to add terrain data to the segments, downloading the data from OSM.

Chapter III: Automatic Personalized Route Generation

As seen in previous chapters, bike tools usually help cyclists find a new route by providing them with a database of trails made by users or by presenting a route planner, so that users are able to create their own route either starting from another trail or from scratch. Both methods, either the database of trails or the route planner, have their advantages and problems.

A collaborated database of trails has the advantage of being curated, as one is able to assume that better trails have better reviews or more rides, is mostly precise in its data and, with a big enough database, it is possible to find trails better matching to the cyclist's preference. On the other hand, trails with more rides having better reviews could also imply that that these trails are located in places accessible to more people, creating a bias for more populated places. This method lacks in places where cyclists are slim to none, even if the place has plenty of dirt roads suitable for mountain biking. A route planner, on the other hand, while more malleable, leaves most of the work to the user, who might not know the area. And except for Komoot's route planner, current route planners offerings do not give a rating for the difficulty of a route. Both problems can be summarized as not finding enough data in a single source and can be solved, or at least attenuated, by using multiple sources from which to gather data.

This project aims to simplify the decision making around choosing a route for a cycling activity, by generating a route from a database of segments and a collaboratively sourced map tool. This route is created by taking into consideration the cyclist's preferences in distance, terrain difficulty and climb difficulty.

A segment is a slice of one or more roads which is significant to cyclists in some way: it might be a place with heavy cyclist traffic, a place where it is easy to gain speed, somewhere hard to travel through or just someone's way to work. Strava records these segments and makes them available in their platform, which can be accessed via an API. Because the API only gives 10 segments for a given search boundary, regardless of the size of the bounding box, we used a recursive square algorithm to get most of the segments in a given area. Segments in Strava are recorded with a name, distance, elevation and elevation grade, climb category, starting and ending locations in the format of a (latitude, longitude) coordinate, and points that form the segment among other attributes.

Using the recursive square method, we download all the segments in a region. The Strava API allows for 600 requests every 15 minutes and a total of 30000 per day, which is usually more than enough for our objectives, unless we take large areas, such as entire states or even countries, but then the processing time for the number of segments would get prohibitive.

To solve the problem we need the distances, elevations and terrain types between all segments and POIs and the same from the starting point of the ride to the start of every segment and POI. Using GraphHopper we calculate the distances between the ending point of one segment and starting point of the other, and it also gives us the elevation difference between a pair of segments and the types of surfaces in this stretch. With the distance and elevation difference we are able to calculate the elevation grade and then the climb category, using the same method Strava does (ROSIE, 2013):

$$N = length \cdot grade \tag{3.1}$$

And then, if *N* is greater than 8000, it is a categorized climb, which we then classify as:

- Cat 1 > 8000;
- Cat 2 > 16000;
- Cat 3 > 32000;
- Cat 4 > 64000;
- Cat 5 or HC (Hors Category) > 80000.

For Strava, the categorization above is ranked in the opposite direction, which we do not follow to remain consistent in 1 being the easiest and 5 being the hardest. This method is similar to the method used by Union Cycliste Internationale (UCI), but less subjective, since the UCI method may classify a segment as a different category depending on its position in a race, for example, a hard climb at the end of the race might be classified as a category 5, when it would be a category 4 if it was at the start of the race. Although the GraphHopper API has a matrix feature that is able to calculate the distance from one group of locations to another, it is very limited in the amount of points it can work with and does not provide road surface data, so we had to use the standard directions API, which provides the route from only one point to another.

Without information about trail width and obstacles in the trail, such as loose rocks, boulders and tree trunks, it is tricky to correctly classify trail difficulty, so based only on terrain surface we used the International Mountain Biking Association's (IMBA) rating[10], to create our own method, going from 1 to 5 in equal steps. Since GraphHopper does not provide the surface smoothness values from OpenStreetMap, we used the equivalent surface type matching IMBA's rating Thread surface. Since we already classified trail elevation using Strava's method, which in itself is based on UCI's method, we ignored IMBA's trail grade rating in the rating of the terrain.

Any paved surface is classified as 1 in level of difficulty. Paved surfaces are: asphalt, concrete:{lanes, plates}, paving_stones, sett, unhewn_cobblestone, cobblestone, metal, wood.

Types of unpaved surface are: compacted, fine_gravel, gravel, pebblestone, dirt, earth (same as dirt), grass, grass_paver, ground, mud (wet ground), sand, woodchips, snow, ice, salt.

Sand and snow surfaces are more dependent of equipment than technique, since very thick wheels are needed not to sink in it, so we try to avoid them. Also, more often than not, riding in the sand is not a pleasurable experience, because of how hard it is to move and, during the winter when there is snow, most avid cyclists simply stay home and ride with the aid of a trainer, which is an equipment that turns the bike into a stationary bicycle. The other surfaces are classified as such, ordered by difficulty:

1. All paved surfaces, equivalent to Hardened or surfaced Thread Surface in IMBA's rating;
2. fine_gravel, grass, grass_paver, compacted, pebblestone, equivalent to Firm and stable;
3. gravel, woodchips, equivalent to Mostly stable with some variability;
4. dirt, earth, ground, equivalent to Widely variable;
5. mud, equivalent to Widely variable and unpredictable;
6. sand, snow, as explained above, are avoided.

In segments with more than one type of surface, if at least one of the multiple surfaces is of paved type, the segment is considered of terrain difficulty 1. In the

[10] https://www.imba.com/resource/trail-difficulty-rating-system

absence of a paved surface, the difficulty considered in an average of the difficulty of the surfaces contained in the segment.

After setting a desired terrain, elevation, and distance, the user needs to set a starting point for this activity, inside the predetermined region. Once all these are set, the problem is to find a route that fits all these numbers. We start by finding the distance, elevation and terrain from the starting point to each of the segments for the region, the same data we already have for all the segments and the ways between them.

A tolerance for the distance was added because it is highly unlikely that a route exists with the exact specified distance. The same applies for the other attributes, terrain and elevation in a slightly different way, as it is very improbable that a route exists with only the same terrain or only the same elevation. Taking this into consideration a penalty was added for going above or below the specified values, in a way that, if a route is above or below any attributes, or attribute ± tolerance in the case of distance, it is penalized. As this work is focused on MTB it is not interesting to add paved streets to the route, so an extra penalty was added for the case when the street is paved. The total value of the penalties must be minimized, resulting in a route as close as possible to the preferences of the user.

3.1. ILP model

The problem is mathematically modeled using an ILP (Integer Linear Programming) formulation. The input data is divided in (i) static data, which is previously collected from OSM and Strava (for example, elevation and type of terrain of various segments), (ii) user defined data, which is informed by the user on each query (for example, starting location, desired elevation and other preferences), and (iii) user generated data, which is generated by the system based on the user's data (for example, distance from the origin point to segments of interest collected from maps).

In this model, POIs are treated as segments with a distance of 0km. Given that some POIs are features with an area (lakes and rivers), as opposed to single coordinates (viewpoints), the distance between POIs and other entities are calculated from the closest points between both entities.

When we add ride distance as a user input of the problem, we have to consider that it is very improbable that a trail exists with that exact distance, so we added a tolerance as well, meaning that the ride distance can go above or below the input by a percentage determined by the value of this tolerance.

The model penalizes terrain difficulty 1 (paved terrains) as the work is focused on MTB, which is done off-road. Elevation and terrain difficulty are set as maximum values, which means any amount above it is penalized as a cyclist might not be able to complete a ride that is too difficult. For distance, both values above and below the target value, plus a tolerance, as a percentage of the distance, are penalized, because a too short route might make a ride boring and the cyclist might not be able to complete a route that is too long. The tolerance value for the distance is also a user input, because while a tolerance of 10\% is acceptable for a 30km ride, it might be too much for 100km.

Input (user) data:

- *D*: the target distance for the tour;
- *T*: the target terrain difficulty for the tour;
- *E*: the target elevation for the tour;

Static (map) data:

- *S*: the set with all the segments available to make the tour;
- *K*: an index set for the order of the segments in the tour;
- x_s^k: 1, if segment s is the k-th segment on the route; 0, otherwise;
- $y_{ss'}$: 1, if segment s is followed by s' on the route, i.e. if user goes from s immediately to s'.

Additionally, we have variables to count the penalties:

- es_s^+: 1, if segment s is used and is above the desired elevation *E*;
- es_s^-: 1, if segment s is used and is below the desired elevation *E*;
- $ep_{ss'}^+$: 1, if segment s is followed by s' on the route, and the path *(s,s')* is above the desired elevation *E*;
- $ep_{ss'}^-$: 1, if segment s is followed by s' on the route, and the path *(s,s')* is below the desired elevation *E*;

- ts_s^+, ts_s^-, $tp_{ss'}^+$, $tp_{ss'}^-$: similar for the desired terrain difficulty
- ts_s^1, $tp_{ss'}^1$: similar for paved terrain (terrain type difficulty = 1).

The mathematical model describing the problem in details is as follows:

$$\min \lambda_1 P_{d^+} + \lambda_2 P_{d^-} + \lambda_3 P_{e^+} + \lambda_4 P_{e^-} + \lambda_5 P_{t^+} + \lambda_6 P_{t^-} + \lambda_7 P_{t^1} \tag{3.2}$$

$$\sum_{s\in\mathcal{S}} y_{0s} = 1 \tag{3.3}$$

$$\sum_{s'\in\mathcal{S}_0} y_{s's} = \sum_{s'\in\mathcal{S}_0} y_{ss'}, \qquad \forall s \in \mathcal{S} \tag{3.4}$$

$$\sum_{s\in\mathcal{S}} y_{s0} = 1 \tag{3.5}$$

$$x_s^1 = y_{0s}, \qquad \forall s \in \mathcal{S} \tag{3.6}$$

$$x_s^{k-1} = 1 \wedge y_{ss'} = 1 \implies x_{s'}^k = 1, \qquad \forall s, s' \in \mathcal{S}, k \in \mathcal{K}\backslash 1 \tag{3.7}$$

$$\sum_{s'\in\mathcal{S}_0} y_{ss'} = \sum_{k\in\mathcal{K}} x_s^k, \qquad \forall s \in \mathcal{S} \tag{3.8}$$

$$x_s^{k-1} = 1 \wedge y_{s0} = 1 \implies \sum_{s'\in\mathcal{S}} x_{s'}^k = 0, \qquad \forall s \in \mathcal{S}, k \in \mathcal{K}\backslash 1 \tag{3.9}$$

$$\sum_{s\in\mathcal{S}} x_s^k \leq \sum_{s\in\mathcal{S}} x_s^{k-1}, \qquad \forall s \in \mathcal{S}, k \in \mathcal{K}\backslash 1 \tag{3.10}$$

$$\sum_{s\in\mathcal{S}} \sum_{k\in\mathcal{K}} d_s x_s^k + \sum_{s\in\mathcal{S}_0} \sum_{s'\in\mathcal{S}_0} d_{ss'} y_{ss'} \leq (1+\alpha)D + P_{d^+} \tag{3.11}$$

$$\sum_{s\in\mathcal{S}} \sum_{k\in\mathcal{K}} d_s x_s^k + \sum_{s\in\mathcal{S}_0} \sum_{s'\in\mathcal{S}_0} d_{ss'} y_{ss'} \geq (1-\alpha)D - P_{d^-} \tag{3.12}$$

$$e_s > E \wedge \sum_{k\in\mathcal{K}} x_s^k = 1 \implies es_s^+ = 1, \qquad \forall s \in \mathcal{S} \tag{3.13}$$

$$e_s < E \wedge \sum_{k\in\mathcal{K}} x_s^k = 1 \implies es_s^- = 1, \qquad \forall s \in \mathcal{S} \tag{3.14}$$

$$e_{ss'} > E \wedge y_{ss'} = 1 \implies ep_{ss'}^+ = 1, \qquad \forall s, s' \in \mathcal{S} \tag{3.15}$$

$$e_{ss'} < E \wedge y_{ss'} = 1 \implies ep_{ss'}^- = 1, \qquad \forall s, s' \in \mathcal{S} \tag{3.16}$$

$$\sum_{s\in\mathcal{S}} d_s es_s^+ + \sum_{s\in\mathcal{S}_0} \sum_{s'\in\mathcal{S}_0} d_{ss'} ep_{ss'}^+ = P_{e^+} \tag{3.17}$$

$$\sum_{s\in\mathcal{S}} d_s es_s^- + \sum_{s\in\mathcal{S}_0} \sum_{s'\in\mathcal{S}_0} d_{ss'} ep_{ss'}^- = P_{e^-} \tag{3.18}$$

$$t_s > T \wedge \sum_{k\in\mathcal{K}} x_s^k = 1 \implies ts_s^+ = 1, \qquad \forall s \in \mathcal{S} \quad (3.19)$$

$$t_s < T \wedge \sum_{k\in\mathcal{K}} x_s^k = 1 \implies ts_s^- = 1, \qquad \forall s \in \mathcal{S} \quad (3.20)$$

$$t_s = 1 \wedge \sum_{k\in\mathcal{K}} x_s^k = 1 \implies ts_s^1 = 1, \qquad \forall s \in \mathcal{S} \quad (3.21)$$

$$t_{ss'} > T \wedge y_{ss'} = 1 \implies tp_{ss'}^+ = 1, \qquad \forall s, s' \in \mathcal{S} \quad (3.22)$$

$$t_{ss'} < T \wedge y_{ss'} = 1 \implies tp_{ss'}^- = 1, \qquad \forall s, s' \in \mathcal{S} \quad (3.23)$$

$$t_{ss'} = 1 \wedge y_{ss'} = 1 \implies tp_{ss'}^1 = 1, \qquad \forall s, s' \in \mathcal{S} \quad (3.24)$$

$$\sum_{s\in\mathcal{S}} d_s ts_s^+ + \sum_{s\in\mathcal{S}_0} \sum_{s'\in\mathcal{S}_0} d_{ss'} tp_{ss'}^+ = P_{t^+} \quad (3.25)$$

$$\sum_{s\in\mathcal{S}} d_s ts_s^- + \sum_{s\in\mathcal{S}_0} \sum_{s'\in\mathcal{S}_0} d_{ss'} tp_{ss'}^- = P_{t^-} \quad (3.26)$$

$$\sum_{s\in\mathcal{S}} d_s ts_s^1 + \sum_{s\in\mathcal{S}_0} \sum_{s'\in\mathcal{S}_0} d_{ss'} tp_{ss'}^1 = P_{t^1} \quad (3.27)$$

The objective is to generate a route as close to the user's preference as possible. The quality of a route is measured by the deviation of the desired distance, elevation and terrain difficulty, which are inputs by the user. Function (3.2) minimizes a weighted penalty for each preference, which are penalized if going above or below the user defined threshold.

Constraints (3.3)-(3.5) control the flow of the route: the route must contain a path from the user's origin to a segment (3.3), and for each segment that the user arrives, the route must include a path to another segment (3.4) or go back to the user's origin point (3.5).

Constraints (3.6)-(3.10) control the sequence of segments visited. By (3.6), the first segment visited is the one the biker goes from its origin point (when $y_{0s} = 1$, $x_s^1 = 1$). After that (3.7), if segment *s* is the *(k-1)*-th visited on the route ($x_s^{k-1} = 1$) and the biker goes from *s* to s' ($y_{ss'} = 1$), then segment s' is the next one to be visited ($x_{s'}^k = 1$). Moreover (3.8) if the biker takes a path from segment *s*, then segment *s* is visited (only once). However (3.9), when the biker goes back to origin, there is no segment visited next ($x_{s'}^k = 0$ for all *s'*). Moreover, constraint (3.10) guarantees that the segments numbering will not be skipped, i.e., if there is no (*k-1*)-th segment, there should not be a *k*-th segment.

The remaining constraints are used to evaluate the route. The left side of constraints (3.11) and (3.12) calculate the total distance of the route, which is the sum of the distances of the segments visited and the paths connecting them. The distance must meet the desired distance *D* set by the user, with a tolerance of α% above and below, or else it will be penalized in objective (3.2). The penalty is due to variables P_{d^+} and P_{d^-}, which are set to the distance above and below, whatever occurs. The desired elevation and terrain difficulty are accounting in a different way because the user set those parameters to the maximum desired level, not the total. By constraints (3.13)-(3.16), binaries *es* and *ep* are set for every used segment or path between them that is above or below the desired elevation. Then, (3.17) and (3.18) calculate P_e^+ and P_e^-, the total distance of the route that is above and below the desired elevation, which are then penalized in the objective function (3.2). The following set of constraints, (3.19)-(3.27) do the same for the terrain level that is above and below the desired level, and additionally for the paved terrains (level 1).

3.2. Post generation evaluation

After running the model through CPLEX[11], a popular commercial Linear Programming model solver, we apply another metric to the three first generated routes fitting the preferences of the cyclist. Inspired by the roundness metric in (STROOBANT et al., 2018), which uses roundness as a user input data, we created a roundness metric similar in concept to the one used in the paper. We consider that, in a good tour, the ways to the middle point of the tour and from there to its starting point are not similar, otherwise the activity might get repetitive.

In the metric proposed by (STROOBANT et al., 2018), the ideal tour is a perfect circle, with roundness error equals to 0, and considering P the polygon containing all the *N* points $p_0 \rightarrow p_1 \rightarrow p_2 \rightarrow ... \rightarrow p_N \rightarrow p_0$ that make the tour and L the total length of the tour. This is done in a separate script and is not considered in the mathematical model of the problem because it would need the model to have all the geographical information of the route, which is a lot of data and would substantially increase its complexity while reducing its efficiency. Besides that, the calculation of the metric requires the use of non-linear equations and would not work on an ILP model.

[11] https://www.ibm.com/analytics/cplex-optimizer

$$C = \frac{1}{L}\int_{P} Q\, dQ \tag{3.28}$$

$$R = \frac{1}{L}\int_{P} d(Q,C) dQ \tag{3.29}$$

$$E = \frac{1}{L}\int_{P} |r - d(Q,P)| dQ \tag{3.30}$$

Considering P as the closed polygon containing all the *N* points that make the tour and *L* as the total length of the tour, we use (3.28) to calculate the geometrical center *C* of the route, a (latitude, longitude) pair that is the average point of all the points that make the route. Then, using (3.29) we set the radius *R* of this circle as the average value of the distance of every point of the route to *C*. Using (3.30) we calculate *E*, the value accumulated via integral of the module of the differences of the estimated radius *R* and the real distance of a point in the route, divided by *L*, meaning the smaller the error, the more similar to a circle a route is. For example, in a circle the distance between every point in the circumference and the center *C* would be the same as *R*, therefore, the error would be 0 or almost 0.

Although the metric is enough to compare the generated routes with similar distances between themselves, it fails when comparing generated routes with routes from other sources, like from Strava activities, since the number or points varies greatly between them. In generated routes it is safe to assume all routes in the same length have a similar number of points, since the same tool was used to create them, but trails obtained by downloading Strava activities in a GPX format have a different number of points.

Chapter IV: Tests and evaluation

In this chapter we present and analyze the results of this work. Using Strava and GraphHopper, we downloaded the data from two different regions in Minas Gerais, Brazil: Viçosa and Guiricema, and Reggio Emilia, in Italy. Those regions were chosen because they are well known by the the author, as they are close to UFV and Unimore, where part of this work was developed. Besides that, the chosen regions of Minas Gerais are places for which both Trailforks and Wikiloc, two of the most popular trail sharing services, do not have public trails, and the region of Reggio Emilia, in Italy, is a mostly asphalted one, without many dirt trails nearby. We set three different sets of preferences, one easy, one medium and one hard, described in Table 1 and compared the three best solutions found for the three regions.

Table 1 - User profiles set for testing.

	Distance ± Tolerance	Elevation	Terrain difficulty
Easy	25km ± 10%	1	2
Medium	50km ± 10%	3	3
Hard	100km ± 5%	5	5

Source: Author.

Two tests were ran for each set of preferences, each with different parameters: all of them use the same penalties for elevation, in which routes are lightly penalized for going below desired elevation and are not penalized for going above it, because most segments are flat, meaning that segments graded above 0 for elevation are very infrequent; for distance, routes in the second set of parameters are more heavily penalized for going below desired distance; as for terrain difficulty, routes are not penalized for going above desired terrain, because most segments are unpaved and segments graded above 2 for terrain difficulty are very infrequent, but routes are penalized in an either lighter or heavier way, with the results for both penalties presented. The penalties are described in Table 2 - Penalties. These parameters were found to yield the best results in a relatively short amount of time of 30 minutes, compared to results yielded by other sets of parameters. Note that tests that took longer than 30 minutes only mean that the excess time was used compiling the information and not actually finding the route, meaning that the time limit was still respected.

Table 2 - Penalties

		Set A	Set B
Distance	Above	2	2
	Below	4	6
Elevation	Above	0	0
	Below	0.5	0.5
Difficulty	Above	0	0
	Below	1	2
Asphalt	-	2	3

Source: Author.

For each test the roundest route in the result set, the one with the smallest roundness error, is plotted in the map. In the figures presented, the red marker is the center of the red circle, used to calculate the roundness coefficient; the purple dashed lines and markers are POIs; blue lines are asphalt; brown lines are unpaved; thin lines are links between segments; and thick lines are Strava segments.

It is important to note that, since the elevation rating also depends on the length of the road going upwards, most segments are rated 0 in elevation. Because of this, segments graded with a 5 in elevation are extremely rare and, unfortunately, there are none in the regions we selected.

4.1. Region of Viçosa

For the region of Viçosa we have 154 segments in total, from which 8 are POIs, consisting of viewpoints, lakes and rivers. Most of these segments are flat or downhill, graded as 0 in elevation, with only 12 segments graded as 1 and 6 segments graded as 2 in elevation. There are no segments graded 3 or above for elevation. The region only has terrain difficulties graded 1 and 2, with 84 segments graded as 1 and 70 segments graded as 2.

The results for the three profiles are described in Table 3.

Table 3 - Results for the Easy, Medium and Hard profiles for the region of Viçosa

	Time		Distance	Elevation	Terrain	Roundness error
Easy A	17:27:50	Route 1	22977.7	1	2	2.271×10^{-3}
		Route 2	23947.37	1	2	2.282×10^{-3}
		Route 3	23102.94	1	2	2.279×10^{-3}
Easy B	05:46:03	Route 1	23324.56	1	2	4.763×10^{-3}
		Route 2	20061.24	1	2	4.699×10^{-3}
		Route 3	21549.81	1	2	3.536×10^{-3}
Medium A	33:57:04	Route 1	44783.35	1	2	5.707×10^{-3}
		Route 2	44783.35	1	2	5.707×10^{-3}
		Route 3	37337.69	0	2	6.845×10^{-3}
Medium B	0.30213	Route 1	40735.32	0	2	5.441×10^{-3}
		Route 2	40421.56	0	2	5.024×10^{-3}
		Route 3	39529.72	0	2	5.047×10^{-3}
Hard A	33:57:65	Route 1	79629.26	1	2	2.043×10^{-2}
		Route 2	77389.91	1	2	1.909×10^{-2}
		Route 3	79554.95	1	2	1.789×10^{-2}
Hard B	33:34:57	Route 1	91423.6	1	2	2.513×10^{-2}
		Route 2	90954.86	1	2	2.505×10^{-2}
		Route 3	94257.35	0	2	1.080×10^{-2}

Source: Author

The results for the first user profile are described in Easy A and Easy B profiles in Table 3.

All of the routes in the Easy profile reached the desired elevation and terrain difficulty and routes 2B and 3B are slightly below the desired distance. The routes with with the lowest roundness error are plotted in Figure 8.

While route Easy 1A visits three POIs, two lakes and one viewpoint, Easy 3B visits four, the same two lakes and a viewpoint plus an additional viewpoint. Both of the routes present similar problems in which both of them visit paved roads that are untagged in the OSM data, that is, streets that have no specified road type, marked as "other" in the data and treated as unpaved, as urban regions with paved roads are usually better mapped than non-urban regions, seen in Figure 9. Again on both sets, on the leftmost part of the route, zoomed in on Figure 10 it is possible to see a segment drawn with two road surface types. This happens because of the way surface types are chosen, in which a segment is considered as unpaved only if the entirety of it is unpaved and is considered as paved otherwise, so while the segment is unpaved, the route from the end of the segment to the next one goes through paved terrain and is considered as paved. This happens because GraphHopper only gives us the surface types of a segment, without mentioning the parts of which type, and considering only the highest terrain difficulty would lead to many paved roads to be considered as unpaved, which is why this is made like this.

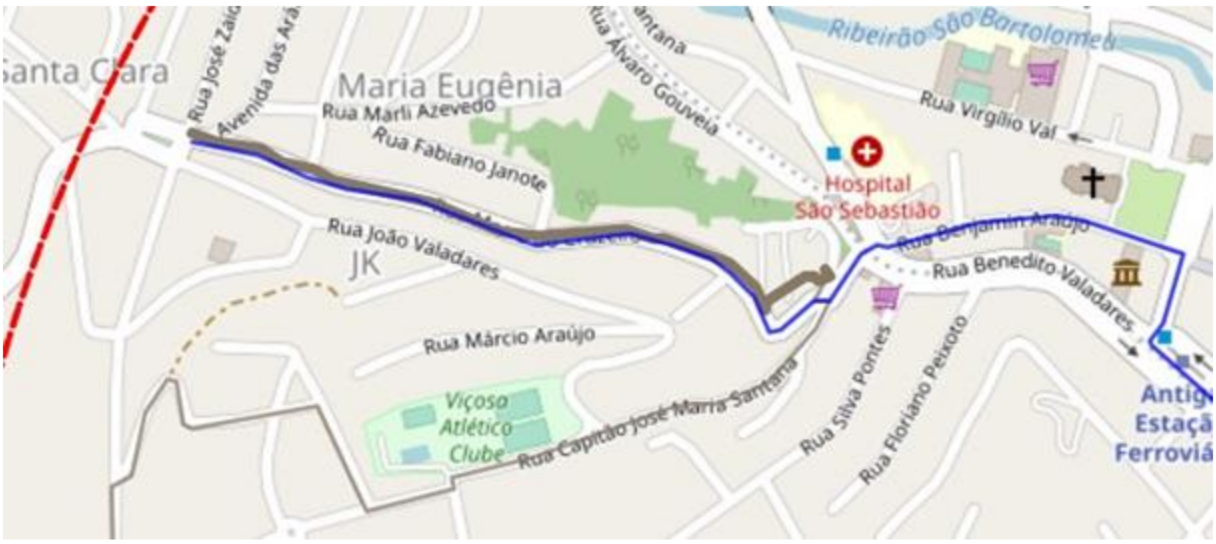
Figure 8. Viçosa's Easy routes with the smallest roundness error, each with a different set of parameters.

Source: Author.

Figure 9. Paved street described as unpaved in the region of Viçosa.

Source: Author.

Figure 10. Route drawn with two surface types.



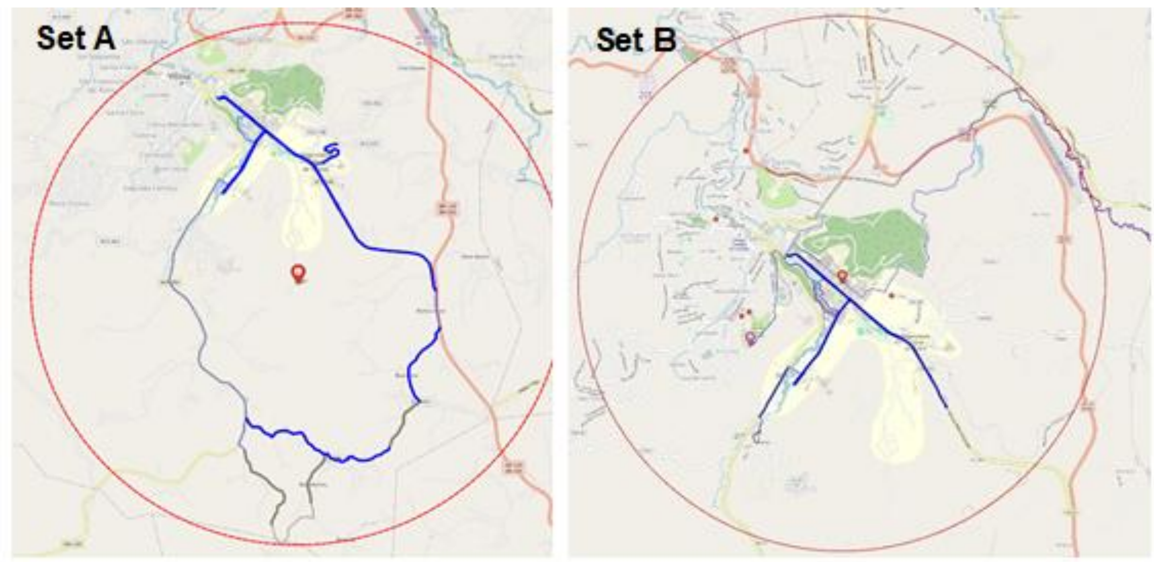
Source: Author.

The results for the second user profile are described in Table 3, in the Medium A and in the Medium B lines.

The routes with the lowest roundness error are plotted in Figure 11. Viçosa's Medium routes with the best roundness coefficient, each with a different set of parameters.. Because of the infrequency of segments with elevation graded as 3 and terrain difficulty graded as 3, all results are under the desired values. While route Medium 1A, which is the same as 1B, visits no POIs, route Medium 2B visits one river, one lake and one viewpoint, but at the cost of visiting less unpaved streets. On Medium 1A, despite the fact that the entirety of the route is represented as paved, a great part of it is made of unpaved roads, mainly in the bottom half of the route. This problem is represented in

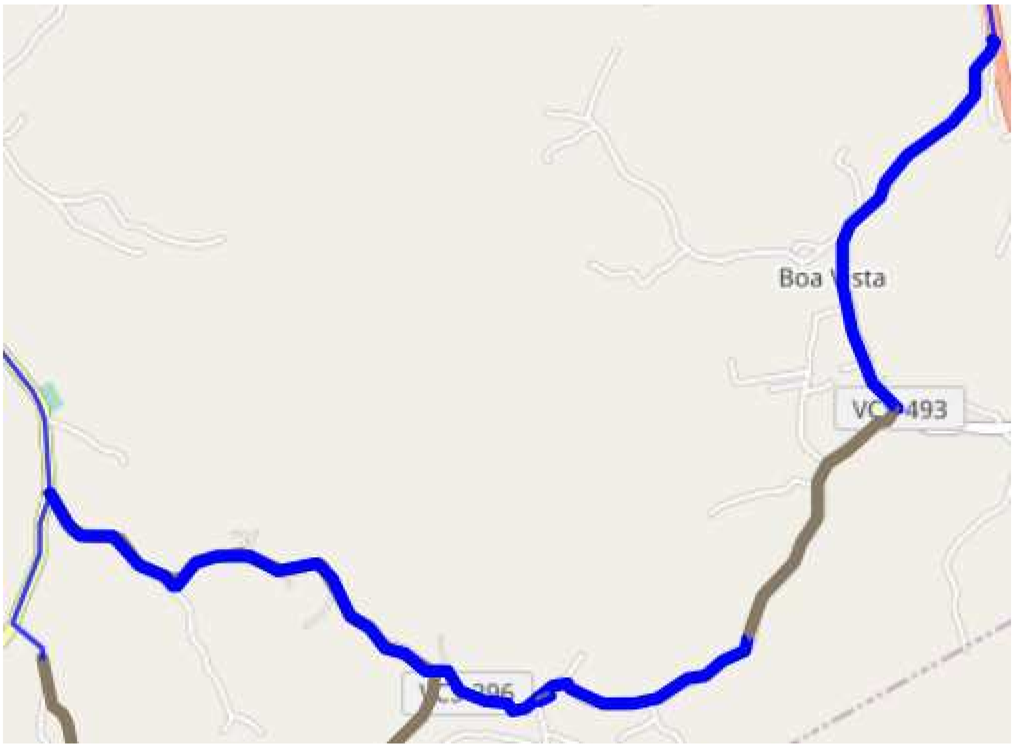
Figure 12, which shows streets represented as paved but that are actually unpaved.

Figure 11. Viçosa's Medium routes with the best roundness coefficient, each with a different set of parameters.



Source: Author.

Figure 12. Unpaved roads represented as paved in the region of Viçosa.

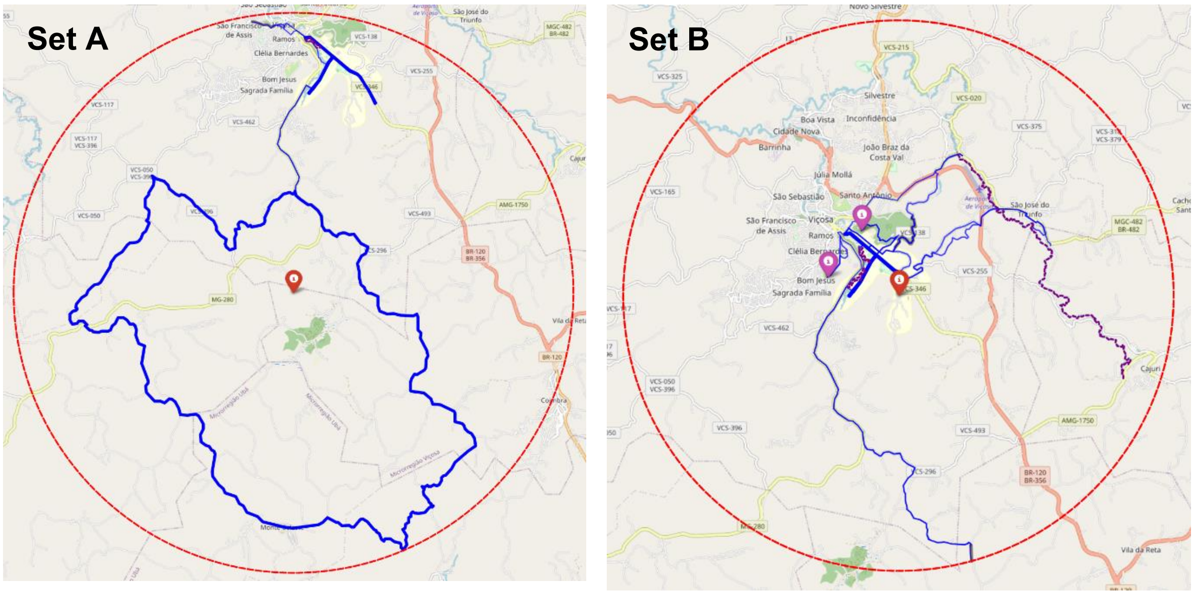


Source: Author.

The results for the third user profile, the hardest one, are described in Table 3, in the Hard A and Hard B lines.

The routes with the lowest roundness error are plotted in Figure 13. Because of the lack of segments with elevation graded as 5 and segments with terrain difficulty graded as 5, all alternatives are under the desired values. For distance, routes from Set A are far below the desired 100Kms and routes from Set B is also slightly below, despite the bigger penalty for it. These routes suffer from the same problems as the routes selected for the Medium profile, being considered as paved even if a great part of the route is in unpaved terrain.

Figure 13. Viçosa's Hard routes with the smallest roundness error, each with a different set of parameters.



Source: Author.

4.2. Region of Guiricema

For the region of Guiricema we have 189 segments in total, from which 15 are POIs, consisting of viewpoints, lakes and rivers. Most of these segments are either flat or downhill, with a 0 in elevation, with only 15 segments with elevation graded as 1, 11 segments with elevation graded as 2 and 3 segments with elevation graded as 3 which, in general, are very rare. There are no segments graded 4 or 5 for elevation. The region is more diverse in terrain difficulties, with 81 paved segments, 106 segments with difficulty graded as 2, 1 segment with difficulty graded as 3 and one segment with difficulty graded as 4. Although the region is more diverse, because the amount of segments would be over one thousand if downloaded in the same way they were for the region of Viçosa, they are also more spread out, in a way that segments are more sparse. Because segments are more sparse, it is more difficult to find routes fitting the desired distance value, especially for shorter distances, a problem which could be solved by downloading all segments, although that would have made the processing costs prohibitively expensive.

The results for the three profiles are described in Table 4.

Table 4 - Results for the Easy, Medium and Hard profiles for the region of Guiricema

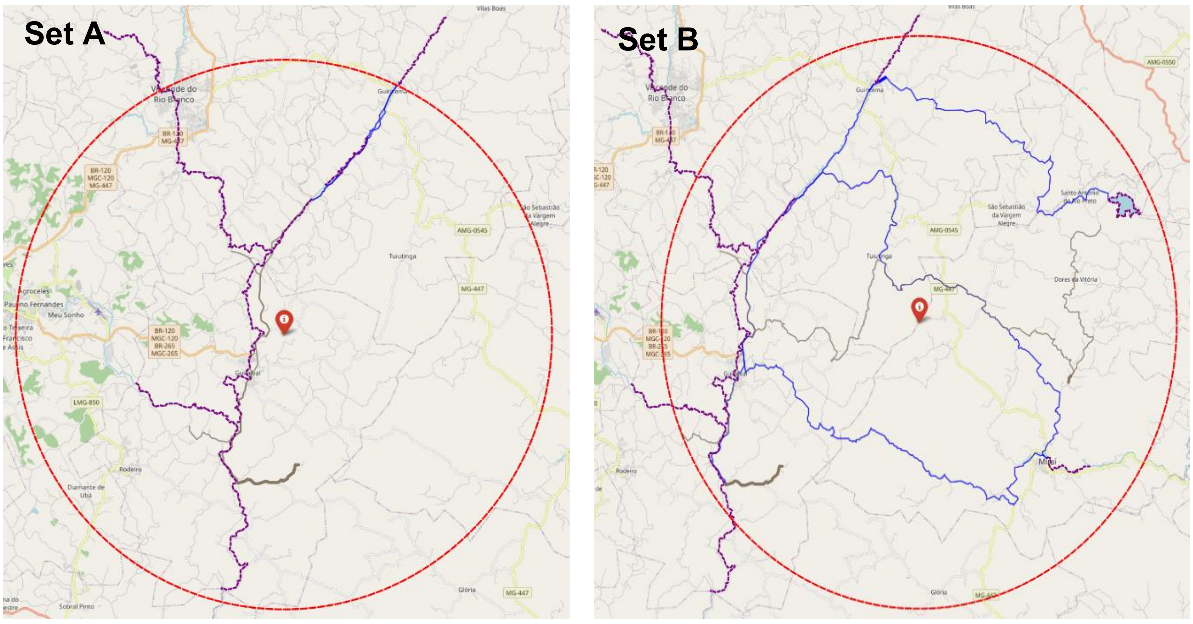
	Time		Distance	Elevation	Terrain	Roundness error
Easy A	11:50:17	Route 1	22086.72	0	2	3.031×10^{-2}
		Route 2	22201.53	2	2	2.562×10^{-2}
		Route 3	36404.49	2	4	2.826×10^{-2}
Easy B	0.526829	Route 1	26548.9	3	2	2.417×10^{-2}
		Route 2	26548.9	3	2	2.417×10^{-2}
		Route 3	26548.9	3	2	2.417×10^{-2}
Medium A	0.858044	Route 1	51169.35	3	3	1.601×10^{-2}
		Route 2	54044.03	3	3	1.594×10^{-2}
		Route 3	46403.25	0	2	2.102×10^{-2}
Medium B	24:59:86	Route 1	43295.32	3	3	1.601×10^{-2}
		Route 2	59123.57	3	3	1.580×10^{-2}
		Route 3	40027.2	3	3	1.720×10^{-2}
Hard A	35:04:61	Route 1	90845.93	0	2	2.229×10^{-2}
		Route 2	94628.56	0	2	2.388×10^{-2}
		Route 3	79087.37	3	2	2.087×10^{-2}
Hard B	31:48:91	Route 1	84438.43	3	2	2.567×10^{-2}
		Route 2	98795.01	2	2	2.518×10^{-2}
		Route 3	99878	3	2	2.518×10^{-2}

Source: Author.

The Easy A and Easy B lines in Table 4 describe the results for the first user profile for the region of Guiricema.

With the exception of Route 1A all of the routes have segments with elevation graded above the desired elevation, most with a 3 and Route 3A has terrain difficulty graded 4, which is far above the desired 2, and a distance 11km above the desired 25km. Because of how sparse the *harder* segments are to find, these were penalized more lightly, resulting in much harder routes in places where these segments are available and close to the starting point.

Figure 14. Guiricema’s Easy routes with the smallest roundness error, each with a different set of parameters.



Source: Author.

Figure 15 - Two visits to the same POI in the region of Guiricema.

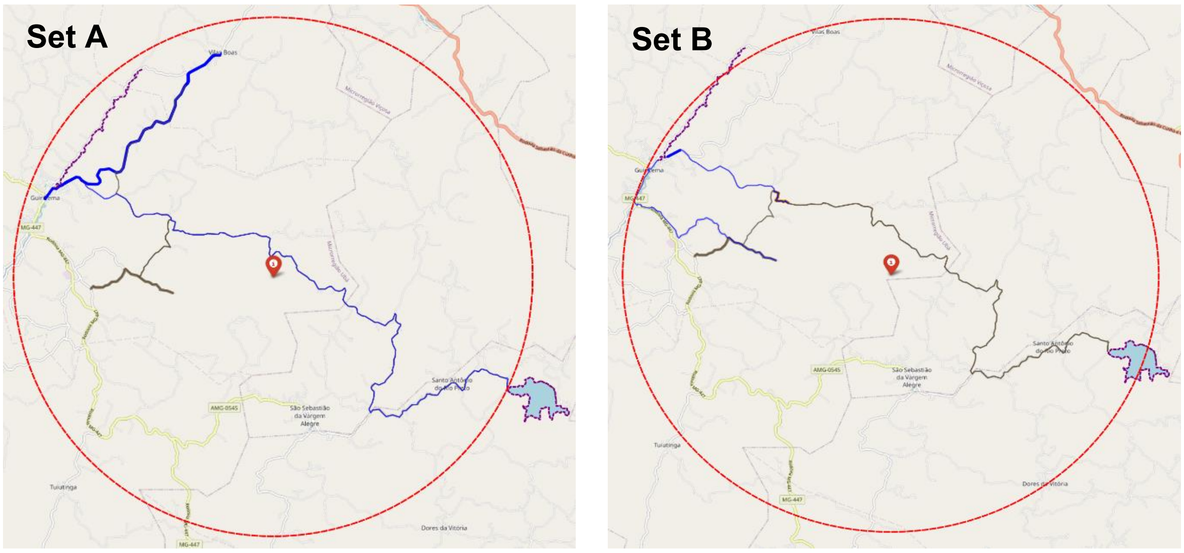
Source: Author.

In both of the routes depicted in Figure 14, the routes with the lowest roundness error, aside from the problems in the previous routes, such as unpaved routes drawn as paved and segments with two types of terrain, the routes also seem to visit the same

POIs twice which happens because, while the river is seen as a single entity by the viewer, it is represented as more than one element in OSM. Figure 15 shows a zoomed in section of Figure 14 A in which this problem happens, but the same also happens in Figure 14 B.

For the second user profile for the region of Guiricema, all the six routes are described in the Medium A and Medium B lines in Table 4.

Figure 16. Guiricema's Medium routes with the smallest roundness error, each with a different set of parameters.



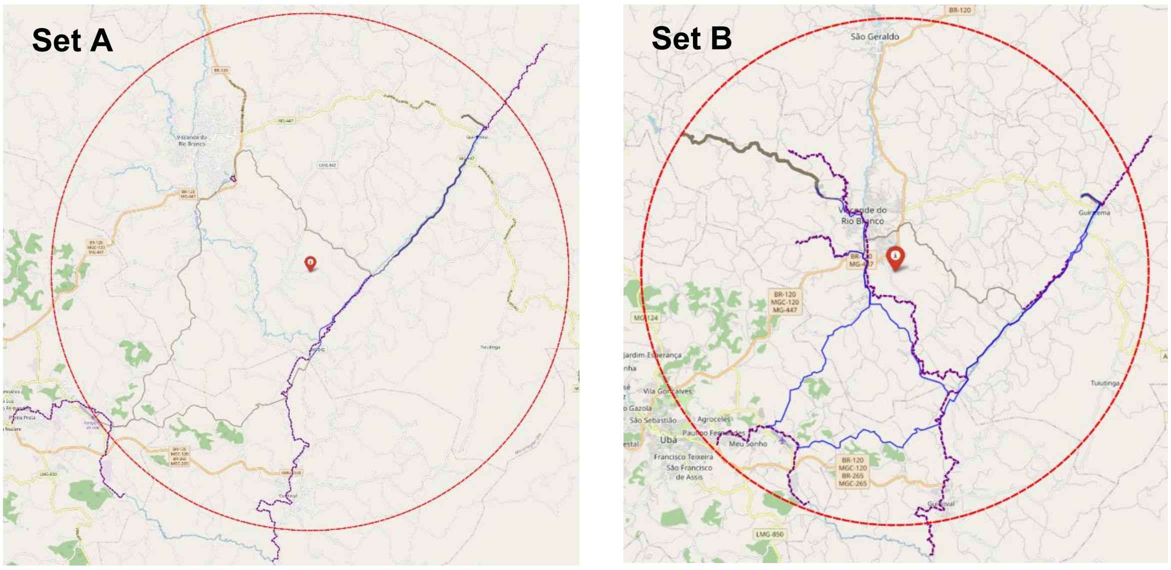
Source: Author.

Both routes, Medium 2A and Medium 2B, described respectively in Figure 16, are very similar but route B represents a road as unpaved as route A represents it as paved, because that segment of road in A has a small paved part. Most of the routes found by the mathematical model fit the parameters of the profile, except for 3A for terrain and elevation and 3B for distance.

For the third user profile for the region of Guiricema, we have the six routes described in Table 4, in the Hard A and Hard B profiles.

The routes with the lowest roundness error are plotted in Figure 17. Due to the absence of segments graded above 3 in elevation, the results did not reach the preferences for the profile, but despite that, it did well for this parameter. These routes show the same recurrent problems of the other profiles: paths drawn with two surface types; unpaved surfaces being considered as paved; paved surfaces considered as unpaved; and two visits to the same POIs. Despite these problems, most of the routes did well in the distance parameter.

Figure 17. Guiricema's Hard routes with the best roundness coefficient, each with a different set of parameters.



Source: Author.

4.3. Region of Reggio Emilia

The region of Reggio Emilia is in a more developed region, where is harder to find unpaved roads and road cycling is more popular. This region has more segments, with 329 of them, 154 being POIs, consisting of mostly drinking water fountains, which are not present in the other regions presented in this work, and also a bicycle repair station; other than that there are some viewpoints, lakes, rivers and waterfalls. Despite the increase in the amount of segments, most still are flat with only 4 with an elevation graded as 3, 23 graded with an elevation of 2, 34 graded with an elevation of 1 and the rest graded with a 0. In terms of terrain there are mostly segments graded with a 2, meaning unpaved, representing 196 of the segments, but as this is a mostly asphalted region this might be because of untagged ways; there are also 8 segments graded with a terrain difficulty of 3 and 6 segments graded with a terrain difficulty of 4, with the rest graded as 1. This region is also more sparse than the region of Viçosa, although it is not as sparse as the region of Guiricema, because otherwise there would be too many segments to process.

The results for the three profiles are described in Table 5.

Table 5 - Results for the Easy, Medium and Hard profiles for the region of Reggio Emilia

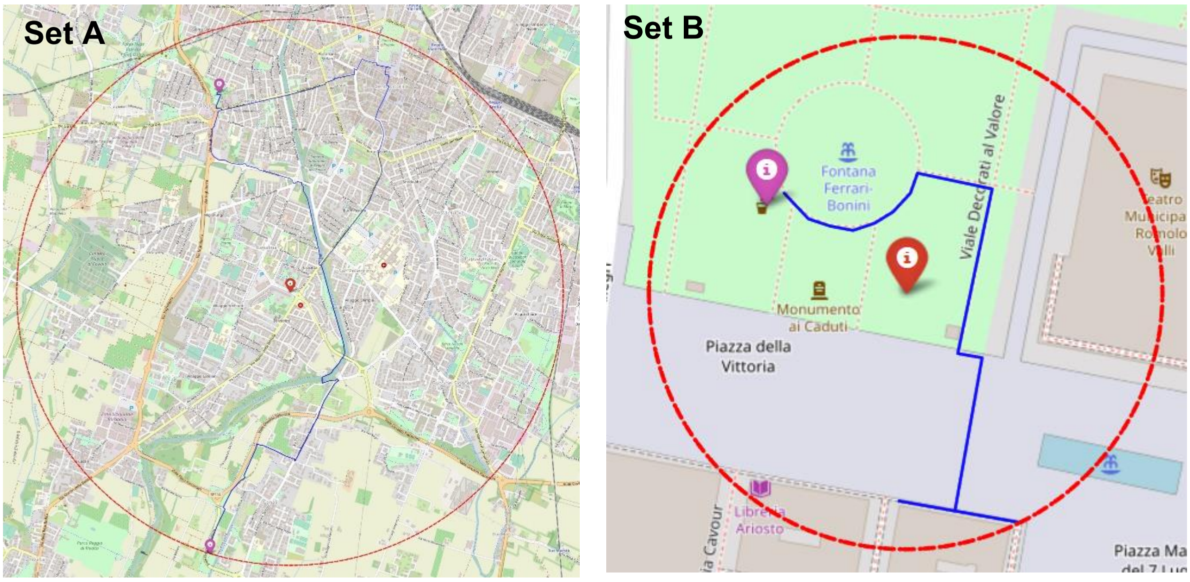
	Time		Distance	Elevation	Terrain	Roundness error
Easy A	11:50:17	Route 1	22226	0	2	1.724×10^{-2}
		Route 2	22396.72	2	2	4.607×10^{-3}
		Route 3	21020.13	2	4	4.561×10^{-3}
Easy B	0.526829	Route 1	23822.83	3	2	1.201×10^{-2}
		Route 2	2217.59	3	2	3.531×10^{-3}
		Route 3	459.71	3	2	1.148×10^{-4}
Medium A	0.858044	Route 1	44959.33	3	3	4.117×10^{-3}
		Route 2	43893.84	3	3	3.933×10^{-3}
		Route 3	43974.34	0	2	2.703×10^{-2}
Medium B	24:59:86	Route 1	44985.4	3	3	5.655×10^{-3}
		Route 2	44815.36	3	3	4.579×10^{-3}
		Route 3	44774.69	3	3	4.281×10^{-3}
Hard A	35:04:61	Route 1	94742.48	0	2	2.553×10^{-2}
		Route 2	94965.41	0	2	3.030×10^{-2}
		Route 3	94942.2	3	2	2.522×10^{-2}
Hard B	31:48:91	Route 1	91980.82	3	2	3.869×10^{-2}
		Route 2	94701.78	2	2	1.465×10^{-2}
		Route 3	91980.82	3	2	3.869×10^{-2}

Source: Author.

The Easy A and Easy B lines in Table 5 describes the results for the first user profile for the region of Reggio Emilia.

As there are nearly no unpaved streets within the distance of 25km from the start, neither of the depicted routes leave the city area and both routes visit viewpoints withing the city. Route Easy 3B, depicted in Figure 18 B is severely under the desired distance, as is the route with the second lowest roundness error, Easy 2B, but Route Easy 1B fits the distance preference, although with a much harder elevation, graded as 3, which might also be a problem in the data, as only either very long segments or *very* steep segments are graded higher when using Strava's algorithm and the route depicted is short and flat. In the A profile the route with the lowest roundness error is Easy 3A, depicted in Figure 18 A, which fits the distance preference but reached a slightly higher elevation, graded as a 2 and a much harder terrain difficulty, graded as a 4.

Figure 18. Reggio Emilia’s Easy routes with the smallest roundness error, each with a different set of parameters.

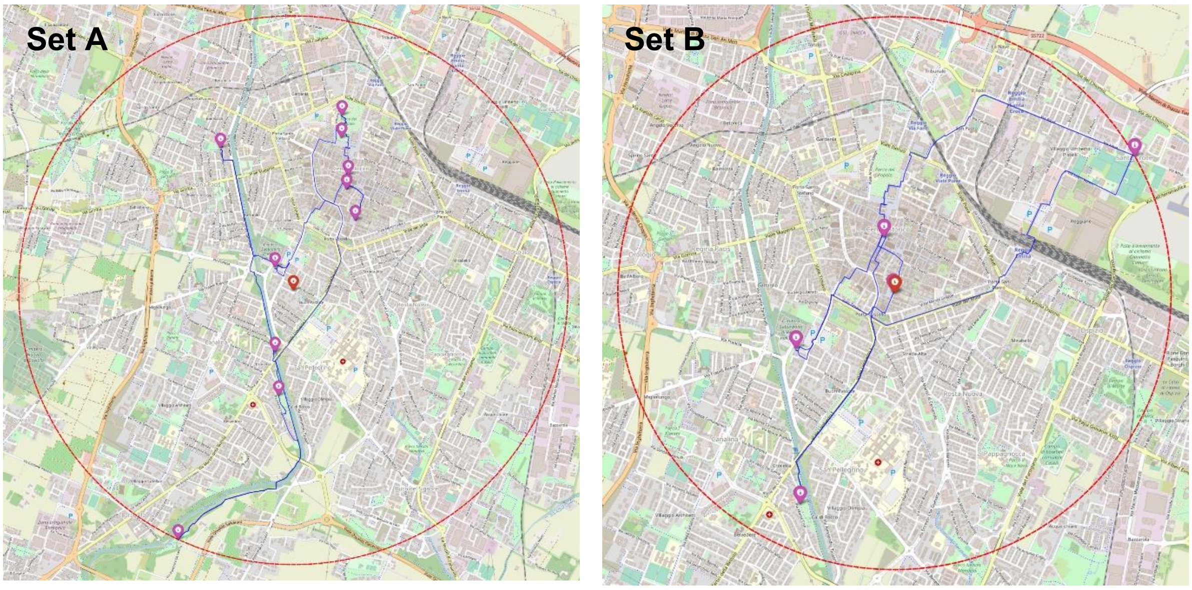


Source: Author.

The Medium user profile for the region of Reggio Emilia is described in lines Medium A and Medium B in Table 5.

Of all the generated routes for the Medium profile, only Medium 3A does not fit the set preferences for elevation and terrain difficulty, with all of them being very slightly under the distance preference, ranging from 45km to 55km.

Figure 19. Reggio Emilia's Medium routes with the smallest roundness error, each with a different set of parameters.



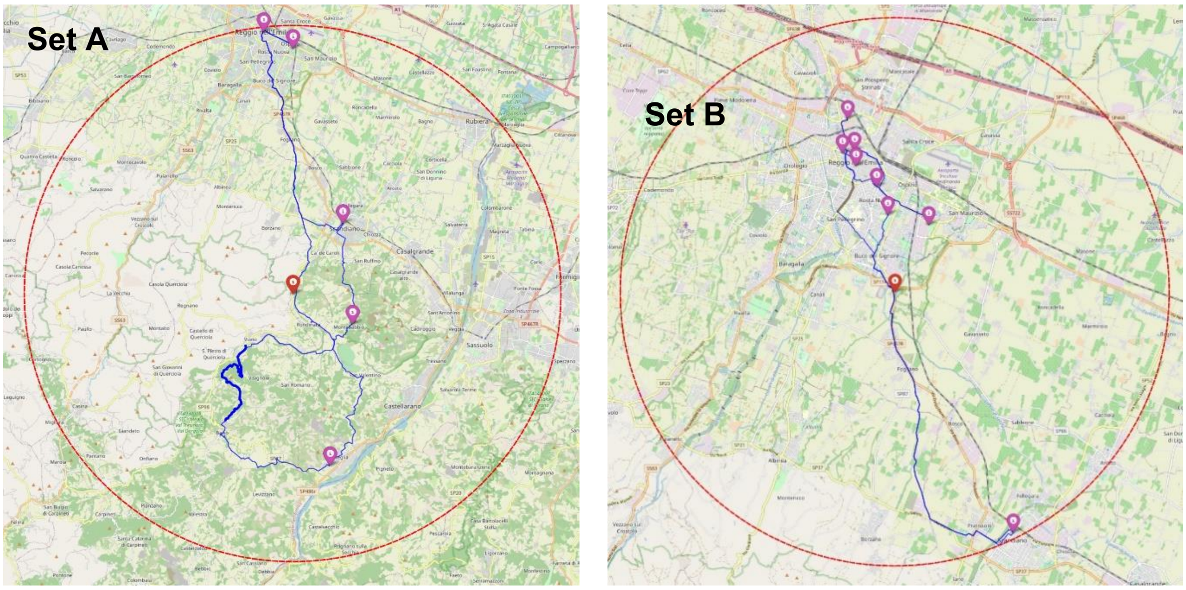
Source: Author.

Both of the routes with the lowest roundness error seem to not leave the urban area of the city, but route Medium 3A, shown in Figure 19 A, seems to be sticking to the riverside area, which is an unpaved area inside the city; it also visits 10 POIs, among water fountains and sightseeing spots in the city. Route Medium 3B, depicted in Figure 19 B, seems more paved and visits a smaller number of POIs, with only 5, which is still a lot compared to routes in other regions.

The Hard user profile for the region of Reggio Emilia is described in lines Hard A and Hard B in Table 5.

All of the generated routes for this profile are slightly under the preference for distance, which is from 95km to 105km, but since there are no segments with graded with a 5 in either elevation or terrain difficulty, all of the routes are under the preference for these parameters.

Figure 20. Reggio Emilia’s Hard routes with the smallest roundness error, each with a different set of parameters.



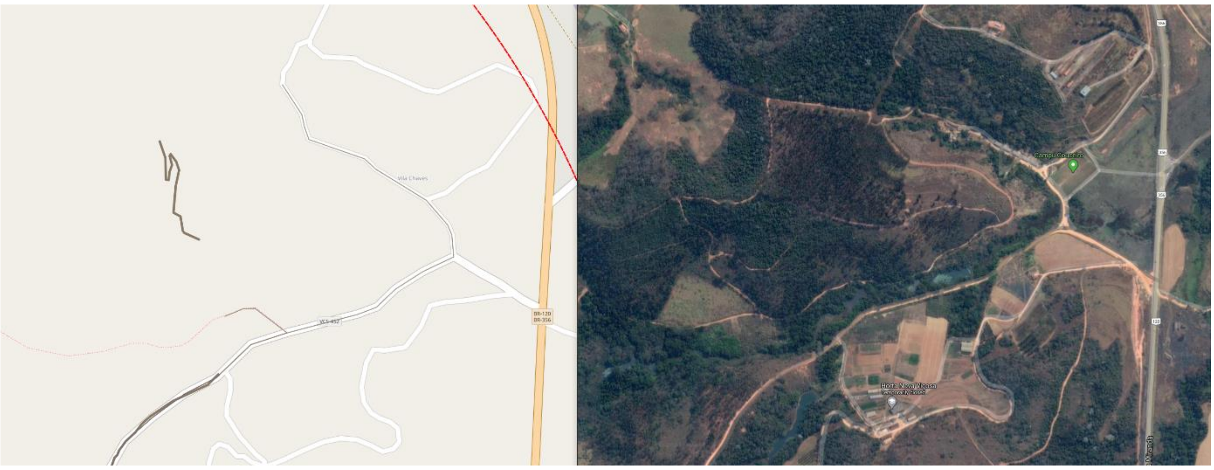
Source: Author.

Both routes with the lowest roundness error are graded with a 2 in terrain difficulty but the route generated with Set A of penalties has a higher elevation, graded as a 3, which is more desirable, compared to a grade 2 elevation for the route generated with Set B. Route Hard 3A, depicted in

Figure 20 A, veers to a less urban and more mountainous region, while visiting 5 POIs, including a water fountain in a city nearby to the city of Reggio Emilia, on the way. Route Hard 2B, depicted in B, while visiting more POIs, 8 of them, stays close to roads and urban areas and ends up being a round trip to a nearby city.

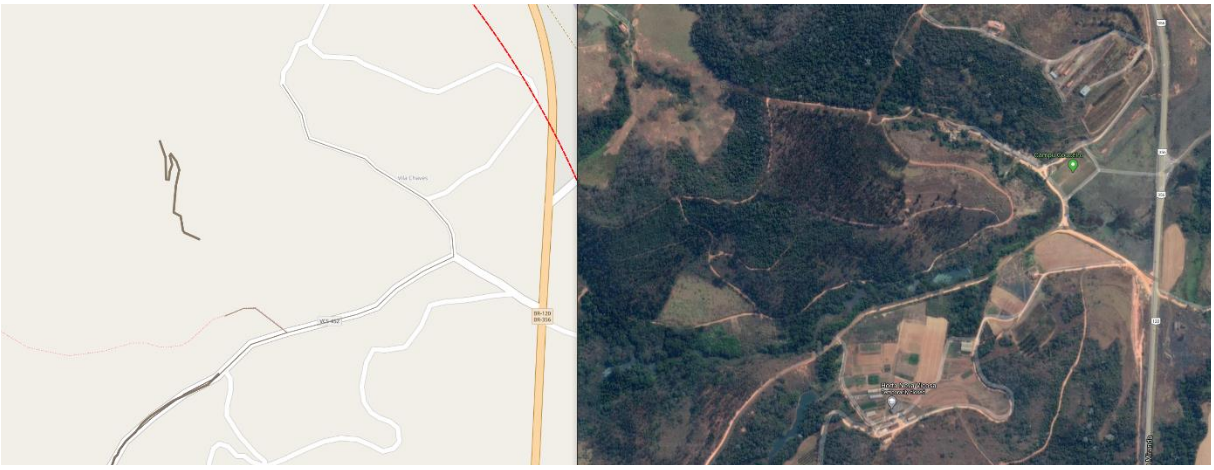
4.4. Other results

Despite not happening for any of the profiles, during the tests a few of the results showed some segments disconnected from the rest of the route, as shown in Figure 21. Unmapped roads in Viçosa.



. This happens because of unmapped roads in the route, meaning that there is no clear path between two segments.

Figure 21. Unmapped roads in Viçosa.



Source: Author.

This can be fixed either by adding the roads to OSM, which might take some time until GraphHopper updates its copy of the data, or by drawing a straight line to the next point when there is no available route.

None of the routes could reach the harder difficulties set in the profiles. This happened because the few harder segments are very rare and the set of segments is mostly populated by paved and *other* (untagged) segments, respectively difficulties 1 and 2, so this was expected.

Chapter V: Conclusions

With the dissemination of Voluntary Geographic Information, the improvement in its quality and the popularization of cycling as a means to become healthier and as a sport, it is possible and important to develop new tools to assist people in finding places to ride. The literature review in this work showed the importance of cycling for the individual health of the cyclist and for the collective financial benefits of society. We also reviewed some of the most popular tools used by cyclists to discover trails and plan rides and discussed their strengths and weaknesses.

This work proposed a method of generating routes using data collected from multiple sources of VGI: Strava segments and OpenStreetMap map data. With user inputted preferences, such as desired distance, elevation and terrain difficulty, it is possible to generate a route close to meeting user requirements, considering the available data. While Strava segments are populated by cyclists with activities in the area, OpenStreetMap data can be added and updated by anyone without needing any specific formation in the area, which allows us to bootstrap new routes in areas where cyclists have not added segments to yet.

We developed a model to generate routes based on user preferences, using segments collected from Strava and POIs collected from OSM. With the aid of GraphHopper to find paths between segments and POIs, an ILP model uses the data to find the routes that better fit the preferences of a user.

For the most part, generated routes managed to satisfy user requirements, considering the available data. In some parts, due to incomplete or mistagged OSM data, trails that were supposed to avoid paved roads still veered into asphalt. In some places the route seemed to end and continue ahead, disconnected from the rest of the trail, in places where roads are not mapped. This means that, while the mathematical model and evaluation worked, the data it worked on was sometimes incomplete — missing road surface data and sometimes missing roads that exist but are not yet mapped. The routes it generates are useful for people who do not know the area and want to explore. The tested routes were correct for the most part, but for experienced cyclists or people who already know the area, exploring by themselves might yield better results. Nonetheless, even with the huge advancements in VGI quality and better tools always coming along with it, wherever there is a group of cyclists, it remains hard to beat word of mouth in discovering new trails.

Despite the problems described previously, using a collaborated map tool such as OSM is an advantage of this project, because the data can be fixed or added wherever it is incomplete or missing. This way, although the results of this work did not always meet every requirement, they can be improved by user collaboration, without requiring changes to the project's implementation. That is, through the collaborative nature of OSM, the project can be improved as it is used.

In the future, work should be done towards better analyzing Points of Interest, differentiating them from segments and dividing them into categories more fitting to each cyclist's preferences. Another layer of VGI can be added too, by allowing users to evaluate the generated routes. As a small addition, instead of simply assuming untagged streets as unpaved, checking the road type (residential, tertiary road, path) could help in more correctly assuming surface type.

The elevation grading method used, which is the same one used by Strava, could have been adapted to better fit the reality of each region and of MTB, since many short but very steep hills are graded low because the distance of the segment has too much weight in the method.

A great improvement would be adding scenery as a user preference, using social network data to identify popular spots and roads, by checking the density of photos taken or posts written in an area, compared to its surroundings. Taking data from social networks might be considered by some a gray area of VGI, since the person whose data is being taken is not aware of it, even if the data was put there voluntarily.

In the optimization side of the project, work could be done to make it fast enough to run on mobile devices. Routes could be generated for some common parameters and starting points, aiming to reduce processing and waiting time. In the mathematical model, penalising segments differently depending on the objective would also help improving the results, e.g. segments graded 3 in elevation are penalised less than a segment graded as 2 if the desired elevation grade for the trail is 4.

BIBLIOGRAPHY

BERGMAN, C.; OKSANEN, J. Conflation of OpenStreetMap and mobile sports tracking data for automatic bicycle routing. **Transactions in GIS**, v. 20, n. 6, p. 848–868, 2016.

BROACH, J.; DILL, J.; GLIEBE, J. Where do cyclists ride? A route choice model developed with revealed preference GPS data. **Transportation Research Part A: Policy and Practice**, v. 46, n. 10, p. 1730–1740, 2012.

CARVALHO, M. L. DE; FREITAS, C. M. DE. Pedalando em busca de alternativas saudáveis e sustentáveis. **Ciência & Saúde Coletiva**, v. 17, n. 6, p. 1617–1628, 2012.

CLIFTON, K. J. et al. Examining consumer behavior and travel choices. 2013.

DE HARTOG, J. J. et al. Do the health benefits of cycling outweigh the risks? **Environmental health perspectives**, v. 118, n. 8, p. 1109–1116, 2010.

Environment: New policy package to clean up Europe's air. , [s.d.].

FIX, P.; LOOMIS, J. The economic benefits of mountain biking at one of its meccas: An application of the travel cost method to mountain biking in Moab, Utah. **Journal of Leisure Research**, v. 29, n. 3, p. 342–352, 1997.

FOODY, G. et al. Mapping and the citizen sensor. **Mapping and the Citizen Sensor**, p. 1–12, 2017.

GOODCHILD, M. F. Citizens as voluntary sensors: spatial data infrastructure in the world of Web 2.0. **International journal of spatial data infrastructures research**, v. 2, n. 2, p. 24–32, 2007.

GRANVILLE, S. et al. Sharing road space: Drivers and cyclists as equal road users. **Scottish Executive Central Research Unit, Edinburgh, Scotland**, 2001.

HRNCIR, J. et al. **Bicycle route planning with route choice preferences.** ECAI. **Anais**...2014

KOEMLE, D. B. A.; MORAWETZ, U. B. Improving mountain bike trails in Austria: An assessment of trail preferences and benefits from trail features using choice experiments. **Journal of outdoor recreation and tourism**, v. 15, p. 55–65, 2016.

KRISMER, N. et al. **Elevation Enabled Bicycle Router Supporting User-Profiles.** GvD. **Anais**...2016

MENGHINI, G. et al. Route choice of cyclists in Zurich. **Transportation research part A: policy and practice**, v. 44, n. 9, p. 754–765, 2010.

MOONEY, P.; MINGHINI, M.; OTHERS. A review of OpenStreetMap data. 2017.

MORRIS, E. A.; GUERRA, E. Mood and mode: does how we travel affect how we feel? **Transportation**, v. 42, n. 1, p. 25–43, 2015.

OKOLLOH, O. Ushahidi, or 'testimony': Web 2.0 tools for crowdsourcing crisis information. **Participatory learning and action**, v. 59, n. 1, p. 65–70, 2009.

PUCHER, J.; BUEHLER, R. Integrating bicycling and public transport in North America. **Journal of Public Transportation**, v. 12, n. 3, p. 5, 2009.

REDDY, S. et al. **Biketastic: sensing and mapping for better biking**. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. **Anais**...2010

ROSIE. **Climb Categorization**, 2013.

STROOBANT, P. et al. Generating constrained length personalized bicycle tours. **4OR**, v. 16, n. 4, p. 411–439, 2018.

TUCKER, B.; MANAUGH, K. Bicycle equity in Brazil: Access to safe cycling routes across neighborhoods in Rio de Janeiro and Curitiba. **International journal of sustainable transportation**, v. 12, n. 1, p. 29–38, 2018.

UNWIN, N. C. Promoting the public health benefits of cycling. **Public Health**, v. 109, n. 1, p. 41–46, 1995.

ZIMMERMANN, M.; MAI, T.; FREJINGER, E. Bike route choice modeling using GPS data without choice sets of paths. **Transportation research part C: emerging technologies**, v. 75, p. 183–196, 2017.